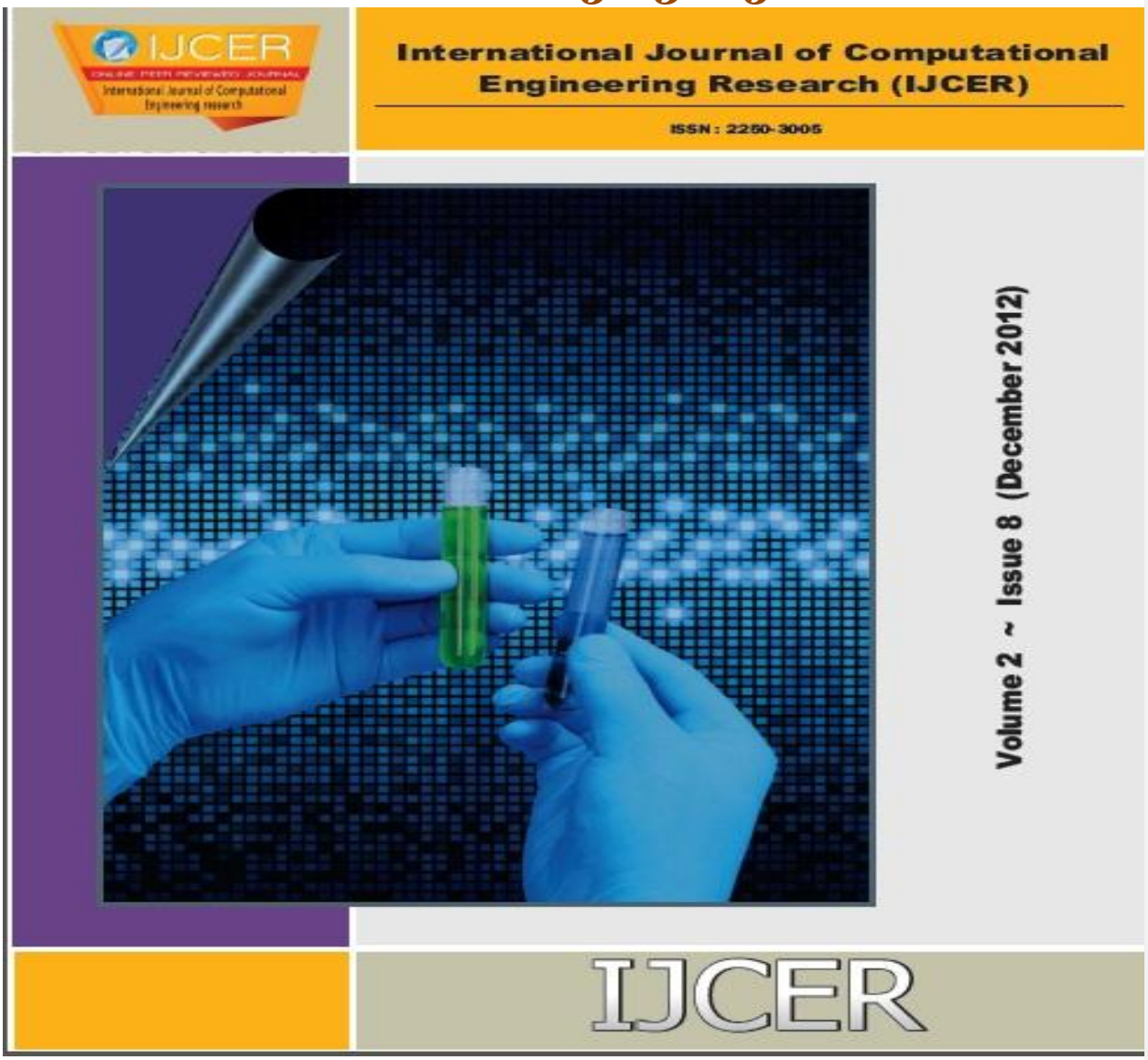
International Journal of computational Engineering Research (IJCER)

ISSN: 2250-3005



VOLUME 2 December 2012 ISSUE 8

Email: ijceronline@gmail.com Url : www.ijceronline.com

International Journal of computational Engineering Research (IJCER)

Editorial Board

CONTENTS :

Estimation of Consciousness of Human Mind Using Wavelet Transform

[1,]Tinku Biswas, [2,]Swarup Sarkar, [3,]Akash Ku. Bhoi, [4,]Surajit Bagchi,

[1,2,3] Department of AE&I Engg, Sikkim Manipal Institute of Technology (SMIT), Majitar
[4]Department of AE&I Engg ,Heritage Institute of Technology, Kolkata

Abstract

This paper introduces a two non-invasive electrode based electroencephalography (EEG) scheme to pick-up the bio-potential (generated by the neuron network of human brain) for the assessment of consciousness level. With the help of a suitable algorithm (developed by us), processed in LabVIEW environment, real-time β-wave (frequency range 13-30 Hz.) is extracted (representing the consciousness level of the brain activities) from the complex bio-signal and reproduced on the computer monitor. The data array is further processed in MATLAB platform using Daubechies wavelet transform (dB 6, level 9) to cross check the results of interpretation of one's awareness level as obtained from LabVIEW environment. The results provided by our proposed device are in good agreements with the actual physiological actions of the subjects' brain and supported by the clinicians.

Key-Words: Electrodes, EEG, β-wave , consciousness ,Level of consciousness, Daubechies wavelet transform, dB6**.**

1. Introduction

EEG is a graphical record of the electrical activity of the brain. Three types of brainwaves are associated with different levels of arousal: theta waves occur during sleep, alpha waves are associated with wakefulness, and beta waves with excitement. EEGs can be used to monitor the effects of exercise since there is a close correlation between certain EEG wave patterns and fatigue or overtraining. They are also used to determine the extent of injuries inflicted to the head (for example, after a knockout in boxing).An electroencephalogram (EEG), also called a brain wave test, is a diagnostic test which measures the electrical activity of the brain (brain waves) using highly sensitive recording equipment attached to the scalp by fine electrodes.[8] This paper is based on the estimation of consciousness or awareness of human mind. In biomedical sense, consciousness is the abnormal generation and propagation of action potential of neurons. The action potential (AP) from neurons has been recorded with microelectrodes. The brain activities are different for different stages of human mind; like alert stage, relax stage, drowsy stage etc.[7] *Propofol induction reduces the capacity for neural information integration: implications for the mechanism of consciuosness and general anesthesia:* The cognitive unbinding paradigm suggests that the synthesis of neural information is attenuated by general anaesthesia. Here, we analyzed the functional organization of brain activities in the conscious and anesthetized states, based on functional segregation and integration. Electroencephalography (EEG) recordings were obtained from 14 subjects undergoing induction of general anaesthesia with propofol. We quantified changes in mean information integration capacity in each band of the EEG. After induction with propofol, mean information integration capacity was reduced most prominently in the γ band of the EEG (p = .0001). Furthermore, we demonstrate that loss of consciousness is reflected by the breakdown of the spatiotemporal organization of γ waves. We conclude that induction of general anaesthesia with propofol reduces the capacity for information integration in the brain. These data directly support the information integration theory of consciousness and the cognitive unbinding paradigm of general anaesthesia.[11] [2]

Background

A non-invasive system was developed by **Konkan Railway Corporation Limited** after the accident which was happened on Sainthia Station, Birbhum on 19th July, 2010. At least 60 people were killed in that accident. The main reason of that accident was the unconscious mind of the pilot. The signalman in-charge at the station claimed to have heard the station master trying to alert the driver of the Uttar Banga Express via walkie-talkie, but got no response. The guard, when questioned said that the driver did not respond to him on the walkie-talkie. The basic principle of their system was measuring the skin impedance between the skin surface electrodes, placed on the pilot's wrist. Depending upon the consciousness of the pilot/driver the skin impedance changes. The change of impedance is picked up by two skin surface electrodes, followed by a signal conditioning circuit. But their system gives erroneous result due to dirt, moisture etc. present in the pilot's hand, causing an increased impedance of the cell giving erroneous indications. The pilots have to drive train for a long time on railway track and the possibilities of presence of dirt and moisture in pilot's hand become greater. This is the main drawback of their system.

2. Theory

The bio-electric potential generated by the neuronal activity of brain is recorded by the electroencephalography (EEG). The brain's electrical charge is maintained by billions of neurons. Neurons are electrically charged (or "polarized") by membrane transport proteins that pump ions across their membranes. Neurons are constantly exchanging ions with the extracellular milieu, for example to maintain resting potential and to propagate action potentials. Ions of similar charge repel each other, and when many ions are pushed out of many neurons at the same time, they can push their neighbours, who push their neighbours, and so on, in a wave. This process is known as volume conduction. When the wave of ions reaches the electrodes on the scalp, they can push or pull electrons on the metal on the electrodes. Since metal conducts the push and pull of electrons easily, the difference in push or pull voltages between any two electrodes can be measured by a voltmeter. Recording these voltages over time gives us the EEG. The neuronal activities responsible for different stages of human brain are different. These neuronal activities were studied in terms of electrical signal to discriminate the different stages. Here, two non-disposal skin surface electrodes were used for collecting the signal from human brain and for differentiating the consciousness of the human mind. An alert person usually displays an unsynchronized high-amplitude EEG signal. [8][15] In this approach the bio-potential is collected by a sensor, reusable non-invasive skin surface electrode. Information lies in the frequency of the waves collected by the electrodes. At rest, relaxed and with the eyes closed, the frequency of these waves is 8-12 Hz (cycles/sec). This 'alpha' activity is believed to reflect the brain in 'idling' mode, because if the person then either opens the eyes, or does mental arithmetic with the eyes closed, these waves disappear, to be replaced by irregular patterns (so-called *desynchronized activity*). In normal sleep there is characteristic higher voltage activity, in patterns which vary according to the level of sleep.

Standard Waveforms

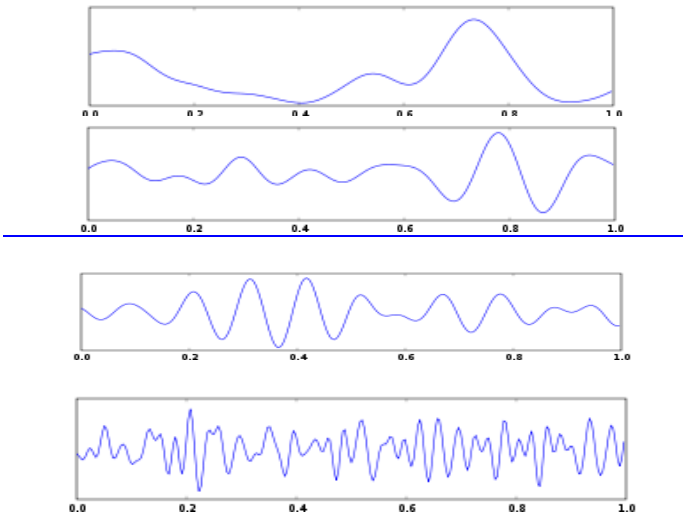


Figure 1(a) delta waves,(b)theta waves,(c) alpha waves,(d) beta waves

The Block Diagram of this system is shown below:

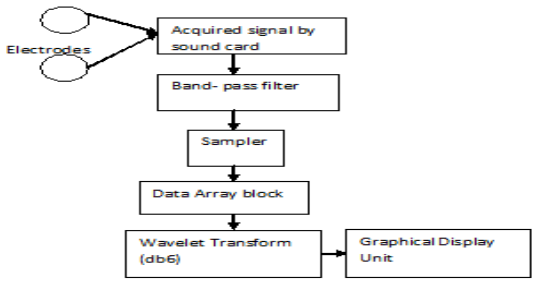


Figure 2: Block diagram of the system

3. Sensors

Brain cells communicate by producing tiny electrical impulses, also called brain waves. These electrical signals have certain rhythms and shapes, and EEG is a technique that measures, records, and analyzes these signals to help make a diagnosis. Electrodes are used to detect the electrical signals. They come in the shape of small discs that are applied to the skin surface. The bio-potential, generated by the neurons, are collected by two non-disposal skin surface Ag-AgCl electrode. One of the electrodes is placed on the surface of subject's forehead with the conducting paste. This electrode is actually responsible for the extraction of signal from forehead. Another electrode is placed on the ear-lobe. This electrode is working as reference electrode.

Figure 3: Figure of used electrodes during the experiment.

4. Signal conditioner

The signal conditioner block consists of an amplifier circuit, followed by a filter circuit. The signals collected from the human brain are in the range of microvolt. This voltage will be first amplified by an amplifier. After that, the amplified signal will be fed to an active band pass filter to eliminate the unwanted noise signals. However, EEG signals are not easily obtained. This is due to the signals' electrical characteristics. They are extremely weak signals, in the range of 1 – 160 μVpp. They are band limited to a very low frequency range, 0Hz - 100Hz for EEG. These signals are so small that they exist in the level of ambient noise. Our objective is concerned about the frequency range 'Beta' (13 - 30 Hz).

i. Filter design

Our target is to design an Active band pass filter whose bandwidth is 13 -30 Hz. f1=13Hz and f2=30Hz

5. Wavelet Transforms (Db6)

The word *wavelet* has been used for decades in digital signal processing and exploration geophysics. The equivalent French word *ondelette* meaning "small wave" was used by Morlet and Grossmann in the early 1980s. A **wavelet** is a wave-like oscillation with amplitude that starts out at zero, increases, and then decreases back to zero. It can typically be visualized as a "brief oscillation" like one might see recorded by a seismograph or heart monitor. Generally, wavelets are purposefully crafted to have specific properties that make them useful for signal processing. As a mathematical tool, wavelets can be used to extract information from many different kinds of data, including - but certainly not limited to - audio signals and images. Sets of wavelets are generally needed to analyze data fully. A set of "complementary" wavelets will deconstruct data without gaps or overlap so that the deconstruction process is mathematically reversible. Thus, sets of complementary wavelets are useful in wavelet based compression/decompression algorithms where it is desirable to recover the original information with minimal loss. A wavelet is a mathematical function used to divide a given function or continuous-time signal into different scale components. Usually one can assign a frequency range to each scale component. Each scale component can then be studied with a resolution that matches its scale. A wavelet transform is the representation of a function by wavelets. The wavelets are scaled and translated copies (known as "daughter wavelets") of a finite-length or fast-decaying oscillating waveform (known as the "mother wavelet"). Wavelet transforms have advantages over traditional Fourier transforms for representing functions that have discontinuities and sharp peaks, and for accurately deconstructing and reconstructing finite, non-periodic and/or non-stationary signals. Wavelet transforms are classified into discrete wavelet transforms (DWTs) and continuous wavelet transforms (CWTs). Note that both DWT and CWT are continuous-time (analog) transforms. They can be used to represent continuous-time (analog) signals. CWTs operate over every possible scale and translation whereas DWTs use a specific subset of scale and translation values or representation grid.

6. Daubechies Wavelets

are a family of orthogonal wavelets defining a discrete wavelet transform and characterized by a maximal number of vanishing moments for some given support. With each wavelet type of this class, there is a scaling function (also called father wavelet) which generates an orthogonal multi-resolution analysis. The procedure starts with passing this signal (sequence) through a half band digital low pass filter with impulse response h[n]. Filtering a signal corresponds to the mathematical operation of convolution of the signal with the impulse response of the filter. The convolution operation in discrete time is defined as follows:

$$x[n] * h[n] = \sum_{k=-\infty}^{\infty} x[k].h[n-k]$$

A half band lowpass filter removes all frequencies that are above half of the highest frequency in the signal. For example, if a signal has a maximum of 1000 Hz component, then half band lowpass filtering removes all the frequencies above 500 Hz. After passing the signal through a half band lowpass filter, half of the samples can be eliminated according to the Nyquist's rule. Half the samples can be discarded without any loss of information. In summary, the low pass filtering halves the resolution, but leaves the scale unchanged. The signal is then sub sampled by 2 since half of the number of samples are redundant. This decomposition halves the time resolution since only half the number of samples now characterizes the entire signal. At every level, the filtering and sub sampling will result in half the number of samples (and hence half the time resolution) and half the frequency band spanned (and hence doubles the frequency resolution). Figure shown below illustrates this procedure, where x[n] is the original signal to be decomposed, and h[n] and g[n] are low-pass and high pass filters, respectively. The bandwidth of the signal at every level is marked on the figure 4.

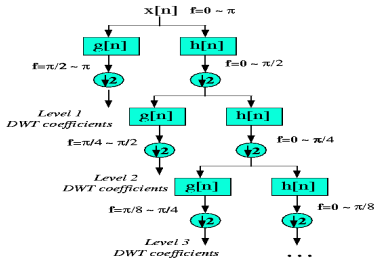


Figure 4

7. SIMULATION TOOLS

For the decomposition of the acquired EEG signal from human brain here, we used Lab VIEW 2009 software. In Lab VIEW platform we used the Daubechies wavelet transform, dB6 to get the de-composited bio-logical signal. We used graphical indicator in the labview platform to show the output responses of the system. There were a number of graphical indicators in the front panel of the labview which were used to plot the signal pattern. MatLab 7.0 is also used to cross check the results of interpretation of one's awareness level as obtained from Lab VIEW environment. [5]

8. Results And Discussions

In this paper, the samples have taken different types of drug to estimate the consciousness of their brain. The following types of drug and alcohol were being used during the experiment:

Heroin: Heroin is an opiate drug that is synthesized from morphine, a naturally occurring substance extracted from the seed pod of the Asian opium poppy plant. Heroin usually appears as a white or brown powder or as a black sticky substance, known as "black tar heroin."

Locally Made Alcohol (Lma): Locally produced moonshine is known in India as tharra, and also (among other names) as desi, latta, gawathi, Haathbhatti, desi daru, hooch, Potli, kothli, dheno, mohua, chullu, Narangi, Neera, kaju, cholai, Saaraayi and santra. It is made by fermenting the mash of sugar cane pulp in large spherical containers made from waterproof ceramic (terra cotta). However, it is dangerous, mainly because of the risk of alcohol or copper formaldehyde poisoning.

Whisky: Whisky or whiskey is a type of distilled alcoholic beverage made from fermented grain mash. Different grains are used for different varieties, including barley, malted barley, rye, malted rye, wheat, and corn. Whisky is typically aged in wooden casks, made generally of charred white oak. Whisky is a strictly regulated spirit worldwide with many classes and types. The typical unifying characteristics of the different classes and types are the fermentation of grains, distillation, and aging in wooden barrels.

Cannabis: Cannabis a genus of flowering plants that includes three putative varieties, *Cannabis sativa*, *Cannabis indica* and *Cannabis ruderalis*. These three taxa are indigenous to Central Asia, and South Asia. *Cannabis* has long been used for fibre (hemp), for seed and seed oils, for medicinal purposes, and as a recreational drug. Industrial hemp products are made from *Cannabis* plants selected to produce an abundance of fiber.

The signal patterns for Heroin addicted subject before taking drug*:*

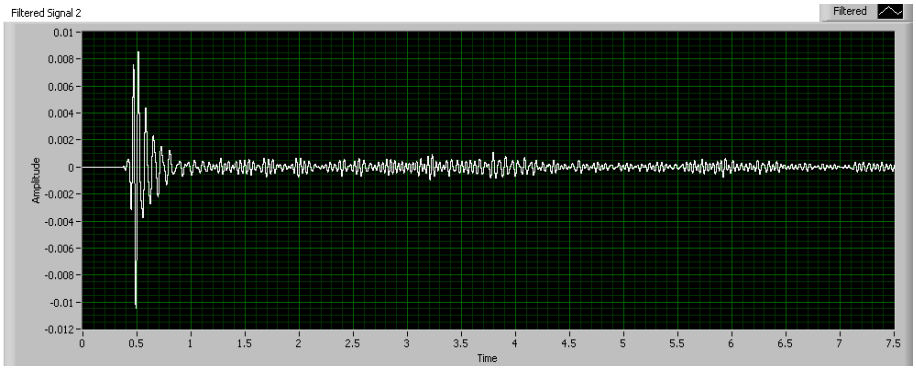


Figure 5(a): signal pattern before taking drug.

The signal patterns for Heroin addicted subject after taking drug:

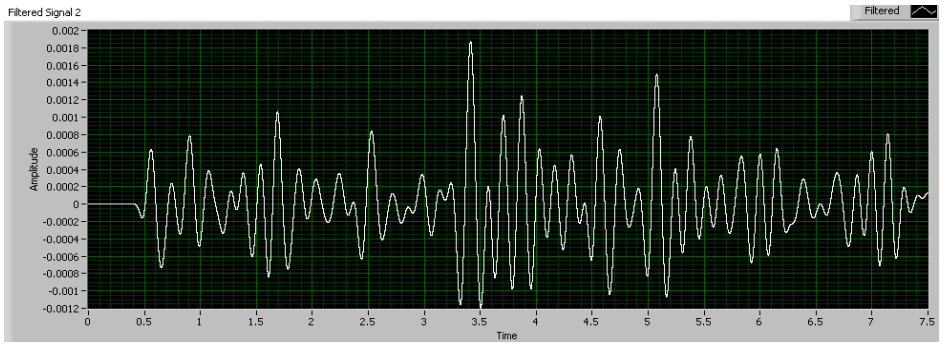


Figure 5(b): signal pattern after taking drug.

From the above two Figures we can see that the peak-peak voltage after taking drug goes down from the signal pattern, generating at normal stage.

The signal patterns for Cannabis addicted subject before smoking cannabis.

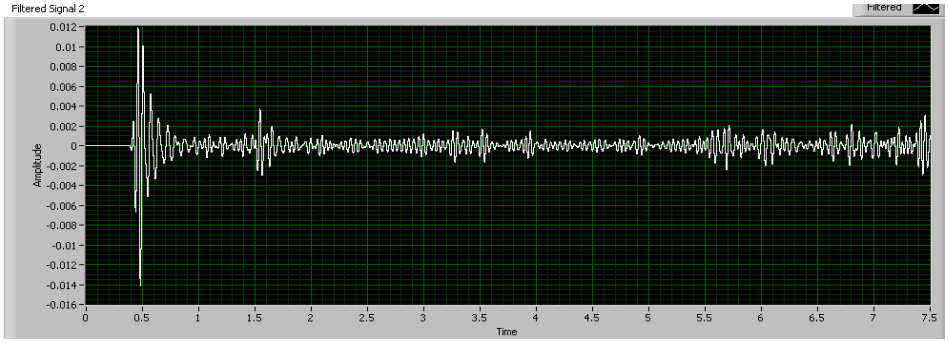


Figure 6(a): The signal patter at normal stage.

The signal patterns for Cannabis addicted subject after smoking cannabis.

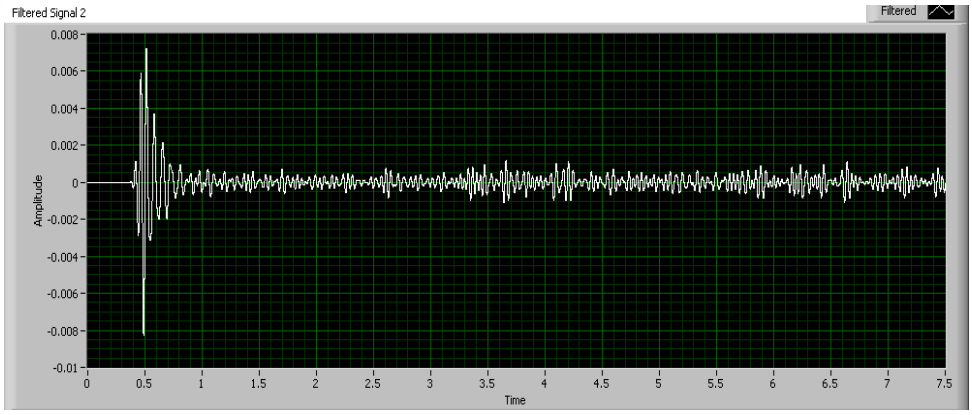


Figure 6(b): The signal patter after smoking cannabis.

In this case also the peak to peak voltages of the signal, related to the conscious stage are going down from the peak to peak voltages of the signal pattern which was extracting after smoking cannabis.

The signal pattern of subject 3 before drinking alcohol:

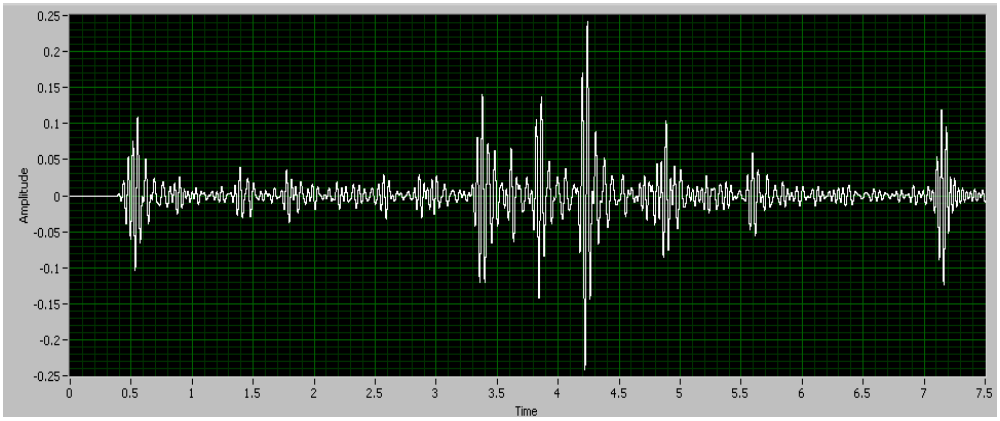


Figure 7(a): signal pattern at normal condition

The signal pattern of subject 3 after drinking alcohol*:*

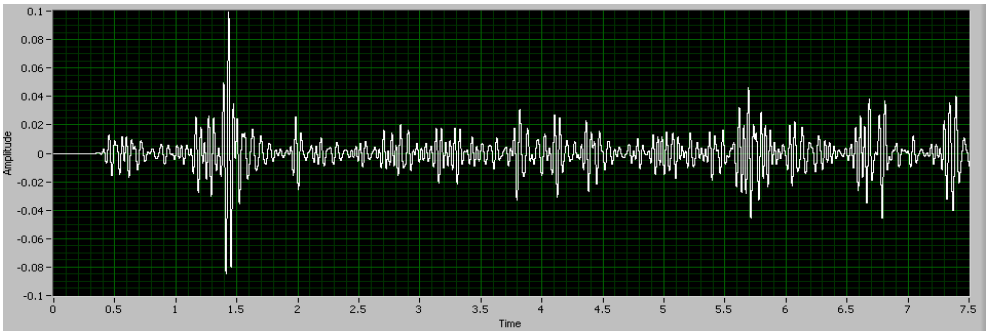


Figure 7(b): signal pattern after drinking alcohol

The peak to peak voltages decrease when the subject was drunk from the peak to peak voltages of the signal pattern, generating at normal condition.

9. CONCLUSION

The method which we adopted provides adequate information about the awareness level of the subject. Results obtained using our method was supported by the neurologists. To improve the performance further we will study the possibilities of implementing algorithm of adaptive filter to this system.The proposed method can be used as a safety measure in railways, airways, and roadways to measures the consciousness level of the pilot/ driver.

REFERENCES

[1]. Computer Standards & Interfaces, Volume 33, Issue 2, February 2011, Pages 136–141, XVI IMEKO TC4 Symposium "Exploring New Frontiers of Instrumentation and Methods for Electrical and Electronic Measurements" and XIII International Workshop on ADC Modelling and Testing Consciousness and Cognition, Volume 18, Issue 1, March 2009, Pages 56–64

[2]. American Academy of Neurology Foundation. 1080 Montreal Avenue, St. Paul, MN 55116. Web site: http://www.neurofoundation.com/

[3]. American Society of Neurophysiological Monitoring. PO Box 60487, Chicago, IL 60660–0487. Web site: http://www.asnm.org/.

[4]. Journal article: University of Victoria, ELEC 499A Report.

[5]. National Instruments, LabVIEW Digital Filter Design Toolkit User Manual

[6]. B. B.Winter and J. G.Webster, "Reduction of interference due to common mode voltage in bio-potential amplifiers," IEEE Transactions on Biomedical Engineering

[7]. Streams and consciousness: visual awareness and the brain Original Research Article Trends in Cognitive Sciences, Volume 2, Issue 1, 1 January 1998, Pages 25-30A.David Milne

[8]. The dreaming brain/mind, consciousness and psychosis Original Research Article Consciousness and Cognition, Volume 20, Issue 4, December 2011, Pages 987-992 Ivan Limosani, Armando D'Agostino, Maria Laura Manzone, Silvio Scarone

[9]. A theoretical basis for standing and traveling brain waves measured with human EEG with implications for an integrated consciousness Original Research Article Clinical Neurophysiology, Volume 117, Issue 11, November 2006,Pages2424-2435Paul L. Nunez, Ramesh Srinivasa

[10]. Decoding visual consciousness from human brain signals Original Research Article Trends in Cognitive Sciences, Volume 13, Issue 5, May 2009, Pages 194-202 John-Dylan Haynes

[11]. Mapping Human Cognition: Thinking, Numerical Abilities, Theory of Mind, Consciousness Brain Mapping: The Systems, 2000, Pages 523-534 Marco Iacobon

[12]. Towards the networks of the brain: from brain imaging to consciousness Original Research Article Neural Networks, Volume 12, Issues 7–8, October–November 1999, Pages 943-959 J.G. Taylor

[13]. Consciousness and metarepresentation: A computational sketch Original Research Article Neural Networks, Volume 20, Issue 9, November 2007, Pages 1032-1039 Axel Cleeremans, Bert Timmermans, Antoine Pasqual

[14]. Signal detection theory, the exclusion failure paradigm and weak consciousness—Evidence for the access/phenomenal distinction? Original Research Article Consciousness and Cognition, Volume 18, Issue 2, June 2009, Pages 551-560 Elizabeth Irvine

[15]. Emotional consciousness: A neural model of how cognitive appraisal and somatic perception interact to produce qualitative experience Original Research Article Consciousness and Cognition, Volume 17, Issue 3, September 2008, Pages 811-834 Paul Thagard, Brandon Aubie

Flue Gas Analysis of a Small-Scale Municipal Solid Waste-Fired Steam Generator

[1,]A. J. Ujam, [2,]F. Eboh

[1,] Department of Mechanical and Production Engineering, Enugu state University of Science and Technology (ESUT), Enugu, Nigeria.

[2,] Department of Mechanical Engineering, Michael Okpara University ofAgriculture, Umudike, Abia State, Nigeria.

Abstract

Flue gas analysis of a small-scale municipal solid waste-fired steam generator has been presented in this work. The analysis was based on the selected design parameters: operating steam pressure of 10 bar, with fuel consumption rate of 500 Kg/h and combustion chamber which utilizes mass burn incineration using water wall furnace. The plant is designed as a possible option for thermal utilization of rural and urban wastes in Nigeria. The average daily generation of MSW was considered in order to assess the availability of the material. The data were collected from Enugu State Waste Management Authority **(**ENSWAMA).This was calculated based on the state population, urbanization and industrialization strengths. Calculation of calorific value of the waste to determine the heat contents was carried out using two methods: Bomb calorimeter and Dulong's formula. Some samples of the garbage were analyzed with bomb calorimeter in the National Centre For Energy Research & Development Laboratory, University of Nigeria Nsukka. This is important because it a direct measure of the temperature requirements that the specific waste will place on the system. The calorific values obtained from this analysis were 12572.308 KJ/kg, 14012.05 KJ/kg, 21833.26 KJ/kg and 20551.01 KJ/kg for paper products, woods, plastics and textiles waste respectively, while the energy content obtained from the elemental composition of waste using Dulong's formula was 15,101 KJ/kg .The maximum temperature of the furnace attained from the energy balance based on this value around the combustion chamber was 833.7 K and the amount of air required per kg of MSW was 8.66kg

Keywords: Solid-Waste, Steam, Temperature, Pressure, Flue gas, Calorific Value, Excess air, Moisture Content, Exergy, Energy, Combustion.

1. Introduction

As a result of high carbon dioxide, CO_2 emission from thermal energy conversion of fossil fuels which is one of the major causes of the greenhouse effect, boiler technologies based on biomass conversion represent a great potential to reduce CO_2 emission since they are based on the utilization of renewal energy source.Furthermore, since conventional energy sources are finite and fast depleting and energy demand is on the increase, it is necessary for scientists and engineers to explore alternative energy sources, such as municipal solid waste (MSW).Biomass is abundantly available on the earth in the form of agricultural residues, city garbage, cattle dung, but is normally underutilized. For an efficient utilization of these resources, adequate design of municipal solid waste- fired steam boiler is necessary in order to extract heat produced in the combustion of waste, considering the calculated high calorific value of MSW and the availability of this material around us. The environmental benefits of biomass technologies are among its greatest assets. Global warming is gaining greater acceptance in the scientific community. There appears now to be a consensus among the world's leading environmental scientists and informed individuals in the energy and environmental communities that there is a discernable human influence on the climate; and that there is a link between the concentration of carbon dioxide (one of the greenhouse gases) and the increase in global temperatures.

Appropriate utilization of Municipal Solid Waste when used can play an essential role in reducing greenhouse gases, thus reducing the impact on the atmosphere. In addition, some of the fine particles emitted from MSW are beneficial. Bottom and fly ash are being mixed with sludge from brewery's wastewater effluent treatment in a composting process, thus resulting in the production of a solid fertilizer. The possibility of selling the bottom and fly ash to the ceramics industry is also being considered, which increases the potentials of MSW fired steam boiler. S.O. Adefemi et al[1] in their work on this subject correlated the concentration of heavy metals in roots of plant from Igbaletere (in Nigeria) dump site with the concentration of heavy metals in the soil samples from the dump site. A. B. Nabegu [2] found out that

solid waste generated by households (62.5%) in Kano metropolis far out weighed that generated by various institutions in the same metropolis (5.8%). In the analysis of Municipal Solid Waste management in Addis Ababa, Nigatu et al [3] observed that part of the reasons for low performance solid waste management was the inadequate and malfunctioning of operation equipment and open burning of garbage. This study thus seeks to analyse an efficient operating and burning system.

2. Combustion Analysis Of Municipal Solid Waste (MSW)

Considering the theoretical combustion reaction for the organic component of the waste, such as carbon, hydrogen and sulphur, Coskun et al [4] gave the equation for stoichiometric combustion as :

$C + (O_2 + \frac{79}{21}N_2) \rightarrow CO_2 + \frac{79}{21}N_2$ (1)

$H + 0.25(O_2 + 3.76N_2) \rightarrow 0.5H_2O + 0.94N_2$ (2)

$S + (O_2 + 3.76N_2) \rightarrow SO_2 + 3.76N_2$ (3)

It is known that nitrogen reacts with oxygen over about 1200^0C to form NO_x. In calculations, the upper limit of the flue gas temperature is assumed as 1200^0 C. Combustion process is assumed as in ideal case (Stiochiometric). So, nitrogen is not considered to react with oxygen during combustion reaction. It limits the intimacy between the fuel molecules and O_2 [4]

Table 1 shows the average daily generation of municipal solid waste in various states of Nigeria.

Table 1 Average daily generation of MSW in Nigeria

S/N	State	Metric Tonne	S/N	State	Metric Tonne	S/N	State	Metric Tonne
1	Abia	11	14	Enugu	8	27	Ogun	9
2	Adamawa	8	15	Gombe	6	28	Ondo	9
3	Anambra	11	16	Imo	10	29	Osun	7
4	Akwa-Ibom	7	17	Jigawa	9	30	Oyo	12
5	Balyesa	8	18	Kaduna	15	31	Plateau	9
6	Bauchi	9	19	Kano	24	32	Rivers	15
7	Benue	8	20	Kastina	11	33	Sokoto	9
8	Borno	8	21	Kebbi	7	34	Taraba	6
9	Cross River	9	22	Kogi	7	35	Yobe	6
10	Delta	12	23	Kwara	7	36	Zamfara	6
11	Ebonyi	7	24	Lagos	30	37	FCT	11
12	Edo	8	25	Nasarawa	6			
13	Ekiti	7	26	Niger	10			

(Source: ENSWAMA, MOE and NPC)

Complete combustion by using excess air can be expressed as follows:

$C + (I + \lambda)(O_2 + 3.76CO_2) \rightarrow CO_2 + (I + \lambda)(3.76N_2) + \lambda O_2$ (4)

$H + (I + \lambda)(O_2 + 3.76N_2) \rightarrow 0.5H_2O + (I + \lambda)(3.76\ N_2) + (0.75 + \lambda)O_2$ (5)

$S + (I + \lambda)(O_2 + 3.76N_2) \rightarrow SO_2 + (I + \lambda)(3.76N_2) + \lambda O_2$ (6)

In combustion reaction, λ is the fraction of excess combustion air, having the relationship, $n = (1 + \lambda)$

where n is the excess air ratio and $\lambda = \frac{Actual \frac{A}{F} ratio - Stoichiometric \frac{A}{F} ratio}{Stoichiometric \frac{A}{F} ratio}$

The mass balance equation can be expressed as showed in figure 1 in the form as,

$m_{in} = m_{out}$ (7)

i.e. The mass of reactants is equal to the mass of products

$m_{fuel} + m_{air} = m_{flue\ gas} + m_{ash} + m_{mst}$ (8)

$m_{fluegas} = m_{air} + (m_{fuel} - m_{ash} - m_{mst})$ (9)

From Eqn. 8

$m_{air} = (m_{fluegas} + m_{ash} + m_{mst}) - m_{fuel}$ (10)

Fig.1 Mass balance in the Furnace

Stiochiometric air amount (n=1) can be calculated as follows;

$m_{air\ steo}$ = O_2 required per kilogram of the fuel/23.3% of O_2 in air

$= m_{O,H}K_H - m_{O,O}K_O + m_{O,S}K_S + m_{O,C}K_C/0.233$ (11)

Where $m_{O,H}$, $m_{O,O}$, $m_{O,S}$, $m_{O,C}$, are the masses of oxygen in hydrogen,oxygen,sulphur and carbon respectively.

$$m_{air,steo} = \frac{8K_H - K_O + K_S + \frac{32}{12}K_C}{0.233} \quad (12)$$

$$m_{air.Steo.} = 34.3348K_H - 4.2918K_O + 4.2918K_S + 11.4449K_C$$

$$m_{air.steo.} = (3K_H - 0.3750K_O + 0.3750K_S + K_C)11.4449 \quad (13)$$

With excess air ratio,

$$m_{air} = (3K_H - 0.3750K_O + 0.3750K_S + K_C)(11.4449)(1+\lambda) \quad (14)$$

Where K denotes the percentage ratio of the element in chemical composition (in %) and m_{air} is the air requirement per kg fuel (kg air/kg fuel). Flue gas amount can be found by Eq. 9

Substituting Eq.13 in Eq. 9, knowing that calculations are done for 1 kg fuel, so the equation can be expressed as follows:

$$m_{fluegas} = (3K_H - 0.3750K_O + 0.3750K_S + K_C)(11.4449) + (1\text{-}K_{ash}\text{-}K_{mst}) \quad (15)$$

Employing the excess air ratio,

$$m_{fluegas} = (3K_H - 0.3750K_O + 0.3750K_S + K_C)(11.4449)(1+\lambda) + (1 - K_{ash} - K_{mst}) \quad (16)$$

Using the elemental composition of waste as shown in figure 1, the calculation of amount of air required and the flue gas produced can be done considering the above equations.

Table 2 Percentage by mass of MSW

Element	C	H	O	S	N	Moisture	Ash
percentage	35.5	5.1	23.9	0.5	2.4	25	7.6

(Source : P.Chattopadhyay, 2006)

2.1 Calculation of Combustion air supply

Considering theoretical combustion reaction for the elemental analysis of MSW showed in table 2.

Carbon (C):

$C + O_2 \rightarrow CO_2$

$12Kg\,C + 32Kg\,O_2 \rightarrow 44Kg\,CO_2$

Oxygen required = $0.355 * (32/12) = 0.947$/Kg MSW

Carbon dioxide produced = $0.355 * (44/12) = 1.302$/Kg MSW

Hydrogen (H):

$H_2 + \frac{1}{2} O_2 \rightarrow H_2O$

$2Kg\ H_2 + 16Kg\ O_2 \rightarrow 18Kg\ H_2O$

$1Kg\ H_2 + 8Kg\ O_2 \rightarrow 9Kg\ H_2O$

Oxygen required = $0.051 \times 8 = 0.408$ Kg/Kg MSW

Steam produced = $0.051 \times 9 = 0.459$ Kg/Kg MSW

Sulphur (S):

$S + O_2 \rightarrow SO_2$

$32Kg\ S + 32KgO_2 \rightarrow 64KgSO_2$
$1KgS + 1KgO_2 \rightarrow 2KgSO_2$
Oxygen required = 0.005 Kg/Kg MSW
Sulphur dioxide produced = $2 \times 0.005 = 0.01$Kg/KgMSW

Table 3 Oxygen Required per Kilogram of MSW

Constituent	Mass fraction	Oxygen required (Kg/Kg MSW)
Carbon (C)	0.355	0.947
Hydrogen (H)	0.051	0.408
Sulphur (S)	0.005	0.005
Oxygen (O)	0.239	- 0.239
Nitrogen (N)	0.024	___
Moisture	0.25	___
Ash	0.190	___
		1.121

O_2 required per Kilogram of MSW = 1.121Kg

Air required per Kilogram of MSW = $\frac{1.121}{0.233} = 4.811$Kg

Where air is assumed to contain 23.3% O_2 by mass
I.e. Stiochiometric air/fuel ratio = 4.811:1
For air supply which is 80% in excess (this has been derived from industry experience according to (Chattopadhyay, 2006) which suggests that 80% of excess air is just enough to optimize the combustion of solid refuse in the mass-burning system.

Actual A/F ratio, $m_{air} = 4.811 + \left(\frac{80}{100} \times 4.811\right) = 8.660/1$

Or alternatively, m_{air} can be found using Eq. (3.14)

2.2 Flue gas analysis

N_2 supplied = $0.767 \times 8.660 = 6.642$Kg
O_2 supplied = $0.233 \times 8.660 = 2.012$Kg
In the product,
$N_2 = 6.642 + 0.024 = 6.666$Kg and excess
$O_2 = 2.012 - 1.121 = 0.891$Kg
The results are tabulated in Table 4

Table 4 Flue Gas Analysis (Volumetric flue gas analysis).

Product	m_i Kg/Kg MSW	$\frac{m_i}{\sum m_i} \times 100\%$	$\bar{m}_i$ Kg/Kmol	$n_i = \frac{m_i}{\bar{m}_i}$	Wet $n_i / \sum n_i$ %	Dry $n_i / \sum n_i$ %
CO_2	1.302	13.958	44	0.030	9.317	10.135
H_2O	0.459	4.921	18	0.026	8.075	___
SO_2	0.01	0.107	64	0.0002	0.062	0.068
O_2	0.891	9.552	32	0.028	8.696	9.459
N_2	6.666	71.462	28	0.238	73.913	80.405
Total	9.328	100		0.322wet	100	100
				0.296dry		

2.3. Calculation of flue gas specific heat capacity

The specific heat values of gases found in flue gas are required to be known in order to obtain the average specific heat capacity (C_p) of flue gas. Taking these values from thermodynamic tables, a model is formed. The reference combustion reaction is required to generate one formulation in energy balance. Since carbon is an element found almost in all fossil fuels, the combustion reaction is considered to be a reference reaction for the model. Then, the specific heat values of all gases are defined depending on carbon dioxide. For that purpose, model coefficients are defined and expressed in detail as follows (Coskun et al., 2009).

$$C_{p,fluegas} = \frac{C_{p,C}}{(a_C + b_N + c_H + d_S)} \times \frac{m_{tot.steo.}}{m_{fluegas}} + fA \qquad (17)$$

a, b, c, d and f are the model coefficients in Eq. (3.17). $C_{p,\text{flue gas}}$ represents the average flue gas specific heat value. $C_{p,C}$ is the specific heat of CO_2.

2.3.1 Estimation of coefficient 'a$_c$'

Calculation method of a_c is given by the following equation:

$$a_c = \frac{a_m}{a_{cp}} \qquad (18)$$

where, a_{cp} can be defined as the specific heat ratio of CO_2 to CO_2. So, a_{cp} equals to 1. a_m can be indicated as the mass ratio of CO_2 to flue gas for n =1.

$$a_m = \frac{m_c}{m_{tot\cdot steo}} = \frac{3.667K_C}{m_{tot.steo.}} \qquad (19)$$

2.3.2 Estimation of coefficient 'b$_N$'

From (Coskun et al., 2009), calculation method of b_N is given by the following equation:

$$b_N = \frac{b_m}{b_{cp}} \qquad (.20)$$

where, b_{cp} can be defined as the specific heat ratio of CO_2 to N_2 for different temperatures. Coefficient b_{cp} is estimated by using heat capacity model. b_m can be defined as the mass ratio of N_2 to total flue gas.

$$b_{cp} = 0.9094 + 1.69\times10^{-4}\times T - \frac{11135}{T^2}$$

$$b_m = \frac{m_N}{m_{tot.steo.}}$$

$$b_m = \frac{0.767(2.9978.K_H - 0.3747\times K_O + 0.3747\times K_S \times K_C)\times(11.445) + K_N}{m_{tot.steo.}} \qquad (.21)$$

2.3.3 Estimation of coefficient 'c$_H$'

Coefficient 'c_H' can be expressed as in the following equation:

$$c_H = \frac{c_m}{c_{cp}} \qquad (22)$$

$$c_{cp} = 0.5657 - 6.68\times10^{-6}\times T - \frac{10465}{T^2} \qquad (23)$$

$$c_m = \frac{m_H}{m_{tot.steo.}} = \frac{8.938\times K_H + K_M}{m_{tot.steo.}} \qquad (24)$$

2.3.4 Estimation of coefficient 'd_S'

Coefficient d_S can be expressed as in the following equation:

$$d_s = \frac{d_m}{d_{cp}} \quad (25)$$

where, d_{cp} can be defined as the specific heat ratio of CO_2 to SO_2 for different temperatures. Coefficient d_{Cp} is estimated by using Vapour Pressure Model. d_m can be defined as the mass ratio of SO_2 to total flue gas [4]

$$d_{cp} = e^{[2.679 - \frac{151.16}{T} - 0.289\ln(T)]} \quad (26)$$

$$d_m = \frac{m_s}{m_{tot.steo.}} = \frac{2 \times K_S}{m_{tot.steo.}} \quad (27)$$

2.3.5 Calculation of coefficient 'f_A'

Coefficient f_A is calculated for access air amount. Coefficient $_{fA}$ can be expressed as in the following equation

$$f_A = f_m . C_{P,A} \quad (28)$$

$$C_{P,A} = 0.7124 \times 1.0000 \times 11^T \times T^{0.051} \quad (29)$$

$$f_m = \frac{m_{air}.steo.(n-1)}{m_{fluegas}} \quad (30)$$

2.3.6 Calculation of $C_{p,C}$

$C_{p,C}$ denotes the specific heat of CO_2. Specific heat value of CO_2 is given by [5] and adopted as a new parabola by using hoerl model.

$C_{P,C} = (0.1874) \times 1.000061^T \times T^{0.2665}$ (31)

3.2.4. Flue gas specific exergy value

The flow exergy of flue gas can be expressed in the ratio form as [6]

$$\psi = (h - h_0) - T_0(s - s_0) \quad (32)$$

where ψ is the flow exergy or the availability of stream flow neglecting K.E & P.E, s is the specific entropy and the subscript zero indicates the properties at the dead state of P_0 and T_0.

Entropy difference can be expressed in the form as

$$s - s_0 = C_p \ln\frac{T}{T_0} - R_{ave.} \ln\frac{P}{P_0} \quad (33)$$

$$R_{ave} = \frac{K_C(0.6927) + K_N(0.2968) + K_H(4.1249) + K_S(0.2596) + K_M(0.4615) + m_{air.steo}(0.2201)}{m_{fluegas}} + \frac{m_{air.steo.}(n-1)(0.287))}{m_{fluegas}} \quad (34)$$

where, R_{ave} is the average universal gas constant value of flue gas. Each gas has different gas constant. So, the average universal gas constants of combustion products are calculated and given in Eq. (35) [6]

$$\psi = C_{p,fluegas}(T - T_0) - T_0(C_{P,fluegas} \ln\frac{T}{T_0} - Rave.\ln\frac{P}{P_0}) \quad (35)$$

$$\psi = C_{p,fluegas}(T - T_0) - T_0 C_{p,fluegas}(\ln\frac{T}{T_0} - \frac{Rave.}{C_{p,fluegas}}\ln\frac{P}{P_0}) \quad (36)$$

$$\psi = C_{p,fluegas}[(T - T_0) - T_0(\ln\frac{T}{T_0} - \frac{Rave.}{C_{p,fluegas}}\ln\frac{P}{P_0})] \quad (37)$$

When $P \equiv P_0$,

General exergy flow equation can be written as:

$$\psi = C_{p,fluegas}[(T - T_0) - T_0(\ln\frac{T}{T_0})] \quad (38)$$

3. Calculation Of Calorific Value Of MSW

The first step in the processing of a waste is to determine its calorific content or heating value. This is a measure of the temperature and the oxygen requirements that the specific waste will be placed on the system[8]. The calorific value of a fuel can be determined either from their chemical analysis or in the laboratory[9]. In the laboratory Bomb Calorimeter is used. The analysis of some sample of wastes from the Energy Centre, UNN using *Bomb Calorimeter* are shown in Table 5

Table 5 Calculation of Calorific value of the fuel using Bomb Calorimeter

Paper product	Wood waste	Plastics waste	Textile waste
Sample wt.,m,=1.060g	Sample wt.,m,=0.974g	Sample wt.,m,=1.023g	Sample wt.,m,=1.065g
Initial Temp. $= 29.986^0 C$ Final Temp. $= 31.009^0 C$ $\Delta T = 1.023^0 C$	Initial Temp. $= 29.933^0 C$ Final Temp. $= 30.981^0 C$ $\Delta T = 1.048^0 C$	Initial Temp. $= 28.743^0 C$ Final Temp. $= 30.457^0 C$ $\Delta T = 1.714^0 C$	Initial Temp. $= 29.015^0 C$ Final Temp. = $30.695^0 C$ $\Delta T = 1.68^0 C$
Unburnt = 2.5+3.0=5.5	Unburnt = 1.3+2.2=3.5	Unburnt = 1.6+2.7=4.3	Unburnt = 2.5+0.8=3.3
Burnt = 10 - 5.5 = 4.5	Burnt = 10 - 3.5 = 6.5	Burnt = 10 - 4.3 = 5.7	Burnt = 10 - 3.3 = 6.7
$\Phi = 4.5 * 2.3 = 10.35$	$\Phi = 6.5 * 2.3 = 14.95$	$\Phi = 5.7 * 2.3 = 13.11$	$\Phi = 6.7 * 2.3 = 15.41$
$V = 2.3$	$V = 2.5$	$V = 3.9$	$V = 3.8$
$E = 13039.308$	$E = 13039.308$	$E = 13039.308$	$E = 13039.308$
$CV_p = (E\Delta T - \Phi - V)/m$	$CV_w = (E\Delta T - \Phi - V)/m$	$CV_p = (E\Delta T - \Phi - V)/m$	$CV_p = (E\Delta T - \Phi - V)/m$
$CV_P = 12572.22 J/g$	$CV_w = 14012.05 J/g$	$CV_P = 21833.26 J/g$	$CV_P = 20551..01 J/g$
= 12572.22KJ/kg	= 14012.05KJ/kg	= 21833.26KJ/kg	= 20551.01KJ/kg

(SOURCE; National Centre For Energy Research & Development (NCERD), UNN.)

For chemical analysis, using *Dulong's formula,* percentage by mass was considered and heat of combustion of Carbon, Oxygen and Hydrogen determined as shown in Table 6

Table 6 Heat of combustion for C, S and H

Combustion	Heat of Combustion
$C+O_2 \rightarrow CO_2$	8075kcal/kg
$S + O_2 \rightarrow SO_2$	2220kcal/kg
$H_2 + \frac{1}{2} O_2 \rightarrow H_2O$	34500kcal/kg

(Source: P.Chattopadhyay, 2006)

Dulong suggested a formula for the calculation of the calorific of the fuel from their chemical composition as

$CV_{msw} = 8075(K_C) + 2220(K_S) + 34500(K_H - K_O/8)$ (39)

where K_C, K_S, K_H and K_O stand for percentage by mass of Carbon, Sulphur, Hydrogen and Oxygen respectively. Substituting the values of K_C, K_S, K_H and K_O from Table 2 will give,

$CV_{msw} = 8075(0.355) + 2220(0.005) + 34500(0.051-0.239/8)$

$CV_{msw} = 3,606.5Kcal/kg$ (40)

$CV_{msw} = 15,101\ KJ/kg$ - - - - - - - - - (1cal = 4.187J)

Figures 2,3 & 4 show the views of the municipal waste steam boiler

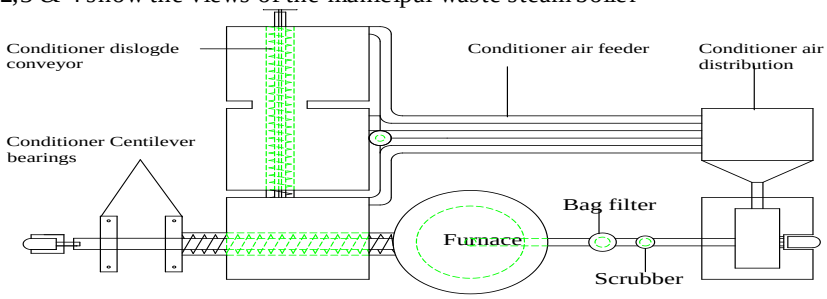


Figure 2 TOP VIEW OF MSW STEAM BOILER

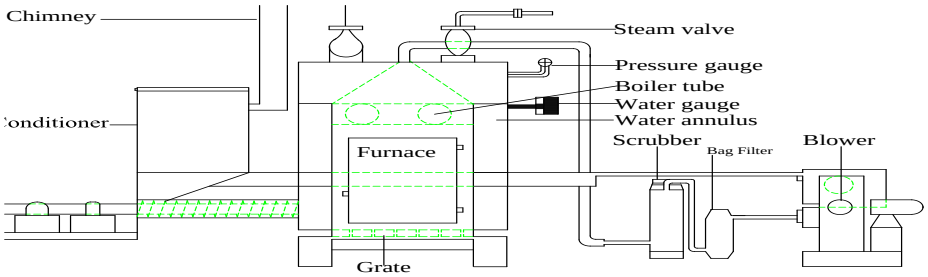


Figure 3 FRONT VIEW OF MSW STEAM BOILER

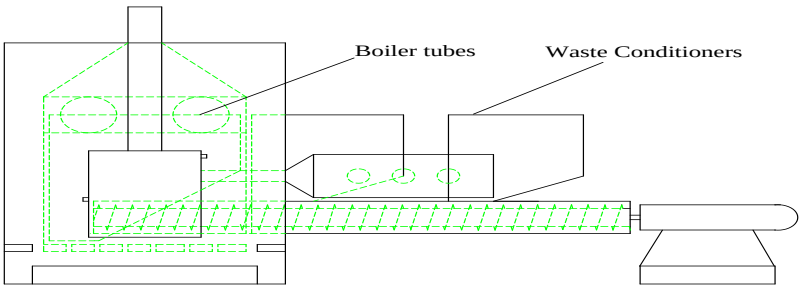


Figure 4 SIDE VIEW OF MSW STEAM BOILER

4.0

Results and Discussion

The Engineering Equation Solver (EES), developed at University of Wisconsin was used to obtain the solution of the equations.

4.1 Parameters for solution of the municipal solid waste-boiler design equations

The results of the calculated parameters for municipal solid waste design equations from the previous section are shown in table 6

Table 6 Parameters for solution of the municipal solid waste-boiler design equations

S/N	Symbols	Calculated data	S/N	Symbols	Calculated data
1	A_c [m^2]	0.1971	31	m_{uf}[kg]	0.326
2	A_{cyl} [m^2]	0.4058	32	O_2 [%]	80
3	A_{inc} [m^2]	0.9553	33	P[N/m^2]	10^6
4	A_{tubes} [m^2]	0.01623	34	Q_{bw}[kJ]	134.6
5	BHP[kW]	0.2587	35	$\dot{Q}_f$ [KW]	2098
6	CP_{fg} [kJ/kg]	1.047	36	$\dot{Q}_{fg}$ [m^3/s]	2.841
7	CV_{msw} [KJ/kg]	15101	37	Q_f [kJ]	10178
8	D_c[m]	0.5529	38	Q_{fg} [kJ]	5235
9	D_{inc}[m]	0.7188	39	Q_{ls}[kJ]	9578
10	D_{oc}[m]	0.8343	40	Q_r[kJ]	6504
S/N	**Symbols**	**Calculated data**	**S/N**	**Symbols**	**Calculated data**
11	D_{tubes}[m]	0.07188	41	Q_s[kJ]	1269
12	E [kg/kg]	4.049	42	Q_{uf} [kJ]	1900
13	eff.[%]	60.52	43	ℓ_{fg} [kg/m^3]	0.4723
14	H[m]	7.02	44	r_1[m]	0.005643
15	H_1 [kJ/kg]	763	45	r_c [m]	0.05634
16	H_2[kJ/kg]	2778	46	S_t [N/m^2]	1.360×10^8
17	h_{fg} [m]	10.59	47	t [m]	0.005947
18	H_{inc} [m]	7.014	48	T_a [K]	298
19	h_o [m]	0.7099	49	T_o [K]	298
20	H_{tubes} [m]	0.7188	50	T_{fg} [K]	833.7
21	h_w [mm]	5	51	T_{mit} [m]	0.0507
22	h_{wmax} [m]	4.158	52	τ_r [s]	1.002
23	K [W/mK]	0.04	53	T_w [m^3]	550
24	m_{air} [kg]	8.66	54	V_{fgc} [m^3]	14.41
25	$\dot{m}_a[kg/s]$	1.203	55	V_{inc} [m^3]	2.835
26	$\dot{m}_{fg}[kg/s]$	1.342	56	V_T [m^3]	7
27	$\dot{m}_{msw}[kg/s]$	0.1389	57	V_{water} [m^3]	1
28	$\dot{m}_{st}[kg/s]$	0.63	58	q[kW/m^2]	2264
29	m_f [kg]	0.674	59	$\dot{Q}_{st}$ [kW]	1269
30	m_{fg} [kg]	9.334	60	q_v[KW/m^3]	739.9
			61	Ψ[KJ/Kg]	209.2

4.2 Influence Of Specific Heat Capacity And Specific Exergy Value Of Flue Gas On Combustion

Considering the chemical composition of municipal solid waste used as a fuel, excess air amount and flue gas temperature which directly affect flue gas specific heat and exergy, variation of flue gas specific heat and exergy with furnace temperature for difference in values of excess air were done as shown in figs.5 and 6. From the figures, with increase in excess ratio, both flue gas specific heat and exergy decrease.

Table 7: Results For Variation Of Flue Gas Specific Heat With Furnace Temperature For Difference In Values Of Excess Air Ratio.

CP_{fgas} (KJ/KgK)	n	T_{fg} (K)	CP_{fgas} (KJ/KgK)	n	T_{fg} (K)	CP_{fgas} (KJ/KgK)	n	T_{fg} (K)
0.8460	1	700	0.7737	1.2	700	0.7128	1.4	700
0.8989	1	1100	0.8220	1.2	1100	0.7573	1.4	1100
0.9491	1	1300	0.8679	1.2	1300	0.7996	1.4	1300
0.9976	1	1500	0.9123	1.2	1500	0.8405	1.4	1500
1.0450	1	1700	0.9558	1.2	1700	0.8806	1.4	1700
1.0920	1	1900	0.9987	1.2	1900	0.9201	1.4	1900
1.1390	1	2100	1.0410	1.2	2100	0.9593	1.4	2100
1.1850	1	2300	1.0840	1.2	2300	0.9984	1.4	2300
1.2310	1	2500	1.1260	1.2	2500	1.0370	1.4	2500
1.2780	1	2700	1.1690	1.2	2700	1.0770	1.4	2700

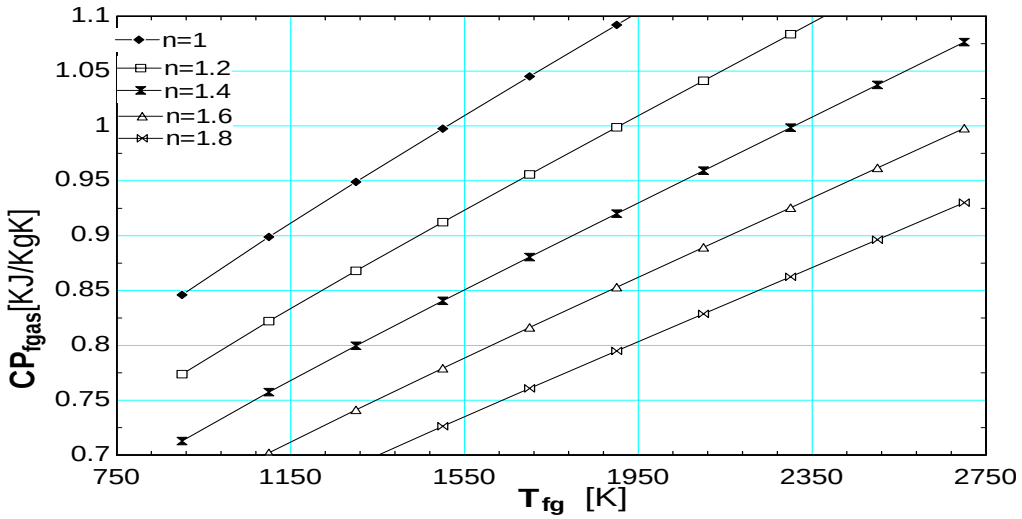


Figure 5: Variation of flue gas specific heat with furnace temperature for difference in values of excess air ratio.

Table 8: Results for Variation of flue gas specific exergy with furnace temperature for difference in values of excess air ratio.

Ψ [KJ/Kg]	n	T_{fg} (K)	Ψ [KJ/Kg]	n	T_{fg} (K)	Ψ [KJ/Kg]	n	T_{fg} (K)
230.6	1	700	210.9	1.2	700	194.3	1.4	700
371.1	1	1100	339.4	1.2	1100	312.6	1.4	1100
534.4	1	1300	488.7	1.2	1300	450.2	1.4	1300
718.7	1	1500	657.2	1.2	1500	605.5	1.4	1500
923.0	1	1700	844.1	1.2	1700	777.6	1.4	1700
1147.0	1	1900	1049.0	1.2	1900	966.0	1.4	1900
1387.0	1	2100	1271.0	1.2	2100	1170.0	1.4	2100
1651.0	1	2300	1510.0	1.2	2300	1391.0	1.4	2300
1931.0	1	2500	1766.0	1.2	2500	1627.0	1.4	2500
2230.0	1	2700	2039.0	1.2	2700	1879.0	1.4	2700

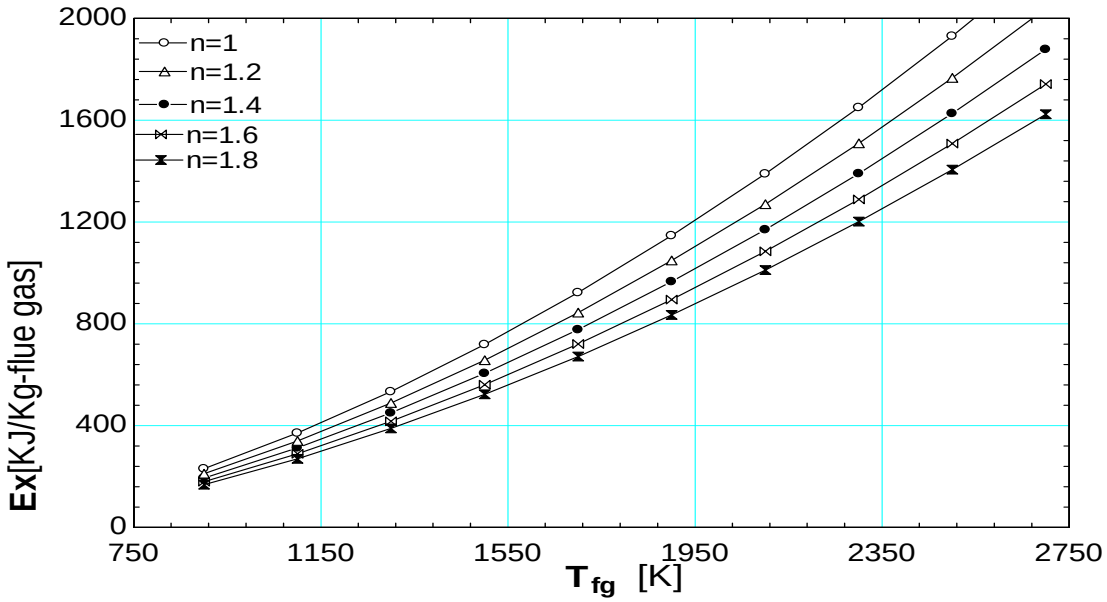


Figure 6: Variation of flue gas specific exergy with furnace temperature for difference in values of excess air ratio.

4.3 Influence Of Moisture Content

Wastes with different moisture contents have different drying characteristics. Those with higher moisture content require a longer drying time and much more heat energy, causing a lower temperature in the furnace; and vice versa. If the moisture content is too high, the furnace temperature will be too low for combustion, such that auxiliary fuel is needed to raise the furnace temperature and to ensure normal combustion. In order to evaluate the effect of moisture content on the combustion process, numerical simulation and analysis were made with ten different values of moisture content .The results of the analysis show that those wastes with a lower moisture content give rise to higher furnace temperatures and larger high-temperature zones during combustion, because the wastes with lower moisture contents have higher heating values and are more combustibles, being easier and faster to burn. Hence, to increase the efficiency of the boiler, refuse conditioner was used in this work to dry the wastes before they were conveyed to the furnace.

Table 9: Result for variation of flue gas Temperature with moisture content for difference in value of excess oxygen.

moisture	Excess O_2 (%)	T_{fg} (K)	moisture	Excess O_2 (%)	T_{fg} (K)	moisture	Excess O_2 (%)	T_{fg} (K)
0.030	5	2206	0.030	25	1941	0.030	45	1740
0.050	5	2115	0.050	25	1862	0.050	45	1670
0.070	5	2024	0.070	25	1782	0.070	45	1600
0.090	5	1932	0.090	25	1702	0.090	45	1529
0.011	5	1839	0.011	25	1622	0.011	45	1458
0.013	5	1749	0.013	25	1541	0.013	45	1387
0.015	5	1651	0.015	25	1460	0.015	45	1315
0,017	5	1557	0.017	25	1370	0.017	45	1243
0.019	5	1461	0.019	25	1295	0.019	45	1171
0.021	5	1365	0.021	25	1213	0.021	45	1098

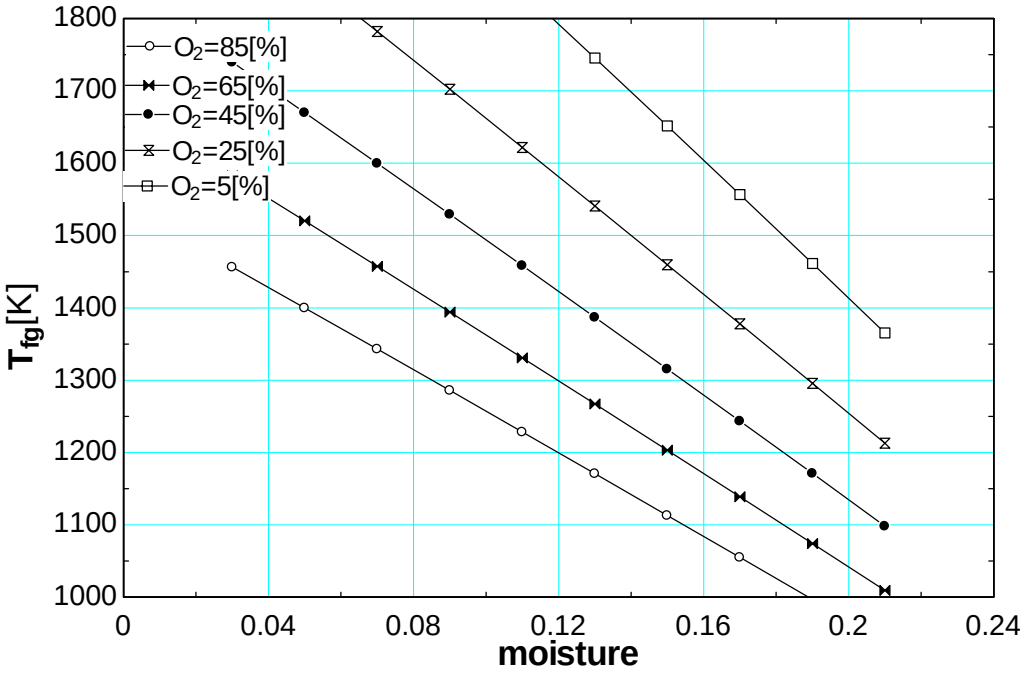


Figure 7: Variation of flue gas Temperature with moisture content for difference in value of excess oxygen.

4.4 Influence of excess air

The temperature in the furnace is closely related to MSW/air ratio. In order to predict the influence of excess air on the combustion in furnace, simulations were performed for different values of excess air. Results show that with the increase of excess air, the temperature of the furnace tends to decrease. To ensure adequate heating and burnout of wastes, a relatively high temperature level in the furnace should be maintained with a corresponding O_2 content.

Table 10 Result for variation of flue gas temperature with calorific value at different values of excess air

T_{fg} (K)	CV_{msw} (KJ/kg)	Excess O_2(%)	T_{fg} (K)	CV_{msw} (KJ/kg)	Excess O_2(%)	T_{fg} (K)	CV_{msw} (KJ/kg)	Excess O_2(%)
759.10	8000	5	716.80	8000	17	681.70	8000	29
817.10	9000	5	769.60	9000	17	730.00	9000	29
875.20	10000	5	822.30	10000	17	778.30	10000	29
933.20	11000	5	875.00	11000	17	826.60	11000	29
991.30	12000	5	927.80	12000	17	874.90	12000	29
1049.00	13000	5	980.50	13000	17	923.20	13000	29
1107.00	14000	5	1033.00	14000	17	971.50	14000	29
1165.00	15000	5	1086.00	15000	17	1020.00	16000	29
1223.00	16000	5	1139.00	16000	17	1068.00	17000	29
1283.00	17000	5	1191.00	17000	17	1116.00	18000	29

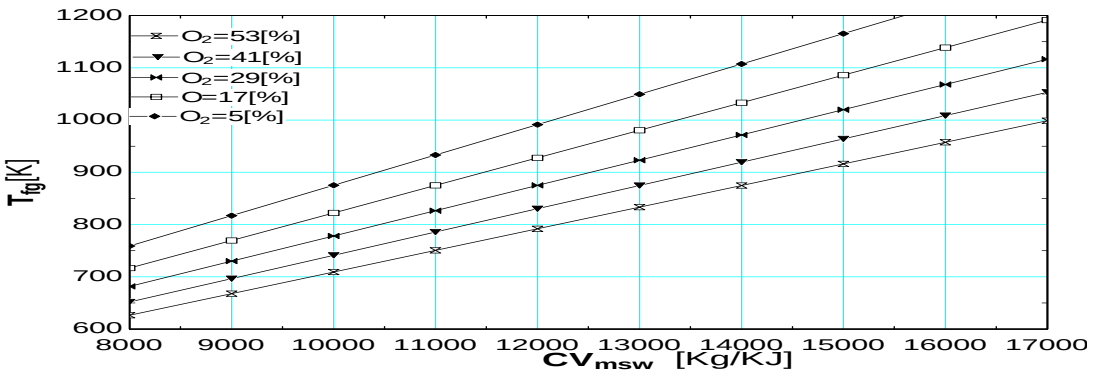


Figure 8 Variation of flue gas temperature with calorific value at different values of excess air.

5. Conclusions

With the rapid development of national economy, the ever-accelerating urbanization and the continued improvement of living standard, the output of the solid waste, particularly Municipal solid waste is constantly increasing. This causes environmental pollution and potentially affects people's health, preventing the sustained development of cities and drawing public concern in all of the society. The continuously generated wastes take up limited land resources, pollute water and air, and consequently lead to serious environmental trouble. Proper waste treatment is therefore an urgent and important task for the continued development of cities In this work, calculation of calorific value of municipal waste has been carried out from the elemental composition of the waste using Dulong's formula. The result of 15,101 KJ/kg obtained agrees with type 1 waste, N.T.Engineering,[10] that contains 25 percent moisture contents from waste classifications. With this heating value, maximum temperature of the flue gas of 833.7K was calculated from the heat balance equation in the furnace. Thermal analysis of the municipal solid waste boiler done with the operational conditions taken into account, showed that the municipal solid waste with higher moisture content has a lower heating value, corresponding to a lower temperature in the furnace and a lower O_2 consumption during combustion, resulting in a higher O_2 content at the outlet. Hence, for an efficient use of municipal solid waste as a fuel for generation of steam in boiler, waste with lower moisture content and adequate excess air supply should be used. In practical operation, the air supply rate and the distribution of the primary air along the grate should be duly adapted for the specific conditions of the wastes. An appropriate excess air ratio can effectively ensure the burnout of combustibles in the furnace, suppressing the formation and the emission of pollutants.

References

[1]. S. O. Adefemi and E. E. Awokunmi (2009), "The Impact of Municipal Solid Waste Disposal
[2]. in Ado Ekiti Metropolis, Ekiti State, Nigeria", African Journal of Environmental Science &
[3]. Technology, Vol.3(8), Pp. 186-189
[4]. A. B. Nabegu (2010), "Analysis of Minicipal Solid waste in kano Metropolis, Nigeria",
[5]. Journal of Human Ecology, 31(2): 111-119
[6]. Nigatu Rigassa, Rajan D. Sundaraa and Bizunesh Bogale Seboka (2011), "Challenges and
[7]. Opportunities in Municipal Solid Waste Management: The case of Addis Ababa City,
[8]. Central Ethiopia", Journal of Human Ecology, 33(3): 179-190
[9]. Coskun, C., Oktay, Z., &Ilten, N. (2009). "A new approach for simplifying the calculation of
[10]. flue gas specific heat and specific exergy value depending on fuel composition".Energy
[11]. Journal, 34; 1898-1902.
[12]. Kyle B. G. (1984)."Chemical and process thermodynamics". Englewood Cliffs:NJ prentice-
[13]. Hall.
[14]. Kotas TJ.(1985)."The exergy method of thermal plant analysis".Great Britain:Anchor
[15]. Brendon.
[16]. Chattopadhyay,P. (2006)."Boiler Operation Engineering" .Tata McGraw-Hill New Delhi.
[17]. Harry M. F. (1998). Standard handbook of hazardous waste treatment and disposal.
[18]. McGraw-Hall, New York
[19]. Rajput.R.K (2008).Thermal Engineering.Laxmi,New Delhi.
[20]. 10.N.T.G.Engineering Ltd.(2009).CT Series Incinerators.Cleveland Trading
[21]. EstatenAlbert Road Darlington.

Mapping Fpga To Field Programmableneural Network Array(Fpnna)

[1],H Bhargav, [2,] Dr. Nataraj K. R

[1,]Assistant Professor,.Vidyavardhaka College of Engineering, Mysore, Karnataka, India.
[2,,]Professor,.SJB Institute of Technology, Bangalore, Karnataka, India.

Abstract

My paper presents the implementation of a generalized back-propagation multilayer perceptron (MLP) architecture, on FPGA, described in VLSI hardware description language (VHDL). The development of hardware platforms is not very economical because of the high hardware cost and quantity of the arithmetic operations required in online artificial neural networks (ANNs), i.e., general purpose ANNs with learning capability. Besides, there remains a dearth of hardware platforms for design space exploration, fast prototyping, and testing of these networks. Our general purpose architecture seeks to fill that gap and at the same time serve as a tool to gain a better understanding of issues unique to ANNs implemented in hardware, particularly using field programmable gate array (FPGA).This work describes a platform that offers a high degree of parameterization, while maintaining generalized network design with performance comparable to other hardware-based MLP implementations. Application of the hardware implementation of ANN with back-propagation learning algorithm for a realistic application is also presented.

Index Terms: Back-propagation, field programmable gate array (FPGA), hardware implementation, multilayer perceptron, neural network, NIR spectra calibration, spectroscopy, VHDL, Xilinx FPGA.

1. INTRODUCTION

In recent years, artificial neural networks have been widely implemented in several research areas such as image processing, speech processing and medical diagnoses. The reason of this widely implementation is their high classification power and learning ability. At the present time most of these networks are simulated by software programs or fabricated using VLSI technology [9]. The software simulation needs a microprocessor and usually takes a long period of time to execute the huge number of computations involved in the operation of the network. Several researchers have adopted hardware implementations to realize such networks [8]&[12].This realization makes the network stand alone and operate on a real-time fashion. Recently, implementation of Field Programmable Gate Arrays (FPGA's) in realizing complex hardware system has been accelerated [7].Field programmable gate arrays are high-density digital integrated circuits that can be configured by the user; they combine the flexibility of gate arrays with desktop programmability. An ANN's ability to learn and solve problems relies in part on the structural Characteristics of that network. Those characteristics include the number of layers in a network, the number of neurons per layer, and the activation functions of those neurons, etc. There remains a lack of a reliable means for determining the optimal set of network characteristics for a given application. Numerous implementations of ANNs already exist [5]–[8], but most of them being in software on sequential processors [2]. Software implementations can be quickly constructed, adapted, and tested for a wide range of applications. However, in some cases, the use of hardware architectures matching the parallel structure of ANNs is desirable to optimize performance or reduce the cost of the implementation, particularly for applications demanding high performance [9], [10]. Unfortunately, hardware platforms suffer from several unique disadvantages such as difficulties in achieving high data precision with relation to hardware cost, the high hardware cost of the necessary calculations, and the inflexibility of the platform as compared to software.In our work, we have attempted to address some of these disadvantages by implementing a field programmable gate array (FPGA)-based architecture of a neural network with learning capability because FPGAs are high-density digital integrated circuits that can be configured by the user; they combine the flexibility of gate arrays with desktop programmability.Their architecture consists mainly of: Configurable Logic Blocks (CLB's) where Boolean functions can be realized, Input output Blocks (IOB's) serve as input output ports, and programmable interconnection between the CLB's and IOB's.

2. MOTIVATION

Features of ANN support evaluation implementations of different implementations of networks by changing parameters such as the number of neurons per layer, number of layers & the synaptic weights. ANNs have three main characteristics: parallelism, modularity & dynamic adaptation. Parallelism means that all neurons in the same layer perform the computation simultaneously. Modularity refers to the fact that neurons have the same structural architecture. It is clear

from these characteristics that FPGAs are well tailored to support implementation of ANNs, since it has a regular structure based on a matrix of parallel configurable units.Implementations in Application Specific Integrated circuits (ASICs) lack flexibility for evaluating the performance of different implementations. This deficiency can be overcome by using Programmable Logic Devices (PLDs) such as FPGAs. FPGAs provide high performance for parallel computation & enhanced flexibility (if compared with ASICs implementation) & are the best candidates for this kind of hardware implementations. If we mount ANN on FPGA, design should be such that there should be a good balance between the response & area restrictions of ANN on FPGA.FPGAs are programmable logic devices that permit the implementation of digital systems. They provide arrays of logical cells that can be configured to perform given functions by means of configuring bit stream. An FPGA can have its behavior redefined in such a way that it can implement completely different digital systems on the same chip. Despite the prevalence of software-based ANN implementations, FPGAs and similarly, application specific integrated circuits (ASICs) have attracted much interest as platforms for ANNs because of the perception that their natural potential for parallelism and entirely hardware-based computation implementation provide better performance than their predominantly sequential software-based counterparts. As a consequence hardware-based implementations came to be preferred for high performance ANN applications [9]. While it is broadly assumed, it should be noted that an empirical study has yet to confirm that hardware-based platforms for ANNs provide higher levels of performance than software in all the cases [10]. Currently, no well defined methodology exists to determine the optimal architectural properties (i.e., number of neurons, number of layers, type of squashing function, etc.) of a neural network for a given application. The only method currently available to us is a systematic approach of educated trial and error. Software tools like MATLAB Neural Network Toolbox [13] make it relatively easy for us to quickly simulate and evaluate various ANN configurations to find an optimal architecture for software implementations. In hardware, there are more network characteristics to consider, many dealing with precision related issues like data and computational precision. Similar simulation or fast prototyping tools for hardware are not well developed.

Consequently, our primary interest in FPGAs lies in their reconfigurability. By exploiting the reconfigurability of FPGAs, we aim to transfer the flexibility of parameterized software based ANNs and ANN simulators to hardware platforms. All these features of ANN & FPGAs have made me think about giving a hardware(FPGA) platform for ANN. Doing this, we will give the user the same ability to efficiently explore the design space and prototype in hardware as is now possible in software. Additionally, with such a tool we will be able to gain some insight into hardware specific issues such as the effect of hardware implementation and design decisions on performance, accuracy, and design size.

3. PREVIOUS WORKS

In the paper published by Benjamin Schrauwen,1, Michiel D'Haene 2, David Verstraeten 2, Jan Van Campenhout in the year 2008 with the title Compact hardware Liquid State Machines on FPGA for real-time speech recognition have proposed that real-time speech recognition is possible on limited FPGA hardware using an LSM. To attain this we first explored existing hardware architectures (which we reimplemented and improved) for compact implementation of SNNs. These designs are however more than 200 times faster than real-time which is not desired because lots of hardware resources are spend on speed that is not needed. We present a novel hardware architecture based on serial processing of dendritic trees using serial arithmetic. It easily and compactly allows a scalable number of PEs to process larger net- works in parallel. Using a hardware oriented RC design flow we were able to easily port the existing speech recognition application to the actual quantized hardware architecture. For future work we plan to investigate different applications, such as autonomous robot control, large vocabulary speech recognition, and medical signal processing, that all use the hardware LSM architectures presented in this work, but which all have very different area/speed trade-offs. Parameter changing without resynthesis will also be investigated (dynamic reconfiguration or parameter pre-run shift-in with a long scan-chain are possibilities).In the paper published by Subbarao Tatikonda, Student Member, IEEE, Pramod Agarwal, Member, IEEE in the year 2008 with the title Field Programmable Gate Array (FPGA) Based Neural Network Implementation of Motion Control and Fault Diagnosis of Induction Motor Drive have proposed A study of fault tolerant strategy on the ANN-SVPWM VSI performed. This Strategy is based on the reconfiguration on the inverter topology after occurrence. The modified topology for the inverter is proposed. Entire system design on FPGA has been suggested which includes the programmable low-pass filter flux estimation, space vector PWM (neural network based), fault diagnosis block and binary logic block. The paper talks about the fault feature extraction and classification and then how the neural networks can be build on the FPGA. Digital circuits models for the linear and log sigmoid been discussed. This work clearly gives the observer no slight change in operation with any fault in one leg. Study suggests that feature extraction is a challenging research topic still to be exploited.

Still this work has many prospects in multilevel inverters, where the better operating algorithms can be proposed with increase in the level of inverters. The number of redundant states is more they have to be exploited in the near future work. The system behaves as no fault in the system at all.

4. NEURAL NETWROK MODEL DESCRIPTION

There are various hardware implementations based on ASIC,DSP & FPGA.DSP based implementation is sequential and hence does not preserve the parallel architecture of the neurons in a layer. ASIC implementations do not offer reconfigurablity by the user. FPGA is a suitable hardware for neural network implementation as it preserves the parallel architecture of the neurons in a layer and offers flexibility in reconfiguration. FPGA realization of ANNs with a large number of neurons is a challenging task. Selecting weight precision is one of the important choices when implementing ANNs on FPGAs. Weight precision is used to trade-off the capabilities of the realized ANNs against the implementation cost. A higher weight precision means fewer quantization errors in the final implementations, while a lower precision leads to simpler designs, greater speed and reductions in area requirements and power consumption. One way of resolving the trade-off is to determine the "minimum precision" required. In this work we consider the weights as 8 -bit fixed-point values. Direct implementation for non-linear sigmoid transfer functions is very expensive. As the excitation function is highly nonlinear we adopt the Look Up Table (LUT) method in order to simplify function computation. The LUT is implemented using the inbuilt RAM available in FPGA IC.

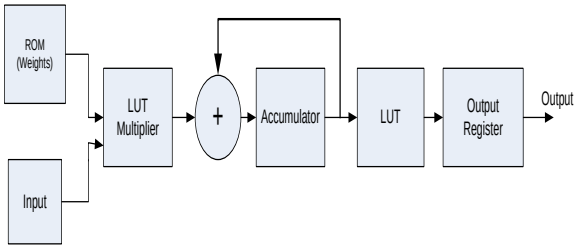


Fig 1. Neuron RTL Block Diagram

The use of LUTs reduces the resource requirement and improves the speed. Also the implementation of LUT needs no external RAM since the inbuilt memory is sufficient to implement the excitation function. The basic structure of the functional unit (neuron) that implements the calculations associated with neuron. Each neuron has a ROM, which stores the weights of the connection links between the particular neuron to the neurons of the previous layer. The multiplier performs high speed multiplication of input signals with weights from ROM. Multiplier is again implemented using an LUT multiplier. Such implementation of a multiplier needs one of the operands to be constant. In this case the other operand addresses the LUT where the result of multiplication is previously stored. Given two operands A&B with n & m bits respectively & B is constant, it is possible to implement their multiplication in LUT of 2^n. Since both multiplier & Activation Function(sigmoid function) are implemented using LUT, cost of implementation is very much reduced. We have tried comparing the model in the reference no. [15], with that of our model with respect to cost of implementation, in the conclusion section. A sixteen bit register is used to hold the weights from the ROM and the input signal from the previous layer. The whole MLP implementation is shown in Fig. 2.The network mainly consists of input registers, control unit, neurons and output register. To provide on neuron output to the next stage at each clock cycle a Multiplexer and a counter is used. The training of the network is done in software and the results loaded into hardware .Weights are updated during the training process, but remain constant during the detection process. The Register Transfer Level design of the system has been carried out using standard VHDL as the hardware description language. This language allows three different levels of description. We have chosen RTL to implement this system. The entire design process was done using the ISE development tool, from Xilinx (ISE development). The system physically implemented on Spartan-3 XC3 S4000 XILINX FPGA device.

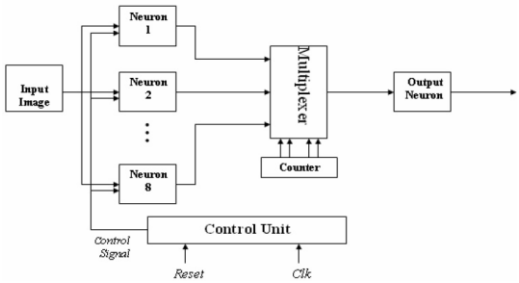


Fig2.The RTL block diagram of MLP neural network

Being a single source for hardware and software expertise, Mistral helps developers save on valuable development time and costs. The software engineers and hardware designers work together in an efficient and seamless manner providing expert design, development and support services.Mistral's professional services include hardware board design, reference designs, driver development, board support packages, embedded applications, codec and DSP algorithms across various domains. These services are delivered through a proven development process, designed specifically for embedded product development.

5. Results:

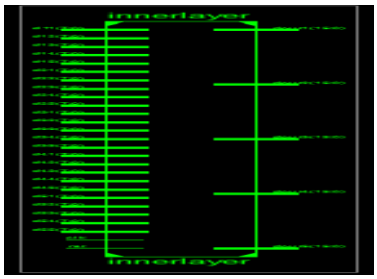


Fig 3. Inner layer top view of neural network.

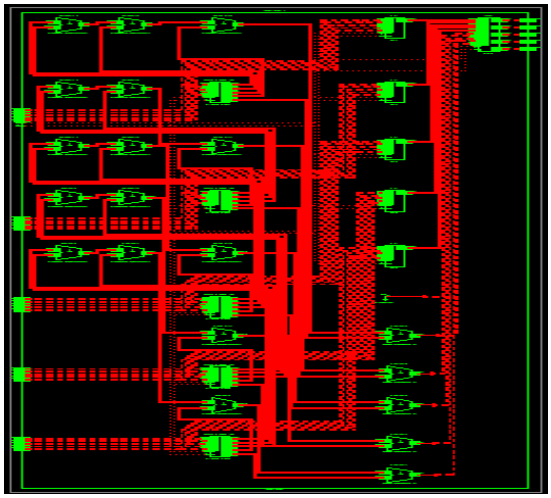


Fig 4. Inner layer RTL Schematic of Neural network.

Device Utilization Summary (estimated values)			
Logic Utilization	Used	Available	Utilization
Number of Slices	329	3584	9%
Number of Slice Flip Flops	430	7168	5%
Number of 4 input LUTs	591	7168	8%
Number of bonded IOBs	282	141	200%
Number of MULT18X18s	5	16	31%
Number of GCLKs	1	8	12%

Fig 5 . Inner layer Design summary

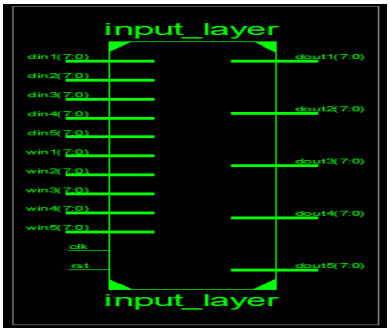


Fig 6. Input layer's top view.

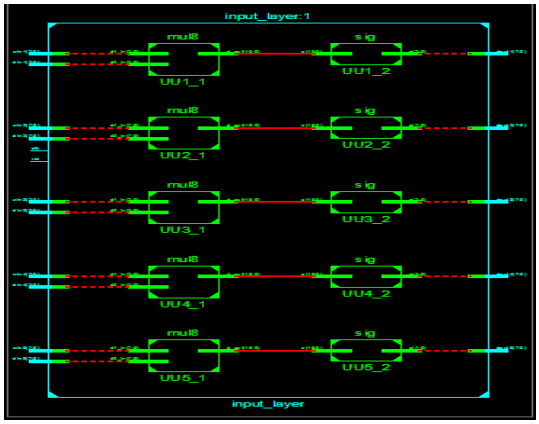


Fig 7.Input layer network of neural network.

Device Utilization Summary (estimated values)			
Logic Utilization	Used	Available	Utilization
Number of Slices	60	3584	1%
Number of 4 input LUTs	105	7168	1%
Number of bonded IOBs	120	141	85%
Number of MULT18X18s	10	16	62%

Fig 8. Input layer Design summary.

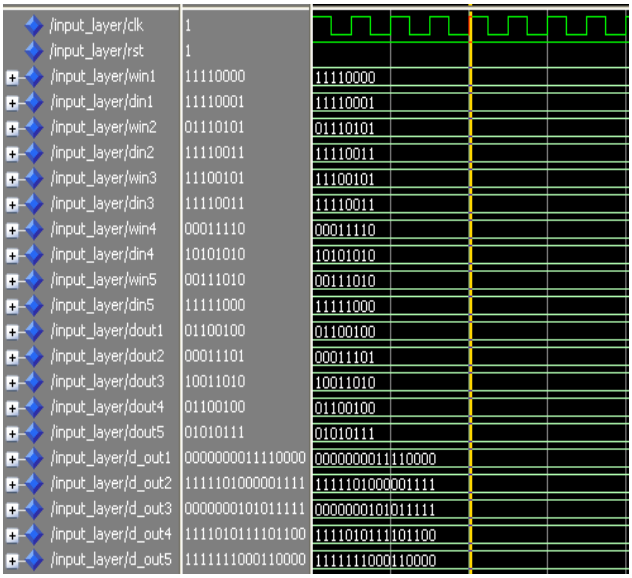


Fig 9. Simulation Results for Input layer of neural network

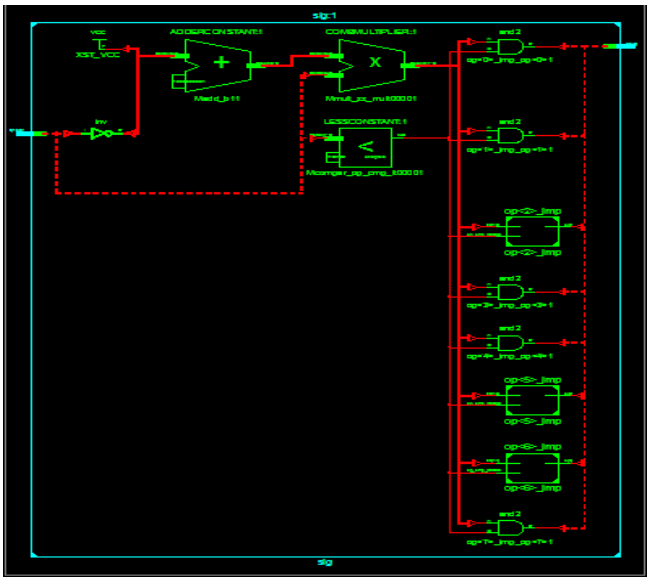


Fig 10. Sigmoid layer RTL Schematic of Neural network

Device Utilization Summary (estimated values)			[-]
Logic Utilization	**Used**	**Available**	**Utilization**
Number of Slices	12	3584	0%
Number of 4 input LUTs	21	7168	0%
Number of bonded IOBs	16	141	11%
Number of MULT18X18s	1	16	6%

Fig 11.Sigmoid layer Design summary

6. COMPARISION RESULTS

The device used to take Comparison results is Spartan XC3S400-5 PQ208.

Logic Utilization	Available	Proposed system used	Previous system used [15]
No of Slices	3584	12	50
No of 4 inputs LUTs	7168	21	94
No of 4 Bonded IOBs	141	16	21
Number of MULT18X18SIOs	16	1	3

Tab 1: Comparison results of Sigmoid function proposed system and previous system.

7. CONCLUSION

It is seen from the comparison table in section VI that previous system [15] used more number of Slices, LUT's and IOB's for sigmoid function. So our sigmoid function system used less number of resources as mentioned above. So we have effectively reduced the area utilization of these neural network systems. This has increased compactness & reliability of our system. In future our system permits us to achieve maximum level of optimization.Therefore one should aim at giving a hardware platform for ANN like FPGA because of the re-configurability of FPGA, we can develop the prototypes of hardware based ANNs very easily. Mapping FPGA into Field programmable Neural Network Arrays can find vast applications in real time analysis.

REFERENCES

[1] I. A. Basheer and M. Hajmeer, "Artificial neural networks: Fundamentals, computing, design, and application," *J. Microbio. Methods*, vol. 43, pp. 3–31, Dec. 2000.

[2] M. Paliwal and U. A. Kumar, "Neural networks and statistical techniques: A review of applications," *Expert Systems With Applications*, vol. 36, pp. 2–17, 2009.

[3] B. Widrow, D. E. Rumelhart, and M. A. Lehr, "Neural networks: Applications in industry, business and science," *Commun. ACM*, vol. 37, no. 3, pp. 93–105, 1994.

[4] A. Ukil, Intelligent Systems and Signal Processing in Power Engineering, 1st ed. New York: Springer, 2007

[5] B. Schrauwen, M. D'Haene, D. Verstraeten, and J. V. Campenhout, "Compact hardware liquid state machines on FPGA for real-time speech recognition," *Neural Networks*, vol. 21, no. 2–3, pp. 511–523, 2008.

[6] C. Mead and M. Mahowald, "A silicon model of early visual processing," *Neural Networks*, vol. 1, pp. 91–97, 1988.

[7] J. B. Theeten, M. Duranton, N. Mauduit, and J. A. Sirat, "The LNeuro chip: A digital VLSI with on-chip learning mechanism," in *Proc. Int. Conf. Neural Networks*, 1990, vol. 1, pp. 593–596.

[8] J. Liu and D. Liang, "A survey of FPGA-based hardware implementation of ANNs," in *Proc. Int. Conf. Neural Networks Brain*, 2005, vol. 2, pp. 915–918.

[9] P. Ienne, T. Cornu, and G. Kuhn, "Special-purpose digital hardware for neural networks: An architectural survey," *J. VLSI Signal Process.*, vol. 13, no. 1, pp. 5–25, 1996.

[10] A. R. Ormondi and J. Rajapakse, "Neural networks in FPGAs," in *Proc. Int. Conf. Neural Inform. Process.*, 2002, vol. 2, pp. 954–959.

[11] B. J. A. Kroese and P. van der Smagt, *An Introduction to Neural Networks*, 4th ed. Amsterdam, the Netherlands: The University of Amsterdam, Sep. 1991.

[12] J. Zhu and P. Sutton, "FPGA implementations of neural networks—A survey of a decade of progress," *Lecture Notes in Computer Science*, vol. 2778/2003, pp. 1062–1066, 2003.

[13] *"MATLAB Neural Network Toolbox User Guide,"* ver. 5.1, The MathWorks Inc., Natick, MA, 2006.

[14] A. Rosado-Munoz, E. Soria-Olivas, L. Gomez-Chova, and J. V. Frances, "An IP core and GUI for implementing multilayer perceptron with a fuzzy activation function on configurable logic devices," *J. Universal Comput. Sci.*, vol. 14, no. 10, pp. 1678–1694, 2008.

[15] *Rafid Ahmed Khali,l* Hardware Implementation of Backpropagation Neural Networks on Field programmable Gate Array (FPGA)

High speed arithmetic Architecture of Parallel Multiplier–Accumulator Based on Radix-2 Modified Booth Algorithm

[1]Harilal, M.Tech, [2,]DURGA PRASAD, M.tech,(Ph.D),
[1,]SKTRMCE,K.
[2,]Associate professor, SKTRMCE,

Abstract:

The sustained growth in VLSI technology is fuelled by the continued shrinking of transistor to ever smaller dimension. The benefits of miniaturization are high packing densities, high circuit speed and low power dissipation. Binary multiplier is an electronic circuit used in digital electronics such as a computer to multiply two binary numbers, which is built using a binary adder. A fixed-width multiplier is attractive to many multimedia and digital signal processing systems which are desirable to maintain a fixed format and allow a minimum accuracy loss to output data. This paper presents the design of high-accuracy modified Booth multipliers using Carry Look ahead Adder. The high accuracy fixed width modified booth multiplier is used to satisfy the needs of the applications like digital filtering, arithmetic coding, wavelet transformation, echo cancellation, etc. The high accuracy modified booth multipliers can also be applicable to lossy applications to reduce the area and power consumption of the whole system while maintaining good output quality. This project presents an efficient implementation of high speed multiplier using the shift and add method, Radix_2, Radix_4 modified Booth multiplier algorithm. The parallel multipliers like radix 2 and radix 4 modified booth multiplier does theComputations using lesser adders and lesser iterative steps. As a result of which they occupy lesser space as compared to the serial multiplier. This very important criteria because in the fabrication of chips and high performance system requires components which are as small as possible.

Key words: Booth multiplier, carry save adder (CSA) tree, computer arithmetic, digital signal processing (DSP), ultiplierand- accumulator (MAC).

1. Introduction

In this paper, we propose a high-accuracy fixed width modified booth multiplier. The functional model design consists ofbooth encoder, partial product generator and compression tree which uses Carry Look ahead Adder. The term "high accuracy"implies that the output produced by the normal 8X8 booth multiplication and the proposed 8X8 booth multiplication are equal. Theterm "fixed width" indicates that the partial product bits are adjusted to fixed width for Carry Look ahead.The result and one operand for the new modulo multipliers use weighted representation, while the other uses thediminished - 1. By using the radix-4 Booth recoding, the new multipliers reduce the number of the partial products to n/2 for n evenand (n+1)/2 for n odd except for one correction term. Although one correction term is used, the circuit is very simple. Thearchitecture for the new multipliers consists of an inverted end-around-carry carry save adder tree and one diminished-1 adder. Booth multipliers using generalized probabilistic estimation bias (GPEB) is proposed. The GPEB circuit can be easily builtaccording to the proposed systematic steps. The GPEB fixed-width multipliers with variable-correction outperform the existingcompensation circuits in reducing error. The GPEB circuit has improved absolute average error reduction, area saving, powerefficiency and accuracy. A truncated multiplier is a multiplier with two n bit operands that produces a n bit result. Truncatedmultipliers discard some of the partial products of a complete multiplier to trade off accuracy with hardware cost. This paperpresents a closed form analytical calculation, for every bit width, of the maximum error for a previously proposed family oftruncated multipliers. The considered family of truncated multipliers is particularly important since it is proved to be the design thatgives the lowest mean square error for a given number of discarder partial products. With the contribution of this paper, theconsidered family of truncated multipliers is the only architecture that can be designed, for every bit width, using an analyticalapproach that allows the a priori knowledge of the maximum error. A 2-bit Booth encoder with Josephson Transmission Lines(JTLs) and Passive Transmission Lines (PTLs) by using cell-based techniques and tools was designed. The Booth encoding methodis one of the algorithms to obtain partial products. With this method, the number of partial products decreases down to the halfcompared to the AND array method. a test chip for a multiplier with a 2-bit Booth encoder with JTLs and PTLs was

fabricated. Thecircuit area of the multiplier designed with the Booth encoder method is compared to that designed with the AND array method. New fixed-width multiplier topologies, with different accuracy versus hardware complexity trade-off, are obtained by varyingthe quantization scheme. Two topologies are in particular selected as the most effective ones. The first one is based on a uniformcoefficient quantization, while the second topology uses a non-uniform quantization scheme. The novel fixed-width multipliertopologies exhibit better accuracy with respect to previous solutions, close to the theoretical lower bound. The electricalperformances of the proposed fixed-width multipliers are compared with previous architectures. It is found that in most of theinvestigated cases the new topologies are Pareto-optimal regarding the area-accuracy trade-off.This paper focuses on variable-correction truncated multipliers, where some partial-products are discarded, to reduce complexity,and a suitable compensation function is added to partly compensate the introduced error. The optimal compensation function, thatminimizes the mean square error, is obtained in this paper in closed-form for the first time. A sub optimal compensation function,best suited for hardware implementation, is introduced. Efficient multipliers implementation based on sub-optimal function isdiscussed. Proposed truncated multipliers are extensively compared with previously proposed circuits. Power efficient 16times 16 Configurable Booth Multiplier (CBM) supports single 16-b, single 8-b, or twin parallel 8-b multiplication operations isproposed. Dynamic range detector detects the dynamic ranges of two input operands. It deactivates redundant switching activitiesin ineffective ranges.The proposed architecture can be used effectively in the area requiring high throughput such as a real-time digitalsignal processing can be expected.

2. Overview Of Mac

In this section, basic MAC operation is introduced. A multipliercan be divided into three operational steps. The first isradix-2 Booth encoding in which a partial product is generatedfrom the multiplicand and the multiplier. The secondis adder array or partial product compression to add all partialproducts and convert them into the form of sum and carry. Thelast is the final addition in which the final multiplication resultis produced by adding the sum and the carry. If the process toaccumulate the multiplied results is included, a MAC consistsof four steps, as shown in Fig. 1, which shows the operationalsteps explicitly.General hardware architecture of this MAC is shown inFig. 2. It executes the multiplication operation by multiplyingthe input multiplier and the multiplicand. This is added tothe previous multiplication result as the accumulation step.

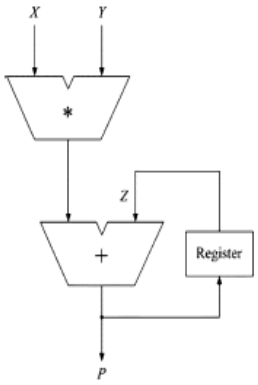


Figure 1. Hardware architecture of general Mac

The -bit 2's complement binary number can be expressedas

$$X = -2^{N-1}x_{N-1} + \sum_{i=0}^{N-2} x_i 2^i, \qquad x_i \in 0,1. \tag{1}$$

If (1) is expressed in base-4 type redundant sign digit form inorder to apply the radix-2 Booth's algorithm, it would be .

$$X = \sum_{i=0}^{N/2-1} d_i 4_i \tag{2}$$

$$d_i = -2x_{2i+1} + x_{2i} + x_{2i-1}. \tag{3}$$

If (2) is used, multiplication can be expressed as

$$X \times Y = \sum_{i=0}^{N/2-1} d_i 2^{2i} Y. \qquad (4)$$

If these equations are used, the afore-mentioned multiplication–accumulation results can be expressed as

$$P = X \times Y + Z = \sum_{i=0}^{N/2-1} d_i 2^i Y + \sum_{j=0}^{2N-1} z_i 2^i. \qquad (5)$$

Each of the two terms on the right-hand side of (5) is calculatedindependently and the final result is produced by addingthe two results. The MAC architecture implemented by (5) is called the standard design.If -bit data are multiplied, the number of the generated partialproducts is proportional to . In order to add them serially,the execution time is also proportional to . The architecture ofa multiplier, which is the fastest, uses radix-2 Booth encodingthat generates partial products and aWallace tree based on CSAas the adder array to add the partial products. If radix-2 Boothencoding is used, the number of partial products, i.e., the inputsto the Wallace tree, is reduced to half, resulting in the decreasein CSA tree step. In addition, the signed multiplication based on2's complement numbers is also possible. Due to these reasons,most current used multipliers adopt the Booth encoding.

3. Proposed Mac Architecture

In this section, the expression for the new arithmetic will bederived from equations of the standard design. From this result,VLSI architecture for the new MAC will be proposed. In addition,a hybrid-typed CSA architecture that can satisfy the operationof the proposed MAC will be proposed.

A. Derivation Of Mac Arithmetic

1) Basic Concept: If an operation to multiply two –bitnumbers and accumulate into a 2 -bit number is considered,the critical path is determined by the 2 -bit accumulation operation.If a pipeline scheme is applied for each step in the standarddesign of Fig. 1, the delay of the last accumulator mustbe reduced in order to improve the performance of the MAC.The overall performance of the proposed MAC is improved byeliminating the accumulator itself by combining it with the CSAfunction. If the accumulator has been eliminated, the criticalpath is then determined by the final adder in the multiplier. Thebasic method to improve the performance of the final adder is todecrease the number of input bits. In order to reduce this numberof input bits, the multiple partial products are compressed into asum and a carry by CSA. The number of bits of sums and carriesto be transferred to the final adder is reduced by adding the lowerbits of sums and carries in advance within the range in whichthe overall performance will not be degraded. A 2-bit CLA isused to add the lower bits in the CSA. In order to efficientlysolve the increase in the amount of data, a CSA architecture ismodified to treat the sign bit.

2) Equation Derivation: The aforementioned concept is appliedto (5) to express the proposed MAC arithmetic. Then, themultiplication would be transferred to a hardware architecturethat complies with the proposed concept, in which the feedbackvalue for accumulation will be modified and expanded for thenew MAC.First, if the multiplication in (4) is decomposed and rearranged,it becomes

(6)

$$X \times Y = d_0 2Y + d_1 2^2 Y + d_2 2^4 Y + \ldots + d_{N/2-1} 2^{N-2} Y.$$

If (6) is divided into the first partial product, sum of the middlepartial products, and the final partial product, it can be reexpressedas (7). The reason for separating the partial product additionas (7) is that three types of data are fed back for accumulation,which are the sum, the carry, and the preadded

$$X \times Y = d_0 2Y + \sum_{i=1}^{N/2-2} d_i 2^{2i} Y + d_{N/2-1} 2^{N-2} Y. \qquad (7)$$

results ofthe sum and carry from lower bits

Now, the proposed concept is applied to in (5). If is firstdivided into upper and lower bits and rearranged, (8) will bederived. The first term of the right-hand side in (8) correspondsto the upper bits. It is the value that is fed back as the sum andthe carry. The second term corresponds to the lower bits and is

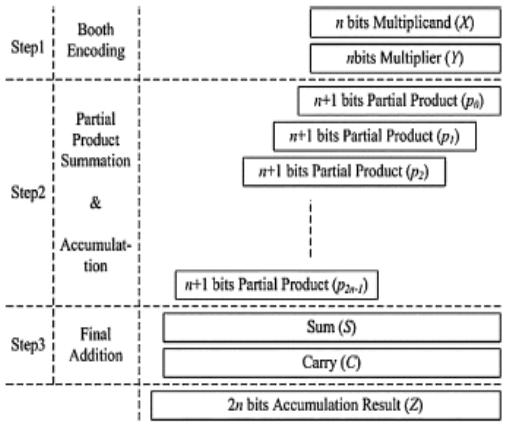


Figure2: Proposed arithmetic architecture of MAC

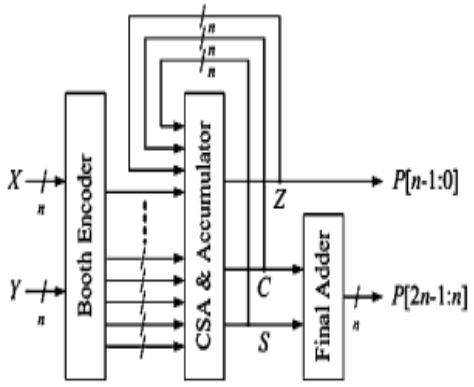


Figure3: Hardware architecture of proposed MAC

the value that is fed back as the addition result for the sum andcarry

$$Z = \sum_{i=0}^{N-1} z_i 2^i + \sum_{i=N}^{2N-1} z_i 2^i. \qquad (8)$$

The second term can be separated further into the carry termand sum term as

$$\sum_{i=N}^{2N-1} z_i 2^i = \sum_{i=0}^{N-1} z_{N+i} 2^i 2^N = \sum_{i=0}^{N-2} (c_i + s_i) 2^i 2^N. \qquad (9)$$

Thus, (8) is finally separated into three terms as

$$Z = \sum_{i=0}^{N-1} z_i 2^i + \sum_{i=0}^{N-2} c_i 2^i 2^N + \sum_{i=0}^{N-2} s_i 2^i 2^N. \qquad (10)$$

If (7) and (10) are used, the MAC arithmetic in (5) can beexpressed as

$$P = \left(d_0 2Y + \sum_{i=1}^{N/2-2} d_i 2^{2i} Y + d_{N/2-1} 2^{N-2} Y \right) + \left(\sum_{i=0}^{N-1} z_i 2^i 2^N + \sum_{i=0}^{N-2} c_i 2^i 2^N + \sum_{i=0}^{N-2} s_i 2^i 2^N \right). \quad (11)$$

If each term of (11) is matched to the bit position and rearranged,it can be expressed as (12), which is the final equationfor the proposed MAC. The first parenthesis on the right is theoperation to accumulate the first partial product with the addedresult of the sum and the carry. The second parenthesis is theone to accumulate the middle partial products with the sumoftheCSA that was fed back. Finally, the third parenthesis expressesthe operation to accumulate the last partial product withthe carry of the CSA.

B. Proposed Mac Architecture

If the MAC process proposed in the previous section is rearranged,it would be as Fig. 3, in which the MAC is organizedinto three steps. When compared with Fig. 1, it is easy to identifythe difference that the accumulation has been merged intothe process of adding the partial products. Another big differencefrom Fig. 1 is that the final addition process in step 3 is notalways run even though it does not appear explicitly in Fig. 3.

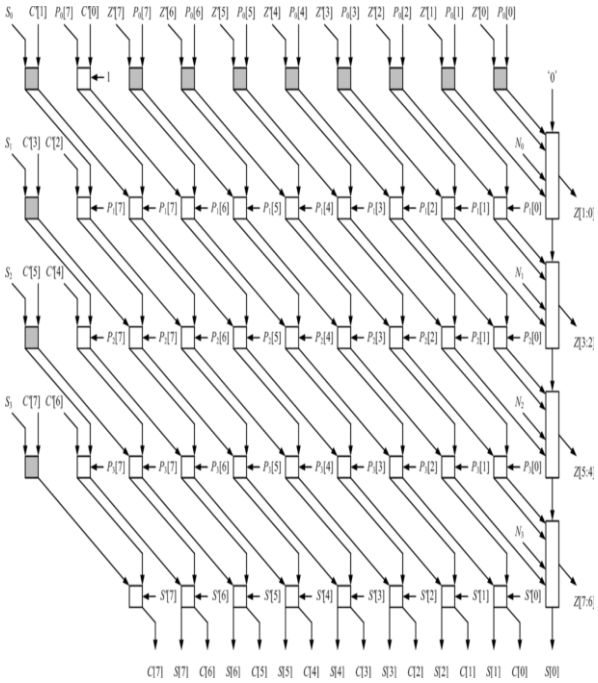


Figure 4.Architecture of the proposed CSA tree.

$$P = \left(d_0 2Y + \sum_{i=0}^{N-1} z_i 2^i \right) + \left(\sum_{i=1}^{N/2-1} d_i 2^{2i} Y + \sum_{i=0}^{N-2} c_i 2^i 2^N \right) + \left(d_{N/2-1} 2^{N-2} Y + \sum_{i=0}^{N-2} s_i 2^i 2^N \right). \quad (12)$$

Since accumulation is carried out using the result from step 2 insteadof that from step 3, step 3 does not have to be run until thepoint at which the result for the final accumulation is needed.The hardware architecture of the MAC to satisfy the processin Fig. 3 is shown in Fig. 4. The -bitMAC inputs, and , areconverted into an -bit partial product by passing throughthe Booth encoder. In the CSA and accumulator, accumulationis carried out along with the addition of the partial products. Asa result, -bit , and (the result from adding the lower bitsof the sum and carry) are generated. These three values are fedback and used for the next accumulation. If the final result forthe MAC is needed, is generated by adding andC in the final adder and combined with that wasalready generated.

C. Proposed Csa Architecture

The architecture of the hybrid-type CSA that complies withthe operation of the proposed MAC is shown in Fig. 5, whichperforms 8 8-bit operation. It was formed based on (12). InFig. 5, is to simplify the sign expansion and is to compensate1's complement number into 2's complement number.and correspond to the th bit of the feedback sum andcarry. is the th bit of the sum of the lower bits for eachpartial product that were added in advance and is the previousresult. In addition, corresponds to the th bit of theth partial product. Since the multiplier is for 8 bits, totally fourpartial products are generated from theBooth encoder. In (11), and correspond toand , respectively. This CSA requires at leastfour rows of FAs for the four partial products. Thus, totally fiveFA rows are necessary since one more level of rows are neededfor accumulation. For an -bit MAC operation, the levelof CSA is . The white square in Fig. 5 represents anFA and the gray square is a half adder (HA). The rectangularsymbol with five inputs is a 2-bit CLA with a carry input.The critical path in this CSA is determined by the 2-bit CLA.It is also possible to use FAs to implement the CSA withoutCLA. However, if the lower bits of the previously generatedpartial product are not processed in advance by the CLAs, thenumber of bits for the final adder will increase. When the entiremultiplier or MAC is considered, it degrades the performance.In Table I, the characteristics of the proposed CSA architecturehave been summarized and briefly compared with other architectures.For the number system, the proposed CSA uses 1's complement, but ours uses a modified CSA array without signextension.Thebiggest difference between ours and the othersis the type of values that is fed back for accumulation. Ours hasthe smallest number of inputs to the final adder.

Table1: Characteristics of CSA

	[6]	[17]	The Proposed
Number System	2's Complement	1's Complement	1's Complement
Sign Extension	Used	Used	Not Used
Accumulation	Result Data of Final Addition	Result Data of Final Addition	Sum and Carry of CSA
CSA Tree	FA, HA	FA, 2 bits CLA	FA, HA, 2 bits CLA
Final Adder	$2n$ bits	$(n+2)$ bits	n bits

Table2: Calculation of Hardware Resources

Component	[6]		[17]		The Proposed	
	General	16 bits	General	16 bits	General	16 bits
FA	$(\frac{n^2}{2}+n)$	964.8	$(\frac{n^2}{2}+2n+3)$	1092.1	$(\frac{n^2}{2}+\frac{n}{2})$	911.2
HA	0	0	0	0	$\frac{3n}{2}$	76.8
2 bit CLA	0	0	$(\frac{n}{2}-1)$	49	$\frac{n}{2}$	56
Accumulator $(2n+1)$ bits CLA		214	-		-	
Final adder	$2n$ bits	197	$(n+2)$ bits	109.5	n bits	97
Total		1375.8		1250.6		1141

4. Implementation And Experiment

In this section, the proposed MAC is implemented andanalyzed. Then it would be compared with some previousresearches. First, the amount of used resources in implementingin hardware is analyzed theoretically and experimentally, thenthe delay of the hardware is analyzed by simplifying Sakurai'salpha power law [20]. Finally, the pipeline stage is defined andthe performance is analyzed based on this pipelining scheme.Implementation result from each section will be compared withthe standard design and Elguibaly's design, each ofwhich has the most representative parallel MBA architecture.

A. Hardware Resource

1) Analysis of Hardware Resource: The three architecturementioned before are analyzed to compare the hardwareresources and the results are given in Table II. In calculatingthe amount of the hardware resources, the resources for Boothencoder is excluded by assuming that the identical ones wereused for all the designs. The hardware resources in Table II arethe results from counting all the logic elements for a general16-bit architecture. The 90 nm CMOS HVT standard cell libraryfrom TSMCwas used as the hardware library for the 16 bits. Thegate count for each designwas obtained by synthesizing the logicelements in an optimal form and the resultwas generated by multiplyingit with the estimated number of hardware resources. Thegate counts for the circuit elements obtained through synthesisare showninTable III, which are based on a two-input NANDgate.Let us examine the gate count for several elements in Table IIIfirst. Since the gate count is 3.2 for HA and 6.7 for FA, FA isabout twice as large as HA. Because the gate count for a 2-bitCLA is 7, it is slightly larger than FA. In other words, even if a2-bit CLA is used to add the lower bits of the partial products inthe proposed CSA architecture, it can be seen that the hardwareresources will not increase significantly.

Table3: Gate size of logic circuit element

Element	Gate Size
Inverter	0.8
2/3/4-NAND	1/1.5/2.5
2/3/4-NOR	1/2/2.2
2/3/4-XOR	2/4/6
2/3/4-AND	1.2/1.5/2
2/3/4-OR	1.2/1.5/2
Half Adder	3.2
Full Adder	6.7
D Flip-Flop	6.2
4×1 MUX	6
8×1 MUX	14.2
2 bits CLA	7
4 bits CLA	20.5

Table4: Estimation of gate size synthesis

	CSA		Booth Encoder	Final Adder		Total (C/L)	
nm	[17]	Proposed		[17]	Proposed	[17]	Proposed
90	1,067	1,009	713	104	97	1,884	1,819
130	1,216	1,158	864	118	110	2,198	2,131
180	1,581	1,484	808	120	114	2,510	2,407
250	2,027	2,001	1,129	141	131	3,297	3,261

As Table II shows, the standard design uses the most hardwareresources and the proposed architecture uses the least. Theproposed architecture has optimized the resources for the CSAby using both FA and HA. By reducing the number of input bitsto the final adder, the gate count of the final adder was reducedfrom 109.5 to 97.*2) Gate Count by Synthesis:* The proposed MAC and [17]were implemented in register-transfer level (RTL) using hardwaredescription language (HDL). The designed circuits weresynthesized using the Design Complier from Synopsys, Inc.,and the gate counts for the resulting netlists were measured andsummarized in Table IV. The circuits in Table IV are for 16-bitMACs. In order to examine the various circuit characteristicsfor different CMOS processes, the most popular four processlibraries (0.25, 0.18, 0.13 m, 90 nm) for manufacturing digitalsemiconductors were used. It can be seen that the finer theprocess is, the smaller the number of gates is.As shown in Table II, the gate count for our architecture isslightly smaller. It must be kept

in mind that ifa circuit is implemented as part of a larger circuit, the number ofgates may change depending on the timing for the entire circuitand the electric environments even though identical constraintswere applied in the synthesis. The results in Table IV were forthe combinational circuits without sequential element. The totalgate count is equal to the sum of the Booth encoder, the CSA,and the final adder.

Table 5: Normalized Capacitance and Gate Delay

Gate	Comment	C_i	T_g
Inverter	-	3	$t+c$
8×1 MUX	4-level logic	4	$35.2+t+c$
D-F/F	Slave delay	4	$16.1+t+c$
1 bit FA	input-to-sum	12	$39.6+t+c$
1 bit FA	input-to-carry	12	$38.7+t+c$
2 bits CLA	input-to-sum	12	$64.9+t+c$
2 bits CLA	input-to-carry	16	$53.9+t+c$
4 bits CLA	input-to-sum	12	$96.8+t+c$
4 bits CLA	input-to-carry	24	$88+t+c$

Table 6: Delay Time analysis and comparison

Step	[6]		[17]		The Proposed	
	General	16 bits	General	16 bits	General	16 bits
Step1	Booth Encoding $52.8n+59.9$	904.7	Booth Encoding $10.6n+81.1$	250.7	Booth Encoding $10.6n+81.1$	250.7
Step2	CSA $25.95n-51.9$	363.3	Hybrid CSA $33.55n-67.1$	469.7	Hybrid CSA $33.55n$	536.8
Step3	Final Addition $57.2n$	915.2	Final Addition $28.6n+57.2$	514.8	Final Addition $28.6n$	457.6
Step4	Accumulation $57.2n$	915.2	- -	-	- -	-
Critical Path	Accumulation $57.2n$	915.2	Final Addition $28.6n+57.2$	514.8	Hybrid CSA $33.55n$	536.8

B. Delay Model

1) Modeling: In this paper, Sakurai's alpha power law isused to estimate the delay. Because CMOS process is used andthe interconnect delay that is not due to gates related to logicoperation is ignored, was used. The delay by simplifyingthe alpha power law was modeled. Order for easy comparisonswith other architectures, the modeled values identicalto are used in this paper. The normalized input capacitanceand gate delay for the hardware building blocks withthese modeled values are shown in Table V.In Table II, is the ratio of the saturation velocity.Andare the load gate capacitance and gate capacitance of theminimum-area transistor, respectively. is the duration timeand is the falling time of the minimum-area inverter due to. Since delay modeling and its simplification process is notthe focus of this paper, it will not be described in detail here.

2) Delay Analysis: The results of delay modeling for theBooth encoder , the CSA , and the final adderusing Table VI are given in (13)–(16). It represents the select logic delay, buffer delay, andMUX delay, respectively.

$$T_f = \left(\frac{n}{4}\right) T_4(\text{carry}) = 28.6n.$$

$$T_b = T_s + (n+2)T_p + T_m \quad (13)$$

$$T_b = 12.3 + (n+2) \times 10.6 + 47.6 = 10.6n + 81.1 \quad (14)$$

$$T_c = \left(\frac{n}{2}\right) T_2(\text{carry}) = \left(\frac{n}{2}\right) 67.1 + 33.5n \quad (15)$$

(16)

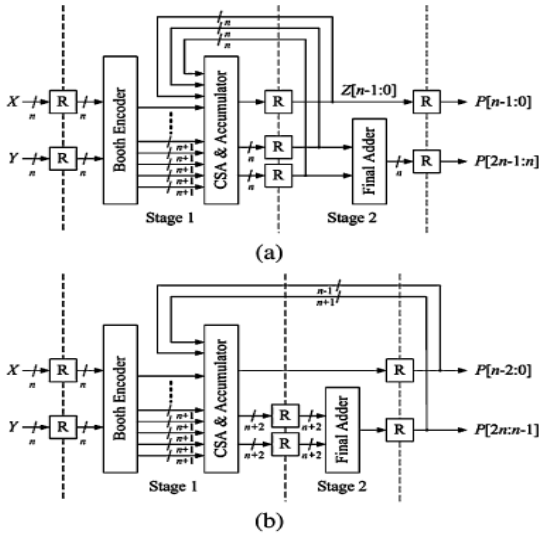


Figure5: pipelined Hardware structre. (a)proposed structure, (b)Elgubaly's structure

The delays in Table VI were obtained using the hardwareresources in Table II and the gate delays in Table V. FromTable VI, it is easily recognizable that the delay of [6] is considerablylarger than others. The proposed architecture uses thesame Booth encoder and the delay is also identicalto. Because the critical path for the CSA tree isdetermined by the 2-bit CLA, the delay is proportional to it.The proposed architecture has one more 2-bit CLA comparedto [17], as shown in Table II where the delay is greater by 67.1.The number of input bits for the final adder is less by one in ourarchitecture and the delay is also faster by 57.2.If pipelining is applied for each step, the critical path for theproposed architecture is 33.55 and it corresponds to the valueof 536.8 for 16-bit MAC. However, if the performance of the actual outputrate is considered, it can be verified that the proposed architectureis superior. The reason will be explained in detail in the nextsection with the pipelining scheme.Because of the difficulties in comparing other factors,only delay is compared. The sizes of bothMACs were 8 8 bitsand implemented by a 0.35mfabrication process. The delay of ours was 3.94, while it was it 4.26 ns, which meansthat ours improved about 7.5% of the speed performance. Thisimprovement is mainly due to the final adder. The architecture should include a final adder with the size of 2 to performan multiplication. It means that the operational bottleneckis induced in the final adder no matter how much delays are reduced in the multiplication or accumulation step, whichis the general problem in designing a multiplier. However, ourdesign uses -bit final adder, which causes the speed improvement.This improvement is getting bigger as the numberof input bits increases.

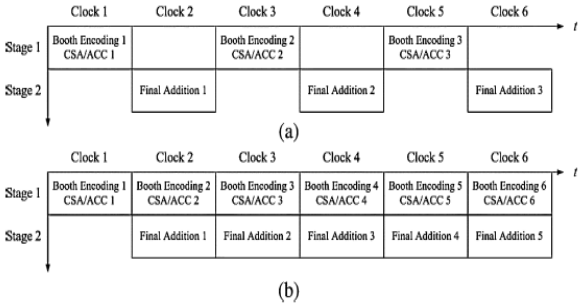
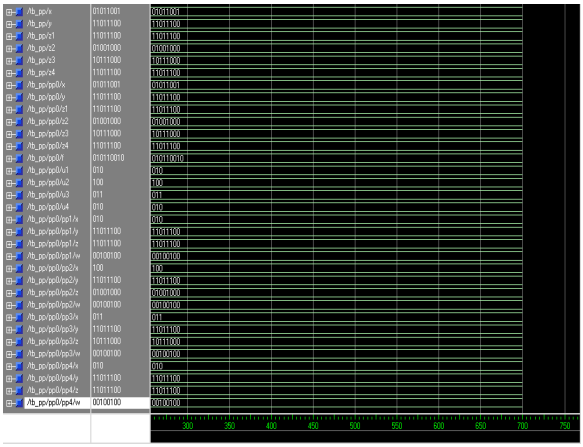
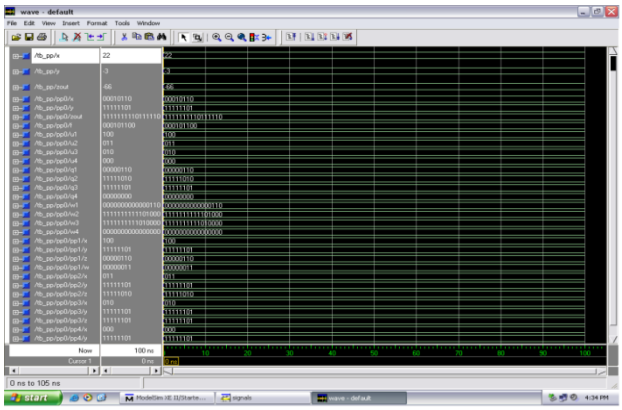


Figure 6: Pipelined operational sequence. (a) Elguibaly's operation. (b) Proposed operation.

Experimental Results:





5. CONCLUSION:

In this paper, a high-accuracy modified Booth multiplier has been proposed. In the proposed multiplier, the booth encoder has reduced the number of partial product array to half the value. The partial product generator has generated the partialproduct array bits. The compression tree has generated the final output product bits. The adder which is used in the implementationof multiplier is Carry Look ahead Adder. The compression tree along with the carry look ahead adder has reduced the hardwareoverhead and power consumption

Future Work:

The current analysis produces high accuracy for the fixed width output product which is of length 2n - bits i.e. n multiplicandand n multiplier produce 2n - bit output product. There is a further need to produce high accuracy for the fixed width of half,quarter, one by eighth and one by sixteenth of the product term bits. The above need is satisfied by means of comparator and asorting network which uses minimum number of logic gates.

REFERENCES :

[1] J.W Chen, R.H Yao, and W.J Wu, "Efficient 2n + 1 Modulo Multipliers," IEEE Transactions on Very Large Scale Integration (VLSI) systems, V. 19, NO. 12, 2011.

[2] Yuan-Ho Chen, Chung-Yi Li, and Tsin-Yuan Chang, *IEEE*, "Area-Effective and Power-Efficient Fixed-Width Booth Multipliers Using GeneralizedProbabilistic Estimation Bias," IEEE Journal on Emerging and Selected topics in Circuits and Systems, V. 1, NO. 3, 2011.

About the Authors:

(1) HARILA L.J,
M.tech, srikottamtulasireddy memorial college of engineering

(2) K. DURGA PRASAD
M.tech,(Phd) Associate professor,
Sri kottamtulasireddy memorial college of engineering

Nonsplit Dom Strong Domination Number Of A Graph

[1,]G. Mahadevan, [2,]Selvam Avadayappan, [3,]M. Hajmeeral

[1,]Department of Mathematics Gandhigram Rural Institute- Deemed University Gandhigram – 624 302
[2,]Department of Mathematics V.H.N.S.N. College, Virudhunagar-626 001
[3,]Department of Mathematics B.S.Abdur Rahman University Vandalur, Chennai-600048

Abstract

A subset D of V is called a dom strong dominating set if for every $v \in V - D$, there exists $u_1, u_2 \in D$ such that u_1v, $u_2v \in E(G)$ and deg $(u_1) \geq$ deg (v). The minimum cardinality of a dom strong dominating set is called dom strong domination number and is denoted by γ_{dsd}. In this paper, we introduce the concept of nonsplit dom strong domination number of a graph. A dom strong dominating set D of a graph G is a nonsplit dom strong dominating set (nsdsd set) if the induced subgraph <V-D> is connected. The minimum cardinality taken over all the nonsplit dom strong dominating sets is called the nonsplit dom strong domination number and is denoted by $\gamma_{nsdsd}(G)$. Also we find the upper bound for the sum of the nonsplit dom strong domination number and chromatic number and characterize the corresponding extremal graphs.

1. Introduction

Let G = (V, E) be a simple undirected graph. The degree of any vertex u in G is the number of edges incident with u and is denoted by d(u). The minimum and maximum degree of G is denoted by $\delta(G)$ and $\Delta(G)$ respectively. A path on n vertices is denoted by P_n. The graph with $V(B_{n,n})=\{u_1,u_2,u_3,\ldots u_n,v_1,v_2,v_3,\ldots v_n\}$ and $E(B_{n,n}) =\{uu_i,vv_i,uv: 1 \leq i \leq n\}$ is called the n-bistar and is denoted by $B_{n,n}$. The graph with vertex set $V(H_{n,n}) = \{v_1,v_2,v_3,\ldots v_n,u_1,u_2,u_3\ldots u_n\}$ and the edge set $E(H_{n,n}) = \{ v_i, u_j, 1\leq i \leq n, n-i+1 \leq j \leq n\}$ is denoted by $H_{n,n}$. The corona of two graphs G_1 and G_2 is the graph $G = G_1 o G_2$ formed from one copy of G_1 and $|V(G_1)|$ copies of G_2 where the i^{th} vertex of G_1 is adjacent to every vertex in the i^{th} copy of G_2.

The Cartesian graph product $G = G_1 \square G_2$ is called the graph product of graphs G_1 and G_2 with disjoint vertex sets V_1 and V_2 and edge set X_1 and X_2 is the graph with the vertex set $V_1 \times V_2$ and $u = (u_1,u_2)$ adjacent with $v = (v_1,v_2)$ whenever [$u_1 = v_1$ and u_2 adjacent to v_2] or [$u_2 = v_2$ and u_1 adjacent v_1]. The book graph B_m is defined as the graph cartesian product $S_{m+1} \times P_2$, where S_m is a star graph and P_2 is the path graph on two nodes. The friendship graph or (Dutch windmill graph) F_n is constructed by joining n copies of the cycle C_3 with a common vertex. The ladder graph can be obtained as the Cartesian product of two path graphs, one of which has only one edge. A graph G is a called a (n x m) flower graph if it has n vertices which form an n-cycle and n-sets of m-2 vertices which form m-cycles around the n-cycle so that each m-cycle uniquely intersects with n-cycle on a single edge.

A (n, k)- banana tree is defined as a graph obtained by connecting one leaf of each of n copies of an k-star graph root vertex that is distinct from all the stars. Recently many authors have introduced some new parameters by imposing conditions on the complement of a dominating set. For example, Mahadevan et.al [14] introduced the concept of complementary perfect domination number.

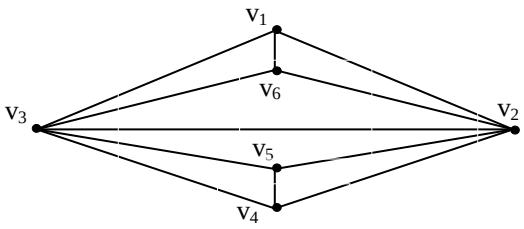
A subset S of V of a non-trivial graph G is said to be an complementary perfect dominating set if S is a dominating set and <V-S> has a perfect matching. The concept of nonsplit domination number of a graph was defined by Kulli and Janakiram [5]. A dominating set D of a graph G is a nonsplit dominating set if the induced subgraph <V-D> is connected. The non split domination number γ_{ns} (G) of G is minimum cardinality of a nonsplit dominating set. The concept of dom strong domination number of the graph is defined in [16]. Double domination introduced by Haynes[18] serves as a model for the type of fault tolerance where each computer has access to atleast two fileservers and each of the fileservers has direct access to atleast one backup fileserver. Sampathkumar and Pushpalatha [15] have introduced the concept of strong weak domination in graphs. A combination of the concepts of double domination and strong weak domination is the concept of domination strong domination where in for every vertex outside the dominating set, there are two vertices inside the dominating set, one of which dominates the outside vertex and the other strongly dominates the outside vertex. In this paper we introduce the concept of non split dom strong domination number of a graph.

2. Non Split Dom Strong Domination Number

Definition 2.1

A dom strong dominating set D of a graph G is a nonsplit dom strong dominating set (nsdsd set) if the induced subgraph <V-D> is connected. The minimum cardinality taken over all the nonsplit dom strong dominating sets is called the non split dom strong domination number and is denoted by $\gamma_{nsdsd}(G)$.

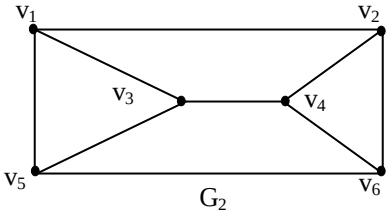
Examples 2.2



G_1

Figure 2.1

In the figure 2.1, $D_1 = \{ v_1, v_2, v_5 \}$ form the nonsplit dom strong dominating set of G_1.



G_2

Figure 2.2

In the figure 2.2, $D_1 = \{ v_1, v_2, v_5, v_6 \}$and $D_2 = \{ v_1, v_2, v_4, v_5, v_6 \}$ form the nonsplit dom strong dominating set of G_2. The minimum cardinality is taken as the nonsplit dom strong domination number for G_2 is 4.

Basic Observations 2.3

The nonsplit dom strong domination number of some of the standard classes of graphs are given below

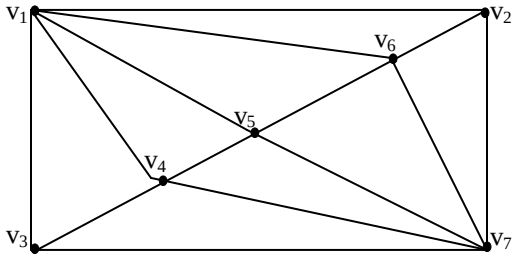
1. $\gamma_{nsdsd}(P_n) = n-1$ for $n \geq 4$,where P_n is a path on n vertices.
2. $\gamma_{nsdsd}(C_n) = n-1$ for $n \geq 4$,where C_n is a cycle on n vertices
3. $\gamma_{nsdsd}(K_n) = 2$ for $n \geq 3$, where K_n is a complete graph on n vertices.
4. $\gamma_{nsdsd}(K_{1,n}) = n+1$, where $K_{1,n}$ is a star graph.
5. $\gamma_{nsdsd}(K_{m,n}) = m + n - 1$ for $m \neq n$ where $K_{m,n}$ is a bipartite graph on m+n vertices
6. $\gamma_{nsdsd}(K_{m,n}) = 4$ for $m = n$ where $K_{m,n}$ is a bipartite graph on m+n vertices
7. $\gamma_{nsdsd}(P) = 8$, where P is the Peterson graph.
8. $\gamma_{nsdsd}(W_n) = n-1$ where W_n is a wheel whose outer cycle has n vertices.
9. $\gamma_{nsdsd}(H_n) = n+1$ where H_n is a Helm graph.
10. $\gamma_{nsdsd}(B_{m,n}) = m+n+1$ where $B_{m,n}$ is a bistar.
11. If G is the corona $C_n o K_1$, then $\gamma_{nsdsd}(G) = 2n-2$ for $n \geq 3$
12. If G is the corona $K_n o K_1$, then $\gamma_{nsdsd}(G) = n+1$ for $n \geq 3$
13. $\gamma_{nsdsd}(B_m) = n - 1$, where B_m is a book graph.
14. $\gamma_{nsdsd}(F_n) = n - 1$, where F_n is a friendship graph.
15. $\gamma_{nsdsd}(L_n) = 2n-2$, where L_n is a ladder graph.
16. $\gamma_{nsdsd}(F_{nxm}) = n(m-1)-2$, where F_{nxm} is a flower graph.
17. $\gamma_{nsdsd}(B_{n,k}) = nk$, where $B_{n,k}$ is a banana tree.

Theorem 2.4 Let G be a graph with no isolates. Then $2 \leq \gamma_{nsdsd}(G) \leq n$ and the bounds are sharp.

Proof Since any dom strong dominating set has at least two elements and at most n elements. Hence for any nonsplit dom strong dominating set has at least two elements and at most n elements. For a star $\gamma_{nsdsd}(K_{1,n}) = n+1$ and for K_n, $\gamma_{nsdsd}(K_n) = 2$.Therefore the bounds are sharp.

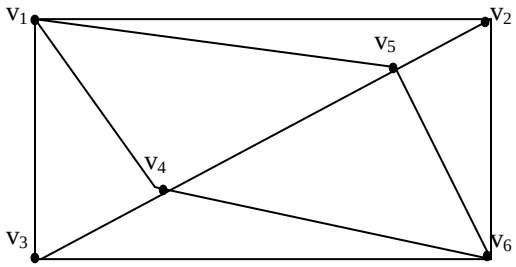
Theorem 2.5 In a graph G, if a vertex v has degree one then v must be in every nonsplit dom strong dominating set of G. That is every nonsplit dom strong dominating set contains all pendant vertices.
Proof Let D be any nonsplit dom strong dominating set of G. Let v be a pendant vertex with support say u. If v does not belong to D, then there must be two points say x,y belong to D such that x dominates v and y dominates v. Therefore x and y are adjacent to v and hence deg $v \geq 2$ which is a contradiction. Since v is a pendant vertex, so v belongs to D.

Observation 2.6 $\gamma(G) \leq \gamma_{dsd}(G) \leq \gamma_{nsdsd}(G)$ and the bounds are sharp for the graph G_3 figure 2.3



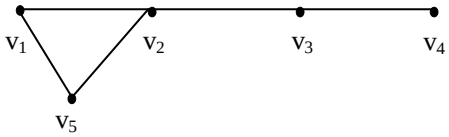
G_3
Figure 2.3

Observation 2.7 For any graph G, $\gamma_{nsdsd}(G) \geq \lceil n/(\Delta+1) \rceil$ and the bound is sharp.
Proof For any graph G, $\lceil n/(\Delta+1) \rceil \leq \gamma$ and also by observation 2.6, the theorem follows. The bound is sharp for the graph G_4 in figure 2.4.



G_4
Figure 2.4

Remark 2.8 Support of a pendant vertex need not be in a nonsplit dom strong dominating set. For the graph G_5 in figure 2.5, $\gamma_{nsdsd}(G_5) = 4$. Here $D_1 = \{ v_1, v_2 v_4, v_5 \}$is a nonsplit dom strong dominating set which does not contains the support v_3.



G_5
Figure 2.5

Observation 2.9 If H is any spanning subgraph of a connected graph G and $E(H) \subseteq E(G)$ then $\gamma_{nsdsd}(G) \leq \gamma_{nsdsd}(H)$.

Theorem 2.10 Let $G \cong C_n$ $(n \geq 5)$**.** Let H be a connected spanning subgraph of G, then $\gamma_{nsdsd}(G) = \gamma_{nsdsd}(H)$.
Proof We have $\gamma_{nsdsd}(G) = n - 1$ and also a connected spanning sub graph of G is a path. Hence <H> is a path so that $\gamma_{nsdsd}(H) = n - 1$.

Observation 2.11 For any cycle C_n and any $v \in V(C_n)$,

$$\gamma_{nsdsd}(C_n - v) = \begin{cases} 2 & \text{if } n = 3, \\ 3 & \text{if } n = 4, \\ n-2 & \text{if } n > 4. \end{cases}$$

Proof Follows from theorem 2.10.

Observation 2.12 If $G \cong K_n$ o K_1, for any complete graph K_n, then $\gamma_{dsd}(G) = \gamma_{nsdsd}(G)$.

Theorem 2.13 For any connected graph G, $\gamma_{nsdsd}(G) = n$ if and only if G is a star.

Proof If G is a star then V is the only nonsplit dom strong dominating set so that $\gamma_{nsdsd}(G) = n$. Conversely, assume that $\gamma_{nsdsd}(G) = n$. We claim that G is a star. Suppose not, let u be a vertex of a maximum degree Δ with $N(u) = \{ u_1, u_2, \ldots u_\Delta \}$. If $< N(u) >$ has an edge $e = u_i u_j$, then $V - \{ u_i \}$ is a nonsplit dom strong dominating set of cardinality n - 1, which is a contradiction. If $< N(u) >$ has no edge then G has an edge e = xy which is not incident with u such that u is adjacent to x. then V-{u} is a nonsplit dom strong dominating set of cardinality n-1 which is a contradiction. Hence G is a star.

Theorem 2.14 For any connected graph G, $\gamma_{nsdsd}(G) = 2$ if and only if there exist u and v such that deg u = deg v = Δ, then deg u and deg v $\geq$ n-2.

Proof Let there exist u and v satisfying the hypothesis. Let D = {u,v}. Let $x \in V-D$, then x is adjacent to both u and v. Since deg u = deg v = Δ, we have deg x $\leq$ deg u and deg x $\leq$ deg v, therefore D is the nonsplit dom strong dominating set. Conversely, let D = {u, v} be a nonsplit dom strong dominating set. Every point $x \in V-D$ is adjacent to both u and v. Therefore deg u $\geq$ n-2, deg v $\geq$ n-2. Also deg x $\leq$ deg u or deg v. Suppose deg u and deg v < Δ then there exists $x \in V - D$ of deg Δ. Therefore D is not a nonsplit dom strong dominating set, which is a contradiction. Hence deg u = deg v = Δ. If deg u is not equal to deg v then deg u = n-1 and deg v = n-2, which is impossible. Therefore deg u = deg v = Δ.

Theorem 2.15 Let G be a graph without isolates and let there exists a γ_{nsdsd} set which is not independent. Then $\gamma(G) + 1 \leq \gamma_{nsdsd}(G)$.

Proof Let D be a γ_{nsdsd} set which is not independent. Let $x \in D$ be such that x is adjacent to some point of D. If $N(x) \cap (V-D) = \Phi$, then as G has no isolates $N(x) \cap D \neq \Phi$. Hence D - { x } is a dominating set. Therefore $\gamma(G) \leq |D-\{ x \}| = \gamma_{nsdsd}(G) - 1$. If $N(x) \cap (V-D) \neq \Phi$. Then for any $y \in N(x) \cap (V-D)$ there exists $z \in D$ such that z is adjacent to y. As x is adjacent to some point of D, D -{ x } is a dominating set. Therefore $\gamma(G) \leq |D-\{ x \}| \leq \gamma_{nsdsd}(G) - 1$.The bound is sharp. $\gamma(K_n) = 1$and $\gamma_{nsdsd}(K_n) = 2$.

Theorem 2.16 $\gamma_{nsdsd}(G) \geq \lceil 2n/(\Delta + 2) \rceil$

Proof Every vertex in V-D contributes two to degree sum of vertices of D. Hence $2|V-D| \leq \sum_{u \in D} d(u)$ where D is a nonsplit dom strong dominating set, so that $2|V-D| \leq \gamma_{nsdsd} \Delta$ which implies $2(|V| - |D|) \leq \gamma_{nsdsd} \Delta$. Therefore $2n - 2\gamma_{nsdsd} \leq \gamma_{nsdsd} \Delta$, which implies $\gamma_{nsdsd}(\Delta+2) \geq 2n$. Hence $\gamma_{nsdsd} \geq \lceil 2n/(\Delta + 2) \rceil$. The bounds are sharp. For K_4, $\gamma_{nsdsd}(K_4) = 2$. $\gamma_{nsdsd}(G) = \lceil 2n/(\Delta + 2) \rceil = 2$.

3. Relation Between The Nonsplit Dom Strong Domination Number And Chromatic Number :

Recently many authors have studied the problem of obtaining an upper bounds for the sum of the one domination parameter and graph theory parameter and characterize the corresponding extremal graphs. In [11], Paulraj Joseph J and Arumugam S proved that $\gamma + k \leq p$. In [12], Paulraj Joseph J and Arumugam S proved that $\gamma + \chi \leq p + 1$. They also characterized the class of graphs for which the upper bound is attained. They also proved similar results for γ and γ_t. In [13], Paulraj Joseph J and Mahadevan G proved that $\gamma_{cc} + \chi \leq 2n-1$ and characterized the corresponding extremal graphs. In [6], Mahadevan G, proved that $\gamma_{pr} + \chi \leq 2n-1$ and characterized the corresponding extremal graphs. He also proved that $\gamma_{ipr} + \chi \leq 2n - 2$ and characterized the corresponding extremal graphs. In [14], Paulraj Joseph J, Mahadevan G and Selvam A. introduced the concept of complementary perfect domination number γ_{cp} and proved that $\gamma_{cp} + \chi \leq 2n-2$, and characterized the corresponding extremal graphs. They also obtained the similar results for the induced complementary perfect domination number and chromatic number of a graph. We find the upper bound for the sum of the nonsplit dom strong domination number and chromatic number and characterize the corresponding extremal graphs

Notations 3.1

$P_k(m_1,m_2)$ where $k \geq 2$, $m_1,m_2 \geq 1$ be the graph obtained by identifying centers of the stars $K_{1,.m1}$ and $K_{1,.m2}$ at the ends of P_K respectively. The graph $C_3(m_1, m_2, 0)$ is obtained from C_3 by identifying the centers of stars $K_{1,.m1}$ and $K_{1,.m2}$ at any two vertices of C_3. The graph $K_n(m_1, m_2, m_3, m_4, m_5,\ldots\ldots m_n)$ denote the graph obtained from K_n by pasting m_1 edges to any one vertex u_i of K_n, m_2 edges to any vertex u_j of K_n, for $i \neq j$, m_3 edges to any vertex u_k for $i \neq j \neq k$, m_4 edges to u_l $i \neq j \neq k \neq l$,…….m_n edges to all the distinct vertices of K_n. $C_n(P_k)$ is the graph obtained from C_n by attaching the end vertex of P_k to any one vertices of C_n. $K_n(P_k)$ is the graph obtained from K_n by attaching the end vertex of P_k to any one vertices of K_n.

Theorem 3.2 For any graph G, $\gamma_{nsdsd}(G) \leq n$.

Theorem 3.3 For any connected graph G, $\chi(G) \leq \Delta(G) + 1$.

Theorem 3.4 For any graph, $\gamma_{nsdsd}(G) + \chi(G) \leq 2n$ and equality holds if and only if $G \cong K_2$.

Proof By theorem 3.2 and 3.3, it follows that $\gamma_{nsdsd}(G) + \chi(G) \leq n + \Delta + 1 \leq n + n - 1 + 1 \leq 2n$. Now we assume that $\gamma_{nsdsd}(G) + \chi(G) = 2n$. This is possible only if $\gamma_{nsdsd}(G) = n$ and $\chi(G) = n$. Since $\chi(G) = n$, G is complete. But for complete graph, $\gamma_{nsdsd}(G) = 2$. Hence $G \cong K_2$. Converse is obvious.

Theorem 3.5 For any graph G, $\gamma_{nsdsd}(G) + \chi(G) = 2n-1$ if and only if $G \cong P_3, K_3$.

Proof If G is either P_3 or K_3, then clearly $\gamma_{nsdsd}(G) + \chi(G) = 2n-1$. Conversely, assume that $\gamma_{nsdsd}(G) + \chi(G) = 2n-1$. This is possible only if $\gamma_{nsdsd}(G) = n$ and $\chi(G) = n-1$ (or) $\gamma_{nsdsd}(G) = n-1$ and $\chi(G) = n$.

Case (i) $\gamma_{nsdsd}(G) = n$ and $\chi(G) = n-1$. Since $\gamma_{nsdsd}(G) = n$, G is a star. Therefore n=3. Hence $G \cong P_3$. On increasing the degree we get a contradiction.

Case (ii) $\gamma_{nsdsd}(G) = n-1$ and $\chi(G) = n$. Since $\chi(G) = n$, G is complete. But for K_n, $\gamma_{nsdsd}(G) = 2$, so that $n = 3$. Hence $G \cong K_3$.

Theorem 3.6 For any graph G, $\gamma_{nsdsd}(G) + \chi(G) = 2n-2$ if and only if $G \cong K_{1,3}, K_3(P_2), K_4$.

Proof If G is any one of the following graphs $K_{1,3}, K_3(P_2), K_4$, then clearly $\gamma_{nsdsd}(G) + \chi(G) = 2n-2$. Conversely, assume that $\gamma_{nsdsd}(G) + \chi(G) = 2n-2$. This is possible only if $\gamma_{nsdsd}(G) = n$ and $\chi(G) = n-2$ (or) $\gamma_{nsdsd}(G) = n-1$and $\chi(G) = n-1$ (or) $\gamma_{nsdsd}(G) = n-2$ and $\chi(G) = n$

Case (i) $\gamma_{nsdsd}(G) = n$ and $\chi(G) = n-2$. Since $\gamma_{nsdsd}(G) = n$, G is a star. Therefore n = 4. Hence $G \cong K_{1,3}$. On increasing the degree we get a contradiction.

Case (ii) $\gamma_{nsdsd}(G) = n-1$ and $\chi(G) = n-1$. Since $\chi(G) = n-1$, G contains a clique K on n-1 vertices. Let $S = \{v\}$ be the vertex other than the clique K_{n-1}.Then v is adjacent to u_i for some i in K_{n-1}. Then $\{v_1, u_i, u_j\}$ is a γ_{nsdsd} set. Hence n = 4. Therefore $K = K_3$. If $d(v_1) = 1$ then $G \cong K_3(P_2)$. On increasing the degree of v_1, no graph exists.

Case (iii) $\gamma_{nsdsd}(G) = n-2$ and $\chi(G) = n$. Since $\chi(G) = n$, $G \cong K_n$. But for K_n, $\gamma_{nsdsd}(G) = 2$. Therefore n = 3. Hence $G \cong K_4$.

Theorem 3.7 For any graph G, $\gamma_{nsdsd}(G) + \chi(G) = 2n-3$ if and only if $G \cong K_{1,4}, K_3(P_3), K_3(2), K_3(P_2, P_2, 0), K_5$, or any one of the graphs in the figure 3.1

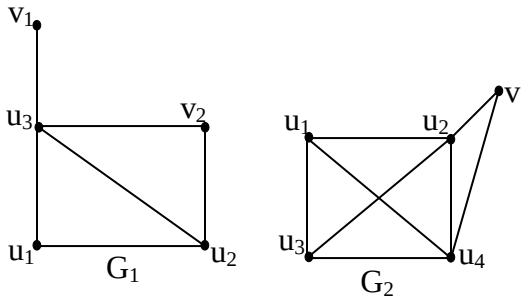


Figure 3.1

Proof If G is any one of the above graphs then clearly $\gamma_{nsdsd}(G) + \chi(G) = 2n-3$. Conversely, assume that $\gamma_{nsdsd}(G) + \chi(G) = 2n-3$. This is possible only if $\gamma_{nsdsd}(G) = n$, $\chi(G) = n-3$ (or) $\gamma_{nsdsd}(G) = n-1$, $\chi(G) = n-2$ (or) $\gamma_{nsdsd}(G) = n-2$, $\chi(G) = n-1$ (or) $\gamma_{nsdsd}(G) = n-3$, $\chi(G) = n$.

Case (i) $\gamma_{nsdsd}(G) = n$ and $\chi(G) = n-3$. Since $\gamma_{nsdsd}(G) = n$, G is a star. Therefore n = 5. Then $G \cong K_{1,4}$. On increasing the degree no new graph exists.
Case (ii) $\gamma_{nsdsd}(G) = n-1$ and $\chi(G) = n-2$. Since $\chi(G) = n-2$, G contains a clique K on n-2 vertices. Let $S = \{v_1, v_2\}$ be the vertices other than the clique K_{n-2} then the possible cases are $\langle S \rangle = K_2$ or $\overline{K_2}$.
Subcase (i) Let $\langle S \rangle = K_2$. Since G is connected, either v_1 or v_2 is adjacent to u_i for some i in K_{n-2}, then $\{v_1, v_2, u_i, u_j\}$ is a γ_{nsdsd} set so that n = 5. Hence $K = K_3$. If $d(v_1) = 2$ and $d(v_2) = 1$, then $G \cong K_3(P_3)$. On increasing the degree, no graph exists.
Subcase (ii) Let $\langle S \rangle = \overline{K_2}$. Since G is connected, v_1 and v_2 is adjacent to u_i for some i in K_{n-2}. Then $\gamma_{nsdsd}(G) = 4$, so that $K = K_3$. If $d(v_1) = d(v_2) = 1$, then $G \cong K_3(2)$. If $d(v_1)=1$ and $d(v_2) = 2$ then $G \cong G_1$. If v_1 is adjacent to u_i and v_2 adjacent to u_j for some $i \neq j$ in K_{n-2} then $\gamma_{nsdsd}(G) = 4$. Hence $K = K_3$. If $d(v_1) = d(v_2) = 1$, then $G \cong K_3(P_2, P_2, 0)$. On increasing the degree, no graph exists.

Case (iii) $\gamma_{nsdsd}(G) = n-2$ and $\chi(G) = n-1$. Since $\chi(G) = n-1$, G contains a clique K on n-1 vertices. Let $S = \{v\}$ be the vertex other than the clique K_{n-1}.If v is adjacent to u_i for some i in K_{n-1}, then $\gamma_{nsdsd}(G) = 3$. Hence n = 4. Therefore $K = K_4$. If $d(v) = 1$, then $G \cong K_4(P_2)$. If $d(v) = 2$, then $G \cong G_2$. On increasing the degree, no new graph exists.
Case (iv) $\gamma_{nsdsd}(G) = n-3$ and $\chi(G) = n$. Since $\chi(G) = n$, $G \cong K_n$. But for complete Graph K_n, $\gamma_{nsdsd}(G) = 2$ so that n = 5. Therefore $G \cong K_5$.

Theorem 3.8 For any graph G, $\gamma_{nsdsd}(G) + \chi(G) = 2n - 4$ if and only if $G \cong K_{1,5}$, $K_3(3)$, $C_4(P_2)$, $S(K_{1,3})$, $K_3(P_3)$, $C_3(1,1,1)$, $K_3(2,1,0)$, $K_4(2)$, $K_4(P_2, P_2, 0,0)$, $K_5(P_2)$. K_6, or any one of the graphs given in the figure 3.2

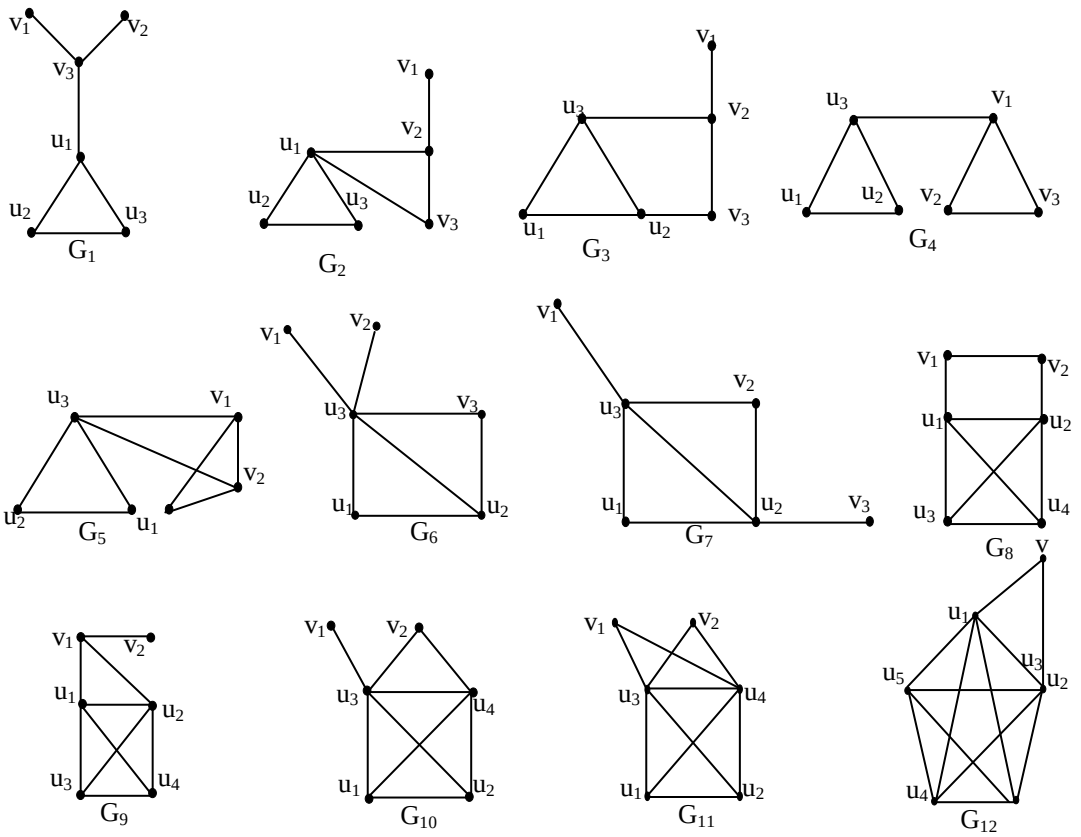


Figure 3.2

Proof Assume that $\gamma_{nsdsd}(G) + \chi(G) = 2n-4$. This is possible only if $\gamma_{nsdsd}(G) = n$ and $\chi(G) = n-4$ (or) $\gamma_{nsdsd}(G) = n-1$ and $\chi(G) = n-3$ (or) $\gamma_{nsdsd}(G) = n-2$ and $\chi(G) = n-2$ (or) $\gamma_{nsdsd}(G) = n-3$ and $\chi(G) = n-1$ (or) $\gamma_{nsdsd}(G) = n-4$ and $\chi(G) = n$.

Case (i) $\gamma_{nsdsd}(G) = n$ and $\chi(G) = n-4$. Since $\gamma_{nsdsd}(G) = n$, G is a star. Therefore n = 6. Then $G \cong K_{1,5}$. On increasing the degree, we get a contradiction.

case (ii) $\gamma_{nsdsd}(G) = n-1$ and $\chi(G) = n-3$.

Since $\chi(G) = n-3$, G contains a clique K on n-3 vertices Let $S = \{v_1, v_2, v_3\}$ be the vertices other than the clique K_{n-3} then $<S> = P_3, K_3, \overline{K_3}, K_2 \cup K_1$

Subcase (i) Let $< S > = P_3$. Since G is connected, the following are the possible cases (i) there exist a vertex u_i of K_{n-3} which is adjacent to any one of end vertices (ii) there exist a vertex u_i of K_{n-3} which is adjacent to other than end vertices. If there exist a vertex u_i of K_{n-3} which is adjacent to any one of end vertices, then $\gamma_{nsdsd}(G) = 5$. Hence n = 6. Therefore $K = K_3$. If $d(v_1) = 2$ and $d(v_2) = d(v_3) = 1$ then $G \cong K_3(P_4)$. If u_i is adjacent to v_2 which is not a pendant vertices then $\gamma_{nsdsd}(G) = 5$. Hence n = 6. Therefore $K = K_3$. If $d(v_1) = d(v_3) = 1$ and $d(v_2) = 3$ then $G \cong G_1$. If $d(v_3) = 2$ and $d(v_1) = 1$ and $d(v_2) = 3$ then $G \cong G_2$. If $d(v_1) = 1$ and $d(v_2) = 3$ and $d(v_3) = 2$ then $G \cong G_3$.

Subcase (ii) Let $< S > = K_3$. Since G is connected, there exist a vertex u_i of K_{n-3} adjacent to anyone of $\{v_1, v_2, v_3\}$. Without loss of generality let v_1 be adjacent to u_i, then $\gamma_{nsdsd}(G) = 5$. Therefore $K=K_3$. If $d(v_1) =3$ and $d(v_2) = d(v_3) = 2$ then $G \cong G_4$. If $d(v_1) = 3$ and $d(v_2) = 3$ and $d(v_3) = 2$ then $G \cong G_5$. On increasing the degree we get a contradiction.

Subcase (iii) Let $< S > = \overline{K_3}$. Since G is connected, let all the vertices of $\overline{K_3}$ be adjacent to vertex u_i. Then $\gamma_{nsdsd}(G) = 5$. Hence n = 6. Therefore $K = K_3$. Let u_1, u_2, u_3 be the vertices of K_3. Let all the three vertices of $\overline{K_3}$ adjacent to u_1. Then $G \cong K_3(3)$. If $d(v_3) = 2$ and $d(v_1) = 1$ and $d(v_2) = 1$ then $G \cong G_6$.On increasing the degree, we get a contradiction. If two vertices of $\overline{K_3}$ are adjacent to u_i and the third vertex adjacent to u_j for some $i \neq j$, then $\gamma_{nsdsd}(G) = 5$. Hence n = 6. Therefore $K = K_3$. Let u_1, u_2, u_3 be the vertices of K_3. Then $G \cong K_3(2, 1, 0)$. If $d(v) = 1$ and $d(v_2) = 2$ and $d(v_3) = 1$ then $G \cong G_7$. On increasing the degree, we get a contradiction. If all the three vertices of $\overline{K_3}$ are adjacent to three distinct vertices of K_{n-3} say u_i, u_j, u_k for $i \neq j \neq k$, then $\gamma_{nsdsd}(G) = 5$. Hence n = 6. Therefore $K = K_3$. Let u_1, u_2, u_3 be the vertices of K_3. Then $G \cong K_3(1,1,1)$. On increasing the degree, we get a contradiction.

Subcase (iv) Let $< S > = K_2 \cup K_1$. Since G is connected, there exist a vertex u_i of K_{n-3} which is adjacent to anyone of $\{v_1, v_2\}$ and v_3. Then $\gamma_{nsdsd}(G) = 4$. Hence n = 6. Therefore $K = K_2$, so that $G \cong S(K_{1,3})$. On increasing the degree, we get a contradiction. Let there exist a vertex u_i of K_{n-3} be adjacent to any one of $\{v_1, v_2\}$ and u_j for some $I \neq j$ in K_{n-3} adjacent to v_3. Hence $\gamma_{nsdsd}(G) = 4$, so that n = 5. Therefore $K = K_2$, which is a contradiction.

If G does not contain a clique K on n-3 vertices, then it can be verified that no new graph exist.

Case (iii) $\gamma_{nsdsd}(G) = n-2$ and $\chi(G) = n-2$. Since $\chi(G) = n-2$, G contains a clique K on n-2 vertices. Let $S = \{v_1, v_2, v_3, v_4\}$ be the vertices other than the clique K_{n-2} then the possible cases are $<S> = K_2, \overline{K_2}$.

Subcase (i) Let $<S> = K_2$. Since G is connected, either v_1 or v_2 is adjacent to u_i for some i in K_{n-2}. Then $\gamma_{nsdsd}(G) = 4$ so that n = 6. Therefore $K= K_4$. Let u_1, u_2, u_3 be the vertices of K_3. Therefore $G \cong K_4(P_3)$. On increasing the degree, then $G \cong G_8, G_9$.

Subcase (ii) Let $<S> = \overline{K_2}$. Since G is connected, both v_1 and v_2 adjacent to u_i for some i in K_{n-2}. Then $\gamma_{nsdsd}(G) = 4$ so that n = 6.Therefore $K= K_4$. Let u_1, u_2, u_3, u_4 be the vertices of K_4. Therefore $G \cong K_4(2)$. If $d(v_1) = 1$ and $d(v_2) = 2$ then $G \cong G_{10}$. On increasing the degree, we get a contradiction. If the two vertices are adjacent to two distinct vertices of K_{n-2}, then $\gamma_{nsdsd}(G) = 4$. Hence n = 6.Theefore $K = K_4$. Then $G \cong K_4(P_2, P_2, 0, 0)$. If $d(v_1) = 2$ and $d(v_2) = 1$ then $G \cong G_{11}$. If $d(v_1) = 2$ and $d(v_2) = 2$ then $G \cong G_{12}$. On increasing the degree, we get a contradiction.

Case (iv) $\gamma_{nsdsd}(G) = n-3$ and $\chi(G) = n-1$.Since $\chi(G) = n-1$, G contains a clique K on n-1 vertices. Let the vertex v_1 is adjacent to u_i for some i in K_{n-1}. Therefore $\gamma_{nsdsd}(G) = 3$, hence n = 6. Therefore $K = K_5$. Then $G \cong K_5(P_2)$. If d (v) = 2 then $G \cong G_{15}$.On increasing the degree, we get a contradiction.

Case (v) Let $\gamma_{nsdsd}(G) = n-4$ and $\chi(G) = n$. Since $\chi(G) = n$, $G \cong K_n$. But for K_n, $\gamma_{nsdsd}(G) = 2$, so that n = 6. Therefore $G \cong K_6$.

References

[1]. Acharya.B.D, and Walikar.H.B,(1979): On Graphs having unique minimum dominating sets, Graph theory news letter, 8.2.

[2]. Acharya.B.D, (1980): The strong domination number of a graph and related concepts, J.Math.Phys.Sci,14 pp 471-475.

[3]. Harary F(1972): Graph theory , Addison Wesley Reading Mass.

[4]. John Clark and Derek Allan Holton (1995): A First Look at Graph Theory, Allied Publisher Ltd .

[5]. Kulli.V.R and Janakiram. B (2000): The nonsplit domination number of a graph, Indian J. Pure and Appl. Math., 31(4) pp 441-447.

[6]. Mahadevan G. (2005): On Domination Theory and related topics in graphs, Ph. D, Thesis, Manonmaniam Sundaranar University, Tirunelveli.
[7]. Mahadevan G, Selvam A, Hajmeeral M (2009): On efficient domination number and chromatic number of a graph I , International Journal of Physical Sciences, vol 21(1)M, pp1-8.
[8]. Mahadevan G, Selvam Avadayappan. A and Hajmeeral M(2010): "Further results on dom strong domination number of the graph", International Journal of Algorithms, Computing and Mathematics, vol 3,no 1 , pp 23-30.
[9]. Mahadevan G, Selvam Avadayappan. A and Hajmeeral.M and Latha Venkateswari.U, (2010): "Further characterization of connected efficient domination number of a graph", International Journal of Combinatorial Graph theory and applications, Vol 3, no 1, Jan-Jun 2010, pp 29-39.
[10]. Namasivayam. P (2008): Studies in strong double domination in graphs, Ph.D., thesis, Manonmaniam Sundaranar University, Tirunelveli, India.
[11]. Paulraj Joseph J and Arumugam S (1992): Domination and connectivity in graphs, International Journal of management and systems,vol 15 No.1, 37-44.
[12]. Paulraj Joseph J. and Arumugam S. (1997): Domination and colouring in graphs, International Journal of Management and Systems, Vol.8 No.1, 37-44.
[13]. Paulraj Joseph J. and Mahadeven G. (2002): Complementary connected domination number and chromatic number of a graph Mathematical and Computational models, Allied Publications, India.342-349.
[14]. Paulraj Joseph J. and Mahadevan G and Selvam A (2006): On Complementary perfect domination number of a graph, Acta Ciencia Indica Vol. XXXI M, No.2, 847, (An International Journal of Physical Sciences).
[15]. Sampathkumar E and Puspalatha.L (1996): Strong weak domination and domination balance in a graph, Discrete math. 161, pp 235-242.
[16]. Swaminathan.V, et.al, (2005): Dom-strong domination and dsd-domatic number of a graph, Proceedings of the national conference on "The emerging trends in Pure and Applied Mathematics", St.Xavier's College, Palayamkottai, pp 150-153.
[17]. Tersa W.Haynes, Stephen T.Hedetniemi and Peter J.Slater(1998): Fundamentals of domination in graphs, Marcel Dekker, Reading mass.
[18]. Teresa W. Hayens (2001): Induced Paired domination in graphs, Arts Combin.57, 111-128.
[19]. Tersa W.Haynes, Stephen T.Hedetniemi and Peter J.Slater(1998): Domination in graphs,Advanced Topics, Marcel Dekker, Reading mass.

Traffic Sign Recognition

[1,] Mr. Chetan J. Shelke, [2,]Dr. Pravin Karde
[1,]Department Of Information Technology P. R. Patil College of Engg. & Tech
[2,]Department Of Information Technology.
Govt Polytechnic Amravati

Abstract

Traffic sign recognition is a difficult task if aim is at detecting and recognizing signs in images captured from unfavorable environments. Complex background, weather, shadow, and other lighting-related problems may make it difficult to detect and recognize signs in the rural as well as the urban areas. Two major problems exist in the whole detection process. Road signs are frequently occluded partially by other vehicles. Many objects are present in traffic scenes which make the sign detection hard (pedestrians, other vehicles, buildings and billboards may confuse the detection system by patterns similar to that of road signs). Color information from traffic scene images is affected by varying illumination caused by weather conditions, time (day night) and shadowing (buildings)

1. Introduction

Traffic sign recognition is important for driver assistant systems, automatic vehicles, and inventory purposes. The best algorithm will be the one that yields the best global results throughout the whole recognition process, which comprises three stages: 1) segmentation; 2) detection; and 3) recognition. Researchers have developed vision-based techniques for traffic monitoring, traffic-related parameter estimation, driver monitoring, and intelligent vehicles, etc. [1]. Traffic sign recognition (TSR) is an important basic function of intelligent vehicles [2], and TSR problems have attracted attention of many research groups since more than ten years ago. Traffic sign recognition is part of the general case of Pattern Recognition. Major problem in pattern recognition is the difficulty of constructing characteristic patterns (templates). This is because of the large variety of the features being searched in the images, such as people faces, cars, etc. On the contrary, traffic signs a) are made with vivid and specific colors so as to attract the driver's attention and to be distinguished from the environment b) are of specific geometrical shapes (triangle, rectangle, circle - ellipse) and c) for each sign there is a specific template. It is therefore rather easy to develop an algorithm in such a way that the computer has "a priori knowledge" of the objects being searched in the image.

2. Related Work

Traffic signs are normally classified according to their color and shape and should be designed and positioned in such a way that they can easily be noticed while driving. Inventory systems must take advantage of these characteristics. However, various questions need to be taken into account in traffic sign-recognition system. For example, the object's appearance in an image depends on several aspects, such as outdoor lighting condition. In addition, deterioration of a traffic sign due to aging or vandalism affects its appearance, whereas the type of sheeting material used to make traffic signs may also cause variations. These problems particularly affect the segmentation step [3], which is usually the first stage in high-level detection and recognition systems. Segmentation can be carried out using color information or structural information. Many segmentation methods have been reported in the literature since the advent of digital image processing. Detection and recognition are two major steps for determining types of traffic signs [4]. Detection refers to the task of locating the traffic signs in given images. It is common to call the region in a given image that potentially contains the image of a traffic sign the region of interests (ROI). Taking advantages of the special characteristics of traffic signs, TSR systems typically rely on the color and geometric information in the images to detect the ROIs. Hence, color segmentation is common to most TSR systems, so are edge detection [5] and corner detection techniques [6]. After identifying the ROIs, we extract features of the ROIs, and classify the ROIs using the extracted feature values. Researchers have explored several techniques for classifying the ideographs, including artificial neural networks (ANNs) [7], template matching [8], chain code [9], and matching pursuit methods [10]. Detection and recognition of traffic signs become very challenging in a noisy environment. Traffic signs may be physically rotated or damaged for different reasons. View angles from the car-mounted cameras to the traffic signs may lead to artificially rotated and distorted

images. External objects, such as tree leaves, may occlude the traffic signs, and background conditions may make it difficult to detec ..0traffic signs. Bad weather conditions may have a detrimental effect on the quality of the images.
The traffic sign-recognition system which was described in detail in [11] consists of four stages as shown in Figure 1.

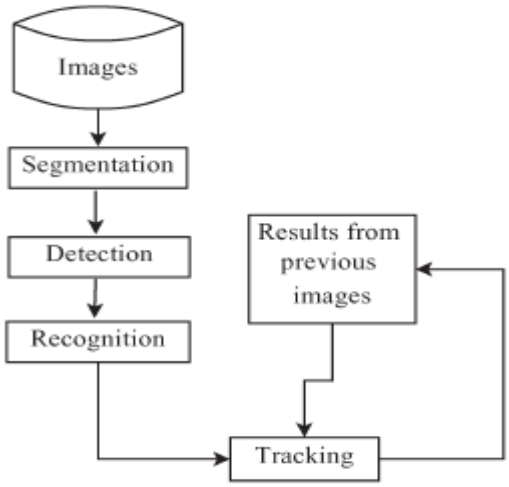


Fig. 1: Traffic sign recognition system.

Segmentation: This stage extracts objects from the background, which are, in this case, traffic signs using color information.

Detection: Here, potential traffic signs are located through shape classification.
Recognition: Traffic sign identification is effected using SVMs.
Tracking: This stage grouped multiple recognitions of the same traffic sign.

3. Region Of Interest (Roi) Detection

Transportation engineers design traffic signs such that people can recognize them easily by using distinct colors and shapes for the signs. Many countries use triangles and circles for signs that carry warning and forbidding messages, respectively. These signs have thick and red borders for visibility from apart. Hence, we may use color and shape information for detecting traffic signs.

4. Color Segmentation

Identifying what pixels of the images are red is a special instance of the *color segmentation* problems. This task is not easy because images captured by cameras are affected by a variety of factors, and the "red" pixels as perceived by human may not be encoded by the same pixel values all the time.Assuming no directly blocking objects, lighting conditions affect the quality of the color information the most. Weather conditions certainly are the most influential factor. Nearby buildings or objects, such as trees, may also affect quality of the color information because of their shadows. It is easy to obtain very dark images, e.g., the middle image in Figure 2, when we are driving in the direction of the sun.

Fig. 2: Selected "hard" traffic signs. The left sign did not face the camera directly, and had a red background. The middle picture was taken in the dusk. The signs in the rightmost image were in the shadow.

As a consequence, "red" pixels can be embodied in a range of values. Hence, it is attempted to define the range for the red color. We can convert the original image to a new image using a pre-selected formula. Let *Ri*, *Gi*, and *Bi* be the red, green, and blue component of a given pixel in the original image. We encode the pixels of the new image by *Ro*, *Go*, and *Bo*. Based on results of a few experiments, we found that the following conversion most effective: $Ro = \max(0, (Ri - Gi) +$

($Ri - Bi$)), $Go = 0$, and $Bo = 0$. After the color segmentation step, only pixels whose original red components dominate the other two components can have a nonzero red component in the new image most of the time.

5. Region Of Interests

Then the red pixels are grouped into separate objects, apply the *Laplacian of Gaussian* (LoG) edge detector to this new image, and use the 8-connected neighborhood principle for determining what pixels constitute a connected object. We consider any red pixels that are among the 8 immediate neighbors of another red pixel *connected*. After grouping the red pixels, we screen the object based on four features to determine what objects may contain traffic signs. These features are areas, height to width ratios, positions, and detected corners of the objects. According to the government's decrees for traffic sign designs, all traffic signs must have standard sizes. Using camera, which is set at a selected resolution, to take pictures of warning signs from 100 meter apart, the captured image will occupy 5x4 pixels. Due to this observation, we ignore objects that contain less than 40 red pixels. We choose to use this threshold because it provided a good balance between recall and precision when we applied the *Detection* procedure to the training data. Two other reasons support our ignoring these small objects. Even if the discarded objects were traffic signs, it would be very difficult to recognize them correctly. Moreover, if they are really traffic signs that are important to our journey, they would get closer and become bigger, and will be detected shortly. The decrees also allow us to use shapes of the bounding boxes of the objects to filter the objects. Traffic signs have specific shapes, so heights and widths of their bounding boxes must also have special ratios. The ratios may be distorted due to such reasons as damaged signs and viewing angles. Nevertheless, we can still use an interval of ratios for determining whether objects contain traffic signs. Positions of the objects in the captured images play a similar role as the decrees. Except driving on rolling hills, we normally see traffic signs above a certain horizon. Due to this physical constraint and the fact that there are no rolling hills in Taiwan, we assume that images of traffic signs must appear in a certain area in the captured image, and use this constraint for filtering objects in images. We divide the bounding boxes of the objects into nine equal regions, and check whether we can detect corners in selected regions. The leftmost image in Figure 3 illustrates one of these patterns by the blue checks. More patterns are specified in the following *Detection* procedure. If none of the patterns is satisfied, chances are very low that the object could contain a triangular sign. Using this principle, system detected the rightmost four signs in Figure 3.

Fig. 3: Using Corners for identifying triangular borders.

Procedure *Detection* (Input: an image of 640x480 pixels; Output: an ROI)
Steps:

1 Color segmentation
2 Detect edges with the LoG edge detector.
3 Remove objects with less than 40 red pixels.
4 Mark the bounding boxes of the objects.
5 Remove objects whose highest red pixel locates below row 310 of the original images, setting the origin (0, 0) of the coordinate system to the upper-left corner.
6 Remove objects with height/width ratios not in the range [0.7, 1.3].
7 Check existence of the corners of each object.
a. Find the red pixel with the smallest row number. When there are many such pixels, choose the pixel with the smallest column number.
b. Find the red pixels with the smallest and the largest column numbers. If there are multiple choices, choose those with the largest row numbers.
c. Mark locations of these three pixels in the imaginary nine equal regions, setting their corresponding containing regions by 1.
d. Remove the object if these pixels do not form any of the patterns listed aside.
8. For each surviving bounding box, extract the corresponding rectangular area from the original image and save it into the ROI list.

Figure 4 illustrates how we detect a triangular sign with the *Detection* procedure. Notice that the sing in (f) is not exactly upright. The tree trunks and red sign behind the sign made our algorithm unable to extract the complete red border. All objects detected by *Detection* are very likely to contain a triangular traffic sign. They will be used as input to the recognition component after the preprocessing step.

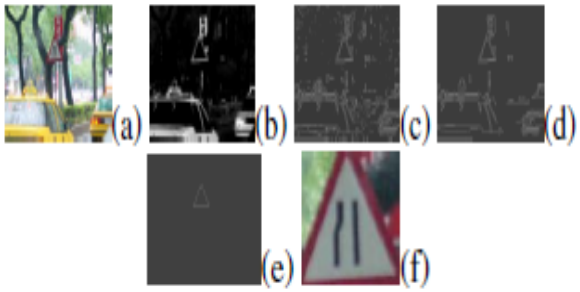


Fig. 4: An illustration of detection steps: (a) the original image; (b) result of color segmentation; (c) result of edge detection; (d) result of removing small objects; (e) results of filtering objects by step 7; (f) the enlarged image of the detected sign.

6. Preprocessing

Procedure *Preprocessing* (Input: an ROI object list; Output: an object list)

Steps:

For each object in the ROI list, do the following:

1. Normalize the object to the standard size 80x70.s
2. Extract the rectangle of 32x30 pixels from (25, 30).
3. Remove remaining red pixels.
4. Convert the object to a gray-level image.

As the first step of the preprocessing, we normalize all objects to the 80x70 standard size. After a simple analysis of the 45 standard triangular signs, we found that the ideographs appear in a specific region in the normalized images. As shown in Figure 5(a), we can extract the ideographs from a particular rectangular area in the image. We extract the ideograph from a pre-selected area of 32x30 pixels from the normalized image. The coordinates of the upper left corner of the extracted rectangle is (25, 30). Notice that, although we have attempted to choose the rectangular area such that it may accommodate distorted and rotated signs, the extracted image may not include all the original ideographs all the time. Figure 5(b) shows that the bottom of the ideograph was truncated. Similarly, the extracted area may contain noisy information. After extracting the rectangular area that might contain the ideograph, we remove red pixels in the extract. We use a more stringent standard for defining "red." Let *R*, *G*, and *B* be the red, green, and blue component of a pixel. A pixel is red if $R>20$, $(RB)> 20$, and $(R-G)>20$. After removing the red pixels, we convert the result into a gray-level image. We adjust pixel values based on the average luminance to increase contrast of the image.We compute the YIQ values of each pixel from its RGB values, set their gray levels to their luminance values, and compute the average gray levels of all pixels. Let the average be α. We invert the colors of the pixels by deducting the amount of $(\alpha-100)$ from the gray levels of all pixels. Then, pixels whose remaining gray levels are smaller than 70 are set to 0, and others are set to 255. However, if using 70 as the threshold gives us less than 10 pixels with value 255 or 10 pixels with value 0, we apply another slightly more complex method. We calculate the average gray level of the pixel values, and use this average, λ, as the cutting point for assigning pixel values in the gray-level image. Pixels whose gray levels are less than λ are set to 0, and others are set to 255. Figure 4(c) shows such a gray-level image.

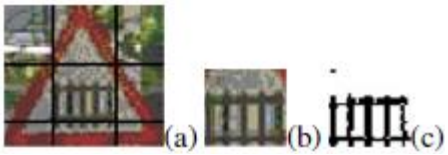


Fig. 5: Preprocessing Steps

7. TRAFFIC SIGN RECOGNITION

After the preprocessing procedure, each object becomes a rectangle of 32x30 pixels. We can use these raw data as features for recognition. In addition, we employ the discrete cosine transform (DCT) and the singular value decomposition (SVD) procedures for extracting the invariant features of the ideographs. DCT is one of the popular methods for decomposing a signal to a sequence of components and for coding images. We concatenate rows of a given object, generated at step 5 in *Preprocessing*, into a chain, and apply the one-dimension DCT over the chain, and use the first 105 coefficients as the feature values. We apply singular value decomposition to the matrices of the objects that are obtained at step 4 in the *Preprocessing* procedure for extracting features of the objects. Let $U\Sigma V^T$ be the singular value decomposition of the matrix that encodes a given object. We employ the diagonal values in Σ as the feature values of the given object. Since the original matrix is 32x30, we obtain 30 feature values from Σ.

8. Conclusion and Future Directions

Implementation of the algorithm in test images showed that it is very effective in the sign location phase. There is a slight weakness in the some phase, in cases of color similarity between signs and other areas of the image. It is sensitive in light condition changes during the image acquisition, because of the effect they have in the color thresholds used in the regions of interest segmentation step. The use of proper thresholds is very important as it affects in a great deal the success of the sign detection and it's final recognition. Based in the experience acquired from the tests, the aspects which should be further researched and be improved in the future are:

1. Recognition of signs of more complex shape.
2. Recognition of two (or more) signs in the same region of interest.
3. Increase of the speed of the algorithm by improving the source code and again, by possible changes in its structure.
4. Increase of the robustness of the algorithm in light condition changes.
5. Merging of the rectangle and triangle-ellipse detection process.

REFERENCES

[1] Hilario Gómez-Moreno, *Member, IEEE*, Saturnino Maldonado-Bascón, *Senior Member, IEEE*, Pedro Gil-Jiménez, and Sergio Lafuente-Arroyo, "Goal Evaluation of Segmentation Algorithms for Traffic Sign Recognition", *IEEE Trans. On Intelligent transportation system Apr. 2009*.

[2] Hsiu-Ming Yang, Chao-Lin Liu, Kun-Hao Liu, and Shang-Ming Huang, "Traffic Sign Recognition in Disturbing Environments", *N. Zhong et al. (Eds.): ISMIS 2003, LNAI 2871, pp. 252–261, 2003*.

[3] Matthias M¨uller, Axel Braun, Joachim Gerlach, Wolfgang Rosenstiel, Dennis Nienh¨user, J. Marius Z¨ollner, Oliver Bringmann, "Design of an Automotive Traffic Sign Recognition System Targeting a Multi-Core SoC Implementation", *978-3-9810801-6-2/DATE10 © 2010 EDAA*.

[4] N. Barnes, A. Zelinsky, and L. Fletcher, "Real-time speed sign detection using the radial symmetry detector," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 2, pp. 322–332, Jun. 2008.

[5] S. Maldonado-Bascon, S. Lafuente-Arroyo, P. Siegmann, H. Gomez-Moreno, and F. J. Acevedo-Rodriguez, "Traffic sign recognition system for inventory purposes," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 4–6, 2008, pp. 590–595.

[6] C. Nunn, A. Kummert, and S. Muller-Schneiders, "A novel region of interest selection approach for traffic sign recognition based on 3D modelling," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 4–6, 2008, pp. 654–659.

[7] C. F. Paulo and P. L. Correia, "Traffic sign recognition based on pictogram contours," in *Proc. 9th WIAMIS*, May 7–9, 2008, pp. 67–70.

[8] B. Cyganek, "Road signs recognition by the scale-space template matching in the log-polar domain," in *Proc. 3rd Iberian Conf. Pattern Recog. Image Anal.*, vol. 4477, *Lecture Notes in Computer Science*, 2007, pp. 330–337.

[9] W.-J. Kuo and C.-C. Lin, "Two-stage road sign detection and recognition," in *Proc. IEEE Int. Conf. Multimedia Expo.*, Beijing, China, Jul. 2007, pp. 1427–1430.

[10] G. K. Siogkas and E. S. Dermatas, "Detection, tracking and classification of road signs in adverse conditions," in *Proc. IEEE MELECON*, Málaga, Spain, 2006, pp. 537–540.

[11] P. Gil-Jiménez, S. Maldonado-Bascón, H. Gómez-Moreno, S. Lafuente- Arroyo, and F. López-Ferreras, "Traffic sign shape classification and localization based on the normalized FFT of the signature of blobs and 2D homographies," *Signal Process., vol. 88, no. 12, pp. 2943–2955, Dec. 2008*.

AUTHORS PROFILE

.

Chetan J Shelke is Asst Professor in P.R.Patil College of Engg.He did his M.E (information technology) in 2011 at Prof. Ram Meghe Institute Of Technology & Research, Badnera.He did his B.E(Computer Science & Engg) from H.V.P.M College of Engg & Tech Amravati in 2007.

Image Compression: An Artificial Neural Network Approach

Anjana B[1], Mrs Shreeja R[2]

[1] Department of Computer Science and Engineering, Calicut University, Kuttippuram
[2] Department of Computer Science and Engineering, Calicut University, Kuttippuram

Abstract

Image compression has become the most recent emerging trend throughout the world. Image compression is essential where images need to be stored, transmitted or viewed quickly and efficiently. The artificial neural network is a recent tool in image compression as it processes the data in parallel and hence requires less time and is superior over any other technique. The reason that encourage researchers to use artificial neural networks as an image compression approach are adaptive learning, self-organization, noise suppression, fault tolerance and optimized approximations. A survey about different methods used for compression has been done. From the above study, recently used network is multilayer feed forward network due to its efficiency. The choice of suitable learning algorithm is application dependent. A new approach by modifying the training algorithm to improve the compression is proposed here. Protection of image contents is equally important as compression in order to maintain the privacy. If any malicious modification occurs either in storage or in transmission channel, such modifications should be identified. So authentication and protection are incorporated into the proposed system to enhance the security.

Keywords: Jacobian, Levenberg-Marquardt, Multilayer perception, Neural network, Radial basis function.

1. Introduction

Image compression has become the most recent emerging trend throughout the world. Some of the common advantages of image compression over the internet are reduction in time of web page uploading and downloading and lesser storage space in terms of bandwidth. Compressed images make it possible to view more images in a shorter period of time. Image compression is essential where images need to be stored, transmitted or viewed quickly and efficiently. Image compression is the representation of image in a digitized form with a few bits maintenance only allowing acceptable level of image quality. A high quality image may require 10 to 100 million bits for representation. The large data files associated with images thus drive the need for extremely high compression ratio to make storage practical. Compression exploits the following facts.

* Imagery data has more redundancy than we can generally find in other types of data.
* The human eye is very tolerant of approximation error in an image. This tolerance has to be exploited in order to produce increased compression at the expense of image quality.

Artificial neural networks are simplified models of the biological neuron system. A neural network is a highly interconnected network with a large number of processing elements called neurons in an architecture inspired by the brain. Artificial neural networks are massively parallel adaptive networks which are intended to abstract and model some of the functionality of the human nervous system in an attempt to partially capture some of its computational strengths. A neural network can be viewed as comprising eight components which are neurons, activation state vector, signal function, pattern of connectivity, activity aggregation rule, activation rule, learning rule and environment. They are considered as the possible solutions to problems and for the applications where high computation rates are required. The BPNN has the simplest architecture of ANN that has been developed for image compression but its drawback is very slow convergence.Image processing is a very interesting and are hot areas where day-to-day improvement is quite inexplicable and has become an integral part of own lives. It is the analysis, manipulation, storage, and display of graphical images. Image processing is a module primarily used to enhance the quality and appearance of black and white images. It enhances the quality of the scanned or faxed document, by performing operations that remove imperfections. Image processing operations can be roughly divided into three major categories, image enhancement, image restoration and image compression. Image compression techniques aim to remove the redundancy present in data in a way, which makes image reconstruction possible. Image compression continues to be an important subject in many areas such as communication, data storage, computation etc.The report begins with an introduction to image compression following the need for the compression. The next section describes some of the underlying technologies for performing the image compression follows its observation and analysis. Last section is the future scope and conclusion.

2. Related works

2.1 Back Propagation Neural Network [1]

The neural network is designed with three layers, one input layer, one output layer and one hidden layer. The input layer and output layer are fully connected to the hidden layer. Compression is achieved by designing the number of neurons at the hidden layer, less than that of neurons at both input and the output layers. Image compression is achieved by training the network in such a way that the coupling weights scale the input vector of N-dimension into a narrow channel of K-dimension with K less than N, at the hidden layer and produce the optimum output value which makes the quadratic error between input and output minimum. Basic neural network used for compression is shown in Figure 1. The basic back-propagation network is further extended to construct a hierarchical neural network by adding two more hidden layers into the existing network.

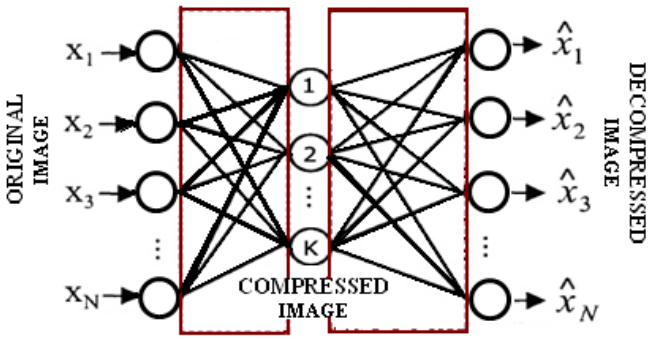


Fig 2.1:- Back Propagation Neural Network

2.2 Hierarchical and adaptive back-propagation neural network [2]

The basic back-propagation network is further extended to construct a hierarchical neural network by adding two more hidden layers into the existing network. All three hidden layers are fully connected. Nested training algorithm is proposed to reduce the overall neural network training time. The neuron weights are maintained the same throughout the image compression process. Hierarchical neural network for compression is shown in Figure 2. Adaptive schemes are based on the principle that different neural networks are used to compress image blocks with different extent of complexity. The basic idea is to classify the input image blocks into a few subsets with different features according to their complexity measurement. A fine tuned neural network then compresses each subset. Prior to training, all image blocks are classified into four classes according to their activity values which are identified as very low, low, high and very high activities. The network results in high complexity.

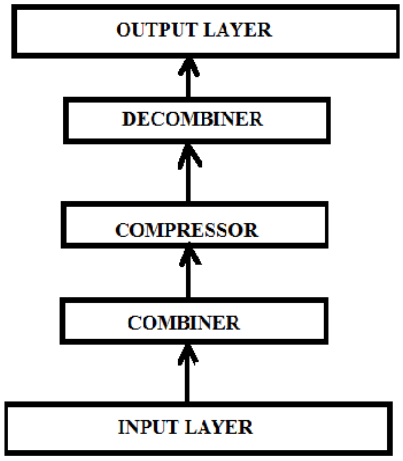


Fig 2.2:- Hierarchical Neural Network

2.3 Multi layer Feed Forward Artificial Neural Network [3], [4]

The network is designed in a way such that N will be greater than Y, where N is input layer/output layer neurons and Y is hidden layer neurons. Divide the training image into blocks. Scale each block and apply it to input layer and get the output of output layer. Adjust the weight to minimize the difference between the output and the desired output. Repeat until the error is small enough. The output of hidden layer is quantized and entropy coded to represent the compressed image. Two categories of optimization algorithms are considered i.e., derivative-based and derivative-free [5]. Derivative based methods include gradient descent, conjugate-gradient,Quasi Newton and Levenberg-Marquardt methods. Gradient descent indicates the direction to move. The conjugate-gradient method reduces oscillatory behavior and adjusts weight according to the previously successful path directions as it uses a direction vector which is a linear combination of past direction vectors and the current negative gradient vector. LM and QN algorithm-based back propagation neural networks are equally efficient. Under derivative free, two of the popular developed approaches are Genetic Algorithm (GA) and Particle Swarm Optimization (PSO).

2.4 Multilayer Perception [6]

Basic multilayer perception (MLP) building unit is a model of artificial neuron. This unit computes the weighted sum of the inputs plus the threshold weight and passes this sum through the activation function usually sigmoid. In a multilayer perception, the outputs of the units in one layer form the inputs to the next layer. The weights of the network are usually computed by training the network using the back propagation algorithm. The basic computational unit, often referred to as a neuron, consists of a set of synaptic weights, one for every input, plus a bias weight, a summer, and a nonlinear function referred to as the activation function. Each unit computes the weighted sum of the inputs plus the bias weight and passes this sum through the activation function to calculate the output value as

$$y_j = f(\sum w_{ji} x_i + \emptyset_i) \quad (1)$$

2.5 Radial Basis Function Network [6]

Radial basis function networks are feed-forward networks. They are trained using a supervised training algorithm. They are typically configured with a single hidden layer of units whose output function is selected from a class of functions called basis functions. The input layer is made up of source nodes (sensory units) whose number is equal to the dimension N of the input vector. The second layer is the hidden layer which is composed of nonlinear units that are connected directly to all of the nodes in the input layer. Each hidden unit takes its input from all the nodes at the input layer. The hidden units contain a basis function, which has the parameters centre and width. Observation and AnalysisThe back propagation neural network is generally used as a basic network through which different variations of image compression schemes can be implemented with different error functions and using overlapped blocks, which include hierarchical and adaptive back propagation neural networks. Later came neural network based adaptive image coding which was basically developed from the mathematical iterations for obtaining the K-L transform conventionally. To improve the compression performance, multi layer feed forward network is used. It uses different optimization methods of which Quasi Newton is better but takes a long time. There are different optimization techniques which can be combined with basic networks in order to improve the compression efficiency. Survey is concluded by giving a brief idea about how the authentication and protection to be incorporated into the neural network to enhance the security.

3. Proposed System

Two different categories for improving the compression methods and their performance have been suggested. In the first case, conventional methods like SPIHT, vector quantization (VQ) etc., can be used with some enhancements. Secondly, apply neural network to develop the compression scheme, so that new methods can be developed and further research possibilities can be explored in future.In this work, image compression using multi layer neural networks has been proposed. In the proposed system, there is a testing set consists of sub images that are not included in the training set. Levenberg-Marquardt algorithm is used for training purpose. Image pixels are normalized before the compression process. If the learning factor ∞ is very large, the LM algorithm becomes the steepest decent. This parameter is automatically adjusted for all iterations in order to secure convergence. Here, a modified version of LM algorithm is proposed that provides a similar performance, while lacks the inconveniences of LM. It is more stable. The MSE between the target image and reconstructed image should be as small as possible so that the quality of reconstructed image should be near to the target image. The proposed method gives high compression ratio.

(a) One to one mapping:
For incorporating protection of the data, one to one property of the neural network can be used. If there are interactions of two parameters, resultant should be a unique value stated as:

$$\emptyset(x_i, y) \neq \emptyset(x_j, y); \forall j if j \neq i; \quad (2)$$

(b) One way property:

For authentication, the property allows to compute output from the input easily while very difficult to compute input from the output. The input P is composed of n elements $[p_1, p_2, \dots \dots, p_n]$ while the output is unique C as:

$$C = f\left(\sum_{j=1}^{n} w_j p_j + b\right) \quad (3)$$

It is easy to compute C from a given P, but difficult to compute P from C.

3.1 Neural Network Compression

The compression process is described below:-

1. Read image pixels and then normalize it to range [0-1].
2. Divide the image into non-overlapping blocks.
3. Rasterize the pixels and apply to the input layer.
4. Compute the outputs of hidden layer units by multiplying the input vector by the weight matrix (V).
5. Store the outputs in a compressed file after renormalization.
6. If there are more image vectors go to (4).
7. Stop.

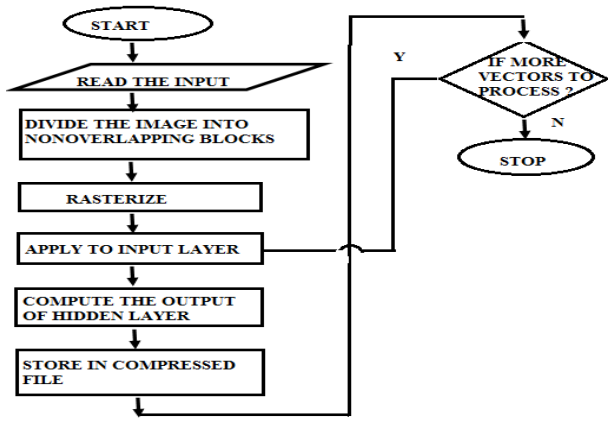


Fig 3.1 Compression

The decompression process is described below:-

1. Take one by one vector from the compressed image.
2. Normalize this vector.
3. The outputs of output layer units by multiplying outputs of hidden layer units by the weight matrix.
4. Derasterize the outputs of output layer units to build the sub image.
5. Return this sub image to its proper location.
6. Renormalize this block and store it in the reconstructed file.
7. If there are more vectors go to (1).

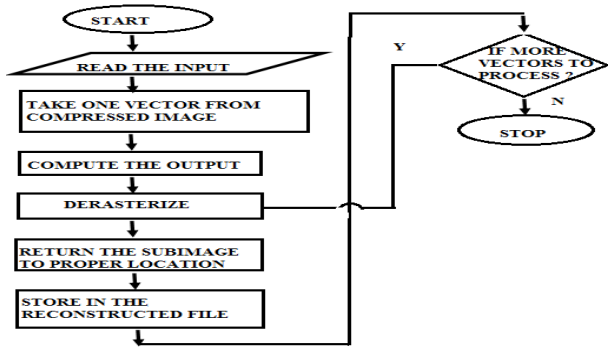


Fig 3.2 Decompression

4. Implementation

4.1 Preprocessing

The neural network requires inputs with real type and the sigmoid function of each neuron requires the input data to be in the range [0-1]. For this reason, the image data values must be normalized. The normalization is the process of linearly transformation of image values from the range [0-255] into another range that is appropriate for neural network requirements. Segmentation is the process of dividing it into non overlapping blocks with equal size to simplify the learning/compressing processes. Image rasterization is the process of converting each sub image from a two dimensional block in to a one dimensional vector, to speed up the learning.

4.2 Neural Network Design

Multilayer feedforward network is used for compressing the images. Neural network is designed in such a way that the numbers of input and output layer neurons are set to 64. Hidden layer neurons are set to 16. The two weight matrices are selected to small random numbers.

4.3 Training

The input image is split up into blocks or vectors of 4X4, 8X8 or 16X16 pixels. These vectors are used as inputs to the network. The network is provide by the expected output, and it is trained so that the coupling weights, {wij}, scale the input vector of N -dimension into a narrow channel of Y -dimension, which is less than N, at the hidden layer and produce the optimum output value which makes the quadratic error between output and the desired one minimum.

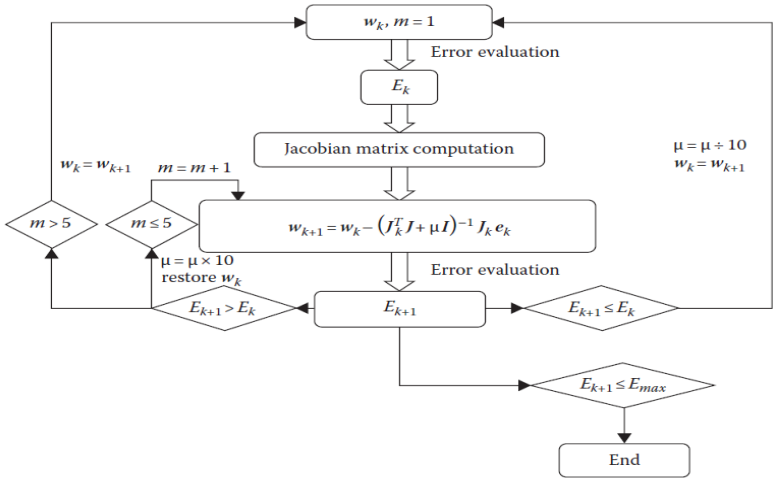


Fig 4.1 LM algorithm

The LM algorithm has got some disadvantages. If the learning factor is very large, the LM algorithm becomes the steepest decent. This parameter is automatically adjusted for all iterations in order to secure convergence. The LM algorithm computes the Jacobin J matrix at each iteration step and the inversion of square matrix. In the LM algorithm must be inverted for all iterations. Hence for large size neural networks, the LM algorithm is not practical. Here, a modified version of LM algorithm is proposed that provides a similar performance, while lacks the inconveniences of LM. A new performance index is introduced,

$$F(w) = \sum_{k=1}^{p} \left[\sum_{p=1}^{p} \left(d_{kp} - o_{kp} \right)^2 \right]^2 \quad (4)$$

where d_{kp} is the desired value of k^{th} output and o_{kp} is the actual value of k^{th} output and the p^{th} pattern is the number of the weights, P is the number of patterns, and K is the number of network outputs. This index represents a global error, will later lead to a significant reduction of the size of a matrix to be inverted at each iteration step [6]. The learning factor, α is modified as 0.01 $E^T E$, where E is a k×1 matrix. If the error is small, then actual output approaches to desired output.

The trained network is now ready to be used for image compression which, is achieved by dividing or input images into normalization and segmentation. To decompress the image; first the compressed image is renormalized then applies it to the output of the hidden layer and get the one vector of the hidden layer output is normalized then it rasterization to represent the reconstruct the image.

MSE and PSNR are the parameters which define the quality of an image reconstructed at the output layer of neural network.

a) Mean Square Error (MSE)

The MSE between the target image and reconstructed image should be as small as possible so that the quality of reconstructed image should be near to the target image. Ideally, the mean square error should be zero for ideal decompression. The compression ratio is defined by the ratio of the data fed to the input layer neurons to the data out from the hidden layer neurons. In a structure 1, 016 neurons were used in the hidden layer. So it will results in the fixed 4:1 compression ratio.

b) Peak Signal to Noise ratio (PSNR)

The term peak signal-to-noise ratio (PSNR) is an expression for the ratio between the maximum possible value (power) of a signal and the power of distorting noise that affects the quality of its representation. The PSNR computes by the following equation:-

$$PSNR = 10\ log_{10}\ {255^2}/{MSE} \quad (5)$$

The compression ratio performance can be computed by,

$$CR=(1 - N_h/N_i) \times 100\% \quad (6)$$

where N_h is the input layer neurons and N_i is the hidden layer neurons.

5. Conclusion

The need for effective data compression is evident in almost all applications where storage and transmission of digital images are involved. Neural networks offer the potential for providing a novel solution to the problem of compression by its ability to generate an internal data representation.Multilayer feed forward network is used due to its efficiency. Learning algorithms has significant impact on the performance of neural networks, and the effects of this depend on the targeted application. The choice of suitable learning algorithms is therefore application dependent. The performance can be increased by modifying the training algorithm which outperforms the existing method.

Protection of image contents is equally important as compression in order to maintain the privacy. If any malicious modification occurs either in storage or in transmission channel, such modifications should be identified. So the authentication and protection can be incorporated into the proposed system in future by utilizing the other properties of the neural network.

References

[1] J. Jiang, Image compression with neural networks–a survey, Signal processing: image Communication, 1999, 737–760.

[2] M. Egmont-Petersen, D. de Ridder, and H. Handels, Image processing with neural networks–a review, Pattern recognition, vol. 35, no. 10, 2002, 2279–2301.

[3] F. Ibrahim, Image compression using multilayer feed forward artificial neural network and dct, Journal of Applied Sciences Research, vol. 6, no. 10, 2010, 1554–1560.

[4] V. Gaidhane, V. Singh, Y. Hote, and M. Kumar, New approaches for image compression using neural network, Journal of Intelligent Learning Systems and Applications, vol. 3, no. 4, 2011, 220–229.

[5] N. Relhan, M. Jain, V. Sahni, J. Kaur, and S. Sharma, Analysis of optimization techniques for feed forward neural networks based image compression, International Journal of Computer Science and Information Technologies, vol. 3, no. 2, 2012.

Effect of Radiation on Flow of Second Grade Fluid over a Stretching Sheet Through Porous Medium With Temperature Dependent Viscosity And Thermal Conductivity

G. C. Hazarika[1] P. K. Mahanta[2],

[1]Department of Mathematics, Dibrugarh University, Dibrugarh-7866004, Assam, India
[2]Department of Mathematics, Namrup College, Dibrugarh-786623, Assam, India

Abstract:

The effect of thermal radiation on boundary layer flow with temperature dependent viscosity and thermal conductivity due to a stretching sheet in porous media is investigated. The Rosseland diffusion approximation is used to describe the radiative heat flux in the energy equation. The sheet is being stretched linearly in the presence of a uniform transverse magnetic field and the flow is governed by the second –order viscoelastic fluid. The partial differential equations governing the flow and heat transfer characteristics are converted into ordinary differential equations by similarity transformations and solved numerically by fourth-order Runge-Kutta shooting method. The effects of various parameters on the velocity and temperature profiles as well as the skin-friction coefficient and Nusselt number has been shown graphically and in tabulated form and discussed in detail.

Keywords: Heat transfer, Porous medium, Radiation, Second order fluid, Stretching sheet, Thermal Conductivity, Variable viscosity

1. Introduction

The study of the flow and heat transfer created by a moving surface is relevant to several applications in the fields of metallurgy and chemical engineering, polymer processing, electro-chemistry, MHD power generators, flight magneto hydro dynamics as well as in the field of planetary magneto spheres, aeronautics and chemical engineering. Sakiadis [1] was the first to study the boundary layer flow due to a moving wall in fluid at rest. The study of flow over a stretching surface has generated much interest in recent years in view of its numerous industrial applications such as extension of polymer sheets, glass blowing, rolling and manufacturing plastic films and artificial fibers. The pioneer work on the boundary layer flows over stationary and continuously moving surfaces was initially done by Blasius [2] and Crane [3]. Ali [4] carried out a study for a stretching surface subject to suction or injection for uniform and variable surface temperatures. Rajgopal et al [5] , Dandapat and Gupta [6], Shit [7] and Reddaiah and Rao [8] extensively studied on various aspects of boundary layer flow problems over a stretching sheet.

In cooling processes, the effect of thermal radiation is also an important factor in non-isothermal systems. Hady and Mohamed [9] studied the MHD mixed convection with thermal radiation in laminar boundary layer flow over a semi-infinite flat plate embedded in porous media. Mansour [10] studied the effects of radiation and forced convection on the flow over a flat plate submersed in a porous medium of a variable viscosity. Mohammadein *et .al* [11] studied the effects of radiation with both first and second-order resistance's due to the solid matrix on some natural convection flows in fluid-saturated porous media. The effect of thermal radiation on mixed convection from horizontal surfaces in saturated porous media was investigated by Bakier and Gorla [12]. Prasad et al [13]) studied the radiation and mass transfer effects on unsteady MHD free convection flow past a vertical porous plate embedded in porous medium: a numerical study. Anjali Devi and Kayalvizhi [14] presented analytical solution of MHD flow with radiation over a stretching sheet embedded in a porous medium.

In most of the studies of this type of problems, the viscosity and thermal conductivity of the fluid were assumed to be constant. However, it is known that the physical properties can changed sufficiently with temperature and when the effects of variable viscosity and thermal conductivity are taken in to account, the flow characteristics are significantly changed compared to the constant property. Hassanien et al [15] revealed that the fluid viscosity and thermal conductivity might function of temperatures as well as the fluid is considering. Recently Sharma and Hazarika [16] studies the effects of variable viscosity and thermal conductivity on heat and mass transfer flow along a vertical plate in the presence of a magnetic field.

Also, most of the practical situations demand for fluids that are non-Newtonian in nature which are mainly used in many industrial and engineering applications. It is well known that a number of fluids such as molten plastic, polymeric liquid, food stuffs etc exhibit non-Newtonian character.

In the present work, thermal radiation effects on heat transfer of second grade fluid over a stretching sheet through porous medium with temperature dependent viscosity and thermal conductivity is investigated. The governing equations are transformed by using similarity transformation and the resultant dimensionless equations are solved numerically using the Runge-Kutta fourth order method with shooting technique. The effects of various governing parameters on the velocity, temperature, skin-friction coefficient and Nusselt number are shown in figures and tables and analyzed in detail. Numerical results are presented for velocity and temperature profiles for different parameters of the problem.

2. Mathematical Formulation

We consider the two-dimensional laminar boundary layer flow of viscous, incompressible, electrically conducting and radiating second grade fluid with temperature dependent viscosity and thermal conductivity past a semi-infinite stretching sheet coinciding with the plane $y=0$ embedded in a uniform porous medium. A uniform magnetic field of strength B_0 is applied in the direction perpendicular to the plate. The transverse applied magnetic field and magnetic Reynolds number are assumed to be very small, so that the induced magnetic field is negligible. Keeping the origin fixed, two equal and opposite forces are applied along the X - axis, so that the sheet is stretched with a velocity proportional to the distance from the fixed origin. Under the above assumptions, the basic boundary layer equations governing the flow and heat transfer of second grade fluid due to the stretching sheet are given by the following equations:

The equation of continuity:

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0 \tag{1}$$

Momentum conservation:

$$\rho_\infty\left(u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}\right)=\frac{\partial}{\partial y}\left(\mu\frac{\partial u}{\partial y}\right)-k_0\left\{u\frac{\partial^3 u}{\partial x\partial y^2}+v\frac{\partial^3 u}{\partial y^3}+\frac{\partial u}{\partial y}\frac{\partial^2 v}{\partial x\partial y}+\frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial y^2}\right\}-\sigma {B_0}^2 u-\frac{\mu}{K'}u \tag{2}$$

Thermal energy conservation:

$$\rho_\infty C_p\left(u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y}\right)=\frac{\partial}{\partial y}\left(k\frac{\partial T}{\partial y_r}\right)+\mu\left(\frac{\partial u}{\partial y}\right)^2-\frac{\partial q_r}{\partial y}-k_0\frac{\partial u}{\partial y}\left[\frac{\partial}{\partial y}\left(u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}\right)\right] \tag{3}$$

Along with the boundary conditions,

$$\begin{aligned} &u=U_w=cx, && v=0, && T=T_w && \text{at} && y=0\\ &u=0, && v=0, && T=T_\infty && \text{as} && y\to\infty \end{aligned} \tag{4}$$

Where u and v are the flow velocity components along x- and y- directions respectively, B_0 is the applied magnetic field, μ_∞ and k_∞ are the constant viscosity and constant thermal conductivity of the free stream of the fluid respectively. T is the temperature of the fluid. μ and k are the coefficient of variable viscosity and variable thermal conductivity respectively of the fluid which are considered to vary as a function of temperature. C_p is the specific heat at constant pressure and k_0 is the coefficient of visco-elastricity. σ is the electrical conductivity. c is the constant stretching rate. T_∞ and ρ_∞ are the free stream temperature and density. K' is the permeability of the porous medium. q_r is the radiation heat flux.

Flowing Lai and Kulacki [17] We assume

$$\frac{1}{\mu}=\frac{1}{\mu_\infty}\left[1+\gamma\left(T-T_\infty\right)\right] \qquad or \qquad \frac{1}{\mu}=a\left(T-T_r\right) \tag{5}$$

$$where \qquad a=\frac{\gamma}{\mu_\infty}, \qquad and \qquad T_r=T_\infty-\frac{1}{\gamma}$$

and

$$\frac{1}{k}=\frac{1}{k_\infty}\left[1+\kappa\left(T-T_\infty\right)\right] \qquad or \qquad \frac{1}{k}=\varepsilon\left(T-T_\infty\right) \tag{6}$$

$$where \qquad \varepsilon=\frac{\kappa}{k_\infty} \qquad and \qquad T_e=T_\infty-\frac{1}{\kappa}$$

Where a, ε, T_r, T_e are constants and their values depend on the reference state and thermal properties of the fluid i.e γ and κ. In general $a>0$ for liquids and $a<0$ for gases (the viscosity and thermal conductivity of liquid/gas usually decrease/increase with increasing temperature).

By assuming Rossenland approximation for radiation, the radiative heat flux q_r is given by

$$q_r=-\frac{4\sigma^*}{3K^*}\frac{\partial T^4}{\partial y} \tag{7}$$

Where σ^* and K^* are the Stefan-Bolzman constant and the mean absorption coefficient respectively. We assume that the temperature differences within the flow are sufficiently small such that T^4 may be expressed as a linear function of the temperature as shown in Chamakha [18]. Expanding T^4 in a Taylor series about T_∞ and neglecting higher order terms we obtain

$$T^4\cong 4T_\infty{}^3T-3T_\infty{}^4 \tag{8}$$

Using (7) and (8), we obtain as

$$\frac{\partial q_r}{\partial y}=-\frac{16\sigma^*T_\infty{}^3}{3K^*}\frac{\partial^2T}{\partial y^2} \tag{9}$$

3. Method of Solution

The mathematical analysis of the problem is simplified by introducing the following dimensionless coordinates in terms of similarity variable η and the similarity function f as

$$u=cxf'(\eta), \qquad v=-\sqrt{c\upsilon}\,f\left(\eta\right), \qquad \eta=\sqrt{\frac{c}{\upsilon}}y, \qquad \theta=\frac{T-T_\infty}{T_w-T_\infty} \tag{10}$$

Where prime denotes the differentiation with respect to η and θ is the dimensionless temperature.

Clearly the continuity equation (1) is satisfied by u and v defined in equation (10). Substituting equation (10) in equations (2) - (3) gives the following equations

$$\left(\frac{\theta-\theta_r}{\theta_r}\right)\left[\left(f'\right)^2-ff''\right]+f'''-\frac{\theta'}{\theta-\theta_r}f''+K_1\left(\frac{\theta-\theta_r}{\theta_r}\right)\left[2ff'''-\left(f''\right)^2-ff^{iv}\right]+\left[M\left(\frac{\theta-\theta_r}{\theta_r}\right)+K\right]f=0 \tag{11}$$

And

$$(4R+3)\theta''+3P_r f\theta'-3P_r E_c\left(\frac{\theta_r}{\theta-\theta_r}\right)(f'')^2-3K_1E_cP_r\left[f'(f'')^2-ff''f'''\right]=0 \tag{12}$$

The transformed boundary conditions are reduce to

$$f'(\eta)=1, \quad f(\eta)=0, \quad \theta(\eta)=1, \quad at \quad \eta=0, \tag{13}$$

$$f'(\eta)\to 0, \quad f''(\eta)\to 0, \quad \theta(\eta)\to 0, \quad as \quad \eta\to\infty, \tag{14}$$

Where prime denotes differentiation with respect to η only and

$K_1=\frac{k_0 c}{\rho_\infty \upsilon}$ is the viscoelastic parameter, $M=\frac{\sigma B_0^{\,2}}{\rho_\infty c}$ is the magnetic parameter,

$\Pr=\frac{\mu c_p}{k}$ is the Prandtl number, $Ec=\frac{U_w^{\,2}}{c_p(T_w-T_\infty)}$ is the Eckert number.

$K=\frac{\upsilon}{K'c}$ is the porosity parameter, $R=\frac{4\sigma^* T_\infty^{\,3}}{kK^*}$ is the radiation parameter and

θ_r is the dimensionless parameter characterizing the influence of viscosity , where

$$\theta_r=\frac{T_r-T_\infty}{T_w-T_\infty}=-\frac{1}{\gamma(T_w-T_\infty)} \tag{15}$$

For engineering purpose, one is usually less interested in the shape of the velocity and temperature profiles then in the value of the skin-friction, heat transfer. The expression for the local skin-friction coefficient C_f and the local Nusselt number Nu defined by:

$$C_f=\frac{\tau_w}{\mu_\infty(cx)\sqrt{\frac{c}{\upsilon}}}=-\left[\frac{\theta_r}{\theta-\theta_r}+2K_1\right]f''(0), \tag{16}$$

$$Nu=\frac{q_w}{k\sqrt{\frac{c}{\upsilon}}(T_w-T_\infty)}=-\theta'(0) \tag{17}$$

Where

$$q_w=-k\left(\frac{\partial T}{\partial y}\right)_{y=0}=-k\sqrt{\frac{c}{\upsilon}}(T_w-T_\infty)\theta'(0),$$

3. Numerical Results and Discussion

The system of differential equations (11) and (12) governed by boundary conditions (13) and (14) are solved numerically by applying an efficient numerical technique based on the fourth order Runge-Kutta shooting method and an iterative method. It is experienced that the convergence of the iteration process is quite rapid. The numerical computations have been carried out for various values of radiation parameter R, visco-elastic parameter K_1, Eckert number Ec, Prandtl number $\Pr$, porosity parameter K, Magnetic parameter M and the dimensionless viscosity parameter θ_r. In order to illustrate the results graphically, the numerical values of dimensionless velocity $f'(\eta)$ and dimensionless temperature $\theta(\eta)$ are plotted in Figures 1 – 14.

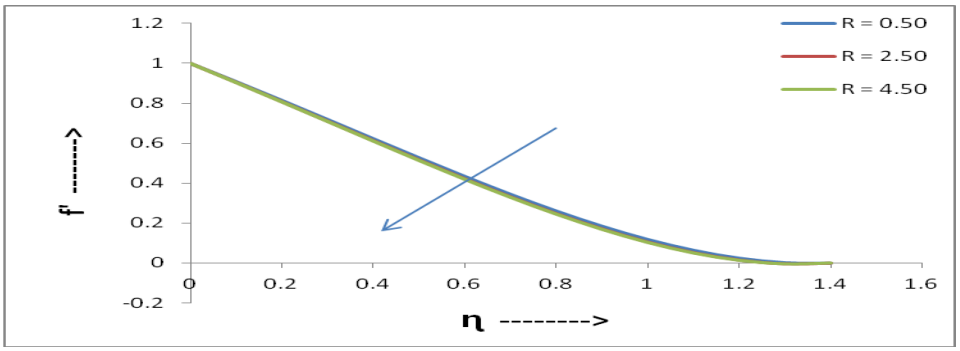


Figure 1. Variation of $f'(\eta)$ with η for different values of R

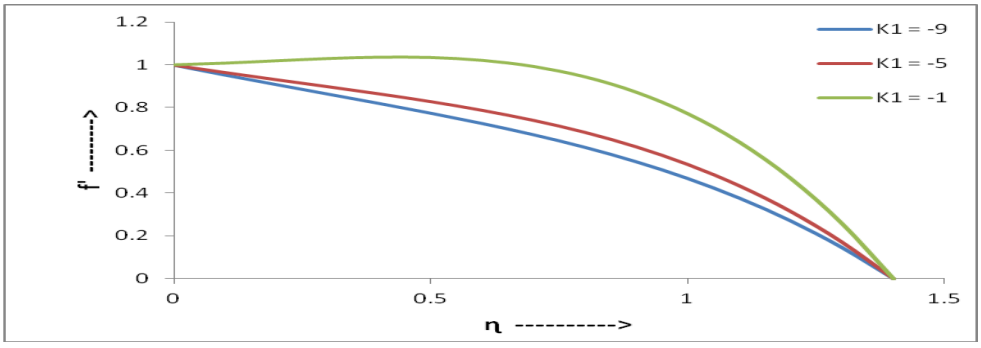


Figure 2. Variation of $f'(\eta)$ with η for different values of K_1

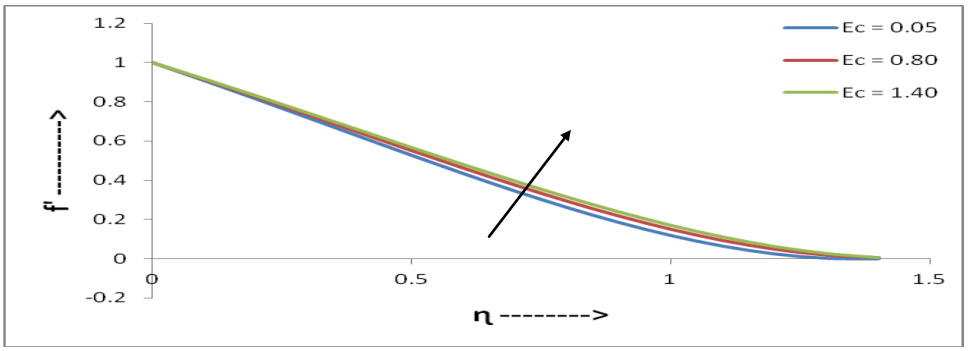


Figure 3. Variation of $f'(\eta)$ with η for different values of Ec

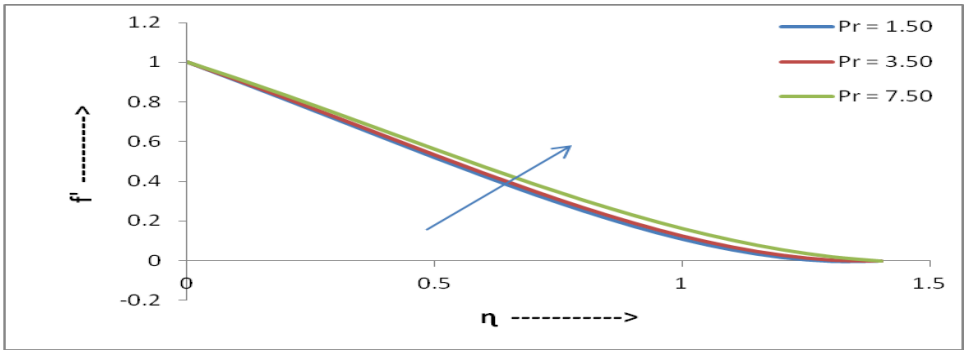


Figure 4. Variation of $f'(\eta)$ with η for different values of $\Pr$

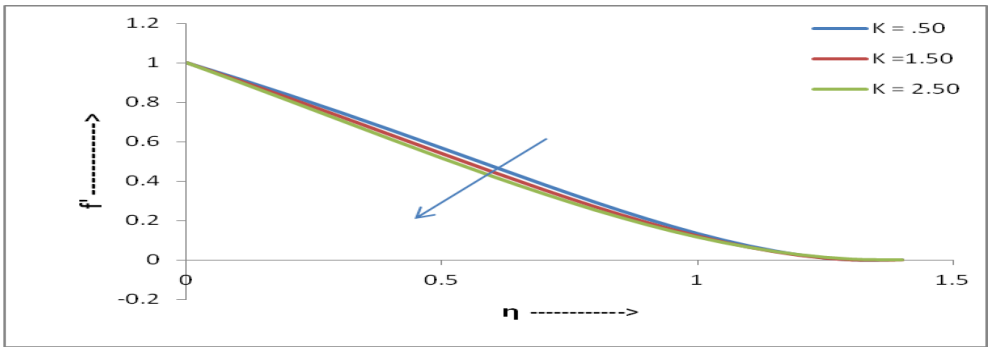


Figure 5. Variation of $f'(\eta)$ with η for different values of K

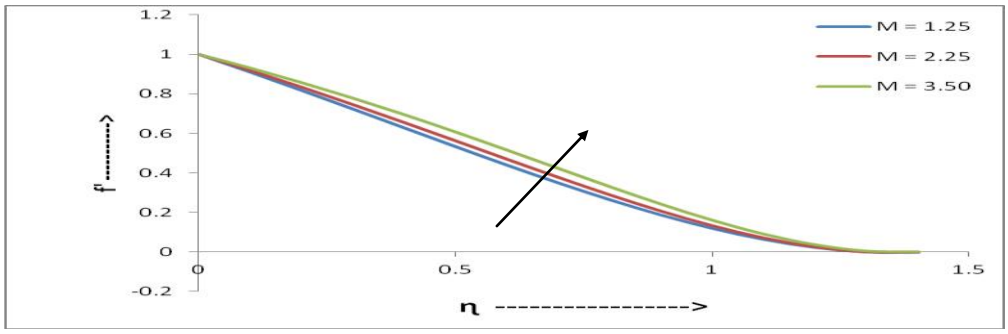


Figure 6. Variation of $f'(\eta)$ with η for different values of M

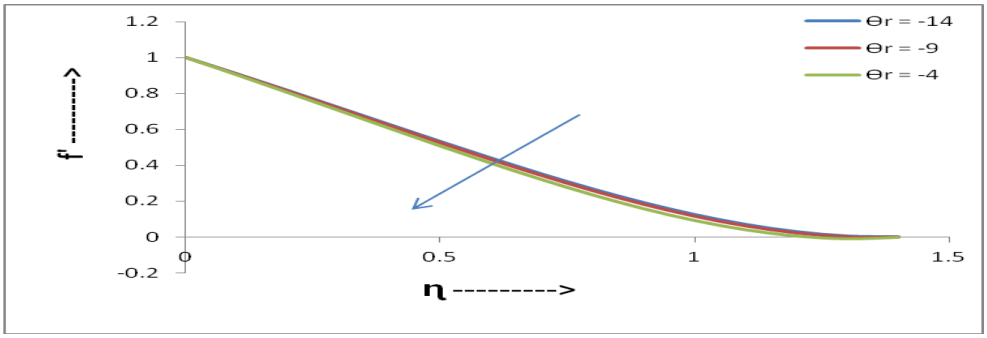


Figure 7. Variation of $f'(\eta)$ with η for different values of θ_r

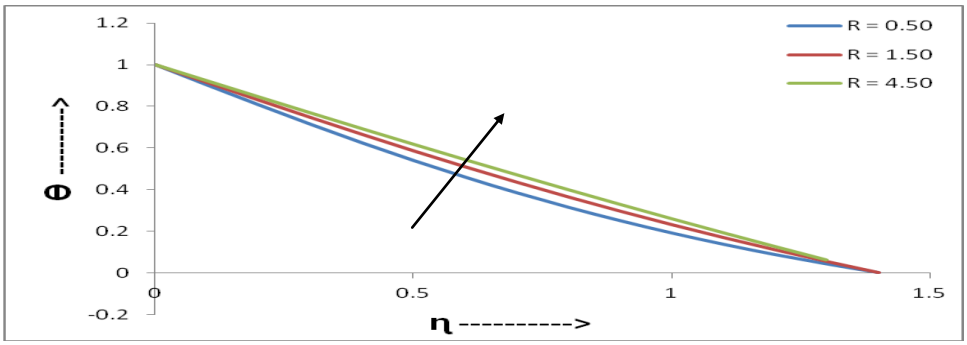


Figure 8. Variation of $\theta(\eta)$ with η for different values of R

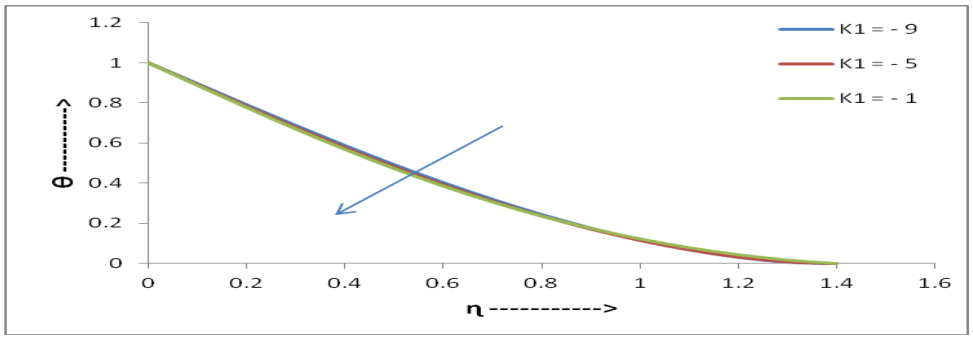


Figure 9. Variation of $\theta(\eta)$ with η for different values of K_1

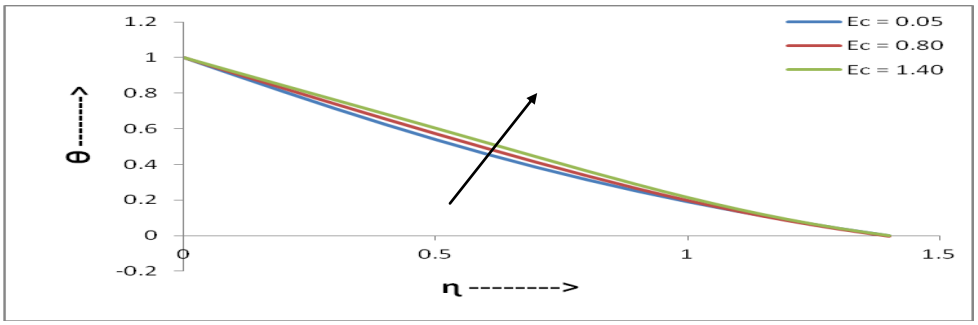


Figure 10. Variation of $\theta(\eta)$ with η for different values of Ec

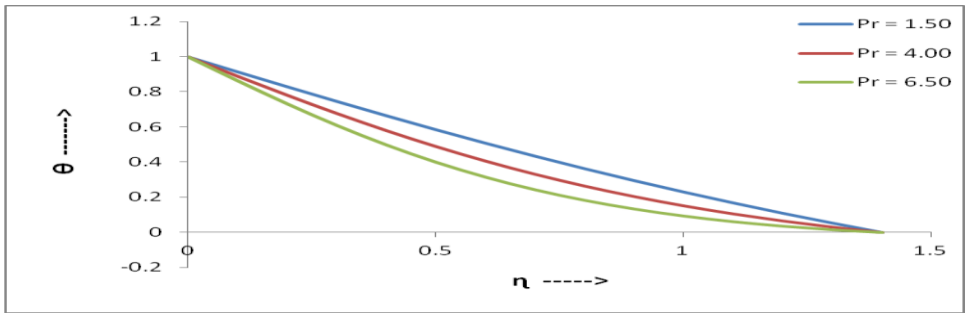


Figure 11. Variation of $\theta(\eta)$ with η for different values of Pr

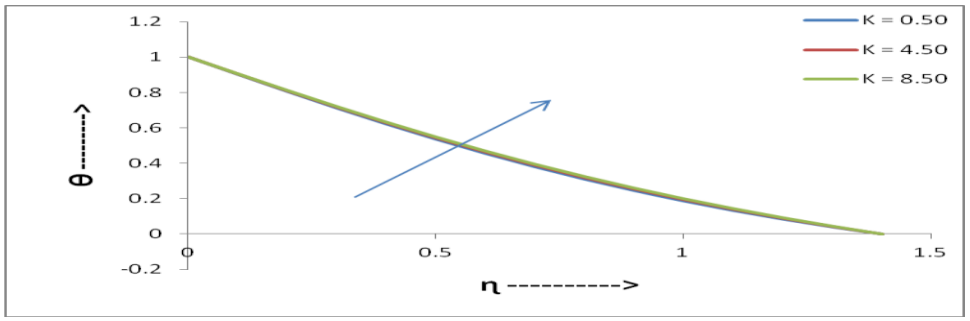


Figure 12. Variation of $\theta(\eta)$ with η for different values of K

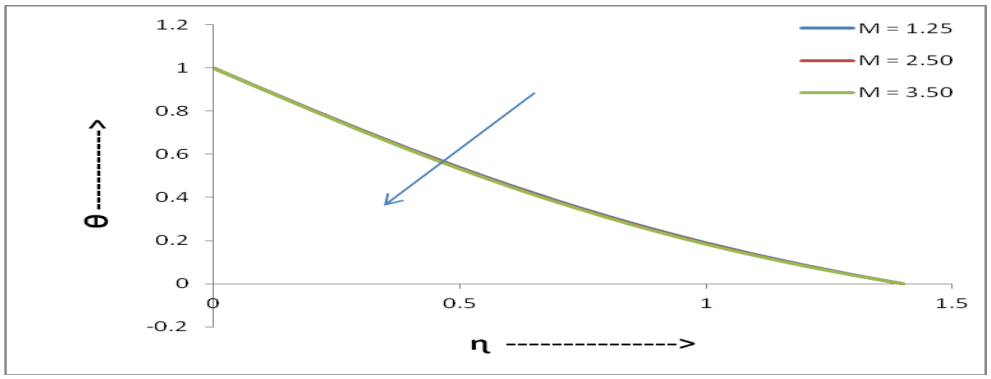


Figure 13. Variation of $\theta(\eta)$ with η for different values of M

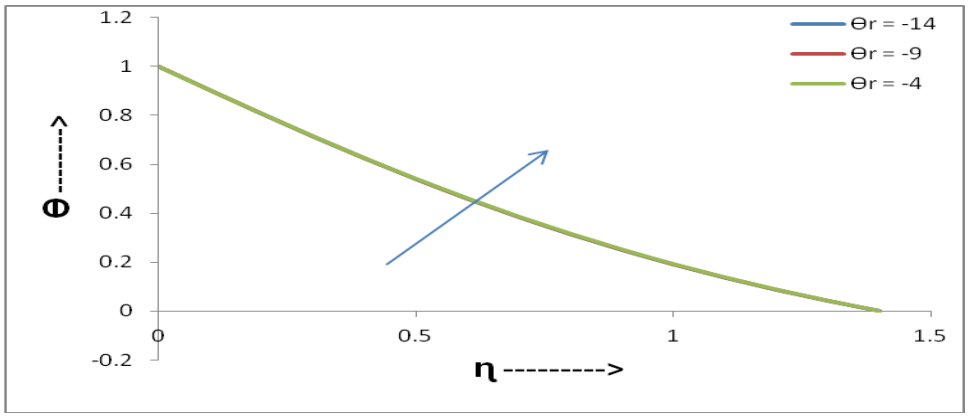


Figure 14. Variation of $\theta(\eta)$ with η for different values of θ_r

For various values of the radiation parameter R, the velocity profiles are plotted in Fig.1. It can be seen that as R increases, the velocity decreases. Fig.2. shows the effect of viscoelastic parameter K_1 on the velocity profiles. It is seen that the velocity increases as the viscoelastic parameter increases. The effect of Eckart number Ec on the velocity field is shown in Fig.3. It is noticed that the velocity profiles increases with the increase of Ekcart number. The velocity profiles for different values of Prandtl number $\Pr$ are illustrated in Fig.4. It is clear that increasing values of $\Pr$ results in increasing velocity. . Fig.5. shows the effect of permeability parameter K on the velocity profiles. It is seen that the velocity decreases as the permeability parameter increases. For various values of the magnetic parameter M , the velocity profiles are plotted in Fig.6. It can be seen that as M increases, the velocity increases. The effect of dimensionless viscosity parameter θ_r on the velocity profiles is shown in Fig.7. It is found that the velocity slightly decreases with an increase in θ_r . The effect of radiation parameter R on the temperature profiles is shown in Fig.8. It is observed that the temperature increases as R increases. Fig.9. shows the temperature profiles for different values of viscoelastic parameter K_1. It is obvious that an increase in K_1 results in decreasing temperature within the boundary layer. The effect of Eckart number Ec on the temperature profiles is depicted in Fig.10. It can be seen that an increase in Ec results in increase of the thermal boundary layer. Figs 11 and 12 noticed that the dimensionless temperature $\theta(\eta)$ decreases with the increase of the Prandtl number Pr and increases with the increasing values of porosity parameter K . It is interesting to note from Fig. 11 that the increase of Prandtl number Pr means decrease of thermal conductivity. The effect of the magnetic parameter M on temperature distribution shown in Fig. 13. From this figure we conclude that the temperature decreases with the increase of the magnetic parameter M . It may also observed from Fig. 14 that the effect of thermal radiation is to enhance the temperature with increase in the fluid viscosity parameter θ_r . It is interesting to note that in the presence of thermal radiation, the effect of viscosity parameter θ_r causes marginal significance.

The important characteristics in the present study are the local skin-friction coefficient C_f and the local rate of heat transfer at the sheet (Nusselt number Nu) defined in equations in (16) and (17).

Table-1. Numerical values of the local skin-friction: $C_f = \dfrac{\tau_w}{\mu_\infty (cx)\sqrt{\dfrac{c}{\upsilon}}} = -\left[\dfrac{\theta_r}{\theta - \theta_r} + 2K_1\right] f''(0),$

R	K	K_1	Pr	θ_r	Ec	$M = 0.0$	$M = 0.2$	$M = 0.4$
0.5	2	1	2.7	-10	.05	1.074144	1.057201	1.039990
2.5						1.111263	1.094620	1.077719
3.5						1.116569	1.099968	1.083112
0.5	**0.5**	1	2.7	-10	.05	0.951911	0.932854	0.913423
	1.5					1.034942	1.017369	0.999501
	2.5					1.112005	1.095630	1.079014
0.5	2	**-6**	2.7	-10	.05	-6.855943	-6.506797	-6.152434
		-4				-4.131017	-3.869648	-3.603017
		-2				-1.644921	-1.473744	-1.297443
0.5	2	1	**1.5**	-10	.05	1.101120	1.084394	1.067409
			2.5			1.078689	1.061782	1.044609
			3.5			1.055836	1.038748	1.021390
0.5	2	1	2.7	**-9**	.05	1.088707	1.071406	1.053830
				-5		1.197815	1.177733	1.157298
				-2		1.486126	1.457645	1.428535
0.5	2	1	2.7	-10	**.05**	1.074144	1.057201	1.039990
					0.15	1.064218	1.047599	1.030719
					0.25	1.054411	1.038109	1.021553

Table- 2. Numerical values of local Nusselt number : $Nu = -\theta'(0)$

R	K	K_1	Pr	θ_r	Ec	$Nu = -\theta'(0)$		
						$M = 0.0$	$M = 0.2$	$M = 0.4$
0.5	2	1	2.7	-10	.05	0.970627	0.971503	0.972407
2.5						0.809208	0.809521	0.809844
3.5						0.786403	0.786638	0.786882
0.5	**0.5**	1	2.7	-10	.05	0.977029	0.978125	0.979260
	1.5					0.972603	0.973544	0.974515
	2.5					0.968787	0.969606	0.970449
0.5	2	**-6**	2.7	-10	.05	1.003040	1.017594	1.032605
		-4				1.033991	1.046184	1.058830
		-2				1.061449	1.071274	1.081550
0.5	2	1	**1.5**	-10	.05	0.853002	0.853466	0.853944
			2.5			0.950695	0.951501	0.952333
			3.5			1.051387	1.052544	1.053736
0.5	2	1	2.7	**-9**	.05	0.970361	0.971254	0.972175
				-5		0.968634	0.969653	0.970706
				-2		0.966323	0.967656	0.969040
0.5	2	1	2.7	-10	**.05**	0.970627	0.971503	0.972407
					0.15	0.959883	0.960266	0.960672
					0.25	0.950152	0.950011	0.949887

Tables 1 and 2 exhibit the numerical values to the local skin-friction C_f and local Nusselt number Nu respectively. It has been observed empirically that for any particular values of R, K, Pr, θ_r and Ec the local skin-friction decreases with the increase in the magnetic parameter M. The skin friction is also decreases with the increase in Ec and the Prandtl number Pr. But the reversal trend is observed in the presence of fluid viscoelasticity K_1, K, θ_r and the thermal radiation R. It is worthwhile to mention here that the rate of heat transfer decreases with the increasing values of R, K, θ_r and Ec. However, the heat transfer rate increases with the increasing values of Prandtl number Pr and the viscoelastic parameter K_1.

4. **Conclusions**

In this paper a theoretical analysis has been done to study the effect of radiation on flow of Second grade fluid over a Stretching sheet through porous medium with temperature dependent viscosity and thermal conductivity. Some conclusions of the study are as below:

a. Velocity increases with the increase in magnetic parameter M, Eckart number Ec, Prandtl number $\Pr$ and viscoelastic parameter K_1.

b. Velocity decreases when radiation parameter R, viscosity parameter θ_r and porosity parameter K increases.

c. Skin friction decreases when magnetic field parameter M, Ec and Prandtl number $\Pr$ increases.

d. Skin friction increases when radiation parameter R, visco elastic parameter K_1 and viscosity parameter θ_r increases.

e. Temperature increases when radiation parameter R is increased. But temperature decreases when Prandtl number $\Pr$ and visco elastic parameter K_1 increases.

f. Nusselt number increases when Prandtl number $\Pr$ and visco elastic parameter K_1 increases.

References

[1] B.C. Sakiadis. Boundary-layer behaviour on continuous solid surface, J. AIChe. 7: 26-28, 1961.
[2] H. Blasius. Grenzschechten in Flussigkeiten mit kleiner Reibung, Z. Math. U. Phys., 56: 1-37(English translation), NACATM 1256, 1908, .
[3] L. J. Crane. Flow past a stretching sheet, Z. Appl. Math. Phys., 21: 645-647, 1970.
[4] M.E. Ali. On thermal boundary layer on a power-law stretched Surface with suction or injection, Int. J. Heat Fluid Flow, 16: 280-290, 1995.
[5] K. R. Rajagopal, T. Y. Na and A. S. Gupta. Flow of a visco-elastic fluid over a stretching sheet, Rheological Acta. , 23: 213-215, 1984..
[6] B. S. Dandapat and A. S. Gupta. Flow and heat transfer in a viscoelastic fluid over a stretching sheet, International Journal of Non-linear Mechanics, 24: 215-219. 1989.
[7] G. C Shit. Hall effects on MHD free convective flow and mass transfer over a stretching sheet, International Journal of Applied Mathematics. 5(8): 22-38, 2009. .
[8] P. Reddaiah and D.R.V. Prasada Rao. Convective heat and mass transfer flow of a viscous fluid through a porous medium in an elliptic duct – by finite element method. International Journal of Applied Mathematics and Mechanics, 8(10): 1-27, 2012.
[9] FM. Hady and R. A. Mohamed. *Applied Mathematics and computation*, 80: 1-26, 1994.
[10] M. A. Mansour. *Applied Mechanics and Engineering*, 2: 405-413, 1997.
[11] A .A. Mohamadien, M A Mansour, Abd El-Gaid S, and R SR.Gorla. *Transport in porous media*, 32: 263-283, 1998.
[12] A. Y. Bakier and RSR Gorla. *Transport in porous media*, 23: 357-363.1996.
[13] VR.Prasad, R . Muthucumaraswamy and B . Vasu. Radiation and Mass transfer effects on unsteady MHD free convection flow past a vertical porous plate embedded in porous medium: a numerical study, Int. J. of Appl. Math and Mech., 6(19): 1-21, 2010.
[14] S.P. Anjali Devi and M. Kayalvizhi. Analytical solution of MHD flow with radiation over a stretching sheet embedded in a porous medium, Int. J. of Appl. Math and Mech., 6(7): 82-106, 2010.
[15] I. A. Hassanein, A. Essawy. and N. M. Morsy. Variable viscosity and thermal conductivity effects on heat transfer by natural convection from a cone and a wedge in porous media, Arch. Mech. 55: 345-356, 2004.
[16] U. Sarma and G. C. Hazarika. Effects of variable viscosity and thermal conductivity on heat and mass transfer flow along a vertical plate in the presence of a magnetic field, Lat. Am. J. Phys. Educ., 5 (1): 100-106. 2011.
[17] F. C. Lai, and F. A. Kulacki. The effect of variable Viscosity on convective Heat and Mass Transfer along a vertical Surface in Saturated Porous media, International Journal of Heat and Mass Transfer., 33: 1028-1031. 1991.
[18] A. J. Chamakha. Hydromagnetic natural convection from an isothermal inclined surface adjacent to a thermally stratified porous medium, International Journal of Engineering Science, 35: 975 – 986. 1997.

Decision Support System for Patient Care

[1,]Kulvinder Singh Mann, [2,]Avneet Kaur, [3,]Mohit Sudhera
[1,2,3,]Punjab Technical University Guru Nanak Dev Engineering College
Dean T& P Ludhiana-141006,India

Abstract

With rapid development of medical information technology, Electronic Health Record(EHR) provide basis for various health services.This paper presents Decision Support System(DSS) for patient care with help of EHR. This paper tells how this DSS can be helpful for both doctor and patient.To check the effectiveness of this DSS a small survey was done.From results of the survey it is evident that quality of healthcare delivery can be improved by using this DSS.The investigation revealed that it prevent handwritten prescription risk,allow quick access during emergencies & can also be helpful for patient in remote area where doctor is not available.By this means,we provide our consumers an overall scene of the patient's personal history,personal health status & future care plans. This paper tells how DSS for patient care is helful for making better decisions by doctor i.e. by spending less time in management and more with patients & by patient i.e. with application of "Six in one module" function in system.

Keyword- Decision Support System ,Electronic Health Record

1. Introduction

The major problems confronting clinics and many hospitals are increasing population, management of large amount of medical records, medical errors & uneasy access of healthcare information.This directly effect patient care & health.To sort such problems can be time consuming if done manually that's why demand for quality and safe health care decision support system softwares are increasing.

With a rapid development of medical informationization,more

Fig. 1. Health record storage systems

and more eyes are attracted to the EHR [1].An Electronic

Health Record(EHR) is basically a patient record that resides in a computer system specifically designed to support healthcare providers by providing accessibility to patient data, medical alerts,reminders,clinical decision support systems, links to medical knowledge,for observations and other aids.Use of Medical Information Technology in Healthcare ,especially Electronic health record, can potentially improve and maintain healthcare.A Decision Support System(DSS) is a computer program application that anatomize data and presents it so that users can make decision more easily.It is an informational application.A Decision Support may present information graphically and may include an expert system or artificial intelligence. Healthcare extends beyond one person, in one department, in one building,in one sector.It is an active process that requires communication, collaboration, and

decision making across care providers and care setting.An EHR as DSS for patient care offer solutions that break down barriers to help you to solve business problems,improve patient safety, strengthen the revenue cycle,help in decision making and enable technology to improve patient care.Also it aimed at satisfying the requirement of Community Health Centre(CHC) and solving the problem of lacking mobility.This paper therefore studies the working process of end users and tries to uncover the real nature of medical work by introducing relevance of EHR.

2. Decision Support System For Patient Care

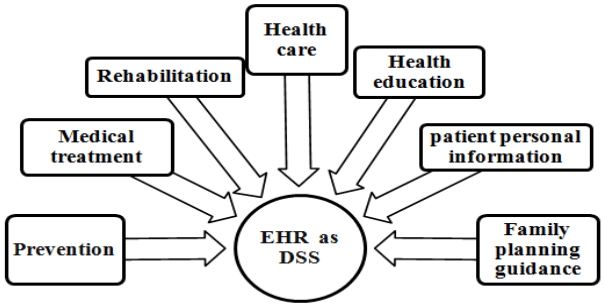


Fig. 2. Healthcare scenario

The Six in One is the emphasis of CHS, including Pre- vention,Medical treatment, Rehabilitation,Health care,Health education and Family Planning guidance.The six functions are not divided,but a comprehensive service combining with all functions in the module of Six in One.We had embedded the Six in One function in the EHR system as DSS.This EHR system could provide the corresponding to the differ- ent individual resident health,so as to be more individua- tion.Like for sudden illness the EHR as DSS, because of the portability and mobility,could collect and record health information.Time could be saved by this.It is also helpful for chronic and non-communicable disease in remote areas by providing diagnosis, treatment and medical alert.Health Education(including family planning guidance) is considered as the main contents of CHS.Then the health protection knowledge could be pervaded. The EHR is defined and divided into 4 categories i.e.Institutional EHR,Shared EHR,Personal EHR and Population EHR.The successful deployment of ap- propriate EHR requires both functional and semantic interop- erability and security and privacy protection with applications of relevant standards HL7 CDA[3],Clinical Decision Sup- port Systems,Evidence- Based Medicine, Individual- Based Medicine.The application of international Health Informatics standards is essential for a successful EHR development. The content of an EHR consist of administrative and clin- ical data.The content should be comprehensive and expres- sive,addressing all aspects of healthcare process for all re- lated disciplines and authorities.The administrative content in- cludes patients name,record number, food preferences,smoking and alcohol consumption [2].The clinical content includes symptoms,drugs prescribed,observations and lab reports. The availability of information is the expected value of an EHR as DSS that is agreed upon by all end users.Our project EHR as DSS is completely patient-centric.It keeps medical information safe and secure.It helps to make better decisions.It allows you to spend less time in management and more time with patients.It can cause reduction in chaos in hospitals or clinics during peak hours.It let you get a specialist and colleagues opinion.It prevent handwritten prescription risk.It allow quick access and response during emergencies.It can alert you to potential adverse drug reactions.It will provide you information confirmed from specialist even if you are present in remote area where there is no doctor. For example when an individual goes to see a doctor for a specific condition or care, a medical record is constructed containing information such as personal and social history,a physical notes made by doctor, consultations,lab or image results from other health care providers and so forth.In our application information typed into such an electronic medical record; paperless records that contain health care and medical information just as paper medical records,but take up such less space and are available in electronic formats,which makes them accessible via palm pilots,desktop application,web application etc. that connect doctors office,hospitals and clinics. Unanimously, EHR as DSS is a platform and technology independent standard.It reduces paper work.It facilitate better patient care. It reduces labour and time.It gives you flexibility as it can be implemented using a variety of software technologies to suit your information needs.

3. Methodology

The Electronic Health Record(EHR) as DSS for patient care is the keystone of a medical information system.In India,the IT adoption in Healthcare is estimated to be only twenty percent and EHR adoption in government healthcare facilities is very slow but the private sector is aggressive in their plans.To gain insight into the functioning of healthcare centers with respect to use of information technology and their effectiveness in health care delivery, a survey was done.Thirty(possible) end users of DSS for patient care were interviewed about their process and the expected value of EHR.The interviews took in 5 different hospitals.Beside medical specialists other medical staff & patients was also interviewed.Questions was based on patient load,medical record formats,hospital infrastructure,daily routines and staffing information.Responses were tabulated and were used to depict results and draw inferences.Factors which end users find relevant for an EHR as DSS are Availability of information,Less administrative work, Analyses,Uniformity of working processes, Reliability,Quality of care,Collaboration with colleagues,Time,& Just being a good doctor.The availability of information is the expected value of an EHR as DSS for patient care that is agreed upon by all end users.

Also during survey large amount of medical information

Fig. 3. Patient's Health-Record Management

collected.EHR as DSS for patient care can be helpful for doctor i.e. by recording patient health status as well as for patient in remote areas or not in condition to visit to any clinic i.e. tells various possibilities of communicable or non-communicable diseases [6] & also various internal or external injuries from symptoms which user will choose from given options.From patient point of view EHR as DSS for patient care is helful for all five categories:Newborn baby; child;teen;adult;senior[5].It can contain lab tests results which can tell the current status of patient.For example diabetic patient test his/her blood sugar from diabetic tester and then can check his/her level from lab results.It includes best diet charts for different categories.Due to portability and security java language can be chosen to make this system more effective and successful.From Doctor point of view it can keep all medical records safe and secure and allow quick access during emergencies[4].

4. Feasibility Result

Feasibility studies aim to objectively and rationally uncover the strengths and weaknesses of an existing business or proposed venture,opportunities and threats as presented by the environment,the resources required to carry through, and ultimately the prospects for success.In its simplest terms, the two criteria to judge feasibility are cost required and value to be attained. Today,healthcare organizations still heavily depend on paper-based medical that are the least secure form of health records,whether in hands of a patient or a medical provider. Paper is easily misplaced,lost,intercepted and read by unauthorized parties[7].In the digital era,Electronic Health Records are replacing paper based records.Use of Information Technology in healthcare, especially electronic health records,can potentially improve healthcare.Our project EHR as DSS for patient care is completely patient-centric.It keeps medical information safe and secure.It helps to make better decisions.It allows you to spend less time in management and more time with patients.It can cause reduction in chaos in hospitals or clinics during peak hours.It let you get a specialist and colleagues opinion.It prevent handwritten prescription risk.It allow quick access and response during emergencies.It can alert you to potential adverse drug reactions. It will provide you information confirmed from specialist even if you are present in remote area where there is no doctor.Also with help of EHR as DSS healthcare awareness of patient increases,demographic information will be available for planning better health care delivery and reduces risk of loosing research work on certain observations[8]. Unanimously,EHR as DSS is a

platform and technology independent standard.It reduces paper work & uses cost effective approaches. It facilitate better patient care.It reduces labour and time.It gives you flexibility as it can be implemented using a variety of software technologies to suit your information needs.The availability of information is the expected value of an EHR as DSS that is agreed upon by all end users i.e. early prevention is the foundation of digital health care system for community health service.

5. Conclusions & Discussions

In this paper we discuss how decision support system is useful for both doctor and patient.First we represent this DSS as patient-centric system.Then we further divided into two modules i.e. first from patient point of view & second from doctor point of view.First one based on decisions made by patients by clicking on symptoms of their disease.Then from graphical decision support system, various preventive measures are provided & even in case of emergencies contact number of specialist is also provided.Basically first module is based on content management.Second one represent all kinds of EHR data as clinical acts under unified structure.This provide some fundamental visualisation forms for each kind of EHR data.This depicts overall situation of patient.It just not only help the clinicians in their daily work but also useful during emergencies. Clincians & Patients told us that it was very useful and helpul for them to view overall health status of any particular patient.However they still have many more requirements[1].They need an integrated viewer which provide more information and more flexible visualization.Also the need artificial intelligence approach instead of graphical approach in content management.We will make more detailed analysis of all clinical acts and on neural and finger-print systems and design more visualization forms to satisy these new requirements in the future[9].

References

[1] DeWar,C.,Bring the EHR to life. Nursing Management,2006(January): p.31-35

[2] Faber, M.G.,Design and Introduction of an Electronic Patient Record: How to involve the Users? Methods of Information in Medicine,2003,42(4): p.371-375

[3] HL7 Reference Information Model,Health Level Seven, Inc.http:// www. hl7.org. Accessed at August 18, 2012

[4] Powsner SM, Tufte ER.Graphical summary of patient status. The Lancet,1994,344:386-389.

[5] BOOTSTRAP: Health Tools,Health Information For Whole Family,Inc., http:// familydoctor.org . Accessed at September 1,2012

[6] Powsner SM, Tufte ER.Graphical summary of patient status. The Lancet,1994,344:386-389.

[7] The state HIPAA privacy and security compliance.AHIMA.Apr.2005.

[8] Institute of Medicine.2003.Key Capabilities of an Electronic Health Record System.Letter Report.

[9] Integrating the Healthcare Enterprise Expands Strategy on Interoperabil- ity & Standards Implementation, November 1,2006

Implementation of an OFDM FFT Kernel for WiMAX

Lokesh C. [1], Dr. Nataraj K. R.[2].

[1] Assistant Professor,Department of Electrical and Electronics Engineering.Vidyavardhaka College of Engineering, Mysore, Karnataka, India.

[2] Professor, Department of Electronics and Communications Engineering.SJB Institute of Technology, Bangalore, Karnataka, India.

Abstract:

In this paper we focus on the OFDM Kernel which refers to the inverse fast Fourier transform and cyclic prefix insertion blocks in the downlink flow and the FFT and cyclic prefix removal blocks in the uplink flow. To support orthogonal frequency-division multiple access (OFDMA) an extension to the OFDM kernel is required that allows each user to be allocated with a portion of the available carriers. This process is referred to as sub channelization. The WiMAX building blocks include bit-level, OFDMA symbol-level, and digital intermediate frequency processing blocks. For bit-level processing, Altera provides symbol mapping/demapping reference designs and support for forward error correction using the Reed- Solomon and Viterbi MegaCore® functions. The OFDMA symbol-level processing blocks include reference designs that demonstrate subchannelization and desubchannelization with cyclic prefix insertion supported by the fast Fourier transform, and inverse fast Fourier transform MegaCore functions. Other OFDMA symbol-level reference designs illustrate ranging, channel estimation, and channel equalization. The digital IF processing blocks include single antenna and multiantenna digital up converter and digital down converter reference designs, and advanced crest-factor reduction and digital predistortion.

Keywords: inverse fast Fourier transform (IFFT), orthogonal frequency-division multiple access (OFDMA), intermediate frequency (IF), forward error correction (FEC), digital up converter (DUC), digital down converter (DDC), crest-factor reduction (CFR), digital predistortion (DPD, WiMAX (Worldwide Interoperability for Microwave Access).

1. Introduction

The Altera® orthogonal frequency division multiplexing (OFDM) kernel can be used to accelerate the development of wireless OFDM transceivers such as those required for the deployment of mobile broadband wireless networks based on the *IEEE 802.16* standard. OFDM is one of the key physical layer components associated with mobile worldwide interoperability for microwave access (WiMAX) and is widely regarded as an enabling technology for future broadband wireless protocols including the 3GPP and 3GPP2 long term evolution standards.

The OFDM kernel has the following key features:

- Support for 128, 512, 1K, and 2K FFT sizes to address variable bandwidths from 1.25 to 20 MHz
- Parameterizable design
- Optimized for efficient use of Cyclone II, Stratix II, and Stratix III device resources

2. Introduction to WiMAX

WiMAX (Worldwide Interoperability for Microwave Access) is a wireless communications standard designed to provide 30 to 40 megabit-per-second data rates, with the 2011 update providing up to 1 Gbit/s for fixed stations. WiMAX refers to interoperable implementations of the IEEE 802.16 family of wireless-networks standards ratified by the WiMAX Forum. Similarly, Wi-Fi, refers to interoperable implementations of the IEEE 802.11 Wireless LAN standards certified by the Wi-Fi Alliance. WiMAX Forum certification allows vendors to sell fixed or mobile products as WiMAX certified, thus ensuring a level of interoperability with other certified products, as long as they fit the same profile.

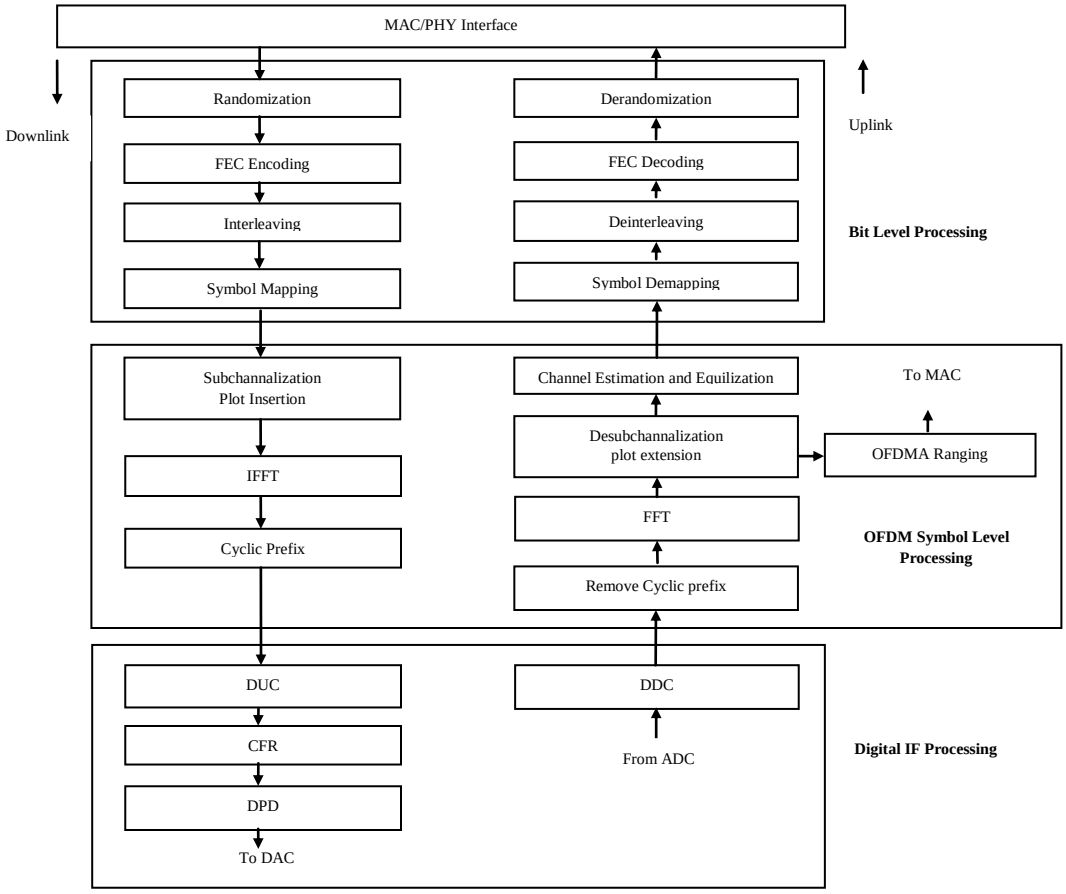


Figure 1. *WiMAX Physical Layer Implementatio*, an overview of the *IEEE 802.16e-2005* scalable orthogonal frequency-division multiple access (OFDMA) physical layer (PHY) for WiMAX basestations.

Altera's WiMAX building blocks include bit-level, OFDMA symbol-level, and digital intermediate frequency (IF) processing blocks. For bit-level processing, Altera provides symbol mapping/demapping reference designs and support for forward error correction using the Reed- Solomon and Viterbi MegaCore® functions. The OFDMA symbol-level processing blocks include reference designs that demonstrate subchannelization and desubchannelization with cyclic prefix insertion supported by the fast Fourier transform, and inverse fast Fourier transform MegaCore functions. Other OFDMA symbol-level reference designs illustrate ranging, channel estimation, and channel equalization. The digital IF processing blocks include single antenna and multiantenna digital up converter and digital down converter reference designs, and advanced crest-factor reduction and digital predistortion.

3. Introduction to OFDM Kernal

The OFDM Kernel refers to the inverse fast Fourier transform (IFFT) and cyclic prefix insertion blocks in the downlink flow and the FFT and cyclic prefix removal blocks in the uplink flow. To support orthogonal frequency-division multiple access (OFDMA) an extension to the OFDM kernel is required that allows each user to be allocated with a portion of the available carriers. This process is referred to as subchannelization. The physical layer is based around OFDM modulation. Data is mapped in the frequency domain onto the available carriers. For this data to be conveyed across a radio channel, it is transformed into the time domain using an inverse fast Fourier transform (IFFT) operation. To provide multipath immunity and tolerance for synchronization errors, a cyclic prefix is added to the time domain representation of the data. Multiple modes are supported to accommodate variable channel bandwidths. This scalable architecture is achieved by using different FFT/IFFT sizes. This reference design supports transform sizes of 128, 512, 1,024, and 2,048.

4. Implementing OFDM Kernal for WiMAX

FPGAs are well suited to FFT and IFFT processing because they are capable of high speed complex multiplications. DSP devices typically have up to eight dedicated multipliers, whereas the Stratix III EP3SE110 FPGA has 112 DSP blocks that offer a throughput of nearly 500 GMACs and can support up to 896 18x18 multipliers, which is an order of magnitude higher than current DSP devices.

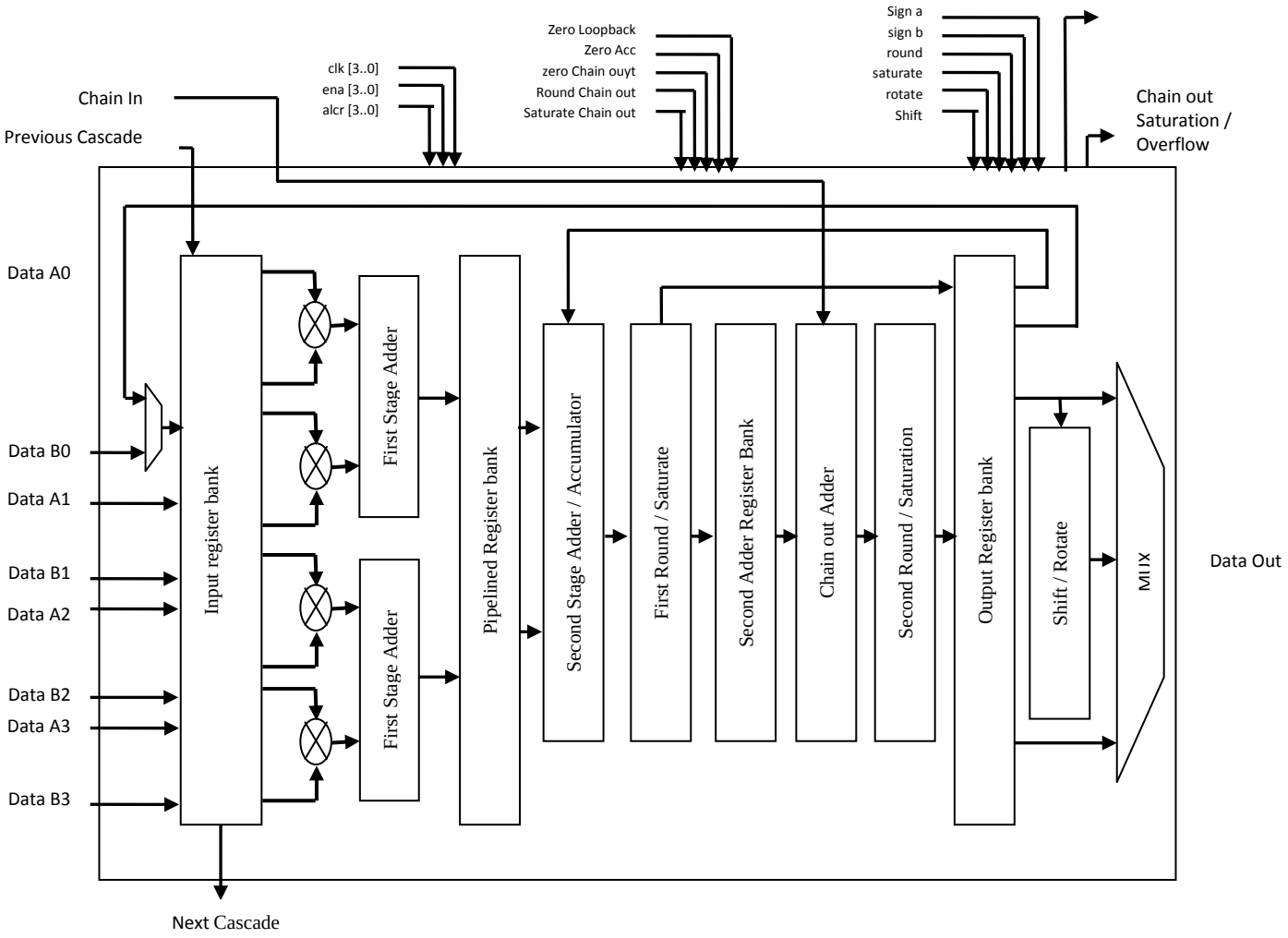


Figure2. Embedded DSP Blocks Architecture in Stratix III Devices

Such a massive difference in signal processing capability between FPGAs and DSP devices is further accentuated when dealing with basestations that employ advanced, multiple antenna techniques such as space time codes (STC), beam forming, and multiple-input multiple-output (MIMO) schemes. The combination of OFDMA and MIMO is widely regarded as a key enabler of higher data rates in current and future WiMAX and 3GPP long term evolution (LTE) wireless systems. When multiple transmit and receive antennas are employed at a basestation, the OFDMA symbol processing functions have to be implemented for each antenna stream separately before MIMO decoding is performed. The symbol-level complexity grows linearly with the number of antennas implemented on DSPs that perform serial operations. For example, for two transmit and two receive antennas the FFT and IFFT functions for WiMAX take up approximately 60% of a 1-GHz DSP core when the transform size is 2,048 points. In contrast, a multiple antenna-based implementation scales very efficiently when implemented with FPGAs. Using Altera devices, we can exploit parallel processing and time-multiplexing between the data from multiple antennas. The same 2×2 antenna FFT/IFFT configuration uses less than 10% of a Stratix II 2S60 device.

5. Functional description of OFDM kernal for WiMAX

Altera provides the reference design as clear text VHDL. The reference design also demonstrates the use of the FFT MegaCore function. To accelerate integration with Altera intellectual property (IP) or other subsystems, the interfaces support the Altera Avalon® Streaming (Avalon-ST) interface specification. Altera has verified the RTL behavior against a fixed point model of the algorithms. The reference design includes RTL testbenches that stimulate the designs using golden reference data from the fixed point model and check for correct functionality. The OFDM kernel handles the FFT operations and cyclic prefix addition and removal. The FFT size is a parameter that we must specify at synthesis time, but we can change the guard interval at run time.

Downlink Transmit

The downlink OFDM kernel module performs an inverse Fourier transform of the frequency domain input data and adds a cyclic prefix to the resulting time domain data. The cyclic prefix addition block contains a controller that buffers the output packets from the FFT, and adds the appropriate proportion of the end of the output packet to the beginning of the output packet. As this requires a fairly significant memory resource, the hardware architecture has been designed so that the embedded memory may be shared with the uplink OFDM kernel if the modem is operating in time division duplex (TDD) mode.

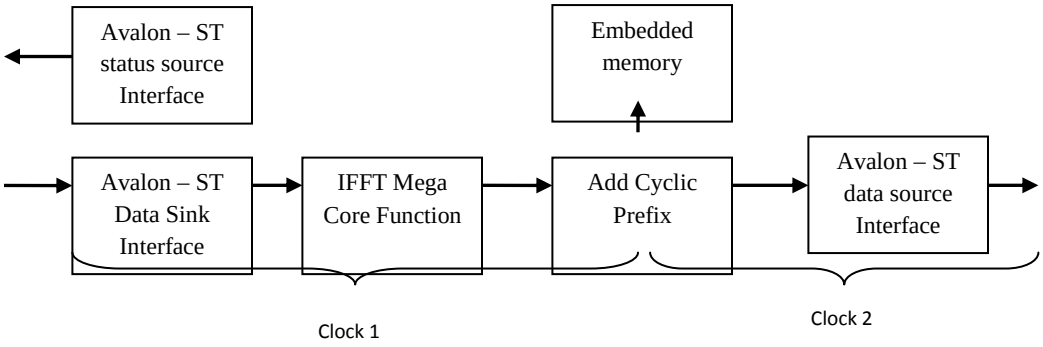


Figure 3 shows a block diagram of the downlink OFDM kernel.

Interface Specifications

The block has two clock domains. In addition, there are two reset ports; one for each clock domain. The reset ports are active low. Figure 4 shows the downlink OFDM kernel interfaces.

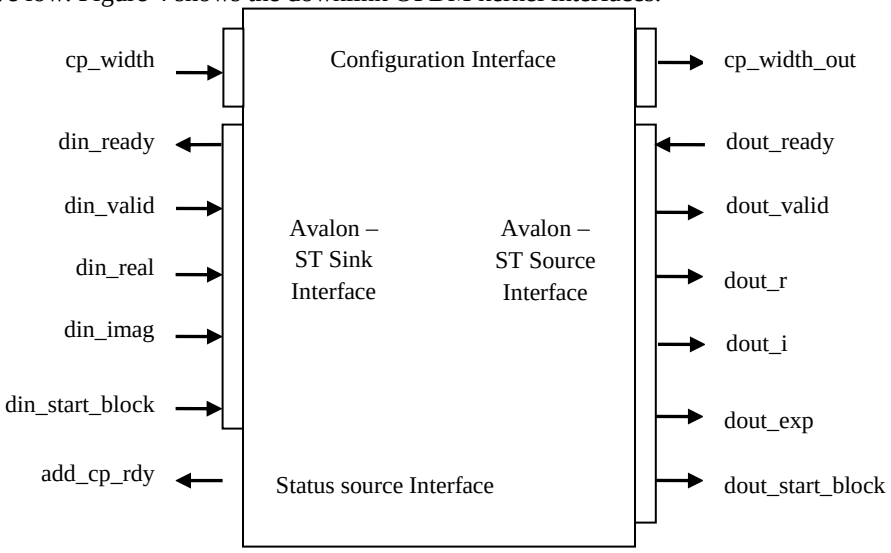


Figure4. Downlink OFDM Kernel Interfaces

The input interface has the following features:

- Avalon-ST data sink and status source
- Ready signal latency of one cycle—the earliest time valid data may be presented to the block after ready has been signaled is one clock cycle

The output interface has the following features:

- Avalon-ST data source
- Ready signal latency of four cycles—the block responds to new data or stops delivering data four cycles after an event on the ready signal
- Support for back pressure and Dynamically changeable cyclic prefix

Uplink Receive

The uplink OFDM kernel module performs an FFT of the time domain input data and removes the cyclic prefix. The Avalon-ST start of packet pulse should specify the start of the cyclic prefix. The remove cyclic prefix block ignores the data during the cyclic prefix and writes the remaining samples to the FFT input buffer.

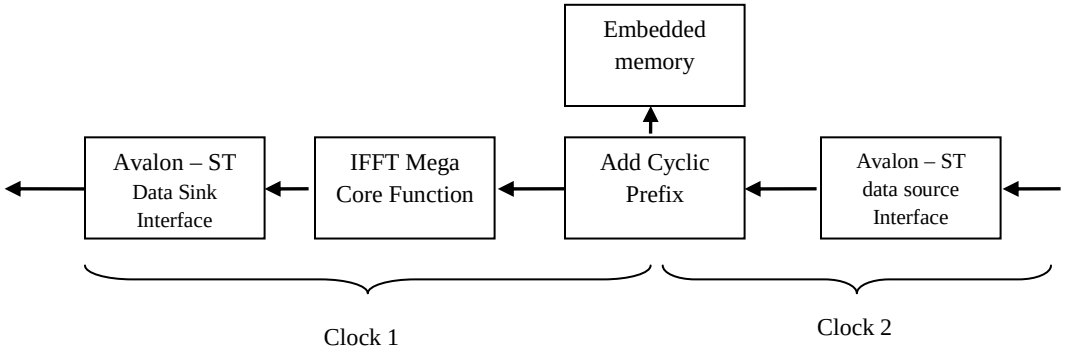


Figure5. Uplink OFDM Kernel Block Diagram

Because the channel characteristics can change, it is possible that the start of the packet pulse is not always after the start of the cyclic prefix time. The hardware has been designed to deal with this scenario but with the constraints that the variation of the pulse must be within the cyclic prefix time and that the start pulse will not be before the preceding symbol has been fully clocked in.

Uplink Interface Specifications

The block has two clock domains. In addition, there are two reset ports; one for each clock domain. The reset ports are active low.

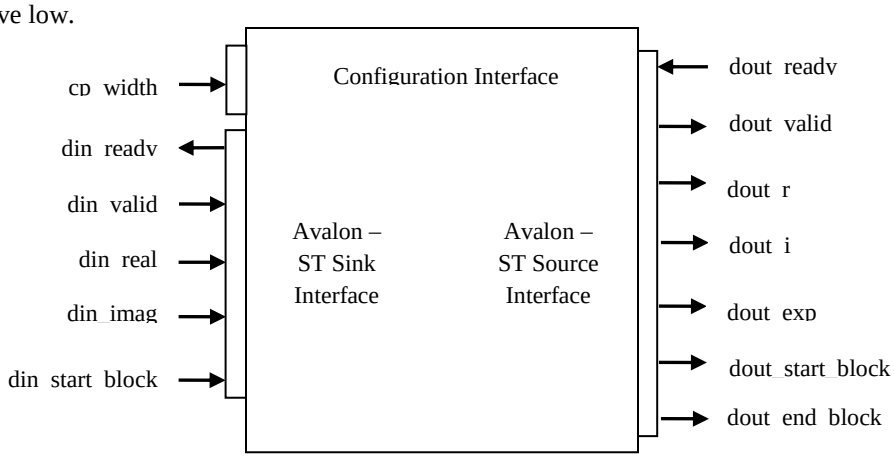


Figure6. Uplink OFDM Kernel Interfaces

The input interface has the following features:

- Avalon-ST data sink
- Ready signal latency of one cycle
- Does not apply back pressure since data is continuous from RF card

The output interface has the following features:

- Avalon-ST data source
- Ready signal latency of one cycle
- Does not accept back pressure from downstream sink
- Dynamically changeable cyclic prefix

FFT MegaCore Function

The FFT MegaCore function is capable of performing both the forward and inverse transform. The hardware architecture is chosen to minimize the resource usage and has the following parameters:

- Burst mode
- Single output engine
- Single instance of engine
- 16-bit internal and data input/output precision widths

In addition, design implements two clock domains so that it is possible to exploit time sharing and minimize resource utilization in the FFT MegaCore function by running Clock 1 faster than Clock 2. The FFT MegaCore function generates block floating point output data and the output dynamic range is maximized for the given input and output data widths.

Clock Requirements

The clocking requirements are as follows:

- The two clock domains must be synchronous
- The minimum Clock 2 frequency is the data sampling frequency given in Table 1. This would lead to a constant output from the FFT MegaCore function

FFT Points	Bandwidth (MHz)	Clock 2 (MHz)
128	1.25	1.429
256	2.5	2.857
512	5	5.714
1,024	10	11.429
2,048	20	22.857

Table 1. Minimum Clock 2 Rate

- The Clock 2 frequency may equal or exceed the Clock 1 frequency
- The Clock 1 requirements are dictated by the FFT MegaCore function and are summarized in Table 2

FFT Points	Required data rate (MHz)	FFT throughput (Cycles/ N block)	Clock 1 minimum Speed (MHz)
128	1.429	858	9.579
256	2.857	1,626	18.146
512	5.714	3,693	41.214
1,024	11.429	7,277	81.220
2,048	22.857	16,512	184.285

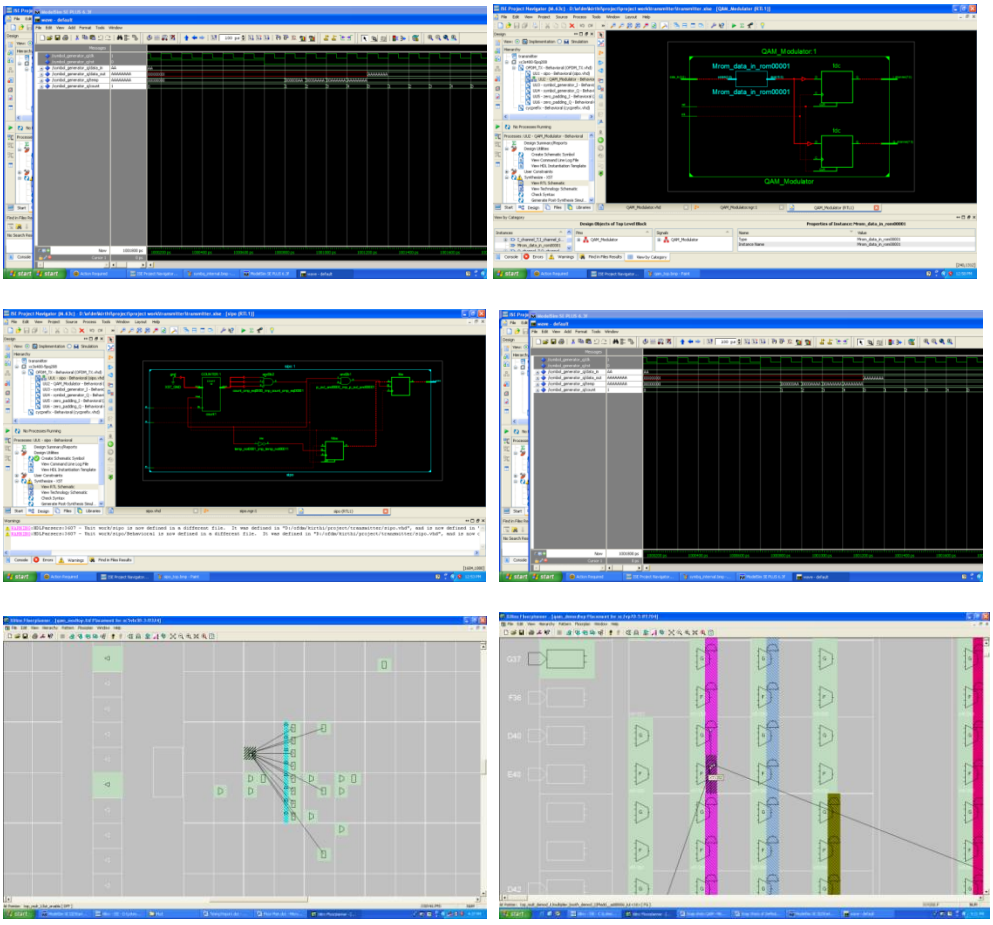
Table 2. Clock 1 Requirements

- Table 2 ignores the cyclic prefix effect, which reduces the Clock 1 speed requirement slightly
- The Clock 1 minimum speed = throughput/*N* × data rate

6. Conclusion

This application note has outlined the advantages of using Altera FPGAs for implementing OFDM systems such as an IEEE 802.16e deployment. A flexible, high-throughput DSP platform needs an FPGA-based implementation platform. In addition, this reference design demonstrates the implementation of a key function that may be exploited to facilitate rapid system deployment.

7. Results



Symbol generator simulation output of an OFDM kernel for WiMAX as obtained in Model Sim®

References

[1] K. Fazel and S. Kaiser, *Multi-Carrier and Spread Spectrum Systems: From OFDM and MC-CDMA to LTE and WiMAX*, 2nd Edition, John Wiley & Sons, 2008, ISBN 978-0-470-99821-2

[2]. M. Ergen, *Mobile Broadband - Including WiMAX and LTE*, Springer, NY, 2009 ISBN 978-0-387-68189-4

[3]. Carl Weinschenk (April 16, 2010). "Speeding Up WiMax". *IT Business Edge*. "Today the initial WiMax system is designed to provide 30 to 40 megabit-per-second data rates."

[4]. Wimax Forum Industry Research Report http://www.wimaxforum.org/sites/wimaxforum.org/files/page/2011/03/Monthly_Industry_Report_March2011.pdf

[5] Synthesis of band-limited orthogonal signals for multi-channel data transmission, *Bell System Technical Journal* 46, 1775-1796.

[6] *MATLAB 7.8.0 (R2009a), Help.* s.l. : MathWorks, Inc., 2009.

[7] van Nee, R. and Prasad, R**.** *OFDM for Wireless Multimedia Communications.* s.l. : Artech House, 2000.

[8] Charan Langton**.** OFDM. *Intuitive Guide to Principles of Communications.* http://www.complextoreal.com/.

[9] Amalia Roca Persiva **.** *Implementation of a WiMAX simulator in Simulink.* Institute of Communications & Radio-Frequency Engineering, Vienna University of Technology. Vienna : s.n., 2007. Master Thesis.

Study and comparison of various point based feature extraction methods in palmprint authentication system

Vinod Kumar D[1], Dr. Nagappan A[2]
[1] Department of ECE & CSE,
[2] Principal, V.M.K.V. Engineering College,
Vinayaka Missions Research Foundation Deemed University
Salem, Tamilnadu, India

Abstract

Biometrics is the word derived from Greek words "Bio" means (life) & "Metrics" means (to measure). Biometrics is the science of measuring human characteristics for the purpose of authenticating or identifying the identity of an individual. Biometrics System is used for automated recognition of an individual. In Information system in particular biometrics is used as a form of identity access management and access control. It is also used to identify individuals in groups that are under surveillance. In this paper, palmprint biometric is used for personal authentication. Various feature extraction methods to be discussed and compared are Forstner operator, SUSAN operator, Wavelet based salient point detection and Trajkovic and Hedley corner detector.

Keywords: Biometric, Corner, Palmprint, Palmprint features.

1. Introduction

The palmprint of a person can be also taken as a biometric. It is a physiological biometrics, palmprint, the inner part of a person's hand, below the fingers to the wrist meets out both the theoretical and practical requirements to be a biometric. Palmprint is universal because every person has palmprint. It is unique because every palmprint is different from other person palmprint even identical twins have different palmprint features. Palmprint is permanent or inseparable from individual as compared to identification items. It is easy to collect and consistent because it does not change much with time. It performs better in term of accuracy, speed and robustness. Palmprint biometric system is more acceptable by public because users can gain access anytime they want without being monitored by a surveillance camera. It is hard to imitate because of its size.

Palmprint is features rich. It consists of geometry features, point features, line features, texture features and statistical features. Palmprint point features are datum points, line intersection points, end line points etc. Point features can be obtained through high resolution palmprint image. In this paper, several point (corner) based feature extraction methods are studied and compared. Results of various feature extraction methods are compared and among them best method is found out.

2. Corner based feature extraction methods

Corners are the points where intensity changes in all directions. In palmprint, palmprint features can be of the form of corner points. These features can be extracted using Forstner operator, SUSAN operator etc. These methods are discussed in detail as follows:

2.1 Förstner Operator

The Förstner Operator developed by Förstner and Gülch in 1987 has been widely adopted in photogrammetry and computer vision over the last two decades. The aim of developing this operator is to create a fast operator for the detection and precise location of distinct points, corners and centres of circular image features. The algorithm identifies interest points, edges and regions using the autocorrelation function A. The derivatives of A are computed and summed over a Gaussian window. Error ellipses are computed and based on the size and shape properties of each ellipse, the interest points found are classified as points, edges or regions. Förstner calculates the size and shape of the error ellipses using two eigenvalues λ1 and λ2 as well as the inversion of A.

The error ellipse size is determined by:

$$w = \frac{1}{\lambda_1 + \lambda_2} = \frac{\det(A)}{trace(A)}, \quad w > 0$$

The roundness of the ellipse is determined by:

$$q = 1-\left(\frac{\lambda_1-\lambda_2}{\lambda_1+\lambda_2}\right)^2 = \frac{4.\det(A)}{trace(A)^2}, \quad 0 \le q \le 1$$

The algorithm classifies each area based on the values of w and q.

- Small circular ellipses define a salient point
- Elongated error ellipses suggest a straight edge
- Large ellipses mark a homogeneous area

Practically Förstner operator has few limitations like high computational cost, relatively slow and impractical for high-level data analysis.

2.2 SUSAN Operator

The SUSAN operator was developed by Smith and Brady (1997) for image processing. It is an edge and corner detector method which is accurate, noise resistant and fast. It is better as compared to other operators and overcome maximum problems faced by other methods, such as high computation time. The SUSAN operator is based on the concept that each point of interest in the image will have associated with it a local area of similar brightness values and that these areas can be used as a guide to help find features of interest such as corners and edges in the image. The SUSAN operator finds areas of similar brightness, and consequently for interest points within a weighted circular window. The central pixel in the search window is denoted as the nucleus. The area within the window that has similar intensity values to the nucleus is computed and referred to as the USAN (Univalue Segment Assimilating Nucleus). A low value for the USAN indicates a corner since the central pixel would be very different from its surroundings. After assessing results and eliminating outliers, the local minima of the SUSANs (smallest USAN) remain as valid interest points. The comparison between pixel brightness values is computed using the following equation:

$$o(\vec{p},\overrightarrow{p_0}) = \begin{cases} 1 \ if \ \left|I(\vec{p})-I(\overrightarrow{p_0})\right| \le t_b \\ 0 \ if \ \left|I(\vec{p})-I(\overrightarrow{p_0})\right| > t_b \end{cases}$$

where, $\overrightarrow{p_0}$ is the position of the nucleus in the two-dimensional image, $\overrightarrow{p}$ is the position of any other point within the circular window, I($\vec{p}$) is the brightness value of any pixel, t_b is the brightness value threshold and o is the output of the comparison. The comparison is calculated for each pixel in the circular window and the total number of pixels with similar brightness values as the nucleus is summarized as:

$$n\left(\overrightarrow{p_0}\right) = \sum_{\vec{p}} o\left(\vec{p},\overrightarrow{p_0}\right)$$

N($\overrightarrow{p_0}$) value is compared with a geometric threshold, g. The algorithm uses a threshold value in order to distinguish between features that make suitable interest points and non-suitable features. To find a corner in the image, the threshold value g is set to half of the maximum value of N($\overrightarrow{p_0}$), n_{max}. If $n_{max} < g$ then it indicates corner existence.

2.3 Wavelet based salient point detection

The wavelet transform is a multi-resolution representation of image variations at different scales. A wavelet is an oscillating and attenuated function. It is known that wavelet representation gives information about the variations in the signal at different scales. The aim is to extract salient points from the image where there is some variation in the signal at any resolution. A high wavelet coefficient at a coarse resolution corresponds to a region with high global variations. A relevant point is found out to represent this global variation by looking at wavelet coefficients at finer resolutions.

The coefficient represents 2p signal points. To select a salient point from this tracking, among these 2p points the one with the highest gradient is chosen. Saliency value is set as the sum of the absolute value of the wavelet coefficients in the track:

$$saliency = \sum_{k=1}^{-j} \left| C^{(k)}(W_{2^j} f(n)) \right|, 0 \le n < 2^j N, -\log_2 N \le j \le -1$$

where, $C(W_{2^j} f(n))$ is the wavelet coefficient, N is the length of the signal.

The tracked point and its saliency value are computed for every wavelet coefficient. A point related to a global variation has a high saliency value, since the coarse wavelet coefficients contribute to it. A finer variation also leads to an extracted point, but with a lower saliency value. The saliency value is thresholded, in relation to the desired number of salient points. The points related to global variations; local variations are obtained. The salient points extracted depend on the wavelet used.

2.4 Trajkovic and Hedley corner detector

Trajkovic and Hedley corner operator was developed by Miroslav Trajkovic and Mark Hedley in 1998. The operator is compared with other operators like Plessey operator etc and found that operator has slightly inferior repeatability rate, but the localization is comparable and improved on junctions. Trajkovic and Hedley proved empirically that their operator is five times faster than the Plessey operator and at least three times faster than all of the operators considered. Being fast method, the computation time is less. This operator is suitable for real-time applications because of minimal computational demands.

The cornerness measurement of Trajkovic operator is calculated by considering a small circular window and all the lines which pass through the center of the circle. Center of the circle is denoted by *C* and an arbitrary line that passes through *C* and intersects the boundary of the circular window at *P* and *P'*. Intensity at a point X is denoted by I_X and is summarized in Figure 1.

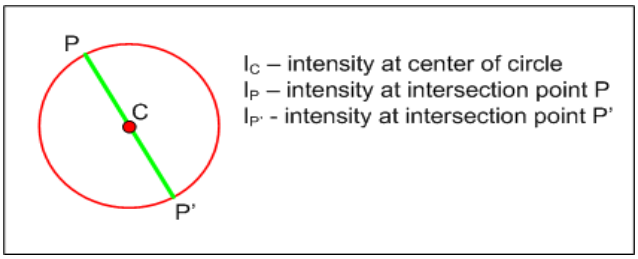


Figure 1: Notation for Trajkovic Operator

The cornerness measure for the Trajkovic operator is then given as:

$$C(x,y) = \min\left((I_P - I_C)^2 + (I_{P'} - I_C)^2 \right), \forall_{P,P'}$$

Different cases are studied to understand the cornerness measurement.

- *Interior Region* – it is clear from figure 2 for interior region that the majority of the circular window is within an interior region (i.e. region of near uniform colour) there will be at least one line where the intensity at the center of the circle I_C is approximately equal to I_P and $I_{P'}$. it is illustrated by the green lines in Figure 4.1, there is in general several lines where I_C is approximately equal to both I_P and $I_{P'}$. It concludes that cornerness measure will be low and robust to noise.
- *Edge* - for the case where the center of circle lies just on an edge there will be exactly one line, shown in green, where I_C is approximately equal to both I_P and $I_{P'}$. Since there is only one line where I_C is approximately equal to both I_P and $I_{P'}$ the cornerness measure along edges is susceptible to noise.
- *Corner* - for the case where the center of the circle is on a corner, for every line at least one of I_P or $I_{P'}$ will not be in the interior region so *should* be different than I_C. Therefore, the cornerness measure will be high and is not particularly robust to noise.
- *Isolated Pixel* - for the case of an isolated pixel, for every line both I_P and $I_{P'}$ will be different than I_C so the cornerness measure will be high. An isolated pixel is likely the result of noise.

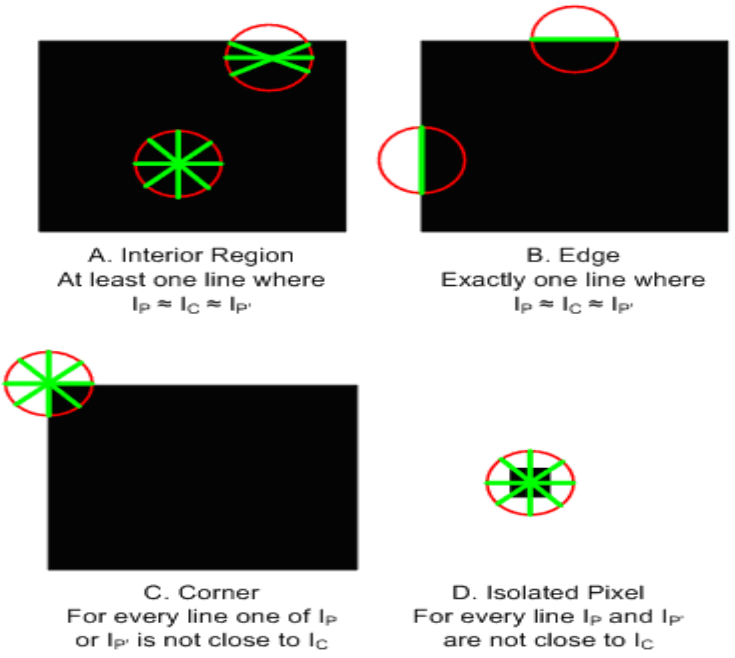


Figure 2: Different cases for the Trajkovic Operator

Based on the analysis, it is observed that Trajkovic Operator will only perform well for relatively clean images.

3. Results and Discussion

The proposed approach for personal identification using palmprint images is rigorously evaluated on palmprint image database. Figure (3, 4, 5, and 6) shows the sample images of palmprint images using several point based feature extraction methods.

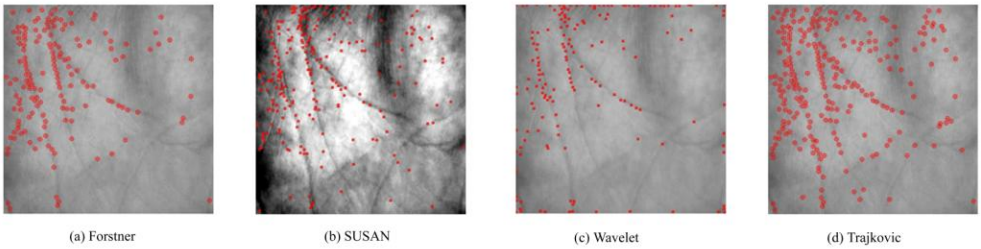


Figure 3: Point based feature extraction results for person 1 sample 1

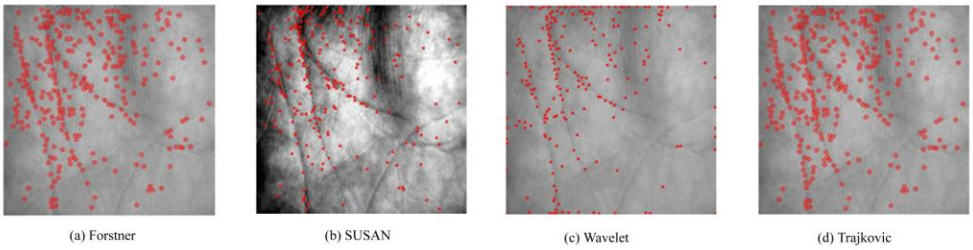


Figure 4: Point based feature extraction results for person 1 sample 2

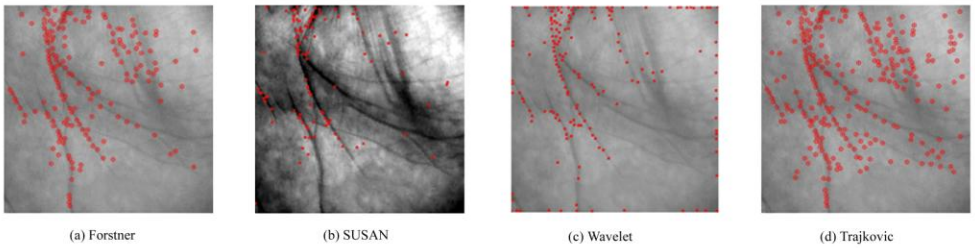


Figure 5: Point based feature extraction results for person 2 sample 1

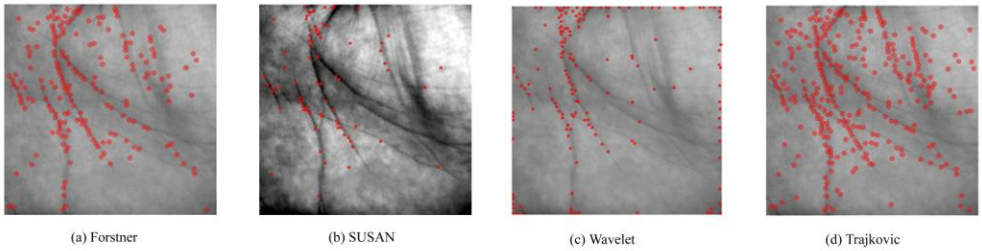


Figure 6: Point based feature extraction results for person 2 sample 2

The FAR, FRR and ROC curves for each feature extraction method is illustrated in following Figures (7, 8, 9, and 10).

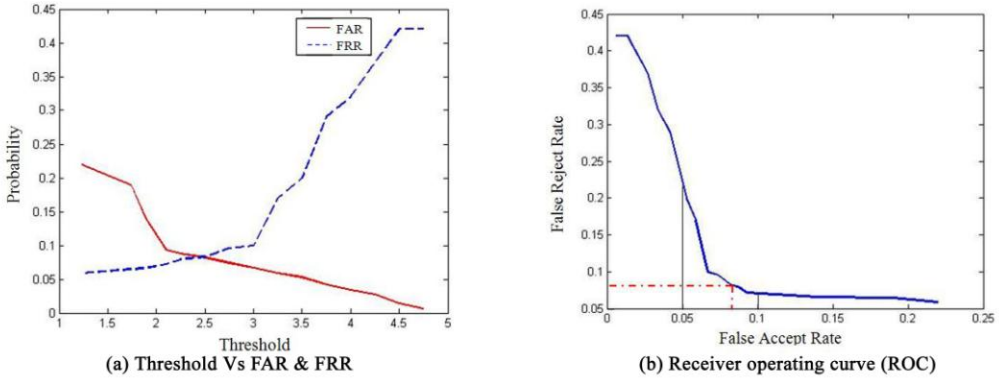


Figure 7: Accuracy plot for SUSAN operator

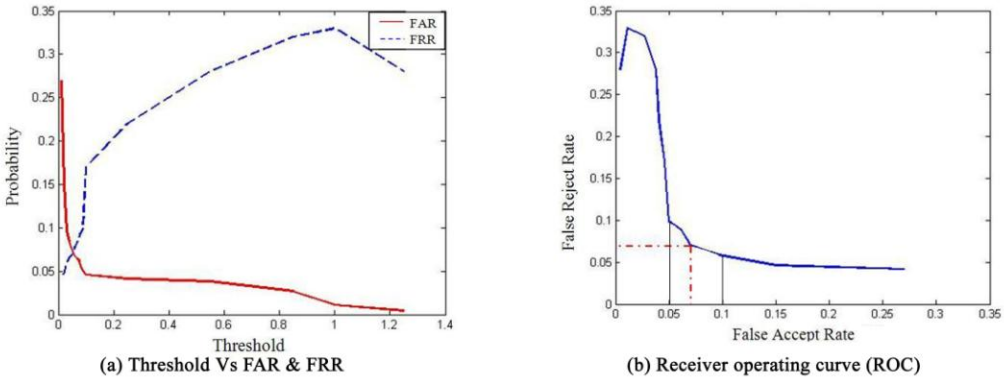
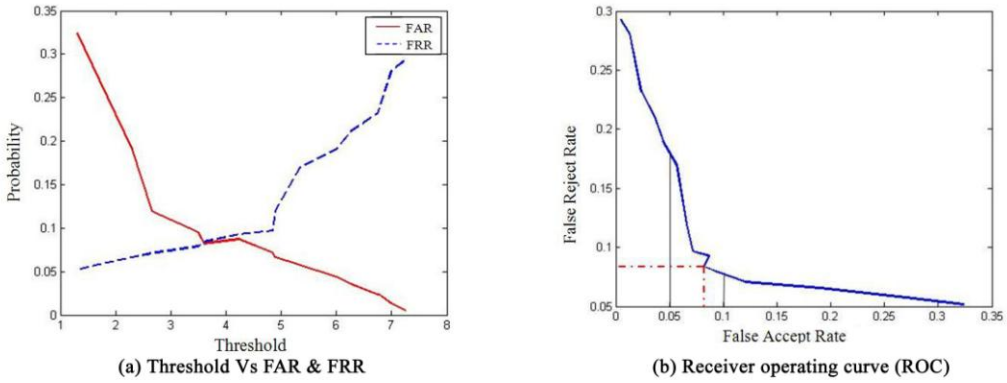
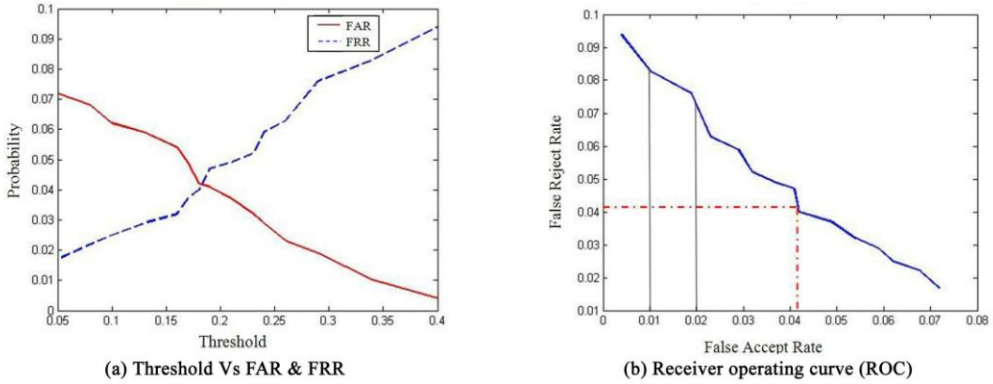


Figure 8: Accuracy plot for Wavelet operator



(a) Threshold Vs FAR & FRR (b) Receiver operating curve (ROC)

Figure 9: Accuracy plot for Trajkovic operator



(a) Threshold Vs FAR & FRR (b) Receiver operating curve (ROC)

Figure10: Accuracy plot for Forstner operator

Table 1: Comparison of FAR, FRR and Accuracy of Point based methods

Method Name	FAR	FRR	Accuracy
SUSAN	0.082	0.083	91.75
WAVELET	0.0702	0.071	92.94
TRAJKOVIC	0.082	0.084	91.7
FORSTNER	**0.042**	**0.04**	**95.9**

FORSTNER perform better than other point based methods. Low FAR of 4.02% is observed.

4. Conclusion

Personal authentication using palmprint is gaining popularity because of palmprint being a feature-rich and tamper-proof biometric. Various characteristics of palmprint make it better biometric than other biometrics, i.e. Universality, Distinctiveness, Permanence, Collectability, Performance, Acceptability and Circumvention. The palmprint is feature rich biometrics with various types of features as geometry features, line features, point features, texture features and statistical features. Here, palmprint authentication using various point based methods is discussed. All the methods are analyzed with programming in MATLAB. It is very important while discussing authentication, that it is not possible to do ideal (100%) authentication. There are chances of person getting false accepted or rejected. All these factors have to be taken into consideration while talking about authentication. It is well known facts that increase in FAR leads to less security or not proper authentication because any person can be accepted as genuine. Same applies with FRR, false rejection leads to more time taken for authentication by a genuine person. There has to be a balance between the both FAR and FRR for proper

authentication. Here, the main aim is to have low FAR as possible. For each type of features extracted and analyzed, best method is chosen. Out of all the various methods discussed, FORSTNER performed best with FAR of 4.02%.

References

[1] A. H. M. Al-Helali, W. A. Mahmmoud, H. A. Ali, "A Fast Personal Palmprint Authentication Based On 3D-Multi Wavelet Transformation", Transnational Journal of Science and Technology, September 2012 edition vol.2, No.8, Pages 1-10.

[2] Ajay K., Wong D. C. M., Shen H. C. and Anil K. J., "Personal Verification Using Palmprint and Hand Geometry Biometric", Proceeding of 4th International Conference on Audio- and Video-Based Biometric Person Authentication, Guildford, UK, 2003, Pages 668–678.

[3] CANNY J.F.: 'A computational approach to edge detection', IEEE Trans. Pattern Anal. Mach. Intell., Vol. 8, No. 6, 1986, Pages 112–131.

[4] Chen J., Zhang C. S. and Rong G., "Palmprint recognition using crease", Proceedings of International Conference on Image Processing 2001, Vol 3, 7-10 October 2001, Thessaloniki, Greece, Pages 234-237.

[5] Chen W. S., Chiang Y. S. and Chiu Y. H., "Biometric Verification by Fusing Hand Geometry and Palmprint", Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing 2007, Vol 2, 26-28 November 2007, Kaohsiung, Taiwan, Pages 403-406.

[6] Doi, J. and Yamanaka, M., "Personal authentication using feature points on finger and palmar creases", Proceedings of 32nd Applied Imagery Pattern Recognition Workshop 2003, 15-17 Oct. 2003, Washington, DC, United States of America, Pages 282-287.

[7] Dongmei Sun, Zhengding Qiu, Qiang Li, "Palmprint Identification using Gabor Wavelet Probabilistic Neural Networks", ICSP2006 Proceedings, Vol. 4, 2006.

[8] Duta N., Anil K. J. and Kanti V. M., "Matching of palmprints", Pattern Recognition Letters, Vol 23, Issue 4, February 2002, Pages 477-485.

[9] Edward Wong Kie Yih, G. Sainarayanan, Ali Chekima,"Palmprint based biometric system: A comparative study on discrete cosine transform energy, wavelet transform energy and sobelcode methods", Biomedical Soft Computing and Human Sciences, Vol. 14, No. 1, 2009, Pages 11-19.

[10] Guang-Ming Lu, Kuan-Quan Wang, David Zhang, "Wavelet Based Independent Component Analysis for Palmprint Identification", Proceedings of the Third International Conference on Machine Learning and Cybernetics, Vol. 6, Shanghai, 26-29 August 2004, Pages 3547-3550.

[11] H B Kekre and V A Bharadi. "Texture Feature Extraction using Partitioned/Sectorized Complex Planes in Transform Domain for Iris & Palmprint Recognition", IJCA Proceedings on International Conference and workshop on Emerging Trends in Technology, ICWET(3), March 2012, Published by Foundation of Computer Science, New York, USA, Pages 18-24.

[12] Han C. C., Cheng H. L., Lin C. L. and Fan K. C., "Personal authentication using palm-print features", Pattern Recognition, Vol 36, Issue 2, February 2003, Pages 371-381.

[13] Harris and M.J. Stephens, "A combined corner and edge detector," in 4th Alvey Vision Conference, Manchester, UK, 1988, Pages 147–151.

[14] Jifeng Dai and Jie Zhou, "Multifeature- Based High Resolution Palmprint Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 33, No. 5, May 2011, Pages 945-957.

[15] Jyoti Malik, Ratna Dahiya & G Sainarayanan, "Fast Complex Gabor Wavelet Based Palmprint Authentication", International Journal of Image Processing (IJIP), Vol 5, Issue 3, Sep 2011, Pages 283-297.

[16] Kong W. K. and Zhang D., "Palmprint texture analysis based on low-resolution images for personal authentication", Proceedings of 16th International Conference on Pattern Recognition 2002, Vol 3, 11-15 August 2002, Quebec city, QC, Canada, Pages 807-810.

[17] Kovesi P. "Image features from Phase Congruency", Videre J. Comput. Vis. Res., Vol. 1, No. 3, Pages 1–26, 1999.

[18] Lei Zhang and David Zhang, "Characterization of Palmprints by Wavelet Signatures via Directional Context Modeling" IEEE Transactions on Systems, Man, And Cybernetics—Part B: Cybernetics, Vol. 34, No. 3, June 2004, Pages 1335-1347.

[19] Li W. X., Zhang D. and Xu Z. Q., "Image alignment based on invariant features for palmprint identification", Signal Processing: Image Communication, Vol 18, Issue 5, May 2003, Pages 373-379.

[20] Li W., Zhang D., and Xu Z., "Palmprint Identification by Fourier Transform", International Journal of Pattern Recognition and Artificial Intelligence 2002, Vol 16, No. 4, 2002, Pages 417-432.

[21] Li, W. and Zhang, D. and Zhang, L. and Lu, G. and Yan, J., "3-D Palmprint Recognition With Joint Line and Orientation Features", IEEE Trans. Systems, Man, and Cybernetics, Part C, 41(2), 2011, Pages 274 -279.

[22] Lu G. M., Wang K. Q. and Zhang, D., "Wavelet based independent component analysis for palmprint identification", Proceedings of 2004 International Conference on Machine Learning and Cybernetics, Vol 6, 26-29 August 2004, Shanghai, China, Pages 3547-3550.
[23] Rafael D. M., Travieso, C. M., Alonso, J. B. and Ferrer, M. A., "Biometric system based in the feature of hand palm", 38th Annual 2004 International Carnahan Conference on Security Technology, 11-14 October 2004, Albuquerque, New Mexico, United States of America, Pages 136-139.
[24] Struc V and N. Pavesic, "Phase Congruency features for palmprint verification", IET Signal Processing, Vol. 3, Issue 4, 2008, Pages 258-268.
[25] Tantachun, S., Pintavirooj, C., Lertprasart P. and Bunluechokchai S., "Biometrics with Eigen-Hand", 1st IEEE Conference on Industrial Electronics and Applications 2006, 23 – 26 May 2006, Singapore, Pages 1-4.
[26] Tao J. W., Jiang W., Gao Z., Chen S. and Wang C., "Palmprint Recognition Based on 2-Dimension PCA", First International Conference on Innovative Computing, Information and Control 2006, ICICIC '06, Vol 1, 30-01 August 2006, Beijing, China, Pages 326-330.
[27] Wang M. and Ruan Q. Q., "Palmprint Recognition Based on Two-Dimensional Methods", The 8th International Conference on Signal Processing, Vol 4, 16-20 November 2006, Guilin, China.
[28] Wei X. Y., Xu D. and Ngo C. W., "Multi-biometrics based on palmprint and hand geometry", Proceeding of the Fourth Annual ACIS International Conference on Computer and Information Science 2005, Jeju Island, South Korea, Pages 495–500.
[29] Wong K. Y. E., G. Sainarayanan and Ali Chekima, "Palmprint Identification Using SobelCode" Malaysia-Japan International Symposium on Advanced Technology 2007, 12-15 November 2007, Kuala Lumpur, Malaysia.
[30] Wong, M., Zhang, D., Kong, W. K. and Lu, G., "Real-time palmprint acquisition system design", IEE Proceedings - Vision, Image and Signal Processing, Vol 152, Issue 5, 7 October 2005, Pages 527-534.
[31] Wu X. Q., Wang K. Q. and Zhang D., "Wavelet based palm print recognition", Proceedings of International Conference on Machine Learning and Cybernetics 2002, Vol 3, 4-5 November 2002, Beijing, China, Pages 1253-1257.
[32] Wu X. Q., Wang K. Q. and Zhang D., "Wavelet Energy Feature Extraction and Matching for Palmprint Recognition", Journal of Computer Science and Technology, Vol 20, No. 3, May 2005, Pages 411-418.
[33] Wu X. Q., Zhang D., Wang K. Q. and Huang B., "Palmprint Classification Using Principal Line", Pattern Recognition, Vol 37, Issue 10, October 2004, Pages 1987-1998.
[34] Xiangqian Wu, Kuan-Quan Wang, David Zhang "Palmprint Recognition Using Fisher's Linear Discriminant" Proceedings of the Second International Conference on Machine Learning and Cybernatics, Vol. 5, 2-5 November 2003, Pages 3150-3154.
[35] Xiao-Yuan Jing and David Zhang, "A Face and Palmprint Recognition Approach Based on Discriminant DCT Feature Extraction", IEEE Transactions On Systems, Man, And Cybernetics—Part B: Cybernetics, Vol. 34, issue 6, June 2004, Pages 2405-2415.

Block Diagram and Formal Mathematical Definition of Steganographic System

Alexey Smirnov
Associate Professor in the Department of Software, Kirovohrad National Technical University, Kirovohrad, Ukraine

Abstract

This paper studies a formal mathematical description and the block diagram of the secret system and, by analogy with the theory of secret systems, introduces the basic elements and mathematical operators, abstractly describing steganographic information protection system.

Keywords: secret systems, steganographic system, formal mathematical definition

1. Introduction

Mathematical foundations of modern cryptography are laid by the famous American scientist C. Shannon [1-3], who, for the first time, using information-theoretic approach, introduced abstract mathematical definition of a secret system and formalized the procedures for data cryptographic transformation of data. These studies gave a significant boost to the development of the individual methods of the theory of information security, cryptography and authentication, digital steganography, and digital signal processing techniques and error-correcting coding [4-12].

This paper studies a formal mathematical description (in terms of C. Shannon) and the block diagram of the secret system and, by analogy with the theory of secret systems, introduces the basic elements and mathematical operators, abstractly describing steganographic information protection system.

2. Block diagram and a formal mathematical definition of cryptographic (secret) system

Abstract secret system is defined as some set of mappings from one space (the set of possible messages) to a different space (the set of possible cryptograms) [1-3].

Let's fix a set of possible messages $M = \{M_1, M_2, \ldots, M_m\}$ and a set of cryptograms $E = \{E_1, E_2, \ldots, E_n\}$. We will also fix a set of mappings:

$$\varphi = \{\varphi_1, \varphi_2, \ldots, \varphi_k\},$$

where:

$$\varphi_i: M \to E,\ i = 1, 2, \ldots, k.$$

If the sets M and E are equivalent, i.e., n = m, then there is an inverse mapping $\varphi_i^{-1}: E \to M$, which assigns each element of the set E to an element of M. Obviously, φ_i and φ_i^{-1} are given reciprocally the same mapping of the sets M and E.

Let's now fix a set of keys $K = \{K_1, K_2, \ldots, K_k\}$, so that for all $i = 1, 2, \ldots, k$ mapping $\varphi_i \in \varphi$ is uniquely specified by the key K_i, that is:

$$\varphi_i : M \xrightarrow{K_i} E .$$

Each specific mapping of φ_i from the set φ corresponds to the way of encryption with a specific key K_i.

Let's fix a set of keys $K^* = \{K_1^*, K_2^*, \ldots, K_k^*\}$, in general to $K \neq K^*$. All the elements of the inverse mappings:

$$\varphi^{-1} = \{\varphi_1^{-1}, \varphi_2^{-1}, \ldots, \varphi_k^{-1}\}$$

are given the appropriate key:

$$\varphi_i^{-1} : E \xrightarrow{K_i^*} M .$$

Each specific mapping φ_i^{-1} of the set φ^{-1} corresponds to the way of decryption using the key K_i^*. If the key K_i^* is known, then the only one answer is possible as the result of decryption – an element of the set M.

Thus, an abstract definition of a secret system includes the following sets of M, E, φ, φ^{-1}, K and K^* (the sets of open texts and cryptograms, sets of direct and inverse mappings, sets of keys). If, in addition, $K \neq K^*$, then the system is asymmetric. On the contrary, if $K = K^*$ – the system is symmetric. Fig. 1 represents a block diagram of a secret system.

A message source generates the flow of messages from the set M. Each message is a specific implementation of some random process describing the work of a message source. Each message $M_j \in M = \{M_1, M_2, \ldots, M_m\}$ corresponds to the probability $P(M_j)$. A distribution of the random process probability is given by set of probability distribution of random variables, i.e. by a set of probabilities:

$$P_M = \{P(M_1), P(M_2), \ldots, P(M_m)\}. \quad (1)$$

Keys' source generates a flow of keys from the set K and/or K^*. Each key $K_i \in K = \{K_1, K_2, \ldots, K_k\}$ corresponds to some probability $P(K_i)$, and each $K_i^* \in K^* = \{K_1^*, K_2^*, \ldots, K_k^*\}$ corresponds to the probability $P(K_i^*)$. Random process of keys' generation is defined by the sets of probabilities:

$$P_K = \{P(K_1), P(K_2), \ldots, P(K_k)\} \quad (2)$$

and

$$P_{K^*} = \{P(K_1^*), P(K_2^*), \ldots, P(K_k^*)\}. \quad (3)$$

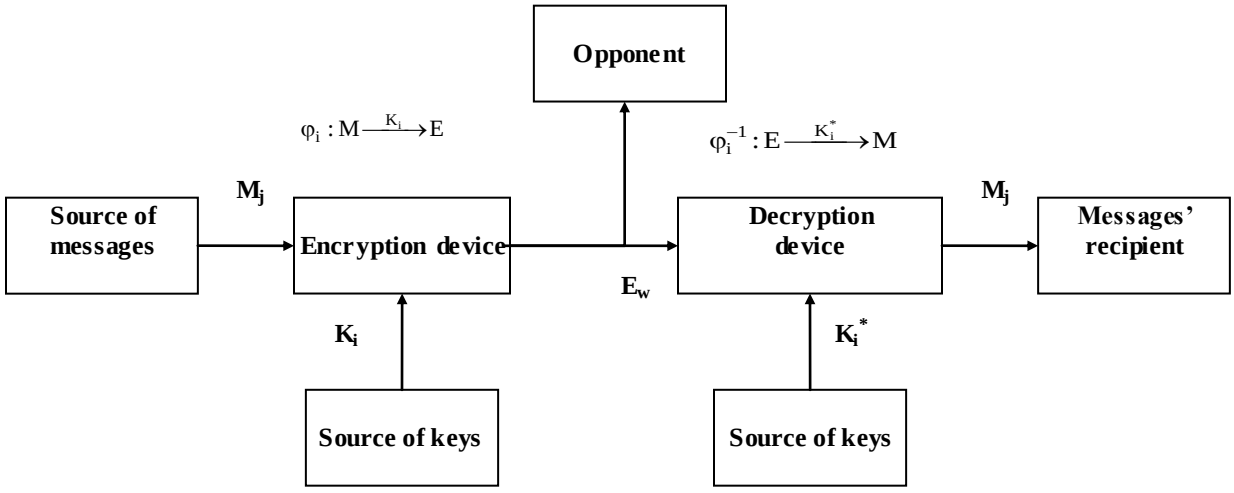


Fig. 1. The block diagram the secret system

Sets of values of a priori probabilities (1 - 3) form a priori knowledge of the opponent about the source of messages and the source of keys, respectively. In fact, these sets characterize the a priori knowledge of the opponent of the possible "weakness" of the secret system.

Selection of K_i determines specific mapping φ_i of the set of mappings φ. With the help of the mapping φ_I which corresponds to the selected key K_i, the cryptogram M_j is formed according to a received message:

$$E_w = \varphi_i(K_i, M_j),$$

$$i \in [1,2,\ldots,k], \; j \in [1,2,\ldots,m], \; w \in [1,2,\ldots,n], \; n \geq m.$$

E_w cryptogram is transmitted to the point of taking on some of the channels and can be intercepted by the opponent. At the receiving end, the original message is restored of cryptogram E_w using reverse mapping φ_i^{-1} (given by the key K_i^*):

$$M_j = \varphi_i^{-1}(K_i, E_w).$$

If the opponent takes over the cryptogram E_l, he can use it to try to calculate the a posteriori probabilities of various possible messages:

$$P_{M|E_w} = \{P(M_1|E_w), P(M_2|E_w), \ldots, P(M_m|E_w)\}, \quad (4)$$

and a variety of possible keys:

$$P_{K|E_w} = \{P(K_1|E_w), P(K_2|E_w), \ldots, P(K_k|E_w)\}, \quad (5)$$

that could be used in the formation of cryptogram E_w.

Sets of a posteriori probabilities (4 – 5) form a posteriori knowledge of the opponent about the keys $K = \{K_1, K_2, \ldots, K_k\}$ and messages $M = \{M_1, M_2, \ldots, M_m\}$ after intercepting a cryptogram E_l. In fact, the sets $P_{K|E_w}$ and $P_{M|E_w}$ are the sets of assumptions, which the corresponding probabilities are assigned to.

3. Block diagram and a formal mathematical definition of steganographic system

By analogy with the theory of secret systems let's consider the basic functional elements and mathematical operators abstractly describing steganographic information protection system.

Let's fix a set of possible messages $M = \{M_1, M_2, \ldots, M_m\}$, the set of possible container $L = \{L_1, L_2, \ldots, L_l\}$, and the set of possible filled containers (steganograms) $E = \{E_1, E_2, \ldots, E_n\}$. Let's also fix a set of mappings:

$$\varphi = \{\varphi_1, \varphi_2, \ldots, \varphi_k\},$$

where:

$$\varphi_i: (M, L) \to E,\ i = 1, 2, \ldots, k.$$

We will define the inverse mapping:

$$\varphi_i^{-1}: E \to (M, L),$$

which each element of the set E assigns to an element of the set M and an element of the set L.

Let's fix a set of keys $K = \{K_1, K_2, \ldots, K_k\}$, so that for all $i = 1, 2, \ldots, k$ mapping $\varphi_i \in \varphi$ is uniquely specified by the key K_i, that is:

$$\varphi_i : (M, L) \xrightarrow{K_i} E.$$

Each specific mapping of φ_i of the set φ corresponds to the way of the message embedding from the set M in the container of the set L with the help of the specific key K_i.

Let's fix the set of keys $K^* = \{K_1^*, K_2^*, \ldots, K_k^*\}$, in general, to $K \neq K^*$. All the elements of the inverse mappings' set:

$$\varphi^{-1} = \{\varphi_1^{-1}, \varphi_2^{-1}, \ldots, \varphi_k^{-1}\}$$

are given the appropriate key:

$$\varphi_i^{-1} : E \xrightarrow{K_i^*} (M, L).$$

Each specific mapping φ_i^{-1} of the set φ^{-1} corresponds to a process of recovering messages from the filled container (and the formation of empty container) with the key K_i^*. If the key K_i^* is known, there is only one possible answer as a result of the extraction operation – an element of the set M and an element of the set L:

$$(M_j, L_l) = \varphi_i^{-1}(E_w, K_i^*).$$

For robust systems the following equality is correct:

$$(M_j, L_l) = \varphi_i^{-1}(E_w + \varepsilon, K_i^*),$$

i.e. slight change of the filled container (for the value ε) will not lead to an incorrect message retrieval.

Fragile steganosystems are characterized with the performance of inequality:

$$(M_j, L_l) \neq \varphi_i^{-1}(E_w + \varepsilon, K_i^*)$$

for an arbitrarily small value ε.

Thus, an abstract definition of steganographic system includes the following sets of M, L, E, φ, φ^{-1}, K and K^* (the sets of open texts, empty containers and steganograms (filled containers), sets of forward and backward mappings, and sets of the corresponding keys).

Fig. 2 represents a block diagram of a steganographic system.

A message source generates a flow of information from a variety of information messages I_j from the set $I = \{I_1, I_2, \ldots, I_m\}$, which, after preliminary converting in a precoder is formed as a message M_j from the set M. A precoder performs a function of preliminary preparation of the informational message to embedding in a container (such as converting an informational message in an array of specially formatted digital data).

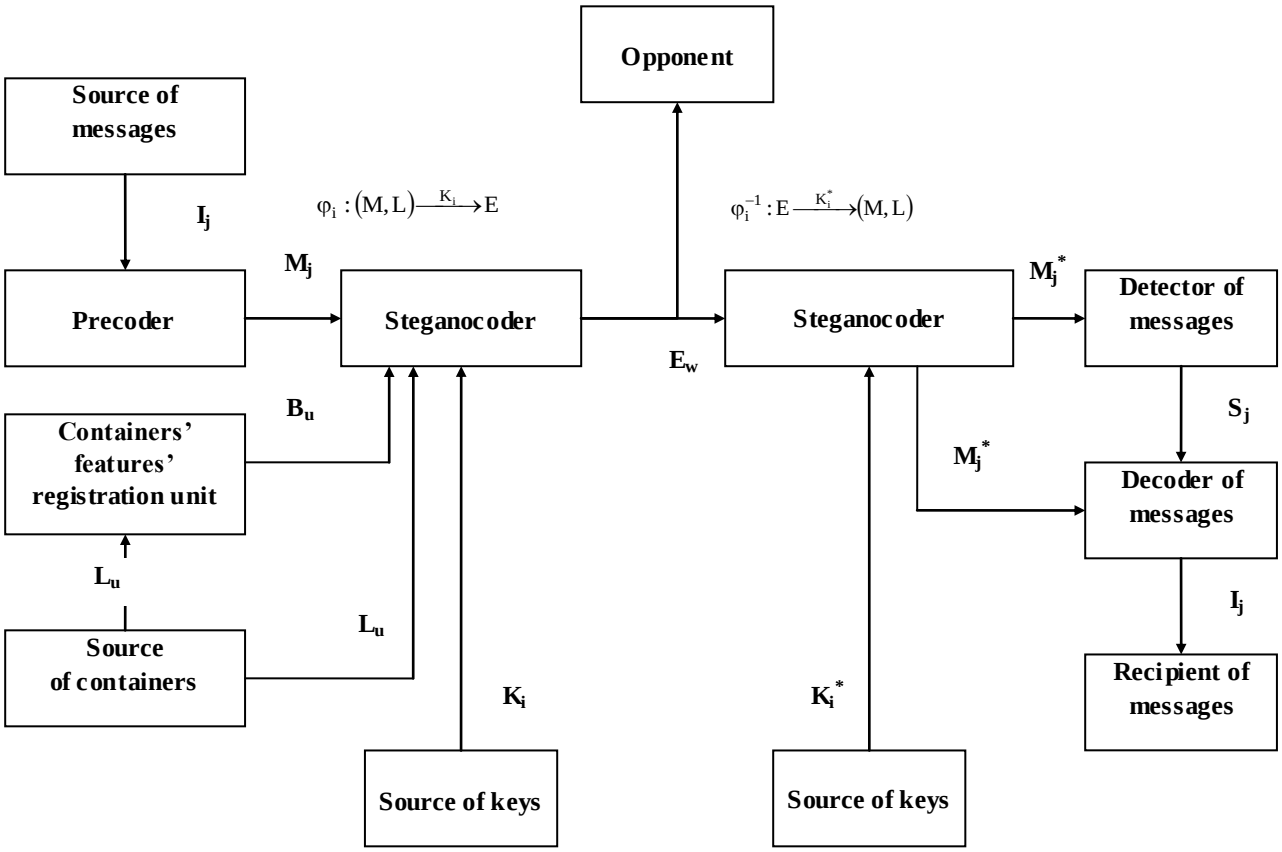


Fig. 2. Block diagram of steganographic system

Each message $M_j \in M = \{M_1, M_2, \ldots, M_m\}$ corresponds to the probability $P(M_j)$. The probability distribution of a random process is given by a cumulative distribution of probability distribution sets of random variables, i.e. the set of probabilities:

$$P_M = \{P(M_1), P(M_2), \ldots, P(M_m)\}. \quad (6)$$

Source of containers generates a flow of empty containers L_u from the set $L = \{L_1, L_2, \ldots, L_l\}$. Work of the source of containers can also be described by some random process, the specific realization of which is the container L_u. In this case we deal with random containers that can be attributed to the corresponding probabilities:

$$P_L = \{P(L_1), P(L_2), \ldots, P(L_l)\}.$$

Much more often, in practice, a different type of containers is used, the formation of which is impossible to describe by a random process. In this case, the source container works on a deterministic rule, asked or authorized (e.g., transmitting) side, or the opponent. In the first case, a so-called selected container, i.e. the container used is not formed by chance, but is chosen by the party responsible for some non-stochastic characteristics. In the second case, the source container is managed by the opponent, and the containers themselves are generated by an attacker and imposed the transmitting side by a deterministic rule. Thus, we have the so-called imposed-on container.

In the simplest case, a lot of empty containers contain only one element, which is used by the transmitting side to embed message and secretly pass it through a communication channel.

L_u shaped container is processed by the containers' features' registration unit. The main function of the containers' features' registration unit is the selection of attributes (features) B_u of incoming container L_u, which will be used for embedding the message to M_j.

A source of keys in steganographic system generates a flow of the set of keys K and/or K^*. Each key $K_i \in K = \{K_1, K_2, \ldots, K_k\}$ corresponds to some probability $P(K_i)$, and each $K_i^* \in K^* = \{K_1^*, K_2^*, \ldots, K_k^*\}$ corresponds to the probability $P(K_i^*)$. Random key generation process is given by the sets of probabilities:

$$P_K = \{P(K_1), P(K_2), ..., P(K_k)\} \quad (7)$$

and:

$$P_{K^*} = \{P(K_1^*), P(K_2^*), ..., P(K_k^*)\}. \quad (8)$$

Sets of values of a priori probabilities (6 - 8) form a priori knowledge of the opponent about the source of messages and the source of keys, respectively. In fact, these sets characterize the a priori knowledge of the opponent on the possible "weakness" of steganographic system.

Key selection K_i determines specific mapping φ_i of the set of mappings φ. With the help of mapping φ_i corresponding to the selected key K_i, following received message M_j and the received container L_u based on identified characteristics B_u of the container L_u, a steganogram (full container) is formed:

$$E_w = \varphi_i(K_i, M_j, L_u),$$

$$i \in [1,2,...,k],\ j \in [1,2,...,m],\ u \in [1,2,...,l],\ w \in [1,2,...,n],\ n \geq m.$$

A steganogram E_w is transferred to the receiving point by a certain channel and may be intercepted by the opponent. At the receiving end with the help of the reverse mapping φ_i^{-1} (given by the key K_i^*) of the steganogram E_w restored the original message and the empty container is restored:

$$(M_j, L_u) = \varphi_i^{-1}(K_i, E_w).$$

When transferring the steganogram E_w through a communication channel and because of the opponent's possible impact on E_w, a transmitted steganogram may become distorted. In this case, the receiving side will get a mixture of a delivered filled container and of a feedback to the container during the transmission through the communication channel: $E_w+\varepsilon$. Performing the operation of a reverse mapping φ_i^{-1} (given by the key K_i^*) will lead, in this case, to a certain evaluation of a transferred message and give an empty container, i.e. we get:

$$(M_j^*, L_u^*) = \varphi_i^{-1}(K_i, E_w + \varepsilon).$$

For fragile steganographic systems, an inequality $M_j^* \neq M_j$ should lead to a message rejection, i.e. at the slightest distortion of the container ($\varepsilon \neq 0$), an extracted assessment M_j^* should not lead to the reading of embedded message (the message M_j is destroyed when $\varepsilon \neq 0$).

Robust steganographic systems are resistant to the impact on a filled container. In the above notations, this means that when $\varepsilon \neq 0$, an extracted assessment M_j^* should be compared to one of the possible messages (ideally, with the message M_j). At the same time, the derived from a communication channel container E_w can contain no embedded message at all, i.e. the extracted from the container assessment M_j^* should not be compared to any of the possible messages. A built-in message detection functions at a receiving side are assigned to messages' detector, which by the received assessment M_j^* decide on the presence or absence of an internal message in the received container E_w. Thus, the estimate of the detector S_j can be interpreted as a binary (yes/no) decision of an error-correcting decoder on the presence or absence of uncorrectable errors. The decoding itself is performed at a messages decoder, the main functions of which are to compare the extracted assessment M_j^* with one of the possible messages M_j and to transform the latter to the informational message I_j provided to the recipients of information.

The opponent may capture the steganogram E_w. In this case, he can use it to try to calculate posteriori probabilities of various possible messages:

$$P_{M|E_w} = \{P(M_1|E_w), P(M_2|E_w), ..., P(M_m|E_w)\} \quad (9)$$

and of a variety of possible keys:

$$P_{K|E_w} = \{P(K_1|E_w), P(K_2|E_w), ..., P(K_k|E_w)\}, \quad (10)$$

that could be used in the formation of the steganogram E_w.

The sets of posterior probabilities (9 - 10) form a posteriori knowledge of the opponent about the keys $K = \{K_1, K_2, ..., K_k\}$ and the messages $M = \{M_1, M_2, ..., M_m\}$ after the interception of the steganogram E_w. In fact, the sets $P_{K|E_w}$ and $P_{M|E_w}$ are sets of assumptions, which are assigned the corresponding probabilities.

4. Conclusions

In this paper we have analyzed and studied the formal mathematical description and a block diagram of a secret system. By analogy with the examined formalization of the theory of secret systems the basic elements and mathematical operators, abstractly describing steganographic information protection system, are introduced. In the introduces formalization a definition of fragile and robust steganosystems has been received, as well as probabilistic indicators characterizing a posteriori knowledge of the opponent on the secret keys and embedded messages. A promising direction for further research is the analysis and theoretical basis of criteria and performance indicators of steganographic security systems, the study of the properties of the known examples of steganosystems by entering the show-makers and the criteria of performance evaluation.

References

[1] C. Shannon, A Mathematical Theory of Information and Cybernetics. – M: FL, 1963. – 829 p.
[2] C. Shannon, Communication in the Presence of Noise. // Information Theory and its Applications. Collection of translations. – Moscow: FIZMATGIZ, 1959. – P. 82-12.
[3] C. Shannon, Communication Theory of Secret Systems // C. Shannon, A Mathematical Theory of Information and Cybernetics. – Moscow: Publishing House of Foreign Literature, 1963. – P.333-402.
[4] Dolgov V. I., Statistical Theory of Receiving Digital Signals. – Kh.: KhHMCSMF, 1989. – 448 p.
[5] Stasev Y. V., Basic Theory of Building Signals. - Kh.: KhMU, 1999. – 87 p.
[6] MacWilliams F. J., Sloane N. J. A., The Theory of Error-Correcting Codes. –M.: Sviaz, 1979. – 744 p.
[7] Naumenko M. I., Stasev Y. V., Kuznetsov O. O., Theoretical Basis and Methods of Algebraic Block Codes. Monography. – Kh.: KhAFU, 2005. – 267 p.
[8] Moldovyan N. A., Moldovyan A. A., Eremeev M.A., Cryptography: From Primitive to the Synthesis of Algorithms. – St.: BHV-Petersburg, 2004. – 448p.
[9] V.M. Sidelnikov, Cryptography and Coding Theory. Materials of the conference "Moscow University and the Development of Cryptography in Russia", Moscow State University. – 2002. – 22 p.
[10] Salomaa A., Public-key Cryptography: Trans. from English, – M.: Mir, 1995. – 318 p.
[11] Konahovich G. F., Puzyrenko A. Y., Computer Steganography. Theory and Practice. – K.: "MK-Press"b 2006. – 288 p., Ill.
[12] Sklar B., Digital Communication. Theoretical Basis and Practical Application. – M. Williams, 2003. – 1104 p.

Author Name

Alexey Smirnov was born in 1977. Graduated from Kharkiv Military University with a degree in "Automated Control Systems" in 1999. Candidate of Technical Sciences. Associate Professor of Department of Software of Kirovohrad National Technical University, Ukraine. Field of interest: information security and routing issues.

A New Geometric Method to Plotting a Family of Functions Logarithm

[1,]B. Nachit, [2,]A. Namir, [3,]M. Bahra, [4,]K. Hattaf, [5,]R. Kasour, [6,]M. Talbi

[1,2] Laboratoire de Technologie de l'Information et Modélisation(LTIM), Faculté des Sciences Ben M'Sik, Université Hassan II-Mohammedia, Casablanca, Maroc
[1,3,5] Cellule d'Observation et de Recherche en Enseignement des Sciences et Techniques (COREST), Centre Régional des Métiers de l'Education et de la Formation Derb Ghalef, Casablanca, Maroc
[4] Département de mathématiques et informatiques, Faculté des Sciences Ben M'Sik, Université Hassan II-Mohammedia, Casablanca, Maroc
[1,5,6] Observatoire de Recherches en Didactique et Pédagogie Universitaire (ORDIPU), Faculté des Sciences Ben M'Sik, Université Hassan II-Mohammedia, Casablanca, Maroc

Abstract:
In this paper, from the study of the family of logarithmic function, we derive a new method to construct the curves: $y = kx + ln(x), k \in IR$. This method will be a new learning situation to present the logarithm function at high school.

Keywords: Algorithm, Family of functions, Function, Logarithm, Register

1. Introduction

The visualization is very important in the teaching of analysis [8]. The notion of function as an object of analysis can intervene with many frames [4] and it is related to other objects (real numbers, numerical sequences ...). This concept also requires the use of multiple registers [5], that are, algebraic Register (representation by formulas); numerical register (table of values), graphical register (curves); symbolic register (table of variations); formal register (notation *f, f (x), fog ...)* and geometrical register (geometrical variables).In addition, Balacheff and Garden [1] have founded two types of image conception among pairs of students at high school, that are, conception curve-algebraic, i.e., functions are seen as particular cases of curves, and a conception algebraic-curve, i.e., functions are first algebraic formulas and will be translated into a curve.The authors Coppe et al. [3] showed that students had more difficulties to translate the table of variations from one function to a graphical representation which shows that students have difficulties to adopt a global point of view about the functions. They have also shown that the algebraic register is predominant in textbooks of the final year at high school. They also noted that the study of functions is based on the algebraic calculation at the final year in high school (limits, derivatives, study of variations...). According to Raftopoulos and Portides [7], the graphical representations make use of point of global and punctual point of view of functions; on the contrary, the properties of the functions are not directly visible from the algebraic formulas. Bloch [2] highlighted that students rarely consider the power of the graphics at the global level and propose teaching sequences supported by a global point of view of the graphical register. The students do not know how to manipulate the functions that are not given by their algebraic representations. And they do not have the opportunity to manipulate the families of functions depending on a parameter.

To study the logarithm three methods are available:
a- From the properties of exponential functions.
b- Put the problem of derivable functions on IR^{+*} such as $f(xy) = f(x)+f(y)$ and admit the existence of primitive for the function $x \longrightarrow 1/x\ (x\neq 0)$.
c- Treat the log after the integration. In this paper, we propose a new method of tracing the family of functions $f_k(x) = kx + ln(x), k \in IR$, without going through the study of functions (boundary limits, table of variations, infinite branches ...) based only on algorithms for tracing tangents.
This method will be used in particular to plot the curve of the logarithm function and the student can from the graphical representation find the properties such as domain of definition, limits, monotony… etc. The idea of this new method is based on our work presented in [6].

2. Description of the method

Let f_k be a family of functions defined the interval $]0, +\infty[$ by

$$f_k(x) = kx + ln(x)$$

With $k \in IR$. The function f_k is strictly increasing on $]0, +\infty[$ when $k \geq 0$. If $k < 0$, then f_k is strictly increasing on $]0, \frac{-1}{k}[$ and is strictly decreasing on $]\frac{-1}{k}, +\infty[$. Moreover, the line $y = kx$ is an asymptotic direction of graph of f_k, and the equation of the tangent at any arbitrary point x_0 is

$$y = (k + \frac{1}{x_o})x - 1 + \ln(x_o)$$

2.1. Tangents of the family of functions f_k at $x_0 = \frac{-1}{k}$

The equation of the tangent at point $x_0 = \frac{-1}{k}$ is

$$y = -1 - \ln(-k)$$

Then, this tangent is line parallel to the axis *(OX)* passing through the point $N(\frac{-1}{k}, -1-\ln(-k))$ that is the intersection of the line $x = \frac{-1}{k}$ and the curve $y = -1 + \ln(x)$.

Hence, we obtain the following algorithm to plot geometrically the tangent at $x_0 = \frac{-1}{k}$

Algorithm:

- Plot the line $y = kx$ and the line $y = -1$;
- Mark the point M which is the intersection of the last two lines;
- Plot the parallel to *(OY)* passing through the point M;
- Mark the point N which is the intersection of this parallel with the curve $y = -1 + \ln(x)$.

Note that if $k < 0$, then the tangent at $x_0 = \frac{-1}{k}$ is the line passing through the point N and the parallel to *(OX)* (see figure 1), and the function f_k admits an extremum at the point N. However, if $k \geq 0$ we have the intersection of the parallel to *(OY)* passing through the point M with the curve $y = -1 + \ln(x)$ is empty, and the function f_k admits no extremum.

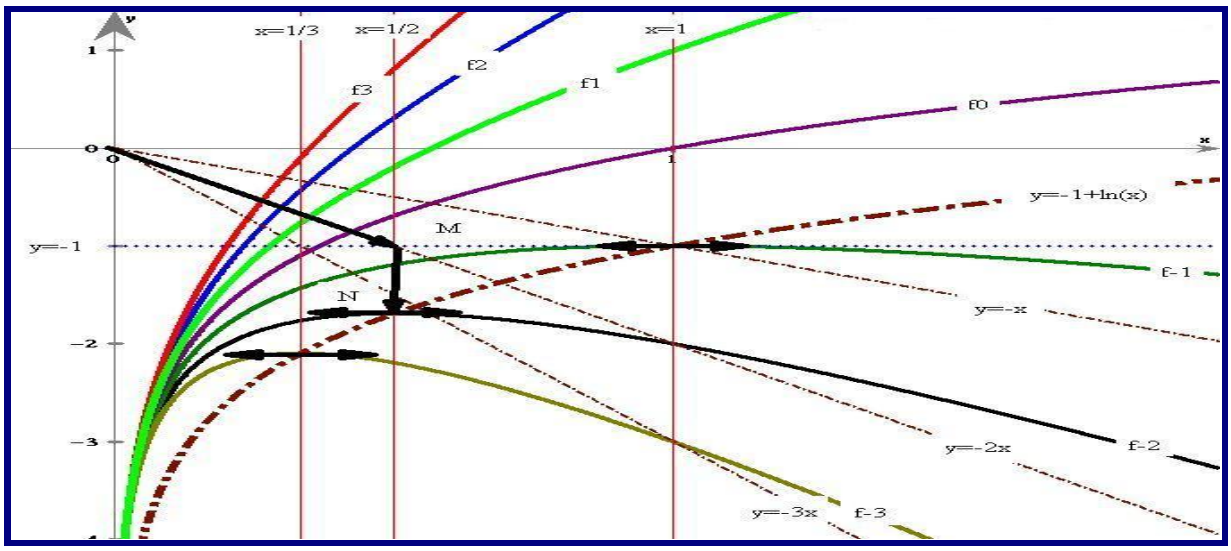


Figure 1. Tangents of the family $f_k(x) = kx + \ln(x)$ at $x_0 = \frac{-1}{k}$

2.2. Tangents of the family of functions f_k at $x_0 = 1$

The equation of the tangent at point $x_0 = 1$ is

$$y = (k + 1)x - 1$$

Then, this tangent is line passing through the point $M(1, k)$ and the point $N(0, -1)$.

Hence, we obtain the following algorithm to plot geometrically the tangent at $x_0 = 1$

Algorithm:

- Plot the line $y = kx$ and the line $x = 1$;
- Mark the point M which is the intersection of the last two lines.

In this case, the tangent at $x_0 = 1$ is the line *(MN)* (see figure 2).

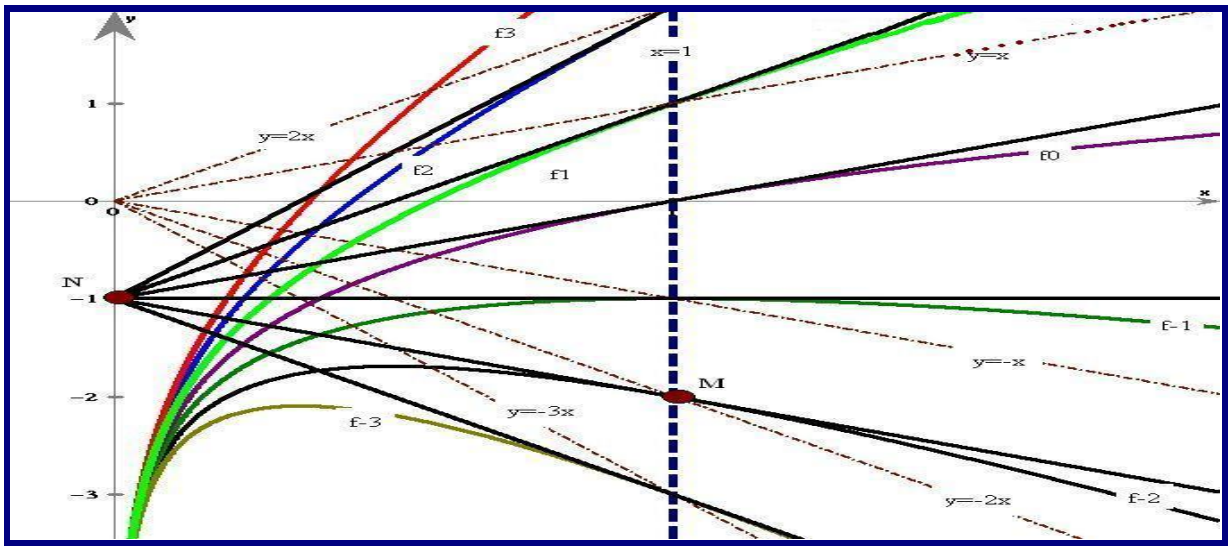


Figure 2. Tangents of the family $f_k(x) = kx + ln(x)$ at $x_0 = 1$

2.3 Tangents of the family of functions f_k at $x_0 = e$

The equation of the tangent at point $x_0 = e$ is

$$y = (k + \frac{1}{e})x$$

Then, this tangent is line passing through the point $N(e, ke + 1)$ and the point $O(0, 0)$.
Hence, we obtain the following algorithm to plot geometrically the tangent at $x_0 = e$

Algorithm:

- Plot the line y = kx and the line x = e;
- Mark the point M which is the intersection of the last two lines;
- Plot the point N which is image of the point M by the translation of vector $t\vec{j}$.

In this case, the tangent at $x_0 = e$ is the line (ON) (see figure 3).

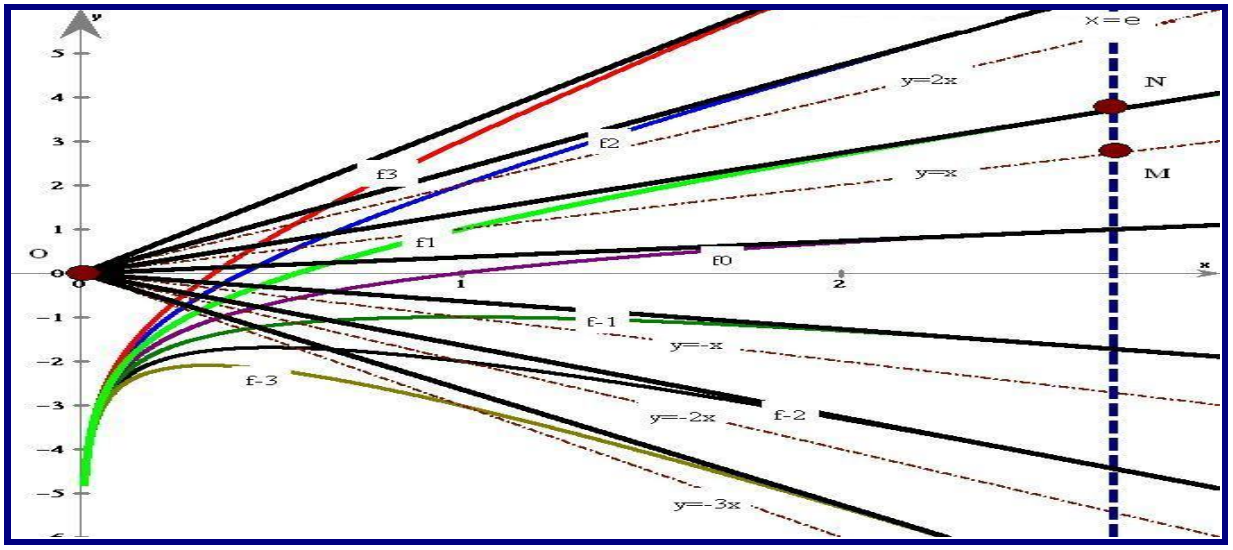


Figure 3. Tangents of the family $f_k(x) = kx + ln(x)$ at $x_0 = e$

2.4 Tangents of the family of functions f_k at $x_0 = \frac{1}{e}$

The equation of the tangent at point $x_0 = \frac{1}{e}$ is

$$y = (k + e)x - 2$$

Then, this tangent is line passing through the point $N(\frac{1}{e}, \frac{k}{e} - 1)$ and the point $P(0, -2)$.

Hence, we obtain the following algorithm to plot geometrically the tangent at $x_0 = \frac{1}{e}$

Algorithm:

- Plot the line $y = kx$ and the line $x = \frac{1}{e}$;
- Mark the point M which is the intersection of the last two lines;
- Plot the point N which is image of the point M by the translation of vector $t - \vec{j}$.

In this case, the tangent at $x_0 = \frac{1}{e}$ is the line (PN) (see figure 4).

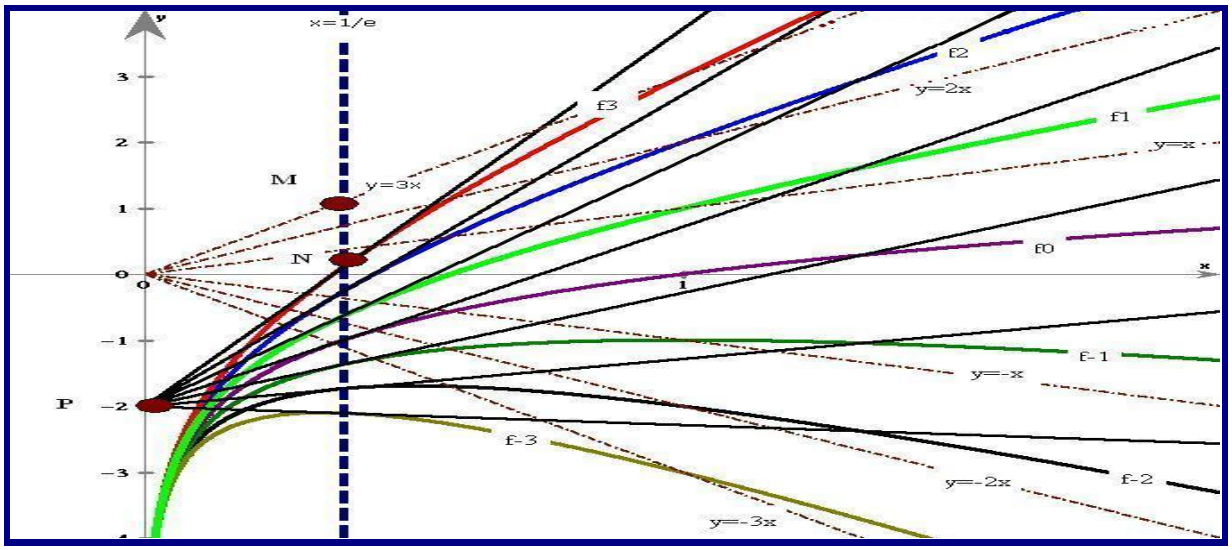


Figure 4. Tangents of the family $f_k(x) = kx + ln(x)$ at $x_0 = \frac{1}{e}$

2.5 Tracing the curves $y = kx + ln(x)$ from bundles of tangents

In this new learning situation, we give to student the algorithms of the tangents of a family of functions f_k, $k \in IR$, at the following points $x_0 = \frac{-1}{k}$, $x_0 = 1$, $x_0 = e$ and $x_0 = \frac{1}{e}$. And we demand to student to plot the curves of f_k, $k \in IR$. In the second part of this situation, the student is informed that this family of function is $f_k(x) = kx + ln(x)$, $k \in IR$, and we demand him to focus on the graphical representation of the function f_0 in order to find the properties of this function (domain of definition, convexity, monotony, boundary limits,...). Thus, this new learning situation can present the logarithm function at high school.

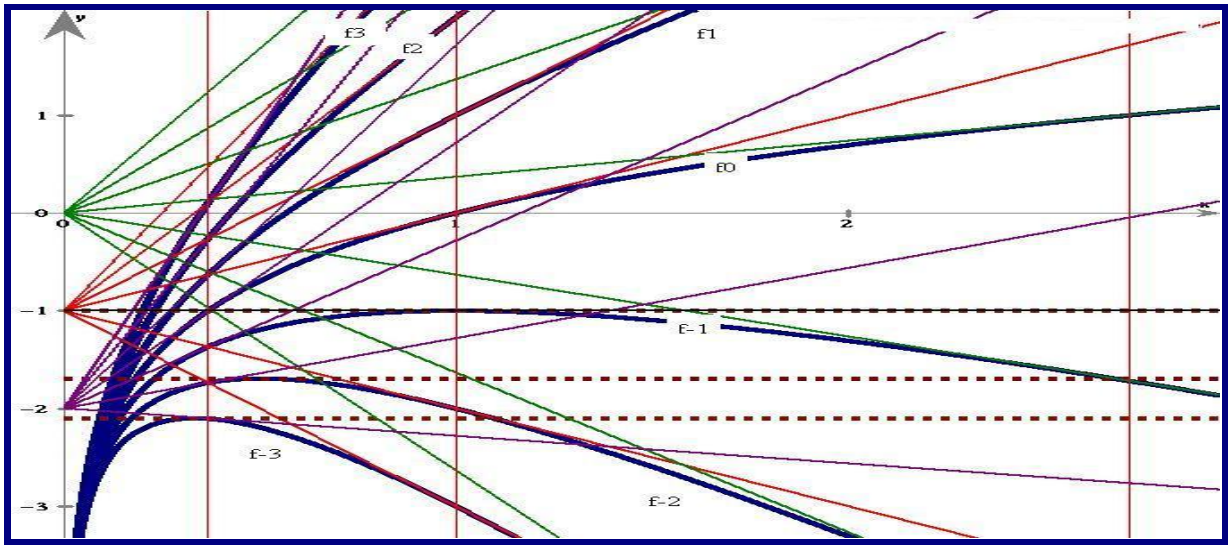


Figure 5. Bundles of tangents of the family $f_k(x) = kx + \ln(x)$, $k \in IR$

3. Conclusion and perspectives

In this paper, we have proposed a new method to present the logarithm function to the students at high school. A quick reading of current and past textbooks shows the absence of this manner of teaching the notion "numerical function" (example: the geometry of the logarithm function). This method is very motivating and after the experiments we did in class, the results were encouraging and the student will be able to find the graphical representation of the family of functions $f_k(x) = kx + ln(x), k \in IR$, and in particular the logarithm function and from this representation, the student deduces the properties of the logarithm function such as domain of definition, convexity, monotony, limits,..., etc.

Other situations learning situations such as the exponential function and the function $x \longrightarrow arctan(x)$ will be the object of the future researches.

References

[1] N. Balacheff and N. Gaudin. Students conceptions: an introduction to a formal characterization. Cahier Leibniz, Numéro 65, Publication de Université Joseph Fourrier, 2002. http://halshs.archives-ouvertes.fr/hal-00190425_v1/

[2] I. Bloch. Teaching functions in a graphic milieu: what forms of knowledge enable students to conjecture and prove. Educational Studies in Mathematics, 52: 3–28, 2003.

[3] S. Coppe, J.L. Dorier and I. Yavuz. De l'usage des tableaux de valeurs et des tableaux de variations dans l'enseignement de la notion de fonction en France en seconde. Recherche en Didactique des Mathématiques. 27(2) :151–186, 2007.

[4] R. Douady. Jeux de cadre et dialectique outil-objet. Recherches en Didactique des Mathématiques. 7 (2) : 5–31, 1986.

[5] R. Duval. Registres de représentation sémiotique et fonctionnement cognitif de la pensée. Annales de didactique et de sciences cognitives. 5:37–65, 1991.

[6] B. Nachit, A. Namir, M. Bahra, R. Kasour and M. Talbi. Une approche 3D du concept de fonction d'une variable réeelle. MathémaTICE. 32 : 1–23, 2012. http://revue.sesamath.net/spip.php?article447.

[7] A. Raftopoulos and D.Portides. Le concept de fonction et sa représentation spatiale, dans symposium FrancoChypriote "Mathematical Work Space", 201–214, Paris, France, 2010

[8] D. Tall. The psychology of advanced mathematical thinking. In D. Tall (Ed.), Advanced Mathematical Thinking (pp. 3–24). Dordrecht: Kluwer Academic Publishers, 1991.

Statistical Distributions involving Meijer's G-Function of matrix Argument in the complex case

[1,] Ms. Samta [2,] Prof. (Dr.) Harish Singh
[1]Research Scholar, NIMS University
[2]Professor Department of Business Administration, Maharaja Surajmal Institute,
Guru Gobind Singh Indraprastha University, Delhi

Abstract:
The aim of this paper is to investigate the probability distributions involving Meijer's g Functions of matrix argument in the complex case. Many known or new Result of probability distributions have been discussed. All the matrices used here either symmetric positive definite or hermit ions positive definite.

I Introduction: The G-Function

The G-functions as an inverse matrix transform in the following form given by Saxena and Mathai (1971).

$$\int_{\tilde{z}>0} |\tilde{z}|^{\delta-m} G_{r,s}^{p,q}\left[\tilde{z}\middle|\begin{matrix}a_1 \dots .. a_r\\ b_1 \dots .. b_r\end{matrix}\right]\overline{dz}$$
$$= \frac{\prod_{j=1}^{p} r_m\,(b_j+\delta)\prod_{j=1}^{q} r_m\,(m-a_j-\delta)}{\prod_{j=p+1}^{s} r_m\,(m-a_j-\delta)\prod_{j=q+1}^{r} r_m\,(b_j+\delta)} \qquad 1.1$$

For p < q or p = q, q ≥ 1, z is a complex matrix and $\tilde{z}>0, Re\left(b_j+\delta\right)>m-1, (j=1,\dots,p) and\ Re\left(a_j+\delta\right)<m-1 (j=1,\dots,q).$ the gamma products are such that the poles of $\prod_{j=1}^{p} r_m\,(b_j+\delta)$ and those of $\prod_{j=1}^{p} r_m\,(m-a_j-\delta)$ are separated.

II. The Distribution

In the multivariate Laplace type integral

$$I = \int_{X>0} etr\,(-\tilde{B}\tilde{X})|\det \tilde{X}|^{a-m}\emptyset(\tilde{X})d\tilde{X}$$

taking $\emptyset(\tilde{X}) = G_{r,s}^{p,q}\left[\tilde{R}\tilde{X}\middle|\begin{matrix}a_1 \dots .. a_r\\ b_1 \dots .. b_r\end{matrix}\right]$

The integral reduces to

$$I = \left|\det(\tilde{B})\right|^{-a} G_{r+1,s}^{p,q+1}\left[\tilde{R}\tilde{B}^{-1}\middle|\begin{matrix}m-a, a_1 \dots .. a_r\\ b_1 \dots . b_s\end{matrix}\right] \qquad 2.1$$

For $Re(-a+\min b_j)>m(j=1,2,3,\dots,p)$ and $\tilde{B}$ is hermitian positive definite matrix and $\tilde{R}$ is an arbitrary complex symmetric m X m matrix.

The result (2.1) is a direct consequence of the result Mathai and Saxena (1971, 1978).

[Notation $etr\left(-\tilde{B}\tilde{R}\right) = \exp\,[tr(X)]$ for exponential to the power $tr(X)$]

Thus the function

$$f(\tilde{x}) = f(\tilde{x}; a, a_1, \dots, b_1, \dots .. b_s;\ \tilde{B}, \tilde{R}) =$$
$$\frac{etr\left(-\tilde{B}\tilde{X}\right)\det x|^{a-m} G_{r,s}^{p,q}\left[\tilde{R}\tilde{X}\middle|\begin{matrix}a_1 \dots .. a_r\\ b_1 \dots .. b_s\end{matrix}\right]}{\left|\tilde{B}\right|^{-a} G_{r+1,s}^{p,q+1}\left[\tilde{R}\tilde{B}^{-1}\middle|\begin{matrix}m-a, a_1 \dots .. a_r\\ b_1 \dots . b_s\end{matrix}\right]}$$

where $Re(-a+\min b_j)>m-1$

$= 0$, else where

provides a probability density function (p.d.f)

2.1 Special Cases

Case (i)

Replacing $\tilde{R} = \tilde{I}$, letting $\tilde{B}$ tends to null matrix and using the result due to Mathai (1977)

$$\int_{\tilde{X}=\tilde{X}'>0} |\det \tilde{X}|^{a-m} G_{r,s}^{m,n}\left[\tilde{X}\middle|\begin{matrix}a_1 \dots . a_r\\ b_1 \dots . b_s\end{matrix}\right] d\tilde{x} = \emptyset_1(a)$$

where

$$\emptyset_1(a) = \frac{\prod_{j=1}^{p} \Gamma_m(b_j+a)\prod_{j=1}^{q}\Gamma_m(m-a_j-a)}{\prod_{j=P+1}^{s}\Gamma_m(m-b_j-a)\prod_{j=q+1}^{r}\Gamma_m(a_j+a)} \quad 2.1.1$$

Where $Re(b_j+a) > m; (j = 1, 2, \dots, m)$

In (2.1.1), we get

$$f(\tilde{X}) = [\emptyset_1(a)]^{-1} = |\det(\tilde{X})|^{a-m} G_{r,s}^{p,q}\left[\tilde{X}\middle|\begin{matrix}a_1 \dots . a_r\\ b_1 \dots . b_s\end{matrix}\right] \quad 2.1.2$$

where $Re(b_j+a) > m; (j = 1, 2, \dots, m)$

$Re(a_j+a) > m; (j = 1, \dots, n), \tilde{X} = \tilde{X}' > 0 = 0$, elsewhere

Case (ii)

Putting p = 1, q = 0, r = 0, s = 1, B = I, then (2.1.1) takes the form

$$I = \int_{X>0} etr(-\tilde{X})|\det \tilde{X}|^{a-m} G_{0,1}^{1,0}[\tilde{R}\tilde{X}|a]d\tilde{X} \quad 2.1.3$$

We know that $G_{0,1}^{1,0}[\tilde{R}\tilde{X}|a] = |\det \tilde{R}|^{a} |\det(\tilde{X})|^{a} e^{-tr\tilde{R}\tilde{X}}$

where $\tilde{X} = \tilde{X}' > 0$ 2.1.4

Using (2.1.4) iin (2.1.3), we have

$$I = |\det R|^{a} \int_{\tilde{X}>0} e^{-tr(\tilde{R}+1)\tilde{X}} \left|\det(\tilde{X})\right|^{a+a-m} d\tilde{X} \quad 2.1.5$$

The integral reduces to

$$= |\det R|^{a} \tilde{\Gamma}_m(a+a)|\det(\tilde{I}+\tilde{R})|^{-(a+a)}$$

(2.1.2) takes the form

$$f(\tilde{X}) = e^{-tr(1+\tilde{R})\tilde{X}} \frac{|\det(\tilde{X})|^{a+a-m}}{\Gamma_m(a+a)|\det(\tilde{I}+\tilde{R})|^{-(a+a)}} \quad 2.1.6$$

$where\ Re(a+a) > m, Re(I+\tilde{R}) > 0, \tilde{X} = \tilde{X}' > 0$

$= 0, elsewhere$

Which is a gamma distribution.

Taking $(a+a) = m, (2.2.6) takes\ the\ form$

$$f(\tilde{X}) = \frac{e^{-tr(1+\tilde{R})\tilde{X}}}{\Gamma_m(m)|\det(\tilde{I}+\tilde{R})|^{-m}} \quad 2.1.7$$

$where\ Re(\tilde{I}+\tilde{R}) > 0, \tilde{X} = \tilde{X}' = 0$

$= 0, elsewhere$

Taking $(a+a) = \frac{n}{2}, (\tilde{I}+\tilde{R}) = \frac{1}{2}\tilde{T}^{-1}$,(2.1.6) yields the wishart distribution with scalar matrix $\tilde{T}$ and n degree of freedom.

$$f(\tilde{X}) = \frac{e^{-tr\left(\frac{\tilde{T}^{-1}\tilde{X}}{2}\right)}|\det \tilde{X}|^{\frac{n}{2}-m}}{\tilde{\Gamma}_m\left(\frac{n}{2}\right)\left|\frac{1}{2}\tilde{T}^{-\frac{1}{2}}\right|^{-\frac{n}{2}}}$$

$$= \frac{2^{-\frac{n}{2}}|\det\tilde{X}|^{\frac{n}{2}-m} e^{-tr\left(-\frac{1}{2}\tilde{T}^{-1}\tilde{X}\right)}}{\Gamma_m\left(\frac{n}{2}\right)|\tilde{T}|^{\frac{n}{2}}} \quad 2.1.8$$

$for\ \tilde{X} = \tilde{X}' > 0, T > 0, m \le n$

$= 0, elsewhere$

Case (iii)

Putting p = 1, q = 0, r = 1, s = 1, then (2.1.1) takes the form

$$\int_{\tilde{X}>0} etr(-\tilde{B}\tilde{X})|\det\tilde{X}|^{a-m} G_{1,1}^{1,0}\left[\tilde{R}\tilde{X}\middle|\begin{matrix}a\\ b\end{matrix}\right] d\tilde{X} \quad 2.1.9$$

We know that

$$G_{0,1}^{1,0}\left[\tilde{R}\tilde{X}\middle|\begin{matrix}a\\ b\end{matrix}\right] = \frac{1}{\Gamma_m(a-b)}\left|\det(\tilde{X})\right|^{a}\left|\det(\tilde{I}-\tilde{X})\right|^{a-b-m} \quad 2.1.10 for\ 0 < \tilde{X} < \tilde{I}, Re(a-b) > m$$

Using (2.2.10) in (2.2.9) we have

$$\tilde{X}\frac{|\det(\tilde{R}|^{b}}{\Gamma_m(a-b)} \int_{0<\tilde{X}<\tilde{I}} etr(-\tilde{B}\tilde{X})\left|\det(\tilde{R})\right|^{b+a-m}|\det(\tilde{I}-\tilde{X})|^{a-b-m} d\tilde{X}$$

2.1.11

The integral reduces to

$$I = \left|\det \tilde{X}\right|^{b} \frac{\tilde{r}_m(b+a)}{\tilde{r}_m(a-b)}\, {}_1\tilde{F}_1\left(b+a; a+a; -\tilde{B}\right) \qquad 2.1.12$$

For $Re(b+a) > m-1, Re(a+a) > m-1$

$= 0, else\ where$

The result (2.1.12) is a direct consequence of the result

$$\int_{e^{-tr(\tilde{X}\tilde{Z})}} \left|det\ \tilde{X}\right|^{\delta - m} |\det(\tilde{I} - \tilde{X})|^{p-\delta-m} d\tilde{X}$$

$$= \frac{\tilde{r}_m(\delta)\tilde{r}_m(p-\delta)}{\tilde{r}_m(P)}\, {}_1F_1[\delta; p; -z] \qquad 2.1.13$$

For $Re(\delta) > m-1, Re(p) > m-1, Re(p-\delta) > m-1$

Thus the function (2.1.1) takes the form

$$f(\tilde{X}) = \frac{\tilde{r}_m(a+a) etr\left(-\tilde{B}\tilde{X}\right) |\det X|^{b+a-m} |\det(\tilde{I} - \tilde{X})|^{(a-b)-m}}{\tilde{r}_m(a-b)\tilde{r}_m(b+a)|\, {}_1F_1|(b+a; a+a; -\tilde{B})}$$

for $R(a-) > m-1, Re(b+a) > m-1, Re\,\tilde{B}>0, \tilde{X} = \tilde{X}' = 0$

$= 0,\ elsewhere$ 2.1.14

Case (iv)

Putting p = 1, q = 1, r = 1, s = 1, a = m – a + b

Then (2.1.1) takes the form as

$$\int_{\tilde{X}>0} etr\,(-\tilde{B}\tilde{X})|\det X|^{a-m} G^{11}_{1,1}\left[\tilde{R}\tilde{X}\middle|_{b}^{m-a+b}\right] d\tilde{X} \qquad 2.1.15$$

We know that

$$G^{11}_{1,1}\left[RX\middle|_{b}^{m-a+b}\right] = \tilde{r}_m(a)\left|\det \tilde{R}\right|^{b}\left|\det(\tilde{X})\right|^{b}\left|\det(I+\tilde{X}\right|^{-a}$$

for $Re(b, a-b) > m-1, \tilde{X} = \tilde{X}' = 0$ 2.1.16

Using (2.1.16) in (2.1.15) we have

$$= \tilde{r}_m(a)\left|\det(\tilde{R})\right|^{b} \times \int_{\tilde{X}>0} etr(-\tilde{B}\tilde{X})\left|\det(\tilde{X})\right|^{a+b-m} |\det\,(I+\tilde{R}\tilde{X}|^{-a} d\tilde{X}.$$

The integral reduces to

$$I = \tilde{r}_m(a)\left|\det(\tilde{R})\right|^{b} \tilde{r}_m(a+b)|B|^{-(a+b)}\, {}_2F_0[a, a+b; -; -\tilde{R}\tilde{B}^{-1}]$$

For $Re\left(\tilde{B}\right) > 0, Re(a+b) > m-1$ 2.1.17

$= 0; elsewhere.$

The result (2.1.17) is a direct consequence of the result

$$\int_{\tilde{X}>0} |\det X|^{a-m}\, e^{-tr(\tilde{B}\tilde{X})}\, {}_1F_0[a; -; -\tilde{R}\tilde{X}]\, d\tilde{X} = \tilde{r}_m(a)|\beta|^{-a}\, {}_2F_0[a; a; -; -\tilde{R}\tilde{\beta}^{-1}]$$

For $Re\left(\tilde{B}\right) > 0, Re(a+b) > m$

Where $\left|\det(\tilde{I} + \tilde{R}\tilde{X})\right|^{-a} = {}_1F_0[a; -; -\tilde{R}\tilde{X}$

Thus the p.d.f. (2.1.1) takes the form

$$f(\tilde{X}) = \frac{etr(-\tilde{B}\tilde{X}|\det(\tilde{X})|^{a+b-m}|\det\,(\tilde{I}+\tilde{R}\tilde{X}|^{-a}}{\tilde{r}_m(a+b)|\beta|^{-(a+b)}\, {}_2F_0[a, a+b; -; -\tilde{R}\tilde{\beta}^{-1}]} \qquad 2.1.18$$

For $Re\left(\tilde{B}\right) > 0, Re(a+b) > m-1, Re\left(\tilde{B}\right) > Re\left(\tilde{R}\right), \tilde{X} = \tilde{X}' > 0$

$= 0; elsewhere.$

Replacing $-\tilde{R}$ with $\tilde{R}$ and $\tilde{B}$ with $\tilde{R}$, then (2.2.18) takes the form as

$$f(\tilde{X}) = \frac{etr(-\tilde{R}\tilde{X}|\det(\tilde{X})|^{a+b-m}|\det\,(\tilde{I}-\tilde{R}\tilde{X}|^{-a}}{\tilde{r}_m(a+b)\,|\det\tilde{R}|^{-(a+b)}\, {}_2F_0[a, a+b; -; I]} \qquad 2.1.19$$

For $Re\left(\tilde{R}\right) > 0, Re(a+b) > m-1, \tilde{X} = \tilde{X}' > 0$

$= 0; elsewhere.$

Case (v)

Putting p = 1, q = 0, r = 0, s = 2 2.1.20

Then (2.1.1) takes the form as

$$I = \int_{\tilde{X}>0} etr\,(-\tilde{B}\tilde{X}\left|\det(\tilde{X})\right|^{a-m} G^{12}_{0,2}[\tilde{R}\tilde{X}|a, b] d\tilde{X} \qquad 2.1.21$$

We know that

$$G^{10}_{0,2}[\tilde{R}\tilde{X}|a, b] = \frac{|det\ \tilde{X}|^{a}|det\ \tilde{X}|^{a}\, {}_0F_1[-, m+a-b; -; \tilde{R}\tilde{X}]}{\tilde{r}_m(m+a-b)} \qquad 2.1.22$$

For $Re(a-b) > -1, \tilde{X} = \tilde{X}' > 0$
Making use of (2.1.22) in (2.1.21) we get

$$I = \frac{\left|det\ \tilde{R}\right|^{a}}{\tilde{r}_m(m+a-b)} \int_{\tilde{X}>0} etr(-\tilde{B}\tilde{X}) \left|\det(\tilde{X})\right|^{a-a+m}$$

$$\times {}_1F_1[-;m+a-b;-\tilde{R}\tilde{X}]d\tilde{X}]$$

$$= \frac{\left|\det \tilde{R}\right|^{a}\tilde{r}_m(a+a)}{\tilde{r}_m(m+a-b)} \times \left[\left|\tilde{\beta}\right|^{-(a+a)} {}_1F_1(a+a;m+a-b;-\tilde{R}\tilde{B}^{-1})\right] \quad 2.1.23$$

$for\ Re\ \tilde{B} > 0, Re(a+a) > m-1$
Thus the p.d.f. (2.2.1) takes the form

$$f(\tilde{X}) = \frac{etr(-\tilde{B}\tilde{X})\left|\det \tilde{X}\right|^{a+a-m}{}_0F_1[-;m+a-b;-\tilde{R}\tilde{X}]}{\tilde{r}_m(a+a)\left|\det \tilde{R}\right|^{-(a+a)}{}_1F_1[a+a;m+a-b;I]} \quad 2.1.24$$

For $Re\ \tilde{B} > 0, Re(a) > m-1, Re\ \tilde{R} > Re\ (\tilde{B}), \tilde{X} = \tilde{X}' > 0$
$= 0; elsewhere.$

Case (vi)

Putting p = 1, q = 1, r = 1, s = 2, then (2.1.1) takes the form

$$\int_{\tilde{X}>0} etr(-\tilde{B}\tilde{X}) \left|\det \tilde{X}\right|^{a-m} G^{12}_{1,2}\left[\tilde{R}\tilde{X}\middle|^{a}_{b,c}\right] d\tilde{X} \quad 2.1.25$$

We know that

$$G^{12}_{1,2}\left[\tilde{R}\tilde{X}\middle|^{a}_{b,c}\right] = \tilde{r}_m(m-a+b)\left|\det \tilde{X}\right|^{a}\left|\det \tilde{X}\right|^{b} \quad 2.1.26$$

$$\times {}_1F_1[m-a+b;m+b-c;-\tilde{R}\tilde{X} \quad 2.1.27$$

For $Re(b-c, b-a) > -1$
Using (2.1.26) in (2.1.27) we get

$$= \tilde{r}_m(m-a+b)\left|\det \tilde{R}\right|^{\beta} \int_{\tilde{X}>0} etr(-\tilde{B}\tilde{X})\left|\det \tilde{X}\right|^{a+\beta-m} \times {}_1F_1(m-a+b;m+b-c;-\tilde{R}\tilde{X}$$

2.1.28

$$= \tilde{r}_m(m-a+b)\left|\det \tilde{R}\right|^{\beta}\tilde{r}_m(a+\beta)\left|\det \tilde{R}\right|^{(a+\beta)}$$

$$\times {}_2F_1[m-a+b;a+\beta;m+b-c;-\tilde{R}\tilde{B}^{-1}] \quad 2.1.29$$

For $Re(\tilde{B}) > 0, Re(a+\beta) > m-1$
The result (2.2.29) is a direct consequence of the result

$$\int_{\tilde{X}>0} \left|\det \tilde{R}\right|^{a-m} e^{-tr(\tilde{B}\tilde{X})} {}_1F_1[a;b;-\tilde{R}\tilde{X}]d\tilde{X}$$

$$= \tilde{r}_m(a)\left|\det(\tilde{B})\right|^{-\beta} {}_2F_1[a;a;b;-\tilde{R}\tilde{B}^{-1}] \quad 2.1.30$$

For $Re(\tilde{B}) > 0, Re(a) > m-1, \tilde{X} = \tilde{X}' > 0\ Re(\tilde{B}) > Re[\tilde{R}]$
Then the p.d.f. (2.1.30) takes the form as

$$f(\tilde{X}) = \frac{etr(-\tilde{B}\tilde{X})\left|\det \tilde{X}\right|^{a+\beta-m}{}_0F_1[\bullet]}{\tilde{r}_m(a+\beta)\left|\det \tilde{X}\right|^{a+\beta-m}{}_2F_1[\bullet\bullet]} \quad 2.1.31$$

Where ${}_1F_1[*] = {}_1F_1[m-a+b, m+b-c;-\tilde{R}\tilde{X}]$
${}_2F_1[**] = {}_2F_1[m-a+b, m+\beta; m+b-c;-\tilde{R}\tilde{B}^{-1}]$
For $Re(\tilde{B}) > 0, Re(a+\beta) > m-1, Re(\tilde{B}) > Re(\tilde{R}), \tilde{X} = \tilde{X}' > 0$
$= 0; elsewhere,$
Replacing $\tilde{B}$ with - $\tilde{R}$, (2.1.31) takes the form

$$f(\tilde{X}) = \frac{etr(-\tilde{R}\tilde{X})\left|\det \tilde{X}\right|^{a+\beta-m}{}_1F_1[+]}{\tilde{r}_m(a+\beta)\left|\det \tilde{R}\right|^{(a+\beta)}{}_2F_1[++]} \quad 2.1.32$$

Where ${}_1F_1[+] = {}_1F_1[m-a+b, m+b-c;-\tilde{\beta}\tilde{X}]$
${}_2F_1[++] = {}_2F_1[m-a+b, a+\beta; m+b-c; I]$
For $Re(\tilde{R}) > 0, Re(a+\beta) > m-1, Re(\tilde{R}) > Re(\tilde{B}), \tilde{X} = \tilde{X}' = 0$
We know that

$${}_2F_1[a,b,c;I] = \frac{\tilde{r}_m(c)\tilde{r}_m(c-a-b)}{\tilde{r}_m(c-a)\tilde{r}_m(c-b)} \quad 2.1.33$$

Making use of (2.2.33) in (2.2.32), we get

$$f(\tilde{R}) = \prod_m etr\ (\tilde{R}\tilde{X})\left|\det \tilde{X}\right|^{a+\beta-m} {}_1F_1[+] \quad 2.1.34$$

Where

$$\prod_m = \frac{\tilde{r}_m(a-c)\tilde{r}_m(m+b-c-a-\beta)}{\tilde{r}_m(m+b-c)\tilde{r}_m(a-c-a-\beta)\tilde{r}_m(a+\beta)\left|\det \tilde{R}\right|^{-(a+\beta)}}$$

$${}_1F_1[+] = {}_1F_1[m-a+b; m+b-c; \tilde{B}\tilde{X}]$$

We know that the Kummer transformation as

$${}_1F_1[a;b;\tilde{B}\tilde{X}] = etr(\tilde{B}\tilde{X}){}_1F_1[b-a;b;-\tilde{B}\tilde{X}] \qquad 2.1.35$$

Making use of (2.2.35) in (2.2.34) we get

$$f(\tilde{X}) = \prod_m etr\,[(\tilde{R}+\tilde{B})\tilde{X}]\left|\det \tilde{X}\right|^{a+b-m}{}_1F_1[\#] \qquad 2.1.36$$

When $\prod_m same\ as\ above$

$${}_1F_1[\#] = {}_1F_1[a-c; m+b-c; \tilde{B}\tilde{X}]$$

Where $Re(a+\beta) > m-1, Re(m+b-c\_ > m-1,$

$Re(a-c-\beta) > m-1, \tilde{X} = \tilde{X}' > 0\ Re(\tilde{R}+\tilde{B}) > 0$

Case (vii)

Putting p = 1, q = 2, r = 2, s = 2, a = -c1, b = -c2, c = a-m, a = -b

The (2.1.1) takes the form

$$\int_{\tilde{X}>0} etr\ (-\tilde{B}\tilde{X})\left|\det \tilde{X}\right|^{a-m} G^{12}_{2,2}\left[\tilde{R}\tilde{X}\middle|_{b-m,}^{a-c_2,} - b\right] \qquad 2.1.37$$

Who know that? $G^{12}_{2,2}\left[\tilde{R}\tilde{X}\middle|_{a-m,}^{1-c_2,} - b\right]$

$$= \frac{\tilde{r}_m(a+c_1)\tilde{r}_m(a+c_2)}{\tilde{r}_m(a+b)}\left|\det \tilde{X}\right|^{a-m} \times {}_2F_1[a+c_1, a+c_2; a+b; -\tilde{R}\tilde{X}] \qquad 2.1.38$$

For $Re(a+c_1, a+c_2, a+b) > m = 1, \tilde{X} = \tilde{X}' > 0$

Making use of (2.2.38) in (2.2.37), we get

$$= \frac{\tilde{r}_m(a+c_1)\tilde{r}_m(a+c_2)\tilde{r}_m(a+a-m)}{\tilde{r}_m(a+b)|\beta|^{a+a-m}} \times {}_2F_1[a+c_1, a+c_2; a+b; -\tilde{R}\tilde{X}]d\tilde{X} \qquad 2.1.39$$

The result (2.1.39) is a direct consequence of the result

$$\int_{\tilde{X}>0} etr\,(-\tilde{B}\tilde{X})\left|\tilde{X}\right|^{a-m}{}_2F_1[a_1, a_2, a; b; -\tilde{R}\tilde{B}^{-1}] = {}_{\tilde{r}\ m}(a)\left|\tilde{B}\right|^{-a}$$

$${}_3F_1[a_1, a_2, a; b; -\tilde{R}\tilde{B}^{-1}] \qquad 2.1.40$$

Thus the p.d.f. (2.1.2) takes the following form

$$f(\tilde{X}) = \frac{etr(-\tilde{B}\tilde{X})|\det \tilde{X}|^{a+a-m-m}{}_2F_1[-]}{\tilde{r}_m(a+a-m)\,|\det \tilde{B}|^{(a+a-m)}{}_3F_1[-]} \qquad 2.1.41$$

Where ${}_2F_1[-] = {}_2F_1[a+c_1, a+c_2; a+b; -\tilde{R}\tilde{X}]$

${}_2F_1[-] = {}_3F_1[a+c_1, a+c_2; a+a-m; a+b; -\tilde{R}\tilde{B}^{-1}]$

For $Re(\tilde{B}) > 0, Re(a-a-m) > m-1, Re(\tilde{B}) > Re(\tilde{R}), \tilde{X} = \tilde{X}' > 0$

$= 0; elsewhere.$

Replacing $-\tilde{R}$ with $\tilde{R}$ and $\tilde{X}$ with$\tilde{R}^{-1}$, (2.2.40) reduces to

$$f(\tilde{X}) = \frac{etr\,(-\tilde{B}\tilde{R})|\det \tilde{R}|^{-(a+a+m+1)}{}_2F_1[+]}{\tilde{r}_m(a+a-m)\,|\det \tilde{B}|^{(a+a+m)}{}_3F_1[++]} \qquad 2.1.42$$

Where ${}_2F_1[+] = {}_2F_1[a+c_1, a+c_2; a+c; I]$

${}_3F_1[++] = {}_3F_1[a+c_1, a+c_2; a+a-m; a+b; -\tilde{R}\tilde{B}^{-1}]$

Making use of (2.1.33) in (2.1.42), we get

$$f(\tilde{X}) = \frac{\prod_m etr\,[(-\tilde{B}\tilde{R}^{-1})\left|\det \tilde{R}\right|^{-a-a-(m+1)}{}_2F_1[+]}{\left|\det \tilde{B}\right|^{-(a+a+m)} \blacksquare {}_3F_1[+]}$$

Here,

$$\prod_m = \frac{\tilde{r}_m(a+b)\tilde{r}_m(b-a-c_1-c_2)}{\tilde{r}_m(a-b-c_1)\tilde{r}_m(b-c_2)\tilde{r}_m(a+a-m)}$$

References

[1] Mathai AM.1997. Jacobeans of matrix Transformations and function of Matrix Argument world Scientific Publishing Co. Pvt ltd.

[2] Mathai Am and Saxena R.K. 1971. Meijer's g Function with matrix argument, Acta, Mexicans ci-tech, 5, 85-92.

[3] R Muirrhead RJ Aspects of multi Variate Statical theory, wicley, New york, 1983.

Enhancing Performance of User Authentication Protocol with Resist to Password Reuse Attacks

[1,] Ms. R.R.Karthiga, [2,] Mr.K.Aravindhan,
[1,] Final year, M.E/CSE, SNS College of Engineering
[2,], Asst Professor/CSE, SNS College of Engineering

Abstract

Text password is the most popular form of user authentication on websites due to its convenience and simplicity. However, user's passwords are prone to be stolen and compromised by different threats and vulnerabilities. Users often select weak passwords and reuse the same passwords across different websites. Typing passwords into untrusted computers suffers password thief threat. The user authentication protocol proposes the oPass enhancement to protect user identity; it requires a long-term password for cell phone protection and account ID for login on all websites. OPass only requires each participating website possesses a unique phone number, and involves a telecommunication service provider in registration and recovery phases for the creation of one-time password. User can recover oPass system with reissued SIM cards and long-term passwords. Opass is efficient and affordable compared with the conventional web authentication mechanisms. Therefore one-time password mechanism that has enhanced security using private key infrastructure to prevent integrity problem due to phishing attack and keyloggers.

Index Terms—Network security, password reuse attack, pass- word stealing attack, user authentication.

I. Introduction

OVER the past few decades, text password has been adopted as the primary mean of user authentication for websites. People select their username and text passwords when registering accounts on a website. In order to log into the website successfully, users must recall the selected passwords. Generally, password-based user authentication can resist brute force and dictionary attacks if users select strong passwords to provide sufficient entropy. However, password-based user authentication has a major problem that humans are not experts in memorizing text strings. Thus, most users would choose easy-to-remember passwords (i.e., weak passwords) even if they know the passwords might be unsafe. Another crucial problem is that users tend to reuse passwords across various websites [2]. Password reuse causes users to lose sensitive information stored in different websites if a hacker compromises one of their passwords. This attack is referred to as the password reuse attack. Therefore, it is important to take human factors into consideration when designing a user authentication protocol. Up to now, researchers have investigated a variety of technology to reduce the negative influence of human factors in the user authentication procedure. Since humans are more adept in remembering graphical passwords than text passwords, many graphical password schemes were designed to address human's password recall problem [3]. Using password management tools is an alternative [6] [8]. These tools automatically generate strong passwords for each website, which addresses password reuse and password recall problems. The advantage is that users only have to remember a master password to access the management tool. Although graphical password is a great idea, it is not yet mature enough to be widely implemented in practice and is still vulnerable to several attacks. Furthermore, they have trouble using these tools due to the lack of security knowledge. Besides the password reuse attack, it is also important to consider the effects of password stealing attacks. Adversaries steal or compromise passwords and impersonate users' identities to launch malicious attacks, collect sensitive information, perform unauthorized payment actions, or leak financial secrets. Phishing is the most common and efficient password stealing attack. Many previous studies have proposed schemes to defend against password stealing attacks. Some researches focus on three-factor authentication rather than password-based authentication to provide more reliable user authentication. In this paper, we propose a user authentication protocol named oPass which leverages a user's cellphone and short message service (SMS) to prevent password stealing and password reuse attacks. The main concept of oPass is free users from having to remember or type any passwords into conventional computers for authentication. Unlike generic user authentication, oPass involves a new component, the cellphone, which is used to generate one-time passwords and a new communication channel, SMS, which is used to transmit authentication messages. oPass presents the following advantages.

1) Anti-malware—Malware (e.g., key logger) that gathers sensitive information from users, especially their passwords are surprisingly common. In oPass, users are able to log into web services without entering passwords on their computers. Thus, malware cannot obtain a user's password from untrusted computers.
2) Phishing Protection—Adversaries often launch phishing attacks to steal users' passwords by cheating users when they connect to forged websites. As mentioned above, oPass allows users to successfully log into websites without revealing passwords to computers. Users who adopt oPass are guaranteed to withstand phishing attacks.
3) Secure Registration and Recovery—In oPass, SMS is an out-of-band communication interface. oPass cooperates with the telecommunication service provider (TSP) in order to obtain the correct phone numbers of websites and users respectively. SMS aids oPass in establishing a secure channel for message exchange in the registration and recovery phases. Recovery phase is designed to deal with cases where a user loses his cellphone. With the aid of new SIM cards, oPass still works on new cell phones.
4) Password Reuse Prevention and Weak Password Avoidance— oPass achieves one-time password approach. The cellphone automatically derives different passwords for each login. That is to say, the password is different during each login. Under this approach, users do not need to remember any password for login. They only keep a long term password for accessing their cellphones, and leave the rest of the work to oPass.
5) Cellphone Protection—An adversary can steal users' cellphones and try to pass through user authentication. However, the cellphones are protected by a long-term password. The adversary cannot impersonate a legal user to login without being detected.

II. Background

oPass adopts the one-time password strategy[9] therefore, the strategy describe the secure features of SMS channel and explain why SMS can be trusted. Finally, introduce the security of 3G connection used in the registration and recovery phases of oPass.

A. One-Time Password

The one-time passwords in oPass are generated by a secure one-way hash function. With a given input, the set of one- time passwords is established by a hash chain through multiple hashing. Assuming we wish to prepare one-time N passwords [1], the first of these passwords is produced by performing hashes on input C.

$$c = \mathcal{H}(P_u \| \mathrm{ID}_s \| \phi).$$

B. SMS Channel

SMS is a text-based communication service of telecommunication systems. oPass leverages SMS to construct a secure user authentication protocol against password stealing attacks. As we know, SMS is a fundamental service of telecom, which belongs to 3GPP standards. SMS represents the most successful data transmission of telecom systems; hence, it is the most widespread mobile service in the world. Besides the above advantages, we chose SMS channel because of its security benefits. Compared with TCP/IP network, the SMS network is a closed platform; hence, it increases the difficulty of internal attacks, e.g., tampering and manipulating attacks. Therefore, SMS is an out-of-band channel that protects the exchange of messages between users and servers.

C. 3G Connection

3G connection provides data confidentiality of user data and signal data to prevent eavesdropping attacks. It also provides data integrity of signal data to avoid tampering attacks. oPass utilizes the security features of 3G connection to develop the convenient account registration and recovery procedures. Users can securely transmit and receive information to the web site through a 3G connection.

III. Problem Definition and Assumptions

In this section, we consider various methods of password stealing. Afterwards, we introduce the architecture of our oPass system and make some reasonable assumptions.

A. Problem Definition

People nowadays rely heavily on the Internet since conventional activities or collaborations can be achieved with network services (e.g., web service). Widely deployed web services facilitate and enrich several applications, e.g., online banking, e-commerce, social networks, and cloud computing. But user authentication is only handled by text passwords for most websites. Applying text passwords has several critical disadvantages. First, users create their passwords by themselves. For easy

Memorization, users tend to choose relatively weak passwords for all websites [2]. Second, humans have difficulty remembering complex or meaningless passwords. Some websites generate user passwords as random strings to maintain high entropy, even though users still change their passwords to simple strings to help them recall it [8]. In addition, phishing attacks and malware are threats against password protection. Protecting a user's password on a kiosk is infeasible when key loggers or backdoors are already installed on it. Therefore, we proposed a user authentication, called oPass, to thwart the above attacks. The goal of oPass is to prevent users from typing their memorized passwords into kiosks.

By adopting one-time passwords, password information is no longer important. A one-time password is expired when the user completes the current session. Different from using Internet channels, oPass leverages SMS and user's cellphones to avoid password stealing attacks. We believe SMS is a suitable and secure medium to transmit important information between cellphones and websites. Based on SMS, a user identity is authenticated by websites without inputting any passwords to untrusted kiosks. User password is only used to restrict access on the user's cellphone. In oPass, each user simply memorizes a long-term password for access her cellphone. The long-term password is used to protect the information on the cellphone from a thief.

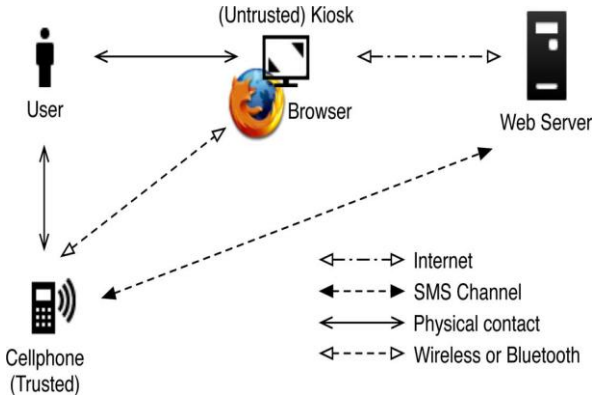


Fig. 1. Architecture of oPass system

Fig 1 describes the architecture (and environment) of the oPass system. For users to perform secure login on an untrusted computer (kiosk), oPass consists of a trusted cellphone, a browser on the kiosk, and a web server that users wish to access. The communication between the cellphone and the web server is through the SMS channel. The web browser interacts with the web server via the Internet. In our protocol design, we require the cellphone interact directly with the kiosk. The general approach is to select available interfaces on the cellphone, Wi-Fi or Bluetooth. The assumptions in oPass system are as follows.

1) Each web server possesses a unique phone number. Via the phone number, users can interact with each website through an SMS channel.
2) The users' cellphones are malware-free. Hence, users can safely input the long-term passwords into cellphones.
3) The telecommunication service provider (TSP) will participate in the registration and recovery phases. The TSP is a bridge between subscribers and web servers. It provides a service for subscribers to perform the registration and recovery progress with each web service. For example, a subscriber inputs her id ID and a web server's id ID to start to execute the registration phase. Then, the TSP forwards the request and the subscriber's phone number to the corresponding web server based on the received ID.
4) Subscribers (i.e., users) connect to the TSP via 3G connections to protect the transmission. 5) The TSP and the web server establish a secure sockets layer (SSL) tunnel. Via SSL protocol, the TSP can verify the server by its certificate to prevent phishing attacks. With the aid of TSP, the server can receive the correct sent from the subscriber.

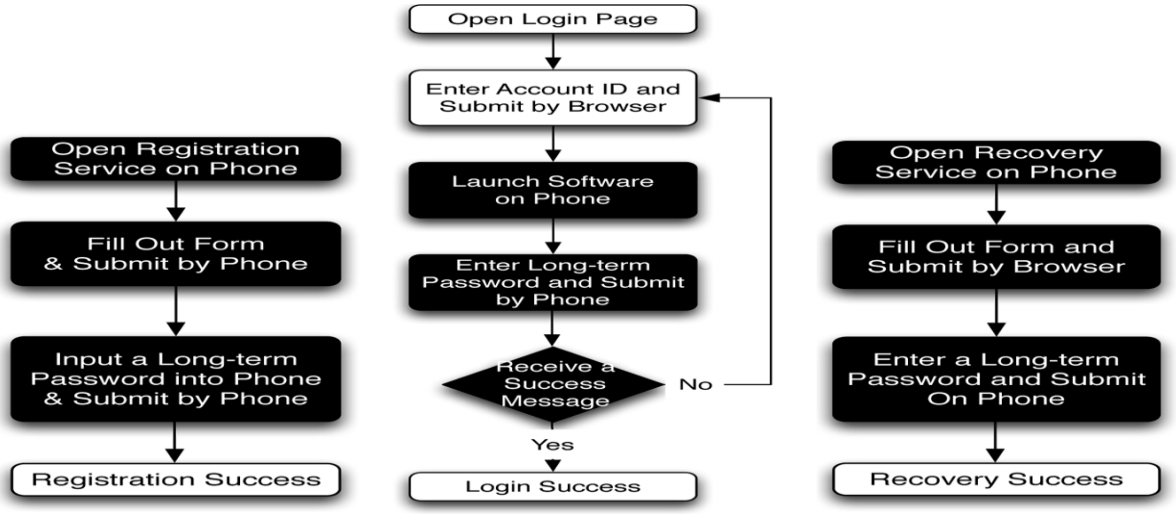


Fig. 2. Operation flows for user in each phase of oPass system respectively. Black rectangles indicate extra steps contrasted with the generic authentication system: a) registration, (b) login and (c) recovery.

IV. OPass

In this section, we present oPass from the user perspective to show operation flows. oPass consists of registration, login, and recovery phases. Fig. 2 describes the operation flows of users during each phase of oPass. In the *registration* phase, a user starts the oPass program to register her new account on the website she wishes to visit in the future. Unlike conventional registration, the server requests for the user's account id and phone number, instead of password. After filling out the registration form, the program asks the user to setup a long-term password. This long-term password is used to generate a chain of one-time passwords for further logins on the target server. Then, the program automatically sends a registration SMS message to the server for completing the registration procedure. The context of the registration SMS is encrypted to provide data confidentiality. oPass also designed a *recovery* phase to fix problems in some conditions, such as losing one's cellphone. Contrasting with general cases, *login* procedure in oPass does not require users to type passwords into an untrusted web browser. The user name is the only information input to the browser. Next, the user opens the oPass program on her phone and enters the long-term password; the program will generate a one-time password and send a login SMS securely to the server. The login SMS is encrypted by the one-time password. Finally, the cellphone receives a response message from the server and shows a success message on her screen if the server is able to verify her identity. The message is used to ensure that the website is a legal website, and not a phishing one. Protocol details of each phase are provided as follows. Table I shows the notations used in the oPass system.

A. Registration Phase

Fig. 3 depicts the registration phase. The aim of this phase is to allow a user and a server to negotiate a shared secret to authenticate succeeding logins for this user. The user begins by opening the oPass program installed on her cellphone. She enters ID (account id she prefers) and ID (usually the website url or domain name) to the program. The mobile program sends ID and ID to the telecommunication service provider (TSP) through a 3G connection to make a request of registration. Once the TSP received the user ID and the server ID, it can trace the user's phone number based on user's SIM card. The TSP also plays the role of third-party to distribute a shared key between the user and the server. The shared key is used to encrypt the registration SMS with AES-CBC. The TSP and the server will establish an SSL tunnel to protect the communication. Then the TSP forwards ID to the assigned server. Server will generate the corresponding information.

For this account and reply a response, including server's identity ID, a random seed, and server's phone number. The TSP then forwards ID and a shared key to the user's cellphone. Once reception of the response is finished, the user continues to setup a long-term password with her cellphone. The cellphone computes a secret credential by the following operation:

To prepare a secure registration SMS, the cellphone encrypts the computed credential with the key and generates the corresponding MAC, i.e., HMAC. HMAC-SHA1 takes input user's identity, cipher text, and IV to output the MAC [10]. Then, the cellphone sends an encrypted registration SMS to the server by phone number as follows:

$$c = \mathcal{H}(P_u \| \mathrm{ID}_s \| \phi).$$

$$\text{Cellphone} \xrightarrow{SMS} S : \mathrm{ID}_u, \{c \| \phi\}_{K_{sd}}, IV, \mathrm{HMAC}_1.$$

Server can decrypt and verify the authenticity of the registration SMS and then obtain with the shared key Server also compares the source of received SMS with to prevent SMS spoofing attacks. At the end of registration, the cellphone stores all information ID, except for the long term password and the secret. Variable indicates the current index of the one-time password and is initially set to 0. With, the server can authenticate the user device during each login.

B. Login Phase

The *login* phase begins when the user sends a request to the server through an untrusted browser (on a iosk).The user uses her cellphone to produce a one-time password and deliver necessary information encrypted with to server via an SMS message. Based on preshared secret credential, server can verify and authenticate user based on Fig. 4 shows the detail flows of the login phase. The protocol starts when user wishes to log into her favorite web server (already registered). However, begins the login procedure by accessing the desired website via a browser on an untrusted kiosk. The browser sends a request to with user's account ID. Next, server supplies the ID and a fresh nonce to the browser. Meanwhile, this message is forwarded to the cellphone through Bluetooth or wireless interfaces. After reception of the message, the cellphone inquires related information from its database via ID, which includes server's phone number and other parameters. The next step is promoting a dialog for her long-term password. Secret shared credential can regenerate by inputting the correct on the cellphone. The one-time password for current login is recomputed using the following operations:

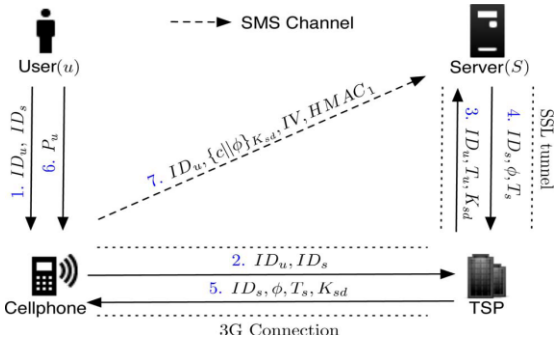


Fig. 3. Registration phase

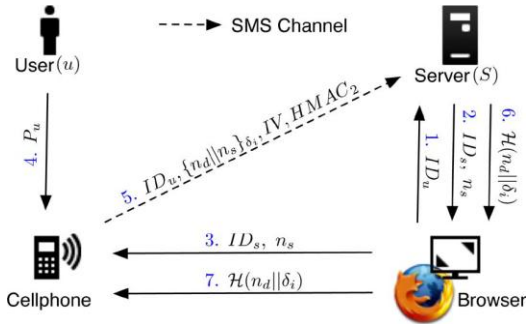


Fig. 4. Login phase

After receiving the login SMS, the server recomputed (i.e.,) to decrypt and verify the authenticity of the login SMS. If the received equals the previously generated, the user is legitimate; otherwise, the server will reject this login request. Upon successful verification, the server sends back a success message through the Internet, to the user device. The cellphone will verify the received message to ensure the completion of the login procedure.

D. Recovery Phase

Recovery phase is designated for some specific conditions; for example, a user may lose her cellphone. The protocol is able to recover oPass setting on her new cellphone assuming she still uses the same phone number (apply a new SIM card with old phone number). Once user installs the oPass program on her new cellphone, she can launch the program to send a recovery request with her account ID and requested server ID to predefined TSP through a 3G connection. As we mentioned before, ID can be the domain name or URL link of server. Similar to registration, TSP can trace her phone number based on her SIM card and forward her account ID and the server ID to server through an SSL tunnel. Once server receives the request probes the account information in its database to confirm if account is registered or not. If account ID exists, the information used to compute the secret credential will be fetched and be sent back to the user. The server generates a fresh nonce and replies a message which consists of ID s. This message includes all necessary elements for generating the next one-time passwords to the user. When the mobile program receives the message, like registration, it forces the user to enter her long-term password to reproduce the correct one-time password (assuming the last successful login before lost her cellphone is). During the last step, the user's cellphone encrypts the secret credential and server nonce to a cipher text. The recovery SMS message is delivered

back to the server for checking. Similarly, the server computers and decrypts this message to ensure that user is already recovered. At this point, her new cellphone is recovered and ready to perform further logins. For the next login, one-time password will be used for user authentication. Fig. 5 shows the detail flows of *recovery* phase.

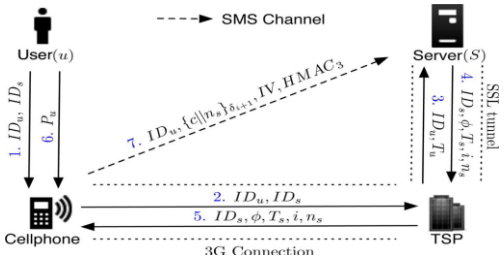


Fig. 5. Recovery phase

V. Security Analysis

All sorts of attacks may happen in such settings, including key logger, malware, and phishing. Hence, we define a threat model of oPass and demonstrate that oPass is secure later. Furthermore, we examine oPass by a cryptographic protocol verifier, ProVerif to guarantees the necessary security features claimed by oPass.

A. Threat Model

In our setting, attackers can intercept, eavesdrop, and manipulate any message transmitted over the Internet and other wireless networks. The attacker can also spoof an SMS to cheat websites.

1) User side—In this category, the malicious threat is originated from user side. Password stealing is an effective method to complete the attacker's goal. The attacker can launch phishing attacks, such as phishing websites and phishing emails, to swindle passwords out of users. A user's computer probably installs some malwares (e.g., key logger and Trojan horses). Furthermore, as we mentioned earlier, users always choose weak passwords which are vulnerable to dictionary attacks. Users also suffer from password reuse attacks. In oPass, the attacker can steal a user's cellphone to log on websites.
2) Server side: The malicious threat in this category is different from user side. Attackers exploit vulnerabilities of oPass to pass the authentication without being detected by the websites. For example, the attacker can intercept and manipulate messages to launch reply and SMS spoofing attacks.

We assume that attackers cannot break basic cryptographic primitives. For instance, the cipher text cannot be decrypted without the corresponding secret key, and hash function is irreversible.

B. Attacks on Registration

The main task of the registration phase is to generate a shared credential for computing one-time passwords between users and websites. The shared credential should be kept secret to guard oPass from attacks. To prove the guaranteed secrecy of credential in the registration phase, we translate our desired feature and system settings to the Horn clauses. Then we verify our aim through performing ProVerif with those clauses. The registration phase, which consists of seven messages, requires 26 rules to be defined. ProVerif assumes that an attacker can intercept every message (includes SMS) between the cellphone, the TSP, and the server. Meanwhile, we hypothesize that the attacker does not know the session key of 3G connection; the attacker also does not know the session key of SSL tunnel. We also make the assumption that the attacker cannot obtain the long-term password because it is directly typed into the malware-free cellphone by the user. ProVerif was able to confirm that the attacker cannot derive the secret credential. Although we only present the result of registration by using ProVerif, the other two phases, login and recovery, provide the same level of security.

C. Attacks on Login

In the login protocol, the attacker can launch attacks to masquerade itself as a legitimate user without being detected. The attacker has no way of obtaining a one-time password for login even if he builds a spoof website to launch a phishing attack because the secret key for encryption and is never transmitted. The attacker also cannot recover the encrypted login SMS. Hence, phishing attacks do not work under oPass. In oPass, users type their accounts into the kiosk and type their long-term passwords into the cellphones. A kiosk that is installed with malwares or key loggers to snatch user passwords is also useless. oPass achieves a one-time password approach to prevent against password reuse attacks. If an attacker steals a user's cellphone and attempts to log into a website that the victim has visited, he will not succeed because he does not know the user's long-term password, so he cannot

generate a legal one-time password for the next round. Even if the attacker interrupts the login procedure after obtaining the login SMS, login will still fail, because the nonce generated by the server does not match. In order to launch a MITM attack, an attacker must fully control (i.e., interception, eavesdropping, and manipulation) all transmission channels. Suppose the attacker launches an MITM attack between the server and the browser (see Fig. 4). If he tampers with field of message 2, the server will detect the MITM attack once it receives a login SMS from the legitimate user. Because SMS channel is an out-of-band channel, the attacker cannot intercept the SMS. It indicates that message 5 will be sent to the legal web server, not the attacker. Hence, oPass resists MITM attacks.

D. Attacks on Recovery

Potential threat in the recovery protocol is whether an attacker who stole a user's cellphone can succeed in guessing the correct long-term password. This attack is referred to as the password guessing attack. The attacker may try to guess the user's to compute the one-time password for login. He only has to masquerade as a normal user and execute the recovery procedure. After receiving the message in Step 5) from the TSP, the attacker enters a guessed and computes a candidate one-time password. The attacker then transmits a login SMS to the server. However, the attacker has no information to confirm whether or not the candidate is correct. Therefore, the protocol prevents a password guessing attack.

E. Further Discussion

1) Issues of Phone Number Authenticity: Phone number is a critical factor of oPass since we adopt the SMS channel for message exchanging. The potential issue is how users ensure that the phone number received is actually from the desired website rather than a forged phishing website. To address this difficulty, *registration* and *recovery* phases involve a telecommunication service provider (TSP). We assume that TSP provides a service (e.g., cellphone application) to support registration and recovery procedures of oPass. Users input the identity of the desired websites (e.g., Facebook) to the TSP's service. TSP will establish an SSL tunnel with the website before forwarding messages sent from users to it. Based on the SSL protocol, TSP can verify the website's certificate to prevent phishing attacks. Therefore, we can ensure that the phone number is actually from the correct website rather than phishing websites. In addition, the SSL tunnel provides data confidentiality. The communication interface between cellphone and TSP is 3G. 3G provides data confidentiality to protect the messages exchange. Hence, the secret key can be securely distributed by the TSP to both the cellphone and the server for registration use. A malicious user cannot decrypt other users' registration SMS unless he compromised their . This mechanism guards against insider attacks. Another potential issue about the phone number is that websites may change their telecom service provider.

2) One-Time Password Refreshment: The hash chain of a one-time password will be consumed entirely. We introduce parameter to solve this problem. The server checks the status of hash chain after receiving a legal login SMS. If the rest of the one-time passwords are less than , the server sends a new seed to the cellphone at Steps 6) and 7) of the login procedure Once the cell phone gets the new seed, it computes a new credential and sends it to the server through the SMS channel. Hence, the user and the server will use the new hash chain for the next login. This facility can be automatically completed without user effort.

3) Resistance to Phishing Attacks: Although we setup a reasonable assumption: user cellphones should be malware-free, the long-term password still suffers from phishing attack by means of a browser in the cellphone. Via social engineering, the user might input his long-term password into a malicious web site through the cellphone's browser. Even though an attacker can obtain, oPass is still secure since the attacker has no enough information to recompute the credential. Message is only transmitted by the server in the registration and recovery phases; most important of all, the transmission through SSL tunnel and 3G connection ensures data confidentiality and privacy.

4) Password Reuse: Password reuse is a serious problem in the present user authentication systems. To repair this problem, oPass adopts an OTP approach. Even if the long-term password is used for every account, the OTP approach still ensures that all logins are independent. Based on the design, is one of inputs to compute the credential. Ideally, different web servers randomize different to compute distinct. Then distinct derives distinct OTP sequence for login. Therefore, oPass users do not reuse same passwords for different accounts since generated OTP sequences guarantee randomness for each login.

5) Weak Password: Regarding the weak password problem, users tend to pick weak passwords because the passwords are easy to remember. In oPass system, a user just remembers a long-term password for all accounts. Unfortunately, user behavior is not easy to change. To help users, oPass adopts a checker to evaluate the security strength of passwords in the registration phase. If the selected password cannot satisfy the preferred security, oPass would suggest a random strong password or allow the users picking a new one again.

VI. Experiment Design

We implemented a prototype of oPass and conducted a user study to analyze the performance and usability.

A. Prototype Implementation

To make the progress smooth, the user only has to key in her long-term password and select a website. Then the remaining operations would perform by the program through clicking a button. All required interactions are eliminated to ensure oPass's efficiency. Currently, a fully functional extension has been implemented on Firefox browsers. Users installed the extension on the browser in a kiosk. The major purpose of the browser extension on kiosks is allowing forwarding data from the web server to the user's cellphone during the login phase. While the user attempts to login on a predefined website, the extension automatically sets up a TCP server socket. After the user clicks the login button on her cellphone, a connection could be built and forward data from the website to the cellphone and vice versa. In the web server implementation, a server program which consists of main server codes (PHP) and setup scripts for database (MYSQL). Server program can be installed and performed on an Apache HTTP server. On the other hand, capacity of sending/receiving SMS via a GSM modem relies on an open source library SMS Lib. For simulating TSP, partial PHP codes and related information were also established by the database.

B. User Study

However, most of them were not familiar with the use of a smart phone, especially typing on phones. Participants completed individual tests which consisted of three processes that included setting up, registering, logging in. Before starting the study, participants were first asked to complete a demographics questionnaire. They were then introduced to the oPass system. They were told that they would be setting up the system, registering an account, and logging in via a cellphone. Further, they were instructed to choose a strong long-term password that should be at least eight digits long. Participants completed one practice test (not included in the analysis data) to ensure that they understood how to operate the system. They then proceeded to complete a formal test which consisted of the following steps.

1) Setting up the system: Different from the ordinary user authentication system, users should install a cellphone soft-ware and a browser extension to setup the oPass system.
2) Registering for an account: Users first open the registration software on the cellphone. Users then fill out a form, which includes an account id, a website's id, and a long-term pass-word, and submit it to the website.
3) Logging into the website: Users first enter their account id into the browser on the kiosk and submit it to the server. Users then type their long-term password into the cell-phone and submit to the server. The login succeeds if a success message is shown on the screen of cellphone. If login fails, participants should try again until they are successful. After the test, the participants also completed a post-test questionnaire in order to collect their opinions.

VIII. Relatedwork

A number of previous researchers have proposed to protect user credentials from phishing attacks in user authentication. The proposed systems leverage variable technologies, for example, mobile devices, trusted platform module (TPM), or public key infrastructure (PKI). However, these solutions were short of considering the negative influence of human factors, such as password reuse and weak password problems. To prevent compromising user credentials, Wuet al .in 2004 proposed an authentication protocol depending on a trusted proxy and user mobile devices. Secure login is authenticated by a token (mobile device) on untrusted computers, e.g., kiosks. To thwart phishing sites, a random session name is sent by SMS from the proxy to the mobile device. The authors declared that security of the proposed system depends on SMS, which are encrypted with A5/1. However, algorithm A5/1 has been broken by Barkan and Biham in 2006. The system is also vulnerable to cellphone theft. On the contrary, oPass encrypts every SMS before sending it out and utilizes a long-term password to protect the cell phone. Another well-known approach is MP-Auth protocol presented by Mannan and Oorschot in 2007[11]. To strengthen password-based authentication in untrusted environments, MP-Auth forces the input of a long-term secret (typically a user's text password) through a trusted mobile device. Before sending the password to an untrusted kiosk, the password is encrypted by a preinstalled public key on a remote server.MP-Auth is intended to guard passwords from attacks raised by untrusted kiosks, including key loggers and malware. In spite of that, MP-Auth suffers from password reuse vulnerability. An attacker can compromise a weak server, e.g., a server without security patches, to obtain a victim's password and exploit it to gain his access rights of different websites. On the other hand MP-Auth assumes that account and password setup is secure.

TABLE-I Comparing opass with previous method

	Attack Prevention			Requirement						
	Phishing	Keylogger	Password reuse	UICC	Physical account setup	Logical account setup	TPM	On-device secret	Trusted proxy	Malware-free mobile
oPass	✓	✓	✓			•			•	•
MP-Auth [43]	✓	✓		•	•					•
Phoolproof [23]	✓	✓		•	•			•		•
Wu *et al.* [41]	✓	✓			•			•	•	•
BitE [44]		✓			–		•	•		•
Garriss *et al.* [25]		✓		•	–		•	•		•
SessionMagnifier [45]		✓			–			•		•

Users should setup an account and password via physical contact, such as banks requiring users to initialize their account personally or send passwords though postal service. In oPass, it addresses above weakness and removes this assumption. OPass achieves one-time password approach to thwart the password reuse problem. Similar to previous works, Parno utilized mobile devices as authentication tokens to build an anti-phishing mechanism is called Phool proof, via mutual authentication between users and websites. To log on the website, a user should provide the issued public key and username/password combination. Again, Phool-proof is still vulnerable to the password reuse problem and needs physical contacts to ensure that account setup is secure. On the other hand, some literature represents different approaches to prevent phishing attacks. Session Magnifier enables an extended browser on a mobile device and a regular browser on a public computer to collaboratively secure a web session. Session Magnifier separates user access to sensitive inter-actions (online banking). For sensitive interactions, the content is sent to the extended browser on the users mobile de-vice for further confirmation from a user. Another avenue is adopting TPM. McCuneet al designed a bump in the ether (BitE) based on TPM [12]. Via BitE, user inputs can be protected under an encrypted tunnel between the mobile device and an application running on a TPM-equipped untrusted computer. To ensure trustworthy computing on kiosks, Garriss et al. invent another system leveraging by TPM and virtual machine (VM) technologies. Table I summarizes comparisons of oPass with previous systems [14].Many of proposed systems require user involvement in certificate confirmation (UICC) in order to setup a secure SSL tunnel. However, prior research concluded that users do not understand SSL and often ignore the SSL warnings caused by illegal certificates [13]. Consequently, users often accept the received certificates without verification. This inattentive behavior causes users to suffer from potential attacks, such as MITM and DNS spoofing attacks. From previous literature, users should pay attention to confirming server certificate validities by themselves. The significant difference between oPass and other related schemes is that oPass reduces the negative impact of user misbehaviors as much as possible. In the oPass system, the SSL tunnel is established between a TSP and a web site server. From the perspective of users, they feel comfortable since there is no further need to verify the server's certificate by themselves. In other words, overhead on verifying server certificates for users is switching to the TSP. The TSP acts as a users' agent to validate server certificates and also establish SSL tunnels correctly. Therefore, oPass still resists phishing attacks even if users misbehave. The account setup process is classified into two types: physical and logical setup. MP-Auth, Phool proof, and Wu et al. schemes all assume that users must setup their accounts physically. They establish shared secrets with the server via a secret (conceals) out-of-band channel. For example, banks often re-quire users to setup accounts personally through physical contact or utilize the postal service. Conversely, oPass deployed an alternative approach, logical account setup, which allows users to build their accounts without physical contact with the server. In the oPass system, we invite a TSP in the registration phase to accomplish as the same security as physical account setup. oPass inherits existing trust relations between the TSP and the subscribers (i.e., users) in the telecommunication system. The user identities were authenticated by the TSP when they applied their cellphone numbers. With this trust relation, users can smoothly setup their accounts via cellphones without physical contact mentioned above. In commercial considerations, it is much easier to promote a new system if we could make the system seamless (only requiring few additional efforts). In oPass, we require a TSP (trusted proxy) to enhance the security. We think this requirement is reasonable and not costly since 3G telecommunication is widely applied. Considering performance, the TSP is only involved in the registration and recovery phases. These two phases would be executed a few times for each use. In conclusion, oPass resists most attacks and has fewer requirements than the other systems.

IX. Conclusion

A user authentication protocol named oPass which leverages cell phone and SMS to thwart password stealing and password reuse attacks. oPass assume that each website possesses a unique phone number. It also assume that a telecommunication service provider participants in the registration and recovery phases. The design principle of oPass is to eliminate the negative influence of human factors as much as possible. Through oPass, each user only needs to remember a long-term password which has been used to protect their cell phone. Users are free from typing any passwords into untrusted computers for login on all websites. OPass is efficient for website authentication to prevent phising, keylogger and malware. SMS delay could increase the execution time and reduce the performance. The performance of oPass can be improved by Round Robin DNS with the help of simultaneous response from the server for multiple users at a time. Internet relay chat protocol can be used for synchronous conferencing of SMS service. There by communication overhead can be reduced because of many transactions.

References

[1] Hung-Min Sun, Yao-Hsin Chen, and Yue-Hsun Lin,"oPass: A user authentication protocol resistant to password stealing and password reuse attacks," ieee transactions on information forensics and security, vol. 7, no. 2, april 2012

[2] S.Gaw and E.W.Felten,"Password management strategies for onlineaccounts," inSOUPS '06: Proc. 2nd Symp. Usable Privacy. Security, New York, 2006, pp. 44–55, ACM.

[3] I. Jermyn, A. Mayer, F. Monrose, M. K. Reiter, and A. D. Rubin, "The design and analysis of graphical passwords," inSSYM'99: Proc. 8thConf. USENIX Security Symp, Berkeley, CA, 1999, pp. 1–1, USENIX Association.

[4] J. Thorpe and P. van Oorschot, "Towards secure design choices for im-plementing graphical passwords," presented at the 20th. Annu. Com-puter Security Applicat. Conf, 2004.

[5] S.Wiedenbeck, J. Waters, J.-C. Birget, A. Brodskiy, and N. Memon,"Passpoints: Design and longitudinal evaluation of a graphical pass-word system,"Int. J. Human-Computer Studies, vol. 63, no. 1–2, pp.102–127, 2005.

[6] B. Pinkas and T. Sander, "Securing passwords against dictionary at-tacks," inCCS '02: Proc. 9th ACM Conf. Computer Communications Security, New York, 2002, pp. 161–170, ACM.

[7] J. A. Halderman, B. Waters, and E. W. Felten, "A convenient method for securely managing passwords," inWWW '05: Proc. 14th Int. Conf. World Wide Web, New York, 2005, pp. 471–479, ACM.

[8] K.-P. Yee and K. Sitaker, "Passpet: Convenient password management and phishing protection," inSOUPS '06: Proc. 2nd Symp. Usable Pri-vacy Security,New York,2006,pp.32–43,ACM.

[9] L. Lamport,, "Password authentication with insecure communication," Commun. ACM, vol. 24, pp. 770–772, Nov. 1981.

[10] H. Krawczyk, "The order of encryption and authentication for pro-tecting communications (or: How secure is SSL?)," inAdvances Cryp-tology—CRYPTO 2001, 2001, pp. 310–331.

[11] M. Mannan and P. van Oorschot,, "Using a personal device to strengthen password authentication from an untrusted computer," Financial Cryptography Data Security, pp. 88–103, 2007.

[12] C. Yue and H. Wang, "SessionMagnifier: A simple approach to secure and convenient kiosk browsing," inProc. 11th Int. Conf. UbiquitousComputing, 2009, pp. 125–134, ACM.

[13] D. Wendlandt, D. G. Andersen, and A. Perrig, "Perspectives: Im-proving ssh-style host authentication with multi-path probing," in Proc. USENIX 2008 Annu. Tech. Conf., Berkeley, CA, 2008, pp. 321–334, USENIX Association.

[14] R. Biddle, S. Chiasson, and P. van Oorschot, "Graphical passwords: Learning from thefirst twelve years," in ACM Computing Surveys, Carleton Univ., 2010.

Implementation of Serial Communication IP for Soc Applications

[1]K. Raghuram, [2]A.Lakshmi sudha,
[1]Assoc.Professor,Pragati Engg College, Kakinada, AP, India.
[2]Post Graduate Student, Pragati Engg College, Kakinada, Ap, India.

Abstract:

The serial communication is very commonly used communication protocol between various peripherals and processor. The current trend is all high speed buses are built with serial communication interface.

The ALTERA's NIOS II soft processor and PowerPC hard processor are widely used in FPGA based CSOC (configurable system on chip) applications. These processers don't have programmable serial links for interfacing with embedded peripherals which are mostly off chip.

In this project it is proposed to implement dynamically configurable serial communication block in Verilog. The developed module shall be interfaced with NIOS II soft processor as a general purpose IO port. The serial interface blocks shall be implemented to handle high data rate serial links and provide parallel interface to the processor.

The Nios II IDE (EDK) shall be used for developing the test application in C programming language. The serial interface blocks which are coded in Verilog shall be synthesized using QUARTUS II EDA tool. The CYCLONE III family FPGA board shall be used for verifying the results on board.

I. INTRODUCTION

In electronic design a semiconductor intellectual property core, IP core, or IP block is a reusable unit of logic, cell, or chip layout design that is the intellectual property of one party. IP cores may be licensed to another party or can be owned and used by a single party alone. The term is derived from the licensing of the patent and source code copyright intellectual property rights that subsist in the design.

IP cores can be used as building blocks within ASIC chip designs or FPGA logic designs It IP cores in the electronic design industry have had a profound impact on the design of systems on a chip. By licensing a design multiple times, IP core licensor spread the cost of development among multiple chip makers. IP cores for standard processors, interfaces, and internal functions have enabled chip makers to put more of their resources into developing the differentiating features of their chips. As a result, chip makers have developed innovations more quickly.

II. Typres of Ip Cores:

The IP core can be described as being for chip design what a library is for computer programming or a discrete integrated circuit component is for printed circuit board design.

A. SOFT CORES:

As the complexity of embedded systems designs increased over time, designing each and every hardware component of the system from scratch soon became far too impractical and expensive for most designers. Therefore, the idea of using pre-designed and pre-tested intellectual property (IP) cores in designs became an attractive alternative. Soft-core processors are microprocessors whose architecture and behavior are fully described using a synthesizable subset of a hardware description language (HDL). They can be synthesized for any Application-Specific Integrated Circuit (ASIC) or Field Programmable Gate Array (FPGA) technology; therefore they provide designers with a substantial amount of flexibility.

The use of soft-core processors holds many advantages for the designer of an embedded system. First, soft-core processors are flexible and can be customized for a specific application with relative ease. Second, since soft-core processors are technology independent and can be synthesized for any given target ASIC or FPGA technology, they are therefore more immune to becoming obsolete when compared with circuit or logic level descriptions of a processor.

Finally, since a soft-core processor's architecture and behavior are described at a higher abstraction level using an HDL, it becomes much easier to understand the overall design. This paper presents a survey of the available soft-core processors that are used to design and implement embedded systems using either FPGAs or ASICs.

B. HARD CORES:

Hard cores, by the nature of their low-level representation, offer better predictability of chip performance in terms of timing performance and area.

Analog and mixed-signal logic are generally defined as a lower-level, physical description. Hence, analog IP (SerDes, PLLs, DAC, ADC, etc.) are provided to chip makers in transistor-layout format (such as GDSII.) Digital IP cores are sometimes offered in layout format, as well.

Such cores, whether analog or digital, are called "hard cores" (or hard macros), because the core's application function cannot be meaningfully modified by chip designers. Transistor layouts must obey the target foundry's process design rules, and hence, hard cores delivered for one foundry's process cannot be easily ported to a different process or foundry. Merchant foundry operators (such as IBM,Fujitsu, Samsung, TI, etc.) offer a variety of hard-macro IP functions built for their own foundry process, helping to ensure customer lock-in.

III. Commercial Cores and Tools:

Nios II, Micro Blaze, Pico Blaze and Xtensa are the leading soft-core processors provided by Altera, Xilinx and Ten silica respectively. In this section, we will discuss the important features of each soft-core processor. Nios II by Altera Corporation: Altera Corporation is one of the leading vendors of Programmable Logic Devices (PLDs) and FPGAs.

They offer the Stratix, Stratix II and Cyclone families of FPGAs that are widely used in the design of embedded systems and digital signal processing (DSP) applications. They also provide associated CAD tools such as Quartus II and System-on-Programmable-Chip (SOPC) Builder that allow designers to synthesize, program and debug their designs and build embedded systems on Altera's FPGAs.

The Nios II Processor is their flagship IP soft-core processor and can be instantiated with any embedded system design. This processor is the successor of Altera's original Nios softcore processor and features major improvements focused on the reduction of logic element (LE) consumption on an FPGA and improved performance.

The Nios II Soft-Core Processor is a general purpose Reduced Instruction Set Computer (RISC) processor core and features Harvard memory architecture. This core is widely used with Altera FPGAs and SOPC Builder. This processor features a full 32-bit Instruction Set Architecture (ISA), 32 general-purpose registers, single-instruction 32x32 multiply and divide operations, and dedicated instructions for 64-bitand 128-bit products of multiplication.

The Nios II also has a performance of more than 150 Dhrystone MIPS (DMIPS) on the Stratix family of FPGAs. This soft-core processor comes in three versions: economy, standard and fast core. Each core version modifies the number of pipeline stages; instruction and data cache memories and hardware components for multiply and divide operations. In addition, each core varies in size and performance depending on the features that are selected.

Adding peripherals with the Nios II Processors is done through the Avalon Interface Bus which contains the necessary logic to interface the processor with off-the-shelf IP cores or custom made peripherals. Micro Blaze and Pico Blaze by Xilinx Incorporated: Xilinx Incorporated are the makers of the Spartan and Virtex families of FPGAs. In addition, they also offer soft IP cores that target their FPGAs.

The fundamental components that build up a HW/SW system are a CPU (Central Processing Unit) which processes the information in a system, On-chip RAM (Random Access Memory) to store the instructions for the CPU and a JATG UART (Joint Test Access Group Universal Asynchronous Receiver Transmitter) for communication with the host computer. These components communicate with each other through the system bus, see figure below.

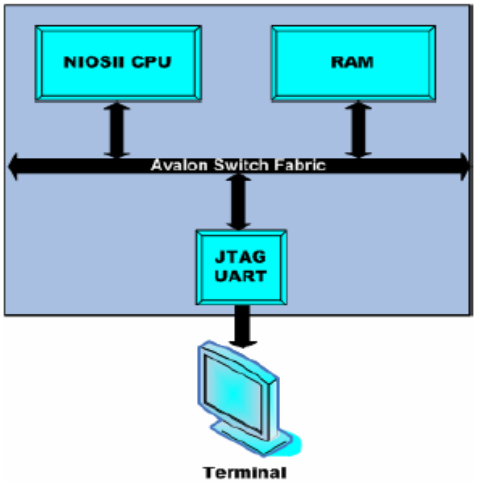


Figure1. system Architecture

The system is generated with the help of SOPC Builder tool. This tool makes easy to specify the system components and their connections and generate a complete system-on programmable- chip (SOPC) in much less time than using traditional, manual integration methods.

IV. Software Development For The System In Niosii Ide:

By this stage the HW structure of the system is complete. To utilize it and verify whether it is working correctly, software has to be created. The programming language that is used is ANSI C. ANSI C (Standard C) is one standardized version of the C programming language. Before the code (software) can be generated and executed, a project has to be built in Nios II IDE("user application project") which in turn needs a system library project ("Hardware Abstraction Layer (HAL) system library project"). The system library is created by Nios II IDE automatically after the user application project is created.

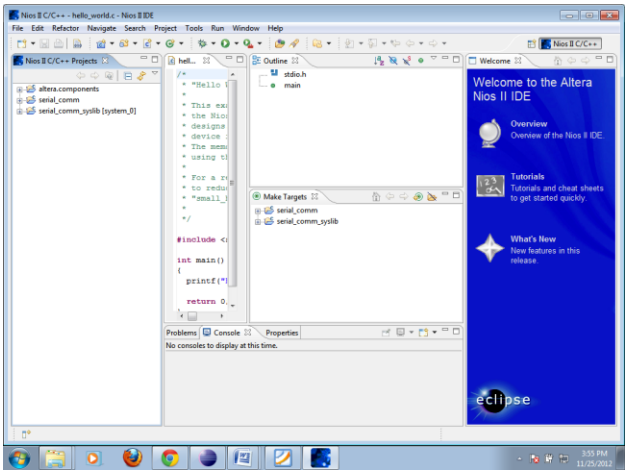


Figure 2. Nios II IDE window.

Flow Summary

Flow Status	Successful - Mon Nov 26 21:58:37 2012
Quartus II Version	9.0 Build 132 02/25/2009 SJ Web Edition
Revision Name	esdsd
Top-level Entity Name	serial_commn_top
Family	Cyclone III
Device	EP3C16F484C6
Timing Models	Final
Met timing requirements	N/A
Total logic elements	1,600 / 15,408 (10 %)
Total combinational functions	1,509 / 15,408 (10 %)
Dedicated logic registers	832 / 15,408 (5 %)
Total registers	832
Total pins	13 / 347 (4 %)
Total virtual pins	0
Total memory bits	43,008 / 516,096 (8 %)
Embedded Multiplier 9-bit elements	0 / 112 (0 %)
Total PLLs	0 / 4 (0 %)

Figure 3 Performance report

PERFORMANCE REPORT:

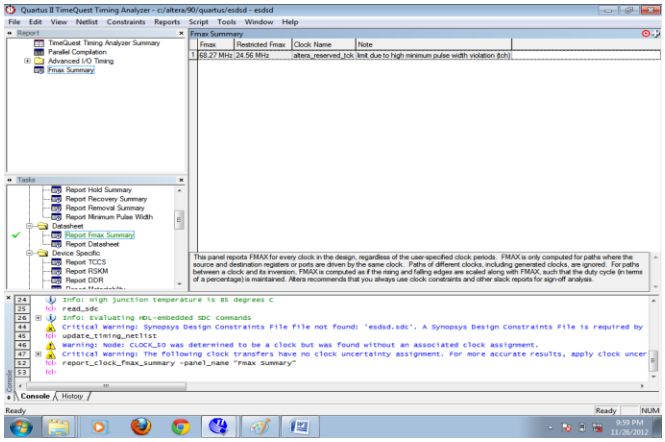


Figure 4. Fmax. Summary report of slow corner.

V. Conclusion

The serial interface blocks are implemented to handle high data rate serial links and provide parallel interface to the processor. The serial interface is interconnected with processor finally through FPGA implementation we verified the functionality. The serial interface blocks which is coded in Verilog is successfully synthesized using QUARTUS II EDA tool. We got the speed rate of around 150MHz clock rate.

Reference

[1] X. Wang and S.G. Ziavras, "Parallel LU Factorization of Sparse Matrices onFPGA-Based Configurable Computing Engines," *Concurrency and Computation:Practice and Experience*, vol. 16, no. 4, (April 2004), pp. 319-343.

[2] Altera Corporation, "Nios Embedded Processor System Development," [OnlineDocument, Cited 2004 February 2], Available HTTP:http://www.altera.com/products/ip/processors/nios/nio-index.html

[3] Xilinx, Inc., "MicroBlaze Soft Processor," [Online Document, Cited 2004 February2],vailableHTTP: http://www.xilinx.com/xlnx/xil_prodcat_product.jsp?title=microblaze

[4] K. Compton and S. Hauck, "Reconfigurable Computing: A Survey of Systems andSoftware," *ACM Computing Surveys*, vol. 34, no. 2 (June 2002), pp. 171-210.

[5] R. K. Gupta and Y. Zorian, "Introducing Core-Based System Design," *IEEE Design andTest of Computers*, vol. 14, no. 4 (October-December 1997), pp 15-25.

[6] Opencores.org Web Site, [Online Document, Cited 2004 February 9]

[7] Altera Corporation, "Excalibur Devices," [Online Document, Cited 2004 February 7],Available HTTP: http://www.altera.com/products/devices/arm/arm-index.html

[8] Xilinx, Inc., "PowerPC Embedded Processor Solution," [Online Document, Cited 2004February 7], Available HTTP:http://www.xilinx.com/xlnx/xil_prodcat_product.jsp?title=v2p_powerpc

[9] Xilinx, Inc., "MicroBlaze Processor Reference Guide," [Online Document], 2003September, [Cited 2004 February 2], Available HTTP:http://www.xilinx.com/ise/embedded/mb_ref_guide.pdf

[10] Xilinx, Inc., "PicoBlaze 8-Bit Microcontroller for Virtex-E and Spartan-II/IIE Devices,"[Online Document], 2003 February, [Cited 2004 February 2], Available HTTP:http://www.xilinx.com/bvdocs/appnotes/xapp213.pdf

[11] V. Betz, J. Rose, and A. Marquardt, *Architecture and CAD for Deep-Submicron FPGAs*,Kluwer Academic Publishers: Norwell, MA, 1999.

[12] Altera Corporation, "Stratix Device Handbook," [Online Document], 2004 January, [Cited 2004 February 3], Available HTTP:http://www.altera.com/literature/hb/stx/stratix_handbook.pdf

A Novel Light-Sensor-Based Information Transmission System for Outdoor tracking tothe Indoor Positioning system

[1]Dr.Shaik Meeravali, [2]S.VenkataSekhar
Department of Electronics and Communication Engineering,
RRS College of Engineering and Technology, Muthangi,
Faculty of Electronics and Communication Engineering,
Jawaharlal Nehru Technological University, Hyderabad, India

Abstract

The objective of this project describes a novel light-sensor-basedInformation transmission system for indoor positioning andNavigationwith particular benefits for mobile and wearable computers. It canseamlessly extend outdoor GPS tracking tothe indoor environment.In a novel manner, fluorescent lightis used as the mediumto transmit information. The userreceives the encoded light information through a photoreceiver. The information ispassed into the wearable or mobile computer after the data aredecoded. This information allows positioning information to begiven to indoor mobile and wearable computers.The proposedsystem can be used in indoor guidance and navigation applications.An embedded system is a combination of software and hardware to perform a dedicated task.Some of the main devices used in embedded products are Microprocessors and Microcontrollers.

Microprocessors are commonly referred to as general purpose processors as they simply accept the inputs, process it and give the output.Itusing PIC16F72, PIC16F877A microcontroller is an exclusive project which is used to find the position identification for the different blocks in the organization by using the Zigbee module. This information is provided by the GPS with the help of the data it receives from the satellites.

Index Terms—Augmented reality (AR), electronic ballast, fluorescentlamp, navigation, wearable computer.

I. Introduction

MOBILE or wearable computers and augmented realitytechnology are finding applications in human positionguidance and navigation [1]. Commonly, GPS sensors havewidely been used with these interactive technologies for navigationand positioning. For example, GPS-based positioningfor wearable computers has been used in the applicationof outdoor augmented reality (AR).AR merges virtual objects or text information into a realenvironment and displays this combination in real time. Unlikevirtual environments, AR supplements reality, rather than completelyreplacing it. This property makes AR particularly wellsuited as a tool to aid the user's perception of and interactionwith the real world. The information conveyed by the virtualobjects helps a user perform real-world tasks.Although AR technology combined with wearable GPS ismature, the information transmission method for wearable GPScannot provide information indoors or in crowded urban areas since the signals from the satellite would be shielded by thearmored concrete structure of the building. One might instead use active badges or beacon architectures, but installing and maintaining such systems involves substantial effort and high expense. Hence, indoor tracking system development becomes useful to seamlessly extend outdoor tracking into indoors. Some forms of indoor positioning, such as magnetic and ultrasonic sensing, are also available, but they are normally for a short range and expensive and require complex hardware installations. Thus, there is a problem that such commercially available sensing systems for indoor tracking of mobile and wearable computers are accurate but impractical and expensive for wide areas.

This project aim is to identifying the different blocks in the organization by using the Zigbee module. And the employer can be able to find where the employee in the organization. All the multinational companies are having more than 50 blocks in a single building those are working for different projects. So it is difficult to find by the new employee to know which block is belongs to which category. For this we are going to develop a new project which is apt for the new employee's to know the different blocks in the organization.

II. Comparison with Other Systems

To extend GPS data for indoor applications, some researchers used computer-vision-based tracking algorithms to perform the tracking. For instance, put fiducialmarkerson the walls and used a marker-based tracking algorithm forindoor tracking of a mobile user. Although this kind of trackingis only software based and there is no need for any specialhardware, except for the paper markers, if we want to use thismethod, we need to have many different markers and put themin every place to cover the whole area, and in the state of artin computer vision tracking systems, we can detect less than100 markers at the same time. Furthermore, all the markersmust be predefined for the users, and the user's mobile devicemust know which position each marker is located, which is notpractical when the user arrives to a new building.

Other proposed methods for indoor tracking aremainly basedon ultrasound, radio frequency, and IR. In addition to these technologies, because of the popularity of wireless networks inrecent years, many works have been done to infer the locationof a wireless client based on Wi-Fi technology on the IEEE802.11 standard. For instance, the Finnish company Ekahau has developed a software-basedWi-Fi location technology.In their system, they only need three wireless stations for theircalculation, and the rests are done in software. Although eachtechnology has its own advantages and disadvantages, in general,there is a tradeoff between the accuracy of the tracking andthe total cost of the system. For example, ultrasound trackingcan be highly accurate, such as the IS-900 system developed bythe Intersense Company, with a price of over 15 000 USD,or it can be designed in a cheap way like the system proposed by Randell and Muller, which costs about 150 USD withan accuracy of 10–25 cm. In Table I, we listed the cost andaccuracy of different indoor tracking systems in comparisonwith our system. As can be seen in this table, the proposedsystem has the lowest tracking performance (on the order ofoutdoor GPS), but it is the cheapest one as well. As a result, oursystem is not suitable for applications that need highly accuratetracking, such as virtual reality applications, and because it is one of the cheapest methods for indoor tracking, it is a goodcandidate for applications such as navigation and guidance(which does not need highly accurate tracking).
In comparison with different technologies for indoor tracking, the proposed system is similar to IR tracking systems such as the method used in [14], which used an IR tracking systemin an AR application.

	Cost in USD	Accuracy	Technology
Intersense IS 900	Over15000	1mm	Ultrasound
Randell's system	150	10-25cm	Ultrasound
Ekahau	100-200	1 m	Wi-Fi
Proposed system	Less than 10	3-4 m	light

III. Hardware System Design

In this section, we will outline the hardware system used forconstructing novel and economical navigation and positioningsystems using fluorescent lamps. The whole system is dividedinto two parts: the transmitter and the receiver. The transmittersends out messages encoded by the fluorescent light whoseflicking is imperceptible to human vision, while the receiverdetects the light using a photo-detector.

In the transmitter section, information can be encoded intothe light through arc frequency variation [see Fig. 1(a)]. Here,we use a fluorescent lamp for our system since, first, it is highlyused in office buildings and, second, nowadays, it is triggeredby electronic ballast circuits, so there is no need to design acostly circuit for controlling the arc frequency of the lamp,and by simple modifications on the current widely cheap andavailable circuit, we can furnish our goal.We add a simple lowcostmicrocontroller chip to control the light frequency from 35 to 40 kHz.

The receiver circuit [Fig. 1(b)], with a photodetector detectingthe fluorescent light, processes the data that are eventuallyfed into the wearable computer. With the information received,the wearable computer can tell the user what the surroundingsituation is.

In the rest of this section, we detail our transmitter andreceiver circuits, and then,we explain the wearable computersystem in terms of how the receiver and other components areintegrated together.

A. Transmitter Circuit

The hardware for the developed transmitter is shown inFig.2, and the schematic circuit diagram is depicted in Fig.3.As shown in this figure, the electronic ballast circuit used for the transmission purpose consists of three parts: the ac–dc rectifier,the dc–ac converter (inverter), and the resonant filter circuit.

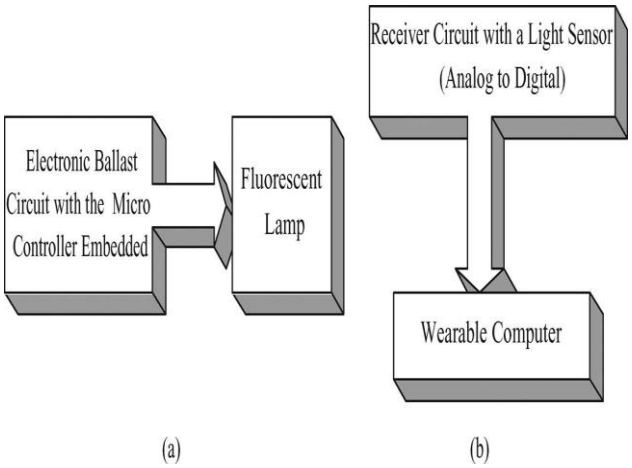


Fig: 1Simple schematic scheme of the system. (a) Transceiver. (b) Receiver.

A Novel Light-Sensor-Based Information Transmission System for Indoor Positioning and Navigation

1. Transmitter section

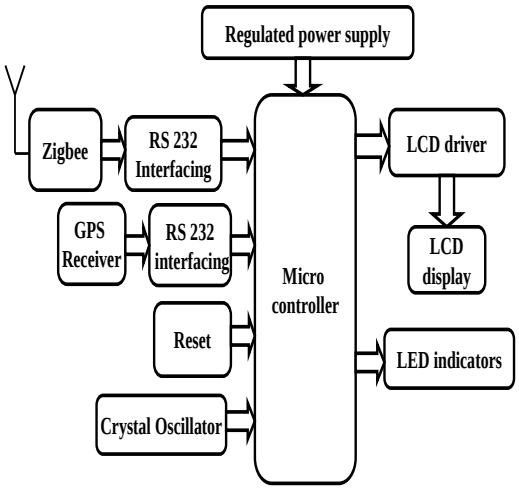


Fig 2: Block diagram of transmitter section

Fig 3: Transmitter hardware.

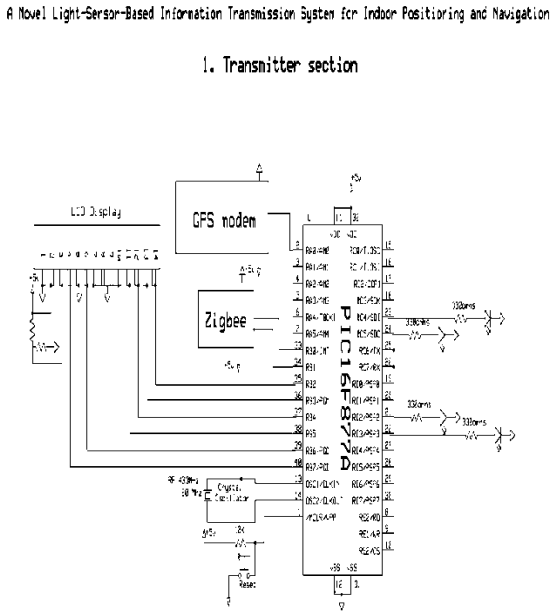


Fig 4: Schematic diagram of transmitter section

B. Receiver Circuit

The receiver detects the fluorescent light and transforms theanalog signals to the digital ones that can be sent to the user's mobile/wearable device. Fig. 5 shows the block diagrams of thereceiver part with a wearable computer. The core part of this receiver system is the receiver circuit, which is shown in Fig. 6. As can be seen in Fig. 5, the main parts of the receiver circuit are as follows:

a) Bandpassfilter:

The bandpass filter is designed to removenoise that is received together with the Manchester-codedinformation in the light.

b) Zero-crossing detector: This block converts the analoginput signal to digital signal. Note that only the frequency of the signal contains information and not itsamplitude.

c) Phase-lockedloop (PLL):

This block converts the incoming digital signal to an analog voltage proportional to thefrequency of incoming signal.

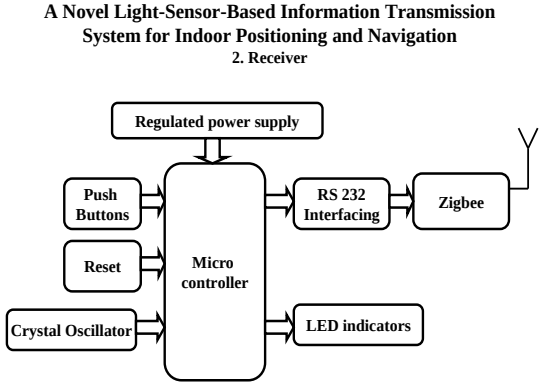


Fig 5: Block diagram of receiver section

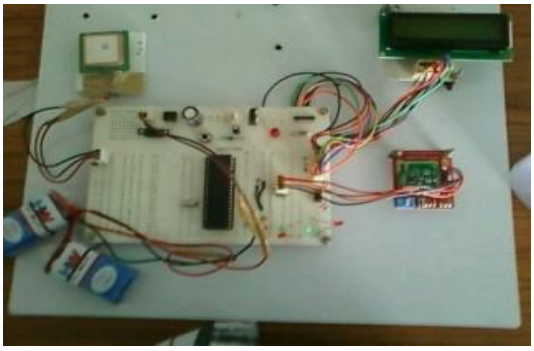


Fig 6: Receiver hardware.

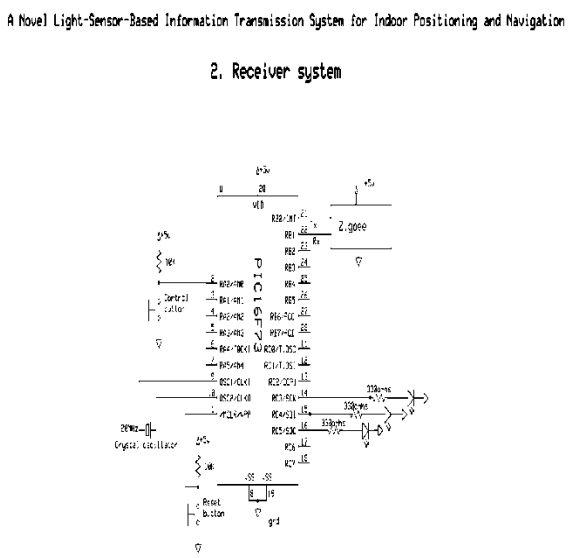


Fig 7: Schematic diagram of receiver section.

IV.EXPERIMENTS OFSYSTEMAPPLICATIONS AND RESULTS

In the previous sections, we discussed the hardware of the indoor navigation system in detail. The indoor system can seamlessly be activated by simply switching from the tra-ditional outdoor GPS system to this indoor system.a GPS signal is no longer received, a fluorescent lamp data code is sought after by the wearable computer. The data codes of the fluorescent lamp are directly tied to a GPS position relative to the outdoor GPS reading. Thus, the indoor fluorescent lamp position is direction correlated to an outdoor GPS position.

Fig. 9 shows the transition from the outdoor environment to the indoors and the screenshots on the HMD at the user's different positions (outdoor and indoors), and Fig. 9 shows the selected screenshots at outdoor and indoor locations. The messages and information appear on the left bottom corner of the HMD, which does not affect the user's eyesight range, providing the user with real-time environmental information. Fig. 10 presents the data flow of the proposed indoor tracking system.

What is more, in the large urban indoor environment, a 3-D digital map stored on the wearable computer can be developed to display on the HMD. The user exactly knowshis or her location by watching his or her place on the map,with the position recognized by the fluorescent light tracking system.

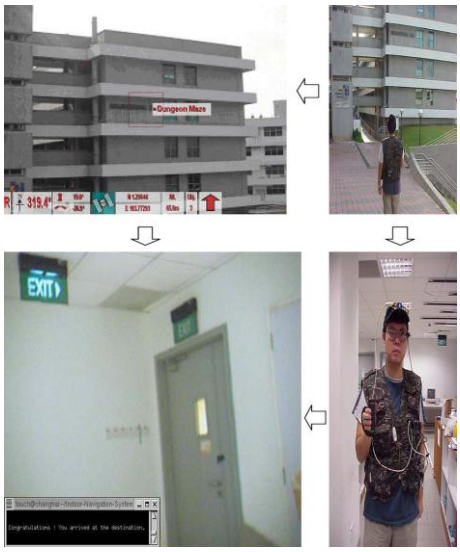


Fig 9: Selected screenshots from the HMD when a user is at outdoor and indoor locations

The project "**A Novel Light-Sensor-Based Information Transmission System for Indoor Positioning and Navigation**" was designed such that to identifying the different blocks in the organization by using the Zigbee and GPS modules.The locations are displayed on LCD when the person reaches those particular locations.Main Applications, This system can be used by blind people in order to know the present location.The system can also be used for tracking the locations

Fig 10(a): Indoor tracking display

Fig 10(b): Indoor tracking display

VI. Conclusion And Future Works

This paper has addressed the problem of enabling economical indoor tracking systems, which are similar to GPS, available with seamless transition function from outdoor GPS tracking environment to indoor. We have focused on the task of indoor navigation and positioning, where the particular aspect of the user's state that is of interest is the user's physical location. Byusing an economical light sensor to build this indoor tracking system, we have been able to infer the user's location in anindoor environment. For example, we can put the system on both sides of the doors of each room in a building; then, a user will receive the proper data by entering or leaving the room.

When data, which are encoded in the fluorescent light, is receivedby the receiver and analyzed by the wearable computer, it will provide location and navigation messages. Specifically, this light-sensor-based navigation and tracking system

is robust and much cheaper than those using electromagnetic ultrasonic sensors. Furthermore, the receiver circuit is light and small, and it can be well suited to wearable computer applications. Aside from thetechnical achievements of our work to date, it is significantto point out that the application of this system provides an innovative and economical form of indoor positioning and navigation method. It must be noted though that the proposed system has limited bandwidth and is therefore more suitable for transmitting text messages to the user's wearable computer rather than images or graphics.

Integrating features of all the hardware components used have been developed in it. Presence of every module has been reasoned out and placed carefully, thus contributing to the best working of the unit. Secondly, using highly advanced IC's with the help of growing technology, the project has been successfully implemented. Thus the project has been successfully designed and tested.

The project "A Novel Light-Sensor-Based Information Transmission System for Indoor Positioning and Navigation" is mainly intended to alert the person through location names displayed on LCD when he enters into a particular location by using GPS and Zigbee module. This system consists of a GPS receiver, Zigbee modules which are interfaced to the micro controller. The location names are displayed on LCD. The micro controller is programmed in such a way that depending on the satellite information of location the predefined location name will be announced and also displays on the LCD and also GPS receiver.

This project can be extended using high efficiency GPS receiver and a GSM module. The GSM module gives the intimation of the person with this system through SMS.

REFERENCES

[1] C. Lee, A. Wollets, H. Tan, and A. Pentland, "A wearable haptic navigation guidance system," in Proc. 2nd Int. Symp. Wearable Comput.,1998.

[2] T. Moore, "An introduction to the global positioning system and its applications," in Proc. Develop. Use Global Positioning Syst., 1994,

[3] B. Thomas, V. Demczuk, W. Piekarski, D. Hepworth, and B. Gunther, "A wearable computers system with augmented reality to support ter-restrial navigation," in Proc. 2nd Int. Symp. Wearable Comput., 1998,

[4] B. Thomas, B. Close, J. Donoghue, J. Squires, P. D. Bondi, and W. Piekarski, "First person indoor/outdoor augmented reality application: ARQuake," Pers. Feb. 2002.

[5] T. Caudell and D. Mizell, "Augmented reality: An application of heads-up display technology to manual manufacturing processes," inProc. Hawaii Int. Conf. Syst. Sci., 1992.

[6] A.R.Golding and N.Lesh,"Indoor navigation using a diverse set of cheap, wearable sensors," in Proc. 3rd Int. Symp. Wearable Comput., 1999.

[7] D. K. Jackson, T. T. Buffaloe, and S. B. Leeb, "Fiat lux: A fluorescent lamp digital transceiver," IEEE Trans. Ind. Appl., vol. 34 , May/Jun. 1998.

[8] S. Bjork, J. Falk, R. Hansson, and P.Ljungstrand, "Pirates! Using the physical world as a game board," inProc. Interact, 2001.

AUTHORS DETAILS:

FIRST AUTHOUR:

Dr. SHAIK MEERAVALI,
Professor and Head, Department of Electronics and Communication Engg, RRS College of Engineering and Technology, Muthangi, Andhra Pradesh, India.

SECOND AUTHOUR:

S. VENKATA SEKHAR, Post Graduate Student,
Department of Electronics and Communication Engg, RRS College of Engineering and Technology, Muthangi, Patancheru, Hyderabad, Andhra Pradesh, India.

Implementation of Berlekamp Algorithm for Error Detection and Correction of Multiple Random Errors Using Reed-Solomon Codes

P. Chiranjeevi[1] (M.Tech), D. Ramadevi[2], Asst.Prof. K. Jeevan Reddy[3], HOD
[1, 2, 3,] Department of ECE, Teegala Krishna Reddy Engineering College/JNTU, India

Abstract:
In the communication systems, RS codes have a widespread use to provide error protection. For burst errors and random errors, RS code has become a popular choice to provide data integrity due to its good error correction capability. This feature has been one of the important factors in adopting RS codes in many practical applications such as wireless communication system, cable modem, computer memory and ADSL systems. Reed Solomon codes are an important sub class of non-binary BCH codes. These are cyclic codes and are very effectively used for the detection and correction of burst errors. Galois field arithmetic is used for encoding and decoding of reed Solomon codes. The design experience will be formulated to form the complete design methodology of the FEC modules at the register-transfer level (RTL). Then we incorporate the knowledge into our RS code generator design flow.

Keywords: *RS codes, random errors, BCH codes, Galois Field, ADSL*

I. Introduction

Digital communication system is used to transport information bearing signal from the source to a user destination via a communication channel. The information signal is processed in a digital communication system to form discrete messages which makes the information more reliable for transmission. Channel coding is an important signal processing operation for the efficient transmission of digital information over the channel. It was introduced by Claude E. Shannon in 1948 by using the channel capacity as an important parameter for error free transmission. In channel coding the number of symbols in the source encoded message is increased in a controlled manner in order to facilitate two basic objectives at the receiver one is Error detection and other is Error correction. Error detection and Error correction to achieve good communication is also employed in devices. It is used to reduce the level of noise and interferences in electronic medium.

The aim of the project is to correct multiple random errors and burst errors that are occur during the transmission of the information by using Reed Solomon codes. The proposed code is designed using verilog coding and the results demonstrate that the reed Solomon codes are very efficient for the detection and correction of burst errors.

II. Proposed Product Codes

The Reed Solomon code is an algebraic code belonging to the class of BCH (Bose-Chaudhry-Hocquehen) multiple burst correcting cyclic codes. The Reed Solomon code operates on bytes of fixed length. Given m parity bytes, a Reed Solomon code can correct up to m byte errors in known positions (erasures), or detect and correct up to m/2 byte errors in unknown positions. This is an implementation of a Reed Solomon code with 8 bit bytes, and a configurable number of parity bytes. The maximum sequence length (code word) that can be generated is 255 bytes, including parity bytes. In practice, shorter sequences are used.

A typical system is shown in the figure 4.1 below:

Fig 1 Typical RS encoder-decoder system

The Reed-Solomon encoder takes a block of digital data and adds extra "redundant" bits. Errors occur during transmission or storage for a number of reasons. The Reed-Solomon decoder processes each block and attempts to correct errors and recover the original data. The number and type of errors that can be corrected depends on the characteristics of the Reed-Solomon code. A Reed-Solomon code is specified as RS (n, k) with s-bit symbols.

This means that the encoder takes k data symbols of s bits each and adds parity symbols to make an n symbol codeword. There are n-k parity symbols of s bits each. A Reed-Solomon decoder can correct up to t symbols that contain errors in a codeword, where 2t = n-k.

The following figure 4.2 shows a typical Reed-Solomon codeword (this is known as a Systematic code because the data is left unchanged and the parity symbols are appended):

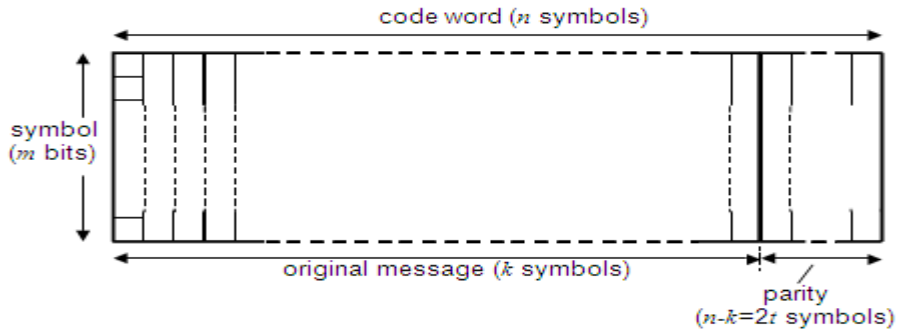


Fig 2 RS codeword format

1. Reed Solomon Encoder

Consider a Reed-Solomon code RS (255,247) with 8-bit symbols. Each codeword contains 255 code word bytes, of which 247 bytes are data and 8 bytes are parity. For this code: n = 255, k = 247, s = 8, 2t=8, t=4. The decoder can correct any 4 symbol errors in the code word: i.e. errors in up to 4 bytes anywhere in the codeword can be automatically corrected.

The k information symbols that form the message to be encoded as one block can be represented by a polynomial M(x) of order k-1, so that:

$$M(x) = Mk\text{-}1xk\text{-}1 + ... + M1x + M0$$

Where each of the coefficients Mk-1, ..., M1, M0 is an m-bit message symbol, that is, an element of GF(2m). Mk-1 is the first symbol of the message.

To encode the message, the message polynomial is first multiplied by xn-k and the result divided by the generator polynomial, g(x). Division by g(x) produces a quotient q(x) and a remainder r(x), where r(x) is of degree up to n-k-1. Thus:

$$\frac{M(x)\times x^{n-k}}{g(x)} = q(x) + \frac{r(x)}{g(x)}$$

The following diagram shows an architecture for a systematic RS(255,247) encoder:

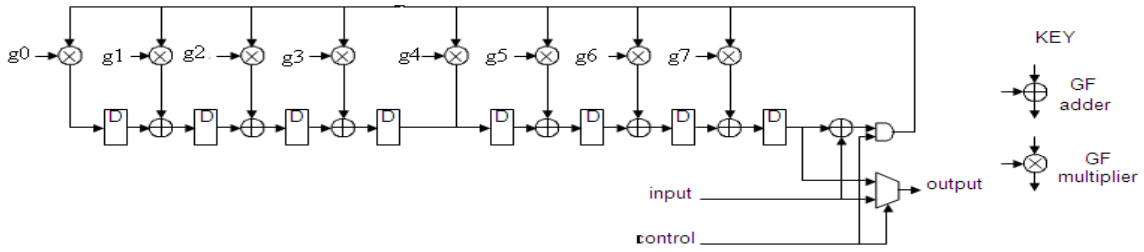


Fig 3 RS (255,247) Encoder

Reed-Solomon codes may be shortened by making a number of data symbols zero at the encoder, not transmitting them, and then re-inserting them at the decoder.

Reed Solomon Decoder

The Reed Solomon decoder tries to correct errors and/or erasures by calculating the syndromes for each codeword. Based upon the syndromes the decoder is able to determine the number of errors in the received block. If there are errors present, the decoder tries to find the locations of the errors using the Berlekamp Massey algorithm by creating an error locator polynomial.

The roots of this polynomial are found using the Chien search algorithm. Using Forney's algorithm, the symbol error values are found and corrected. For an RS (n, k) code where n - k = 2t, the decoder can correct up to t symbol errors in the code word. Given that errors may only be corrected in units of single symbols (typically 8 data bits).

The functional flow chart of the Reed Solomon Decoder is shown in the figure 5.1 below:

The blocks of the Reed Solomon Decoder are:

1) Syndrome calculation block
2) Error Location Determination block
3) Error value calculation block
4) Error correction block

The purpose of the decoder is to process the received code word to compute an estimate of the original message symbols.

The RS decoder block is shown in the figure below:

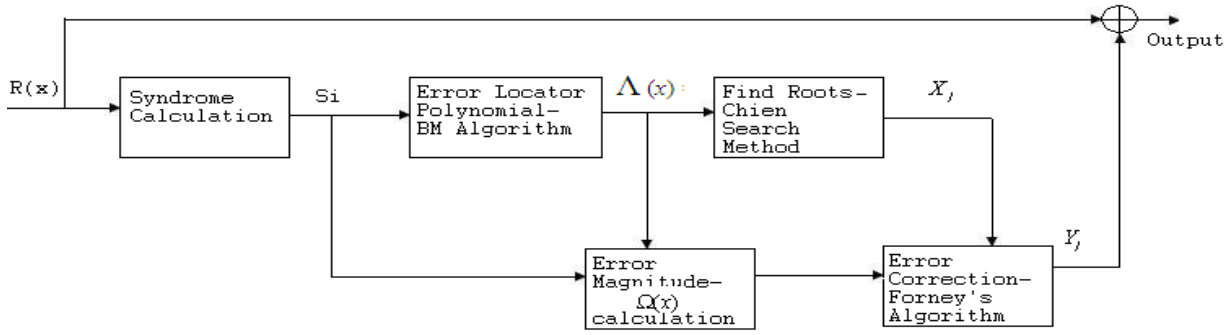


Fig 4. RS Decoder block

There are three main blocks to the decoder first is syndrome generator, then Barlekamp Massey algorithm and the Chien/Forney block. The output of the Chien/Forney block is an estimate of the error vector. This error vector is then added to the received codeword to form the final codeword estimate. Note that the error value vector Y comes out of the Chien/Forney block in reverse order, and it must pass through a LIFO/FIFO block before it is added to the received codeword R(x).

III. Performance

The proposed code can correct up to 16 symbol errors by column wise decoding. In the proposed code since two shortened RS codes are used it can correct up to 16 symbol errors.

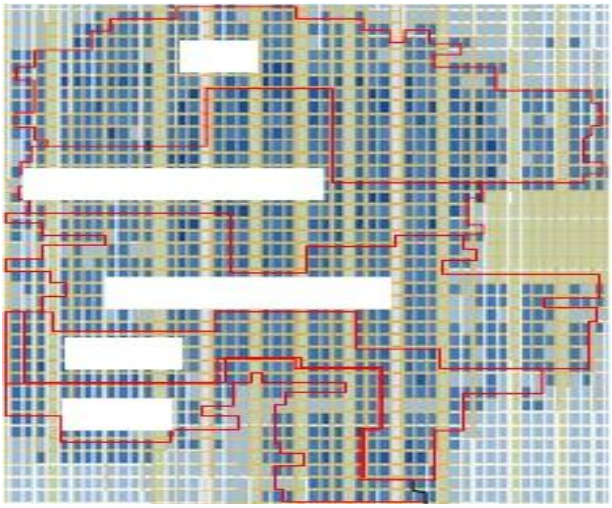


Fig 5 Layout of FPGA chip for the proposed RS code

IV. Results

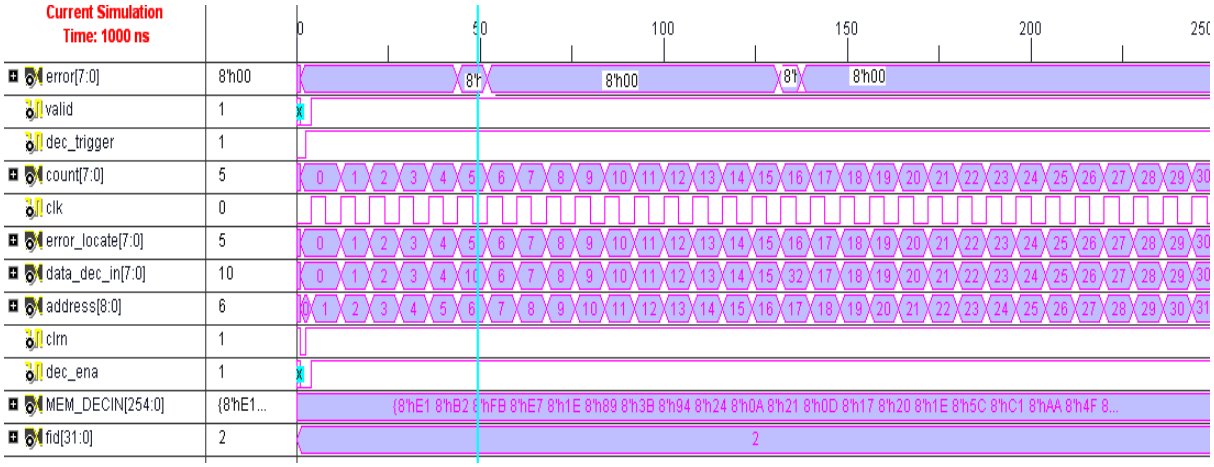


Fig 6 Result of RS(255,231) code

V. Conclusion

In this thesis, error detection and correction techniques have been used which are essential for reliable communication over a noisy channel. The effect of errors occurring during transmission is reduced by adding redundancy to the data prior to transmission. The redundancy is used to enable a decoder in the receiver to detect and correct errors. Cyclic Linear block codes are used efficiently for error detection and correction. The encoder splits the incoming data stream into blocks and processes each block individually by adding redundancy in accordance with a prescribed algorithm. Likewise, the decoder processes each block individually and it corrects errors by exploiting the redundancy present in the received data.

In this work, architectures were modeled using HDL and the functional simulation was carried out using Xilinx ISE 12.4 simulator. The compilation, synthesis and place and route, timing are done with RTL and SOC Encounter of Cadence Tool. Synthesis of the architectures was carried out on SPATAN 3 board with XC3S4000-4PQ208 target devices.

References

[1] M. Mehnert, D. F. von Droste, and D. Schiel, "VHDL Implementation of a (255, 191) Reed Solomon Coder for DVB-H," *IEEE 10th International Symposium on Consumer Electronics (ISCE), pp. 1-5*, 2006.

[2] S. B. Wicker and V. K. Bhargava, eds., Reed-Solomon Codes and their Applications. *New York: IEEE Press*, 1994.

[3] S. Lin and D. J. Costello, *Error Control coding (2nd edition)*. Upper Saddle River, NJ: Prentice-Hall, 2004.

[4] R. Koetter and A. Vardy, \Algebraic Soft-Decision Decoding of Reed-Solomon Codes," *IEEE Transactions on Information Theory, vol. 49, no. 11*, pp. 2809{2825, 2003.

[5] V. Guruswami and M. Sudan, \Improved Decoding of Reed-Solomon and Algebraic-Geometry Codes," *IEEE Transactions on Information Theory, vol. 45, no. 6*, pp. 1757{1767, 1999.

[6] R. J. McEliece, \The Guruswami-Sudan Decoding Algorithm for Reed-Solomon Codes," *Tech. Rep. 42-153, JPL Interplanetary Network Progress Report*, 2003.

A Method for Hiding Secret Messages using Minimum-Redundancy Codes

Srinivas.CH[1], D.Prabhakar[2], Jayaraman.K[3], Gopala Krishna.M[4]

[1,2]Associate Professor, Dept. of ECE, MIC College of Technology, Vijayawada, AP, India
[3]Assistant Professor, Dept. of ECE, S.M.K.FOMRA INSTITUTE OF TECHNOLOGY, Chennai, India
[4]Assistant Professor, Dept. of ECE, MIC College of Technology, Vijayawada, AP, India

Abstract:

This paper demonstrates an effective lossless data hiding scheme using minimum Redundancy codes. The binary secret message is concurrently embedded and encoded with a cover medium such as a video file, an audio file, or even a text file. The proposed scheme not only provides good data hiding capacity and data recovery capability, but also being efficient in space saving. Each symbol in a cover medium can carry one secret bit, and the cover medium can be reversed. And the experimental results show that the redundancy code can saves up to 38% of space compared with the cover medium. In this paper, the symbol or sequence of symbols associated with a given message will be called the "message code." The entire number of messages which might be transmitted will be called the "message ensemble." The mutual agreement between the transmitter and the receiver about the meaning of the code for each message of the ensemble will be called the "ensemble code."

Keywords: data hiding, redundancy codes, Secret messages, method.

I. Introduction

An optimum method of coding an ensemble of messages consisting of a finite number of members is developed. A minimum-redundancy code is one constructed in such a way that the average number of coding digits per message is minimized.

Server data formats are used to be the cover medium in data hiding, e.g. audio files, video files, image files, text files, and so on. Although the data structure of text files is similar to image files than the other data format mentioned above, most of image data hiding schemes are not suitable for text files. The main reason is that most image data hiding schemes embed secret information into cover image by slightly perturbing the pixel values. Since gray-scale or color images can tolerant a small amount modifications of pixel values, it will cause no perceptible distortions. On the contrary, any changes in the text file might lead to meaningless content.

Few studies have referred to hiding secret messages in text files. In [1], the data was embedded by modifying the inter-character space, but it resulted in some distortions in the shape of words. In [3], a technique was proposed for copyright protection that marks the text file by shifting lines up or down and words right or left; however, the technique might change the typesetting of the text file accordingly. In addition to the security problem, bandwidth consumption is also an important concern. The size of transmitted files can be reduced by either of two categories of data compression technology: lossless and lossy technologies. The lossy data compression technology is widely used in images, but it may be unsuitable for text files because any loss of data may lead the content meaningless.

II. Derived Coding Requirements

For an optimum code, the length of a given message code can never be less than the length of a more probable message code. If this requirement were not met, then a reduction in average message length could be obtained by interchanging the codes for the two messages in question in such a way that the shorter code becomes associated with the more probable message. Also, if there are several messages with the same probability, then it is possible that the codes for these messages may differ in length. However, the codes for these messages may be interchanged in any way without affecting the average code length for the message ensemble. Therefore, it may be assumed that the messages in the ensemble have been ordered in a fashion such that

$$P(1)\geq P(2)\geq \cdots P(N-1)\geq P(N)$$

and that, in addition, for an optimum code, the condition

$$L(1)\leq L(2)\leq \cdots L(N-1)\leq L(N)$$

holds. This requirement is assumed to be satisfied throughout the following discussion. It might be imagined that an ensemble code, could assign q more digits to the Nth message than to the (N-l)st message. However, the first L(N- 1) digits of the Nth message must not be used as the code for any other message. Thus the additional q digits would serve no useful purpose and would unnecessarily increase Lav. Therefore, for an optimum code it is necessary that L(N) be equal to L(N-1).

The kth prefix of a message code will be defined as the first k digits of that message code. Basic restriction (b) could then be restated as: No message shall be coded in such a way that its code is a prefix of any other message, or that any of its prefixes are used elsewhere as a message code.

Imagine an optimum code in which no two of the messages coded with length L(N) have identical prefixes of order L(N) -1. Since an optimum code has been assumed, then none of these messages of length L(N) can have codes or prefixes of any order which correspond to other codes. It would then be possible to drop the last digit of this entire group of messages and thereby reduce the value of Lav. Therefore, in an optimum code, it is necessary that at least two (and no more than D) of the codes with length L(N) have identical prefixes of order L(N)-1.

The procedure is applied again and again until the number of members in the most recently formed auxiliary message ensemble is reduced to two. One of each of the binary digits is assigned to each of these two composite messages. These messages are then combined to form a single composite message with probability unity, and the-coding is complete.

III. Hiding Terminology

As we have noted previously, there has been a growing interest, by different research communities, in the fields of steganography, digital watermarking, and fingerprinting. This led to some confusion in the terminology. We shall now briefly introduce the terminology which will be used in the rest of the paper and which was agreed at the first international workshop on the subject [4], [9] (Fig. 1).

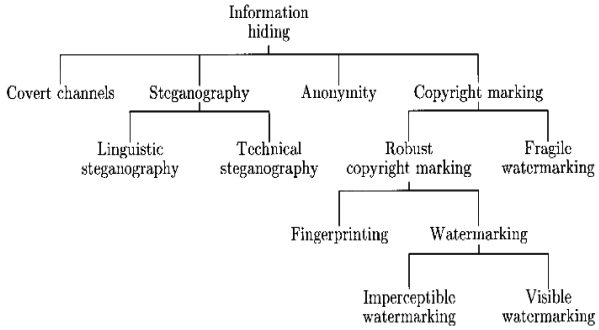


Fig.1: A classification of information-hiding techniques

The general model of hiding data in other data can be described as follows. The embedded data are the message that one wishes to send secretly. It is usually hidden in an innocuous message referred to as a cover text, cover image, or cover audio as appropriate, producing the stegotext or other stego-object. A stego-key is used to control the hiding process so as to restrict detection and/or recovery of the embedded data to parties who know it (or who know some derived key value).

As the purpose of steganography is having a covert communication between two parties whose existence is unknown to a possible attacker, a successful attack consists in detecting the existence of this communication. Copyright marking, as opposed to steganography, has the additional requirement of robustness against possible attacks. In this context, the term "robustness" is still not very clear; it mainly depends on the application. Copyright marks do not always need to be hidden, as some systems use visible digital watermarks, but most of the literature has focused on invisible (or transparent) digital watermarks which have wider applications. Visible digital watermarks are strongly linked to the original paper watermarks which appeared at the end of the thirteenth century to differentiate paper makers of that time. Modern visible watermarks may be visual patterns (e.g., a company logo or copyright sign) overlaid on digital images.

In the literature on digital marking, the stego-object is usually referred to as the marked object rather than stego-object. We may also qualify marks depending on the application. Fragile watermarks1 are destroyed as soon as the object is modified too much. This can be used to prove that an object has not been "doctored" and might be useful if digital images are used as evidence in court. Robust marks have the property that it is infeasible to remove them or make them useless without destroying the object at the same time. This usually means that the mark should be embedded in the most perceptually significant components of the object. Fingerprints (also called labels by some authors) are like hidden serial numbers which enable the intellectual property owner to identify which customer broke his license agreement by supplying the property to third parties. Watermarks tell us who is the owner of the object.

IV. Steganographic Technique

We will now look at some of the techniques used to hide information. Many of these go back to antiquity, but unfortunately many modern system designers fail to learn from the mistakes of their predecessors. By the sixteenth and seventeenth centuries, there had arisen a large literature on steganography and many of the methods depended on novel means of encoding information.

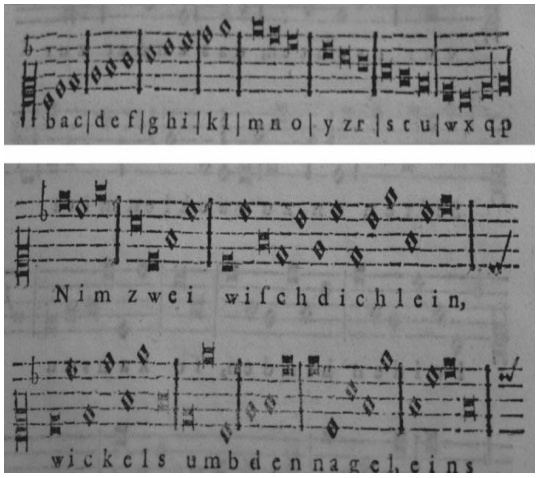


Fig. 2. Hiding information into music scores: Schott simply maps the letters of the alphabet to the notes

Schott (1608–1666) explains how to hide messages in music scores: each note corresponds to a letter (Fig. 4). Another method, based on the number of occurrences of notes and used by Bach, is mentioned in [10]. Schott also expands the "Ave Maria" code proposed by Trithemius (1462–1516) in *Steganographiæ*, one of the first known books in the field. The expanded code uses 40 tables, each of which contains 24 entries (one for each letter of the alphabet of that time) in four languages: Latin; German; Italian; and French. Each letter of the plain text is replaced by the word or phrase that appears in the corresponding table entry and the stego-text ends up looking like a prayer or a magic spell. It has been shown recently that these tables can be deciphered by reducing them modulo 25 and applying them to a reversed alphabet. In [2], Wilkins (1614–1672), Master of Trinity College, Cambridge, shows how "two

Musicians may discourse with one another by playing upon their instruments of musick as well as by talking with their instruments of speech" [2, ch. XVIII, pp. 143–150]. He also explains how one can hide secretly a message into a geometric drawing using points, lines or triangles. "The point, the ends of the lines and the angles of the figures do

each of them by their different situation express a several letter" [2, ch. XI, pp. 88–96].

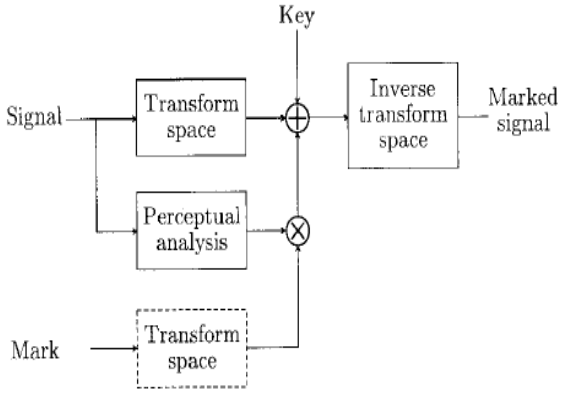


Fig. 3. A typical use of masking and transform space for digital watermarking and fingerprinting.

V. Conclusion

In this paper we gave an overview of information hiding in general and steganography in particular. We looked at a range of applications and tried to place the various techniques in historical context in order to elucidate the relationships between them, as many recently proposed systems have failed to learn from historical experience. We then described a number of attacks on information hiding systems, which between them demolish most of the current contenders in the copyright marking business. We have described a tool, StirMark, which breaks many of them by adding subperceptual distortion, and we have described a custom attack on echo hiding.

This led us to a discussion of marking in general. We described some of the problems in constructing a general theory and the practical requirements that marking schemes and steganographic systems may have to meet. We advanced the suggestion that it is impractical to demand that

VI. Acknowledgements

The authors would like to thank the anonymous reviewers for their comments which were very helpful in improving the quality and presentation of this paper.

References:

[1]. Tacticus, How to Survive Under Siege/Aineias the Tactician (Clarendon Ancient History Series). Oxford, U.K.: Clarendon, 1990, pp. 84–90, 183–193.
[2]. J. Wilkins, Mercury: Or the Secret and Swift Messenger: Shewing, How a Man May with Privacy and Speed Communicate His Thoughts to a Friend at Any Distance, 2nd ed. London, U.K.: Rich Baldwin, 1694.
[3]. D. Chaum, "Untraceable electronic mail, return addresses and digital pseudonyms," Commun. ACM, vol. 24, no. 2, pp. 84–88, Feb. 1981.
[4]. R. J. Anderson, Ed., Information Hiding: 1st Int.Workshop (Lecture Notes in Computer Science), vol. 1174. Berlin, Germany: Springer-Verlag, 1996.
[5]. S. Roche and J.-L. Dugelay, "Image watermarking based on the fractal transform," in Proc. Workshop Multimedia Signal Processing, Los Angeles, CA, 1998, pp. 358–363.
[6]. J.-P. M. G. Linnartz, "The "ticket" concept for copy control based on embedded signalling," in Computer Security—5th Europ. Symp. Research in Computer Security, (ESORICS'98) (Lecture Notes in Computer Science), vol. 1485, J.-J. Quisquater, Y. Deswarte, C. Meadows, and D. Gollmann, Eds. Berlin, Germany: Springer, 1998, pp. 257–274.
[7]. M. L. Miller, I. J. Cox, and J. A. Bloom, "Watermarking in the real world: An application to DVD," in Multimedia andSecurity—Workshop at ACM Multimedia'98 (GMD Report), vol.41, J. Dittmann, P. Wohlmacher, P. Horster, and R. Steinmetz, Eds. Bristol, U.K.: ACM, GMD—Forschungszentrum Informationstechnik GmbH, 1998, pp. 71–76.
[8]. J. C. Benaloh, Verifiable Secret-Ballot Elections, Ph.D. dissertation, Yale University, New Haven, CT, YALEU/DCS/TR-561,1987.
[9]. B. Pfitzmann, "Information hiding terminology," in Lecture Notes in Computer Science, vol. 1174. Berlin, Germany: Springer-Verlag, 1996.

Authors Profile:

Srinivas.CH, working as Associate professor in MIC college of Technology, has 10 years of industrial and teaching experience. He received his M.E degree in ECE from College of Engineering, Guindy, Anna University, Chennai in 2003.

D.Prabhakar, working as Associate professor in MIC College of Technology, has 8 years of Teaching Experience. He received his M.Tech degree in Radar & Microwave Engineering from Andhra University in 2003.

Jayaraman krishnamoorthy, working in Rajiv Gandhi Salay IT Highway (OMR), has worked as Assistant Professor in Shree Motilal Kanhaiyalal FOMRA INSTITUTE OF TECHNOLOGY. He received his M.Tech degree in VLSI from SATHYABAMA UNIVERSITY, Chennai.

Gopala Krishna.M, working as Assistant professor in MIC College of Technology, has 4 years of Teaching Experience. He received his M.Tech degree in VLSI.

Image Mining Method and Frameworks

[1]Shaikh Nikhat Fatma
Department Of Computer, Mumbai University, Pillai's Institute Of Information Technology, New Panvel

Abstract:
Image mining deals with the extraction of image patterns from a large collection of images. Clearly, image mining is different from low-level computer vision and image processing techniques because the focus of image mining is in extraction of patterns from large collection of images, whereas the focus of computer vision and image processing techniques is in understanding and / or extracting specific features from a single image. While there seems to be some overlaps between image mining and content-based retrieval (both are dealing with large collection of images), image mining goes beyond the problem of retrieving relevant images. In image mining, the goal is the discovery of image patterns that are significant in a given collection of images.

Keywords— Image mining (IM); function-driven; knowledge driven; information driven; knowledge remounting

I. INTRODUCTION

Image mining deals with extraction of implicit knowledge, image data relationship or other patterns not explicitly stored in images and uses ideas from computer vision, image processing, image retrieval, data mining, machine learning, databases and AI. The fundamental challenge in image mining is to determine how low-level, pixel representation contained in an image or an image sequence can be effectively and efficiently processed to identify high-level spatial objects and relationships. Typical image mining process involves pre-processing, transformations and feature extraction, mining (to discover significant patterns out of extracted features), evaluation and interpretation and obtaining the final knowledge. Various techniques from existing domains are also applied to image mining and include object recognition, learning, clustering and classification, just to name a few. Association rule mining is a well-known data mining technique that aims to find interesting patterns in very large databases. Some preliminary work has been done to apply association rule mining on sets of images to find interesting patterns[4,5,7]. The fundamental challenge in image mining is to determine how low-level, pixel representation contained in a raw image or image sequence can be efficiently and effectively processed to identify high-level spatial objects and relationships. In other words, image mining deals with the extraction of implicit knowledge, image data relationship, or other patterns not explicitly stored in the image databases.
Research in image mining can be broadly classified into two main directions. The first direction involves domain-specific applications where the focus is in the process of extracting the most relevant image features into a form suitable for data mining. The second direction involves general applications where the focus is on the process of generating image patterns that maybe helpful in the understanding of the interaction between high-level human perceptions of images and low level image features. The latter may lead to improvements in the accuracy of images retrieved from image databases.

In the remaining paper in section II there is an explanation for the image mining process, in section III we take the review of how Image Mining is actually done actually work. In section IV we take an example. In section V we take an over view of Image Mining Frameworks. In section VI we make a conclusion for Image Mining and the last section includes all references for this paper.

II. THE IMAGE MINING PROCESS

Figure 1 shows the image mining process. The images from an image database are first preprocessed to improve their quality. These images then undergo various transformations and feature extraction to generate the important features from the images. With the generated features, mining can be carried out using data mining techniques to discover significant patterns. The resulting patterns are evaluated and interpreted to obtain the final knowledge, which can be applied to applications.

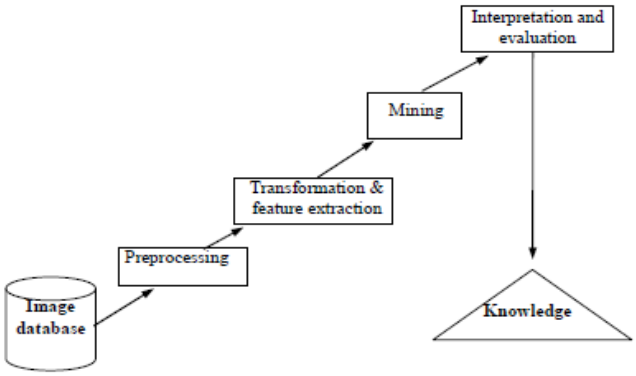


Figure 1: The image mining process

It should be noted that image mining is not simply an application of existing data mining techniques to the image domain. This is because there are important differences between relational databases versus image databases:

(a) Absolute versus relative values

In relational databases, the data values are semantically meaningful. For example, age is 35 is well understood. However, in image databases, the data values themselves may not be significant unless the context supports them. For example, a grey scale value of 46 could appear darker than a grey scale value of 87 if the surrounding context pixels values are all very bright.

(b) Spatial information (Independent versus dependent position)

Another important difference between relational databases and image databases is that the implicit spatial information is critical for interpretation of image contents but there is no such requirement in relational databases. As a result, image miners try to overcome this problem by extracting position-independent features from images first before attempting to mine useful patterns from the images.

(c) Unique versus multiple interpretation

A third important difference deals with image characteristics of having multiple interpretations for the same visual patterns. The traditional data mining algorithm of associating a pattern to a class (interpretation) will not work well here. A new class of discovery algorithms is needed to cater to the special needs in mining useful patterns from images.

III. METHOD FOR IMAGE MINING

In this section, we present the algorithms needed to perform the mining of associations within the context of images. The four major image mining steps are as follows:

1. Feature extraction. Segment images into regions identifyable by region descriptors (blobs). Ideally one blob represents one object. This step is also called segmentation.

2. Object identification and record creation. Compare objects in one image to objects in every other image. Label each object with an id. We call this step the preprocessing algorithm.

3. Create auxiliary images. Generate images with identified objects to interpret the association rules obtained from the following step.

4. Apply data mining algorithm to produce object association rules.

The idea of this method is selecting a collection of images that belong to a specific field (e.g. weather), after the selection stage we will extract the objects from each image and indexing all the images with its objects in transaction database, the data base contain image identification and the objects that belong to each images with its features. After creating the transaction data base that contains all images and its feature we will use the proposed data mining methods to associate rules between the objects. This will help us for prediction (e.g. if image sky contain black clouds then it will rain (65%))[5].
The following block diagram presents the proposed IM method:

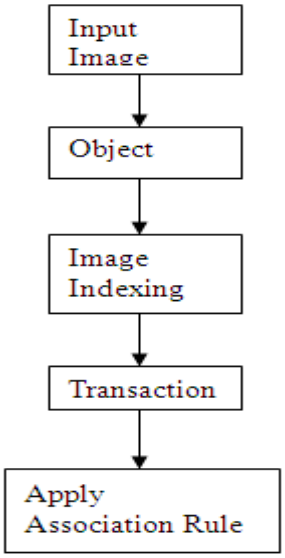


Figure 2: Block diagram of image mining method

(1) Select a collection of images that belong to the same field (E.g. medical images, geographical images, persons images, etc.)

(2) Image Retrieval. Image mining requires that images can be retrieved according to some requirement specifications. In the proposed work we comprise image retrieval by derived or logical features like objects of a given type or individual objects or persons using edge detection techniques[8].

After we extract object we will encoded it as follows:
O1: circle.
O2: triangle.
O3: square.

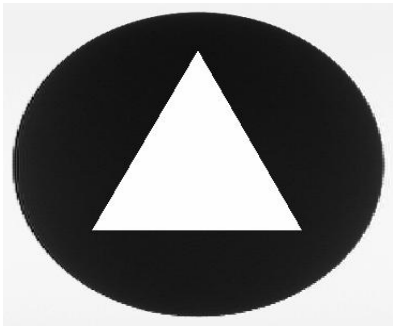


Figure 3: Example of an image

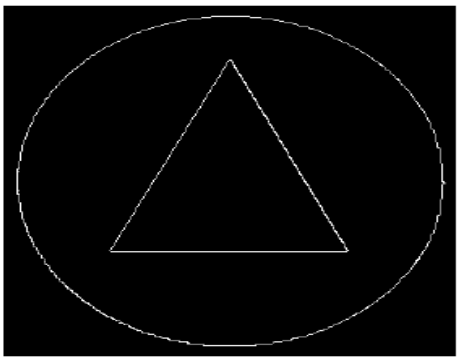


Figure 4: Object extraction using edge detection

(3) Image Indexing. Image mining systems require a fast and efficient mechanism for the retrieval of image data. Conventional database systems such as relational databases facilitate indexing on primary or secondary key(s). We will create two databases:
The first one contains all the objects that have been extracting from the images and its features[5].

Table 1: First database contains the objects and its features

object	color	size	shape	V-next to	H-next to	overlap	include
O1	Blue	*	Square	[red,circle,small]	[red,circle,*]	[red,circle,*]	34%
O2	yellow	Large	Circle	[blue,square,*]	[yellow,*, large]	[green,*,*]	40%
O3	:	:	:	:	:	:	:
...	:	:	:	:	:	:	:
On	:	:	:	:	:	:	:

Therefore the association rule with spatial relationships could be:
V-Next-to ([red, circle, small], [blue, square, *]) ^ H-Next-to ([red, circle, *], [yellow, *, large]) → Overlap([red, circle, *], [green, *, *]) (34%).
In this example, only three dimensions were needed and we made use of the wildcard * to replace absent values.

The second Database contains all the images and the objects that belong to each image.

Table 2: Second Database Contains Each Image and its Objects

Image ID	Objects
I1	{O2, O2, O1}
I2	{O2, O2, O3}
I3	{O1, O3, O2}
I4	{O3, O2}
.	.
.	.
etc	etc

(4) Finally, the last step is applying the proposed mining techniques using the data of the images that has been index to the database.

(5) After that we will use the first a proposed algorithm to find the frequent item sets from the specific table and the result will be the following:

$$\{O_2, O_2, O_4, O_4\},\ \{O_2, O_3, O_4\},\ \{O_2, O_2, O_4\},\ \{O_2, O_4, O_4\},\ \{O_2, O_3\},\ \{O_3, O_4\},\ \{O_2, O_2\},\ \{O_4, O_4\}$$

(6) The final step we will use the second proposed algorithm to find association rules between the objects and we will have the following results:
(1) {O4,O4}→ {O2,O2} [100%]
(2) {O2,O4,O4} → {O2} [100%]
(3) {O3,O4} → {O2} [100%]
(4) {O3} → {O2,O4} [100%]
(5) {O2,O2} → {O4} [100%]
(6) {O4,O4} → {O2} [100%]
(7) {O3} → {O2} [100%]
(8) {O3} → {O4} [100%]

IV. EXAMPLE

A simple example illustrating how image mining algorithms work with n=10. The original images and their corresponding blobs are shown on Figure 4 and Figure 5. Association rules corresponding to the identified objects are also shown . 10 representative images are chosen from the image set created for the experiments.

Figures 5 and 6 shows the original image at the left with several geoemetric shapes and white background. These images are labeled with an image id. These images are the only input data for the program; no domain knowledge is used. Then a series of blob images are shown, each containing one blob. These images are labeled with the id obtained by preprocessing . Each blob has a close (most times equal) position to its corresponding geometric shape. There are some cases in which one blob corresponds to several geometric shapes. For instance in image 013 object 2 corresponds to the triangle and object 3 corresponds to the circle. In image 131 object 4 corresponds to all the shapes in the image.

The data mining was done with a 20% support and 70% confidence. The output is a set of association rules whose support and confidence are above these thresholds.

The 66 rules obtained by the program are shown . Let us analyse some of these rules. The first rule { 3}➔{ 2} means that if there is circle in the image then there is also a triangle. In fact with these simple images, there was never a circle without a triangle. In this case the rule actually has a higher support and a higher confidence than that obtained by the program (50% and 83 % respectively). This happened because the circle had two different object identifiers: 3 and 7. The rule {2,3,5}➔ 8 says if there is a circle, a triangle and an hexagon then there is also a square. Once again images containing the first three shapes always contained the square. Another interesting rule is {3,11}➔ 2 .In this case the rule says that if there are a circle and an ellipse then there is also a triangle; once again the rule is valid; note that this rule has a low support.

It important to note that several incorrect or useless blob matches such as 9, 10, 13, 14, 16 are altered out by the 30% support. That is the case for images 029, 108, 119, 131 and 144. There are no rules that involve these identified objects (matched blobs).

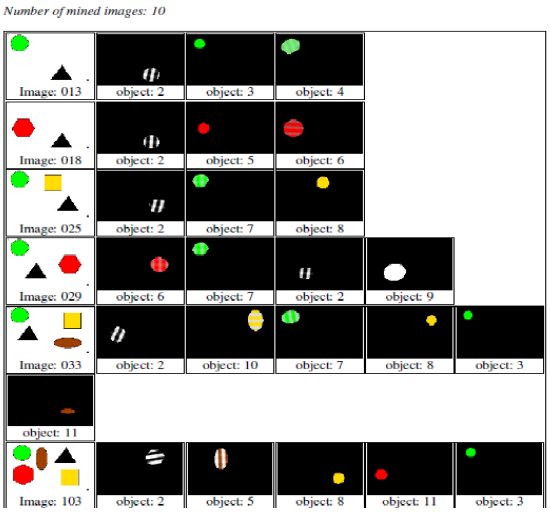


Figure 5: First part of images and blobs

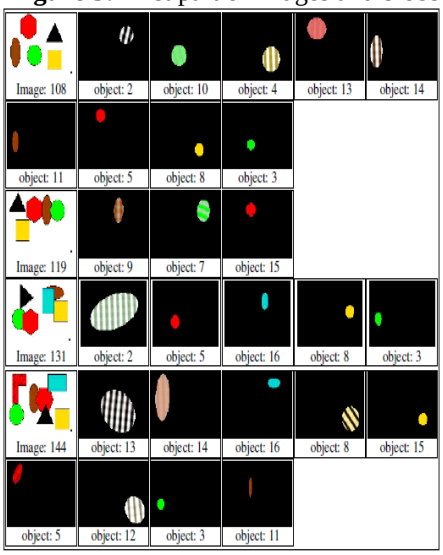


FIGURE 6: SECOND PART OF IMAGES AND BLOBS

RULES GENERATED

```
Parameters:
Support:                 20%
Confidence:              70%
Number of records:       10
Number of associations:  30

{3} ⇒ {2}   s= 50%, c= 83%        {3} ⇒ {8}   s= 50%, c= 83%
{5} ⇒ {2}   s= 40%, c= 80%        {5} ⇒ {3}   s= 40%, c= 80%
{5} ⇒ {8}   s= 40%, c= 80%        {5} ⇒ {3,8}   s= 40%, c= 80%
{7} ⇒ {2}   s= 30%, c= 75%        {8} ⇒ {2}   s= 50%, c= 83%
{8} ⇒ {3}   s= 50%, c= 83%        {11} ⇒ {2}   s= 30%, c= 75%
{11} ⇒ {3}   s= 40%, c=100%       {11} ⇒ {5}   s= 30%, c= 75%
{11} ⇒ {8}   s= 40%, c=100%       {11} ⇒ {2,3}   s= 30%, c= 75%
{11} ⇒ {2,8}   s= 30%, c= 75%     {11} ⇒ {3,5}   s= 30%, c= 75%
{11} ⇒ {3,8}   s= 40%, c=100%     {11} ⇒ {5,8}   s= 30%, c= 75%
{11} ⇒ {3,5,8}   s= 30%, c= 75%   {11} ⇒ {2,3,8}   s= 30%, c= 75%
{2,3} ⇒ {8}   s= 40%, c= 80%      {2,5} ⇒ {3}   s= 30%, c= 75%
{2,5} ⇒ {8}   s= 30%, c= 75%      {2,5} ⇒ {3,8}   s= 30%, c= 75%
{2,8} ⇒ {3}   s= 40%, c= 80%      {2,11} ⇒ {3}   s= 30%, c=100%
{2,11} ⇒ {8}   s= 30%, c=100%     {2,11} ⇒ {3,8}   s= 30%, c=100%
{3,5} ⇒ {2}   s= 30%, c= 75%      {3,5} ⇒ {8}   s= 40%, c=100%
{3,5} ⇒ {11}   s= 30%, c= 75%     {3,5} ⇒ {2,8}   s= 30%, c= 75%
{3,5} ⇒ {8,11}   s= 30%, c= 75%   {3,8} ⇒ {2}   s= 40%, c= 80%
{3,8} ⇒ {5}   s= 40%, c= 80%      {3,8} ⇒ {11}   s= 40%, c= 80%
{3,11} ⇒ {2}   s= 30%, c= 75%     {3,11} ⇒ {5}   s= 30%, c= 75%
{3,11} ⇒ {8}   s= 40%, c=100%     {3,11} ⇒ {2,8}   s= 30%, c= 75%
{3,11} ⇒ {5,8}   s= 30%, c= 75%   {5,8} ⇒ {2}   s= 30%, c= 75%
{5,8} ⇒ {3}   s= 40%, c=100%      {5,8} ⇒ {11}   s= 30%, c= 75%
{5,8} ⇒ {2,3}   s= 30%, c= 75%    {5,8} ⇒ {3,11}   s= 30%, c= 75%
{5,11} ⇒ {3}   s= 30%, c=100%     {5,11} ⇒ {8}   s= 30%, c=100%
{5,11} ⇒ {3,8}   s= 30%, c=100%   {8,11} ⇒ {2}   s= 30%, c= 75%
{8,11} ⇒ {3}   s= 40%, c=100%     {8,11} ⇒ {5}   s= 30%, c= 75%
{8,11} ⇒ {2,3}   s= 30%, c= 75%   {8,11} ⇒ {3,5}   s= 30%, c= 75%
{2,8,11} ⇒ {3}   s= 30%, c=100%   {3,8,11} ⇒ {2}   s= 30%, c= 75%
{3,8,11} ⇒ {5}   s= 30%, c= 75%   {5,8,11} ⇒ {3}   s= 30%, c=100%
{3,5,11} ⇒ {8}   s= 30%, c=100%   {2,5,8} ⇒ {3}   s= 30%, c=100%
{3,5,8} ⇒ {2}   s= 30%, c= 75%    {3,5,8} ⇒ {11}   s= 30%, c= 75%
{2,3,11} ⇒ {8}   s= 30%, c=100%   {2,3,8} ⇒ {5}   s= 30%, c= 75%
{2,3,8} ⇒ {11}   s= 30%, c= 75%   {2,3,5} ⇒ {8}   s= 30%, c=100%

Number of rules generated:  66
```

V. IMAGE MINING FRAMEWORKS

Early work in image mining has focused on developing a suitable framework to perform the task of image mining. The image database containing raw image data cannot be directly used for mining purposes. Raw image data need to be processed to generate the information that is usable for high level mining modules. An image mining system is often complicated because it employs various approaches and techniques ranging from image retrieval and indexing schemes to data mining and pattern recognition. A good image mining system is expected to provide users with an effective access into the image repository and generation of knowledge and patterns underneath the images. Such a system typically encompasses the following functions: image storage, image processing, feature extraction, image indexing and retrieval, patterns and knowledge discovery.

1. Function-Driven Image Mining Framework Model

Function-driven Image Mining Framework Model is usually organized by modules with different functions. It divides the function model into two modules.

1) Data obtaining, pre-treatment and saving module:

Which is mainly used for image pick-up, original image storage and searching[8].

2) Image mining module:

Which is used for mining image model and meanings.

There are 4 function modules included in this system.

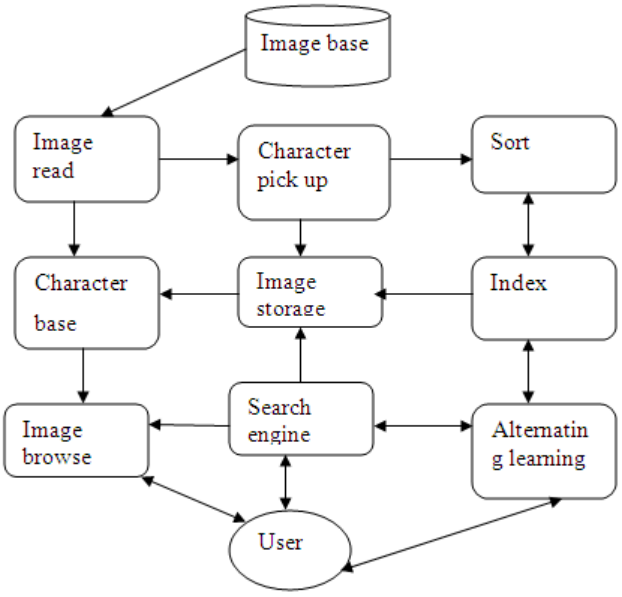


Figure 6: Function-driven image mining framework model

1) Image data acquisition: Get image from multimedia database.
2) Pre processor : Get image character[10].
3) Searches engine: use the image characters for matching inquire.
4) Knowledge discovery module: Mining image.

The Diamond Eye [7] is an image mining system that enables scientists to locate and catalog objects of interest in large image collections. This system employs data mining and machine learning techniques to enable both scientists and remote systems to find, analyze, and catalog spatial objects, such as volcanos and craters, and dynamic events such as eruptions and satellite motion, in large scientific datasets and real-time image streams under varying degrees of a priori knowledge. The architecture of the Diamond Eye system is also based on module functionality.

2. Information Driven Image Mining Framework Model

While the function-driven framework serves the purpose of organizing and clarifying the different roles and tasks to be performed in image mining, it fails to emphasize the different levels of information representation necessary for image data before meaningful mining can take place. Zhang et. al. proposes an information-driven framework that aims to highlight the role of information at various levels of representation (see Figure 7). The framework distinguishes four levels of information given below.This model emphasize different roles of different image arrangement, that incarnate description mechanism of vary arrangement of image data, mark of 4 layers [4].

The Four Information Levels

We will describe the four information levels in our proposed framework.

1. Pixel Level

The Pixel Level is the lowest layer in an image mining system. It consists of raw image information such as image pixels and primitive image features such as color, texture, and edge information[10].

2. Object Level

The focus of the Object level is to identify domain-specific features such as objects and homogeneous regions in the images. An object recognition module consists of four components: model database, feature detector, hypothesizer and hypothesis verifier. The model database contains all the models known to the system. The models contain important features that describe the objects. The detected image primitive features in the Pixel Level are used to help the hypothesizer to assign likelihood to the objects in the image.

The verifier uses the models to verify the hypothesis and refine the object likelihood. The system finally selects the object with the highest likelihood as the correct object.

To improve the accuracy of object recognition, image segmentation is performed on partially recognized image objects rather than randomly segmenting the image. The techniques include: "characteristic maps" to locate a particular known object in images, machine learning techniques to generate recognizers automatically , and use a set of examples already labelled by the domain expert to find common objects in images . Once the objects within an image can be accurately identified, the Object Level is able to deal with queries such as "Retrieve images of round table" and "Retrieve images of birds flying in the blue sky". However, it is unable to answer queries such as "Retrieve all images concerning Graduation ceremony" or "Retrieve all images that depicts a sorrowful mood."

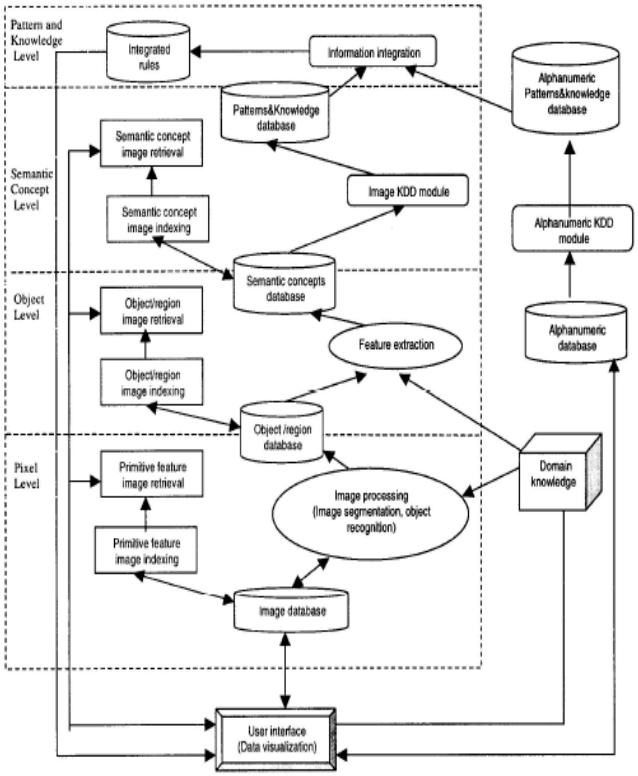


Figure 7: Information driven image mining framework model

3. Semantic Concept Level

While objects are the fundamental building blocks in an image, there is "semantic gap between the Object level and Semantic Concept level. Abstract concepts such as happy, sad, and the scene information are not captured at the Object level. Such information requires domain knowledge as well as state-of-the-art pattern discovery techniques to uncover useful patterns that are able to describe the scenes or the abstract concepts. Common pattern discovery techniques include: **image classification, image clustering, and association rule mining.**

With the Semantic Concept Level, queries involving high-level reasoning about the meaning and purpose of the objects and scene depicted can be answered. Thus, we will able to answer queries such as: "Retrieve the images of a football match" and "Retrieve the images depicting happiness". It would be tempting to stop at this level. However, careful analysis reveals that there is still one vital piece of missing information – that of the domain knowledge external to images. Queries like: "Retrieve all medical images with high chances of blindness within one month", requires linking the medical images with the medical knowledge of chance of blindness within one month. Neither the Pixel level, the Object level, nor the Semantic Concept level is able to support such queries.

4. Pattern and Knowledge Level

At this level, we are concerned with not just the information derivable from images, but also all the domain-related alphanumeric data. The key issue here is the integration of knowledge discovered from the image databases and the alphanumeric databases. A comprehensive image mining system would not only mine useful patterns from large collections of images but also integrate the results with alphanumeric data to mine for further patterns. For example, it is useful to combine heart perfusion images and the associated clinical data to discover rules in high dimensional medical records that may suggest early diagnosis of heart disease. IRIS, an Integrated Retinal Information System, is designed to integrate both patient data and their corresponding retinal images to discover interesting patterns and trends on diabetic retinopathy.

BRAin-Image Database is another image mining system developed to discover associations between structures and functions of human brain . The brain modalities were studied by the image mining process and the brain functions

(deficits/disorders) are obtainable from the patients' relational records. Two kinds of information are used together to perform the functional brain mapping. Discovering knowledge from data stored in alphanumeric databases, such as relational databases, has been the focal point of much work in data mining. However, with advances in secondary storage capacity, coupled with a relatively low storage cost, more and more nonstandard data is being accumulated. One category of "non-standard" data is image data (others include free text, video, sound, etc). There is currently a very substantial collection of image data that can be mined to discover new and valuable knowledge. The central research issue in image mining is how to pre-process image sets so that they can be represented in a form that supports the application of data mining algorithms.

A common representation is that of feature vectors were each image is represented as vector. Typically each vector represents some subset of feature values taken from some global set of features. A trivial example is where images are represented as primitive shape and colour pairs[9].
Thus the global set of tuples might be:

{{blue square}, {red square}, {yellow square}, {blue circle}, {red circle}, {yellow circle}}

which may be used to describe a set of images:

{{blue square}, {red square}, {red circle}}
{{red square}, {yellow square} {blue circle}, {yellow circle}}
{(red box}, {red circle}, {yellow circle}}

However, before this can be done it is first necessary to identify the image objects of interest (i.e. the squares and circles in the above example). A common approach to achieving this is known as "segmentation". Segmentation is the process of finding regions in an image (usually referred to as objects) that share some common attributes (i.e. they are homogenous in some sense)[9].
The process of image segmentation can be helped /enhanced for many applications if there is some application dependent domain knowledge that can be used in the process. In the context of the work described here the author's are interested in MRI "brain scans', and in particularly a specific feature within these scans called the Corpus Callosum.

An example image is given in Figure 8. The Corpus Callosum is of interest to researchers for a number of reasons:

1. The size and shape of the Corpus Callosum are shown to be correlated to sex, age, neuro degenerative diseases (such as epilepsy) and various lateralized behaviour in people.
2. It is conjectured that the size and shape of the Corpus Callosum reflects certain human characteristics (such as a mathematical or musical ability).
3. It is a very distinctive feature in MRI brain scans.

Several studies indicate that the size and shape of the Corpus Callosum in human brains are correlated to sex , age , brain growth and degeneration, handedness and various types of brain dysfunction.

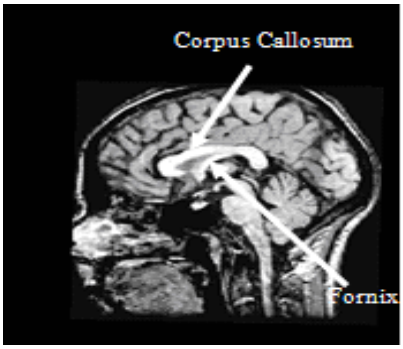


Figure 8: Corpus callosum in a midsagittal brain MRI image.

In order to find such correlations in living brains, Magnetic Resonance Imaging (MRI) is regarded as the best method to obtain cross-sectional area and shape information about the Corpus Callosum. In addition, MRI is fast and safe, without any radiation exposure to the subject such as with x-ray CT. Since manual tracing of Corpus Callosum in MRI data is time consuming, operator dependent, and does not directly give quantitative measures of cross-sectional areas or shape, there is a need for automated and robust methods for localization, delineation and shape description of the Corpus Callosum [9].

The four information levels can be further generalized to two layers: the Pixel Level and the Object Level form the lower layer, while the Semantic Concept Level and the Pattern and Knowledge Level form the higher layer. The lower layer contains raw and extracted image information and mainly deals with images analysis, processing, and recognition. The higher layer deals with high-level image operations such as semantic concept generation and knowledge discovery from image collection. The information in the higher layer is normally more semantically meaningful in contrast to that in the lower layer. It is clear that by proposing a framework based on the information flow, we are able to focus on the critical areas to ensure all the levels can work together seamlessly. In addition, with this framework, it highlights to us that we are still very far from being able to fully discovering useful domain information from images.

3. Knowledge Driven Image Mining Framework Model

Function-driven model is formed from image mining application, and information-driven model is considered from different layer. The essential of image mining is to find knowledge, the above two model doesn't consider the using of mining knowledge, besides, in the whole course, user is on a passive position to receive the mining module and knowledge. Due to the image data itself is an unstructured or semi structure data, so the remounting may happen in image mining, how to mine the maximum knowledge from the mining course. We should know the knowledge user wanted is the knowledge significant[6].

1) **Image choosing:** The aim of image choosing is to confirm the object of image mining which is an original image data in image database as the user's requirement.

2) **Image disposal:** It refers to digital image management and image identification. For example, to remove noises from the image or to proof read the anamorphic image, to recover the low information image[10].

3) **Character pickup:** The character information, such as color , shape, Position are picked up, and stored. Character base is very important because it should support the inquire of image data.

4) **Character choosing Optimize:** The storage of the image character may be overabundant and this factor may affect the operation of the key mining approach so the character choose should be taken before the image mining. If we mark a character with eigenvector, this approach we call dimension decrease. Besides character choosing, sometimes, we should optimize the choosing, including data noise decrease, sequence data dispersion and dispersing data continuum.

5) **Image mining:** Use image mining to mine the data in the image to find related modules. At present, commonly used ways are all from traditional data mining area, such as stat. analyses, associate rule analyses, machine learning. etc.

6) **Explain and comment combining:** In module/knowledge base, it stores the knowledge units which represent image logic concept; we need integrate data to find more potential modules or knowledge. When mining the module, all redundant or useless modules should be removed, the useful modules converse to the knowledge which can be understand by the user.

7) **Image sample training:** Through the image sample training, the validity and veracity can be highly improved[6].

8) **Alternating learning:** Users can learn domain knowledge by system mining, and also can input the domain knowledge to the system, which includes how to split the non figurative knowledge into knowledge unit[6].

9) **Domain knowledge:** In the course of mining, all the former approaches, models or the episteme can be used in discovery of the new system[6].

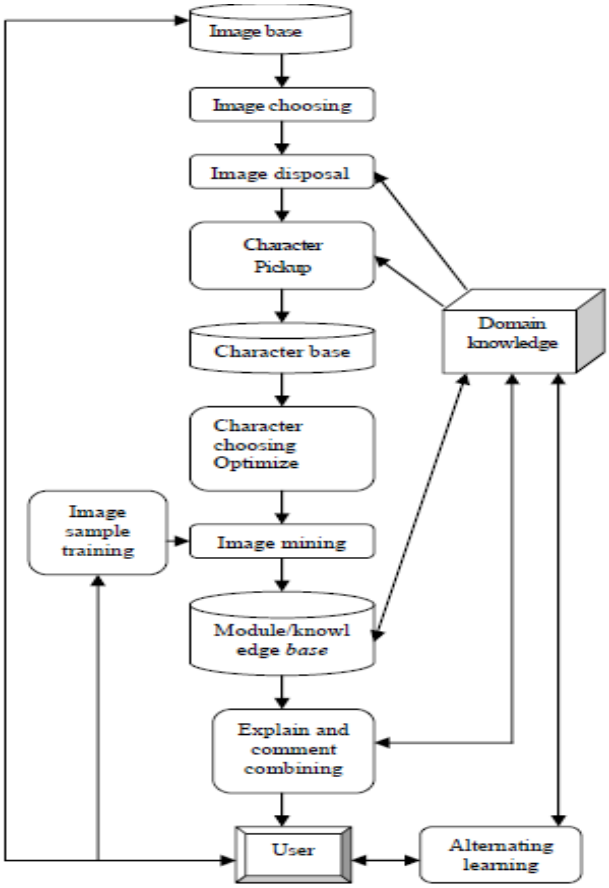


Figure 9: Knowledge-driven image mining framework

VI. CONCLUSIONS

Image mining is the advanced field of Data mining technique and it has a great challenge to solve the problems of various systems. The main objective of the image mining is to remove the data loss and extracting the meaningful information to the human expected needs. It retrieves the most matching images from the collection of the images, with respect to the query image. The framework models represents the first step towards capturing the different levels of information present in image data and addressing the question of what are the issues and challenges of discovering useful patterns/knowledge from each level.

References

[1] J. Han and M. Kamber. Data Mining: Concepts and Techniques, Morgan Kaufmann, USA, 2001.
[2] Margaret H. Dunham, "Data mining: Introductionary and Advanced Topics", Southern Methodist University.
[3] R. Gonzalez and R. Woods. Digital Image Processing, Addison-Wesley Publications Co, March 1992.
[4] Ji Zhang, Wynne Hsu, Mong Li Lee. An Information-driven Framework for Image Mining, in Proceedings of 12th International Conference on Database and Expert Systems Applications (DEXA), Munich, Germany, 2001.
[5] Hilal M. Yousif, Abdul- Rahman Al-Hussaini, Mohammad A. Al-Hamami. Using Image Mining to Discover Association Rules between Image Objects.
[6] Yu Changjin, Xia Hongxia. The Investigation of Image Mining Framework, WUHAN University of Technology Wuhan, China.
[7] Joseph Roden, Michael Burl and Charless Fowlkes. The Diamond Eye Image Mining System, Jet Propulsion Laboratory.
[8] J.Zhang, W.Hsu and M.L.Lee, "Image Mining: Issues, Frameworks and Techniques", Proc. of Second International Workshop on Multimedia Data Mining (MDM/KDD'2001), San Francisco, CA, USA, August, 2001.
[9] Ashraf Elsayed, Frans Coenen, Marta García-Fiñana and Vanessa Sluming, Segmentation for Medical Image Mining: A Technical Report, The University of Liverpool, Liverpool L69 3BX, UK.
[10] Dr.V.Mohan, A.Kannan, "Color Image Classification and Retrieval using Image mining Techniques", International Journal of Engineering Science and Technology Vol. 2(5), 2010, 1014-1020.
[11] M. Antonie, O.R. Zaiane, A. Coman. Application of Data Mining Techniques for Medical Image Classification. In Proceedings of the Second International Workshop on Multimedia Data Mining (MDM/KDD'2001), San Francisco, CA, USA, August, 2001.

A Study On An Interest & Attitude Of The Student Of Khargone Taluka's Urban & Rural Higher Secondary Schools In English Curriculum

[1]Dr. Shri Krishna Mishra (Principal) [2]Mr. Badri Yadav (Asst. Professor)
Mobile- 09669696249, 09753222771
Shri Kanwartara Institute for Teacher's Training,
Shri Nagar Colony, Mandleshwar, Dist.Khargone (M.P.) 451221

ABSTRACT

The main aim of the present research paper to do a study on an interest and attitude of the students of Khargone taluka's urban and rural high secondary schools in English curriculum. In the present research paper the investigator tried the compare the attitude of rural and urban students toward English curriculum. But this results are not the last truth. I hoped that in the suggestions and the results will be helpful for the development of the situation of the English curriculum in the schools in present time.

Key words: Interest & attitude of the student, A study on an interest & attitude, A study of Urban & Rural.

INTRODUCTION-

Education takes the personality of person on great height. In 1964-66 education management told "In present time India's future is developing in the classrooms. For the development of the students we must give proper training to the trainee so with help of it they can do welfare for their country and also for their society. Because the development of the society is depended on the education. And this type of education can be given in the schools. So it is rightly said that, in the schools undeveloped live person is turned in to the well developed in the right personality.

The best tool of the completing the objectives of education is curriculum. Which the student get from the school. In the Indian constitution to get the expected objectives the curriculum is constructed according to the need and interest of the student. According to STANG, "if we have interest towards it, we appreciate it. If have no interest towards it we run away from it. Now we focused on the relation of language with the education. Man is a social animal, he always exhibit his heart's thoughts to others and always ready to know about other's thoughts of the hearts. Language is the medium by which the person do conversation of his thoughts and emotions with his speech. Language is the symbol by it we can do exchange. When we come in the contact with any thing there immerges lots of rays and touches to our senses, we fill some thing extra. Then it reach to the mind by muscles, then it comes to the our tongue as speech. This is the proper present sound symbol is the language, and the education of the proper use of the language is the right education. Through this type of language there is the development of education.

So we can say that education & language are two sides of the coin. If the language is the principal, the education is the invention. If the language is the knowledge, the education is the art. If the language is the permanent, the education is changeable. If the language is the word, the education is the sentence. In this way both are interrelated with each other. So it is important for us to give the proper knowledge of the language.

But today we can see that students & teachers are underestimate the English language. But the English is the one language which connect all the nations of the world.

Place of the English language in India as a foreign language.

There are different nations in the world which have their own cultures. They have their own languages. In ancient time there is no any specific language for communication between two or more nations. Many peoples went to different countries for business they use some common languages for exchange. After long time English language become one largely used language. English language is used in our country from the time of the Britishers, they already go back in their country but left the English language for us. Effect of the foreign countries now our Indian people use it as a status symbol.

In our nation English is accepted as the foreign language. But the public used it as a fashion. Speaking English is the new trend for the people. They think that speaking English is the impression of good personality. English language possesses one of the great place in the India. It is not the mother tongue but it is the language of communication for foreigners. Now English language become the main useful language for us. We can see lots of use of the English around us, we can see big hording, news papers, notice boards etc in English. Now it become local language for us.

Justification of the problem.

There are some points for justification of the problem.

- According to managements :- Curriculum is not stoppable thing, it always changing matter according to social need.
- According to publishers – they can selecting the subject matter it become useful. So the attitude of the students can get adjustment towards the curriculum.
- According to the teachers :- to find out which type of the students attitude towards the English curriculum there can be proper arrangement of the methods and techniques.
- According to the students :- On the basis of the English student's attitude, importance and usefulness in the subject matter of the English curriculum can be developed.
- According to the investigators :- Because of this the investigators can easily make tools for measurement of attitude.

Keeping all these points in the mind we can say there are some mistakes in the English curriculum. Which is in the method of the teacher or in the students? But it Cannot be shown by any logic. For this the investigator must do proper and complete Study.

Statement of the problem

'A study on an interest & attitude of the students of Khargone taluka's urban & rural high secondary schools in English Curriculum'

Intentions of the Research

Before researching on any problem it is needed to determine the objectives of the problem. Because without this determination we cannot achieve the goal. Thus, there are some objectives.

- To identify the interest of the students in the English language.
- To find out the attitude of the students towards the English language.
- To compare the attitudes of male & female students towards English curriculum.
- To compare the attitudes of rural & urban students towards English curriculum.
- To find out the learning difficulties of students.
- To give suggestions for improvement of the English curriculum.

Criteria of the Research

Because of the time limitation & available sources we limit our study to one small part of the taluka.

- It is limited only for the 9^{th} standard.
- It is limited in the Khargone taluka's four higher secondary schools.
- It is limited in the two schools of the rural and two school of urban area.
- There were only 16 male & 16 female students are covered under the research.

Research Hypothesis.

Investigator is used zero hypothesis in the present research.

- The identification of the interest of the students in the English language will be taken care of ……..
- The positive attitude of the students towards the English language will be useful.
- The comparison of the attitudes of male & female students towards English curriculum will be helpful for encouragement.
- The comparison of the attitudes of rural & urban students towards English curriculum will be useful for encouragement.
- The identification of the learning difficulties of students will be useful.
- The suggestions for improvement of the English curriculum will be useful to improvement of attitude of the students towards the English curriculum.

Researches done in India.

In the India there are many research were done during fifty years.

(A) **H.L. SHRIVASTRA** – (1950) :- He had done a study on attitude of the teachers towards their profession, he had found that attitude of the teachers towards their profession is important.

(B) **NARAYAN SINGH** – (1986) :- He had done a study of the teacher towards History education. He had found that the teacher of history thought that their place in the society is title.

Researches done in abroad.

Present section indicates some researches were done in the abroad there are as under.

(A) **HARPER MAHODAY** – (1927) :- He had done a study on attitude of the U.S. teachers. He had found that teachers were never clear on the problems of the society. They always followed their old methods of teaching. They heisted to accept the new methods.

(B) **NANDAN TARESA** – (1994) :- He had done a study on the attitude of the teachers of the science. He used questioner for the research, than he had found that most of the schoolyard teachers attitudes are positive.

Figure of the sample

The area of the present research paper is limited because of the time limitation & available sources. We limited our study to one small block of the khargone taluka.

- It is limited only for the 9th standard.
- It is limited in the khargone taluka's four higher secondary school.
- It is limited in the two schools of the rural and two school of urban area.
- There were only 16 male & 16 female students are covered under the research.

Keeping all the points in the mind there were 32 male & 32 female students selected as the sample.

S.No.	Name of the school	STD.	Total
URBAN			
1	Shri Krishna Colony Khargone	9	16
2	Vidya Vijay Mandir, Khargone	9	16
RURAL			
1	Govt. Boys High School Khargone	9	16
2	Govt. Girls High School Khargone	9	16
	Total Students.	-	**64**

Collection of figures of the research paper

The research paper is moving ached towards it's target from the completing its steps and stages. So, the investigator have starting the collection of the data after selecting of the sample.

This work has been completed with the investigators valuation. After completing the test all the test all the data are collected.

Data collection

After collecting the data the main work his to do the data analysis. Afte data collection the investigator had done the data analysis with the help of the attitude of the students towards the English curriculum. After data analysis the investigator got the proper way towards the solution of the problem.

Use of the numerical technique.

In the present research paper the investigator had used some numerical techniques, which are shown as under.

Mean – Data collection.

"Sum of a set of data divided by the number of subjects in the set is called mean".

Mean is helpful to decide that which type of the mean so any group is and what kind of its result. To get the mean that is the formulas,

$$X = A.M. + \frac{E\ Fd}{N} \times i$$

Where A.M. = assumed mean
D = deviation from the assumed mean
I=class interval
N=total frequencies
standard deviation.

"Standard deviation is the square root of sum total of square of the deviation divided by the sum total."

Percentage.

Formation of the percentages.,

$$\text{Percentage} = \frac{\text{Integer} \times 100}{\text{Total number}}$$

Statistical analysis of data collection :-

After completing the work of the data collection the investigator concentrated on the data analysis. The objectives of the result of the problems. For the data analysis he used the nominal technique. In it he used the mean, standard deviation.

Percentage and 't' value. The use of the percentage for the measurement of the Attitude of the students towards English curriculum and learning difficulties. According to this 40% to 60% were average, above than 60% were positive and less than 40% negative attitude. Than the use of the mean, standard deviation and 't' value for the comparisons of the attitude of the students towards English curriculum.

1 Attitude of the students towards English curriculum. This is shown in the table no. 1

Table No. – 1
Attitude of the students towards English curriculum.

Students	N	No. of statements	Total Marks	Average marks	Per (%)
Male	32	41	205	146.50	71.46
Female	32	41	205	141.91	69.22
Total	**64**	**41**	**205**	**144.20**	**70.34**

Percentage of attitude of the students towards the English curriculum

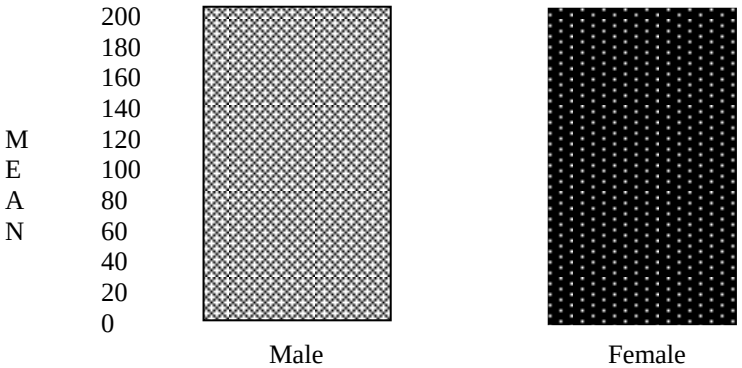
There was taken a test of the attitude towards English curriculum of all the students. We can see from the table no. 1 that the attitude towards English curriculum was 70.34% which is above 60%. So, the attitude of the rural male and female student towards English curriculum was positive. The attitude of the rural male and female students towards English curriculum were in squencly 71.46% and 69.22% which is above 60%. So, the attitude towards English curriculum was positive. We can see from the table no. 1 that the attitude of male students was more than the female students towards English curriculum. For the proof of it the investigator had calculated the Mean, S.D. 't' value which was shown in the table no. 1

2 comparison of attitude of male & female students towards English curriculum.

Table No. – 2
The mean, standard deviation, 't' value of the attitude of male & female

Sex	N	Mean	S.D.	't' value
Male	32	146.50	6.91	2.32
Female	32	141.91	8.16	

Df = 32, 0.05 base 't' value 2.04



Picture no. – 2
Comparison of mode of attitude of male & female students

The investigator had done the comparison of attitude of male & female students towards English curriculum. The investigator found the distance in the both 't' values. The investigator found 2.32 't' value of the attitude which was above the base value. We can see that the attitude of male students was more than the female students towards English curriculum. This was shown in the table no. 2

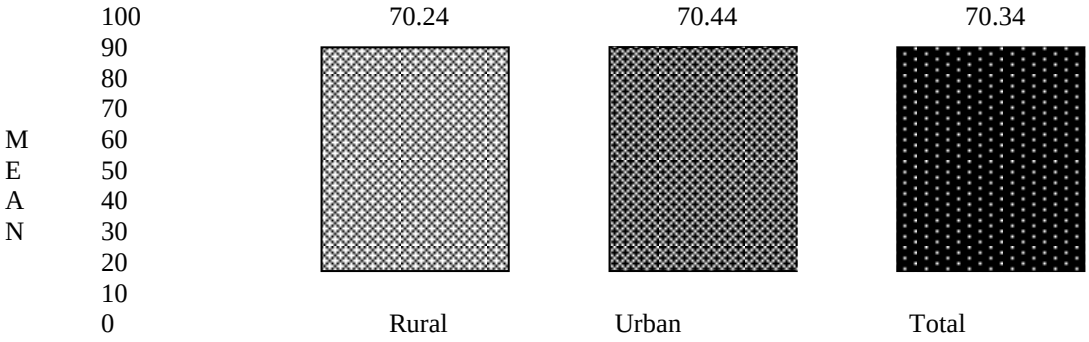
3, Attitude of rural & urban students towards English curriculum.

To get above objective the investigator had collected the figures from the attitude valuation. Than he founded the percentages of the figures, which was shown in the table no. 3.

Table No. – 3

Attitude of rural & urban students towards English curriculum

Area	N	No. of statements	Total marks	Average marks	Per (%)
Rural	32	41	205	144.00	70.24
Urban	32	41	205	144.41	70.44
Total	64	41	205	142.20	70.34



Picture no. – 3

Percentages of attitude of rural & urban students towards English curriculum.

There was the research on the attitude of rural & urban students towards English curriculum. The attitude of the rural & urban students towards English curriculum were in sequencly 70.24% and 70.44% which was above 60%. The attitude of the rural urban and rural students towards English curriculum were positive. There no difference between the attitude of the rural urban and rural students towards English curriculum

4. comparison of attitudes of rural & urban students towards English curriculum

To get above objective the investigator had collected the figures of attitude of rural & urban students towards English curriculum on the basis of them he counted the mean, S.D., and 't' value, which was shown in the table no. 4

The investigator had done the comparison of attitude of rural & urban students towards English curriculum. The investigator did not get distance between 't' value of attitudes of rural & urban students towards English curriculum. That matter was shown in the picture no. 4 so, we can say that the attitudes of rural & urban students towards English curriculum were same.

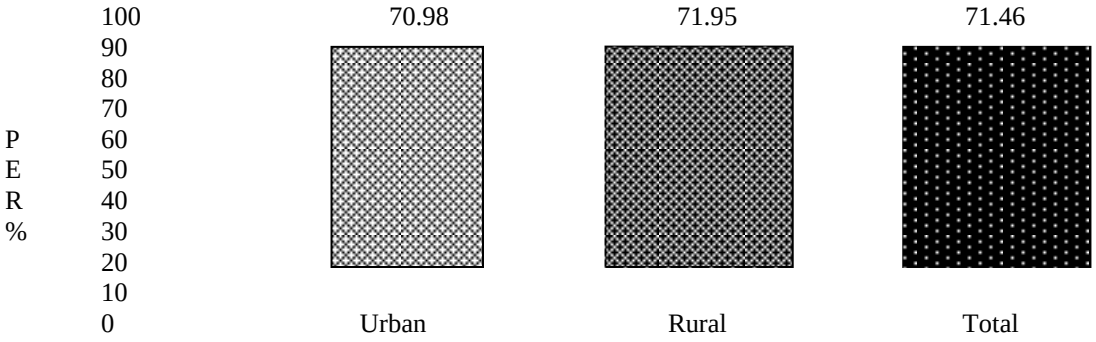
5 attitudes of rural & urban male students towards English curriculum.

To achieve above goal the investigator gave the questioner to the rural & urban students, he got the data than he found the percentages of them. Which was shown in the table no. 5.

Table No. – 5

Attitude of rural & urban male students towards English curriculum

Area	N	No. of statements	Total marks	Average marks	Per (%)
Urban statements	16	41	205	145.50	70.98
Rural statements	16	41	205	147.50	71.95
Total	30	41	205	146.5	71.46



Picture no. – 5

Percentage of rural & urban male students.

There was the test of rural & urban male students attitudes English curriculum.

The attitudes of the rural urban and rural students attitudes English curriculum were in

Sequencly 70.24% and 70.44% which was above 60%. Were in squencly 71.95% and 70.98% which was above 60%. So the attitude of the rural and urban students attitudes English curriculum were positive. There is nothing distance between the attitude of the rural and urban students attitudes English curriculum.

6. comparison of attitudes of rural & urban male students towards English curriculum.

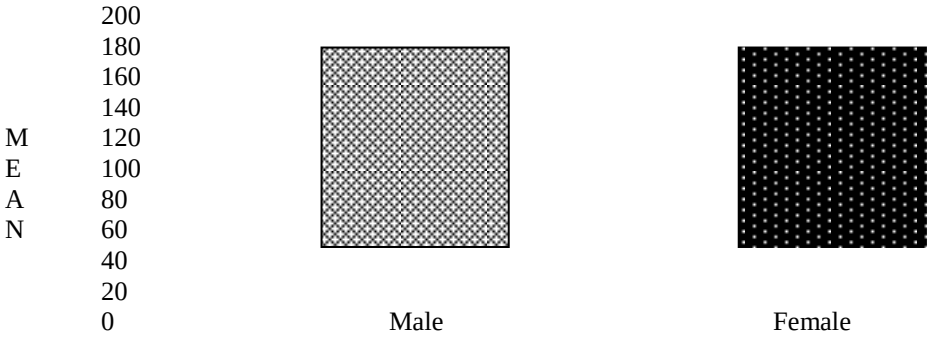
There was the test of rural & urban male students towards English curriculum for the calculation of the mean, S.D. and 't' value. Which was shown in the table no. 6.

Table No. – 6

The mean, standard deviation, 't' value of the rural & urban male students

Area	N	Mean	S.D.	't' value
Urban	16	145.50	7.05	0.82
Rural	16	147.50	6.76	

Df = 32, 0.05 base 't' value 2.04



Picture no. – 6.

Comparison of mode of rural & urban male students

There was the comparison of attitude of the rural & urban male students towards English curriculum. The investigator done the comparison of attitudes of rural & urban male students towards English curriculum. The 't' value was 0.82 which was above the base of the value 0.05. There was no distance of attitude of the rural & urban male students towards English curriculum. This shown in the picture no. 3. The comparison of mode of the rural & urban male students toward English curriculum was done by the investigator. So, we can say that the attitude of the rural & urban male students towards English curriculum were same.

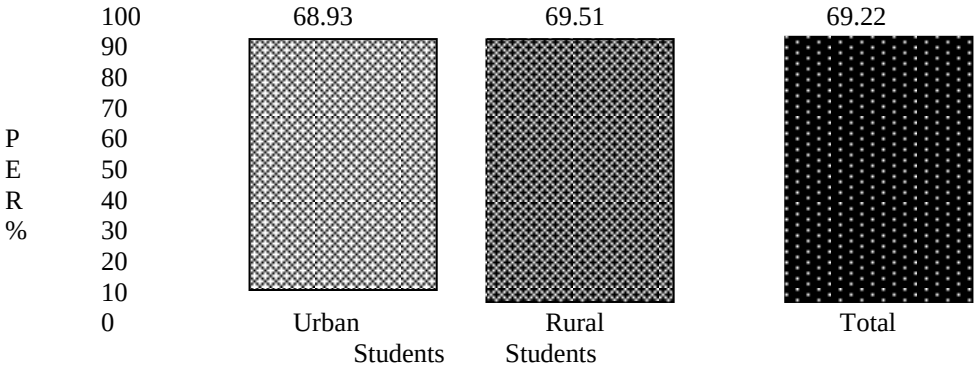
7. Attitudes of rural & urban female students towards English curriculum.

To achieve above goal the investigator gave the questioner to the rural & urban students, he got the data than he found the percentages of them. Which was shown in table no. 7.

Table No. – 7

Attitude of rural & urban female students towards English curriculum

Area	N	No. of statements	Total marks	Average marks	Per (%)
Urban students	16	41	205	141.31	68.93
Rural students	16	41	205	142.50	69.51
Total	32	41	205	141.91	69.22



Picture no. – 7
Percentage of rural & urban female students.

There was the chaking of the attitudes of rural & urban female students towards English curriculum. The investigator found that the attitude of the rural & urban female students towards English curriculum. Sequencly 69.51% and 68.93% which was above 60%. So. The investigator found that the attitude of the rural & urban female students towards English curriculum were positive.

On the base of the percentage the investigator found that the attitude of the rural & urban female students towards English curriculum were same which was shown in table no. 7.

8 comparison of attitude of rural & urban female students towards English curriculum.

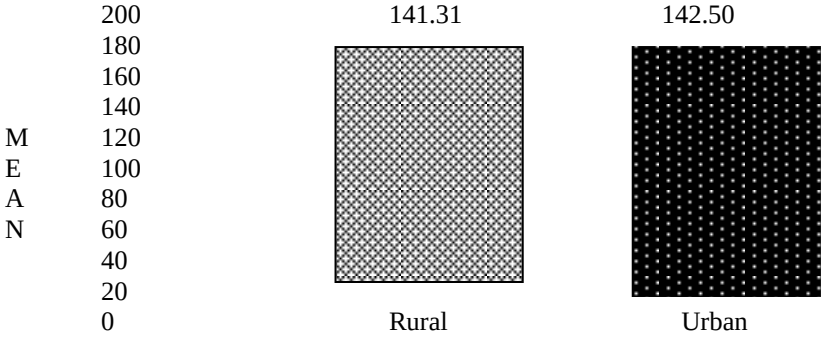
The investigator found the mean, S.D. and 't' value of the attitude of the rural & urban female students towards English curriculum.

Table No. – 8

The mean, standard deviation, 't' value of rural & urban female students

Area	M	Mean	S.D.	't' value
Rural students	16	141.31	9.02	0.36
Urban students	16	142.50	9.87	

Df = 16, 0.05 base 't' value 2.04



Picture no. – 8
Comparison of mode of rural & urban female students.

The investigator had done the comparison of attitude of rural & urban female students. With the help of that figures he got the 't' value. The investigator had found 't' value 0.36 which was lower than the base value. So, there was no difference between attitudes of rural & urban female students towards English curriculum. Which was shown in the picture no. 6 so, we can say that the attitudes of rural & urban female students towards English curriculum were same.

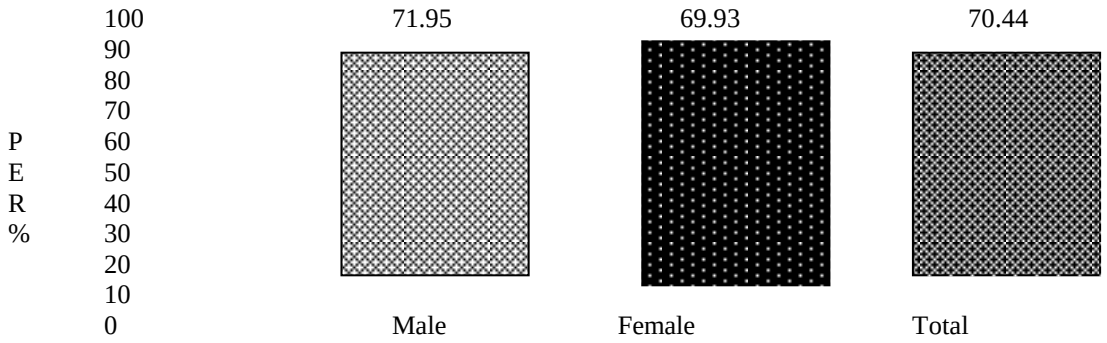
9 attitudes of rural male female students towards English curriculum.

To achieve above goal the investigator gave the questioner to the rural & urban female students, he got the data than he found the percentages of them. Which was shown in the table no. 9.

Table No. – 9

Attitude of rural male female students towards English curriculum

Students	N	No. of statements	Total marks	Average marks	Per (%)
Male	16	41	205	147.50	71.95
Female	16	41	205	141.31	68.93
Total	32	41	205	144.41	70.44



Picture no. – 9

Percentages of rural male & female students.

The investigator had check the attitudes of rural male & female students towards English curriculum. The investigator found that the attitudes of the rural male female students towards English curriculum were in sequencly 71.95% and 69.93% which was above 60%. So, the investigator found that the attitudes of the rural male female students towards English curriculum were positive. On the base of the percentage the investigator found that the Attitudes of the rural male female students towards English curriculum were same. Which was shown as under.

3 comparison of attitudes of rural male female students towards English curriculum

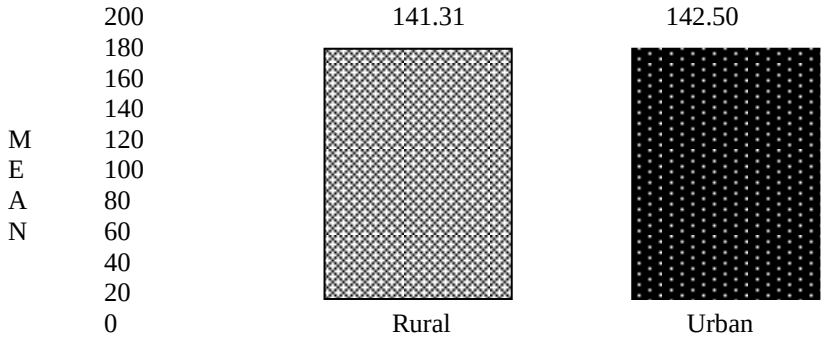
To achieve above goal the investigator gave the questioner to the rural male and female students, he got the data than he found the mean, S.D., and 't' value of rural male & female students. Which was shown in table no. 10

Table No. – 10

The mean, S.D., 't' value of rural male & female students

Students	N	Mean	S.D.	't' value
Male	16	147.50	7.05	2.17
Female	16	141.31	9.02	

Df = 16, 0.05 base 't' value 2.12



Picture no. – 10
Comparison of mode of Attitudes of rural & urban female students.

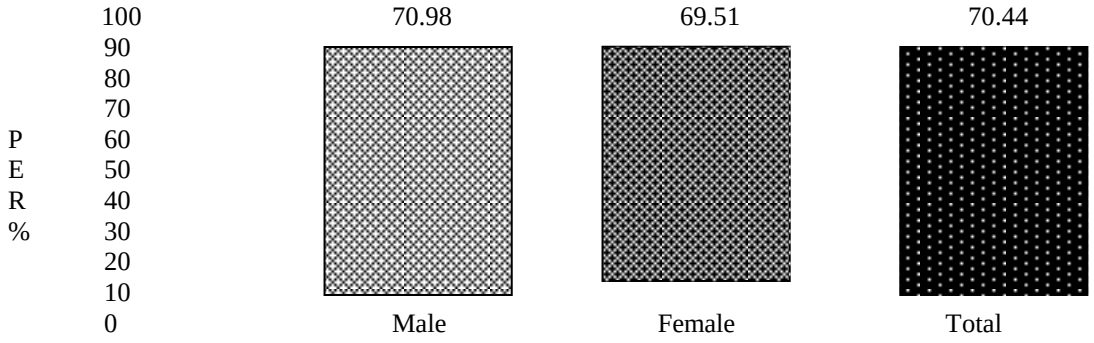
The investigator had compare the attitude of rural male female students towards English curriculum. The investigator found that the attitude of the rural male & female students towards English curriculum were quite differ. So, the investigator found that the Attitudes of rural male female students towards English curriculum were positive. On the base of the percentage the investigator found that the Attitudes of rural male female students towards English curriculum were more. Which was shown as under.

11. attitudes of urban students towards English curriculum.

To achieve above goal the investigator gave the questioner to the urban male and female students, he got the data than he found the mean, S.D., and 't' value of rural male & female students. Which was shown in table no.11 **Table No. – 11**

Attitude of urban students towards English curriculum

Students	N	No. of statements	Total marks	Average marks	Per (%)
Male	16	41	205	142.50	70.98
Female	16	41	205	142.50	69.51
Total	32	41	205	144.0	70.24



Picture no. – 11
Attitudes of urban students towards in percentage.

The investigator had check the attitudes of rural male & female students towards English curriculum. The investigator found that the attitudes of urban male & female students towards English curriculum were in sequencly 70.98% and 69.51% which was above 60%. So, the investigator found that the attitudes of the urban male & female students towards English curriculum were positive. On the base of the percentage the investigator found that the Attitudes of urban male & female students towards English curriculum were same. Which was shown as under.

12 comparison of attitudes of urban male female students towards English curriculum

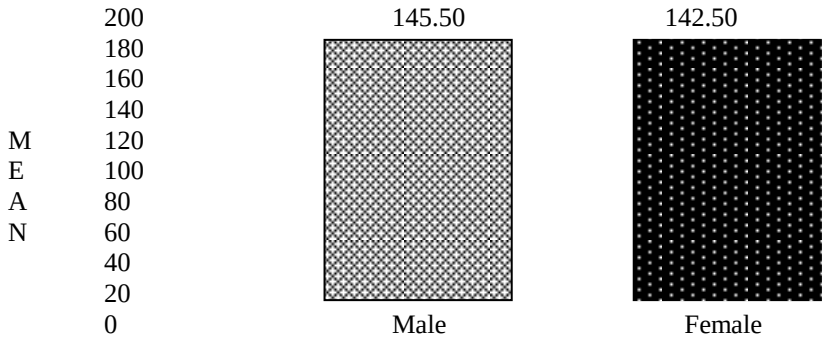
To achieve above goal the investigator gave the questionary to the urban male and female the students, he got the data than he found the Mean, S.D., and 't' value of rural male & female students. Which was shown in table no. 12.

Table No. – 12

Comparison of mode, S.D. and value of students

Students	N	Mean	S.D.	't' value
Male	16	145.50	6.76	1.01
Female	16	142.50	9.87	

Df = 16, 0.05 base 't' value 2.12



Picture no. – 12
Comparison of mode of urban Male & female students.

The investigator had compare the attitude of urban male & female students towards English curriculum. The investigator found that the attitudes of urban male & female students towards English curriculum were no difference. So, the investigator found that the Attitudes of rural male & female students towards English curriculum were positive On the base of the percentage the investigator found that the Attitudes of urban male & female students towards.

13 Find out the learning difficulties of students.

To find out the learning difficulties of students the investigator used the questioner. He gave it to the students and asked them to answer properly. Than he collected all the data calculated the percentages of the data. Which was shown in table no. 13.

Table No. – 13
The learning difficulties of students

Question No.	Answer in Yes		Answer in No	
	Figure	Per%	Figure	Per%
1	54	84.38	10	15.63
2	57	89.06	07	10.94
3	38	59.38	26	40.63
4	50	78.13	14	21.88
5	47	73.44	17	26.56
6	55	85.94	09	14.06
7	44	68.75	20	31.25
8	43	67.19	21	32.81
9	46	71.88	18	28.13

- We can see that from the above table no. 13. that 54 students from the 64 students answers were ‘yes’ in the question about hardness of language. It means 84.38% faced difficulties the language.
- 57 students give ‘yes’ answer to the question about the clarity of the curriculum. It means 89.06% students faced the difficulties because of it.
- There were 38 students agree that the teacher could not teach properly. Means there was 59.38% difficulties during English language learning because of the teachers.
- There were 50 students knowledge weak in the English language. So, 78.13% difficulties came from it.
- There were 47 students who were not agree with the use of the teaching aids, during the teaching. There are 73.10% difficulties from the avidness of the proper use of the teaching aids.
- There were 55 students who think that they could not get opportunity to speak in English. Means 85.94% difficulties come from it.
- 68.75% difficulties of he learning of the English language come from the no encouragement of the teachers.
- There are 43 students agree to get difficulties in the learning English.
- 69% difficulties come because of the hardness of the English language.

Findings-

(a) The attitude of the students towards English curriculum.
The attitude of the students towards English curriculum was suitable.
The attitude of the all students was 70.34%. So, the attitude of the students towards English curriculum positive.

(b) Comparison of the attitude of the male and female students towards English curriculum.
The attitude of the male students towards English curriculum was 71.46% and the attitude of the male students towards English curriculum was 69.22% After the comparison, we can see that the attitude of the male students was more then the female students attitude towards English curriculum.

(c) Result of the male and female students on the bases of the mean, S.D. and 't' value.
The mode of the attitude of the male students towards English curriculum was 146.50 and the mode of the attitude of the female students towards English curriculum is 141.91. There was the remarkable distance between the mode of both.
While the 't' value of the attitude of the male and female students towards English curriculum was 2.32 which was up from the base of the value 0.05.

(b) The attitude of the rural and urban students towards English curriculum.
The attitude of the rural and urban students towards English curriculum were in sequencly 70.44% and 70.24. There were no difference between the attitude of the rural and urban students towards English curriculum.

(e) Results of the rural and urban students towards English curriculum on the bases of the mean, S.D. and 't' value.
The mode of the rural and urban students towards English curriculum were in sequencly 147.50 and 145.50. There was no difference between them 0 the bases of the mean.
There was no difference between them 0 the bases of the 't' value. Because the 't' value was 0.21 which is low from the value base 0.05.

(f) Results of attitude of the rural students on the bases of the mean S.D. and 't' value.
The mode of the rural male and female students towards English curriculum were in sequencly 147.50 and 141.31.
The comparison of the 't' value of the rural male and female students towards English curriculum showed that there was different in the attitude. Because 't' value was 2.17 was more then value base 0.05.

(g) Results of attitude of the urban students on the bases of the Mean, S.D. and 't' value.
The mode of the urban male and female students towards English curriculum were in sequencly 145.50 and 142.50.
The comparison of the 't' value of the urban male and female students towards English curriculum showed that there was no different in the attitude. Because 't' value was 1.01 was less than value base 0.05.

(H) difficulties in the way of the students during the learning of the English curriculum.
There are six types of difficulties faced by the students. More than 89.06% students face difficulty in reading of the text book.
More than 59.38% students face difficulty because of the teaching methods of the teachers.

SUGGESTION :-

Suggestions for the teachers

(A) The teacher should use more and more new educational techniques and methods during teaching.
(B) The teacher should use more and more helping tools to clear the subject matter.
(c) The teacher should use teach the hard words very easily to the students during the classroom teaching.
(D) The behavior of the teacher should be equal toward the all students.

Suggestions for the students.

The students should take more and more interest in the learning of the English language.

Suggestions for the administrators

(A) The administrators the curriculum should do the evaluation on the curriculum time to time. Than send their suggestions to the curriculum committee.
(B) They should the research on according to need and attitude of the students.
(C) They should Do the remedial work on the problems of the curriculum.

Conclusion. -

The main aim of the present research paper to do a study on an interest and attitude of the students of Khargone taluka's urban and rural high secondary schools in English curriculum.

In the present research paper the investigator tried the compare the attitude of rural and urban students toward English curriculum. But this results are not the last truth.

I hoped that in the suggestions and the results will be helpful for the development of the situation of the English curriculum in the schools in present time.

References

[1]. Brahmbhatt, J.C., A study of pupa ration of language programmed in English.
[2]. Data, C., effect of maximizing control clues. A pragmatic study, Ph.D. FLT, CIEFL, 1985.
[3]. "Agarwal, M., A factorial study of attitude of students towards some social problem, Ph.D Edu. Jammu U., 1984
[4]. Setee, E.D and Ross, E.L, (1987)
[5]. "A case study in applied education in rural India" Community development journal, 22 (2): 120-129, oxford university press.
[6]. Madhya Pradesh Rajya shiksha Kendra 2007, state curriculum framework, Bhopal.
[7]. NCERT, 2005, National curriculum framework-2005, New Delhi India.
[8]. Jai prakash, A comparative study of urban, rural and tribal higher secondary students of Madhya Pradesh with the reference to their general mental ability and interest pattern, Dept. of pay, Sag. U., 1972
[9]. (ICSSR financed)-2
[10]. Joseph, K.S., Enrolling a strategy for teaching English grammar at high school level, Ph.D. Edu MSU 1983-4
[11]. Kamlesh, A., A comparative study of self concept, adjustment, interest and motivation among the scheduled caste and non schedule caste students. MD. Edu., Kan. O. 1981-3.

Improving Detection Performance of Cognitive Femtocell Networks

Ms.Madhura Deshpande[1], Dr.S.D.Markande[2]
[1](Electronics & Telecommunication Department, SKNCOE, Pune, India)
[2](Principal, NBNSSOE, Pune, India)

Abstract :
Femtocell is envisioned as a highly promising solution for indoor wireless communications. The spectrum allocated to femtocells is traditionally from the same licensed spectrum bands of macrocells. The capacity of femtocell networks is highly limited due to finite number of licensed spectrum bands and the interference with macrocells and other femtocells. A radically new communication paradigm is proposed by incorporating cognitive radio in femtocell networks. The cognitive radio enabled femtocells are able to access spectrum bands not only from macrocells but also from other licensed systems (e.g. TV), provided the interference from femtocells to the existing systems is not harmful. It results in more channel opportunities. Thus, the co-channel interference can be greatly reduced and the network capacity can be significantly improved. Further, detection performance can be improved by decreasing the collision probability with the help of double threshold energy detection.

Keywords: Co-channel interference, Cognitive radio, Collision probability, Double threshold energy detection, Femtocell network, IEEE 802.22, Licensed user.

I INTRODUCTION

In mobile wireless networks, the demand for higher data rates and lower power consumptions is continuously increasing, while the capacity provided by the existing macro cell networks is limited.

The Studies on wireless usage show that more than 50 percent of all voice calls and more than 70 percent of data traffic originate from indoors. Voice networks are engineered to tolerate low signal quality, since the required data rate for voice signals is very low, on the order of 10 kb/s or less. Data networks, on the other hand, require much higher signal quality in order to provide the multimegabit per second data rates. For indoor devices, particularly at the higher carrier frequencies, high data rates are very difficult to achieve. This motivates for the femtocell approach [1].

Femtocells, also called home base stations (BSs), are short-range low-cost low-power BSs installed by the consumer for better indoor voice and data reception. The user-installed device communicates with the cellular network over a broadband connection such as digital subscriber line (DSL), cable modem, or a separate radio frequency (RF) backhaul channel. While conventional approaches require dual-mode handsets to deliver both in-home and mobile services, an in-home femtocell deployment promises fixed mobile convergence with *existing* handsets. Compared to other techniques for increasing system capacity, such as distributed antenna systems and microcells, the key advantage of femtocells is that there is very little upfront cost to the service provider [1].

In macro cell networks, traditional spectrum allocation methods are mostly based on coloring methods where no neighboring cells can use the same spectrum at the same time. Since the number of femtocells could be much higher than the number of macro cells in a certain area, this kind of spectrum allocation requires more spectrum bands. This will lead to inefficient and unfair spectrum utilization. This motivates the study to improve the spectrum utilization and cell capacity.

In the meantime, it has been shown that spectrum is not efficiently used by licensed (primary) users/systems according to the exclusive spectrum allocation regulation. In recent years, cognitive radio (CR) technology has been developed to allow unlicensed users to exploit spectrum opportunities from primary systems to enhance the spectrum utilization greatly [1]. It then inspires to incorporate the CR technology into femtocell networks, where the CR-enabled femtocell users (FUs) and FBSs can identify and utilize the spectrum opportunities from the licensed systems such as macro cell networks and TV broadcast systems.

II. Related work

According to Federal Communications Commission (FCC), temporal and geographical variations in the utilization of the assigned spectrum range from 15% to 85%. Although the fixed spectrum assignment policy generally served well in the past, there is a dramatic increase in the access to the limited spectrum for mobile services in the recent years. This increase is straining the effectiveness of the traditional spectrum policies.

The limited available spectrum and the inefficiency in the spectrum usage necessitate a new communication paradigm to exploit the existing wireless spectrum opportunistically. *Dynamic spectrum access* (DSA) is proposed wherein unlicensed (secondary) systems are allowed to opportunistically utilize the unused licensed (primary) bands, commonly referred to as “spectrum holes”, without interfering with the existing users [2].

2.1 Cognitive Radio Technology

A "Cognitive Radio" is a radio that can change its transmitter parameters based on interaction with the environment in which it operates [2]. The ultimate objective of the cognitive radio is to obtain the best available spectrum through cognitive capability and reconfigurability. The Fig.1 shows cognitive cycle. The steps in cognitive cycle are as follows :–

1. *Spectrum Sensing*: A cognitive radio monitors the available spectrum bands, captures their information and then detects the spectrum holes.
2. *Spectrum Analysis*: The characteristics of the spectrum holes that are detected through spectrum sensing are estimated.
3. *Spectrum Decision*: A cognitive radio determines the data rate, the transmission mode and the bandwidth of the transmission. Then, the appropriate spectrum band is chosen according to the spectrum characteristics and user requirements.

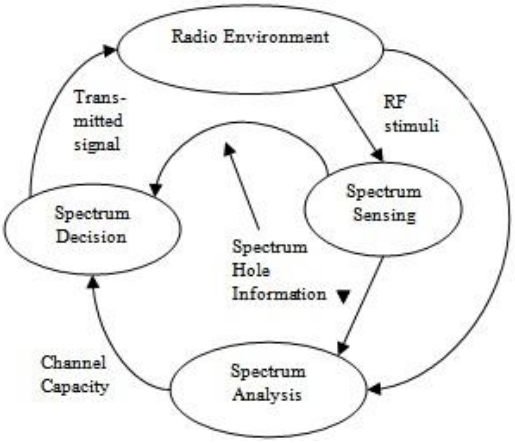


Figure 1: Cognitive Cycle [2]

The two main characteristics of cognitive radio are:–

- *Cognitive Capability*: Cognitive capability refers to the ability of the radio technology to capture or sense the information from its radio environment. Through this capability, the portions of the spectrum that are unused at a specific time or location can be identified. Consequently, the best spectrum and appropriate operating parameters can be selected.
- *Reconfigurability*: The cognitive capability provides spectrum awareness whereas reconfigurability enables the radio to be dynamically programmed according to the radio environment. More specifically, the cognitive radio can be programmed to transmit and receive on a variety of frequencies and to use different transmission access technologies supported by its hardware design.

In 2004, Federal Communications Commission (FCC) indicated that unutilized TV channels in both Very High Frequency (VHF) and Ultra High Frequency (UHF) bands could be used for fixed broadband access. From then on there has been overwhelming interest from research community to develop a standard for Wireless Regional Area Networks (WRAN) systems operating on TV white space using CR technology. IEEE 802.22 WRAN is the first standard developed using CR technology that operates on TV white spaces and focuses on constructing fixed point-to-multipoint WRAN that will utilize VHF/UHF TV bands between 54 MHz and 862 MHz. IEEE 802.22 WRAN systems share the geographically unused TV spectrum on non-interfering basis in rural environment where it is difficult to provide broadband access [3]. IEEE 802.22 is developed to utilize unused TV bands without providing harmful interference to incumbent users.

2.2 Cognitive Femtocell Networks

The surest way to increase the system capacity of a wireless link is by getting the transmitter and receiver closer to each other. It creates the dual benefits of higher-quality links and more spatial reuse. A less expensive alternative is the recent concept of femtocells. The cognitive radio femtocell network works as follows:–

- System initialization: In the beginning, whenever a Femtocell Base Station (FBS) turns on, it first senses the spectrum environment to initialize an available spectrum list. The FBS is responsible to allocate spectrum to its users, and inform them the suitable uplink transmission power. Synchronization between neighboring FBSs is not obligatory in Cog-Fem, but it is an option if any FBS wants to synchronize with its neighbors. The synchronization can be implemented by listening to neighboring femtocells information to obtain the frame length and structure.
- Number of transceivers: Each FBS is equipped with two transceivers. One is called a *sensing radio*, used for spectrum sensing, while the other one is called a *cognitive radio*, used for data communication of both intra-femtocell and inter-femtocell on the selected channels, so that, FBS can do spectrum sensing and data transmission simultaneously.

- Spectrum sensing and primary system protection: Each FBS is able to sense the available spectrum. Whenever an FBS detects the return of a Primary User (PU), it will then stop transmission, and inform its FUs and the neighboring FBSs about the existence of the PU. It then update the available channel list, and run the spectrum sharing algorithms to select new channels and allocate new time-sub-channel blocks for its FUs.
- Control channel: There are two kinds of control channels. One is called *inter-femtocell* control channel, whereby each FBS can communicate with each other. The other one is called *intra-femtocell* control channel, whereby each user in a femtocell can communicate with its FBS to obtain the channel information and allowed transmission power. This control channel could be a dedicated control channel or a rendezvous channel which can be selected according to some metrics such as channel availability. Since every FBS has a broadband connection to the Internet, in spite of using the inter-femtocell control channel, neighboring FBSs can communicate with each other through the broadband connection. Similarly, an additional FBS controller in the Internet can be helpful for the management of FBSs.
- Handover between Macro cell and Femtocell: Whenever an FU moves into a femtocell from a macro cell, it can detect the existence of an FBS by listening to the control channel information, and decide to switch into the femtocell network. By contrast, whenever an FU moves out of a femtocell, it can detect that the strength from FBS is weaker than the strength from macro cell BS (MBS), then it decides to switch into the macro cell network [6].

2.3 Improving Detection Performance of Cognitive Femtocell Networks

The single threshold energy detection may cause serious interference to the primary user. In order to increase the efficiency of the network, double threshold energy detection is proposed. Another detection threshold is added within the conventional single-threshold energy detection algorithm, and it becomes a double-threshold energy detection algorithm with two detection thresholds ($Vth0$ and $Vth1$). The primary user will be detected if and only if $V > Vth1$, and will not be presented if and only if $V < Vth0$, corresponding to $H1$ and $H0$, respectively. When the detected energy V is in ($Vth0$, $Vth1$], this result is invalid because of easy to mistaken. It needs re-detection.

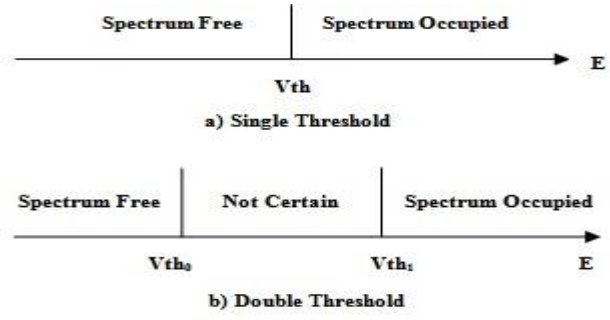


Figure 2: Energy Detection Decision [5]

The probability of collision between the cognitive user and the primary user is $p_c = p\{V \text{ '}< V_{th0} \mid H_1\}$. It is the probability of the primary user which is not detected, but in fact it is existed, and this unoccupied spectrum will be allocated to the cognitive user. It indicates the interference of the cognitive user to the primary user because of the uncertainty of the spectrum detection. The larger the probability of collision between the primary user and the cognitive user, more serious will be the interference of the cognitive user to the primary user. On the contrary, there is less interference [5]. Double threshold energy detection algorithm can decrease the collision probability between primary user and femtocell user effectively. It avoids the femtocell user interfering the primary user.

III. DOUBLE THRESHOLD ENERGY DETECTION ALGORITHM – THE PROPOSED TECHNIQUE

In the proposed technique, the double threshold energy detection algorithm is used to reduce the collision probability between primary user and femtocell user. The received RF power is given to the FPGA platform through the power sensing unit and converter. The block diagram for transmitter and receiver is shown in Fig.3 and 4:–

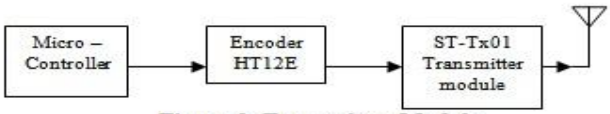


Figure 3: Transmitter Module

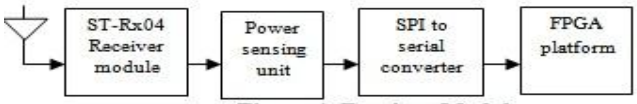


Figure 4: Receiver Module

3.1 Transmitter Module

The input data is generated using microcontroller and given to the encoder.

The encoder is capable of encoding information which consists of N address bits and 12-N data bits. Each address/data input can be set to one of the two logic states. The programmed addresses/data are transmitted together with the header bits via an RF transmission medium upon receipt of a trigger signal.

The HT12E encoder begins a 4-word transmission cycle upon receipt of a transmission enable (active low). This cycle will repeat itself as long as the transmission enable is held low. Once the transmission enable returns high, the encoder output completes its final cycle and then stops.

The encoded data is transmitted through the RF transmitter i.e. ST-Tx01 receiver module.

3.2 Receiver Module

The transmitted data is received through the ST-Rx04 receiver module and then it is given to the power sensing unit for detection RF signal.

The received RF power is converted in voltage output using IC AD8318.

The AD8318 is a demodulating logarithmic amplifier, capable of accurately converting an RF input signal to a corresponding decibel-scaled output voltage. It employs the progressive compression technique over a cascaded amplifier chain, each stage of which is equipped with a detector cell. The AD8318 maintains accurate log conformance for signals of 1 MHz to 6 GHz and provides useful operation to 8 GHz. The input range is typically 60 dB (re: 50 Ω) with error less than ±1 dB. The AD8318 is specified for operation up to 8 GHz. As a result, low impedance supply pins with adequate isolation between functions are essential. In the AD8318, VPSI and VPSO, the two positive supply pins, must be connected to the same positive potential. The VPSI pin biases the input circuitry, while the VPSO pin biases the low noise output driver for VOUT. Separate commons are also included in the device. CMOP is used as the common for the output drivers. Pin CMIP and Pin CMOP should be connected to a low impedance ground plane.

This voltage output is given for analog to digital conversion to the IC 7887 which is a 12-bit ADC.
This digital data is given to the SPI to serial converter. Then, the data is serially transmitted to the FPGA platform.

The double threshold energy detection algorithm is developed on the FPGA platform to reduce the collision probability between the primary user and femtocell user.

The simulation result will be analyzed to see the improvement in the collision probability.

4. CONCLUSION

The review of cognitive femtocell network can be concluded as follows:-

- Cognitive radio is a solution for spectral crowding problem. It introduces the opportunistic usage of frequency bands that are not heavily occupied by licensed users.
- Cognitive radio can be incorporated in femtocell networks to solve the indoor coverage problem and to improve the system performance.
- The cognitive femtocells can achieve almost twice average capacity than the coloring method.
- Co-channel interference can be reduced significantly by using cognitive radio in femtocell networks.
- It will be interesting to see how the double threshold energy detection algorithm is useful in significantly reducing the collision probability.

REFERENCES

[1] J.Xiang, Y.Zhang, T.Skeie, L.Xie, "Downlink spectrum sharing for cognitive radio femtocell networks", IEEE systems journal, Vol.4, No.4, Dec 10.

[2] Akyildiz IF, Lee W, Vuran MC, Mohanty S. "Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey". Int J Comput Telecommun Network 2006: 2127-59.

[3] C.Stevenson, G.Chouinard, Z. Lei, W.Hu, S.Shellhammer, W.Caldwell, "IEEE 802.22: The First Cognitive Radio Wireless Regional Area Network Standard", IEEE communications magazine, Jan.09.

[4] C.Sun, Y.Alemseged, H.Tran and H.Harada, "Cognitive Radio Sensing Architecture and A Sensor Selection Case Study", IEEE communications magazine, 2009.

[5] J.Wu, T.Luo, G.Yue, "An Energy Detection Algorithm Based on Double-threshold in Cognitive Radio Systems", ICISE, 09.

[6] V.Chandrasekhar, J.Andrews, A.Gatherer, "Femtocell networks: A survey", IEEE communications magazine, Sept.08

[7] Lamiaa Khalid, Alagan Anpalagan, "Emerging cognitive radio technology: Principles, challenges and opportunities": 2009: 358-366.

Exergy Requirements for the Manufacturing of Carbon Nanotubes

Renish M vekariya[1], Rakesh P Ravani[2]
1. HOD, A.I.E.T, MECH Dept., Rajkot-GJ, India,
2. Lec.VVP, MECH, Rajkot-GJ, India,

Abstract

The purpose of this paper is to address both the high values, and the large variation in reported values for the energy requirements for the production of carbon nano tubes. The paper includes an estimate of the standard chemical exergy for single walled carbon nano tubes, as well as a historical look at how the minimum physical flow exergy improved as the HiPco process developed.

***Keywords*—** Carbon Nanotubes, Exergy Analysis, SWNT

I. Introduction

Early estimates by Isaacs et al [1] indicated the potentially very large value of the specific energy requirements for carbon single walled nano tubes (SWNT). More recently, energy estimates have been performed for a variety of carbon fibers (SWNTs, multiwall carbon nano tubes – MWNT, and carbon fibers) and a variety of manufacturing processes (Arc, CVD and HiPco) [2]–[6]. These studies show considerable variation in energy estimates (as much as 3 orders of magnitude), and almost two orders of magnitude variation between nominally identical processes. In this paper we review the available data and then look further into the so called HiPco process, (for high pressure carbon mono xide process) [7]-[9] to attempt to explain the large variation in specific energy requirements.

II. Data Summary

In Table 1 we summarize data for various carbon nanofiber production methods (synthesis only) from the literature [1 - 4]. The Synthesis Reaction Carbon Yield (SRCY) is the amount of process carbon needed to produce carbon nano tubes/fibers. It is based upon the flow rates of the process carbon and the carbon product output. Note that these values differ by almost three orders of magnitude, while the specific electrical energy estimates for the synthesis reaction differ by more than three orders of magnitude. The purification step after synthesis may add up to 50% more to the value given in column four [2].
Note: further that estimates for the HiPco process vary by almost two orders of magnitude. We will show that this variation is due in part to the changing nature of the HiPco process as it has been improved. Further, important details of the process are not generally available, and important assumptions in various analyses may differ, and in some cases are not reported.

TABLE 1
Results From The Literature For The Specific Work Input In The Form Of Electrical Energy Per Mass For The Production Synthesis Process For Carbon Nano-Tubes And Fibers.

Process/Product	Source	SRCY1	GJ/kg2	Ref.
HiPco/SWNT	CO	50%	465.8	[1]
HiPco/SWNT	CO	0.08%	31.8	[2]
HiPco/SWNT	CO	NA	5.8	[3]
Arc/SWNT	Carbon Anode	4.5%	458.7	[1]
Arc/SWNT	Carbon Anode	4.5%	83.7	[2]
CVD/SWNT [	CH4	2.95%	915.8	[1]

III. Standard Chemical Exergy Of Swnt

The specific standard chemical exergy of a chemical compound is the minimum (reversible) work per mass to produce this component starting from the identified chemical components of the reference environment at the "dead state". For the production of SWNTs the process would start from the carbon stored in the atmosphere as CO2 gas and include the following steps; 1) the concentrating of the CO2 from its reference concentration in the atmosphere to pure, 2) the reduction of CO2 into its chemical constituents, carbon graphite) and oxygen,

1 SRCY = Synthesis Reaction Carbon Yield
2 Electricity for synthesis reaction only, does not include loss at utility
3 CNF = carbon nano fiber

(3) the separation of a graphene layer from the graphite, and 4) the bending of the graphene layer into a carbon SWNT. The first two steps can be obtained from previous results for the standard chemical exergy for graphite as 410.26 kJ/mol or 34.16 kJ/g [10]. The work of cohesion to reversibly separate a layer from a bulk material is two times the surface energy for the new surface [11]. Abrahamson [12] has provided a review of the surface energy of graphite, and estimates it at 25kJ/mol or 2.08kJ/g. Finally, Lu [13] has analyzed the anisotropic bending of a graphene layer and estimated the (fully elastic) moment – curvature behavior. This bending stiffness is due to the bond angle effect on interatomic interactions. From this we may estimate the strain energy for the bending of a flat sheet 0.34nm thick to a tube with an outside diameter of 1.2 nm as 0.78 kJ/g. Putting this together we estimate the specific standard chemical exergy of a carbon SWNT as,

$$ex,o,SWNT = 34.16 + 4.16 + 0.78 = 39.1 \text{ kJ/g.}$$

This is the minimum reversible work to produce SWNT at the "reduced dead state" (*To, po*) from components of the environment at the ultimate dead state, or just the "dead state". Note that this value is diameter dependent.

IV. Analysis of the Hipco Process

The HiPco process was developed by Prof. Richard Smalley's group at Rice University in the late 1990's. The process is based on the so called disproportionation (Boudouard) reaction as given below. This reaction, under appropriate conditions and in the presence of a suitable catalyst can produce carbon SWNT. Note that the reaction given in equation (1) is spontaneous and exothermic, yielding (at standard conditions) an exergy output of 5.06 kJ/g of SWNT.

$$CO + CO \rightarrow C(s) + CO_2 \quad (1)$$

However, to produce SWNTs this reaction is carried out at elevated temperatures and pressures (~1000oC, ~30 atm) requiring significant exergy inputs, currently several orders of magnitude larger than the chemical exergy change. In what follows we calculate the minimum physical exergy required to create the conditions necessary to produce SWNT as reported in a series of publications by the Smalley team [7]– [9] and others [2], [3]. During the approximately 9 year period of development covered by these publications, we will see that the process has been significantly improved, reducing the exergy requirement by a factor of 34. The process, illustrated in Fig. 1 taken from their 2007 patent [9] shows recycling flows of CO gas that are repeatedly exhausted and then reheated and repressurized. Our analysis focuses on a highly idealized version of the process that looks only at the CO gas flows through the reaction chamber and compressor. These flows are treated separately as steady state open systems, with only a work input (no heat inputs). The gas flow requirements are governed by the flow rates and production rates of the process. We assume that the recycled CO is hot (100C) and at atmospheric pressure before being returned to the process conditions (1000 C and 30 atm). This calculation can be done assuming ideal gas behavior and using (2), see Gutowski and Sekulic [14].

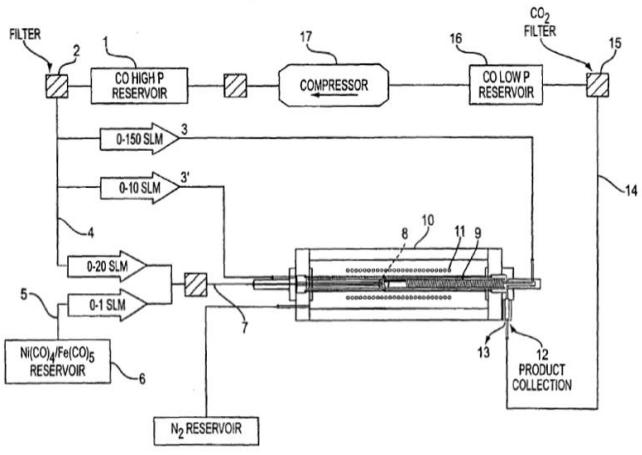


Fig. 1. Schematic representation of the HiPCo process [Ref].

$\Delta e_{x,physical}=C_p(T-T_0)-C_P T_0 \ln(T/T_0)+T_0 R\ln(P/P_0)$

This is the change in the specific physical flow exergy for an ideal gas. To convert the results in [3] to the minimum exergy for comparison purposes, we used an assumed efficiency of 0.75 as discussed in their paper.
The key operating parameters and results for the minimum physical exergy to produce SWNT by the HiPco process are given in Table 2 for different times in the evolution of this process. These results are also plotted in Fig. 2 at the end of this paper. The results clearly show the improvement in the process and imply the difficulty in making energy estimates for new rapidly changing technologies.

V. Comparison with Other Manufacturing Processes

In a previous publication [6], we reviewed a wide range of manufacturing processes. Here we add the new data on carbon nano tube manufacturing (synthesis only) to our summary plot given in Fig. 3 at the end of this paper. We add two types of data: 1) the historical data for the improvement of the HiPco process as given in Table 2 and some of the data from the literature as given in Table 1.

TABLE 2
Estimated Specific Energy Requirements And Process Rates For Synthesis Of Cnts

Process Name	Product	Process Rate (kg/h)	Synthesis Energy Requirements (J/kg)	Reference
HiPco Process	SWNT	4.50E-04	3.18E+10	Healy et al (2008) AND Isaacs et al (2008)
Arc Ablation	SWNT	8.10E-05	8.73E+10	
CVD Process	SWNT	9.80E-06	2.76E+11	
Vapor-Grown CNF Process	CNF	1.30E-02	3.13E+09	Khanna et al (2008)
Vapor-Grown CNF Process	CNF	1.80E-02	2.22E+09	
Vapor-Grown CNF Process	CNF	5.20E-02	7.61E+08	
HiPco Process	SWNT	1.38E-06	5.34E+11	Nikolaev et al (1999)
HiPco Process	SWNT	4.50E-04	8.11E+10	Bronikowski et al (2001)
HiPco Process	SWNT	4.50E-04	2.41E+10	Smalley et al (2007)

Compared to other processes, the carbon nanotube (CNT) manufacturing data shows several noteworthy trends. First, while the specific electrical energy requirements are quite large, (generally exceeding 1 GJ/kg) they are not the largest we have seen. In general, they seem comparable to semi conductor processes. Secondly, the power requirements for CNT are generally on the low side of manufacturing processes, at least currently, for the current modest scales of production. Of course, this can, and is changing as various companies announce the openings of large scale production facilities. The data also clearly show how the HiPco process has improved over recent times.

VI. The Degree of Perfection for Swnt Production Processes

For resource accounting purposes, the so called "Degree of Perfection" can be a useful metric for evaluating manufacturing processes. We define the Degree of Perfection for manufacturing processes as the ratio of the standard chemical exergy of the output product(s) divided by the sum of the input exergies including the exergy equivalents of any work and/or heat inputs as well as the standard chemical exergies of all material inputs. See [10], [14].

$$n_p=\frac{Ex0,product}{\sum Ex,input}$$

The degree of perfection is a second law efficiency measure that can be used to identify opportunities for improvement and to compare with other processes. To illustrate, consider the 2004 version of the idealized HiPco process with ideal reversible thermal treatment and pressurization stages, but operated in an open loop i.e. without recycling of the CO gas. The degree of perfection would be given by (39.1 kJ/g) / (12,500g x 9.82 kJ/g +12,040 kJ/g) = 2.9 x 10-4. This low value is due to the requirement for large amounts of input CO and complete destruction of the physical exergy that was previously invested into the CO gas stream. Now if CO recycling can reduce the input CO from 12,500 grams to 4.67 grams (the minimum stoichiometric quantity as given by (1)) this would improve the degree of perfection to (39.1kJ/g)/ (4.67g x 9.82kJ/g + 12,040kJ/g) = 3.2 x 10-3. This is about an order of magnitude improvement, but still low because of the complete loss of the invested physical exergy. A further improvement could involve the preheating of the incoming CO gas stream using some of this lost exergy. In general, the degree of perfection measure for the performance of other synthesis reactions for CNT are also quite low (on the order of 10-3 to 10-4) owning to; 1) the one time use of large quantities of high exergy material inputs – primarily the carbon source inputs, and 2) the high physical exergy requirements – most processes are performed at high temperatures. However, these low values are still quite highcompared to some semi conductor processes, which can be in the range of 10-5 and 10-6 for SiO2 processes, see [5], [6]. The main difference is due to the relatively high standard chemical exergy of CNTs.

VII. Closing Comments

One purpose of this paper was to examine the minimum exergy requirements to make carbon nanotubes. We find that while the exergy requirements are high, they are falling at a rather fast rate due to process improvements and could fall still more in the future. This exercise underlines the challenges of trying to perform a Life Cycle Assessment of an evolving technology. Early in the life of a new technology one may be following a moving target. We believe that the results given here help explain some of the variation seen in early LCI reports as given in Table 1.

Note that the values reported here are only for the synthesis part of the nanotube production process. Furthermore these values are for minimum exergy requirements, not actual. To make a full estimate one would have to consider: 1) the exergy required to make the input materials. Because of the purity requirements for some of these inputs, this is likely to be very large. For example according to Williams et al [16] for gases with purities in the 0.1 ppm level, the energy required for purification can in the range of 20 – 200 GJ/g. (These are very high energy requirements indeed!) In addition, 2) the carbon nanotube purification step needs to be added, including the effect of yield losses. And 3) the minimum synthesis values given here need to be increased to account for losses in the synthesis step. Finally, 4) additional energy required for infrastructure needs to be added (environmental conditioning etc), and 5) losses at the electric utility need also to be added. Taking this into account it is quite reasonable to expect an order of magnitude estimate of the embodied energy requirements for carbon nanotubes to be in the region of 0.1-1.0 TJ/kg. Such a high value compared to other materials would make this one of the most energy intensive materials known to humankind. See for example [1], [18]. Ironically this enormous specific energy requirement constitutes only a very small fraction of the manufacturing costs (< 1% as discussed by Healy [2] and Isaacs [15]). For example, say the energy cost for making carbon nanotubes is on the order of 36GJ of electricity per kilogram or 36MJ/g.

This is equal to 10 kWh/g. Now at 7 cents a kilowatt hour this yields a cost of 70 cents per gram. But carbon nanotubes can sell for around $300/g. In other words, the electricity cost in this case is on the order of 0.2% of the price, and according to a recent cost study, energy costs for all manufacturing processes for nanotubes result in about 1% of the cost [15]. It appears that new manufacturing processes can produce novel products with high demand resulting in a value that far exceeds the energy (electricity) cost. At the same time however, since our current electricity supply comes primarily from fossil fuels, most of the environmental impacts associated with these materials (e.g. global warming, acidification, mercury emissions) are related to this use of electricity [15]. How can we reconcile this inconsistency? One comment would be that the current price for carbon nanotubes may well be inflated due to the rather substantial government funds for nanotechnology research worldwide. Another comment, of course, is that, from an environmental perspective, electricity from fossil fuels is vastly underpriced. That is, the environmental and health externalities associated with the use of fossil fuels are not included in the price of electricity.

References

[1] J.A. Isaacs, A. Tanwani, and M. L. Healy."Environmental assessment of SWNT production." Proceedings of the 2006 IEEE International Symposium on Electronics and the Environment. 8-11 May 2006: pp. 38-41.

[2] M.L Healy, L. J. Dahlben and J. A. Isaacs,"Environmental assessment of single-walled carbon nanotube processes", Journal of Industrial Ecology, Vol. 12, No. 3, June 2008, , pp. 376-393.

[3] D. Kushnir, and Bjorn A. Sanden, "Energy requirements of carbon nanoparticle production", Journal of Industrial Ecology, Vol. 12, No. 3, June 2008, pp. 360 -375.

[4] V. Khanna, B. R. Bakshi and L. James Lee, "Carbon nanofiber production – life cycle energy consumption and environmental impact", Journal of Industrial Ecology, Vol. 12, No. 3, June 2008, pp. 394 -410.

[5] M. Branham, "Semiconductors and sustainability: energy and materials use in the integrated circuit industry". Department of Mechanical Engineering, M.I.T. MS Thesis, 2008.

[6] T.G. Gutowski, M. S. Branham, J. B. Dahmus, A. J. Jones, A. Thiriez and D. Sekulic, "Thermodynamic analysis of resources used in manufacturing processes"**,** Environmental Science and Technology, 43, January 29, 2009, pp 1584-90.

[7] M.J. Bronikowski, P. A. Willis, D. T. Colbert, K.A. Smith and R. E. Smalley, "Gas-phase production of carbon single-walled nanotubes from carbon monoxide via the HiPco process: A parametric study", J. Vac. Sc. Technol. A 19(4), American Vacuum Society, Jul/Aug. 2001, pp. 1800-1805.

[8] P. Nikolaev, M.J. Bronikowski, R. K. Bradley, F. Rohmund, D.T. Colbert, K.A. Smith and R. E. Smalley, "Gas-phase catalytic growth of single-walled carbon nanotubes from carbon monoxide", Chemical Physics Letters 313, , Elsevier, November 1999, pp. 91-97.

[9] R.E. Smalley, K.A. Smith, D.T. Colbert, P. Nikolaev, M.J. Bronikowski, R.K. Bradley, and F. Rohmund," Single Wall Carbon Nanotubes from High Pressure CO". U.S. Patent No. US7,204,970, April 17, 2007.

[10] J. Szargut,, D. R. Morris and F. R. Steward, Exergy Analysis of Thermal Chemical and Metallurgical Processes, Hemisphere Publishing Corporation, 1988. [11] Cherry B. W., Polymer Surfaces, Cambridge University Press 1981

[12] J.Abrahamson, "The surface energies of graphite", Carbon, Vol. 11, No. 4-E, Pergamon Press, 1975. , pp. 357-362.

[13] Qiang Lu and Rui Huang, "Nonlinear mechanics of single-atomic-layer graphene sheets", International Journal of Applied Mechanics, Vol. 1. No. 3, Imperial College Press, 2009, pp. 443-467.

[14] T.G.Gutowski and D.P. Sekulic, "The thermodynamic analysis of manufacturing processes", chapter in Thermodynamics and the Destruction of Resources, B.R. Bakshi, T.G. Gutowski and D.P. Sekulic Cambridge University Press, to appear 2010 IEEE, International Symposium on Sustainable Systems and Technologies, Washington D.C., May 16-19, 2010

[15] J.A.Isaacs, A. Tanwani, M.L. Healy, "Economic assessment of single-walled carbon nanotube processes", J. Nanopart Res. DOI 10.007/s11051-009-9673-3, Research Paper published on line, Springer, June 2009.

[16] E. Williams, N. Krishnan, and S. Boyd Case Studies in Energy Use to Realize Ultra- High Purities in Semiconductor Manufacturing, ISEE**,** IEEE 2008

[17] V. Smil, Energy in Nature and Society, MIT Press 2008.

[18] M.F. Ashby, Materials and the Environment – Eco- Informed Material Choice, Elsevier Inc., 2009. Fig. 2 Evolution of the HiPco process development over time in terms of the calculated minimum theoretical physical exergy requirements and estimated actual exergy over a nine year period. IEEE, International Symposium on Sustainable Systems and Technologies, Washington D.C., May 16-19, 2010 Fig. 3a. Energy intensity (J/kg) Vs process rate (kg/hr) for 20 different processes. Data and References cited in the figure are from Ref. [6]. Fig 3b. Energy intensity (J/kg) Vs process rate (kg/hr) for the production for Carbon Nano-fibers. See Table 3 and text.

Experimental Study of Partial Replacement of Fine Aggregate with Waste Material from China Clay Industries

[1] A.Seeni, [2]Dr.C.Selvamony, [3]Dr.S.U.Kannan, [4]Dr.M.S.Ravikumar
[1] Research Scholar, Anna University – Chennai, Tamilnadu, India.
[2, 3, 4] Professor, Department of Civil Engineering
Sun College of Engineering & Technology
Tamilnadu, India.

Abstract –
The utilization of industrial and agricultural waste produced by industrial process has been the focus of waste reduction research for economical, environmental and technical reasons. This is because over 300 million tones of industrial waste are being produced per annual by agricultural and industrial process in India. The problem arising from continuous technological and industrial development is the disposal of waste material. If some of the waste materials are found suitable in concrete making not only cost of construction can be cut down, but also safe disposal of waste material can be achieved. The cement of high strength concrete is generally high which often leads to higher shrinkage and greater evaluation of neat of hydration besides increase in cost. A partial substitution of cement by an industrial waste is not only economical but also improves the properties of fresh and hardened concrete and enhance the durability characteristics besides the safe disposal of waste material thereby protecting the environment form pollution This paper deals with partial replacement of fine aggregate with the industrial waste from China Clay industries. The compressive strength, split tensile strength and flexural strength of conventional concrete and fine aggregate replaced concrete are compared and the results are tabulated.

Keywords- China Clay, Compressive strength, Concrete, Fine aggregate, Flexural strength, Industrial waste, Split tensile strength,

I Introduction

Portland cement concrete is made with coarse aggregate, fine aggregate, Portland cement, water and in some cases selected admixtures (mineral & chemical). In the last decade, construction industry has been conducting research on the utilization of waste products in concrete, each waste product has its own specific effect on properties of fresh and hard concrete. Conservation of natural resources and preservation of environment is the essence of any development. The problem arising from continuous technological and industrial development is the disposal of waste material. If some of the waste materials are found suitable in concrete making, not only cost of construction can be cut down, but also safe disposal of waste materials can be achieved. The use of waste products in concrete not only makes it economical but also solves some of the disposal problems.

Objectives and Scopes:

1. To effectively utilize the waste material from the china clay industries.
2. To reduce the problem of disposal of industrial waste.
3. To prove that the industrial waste from china clay industries can be a replacement for fine aggregate.
4. To study the physical and chemical properties of industrial waste and are the ingredients in concrete.
5. To replace the fine aggregate by industrial waste in different ratio such as 10%, 20%, 30%, 40%, and 50% in M30 mix concrete
6. To determine the compressive strength and Split tensile strength and compare it with the conventional concrete.

II TESTING PROGRAMME

In the present study various tests on material such as cement, fine aggregate, coarse aggregate and the waste material from china clay industries were performed as per the Indian Standards.

Material Used

1. Fine aggregate:

a) Sand: River sand was used as fine aggregate. The size of the sand used is 4.75 mm and down size. The properties of fine aggregate investigated are presented in table 1

Table 1 Properties of Fine aggregate

Sl.No	Property	Value
1	Specific Gravity	2.8
2	Fineness Modulus	3.1
3	Water Absorption	0.5%
4	Surface Texture	Smooth

Table 2 Properties of Industrial waste

Sl.No	Property	Value
1	Specific Gravity	2.7
2	Fineness Modulus	2.7
3	Water Absorption	0.5%
4	Surface Texture	Smooth

b) Waste material from China Clay Industry: This material procured from the local china clay products industry was used as partial replacement for river sand. The properties of the material investigated are presented in table 2. The size of the material used is 4.75 mm and down size.

2 Coarse aggregate:

Machine crushed granite obtained from a local quarry was used as coarse aggregate. The properties of the coarse aggregate are shown in table 3

Table 3 Properties of Coarse Aggregate

Sl.No	Property	Value
1	Specific Gravity	2.8
2	Fineness Modulus	7.5
3	Water Absorption	0.5
4	Particle Shape	Angular
5	Impact Value	15.2
6	Crushing Value	18.6

3. Water:

Water used in this project is potable water.

4 Cement:

Portland Pozzolanic Cement of 43 grade was purchased from the local supplier and used throughout this project. The properties of cement used in the investigation are presented in table 4.

Table 4. Properties of Cement

Sl.No	Property	Value
1	Specific gravity	3.15
2	Fineness	97.8
3	Initial Setting Time	45 min
4	Final Setting Time	385 min
5	Standard Consistency	30%
6	Fineness Modulus	6%

III Preparation of Specimens

Based on the above results the water quantity, cement, fine aggregate and coarse aggregate required for design mix of M30 were calculated based on the procedure given in IS code method in IS :2009. The final mix ratio was 1:1.462:2.695 with water cement ratio of 0.44. The measurement of materials was done by weight using electronic weighing machine. Water was measured in volume. Concrete was placed in moulds in layers. The cast specimens were removed from moulds after 24 hours and the specimens were kept for water curing.

Figure 1 Specimen Moulds

The details of mix designation and specimens used in experimental program are given in table 5.

Table 5 Mix Details

Sl. No	Mix Designation	Cement	Fine Aggregate		Coarse Aggregate	No: of Specimens		
			Sand	Industrial waste		Cube	Cylinder	Prism
1	M0	100%	100%	0%	100%	3	3	3
2	M1	100%	90%	10%	100%	3	3	3
3	M2	100%	80%	20%	100%	3	3	3
4	M3	100%	70%	30%	100%	3	3	3
5	M4	100%	60%	40%	100%	3	3	3
7	M6	100%	50%	50%	100%	3	3	3

Figure 2 Casted Specimens

IV Testing of specimens:

For each batch of concrete, 3 cubes of 150mm x 150mm x 150mm size were tested to determine compressive strength of concrete, 3 cylinders of 150mm diameter and 300 mm length were tested to determine split tensile strength of concrete and three prisms of 100mm x 100mmx 500mm were tested to determine flexural strength of concrete.

V Results and discussions:

Table 6 Experimental Test Results at 28 days curing

Sl.No	Mix Designation	Compressive Strength N/mm^2	Split Tensile Strength N/mm^2	Flexural Strength N/mm^2
1	M0	31.5	3.35	5.32
2	M1	34	3.42	5.46
3	M2	36	3.63	5.66
4	M3	37.5	3.85	5.74
5	M4	33	3.45	5.28
6	M5	31	3.15	4.88

From the above table it is found that the compressive strength of the control concrete was 31.5 N/mm^2. The compressive strength was found to be maximum at 30% (37.5N/mm^2) replacement of fine aggregate by industrial waste which was greater than the conventional concrete. The compressive strength reduced beyond 30% replacement. Thus it is evident that fine aggregate can be replaced by the waste material from china clay industries up to 30%.

Similarly the split tensile strength and flexural strength was also found to be maximum at 30% (3.85 N/mm^2 and 5.74 N/mm^2) replacement which was greater than the conventional concrete (3.35 N/mm^2 and 5.32N/mm^2).

The graphs showing the compressive strength, split tensile strength and flexural strength of the different mixes at 28 days of curing are shown in figures 3, 4 and 5 respectively.

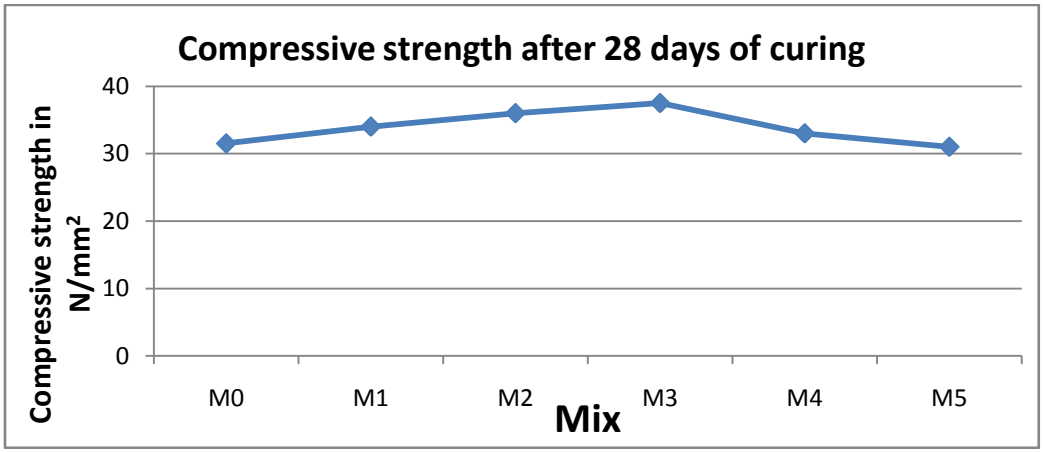


Figure 3 Compressive strength after 28 days of curing

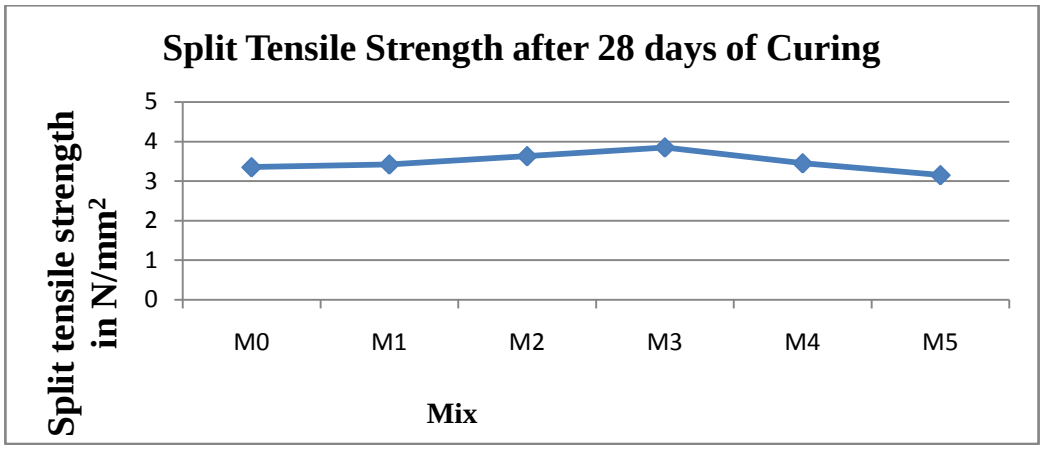


Figure 4 Split Tensile Strength after 28 days of Curing

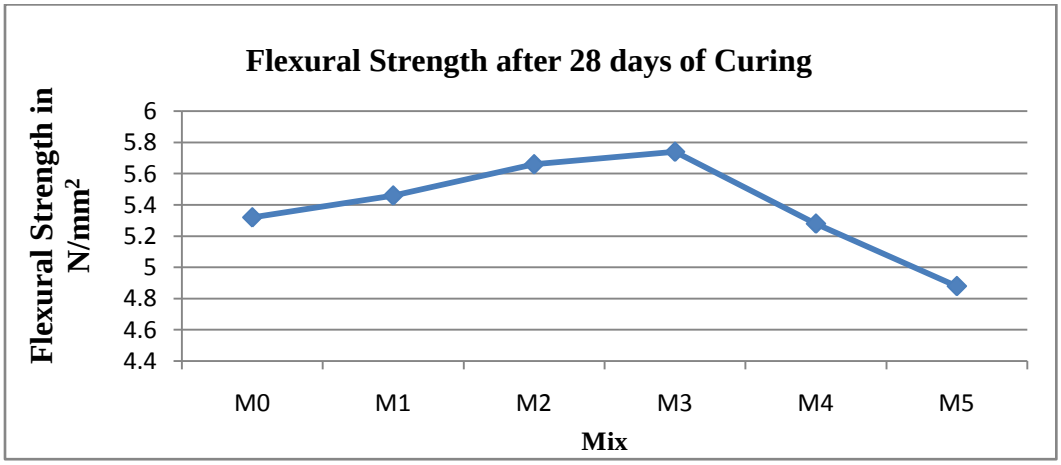


Figure 5 Flexural Strength after 28 days of Curing

VI Conclusion

From the results of experimental investigations conducted it is concluded that the waste material from china clay industries can be used as a replacement for fine aggregate. It is found that 30% replacement of fine aggregate by industrial waste give maximum result in strength and quality aspects than the conventional concrete. The results proved that the replacement of 30% of fine aggregate by the industrial waste induced higher compressive strength, higher split tensile strength and higher flexural strength. Thus the environmental effects from industrial waste can be significantly reduced. Also the cost of fine aggregate can be reduced a lot by the replacement of this waste material from china clay industries.

References

[1]. Aggarwal.P, Aggarwal.Y, Gupta.S.M [2007] "Effect of bottom ash as replacement of fine aggregate in concrete", Asian journal of civil engineering [Building and housing] Vol.8, No.1, PP.49-62.

[2]. Gurpreet Singh and Rafat siddique [2011] "Effect of waste foundry sand [WFS] as partial replacement of sand on the strength, ultrasonic pulse velocity and permeability of concrete", International journal of construction and building materials Vol.26, PP.416-422.

[3]. IS 10262 - 2009 Recommended guidelines for concrete mix design

[4]. John zachar and Tarun R.naikin [2007] "Replacement of fine aggregate with foundry sand", Milwaukee journal – sentinel Vol.3, PP.18-22.

[5]. Mahmoud solyman [2006] "Classification of Recycled Sands and their Applications as Fine Aggregates for Concrete and Bituminous Mixtures", journal article: DOI universitat kassel, Vol.5, pp.1-196.

Robust LMI-Based Controller Design using H_∞ and Mixed H_2/H_∞ for Semi Active Suspension System

Saeed M. Badran
Electrical Engineering Department
Faculty of Engineering, Al-Baha University Al-Baha, Kingdom of Saudi Arabia

Abstract:
Control of vehicle suspension systems has been the focus of extensive work in the past two decades. Many control strategies have been developed to improve the overall vehicle performance, including both ride quality and stability conditions. In this paper, we use the H_∞ and mixed H_2/H_∞ techniques with a semi active suspension system to increase the passenger's ride comfort performance A two degree of freedom dynamic model of a vehicle semi-active suspension system was presented. The role of H∞ is used to minimize the disturbance effect on the system output whereas H_2 is used to improve the transients against random disturbances. The capability of the system of improving comfort of operators has been evaluated through simulations carried out with a validated model of the entire vehicle. The results of this study revealed that the use of mixed H_2/H_∞ with pole placement for a semi-active suspension system showed a great improvement compared with H_∞ systems. In addition, the results of the simulation showed that ride comfortable of the vehicle can be improved effectively by using the semi-active suspension with mixed H_2/H_∞ control method, and also the mixed H_2/H_∞ control method more effective than the H_∞ control method. Finally, this paper showed a robust use of both H_∞ and mixed H_2/H_∞ problem which can be solved using linear matrix inequality (LMI) techniques.

Keywords*:* H_∞ control; mixed H_2/H_∞ control; robust control; semi-active suspension; ride comfort; simulation; pole placement.

I Introduction

All motor manufactures are currently engaged in research and development to ensure that they remain at the competitive edge, in terms of both vehicle performance and perceived human factors such as comfort and drivability. Conventional vehicle suspension systems consists of a passive spring in parallel with a damper, their main functions being to support the body mass and to provide both passenger comfort and road holding. These have a number of limitations due to the fixed nature of the components used and requirements for the vehicle to function over a wide variety of operating conditions. A suspension system is also required to react to changes in vehicle load, a constraint which requires a stiff suspension. The introduction of active elements into the suspension allows the compromise to redefined, providing an all round improvement in performance. The topic of this paper is the using of H_∞ and mixed H_2/H_∞ control with a semi active suspension system. The design of control algorithms for semi-active vehicle suspensions has been an active research field for over forty years [1,2]. Numerous control algorithms have been developed for semi-active suspensions [1,3]. The principle of semi-active damping is the control of variable dampers for the purpose of vibration isolation. Semi-active damping has been shown to significantly improve vibration isolation in comparison to passive damping for a range of mechanical and civil engineering applications see for example [4–8]. The semi-active suspension of vehicles uses the damping components that can be controlled and the closed loop system, which can regulate the damping force according to the feedback signals generated by suspension working space ,and acceleration of the car body, so that the damping suspension stay in the best condition and improve the ride comfort ability. Many active suspension control approaches have been proposed such as Linear Quadratic Gaussian (LQG) control, adaptive control, and non-linear control to overcome these suspension systems problems [9-11]. Stability represents the minimum requirement for control systems. However, in most cases, a good controller should act sufficiently fast with well-damped response beside the disturbance attenuation on selected system outputs. If the controller design is not robust against disturbance and parameters change, the system may become unstable [12,14]. Mixed H_2/H_∞ robust control alleviates such handicap [15-19]. Linear matrix inequality (LMI) [20] is one of the most effective and efficient tools in controller design. Many LMI-based design methods of static out feedback (SOF) design have been proposed over the last decade. The main advantage of the H_∞ control is that it provides maximum robustness to the most destabilizing uncertainty, which is modeled as disturbance input. The H_2 performance criterion introduced above is extended with an H_∞ criterion for the body mass acceleration. This idea leads to an attempt of the mixed H_2/H_∞ control design scheme.

II. Mathematical Model Formulation

A two-degree-of-freedom "quarter-car" vehicle suspension system is shown in Figure 1. An advantage of this model is that many published results are available, which makes it easy to verify and compare the results which those other researches. It represents the vehicle system at each wheel i.e. the motion of the axle and of the vehicle body at any one of the four wheels of the vehicle. The suspension itself is shown to consist of a spring k_s, a damper c_s and an active force actuator u. The active force u can be set to zero in a passive suspension. The sprung mass m_s represents the quarter car equivalent of the vehicle body mass. The unsprung mass mu represents the equivalent mass due to the axle and tire. The vertical stiffness of the tire is represented by the spring k_t. The variables z_b , z_w and z_o represent the vertical displacements from static equilibrium of the sprung mass, unsprung mass and the road respectively[21]. In this paper it was assumed that only the suspension deflection could be measured and used by the controllers.

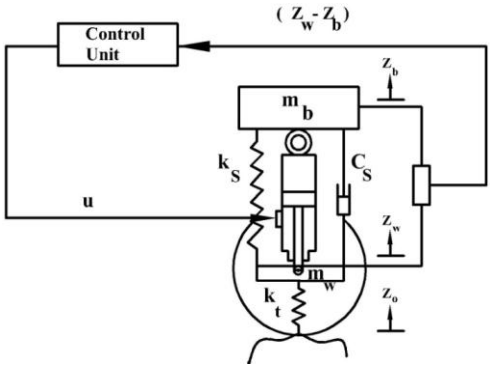


Figure 1. Semi-active suspension system.

$$m_b\ddot{z}_b + C_s(\dot{z}_b - \dot{z}_w) + k_s(z_b - z_w) = u \quad (1)$$

$$m_w\ddot{z}_w + k_t(z_o - z_w) - c_s(\dot{z}_b - \dot{z}_w) - k_s(z_b - z_w) = -u \quad (2)$$

Assume the following

$x_1 = Z_o - Z_w$,

$x_2 = \dot{Z}_b$,

$x_3 = Z_u - Z_r$, $x_4 = \dot{Z}_u$

Where :

$x_{1=}z_s - z_w$ is the suspension deflection (rattle space)

$x_{2=}\dot{z}_b$ is the absolute velocity of sprung mass

$x_{3=}z_o - z_w$ is the tire deflection

$x_{4=}\dot{z}_w$ is the absolute velocity of unsprung mass

The state equations of the sample power system can be written in the vector-matrix differential equation form as:

$$\dot{x} = Ax + B_1 z_o + B_2 u$$

The system matrix A, the control matrix B_1, and the road input matrix B_2 are, respectively, denoted as

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{(k_t+k_s)}{m_b} & 0 & 0 & 1 \\ \frac{k_s}{m_w} & \frac{c_s}{m_w} & -\frac{k_t}{m_w} & -\frac{(c_s)}{m_w} \end{bmatrix} \quad B_1 = \begin{bmatrix} 0 \\ \frac{1}{m_b} \\ 0 \\ -\frac{1}{m_w} \end{bmatrix}, \text{ and } \quad B_2 = \begin{bmatrix} 0 \\ 0 \\ -1 \\ -\frac{1}{m_w} \end{bmatrix}$$

The suspension parameters are shown the Table 1.

Table 1. Quarter car parameters

Parameters	Symbols	Quantities
Body mass	m_b	250 kg
Wheel mass	m_w	50 kg
Stiffness of the body	k_s	16 kN/m
Stiffness of the wheel	k_t	160 kN/m
Stiffness of the damper	c_s	1.5 kN.s/m

III. Input Profile Excitation

The representation of the road profile is vital for vehicle dynamic simulations because it is the main source of excitation. An accurate road model is as important as a good vehicle model. The Excitation input from the road is transmitted to the vehicle floor. For the simplification of the dynamic modeling, it is assumed that there exists only the vertical motion of the vehicle. Both pitching and rolling motions are ignored in this study. The reduction of forces transmitted to the road by moving vehicles (particularly for heavy vehicles) is also an important issue responsible for road damage. Heavy vehicle suspensions should be designed accounting also for this constraint. In this work, A periodic road excitation input has been used for simulation of suspension systems. The periodic input is used for smooth road in order to evaluate ride comfort as shown in Figure 2. It is widely recognized that the road surfaces approximate to Gaussian processes, having a power spectral density (PSD) of the form [22]:

$$PSD(f) = \frac{R_C V^{n-1}}{f^n} \qquad (3)$$

Where:

R_C : Road Roughness Coefficient.

f : Road Excitation Frequency, Hz.

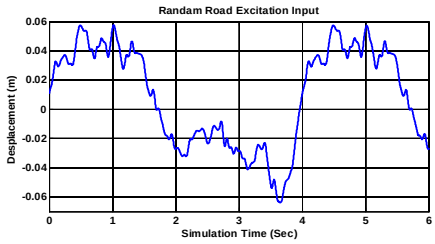


Figure 2. Road excitation.

IV. Robust H_∞ Controller (RH_∞)

In a typical H_∞ design problem, the nominal plant model represented by its transfer function P(s) is usually known and the design problem for an output feedback control is formulated as a standard H_∞ problem, as described by the block diagram of Figure 3. P(s) represents the plant and K(s) the controller transfer function in Laplace domain. The controller is aimed to be designed using the H_∞ design technique. In the block diagram, w represents the external disturbances, z the regulated outputs and y the measured outputs. The vector u consists of the controlled inputs[23].

Let:

$$P(s): \begin{cases} \dot{x} = Ax + B_1 w + B_2 u \\ z_1 = C_1 x + D_{11} w + D_{12} u \\ y = C_2 x + D_{21} u \end{cases} \qquad (4)$$

Controller:

$$K(s): \begin{cases} x_K = A_K x_K + B_K y \\ u = C_K x + D_K y \end{cases} \quad (5)$$

be state-space realizations of the plant P(s) and controller K(s), respectively, and let

$$\begin{cases} \dot{x}_{CL} = A_{CL} x_{CL} + B_{CL} w \\ z = C_{CL} x_{CL} + D_{CL} w \end{cases} \quad (6)$$

be the corresponding closed-loop state-space equations with

$$X_{CL} = [X \quad X_K]^T$$

The design objective for finding K(s) is to optimize the H_∞-norm of the closed-loop transfer G(s) from (w) to (z), i.e.,

$$G(s) = C_{CL}(s - A_{CL})^{-1} B_{CL} + D_{CL} \quad (7)$$

and $|G(s)_{ZW}| < \gamma$

using the LMI technique. □ is a specific number. This can be fulfilled if and only if there exists a symmetric matrix X such that the following LMIs are satisfied.

$$\begin{bmatrix} A_{CL}X + XA_{CL}^T & B_{CL} & XC_{CL}^T \\ B_{CL}^T & -1 & C_{CL}^T \\ C_{CL}X & D_{CL} & -\gamma^2 I \end{bmatrix} < 0 \quad (8)$$

$$X > 0$$

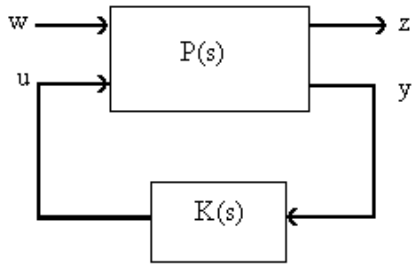


Figure 3. Block diagram of output feedback.

It represents the system disturbance re t-case disturbance on the output. LMI toolbox can be used for such controller design [24]. Where,

$$A_{CL} = \begin{bmatrix} A + B_2 D_K C_2 & B_2 C_K \\ B_K C_2 & A_K \end{bmatrix}$$

$$B_{CL} = \begin{bmatrix} B_1 + B_2 D_K D_{21} \\ B_K D_{21} \end{bmatrix}$$

$$C_{CL} = \begin{bmatrix} (C_1 + D_{12} D_K C_2) & D_{12} C_K \end{bmatrix}$$

$$D_{CL} = \begin{bmatrix} D_{11} + D_{12} D_K D_{21} \end{bmatrix}$$

LMI constraints defined by (8) can be derived from: Stability condition based on Lyapunov energy function;

$$V(X) = x^T X x > 0 \quad (9)$$

$$\frac{dV}{dt} = x^T\left(A^T X + XA\right)x + x^T(XB) + u + u^T(B^T X)x < 0 \quad (10)$$

From Eq. (10) the stability LMI constraints are;

$$\begin{pmatrix} A_{CL}^T X + XA_{CL} & XB_{CL} \\ B_{CL}^T & -\gamma^2 I \end{pmatrix} < 0 \quad (11)$$

$$X > 0$$

Minimization of the disturbance effect condition on the selected outputs based on infinity norm (H_∞) that equals;

$$yTy - \gamma^2 uTu < 0 \quad (12)$$

From Eq. (12) the disturbance effect under LMI constraints is;

$$\begin{pmatrix} C_{CL}^T C_{CL} & C_{CL}^T D_{CL} \\ D_{CL}^T C_{CL} & D_{CL}^T D_{CL} \end{pmatrix} < 0 \quad (13)$$

From Eqs.(11) and (13) LMI constraints become;

$$\begin{pmatrix} A_{CL}^T X + XA_{CL} + C_{CL}^T C_{CL} & XB_{CL} + C_{CL}^T D_{CL} \\ B_{CL}^T P + D_{CL}^T C_{CL} & D_{CL}^T D_{CL} - \gamma^2 I \end{pmatrix} < 0 \quad (14)$$

$X = X^T > 0$ (Positive definite matrix)

According to the Schur complement LMI constraints defined by (14) become as given in (8).

5. Mixed H_2/H_∞ Controller Design

The H_2 and H_∞ control strategies based on the LMI were derived, respectively. Now we will combine these two constraints into one design expression. The mixed H_2/H_∞ control problem is to minimize the H_2 norm of T_{z2w}over all state feedback gains k such that what also satisfies the H_∞ norm constraint. Mixed H_2/H_∞-synthesis with regional pole placement is one example of multi-objective design addressed by the LMI. The control problem is sketched in Fig. 4. The output channel z_∞ is associated with the H_∞ performance while the channel z_2 is associated with the H_2 performance (LQG aspects)[25].

A. System Representation

Figure 4. shows the standard representation of the robust output-feedback control block diagram where P(s) is the plant and K(s) represents the controller that is usually of the same order as the plant let:

$$P(s): \begin{cases} \dot{x} = Ax + B_1 w + B_2 u \\ z_\infty = C_\infty x + D_{\infty 1} w + D_{\infty 2} u \\ z_2 = C_2 x + D_{21} w + D_{22} u \\ y = C_y x + D_{y1} w + D_{y2} u \end{cases} \quad (15)$$

$$K(s): \begin{cases} \dot{\zeta} = A_K \zeta + B_K y \\ u = C_K \zeta + D_K y \end{cases} \quad (16)$$

and let

$$CL:\begin{cases}\dot{x}_{cl} = A_{cl}x_{cl} + B_{cl}w \\ z_{\infty} = C_{cl\infty}x_{cl} + D_{cl\infty}w \\ z_{2} = C_{cl2}x_{cl} + D_{cl2}w\end{cases} \qquad (17)$$

be the corresponding closed-loop state-space equations with $x_{cl} = [x \quad \zeta]^t$

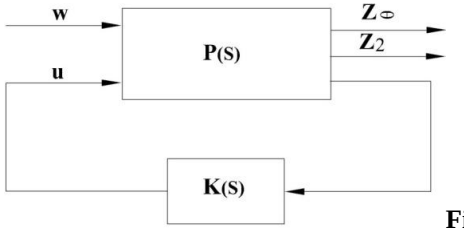


Figure 4. Output feedback block diagram.

B. Pole-Placement Technique[25]

The concept of LMI region [26]. is useful to formulate pole-placement objectives in LMI terms. They are convex subsets D of the complex plane C characterized by

$D = \{z \in C \text{ such that } f_D(z) = L + Mz + M^t z < 0\}$ where M and L=L^t are fixed real matrices,

$L = L^t = [\lambda_{ij}]$ and $M = [\mu_{ij}]$ where $1 \le i,j \le m$

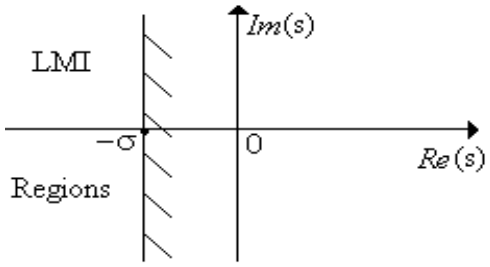


Figure 5. Pole-placement region.

z= x+iy a complex number. More practically, LMI regions include relevant regions such as sectors, disks, conics, strips, etc., as well as any intersection of the above. Only a shift in the left-hand side plane, shown in Figure 5. is considered. Its characteristic function with Re(z) = x<-□ , is $f_D(z) = z + \bar{z} + 2\sigma < 0$, thus L=2□ , M=1.

From a Theorem in [25], the pole-placement constraint is satisfied if and only if there exists X_p >0 such that $[\lambda_{ij}X_p + \mu_{ij}A_{cl}X_p + \mu_{ji}X_pA_{cl}^t] < 0$ with $1 \le i,j \le m$

V. Results and Discussions

The digital simulation results are obtained using MATLAB Platform. The aim of a suspension system for automotive applications is to isolate the passengers or load from vibrations generated by uneven roads. The suspension working space must not be too large because the working space for the suspension mechanism is limited. In this paper some paramters were investigateted its effect on the suspension systems performance. With H_∞ controller technique it is observed that parameter gama **(**□ *)* has most significant effect on the dynamic performance firstly the effect of the tuning variables of the LMI algorithms on the suspension performance are shown in table 2. which illustrates the root mean square value (RMS) of suspension working space, body acceleration, and dynamic tire load. It is clear that the parameter gama □ has a large effect on the system dynamic responses. From the table it can be noted that the optimal value of gama is 105. Figure 6 illustres the effect of gama on the suspension working space.

Table 2. Effect of the parameter (□ □ □on the Suspension Performance.

Case No.	□	SWS (m)	BAC (m/s^2)	WAC (m/s^2)	DTL (N)
1	50	0.0181	2.88	5.326	878.6
3	105	0.0087	1.48	6.340	691.9
4	120	0.0197	2.6258	5.6331	809.8
5	150	0.0207	2.6312	5.652	812

- SWS : Suspension working space (m).
- BAC : Body acceleration (m/s^2).
- WAC : Wheel acceleration (m/s^2).
- DTL : Dynamic tire load (N).

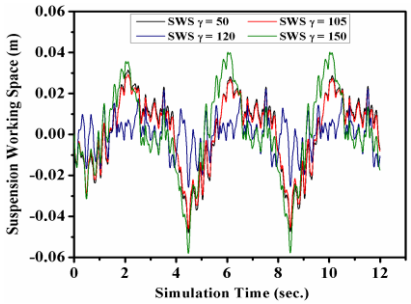


Figure 6. Suspension working space with different gama γ values

With H_2/H_∞ technique there are some parameters effected on the performance of suspension system and investigated as shown in Table 3. Shows the different values of the tuning variable of H_2/H_∞.

Table 3. RMS of suspension system performance.

Parameters	Values	SWS (m)	BAC(m/s2)	DTL (N)
□	100	0.0117	2.14	616.4
	200	0.0085	1.82	594.4
□	3	0.0117	5.45	1635
	50	0.0085	1.82	594.4
□	2	0.0117	2.72	799.5
	20	0.0085	1.82	594.4
□	0.1000	0.0117	1.9	1635
	0.0001	0.0085	1.82	594.4

Table 4. RMS Analysis random excitation.

System		SWS (m)	BAC (m/s^2))	DTL (N)
Passive System		0.0176	3.09	938
Semi Acti	H_∞	0.0150	2.72	855
	H_2/H_∞	0.0069	1.48	711.8
Improvement%		**54**	**46**	**17**

The vehicle body acceleration is an important index while evaluating vehicle ride comfort. The proposed of active suspension system with mixed H_2/H_∞ controller is effective in reducing vehicle body acceleration. Table 3. shows the RMS values of suspension working space, body acceleration, and dynamic tire load. The simulation results show that the vehicle body acceleration reduced from 2.72 m/s2 to 1.48 m/s2, and the suspension working space reduced from 0.015 mm to 0.0069 mm, and the tire dynamic load reduced from 855 N to 711.8 N, so the improvement are (54% , 46% ,and 17% respectively).

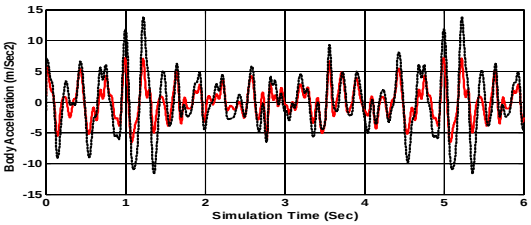


Figure 7. Body acceleration.

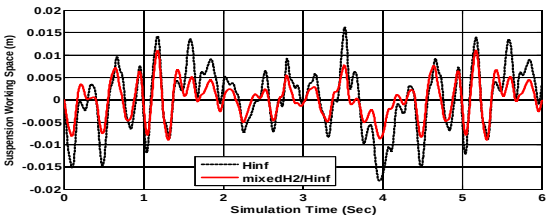


Figure 6. Suspension working space.

Simulation results indicate that the proposed of semi- active suspension system proves to be effective in improving riding comfort and holding ability as shown from Figures (7-9) which illusterate the comparsion between the two methods of controls.

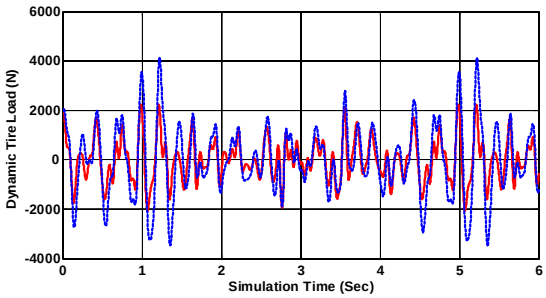


Figure 8. Dynamic tire load.

VI. Conclusions

A qurter car model (2DOF) is devolped in order to investigate the influnce of different control techneques on the suspension system performance, H_∞ and H_2/H_∞ . From the simulations results we can clearly see that the sime-active controlled suspension with both H_∞ and H_2/H_∞ control techniques offers a much better suspension performances than the passive system as compared in time domain, and a comparison between the two techniques H_∞ and H_2/H_∞ was done. It can noted that the mixed technique control method is offer a much better perfomance than H_∞ technique.

References

[1] D. Karnopp, M.J.Crosby, R.A. Harwood, Vibration control using semi- active force generators, ASME Journal of Engineering for Industry 96 (2)(1974) 619–626.

[2] R.S. Sharp, D.A. Corolla, Road vehicle suspension system design-a review, Vehicle System Dynamics 16 (3)(1987)167–192.

[3] N.A.Jalil,Comparativestudyandanalysisofsemi-activevibration-controlsystems, Journal of Vibration and Acoustics 124 (2002)593–605.

[4] S.B. Choi, M.-H.Nam,B.- K.Lee, Vibration control of a MR seat damper for commercial vehicles, Journal of Intelligent Material Systems and Structures 11 (12)(2000)936–944.

[5] M.Ahmadian,C.A.Pare,Aquarter-carexperimentalanalysisofalternativesemiactivecontrolmethods, Journal of Intelligent Material Systems and Structures 11 (8)(2000)604–612.

[6] H.Li,M.Liu,J.Li,X.Guan,J.Ou,VibrationcontrolofstaycablesoftheShandongBinzhouyellowriverhighwaybridgeusingmagnetorheologicalfluid dampers, Journal of Bridge Engineering 12 (4)(2007)401–409.

[7] E.A.Johnson,G.A.Baker,B.F.SpencerJr.,Y.Fujino,Semiactivedampingofstaycables, Journal of Engineering Mechanics 133 (1)(2007)1–11.

[8] S.J. Dyke, B.F. Spencer-Jr., M.K. Sain, J.D. Carlson, An experimental study of mr dampers for seismic protection, Smart Materials and Structures 7 (5)(1998) 693–703.

[9] Gordon, T. J., Marsh, C., and Milsted, M. G., "A Comparison of Adaptive LQG and Non-linear Controllers for Vehicle Suspension Systems," Veh. Syst. Dyn., 20, 1991, pp. 321–340.

[10] Alleyne, A., and Hedrick, J. K., "Non-linear Adaptive Control of Active Suspensions," IEEE Trans. Control Syst. Technol., 3(1), 1995,pp. 94–101.

[11] Ben Gaid, M., Cela, A.,Kocik, R., "Distributed control of a car suspension system," COSI - ESIEE - Cit´e Descartes,

[12] P.M. Anderson and A. A. Fouad, Power System Control and Stability, IEEE Press, 1993.

[13] O. I. Elgerd, Electric Energy System Theory, An Introduction, McGraw-Hill, 1982.

[14] P.A.W. Walker and O.H. Abdallah, "Discrete Control of an A.C. Turbo generator by Output Feedback," Proceedings of the IEE, Control & Science, Vol. 125, No. 9, Oct. 1978, pp. 1031-38

[15] T.C. Yang, "Applying H_∞ optimization methods to power system stabilizer design parts 1 &2 ," Int. J. Elect. Power Energy Syst., vol. 19,n0. 1,pp.29-43,1997.

[16] R. Asgharian, " A robust H_∞ power system stabilizer with no adverse effect on shaft tensional modes," IEEE Trans. Energy Conversion, vol. 9, no. 3, 1994, pp.475-481

[17] C. Scherer, P. Gahinet, and M. Chilali, "Multi-objective output-feedback control via LMI optimization," IEEE Trans. Automat. Contr., vol. 42, pp. 896.911, 1997.

[18] A. Bensenouci and A.M. Abdel Ghany, "Mixed H_∞/H_2 with Pole-Placement Design of Robust LMI-Based Output Feedback Controllers for Multi-Area Load Frequency Control" The IEEE International Conference on Computer as a Tool, EUROCON 2007, Warsaw, Poland, September 9-12, 2007.

[19] M. Saeed Badran and A. S. Emam " H_∞ and Mixed H2/H_∞ with Pole-Placement Design Via ILMI Method for Semi-Active Suspension System"

[20] Boyd, S., L. El Ghaoui, E. Feron, V.Balakrishnan, Linear Matrix Inequalities in Systems and Control Theory, SIAM books, Philadelphia, 1994.

[21] R.S. Sharp and S.A. Hassan, "On the performance capabilities of active automobile suspension systems of limited bandwidth", Vehicle System Dynamics, 16:213–225, 1987.

[22] A. Rowan ," Application of Electronically Controlled Suspension Systems to Military Vehicles," M.Sc.Thesis, Faculty of engineering-Mattaria, Helwan University, Cairo, Egypt 2004.

[23] A. Bensenouci, and A.M. Abdel Ghany, "Performance Analyses and Comparative Study of LMI-Based Iterative PID Load- Frequency Controllers of a Single-Area Power System," WSEAS (World Scientific and Engineering Academy and society) on Power Systems Journal, Issue 2, Vol.5, April 2010, pp.85-97.

[24] A. M. Abdel Ghany and A. Bensenouci, "Robust Output Feedback Control Design using H_∞/LMI and SA/Lead-Lag for an ANN-Adaptive Power System Stabilizer," The 15th, Mediterranean Conference on Control and Automation, MED'07, June 27-29, 2007, Athens, Greece

[25] M. Chilali and P. Gahinet, "H_∞ design with pole placement constraints: An LMI approach," IEEE Trans. Automat. Contr., vol. 41, no. 3, March 96, pp. 358–67.

[26] A. S. Emam, A.M. Abdel Ghany, Enhancement of Ride Quality of a Quarter Car by Using H_∞ Design of a Robust LMIOutput Feedback Controller.Ain shams journal of mechanical engineering, Vol.2, October,2010, pp.35-43.

Analysis of Deep Beam Using Cast Software and Compression of Analytical Strain with Experimental Strain Results

Kale Shrikant M.[1] Prof.Patil.S.S.[2] Dr. Niranjan B.R.[3]

Abstract:
Analysis of deep beam by using CAST software based on strut and tie method. as per ACI 318-05(Appendix -A).Design and casting of several deep beam using STM. Testing of deep beams in heavy structures laboratory for two point loading condition. Measurement of strain, load and deflection under controlled condition. Comparison of analytical flexure strain with experimental results.

Keywords: Analysis of deep beam, CAST (computer aided Strut and Tie) Software, Deep beam, Strut and tie Method(STM), Strain measurement, Strain gauge, Experimentation.

I Introduction

Strut-and-tie modeling (STM) is an approach used to design discontinuity regions (D-regions) in reinforced and prestressed concrete structures. A STM reduces complex states of stress within a D-region of a reinforced concrete deep beam into a truss comprised of simple, uniaxial stress paths. Each uniaxial stress path is considered a member of the STM. Members of the STM subjected to tensile stresses are called ties and represent the location where reinforcement should be placed. STM members subjected to compression are called struts. The intersection points of struts and ties are called nodes. Knowing the forces acting on the boundaries of the STM, the forces in each of the truss members can be determined using basic truss theory. Strain obtained analytical by software was compared with strain recorded experimentally.

II Computer Aided Strut-And-Tie (CAST) Analysis

A research programme was recently conducted to advance the STM for overcoming the aforementioned challenges. In addition to making the design and analysis process using the STM more efficient and transparent, the research aimed to extend the basic use of the STM from a design tool to an analysis tool that can be used for evaluating member behavior and thereby making it possible to evaluate/validate/extend design code provisions (e.g. dimensioning rules and stress limits) of deep beam. By using a computer-based STM tool called CAST (computer aided strut-and-tie) was developed by Tjhin and Kuchma at the University of Illinois at Urbana-Champaign (2002). This tool is the subject of this paper. CAST facilitates the instruction activities for analysis of reinforced concrete deep beam by STM. This paper considers D-regions that can be reasonably assumed as plane (two-dimensional) structures with uniform thickness and the state of stress is predominantly plane (plane stress condition). Two point loading acting on the D-regions is limited to static monotonic, but can be extended to account for the degradation effects of repeated loading. Only strut-and-tie models that consist of unreinforced struts and non-prestressed reinforcement ties are considered. The primary failure modes of the D-regions are the yielding of ties, crushing of struts or nodal zones and diagonal splitting of struts. Failures due to reinforcement anchorage and local lateral buckling are not considered.

III Analytical modeling of RC Deep Beam

The strut-and-tie model was analyzed using CAST software. Experimental and analytical deep beam model was having 0.7 m length, 0.4 m depth and 0.15 m thick. The materials properties obtained from material tests will used for concrete and reinforcing steel in the models. By doing so, the strength reduction factor φ was set to unity. The supports where modeled as a vertical reaction on the left support and a vertical and horizontal reaction on the right support The software's capacity prediction feature was used to estimate the capacity using the provided steel reinforcement, concrete struts and nodal zones.

Additionally, the software has a feature that allows analysis of the nodes to ensure that geometry and stress limits are not exceeded. The estimated capacity according to CAST, the failure would occur by yielding of the diagonal tie. This is desirable in STM because it allows the member to fail in a ductile manner as the reinforcing bars yield first before failure, as opposed to brittle failure of the concrete strut

IV Strut-and-Tie Method: Design Steps

The design process using the Strut-and-Tie Method involves steps described below. These steps are illustrated using the design example of a deep beam.

1. Define the boundaries of the D-Region and determine the boundary forces (the ultimate design forces) from the imposed local and sectional forces. Boundary forces include the concentrated and distributed forces acting on the D-Region boundaries. Boundary forces can also come from sectional forces (moment, shear, and axial load) at the interface of D- and B-Regions. Body forces include those resulted from D-Region self-weight or the reaction forces of any members framing into the D-Region.
2. Sketch a Strut-and-Tie Model and solve for the truss member forces.
3. Select the ordinary reinforcing steel and prestressing steel that are necessary to provide the required Tie capacity and ensure that they are properly anchored in the Nodal Zones.
4. Evaluate the dimensions of the Struts and Nodes such that the capacity of all Struts and Nodes is sufficient to carry the truss member forces.
5. Provide distributed reinforcement to ensure ductile behavior of the D-Region.

Since equilibrium of the truss with the boundary forces must be satisfied (step 2) and stresses everywhere must be below the limits (step 3 and 4), one can see that the Strut-and-Tie Method is a lower-bound (static or equilibrium) method of limit analysis.

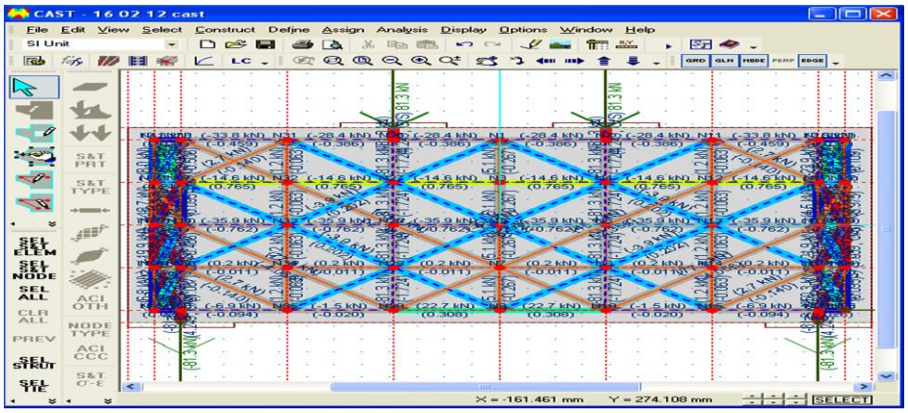


Figure 1. Forces in members by CAST analysis

Above figure shows the forces in strut and tie developed in deep beam using CAST software similarly strain and stress are obtained in graphical as well as tabulated form.

V Experimental work

In experimental investigation of deep beam we have taken same size of deep beam of total length 700 mm , depth 400 mm and width 150 mm. which were casted in concrete technology labarotaty and curring was carried out for 28 days . M25 gread of concrete were used for deep beam. For application of load we have used 1000 kN capacity hydraulic heavy testing machine. To measure deflection dial gauge where placed at central position of bottom of deep beam. to measure strain along mid span we have used strain gauge at equally spacing from top to bottom.

Figure 2. Test setup for Deep beam

Figure 3.Strain measurement of deep beam

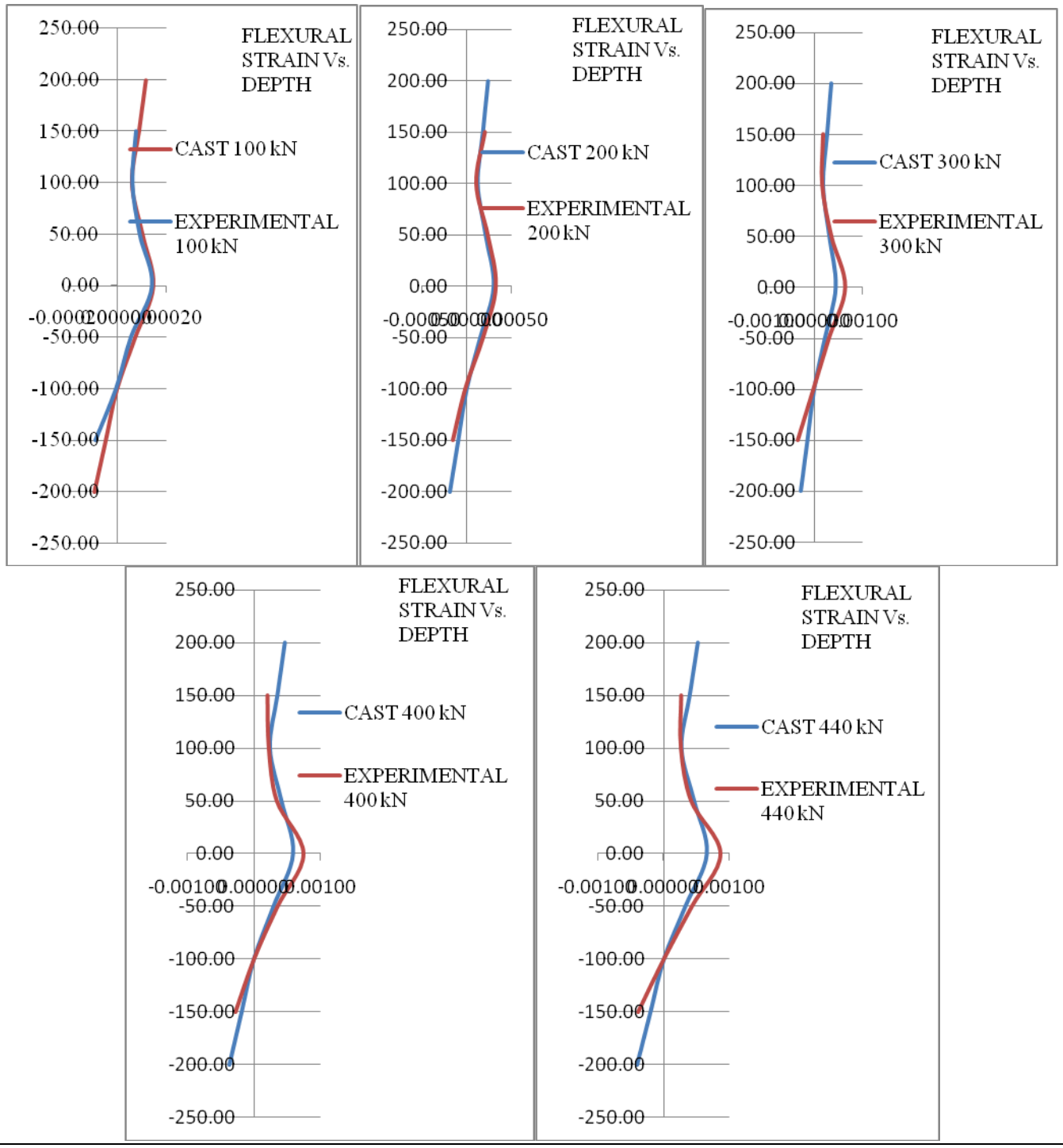
Table 1. Experimental Strain (in mm)

Load → / Depth ↓	100 kN	200 kN	300 kN	400kN	440 kN
150	0.00008	0.00011	0.00022	0.00029	0.00032
100	0.00006	0.00011	0.00026	0.00036	0.00045
50	0.00009	0.00023	0.00052	0.00096	0.00121
0	0.00014	0.00035	0.00072	0.00140	0.00165
-50	0.00005	0.00019	0.00038	0.00080	0.00092
-100	-0.00001	-0.00001	-0.00002	0.00000	0.00000
-150	-0.00009	-0.00015	-0.00035	-0.00038	-0.00039

Table 2. Analytical strain (in mm)

Load	100 kN	200 kN	300 kN	400kN	440 kN
Depth					
200.00	0.00005	0.00010	0.00014	0.00018	0.00020
150.00	0.00005	0.00010	0.00015	0.00021	0.00023
100.00	0.00006	0.00009	0.00016	0.00023	0.00026
50.00	0.00012	0.00021	0.00043	0.00071	0.00084
0.00	0.00019	0.00033	0.00070	0.00120	0.00142
-50.00	0.00011	0.00018	0.00038	0.00065	0.00077
-100.00	0.00002	0.00003	0.00006	0.00010	0.00012
-150.00	-0.00002	-0.00006	-0.00009	-0.00011	-0.00012
-200.00	-0.00007	-0.00015	-0.00024	-0.00032	-0.00036

In above table 1, shows experimental strain at mid span of deep beam at definite incremental loading at various depth to understand the nature of strain. Analytical strain obtained By using cast software are tabulated in table 2.



Graph 1. Comparison of Flexural strain vs. depth

Above graph shows the comparison of experimental strain and analytical strains recorded at 100 kN, 200 kN, 300 kN, 400kN, 440kN. Experimentally deep beam designed by strut and tie method for two point loads of 50 kN (2 X 50 kN =100 kN) deep beam failed at 440 kN (220 kN each).Experimental and analytical results are almost same up to 200 kN (two point load each of 100 kN)

6. Conclusion

1. strut and tie method is useful to understand flow of stress.
2. CAST software gives good results which matches with experimental results.
3. Strain in deep beam is non linear along its vertical axis.
4. At design load experimental as well as analytical strain at mid span of deep beam matches with each other. With further increase in load, experimental strain goes on increasing at bottom and mid depth of deep beam. (Reference Graph 1)

References

[1]. IS : 456–2000. 'Plain and Reinforced Concrete — Code of Practice'. Bureau of Indian Standards, Manak Bhavan, New Delhi, India.

[2]. ACI 318–05. 'Building Code Requirements For Structural Concrete and Commentary' American Concrete Institute, Detroit, USA.

[3]. J Schliach and K Schafer.' Design and Detailing of Structural Concrete using Strut–and–Tie Models'. The Structural Engineer,vol 69, 1991, 113.

[4]. AASHTO, "AASHTO LRFD Bridge Specifications for Highway Bridges" (2001 Interim Revisions), American Association of Highway and Transportation Officials, Washington, D.C., 1998.

[5]. Mr. Varghese and Mr. Krishnamoorthy,(1966),Streingth and Behaviour of Deep Reinforced Concrete Beams, Indian Concrete Journal, 104-108.

[6]. Matamoros and Wong ,(2003), Design of simply supported Deep beam using strut -and –tie models,ACI Structural journal,704-712.

[7]. Quintero-Febres, Parra-Montesinos and Wight ,(2006), Strength of Struts in deep Concrete Members Designed Using Strut and Tie Method, ACI Structural journal, 577-586.

[8]. Park and pauly, Reinforced Concrete Structures,A wiely-Interscience Publication.

[9]. P. Nagarajan, Dr.T.M.M.Pillai and Dr.N.Ganesan, (2007), Design of Simply Supported Deep Beams using IS 456:2000 and Strut and Tie Method, IE (I) Journal-CV, 38-43.

[10]. Michael D. Brown, Cameron L. Sankovich, Oguzhan Bayrak,James O. Jirsa,John E. Breen, Sharon L. Wood, (2006), the technical report on Design for Shear in Reinforced Concrete Using Strut -and-Tie Models.

[11]. James k Wight and Gustavo J.Parra-Montesinos,(2003), Strut-And-Tie Model For Deep Beam Design, Concrete international,63-70.

[12]. Perry Adebar and Zongyu Zhou,(1993),Bearing Strength of Compressive Struts Confined by Plain Concrete,ACI Structural Jouranal,534-541.

[13]. Chung, W., and Ahmad, S.H., 1994, "Model for Shear Critical High-Strength Concrete Deep Beams," ACI Structural Journal, Vol. 91, No. 1, pp. 31-41.

[14]. Laupa, A., Siess, C.P., and Newmark, N.M., 1953, "The Shear Strength of Simple-Span Reinforced Concrete Beams without Web Reinforcement,"Structural Research Series No. 52, University of Illinois, Urbana.

Adopting Trusted Third Party services for Multi-level Authentication accessing cloud

Vivekananth.P (1st) Dr.Ritish Khanna (2nd)
[1] Research Scholar
Department of Computer Science
CMJ University Shillong Meghalaya
[2] Faculty Isara institute of management and professional studies Delhi

Abstract*:*
Cloud computing is an emerging, on-demand and internet-based technology. It provides variety of services over internet such as, software, hardware, data storage and infrastructure. This technology has been used by worldwide customers to improve their business performance. At present, authentication is done in several ways: such as, textual, graphical, bio-metric, 3D password and third party authentication. The framework have seven processes, including: processes-selecting relevant critical areas, strategy and planning, risk analysis, risk assessment, risk mitigation, assessing and monitoring program, and risk management review. To develop the dynamic Trusted Third Party auditing key role of public auditability in the cloud storage security API. The main objective of this research question is to understand information security threats relevant in cloud computing. Considering the security and efficiency in data sharing, we introduced the trusted third party (TTP). In this paper aims at identifying SLA based information security metrics relevant in cloud computing. In this research investigates different Objects relevant in cloud policy models and each of them can be used as a component in the proposed framework. A role represents a specific function within an organization and can be seen as a set of actions or responsibilities associated with this function.

Keywords— Secure Data collection, Cloud Storage Services, Secure distributed storage, Third Party Auditor, Insider Access, Service Level Agreement

I Introduction

Cloud services, which are deployed as self-contained components, are normally partial solutions that must be composed to provide a single virtualized service to Cloud consumers. This composition of services should be carried out in a dynamic and automated manner to promptly satisfy consumer requirements. Data storage correctness or some time more generally referred as data integrity verification is one of chief Cloud security problems [1]. Data can be altered by unauthorized entity without intimating to data owner.

Cloud Computing provides an optimal infrastructure to utilise and share both computational and data resources whilst allowing a pay-per-use model, useful to cost-effectively manage hardware investment or to maximise its utilisation. Cloud computing also offers transitory access to scalable amounts of computational resources, something that is particularly important due to the time and financial constraints of many user communities [2].

Data storage correctness schemes can be classified into two categories (a) Without Trusted Third Party (TTP) and (b) with TTP, based on who makes the verification. In case of TTP, an extra Third Party Auditor (TPA), some time in form of extra hardware or cryptographic coprocessor is used. This hardware scheme provides better performance due to dedicated hardware for the auditing process but has some drawbacks such as single TTP resulting into bottleneck in the system, mutually agreeing on a common TTP where there are thousands of users across the globe. Due to such kind of reasons, we prefer an approach where the functionalities of TPA is integrated in form of client application and the application can be downloaded by cloud user from cloud server [3]. This client application provides all the cryptographic functionalities to achieve the goals of integrity, authentication and confidentiality. As this is a software approach, the performance of the overall system may not be comparable to dedicated hardware kind of TTP alternatives. To improve performance, we emphasize offline execution of computationally costly cryptographic algorithms [4].

The auditing community is aware that current practices for auditing cloud environments are inadequate. As compliance grows in importance, enterprises implementing clouds need to satisfy their auditors' concerns; especially since creating an identity for an individual virtual machine and tracking that virtual machine from creation to deletion creates challenges for even the most mature virtualized environments[5][6].

The clouds have different architecture based on the services they provide. The data is stored on to centralized location called data centers having a large size of data storage. The data as well as processing is somewhere on servers. So, the clients have to trust the provider on the availability as well as data security. The SLA is the only legal agreement between the service provider and client [7]. The only means the provider can gain trust of client is through the SLA, so it has to be standardize.

A key part of a Service Level Agreement deals with monitoring and measuring service level performance. Essentially, every service must be capable of being measured and the results analysed and reported. The benchmarks, targets and metrics to be utilized must be specified in the agreement itself. The service performance level must be reviewed regularly by the two parties.

Data Protection

Data stored in the cloud typically resides in a shared environment collocated with data from other customers. Organizations moving sensitive and regulated data into the cloud, therefore, must account for the means by which access to the data is controlled and the data is kept secure.

Any progress must first occur in a particular domain in our work focuses on an important class of widely used applications that includes e-mail, personal financial management, social networks, and business tools such as word processors and spreadsheets. The following criteria define this class of applications:

Provide services to a large number of distinct end users, as opposed to bulk data processing or workflow management for a single entity;

Use a data model consisting mostly of sharable units, where all data objects have access control lists (ACLs) with one or more users; and Developers could run the applications on a separate computing platform that encompasses the physical infrastructure, job scheduling, user authentication, and the base software environment, rather than implementing the platform themselves.

Cloud Data Protection as a Service

Currently, users must rely primarily on legal agreements and implied economic and reputation harm as a proxy for application trustworthiness. As an alternative, a cloud platform could help achieve a robust technical solution by making it easy for developers to write maintainable applications that protect user data in the cloud, thereby providing the same economies of scale for security and privacy as for computation and storage; and enabling independent verification token of the secure data seeds with platform's operation and the runtime state of applications on it, so users can gain confidence that their data is being handled properly. With multiple providers and competition, users can regain control over their data. With a clear, universal application programming interface (API) to token for secure data seeds cloud services and the user's authorization, Cloud Data Protection as Service providers would be able to access and manipulate the data on another Cloud Data Protection as service provider. Such an API would also enable pervasive systems to run seamlessly between providers and allow interaction between users on different providers.

In a cloud setting, the unit of access control is typically a sharable piece of user data for example, a document in a collaborative editor. Ideally, the system offers some analogous confinement of that data, restricting its visibility only to authorized users and applications while allowing broad latitude for what operations are done on it [7]. This can make writing secure data seeds systems easier for programmers because confinement makes it more difficult for buggy code to leak data or for compromised code to grant unauthorized access to data.

II. RELATED WORK

Cloud infrastructure management networks are how cloud providers access the infrastructure and manage the different components within that infrastructure. Only authorized administrators should have access to this network because control of the management interfaces of the individual virtualization hosts allows for complete control of all of the virtual machines on that host. Root access on this interface is analogous to having the keys to a physical rack of servers within a data center. Administrator access to the central management console that manages all of the different virtualization hosts within the cloud is analogous to having the keys to the datacenter and every rack within that datacenter [9]. Therefore, protection of these interfaces is of paramount importance, and a customer should never need direct access to any of the systems within this network.

The reason for isolating this traffic is two-fold. First, both VMware VMotion and IP storage traffic need very fast data rates for optimal performance. Furthermore, traffic travels over the network in clear text and is susceptible to an attacker sniffing sensitive information off the network. By fully isolating this network, an attacker would need physical access to this network to have any chance of successfully compromising this data.

With so much remote execution, cloud computing requires robust credential management that enable secure logins to multiple cloud services from multiple devices in a seamless manner. The password schemes currently employed are a burden on users and have practically forced users into poor practices. Generally, users can remember a small number of passwords, yet each Web resource generally requires users to develop a unique set of credentials [10]. Services such as OpenID, which allow users to have a single set of credentials for multiple sites, are powerful, but may be inappropriate for sensitive institutions such as banks or government sites. Users may instead be able to use one-time-password devices, but they would need to have a unique device for each remote site to prevent one site from being able to use the credentials to authenticate to another.

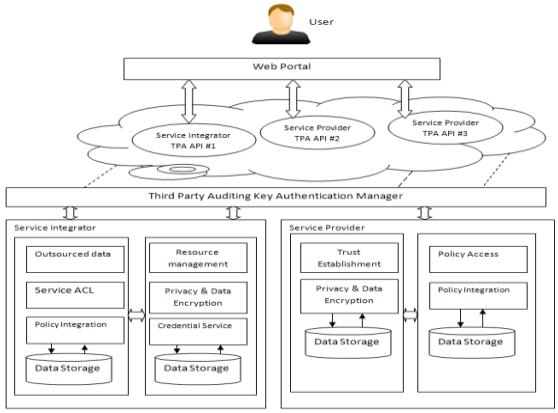
Even though these cloud computing components and characteristic provide compelling solutions to IT problems and many advantages, cloud computing is not risk-free or completely secure. Management is responsible for taking care of security risks to protect systems and data. Governance, risk and control of cloud computing are therefore critical in the performance of any assurance management process. Governance is enforced through the implementation of policies and procedures. These policies and procedures should be based on best practices and should be aligned between business and IT objectives. Risk identification and analysis is important to priorities the implementation (extent and time frame) of governance and controls, as well as to establish scope for reviewing or auditing cloud computing environments [11][12]. Based on the identification and analysis of risks, controls should be designed and implemented to ensure that necessary actions are taken to address risks and to achieve business and IT objectives [13]. This research aims to provide some guidelines to assist management with the identification of risks and recommendations for the mitigation of cloud computing security risks.
The author presented a P2P backup scheme in which blocks of a data file are dispersed across m+k peers using an (m+k,m)-erasure code. Peers can request random blocks from their backup peers and verify the integrity using separate keyed cryptographic hashes attached on each block. Their scheme can detect data loss from free riding peers, but does not ensure all data is unchanged.[13] proposed to verify data integrity using RSA-based hash to demonstrate uncreatable data possession in peer-to peer file sharing networks. However, their proposal requires exponentiation over the entire data file, which is clearly impractical for the server whenever the file is large. [14] Author proposed allowing a TPA to keep online storage honest by first encrypting the data then sending a number of precomputed symmetric-keyed hashes over the encrypted data to the auditor [15].

III. Implementation of Auditing and Data Scalability Framework

We implement the scalability framework we have defined for composite Web services. We proposed a new model–driven methodology for the security testing of cloud environments, to support batch auditing for TPA upon delegations from multi-user. We also proposed new scheme that enables the data owner to delegate tasks of Cloud storage data file encryption and trusted user secret key update to cloud servers without disclosing security data contents or user access privilege information. The trusted third party auditing process will bring in no new vulnerabilities towards user cloud storage data privacy. Our new method combined the secret key based dynamic secure authenticator with random masking to achieve the privacy-preserving public cloud data auditing system. We use a TPM-based trusted cloud storage SSL based third party auditing with secured key authentication

With the help of the loader, our trusted hypervisor finally takes over the machine. Users can use TPM-based attestation to verify the software stack running on the physical machine. The code is divided in three main parts: (1) data unit manager that stores the definition and information of the data units that can be accessed.

During our experiments we observed a significant number of read operations on individual clouds that could not be completed due to some error. The first thing that can be observed from the table is that the number of measurements taken from each location is not the same. Cloud storage concern the user does not have control over data until he has been gain access. To provide control over data in the cloud data-centric security is needed. Before accessing the data it should satisfy the policy rules already defined. So cloud should enforce this scheme by using cryptographic approaches.



In this Paper use the RSA algorithm as a basis to provide Cloud data-centric security for shared storage remote cloud data:

C Select two prime numbers

C Calculate n = p*q.

C Calculate f(n) = (p-1)(q-1)

C Select e such that e is relatively prime to f(n) and less than f(n).

C Determine d such that de congruent modulo 1 (mod f(n)) and d<f(n).

C Public key = {e, n}, Private key = {d, n}

C Cipher text c = message e mod n

C Plain text p = ciphertext d mod n

Implementation of dynamic packet filtering

Restriction of all inbound and outbound traffic to that information specified in the documented and maintained list of ports and services

Prevention of direct wireless access to the cloud infrastructure

Prevention of internal address direct access to external interfaces

Install perimeter firewalls between confidential and configuration data and external interfaces where supported by the cloud host.

Installation of personal firewall software, solutions on external devices, such as computers, mobile computers, mobile devices, and so on, that interface with the cloud environment where supported by your cloud host.

Implement IP masks to prevent internal systems from being presented and identified to external entities.

Implement a firewall to isolate confidential information, and ensure that all confidential information is stored behind the firewall. During retrieval of data, it is decrypted after checking the generated private key with existing private key

Before providing managed services, a security-as-a-service provider must overcome many concerns. Security services must be independent of any platform, adaptable to constantly changing environments, and supportive of a virtualized environment. To meet all these seemingly divergent requirements, security as a service must maintain compatibility with the service offerings provided in the virtualized environment.

Input: (i) call_for_proposals from CAs or other BAs

Output: (i) Instantiation of a Request_evaluator behavior

```
BlockReceive(call_for_proposals(Reqi))
if (not visitedFor(Reqi)) then
Prepare and Send Proposal
BlockReceive(reply, timeout)
if (reply = accept_proposal) then
Instantiate a Request_evaluator(Reqi) behavior
Else
Start over
else
Send refuse message
Start over
```

Users are then made members of roles, thereby acquiring the roles' authorizations. User access to resources is controlled by roles; each user is authorized to play certain roles and, based on his own role he can perform accesses to the resources and operate them correspondingly. As a role organizes a set of related authorizations together, it can simplify the authorization management.

Key Generation Algorithm

Choose a and b: two distinct prime numbers.
Compute m = a b, Where m is used as the modulus for public and private keys.
Compute Φ (m) = (a-1)(b-1), Where Φ is function.
Choose an integer E such that, $1 < E < \Phi(m)$ and common divisor of (E, Φ (m)) = 1.
Determine D = 1/E mod Φ (m).
All the above values of public key and private key must be kept secret.

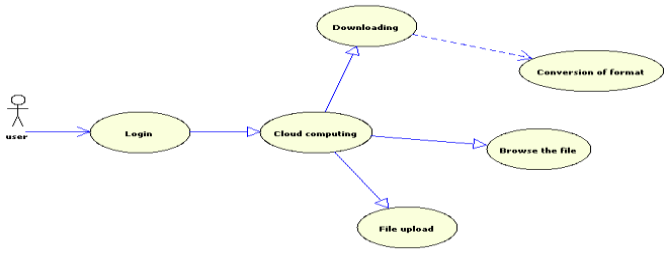
Encryption Algorithm

Sender A transmits her public key (m, E) to recipient B for the process of encryption data.
Represent the plaintext message as a positive integer n.
Computes the cipher c = nE (mod m).
Sends the cipher text c to recipient B

When the data owner redefines a certain set of attributes for the purpose of user revocation, he also generates corresponding proxy re-encryption keys and sends them to Cloud Servers. Cloud Servers, given these proxy re-encryption keys, can update user secret key components and re-encrypt data files accordingly without knowing the underlying plaintexts of data files. This enhancement releases the data owner from the possible huge computation overhead on user revocation. The data owner also does not need to always stay online since Cloud Servers will take over the burdensome task after having obtained the PRE keys.

Implementation Steps

The simulator is written in Java and runs on a Windows Vista Core 2 CPU 2.16 GHz machine. The set-up of parameters of message exchanges and network characteristics are as follows. The cloud deployment consists of 20 server nodes that are potentially able to host Web services. These 20 nodes are connected by a network such that the network has about 10%-50% network connectivity. In the simulation, we simulate the network of the cloud by randomly connecting one server node to the other in the network with a certain probability, which is equal to the network connectivity. For example, if the network connectivity is 20%, each node is directly connected to 30% of the other cloud nodes. We assume that there are 20 different types of request messages and response messages respectively exchanged between pairs of services during their interactions in composite services.



To ensure the security and dependability for cloud data storage under the aforementioned adversary model, we aim to design efficient mechanisms for dynamic data verification and operation and achieve the following goals: (1) Storage correctness: to ensure users that their data are indeed stored appropriately and kept intact all the time in the cloud. (2) Fast localization of data error: to effectively locate the malfunctioning server when data corruption has been detected. (3) Dynamic data support: to maintain the same level of storage correctness assurance even if users modify, delete or append their data files in the cloud. (4) Dependability: to enhance data availability against Byzantine failures, malicious data modification and server colluding attacks, i.e. minimizing the effect brought by data errors or server failures. (5) Lightweight: to enable users to perform storage correctness checks with minimum overhead.

We analyse the security strength of our schemes against server colluding attack and explain why blinding the parity blocks can help improve the security strength of our proposed scheme. With the appropriate runtime extraction the user-interface is able to migrate from the user's desktop to their mobile device and back again, without losing state. The ability to store data either locally or remotely in a transparent fashion will greatly help address issues raised in our previous work in personal data storage on the Internet. The control and implementation of policies is a business imperative that must be met before there is general adoption of cloud computing by the enterprise. SOA is derived from architecture and a methodology. Since cloud computing is typically driven from the view of business resources that are needed, there is a tendency to ignore the architecture. The second area that SOA brings to cloud computing is an end-to-end architectural approach.

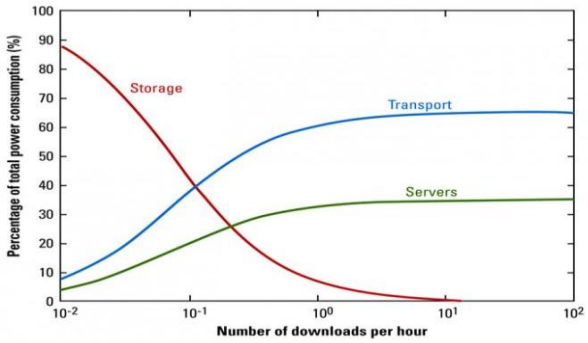
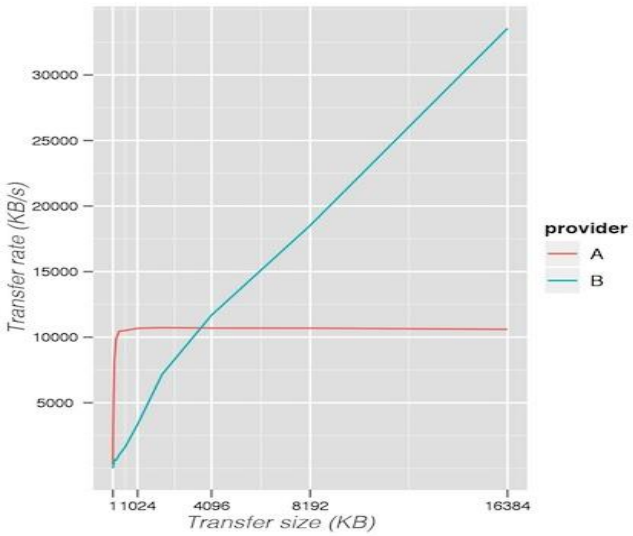


Figure shows that as the number of file transfers between the desktop and the cloud increases, the percentage of total power consumed in the transfer process increases. Says the report, "For a private cloud storage security service, at a download rates above one download per hour, servers consume 35%, storage consumes less than 7%, and the remaining 58% of total power is consumed in transport. These results suggest that transport dominates total power consumption at high usage levels for public and private cloud storage security services. The energy consumed in transporting data between users and the cloud is therefore an important consideration when designing an energy efficient cloud storage security service. Energy consumption in servers is also an important consideration at high usage levels. The percentage of total power consumed in servers is greater in private cloud computing than that in public cloud computing. In both public and private cloud storage security services, the energy consumption of storage hardware is a small percentage of total power consumption at medium and high usage levels. The proposed scheme is more suitable for the privacy-preserving of mass users.



The data is to be encrypted and compressed in multi-server. In encryption and compression the data that has to stored in a cloud can not be stored in a text format due to security reasons so it must be transformed into an encrypted format. The data also has to be compressed for secure transmission.

Conclusion

Finally, we investigated the problem of data security in cloud data storage, which is essentially a distributed storage system. To ensure the correctness of users' data in cloud data storage, we proposed an effective and flexible distributed scheme with explicit dynamic data support, including block update, delete, and append. We rely on erasure-correcting code in the file distribution preparation to provide redundancy parity vectors and guarantee the data dependability. By utilizing the homomorphic token with distributed verification of erasure coded data, our scheme achieves the integration of storage correctness insurance and data error localization, i.e., whenever data corruption has been detected during the storage correctness verification across the distributed servers, we can almost guarantee the simultaneous identification of the misbehaving server(s). Security design from the ground-up that promotes digitally signing each component-to-component call to allow the authorisation of all content executed by the user. When the data owner redefines a certain set of attributes for the purpose of user revocation, he also generates corresponding proxy re-encryption keys and sends them to Cloud Servers. Cloud Servers, given these proxy re-encryption keys, can update user secret key components and re-encrypt data files accordingly without knowing the underlying plaintexts of data files. When submitting their location information to the cloud, a blind user (and, in fact, any other user) could have security concerns that a malicious party could use this information to locate the user and harm or exploit the user for his own benefit.

REFERENCES

[1] C.Wang et al.,"Ensuring Data Storage Security in Cloud Computing," Proc. IWQoS '09, July 2009

[2] Q. Wang et al., "Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing," Proc. ESORICS '09, Sept. 2009, pp. 355–70

[3] C. Erway et al., "Dynamic Provable Data Possession," Proc. ACM CCS '09, Nov. 2009, pp. 213–22.

[4] C. Wang et al., "Privacy-Preserving Public Auditing for Storage Security in Cloud Computing," Proc. IEEE INFOCOM '10, Mar. 2010

[5] L. Carter and M. Wegman, "Universal Hash Functions," Journal of Computer and System Sciences, vol. 18, no. 2, pp. 143–154, 1979.

[6] J. Hendricks, G. Ganger, and M. Reiter, "Verifying Distributed Erasure coded Data," Proc. 26th ACM Symposium on Principles of Distributed Computing, pp. 139–146, 2007

[7] J. S. Plank and Y. Ding, "Note: Correction to the 1997 Tutorial on Reed-Solomon Coding," University of Tennessee, Tech. Rep. CS-03- 504, 2003.

[8] Q. Wang, K. Ren, W. Lou, and Y. Zhang, "Dependable and Secure Sensor Data Storage with Dynamic Integrity Assurance," Proc. of IEEE INFOCOM, 2009

[9] R. Curtmola, O. Khan, R. Burns, and G. Ateniese, "MR-PDP: Multiple- Replica Provable Data Possession," Proc. of ICDCS '08, pp. 411–420, 2008.

[10] D. L. G. Filho and P. S. L. M. Barreto, "Demonstrating Data Possession and Uncheatable Data Transfer," Cryptology ePrint Archive, Report 2006/150, 2006, http://eprint.iacr.org/.

[11] M. A. Shah, M. Baker, J. C. Mogul, and R. Swaminathan, "Auditing to Keep Online Storage Services Honest," Proc. 11th USENIX Workshop on Hot Topics in Operating Systems (HOTOS '07), pp. 1–6, 2007

[12] Dunlap, Kevin, and Rasmussen, Neil, "The Advantages of Row and Rack-Oriented Cooling Architectures for Data Centers", American Power Conversion, TCO

[13] Vouk, M.A. Cloud Computing - Issues, research and implementations, IEEE Information Technology Interfaces 30th International Conference, page(s): 31~40, June, 2008.

[14] Maithili Narasimha and Gene Tsudik. DSAC: integrity for outsourced databases with signature aggregation and chaining. Technical report, 2005

[15] C. Martel, G. Nuckolls, P. Devanbu, M. Gertz, A. Kwong, and S. Stubblebine. A general model for authenticated data structures. Technical report, 2001

Optimized DES Algorithm Using X-nor Operand Upto 4 Round on Spartan3

[1] PoojaRathore, [2] Jaikarn Singh, [3] MukeshTiwari, [4]Sanjay Rathore
[1, 2, 3, 4] Dept. of ECE, SSSIST
Sehore, (MP) – INDIA.
Department of Electronics and Communication Engineering, SSSIST, Sehore, (MP)

Abstract —
In this paper, linear cryptanalysis is a known-plaintext attack that uses a linear relation between input-bits, output-bits, and key-bits of an encryption algorithm that holds with a certain probability. If enough plaintext-ciphertext pairs are provided, this approximation can be used to assign probabilities to the possible keys and to locate the most probable one. Along with the society relies on more and more greatly to the computer, people also attach more and more importance to the security problem in the application. The cryptography is continuously safeguarding the safe effectively protective screen of system. Owing to the fact that to break the key using mathematics technology is very difficult, people put forward the side-channel attack method in recent years.

Keywords — Encryption; Key; Modalism; S-boxes.

Introduction

Data Security is an important parameter for the industries.It can be achieved by Encryption algorithms which are used to prevent unauthorized access of data. Cryptography is the science of keeping data transfer secure, so that eavesdroppers(or attackers) cannot decipher the transmitted Message. In this paper the DES algorithm is optimized upto 4 round using Xilinx software and implemented on Spartan 3 Modelsim. The paper deals with various parameters such as variable key length, key generation mechanism, etc. used in order to provide optimized results.

The DES Algorithm IllustratebyJ. Orlin Grabbe

The DES (Data Encryption Standard) algorithm is the most widely used encryption algorithm in the world. For many years, and among many people, "secret code making" and DES have been synonymous. And despite the recent coup by the Electronic Frontier Foundation in creating a $220,000 machine to crack DES-encrypted messages, DES will live on in government and banking for years to come through a life- extending version called "triple-DES." How does DES work? This article explains the various steps involved in DES-encryption, illustrating each step by means of a simple example. Since the creation of DES, many other algorithms (recipes for changing data) have emerged which are based on design principles similar to DES. Once you understand the basic transformations that take place in DES, you will find it easy to follow the steps involved in these more recent algorithms.But first a bit of history of how DES came about is appropriate, as well as a look toward the future.

The National Bureau of Standards Coaxes the Genie from the BottleOn May 15, 1973, during the reign of Richard Nixon, the National Bureau of Standards (NBS) published a notice in the Federal Register soliciting proposals for cryptographic algorithms to protect data during transmission and storage. The notice explained why encryption was an important issue.Over the last decade, there has been an accelerating increase in the accumulations and communication of digital data by government, industry and by other organizations in the private sector. The contents of these communicated and stored data often have very significant value and/or sensitivity. It is now common to find data transmissions which constitute funds transfers of several million dollars, purchase or sale of securities, warrants for arrests or arrest and conviction recordsbeing communicated between law enforcement agencies, airline reservations and ticketing representing investment and value both to the airline and passengers, and health and patient care records transmitted among physicians and treatment centers.

The increasing volume, value and confidentiality of these records regularly transmitted and stored by commercial and government agencies has led to heightened recognition and concern over their exposures to unauthorized access and use. This misuse can be in the form of theft or defalcations of data records representing money, malicious modification of business inventories or the interception and misuse of confidential information about people. The need for protection is then apparent and urgent.

It is recognized that encryption (otherwise known as scrambling, enciphering or privacy transformation) represents the only means of protecting such data during transmission and a useful means of protecting the content of data stored on various media, providing encryption of adequate strength can be devised and validated and is inherently integrable into system architecture. The National Bureau of Standards solicits proposed techniques and algorithms for computer data encryption. The Bureau also solicits recommended techniques for implementing the cryptographic function: for generating, evaluating, and protecting cryptographic keys; for maintaining files encoded under expiring keys; for making partial updates to encrypted files; and mixed clear and encrypted data to permit labelling, polling, routing, etc. The Bureau in its role for establishing standards and aiding government and industry in assessing technology, will arrange for the evaluation of protection methods in order to prepare guidelines.

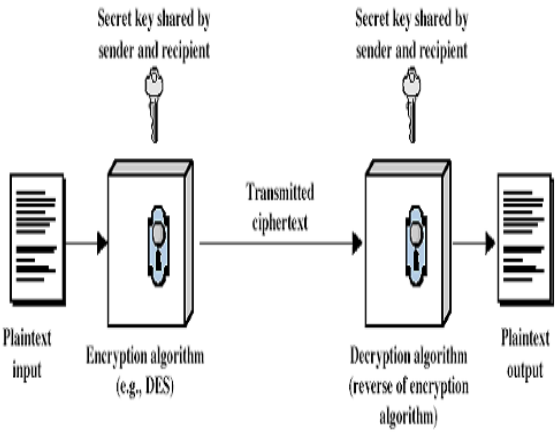
NBS waited for the responses to come in. It received none until August 6, 1974, three days before Nixon's resignation, when IBM submitted a candidate that it had developed internally under the name LUCIFER. After evaluating the algorithm with the help of the National Security Agency (NSA), the NBS adopted a modification of the LUCIFER algorithm as the new Data Encryption Standard (DES) on July 15, 1977.

DES was quickly adopted for non-digital media, such as voice-grade public telephone lines. Within a couple of years, for example, International Flavors and Fragrances was using DES to protect its valuable formulas transmitted over the phone ("With Data Encryption, Scents Are Safe at IFF," Computerworld 14, No. 21, 95 (1980).)

Meanwhile, the banking industry, which is the largest user of encryption outside government, adopted DES as a wholesale banking standard. Standards for the wholesale banking industry are set by the American National Standards Institute (ANSI). ANSI X3.92, adopted in 1980, specified the use of the DES algorithm.

Cryptography: Overview

An overview of the main goals behind using cryptography will be discussed in this section along with the common term used in this field.



Encrption / Decryption

Cryptography is usually referred to as "the study of secret", while nowadays is most attached to the definition of encryption. Encryption is the process of converting plain text "unhidded" to a cryptic text "hidded" to secure it against data thieves.This process has another part where cryptic text needs to be decrypted on the other end Some Preliminary Examples of DES works on bits, or binary numbers--the 0s and 1s common to digital computers. Each group of four bits makes up a hexadecimal, or base 16, number. Binary "0001" is equal to the hexadecimal number "1", binary "1000" is equal to the hexadecimal number "8", "1001" is equal to the hexadecimal number "9", "1010" is equal to the hexadecimal number "A", and "1111" is equal to the hexadecimal number "F".

DES works by encrypting groups of 64 message bits, which is the same as 16 hexadecimal numbers. To do the encryption, DES uses "keys" where are also apparently 16 hexadecimal numbers long, or apparently 64 bits long. However, every 8th key bit is ignored in the DES algorithm, so that the effective key size is 56 bits. But, in any case, 64 bits (16 hexadecimal digits) is the round number upon which DES is organized.

For example, if we take the plaintext message "8787878787878787", and encrypt it with the DES key "0E329232EA6D0D73", we end up with the ciphertext "0000000000000000". If the ciphertext is decrypted with the same secret DES key "0E329232EA6D0D73", the result is the original plaintext "8787878787878787".

This example is neat and orderly because our plaintext was exactly 64 bits long. The same would be true if the plaintext happened to be a multiple of 64 bits. But most messages will not fall into this category. They will not be an exact multiple of 64 bits (that is, an exact multiple of 16 hexadecimal numbers).

For example, take the message "Your lips are smoother than vaseline". This plaintext message is 38 bytes (76 hexadecimal digits) long. So this message must be padded with some extra bytes at the tail end for the encryption. Once the encrypted message has been decrypted, these extra bytes are thrown away. There are, of course, different padding schemes -- different ways to add extra bytes. Here we will just add 0s at the end, so that the total message is a multiple of 8 bytes (or 16 hexadecimal digits, or 64 bits).

The plaintext message "Your lips are smoother than vaseline" is, in hexadecimal, "596F7572206C6970 732061726520736D 6F6F746865722074 68616E2076617365 6C696E650D0A".
(Note here that the first 72 hexadecimal digits represent the English message, while "0D" is hexadecimal for Carriage Return, and "0A" is hexadecimal for Line Feed, showing that the message file has terminated.) We then pad this message with some 0s on the end, to get a total of 80 hexadecimal digits:
"596F7572206C6970 732061726520736D 6F6F746865722074 68616E2076617365 6C696E650D0A0000".

If we then encrypt this plaintext message 64 bits (16 hexadecimal digits) at a time, using the same DES key "0E329232EA6D0D73" as before, we get the ciphertext:
"C0999FDDE378D7ED 727DA00BCA5A84EE 47F269A4D6438190 9DD52F78F5358499 828AC9B453E0E653".

This is the secret code that can be transmitted or stored. Decrypting the ciphertext restores the original message "Your lips are smoother than vaseline". (Think how much better off Bill Clinton would be today, if Monica Lewinsky had used encryption on her Pentagon computer!)
DES is a block cipher--meaning it operates on plaintext blocks of a given size (64-bits) and returns ciphertext blocks of the same size. Thus DES results in a permutationamong the 2^64 (read this as: "2 to the 64th power") possible arrangements of 64 bits, each of which may be either 0 or 1. Each block of 64 bits is divided into two blocks of 32 bits each, a left half block L and a right half R. (This division is only used in certain operations.)

Example: Let M be the plain text message M = 0123456789ABCDEF, where M is in hexadecimal (base 16) format. Rewriting M in binary format, we get the 64-bit block of text:

M =	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
L =	0000		0001		0010		0011		0100		0101		0110		0111	

R = 1000 1001 1010 1011 1100 1101 1110 1111

The first bit of M is "0". The last bit is "1". We read from left to right.
DES operates on the 64-bit blocks using key sizes of 56- bits. The keys are actually stored as being 64 bits long, but every 8th bit in the key is not used (i.e. bits numbered 8, 16, 24, 32, 40, 48, 56, and 64). However, we will nevertheless number the bits from 1 to 64, going left to right, in the following calculations. But, as you will see, the eight bits just mentioned get eliminated when we create subkeys.

Example: Let K be the hexadecimal key K = 133457799BBCDFF1. This gives us as the binary key (setting 1 = 0001, 3 = 0011, etc., and grouping together every eight bits, of which the last one in each group will be unused):
K = 00010011 00110100 01010111 01111001 10011011 10111100 11011111 11110001
The DES algorithm uses the following steps:
Step 1: Create 4 sub-keys, each of which is 48- bits long. The 64-bit key is permuted according to the following table, PC-1. Since the first entry in the table is "57", this means that the 57th bit of the original key K becomes the first bit of the permuted key K+. The 49th bit of the original key becomesthesecond bit of the permuted key. The 4th bit of the original key is the last bit of the permuted key. Note only 56 bits of the original key appear in the permuted key. Example: From the original 64-bit key

K =11111111111111100000000000000001010101 01010101001010101010101 we get the 56-bit permutation
K+=00110011110000110011001111100001100111100001100110011
Next, split this key into left and right halves, C0 and D0, where each half has 28 bits. Example: From the permuted key K+, we get
C0 = 0011001111000011001100111100
D0 = 0011001111000011001100110011

Table 1: PC-1

57	49	41	33	25	17	9
1	58	50	42	34	26	18
10	2	59	51	43	35	27
19	11	3	60	52	44	36
63	55	47	39	31	23	15
7	62	54	46	38	30	22
14	6	61	53	45	37	29
21	13	5	28	20	12	4

With C0 and D0 defined, we now create sixteen blocks Cn andDn, 1<=n<=4. Each pair of blocks Cn and Dn is formed from the previous pair Cn-1 and Dn-1, respectively, for n = 1, 2, … 4, using the schedule of "left shifts" of the previous block. To do a left shift, move each bit one place to the left, except for the first bit, which is cycled to the end of the block. This means, for example, C3 and D3 are obtained from C2 and D2, respectively, by two left shifts, and C4 and D4 are obtained from C3 and D3, respectively, by one left shift. In all cases, by a single left shift is meant a rotation of the bits one place to the left, so that after one left shift the bits in the 28 positions are the bits that were previously in positions 2, 3,..., 28, 1. Example: From original pair C0 and D0 we obtain:
C0 = 0011001111000011001100111100
D0 = 0011001111000011001100110011
C1 = 1110000110011001010101011111
D1 = 0110011110000110011001100110
We now form the keys Kn, for 1<=n<=4, by applying the following permutation table to each of the concatenated pairs CnDn. Each pair has 56 bits, but PC-2 only uses 48 of these.

Table 2: PC-2

14	17	11	24	1	5
3	28	15	6	21	10
23	19	12	4	26	8
16	7	27	20	13	2
41	52	31	37	47	55
30	40	51	45	33	48
44	49	39	56	34	53
46	42	50	36	29	32

Therefore, the first bit of Kn is the 14th bit of CnDn, the second bit the 17th, and so on, ending with the 48th bit ofKn being the 32th bit of CnDn
Step 2: Encode each 64-bit block of data

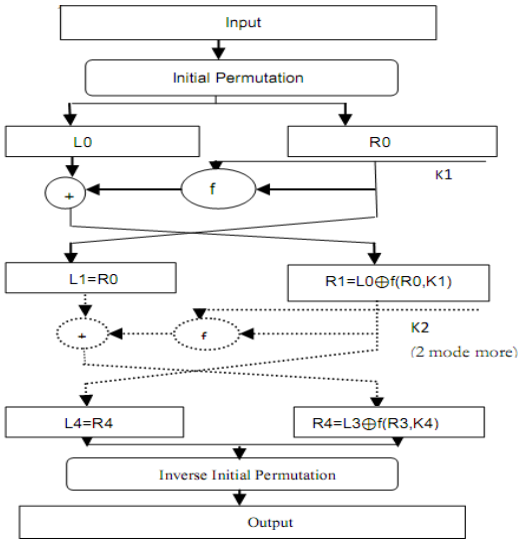
There is an initial permutation IP of the 64 bits of the message data M. This rearranges the bits according to the following table, where the entries in the table show the new arrangement of the bits from their initial order.

Table 3: IP

58	50	42	34	26	18	10	2
60	52	44	36	28	20	12	4
62	54	46	38	30	22	14	6
64	56	48	40	32	24	16	8
57	49	41	33	25	17	9	1
59	51	43	35	27	19	11	3
61	53	45	37	29	21	13	5
63	55	47	39	31	23	15	7

The 58th bit of M becomes the first bit of IP. The 50th bit of M becomes the second bit of IP. The 7th bit of M is the last bit of IP. Example: Applying the initial permutation to the block of text M, given previously, we get
M = 0000 00010010 00110100 01010110 01111000 10011010101111001101111011111
IP=1100110000000000110011001111111111110000 10101010111100001010101010
Here the 58th bit of M is "1", which becomes the first bit of IP. The 50th bit of M is "1", which becomes the second bit of IP. The 7th bit of M is "0", which becomes the last bit of IP. Next divide the permuted block IP into a left half L0 of 32 bits, and a right half R0 of 32 bits.

Example: From IP, we get L0 and R0
L0 = 11001100000000001100110011111111
R0 = 11110000101010101111000010101010
We now proceed through 4 iterations, for 1<=n<=4, using a function f which operates on two blocks--a data block of 32 bits and a key Kn of 48 bits--to produce a block of 32 bits. Let + denote XOR addition, (bit-by-bit addition modulo 2). Then for n going from 1 to 4 we calculate Ln = Rn-1 Rn = Ln-1 + f(Rn-1,Kn)n This results in a final block, for n = 4, of L4R4. That is, in each iteration, we take the right 32 bits of the previous result and make them the left 32 bits of the current step. For the right 32 bits in the current step, we XOR the left 32 bits of the previous step with the calculation f . Example: For n = 1, we have
K1=000110110000001011101111111111000111000001110010
L1 = R0 = 1111 0000 1010 1010 1111 0000 1010 1010
R1 = L0 + f(R0,K1) It remains to explain how the function f works. To calculate f, we first expand each block Rn-1 from 32 bits to 48 bits. This is done by using a selection table that repeats some of the bits in Rn-1 We'll call the use of this selection table the function E. Thus E(Rn-1) has a 32 bit input block, and a 48 bit output block. Thus the first three bits of E(Rn-1) are the bits in positions 32, 1 and 2 of Rn-1 while the last 2 bits of E(Rn-1) are the bits in positions 32 and 1.Example: We calculate
E(R0) from R0 as follows:
R0 = 1111 0000101010101111000010101010
E(R0)=011110100001010101010101011110100001
010101010101
(Note that each block of 4 original bits has been expanded to a block of 6 output bits.) Next in the f calculation, we XOR the output E(Rn-1) with the key Kn: Kn + E(Rn-1). Example: For K1 , E(R0), we have
K1=000110110000001011101111111111100011 1000001110010
(R0)=011110100001010101010101011110100001 010101 010101
K1+E(R0)=100101010001100001010101011101101000010111000111
To this point we have expanded Rn-1 from 32 bits to 48 bits, using the selection table, and XORed the result withn the key Kn . We now have 48 bits, or eight groups of six bits. We now do something strange with each group of six bits: we use them as addresses in tables called "S boxes". Each group of six bits will give us an address in a different S box. Located at that address will be a 4 bit number. This 4 bit number will replace the original 6 bits. The net result is that the eight groups of 6 bits are transformed into eight groups of 4 bits (the 4-bit outputs from the S boxes) for 32 bits total. Write the previous result, which is 48 bits, in the form:
Kn + E(Rn-1) =B1B2B3B4B5B6B7B8, where each Bi is a group of six bits. We now calculateS1(B1)S2(B2)S3(B3)S4(B4)S5(B5)S6 (B6)S7(B7)S8(B8) where Si(Bi) refers to the output of the i-th S box. To repeat, each of the functions S1, S2,..., S8, takes a 6-bit block as input and yields a 4-bit block as output. The table to determine S1 is shown and explained below: If S1 is the function defined in this table and B is a block of 6 bits, then S1(B) is determined as follows: The first and last bits of B represent in base 2 a number in the decimal range 0 to 3 (or binary 00 to 11). Let that number be i. The middle 4 bits of B represent in base 2 a number in the decimal range 0 to 15 (binary 0000 to 1111). Let that number be j. Look up in the table the number in the i-th row and j-th column. It is a number in the range 0 to 15 and is uniquely represented by a 4 bit block. That block is the output S1(B) of S1 for the input B. For example, for input block B = 011101 the first bit is "0" and the last bit "1" giving 01 as the row. This is row 1. The middle four bits are "1110". This is the binary equivalent of decimal 13, so the column is column number 13. In row 1, column 13 appears 5. This determines the output; 5 is binary 0011, so that the output is 0101. Hence S1(011101) = 0011.



DES Fiestel Diagram

Example: For the first round, we obtain as the output of the eight S boxes:
K1+E(R0)=100101010001100001010101011101101000010111000111
S1(B1)S2(B2)S3(B3)S4(B4)S5(B5)S6(B6)S7(B7)S8(B8)=010111001000 00101011 010110010111
The final stage in the calculation of f is to do a permutation P of the S-box output to obtain the final value of f: f = P(S1(B1)S2(B2)...S8(B8)) P yields a 32-bit output from a 32-bit input by permuting the bits of the input block. Example: From the output of the eight Sboxes:
S1(B1)S2(B2)S3(B3)S4(B4)S5(B5)S6(B6)S7(B7)
S8(B8) = 010111001000 00101011 01011001 0111
we get f = 0010 00110100101010100110111011
R1=L0+f(R0,K1)= 11001100000000000110011001
111 1111+ 00100011010010101010100110111011
= 11101111010010100110010101000100
In the next round, we will have L2 = R1, which is the block we just calculated, and then we must calculate R2 =L1 + f(R1, K2), and so on for 4 rounds. At the end of the sixteenth round we have the blocks L4 and R4. We then reverse the order of the two blocks into the 64-bit block R16L16 and apply a final permutation IP-1 as defined by the following table:

Table 4: IP^ (-1)

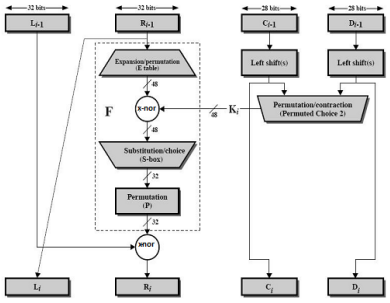
40	8	48	16	56	24	64	32
39	7	47	15	55	23	63	31
38	6	46	14	54	22	62	30
37	5	45	13	53	21	61	29
36	4	44	12	52	20	60	28
35	3	43	11	51	19	59	27
34	2	42	10	50	18	58	26
33	1	41	9	49	17	57	25

That is, the output of the algorithm has bit 40 of the pre output block as

How "FOUR-ROUND DES" Works

This section briefly gives an overview of Four-Rounded DES Algorithm. Four-Rounded DES is composed of substitutions and permutations which take place in four rounds. It is a symmetric cryptosystem which means that both the parties use the same key. Hence, the key must be kept secret.

In DES, 64-bit data is divided into left and right halves. In each round, a main function F is applied on right half of the data and a sub-key (Ki) of 48 bits. During this process eight S-boxes are used which convert each 6-bit block into 4-bit block generating 32-bit data. Finally, the left half of data is X-NORed with 32-bit output of the main function.



Single Round of DES Algorithm

DES Modes of Operation

The DES algorithm turns a 64-bit message block M into a 64-bit cipher block C. If each 64-bit block is encrypted individually, then the mode of encryption is called Electronic Code Book (ECB) mode. There are two other modes of DES encryption, namely Chain Block Coding (CBC) and Cipher Feedback (CFB), which make each cipher block dependent on all the previous messages blocks through an initial X-NOR operation. Single Round of DES Algorithm is as:

Cracking DES

Before DES was adopted as a national standard, during the period NBS was soliciting comments on the proposed algorithm, the creators of public key cryptography, Martin Hellman and Whitfield Diffie, registered some objections to the use of DES as an encryption algorithm. Hellman wrote: "Whit Diffie and I have become concerned that the proposed data encryption standard, while probably secure against commercial assault, may be extremely vulnerable to attack by an intelligence organization" (letter to NBS, October 22, 1975).

Diffie and Hellman then outlined a "brute force" attack on DES. (By "brute force" is meant that you try as many of the 2^56 possible keys as you have to before decrypting the ciphertext into a sensible plaintext message.) They proposed a special purpose "parallel computer using one million chips to try one million keys each" per second, and estimated the cost of such a machine at $20 million.

Fast forward to 1998. Under the direction of John Gilmore of the EFF, a team spent $220,000 and built a machine that can go through the entire 56-bit DES key space in an average of 4.5 days. On July 17, 1998, they announced they had cracked a 56-bit key in 56 hours. The computer, called Deep Crack, uses 27 boards each containing 64 chips, and is capable of testing 90 billion keys a second.

Despite this, as recently as June 8, 1998, Robert Litt, principal associate deputy attorney general at the Department of Justice, denied it was possible for the FBI to crack DES: "Let me put the technical problem in context: It took 14,000 Pentium computers working for four months to decrypt a single message . . . . We are not just talking FBI and NSA [needing massive computing power], we are talking about every police department."

Responded cryptograpy expert Bruce Schneier: " . . . the FBI is either incompetent or lying, or both." Schneier went on to say: "The only solution here is to pick an algorithm with a longer key; there isn't enough silicon in the galaxy or enough time before the sun burns out to brute- force triple-DES" (Crypto-Gram, Counterpane Systems, August 15, 1998).

Conclusion and Future Works

In this paper, for the cryptanalysis of Data Encryption Standard is presented. It shows that it is an effective approach for cryptanalysis of four-rounded DES using X-Nor. The cost function used in this paper is generic and can be used for the cryptanalysis of other block ciphers In the future, it also performed under different operands and even by altering them.

References

[1]. "Cryptographic Algorithms for Protection of Computer Data During Transmission and Dormant Storage," Federal Register 38, No. 93 (May 15, 1973).

[2]. Data Encryption Standard, Federal Information Processing Standard (FIPS) Publication 46, National Bureau of Standards, U.S. Department of Commerce, Washington D.C. (January 1977).

[3]. Carl H. Meyer and Stephen M. Matyas, Cryptography: A New Dimension in Computer Data Security, John Wiley & Sons, New York, 1982.

[4]. Dorthy Elizabeth Robling Denning, Cryptography and Data Security, Addison-Wesley Publishing Company, Reading, Massachusetts, 1982.

[5]. D.W. Davies and W.L. Price, Security for Computer Networks: An Introduction to Data Security in Teleprocessing and Electronics Funds Transfer, Second Edition, John Wiley & Sons, New York, 1984, 1989.

[6]. Miles E. Smid and Dennis K. Branstad, "The Data Encryption Standard: Past and Future," in Gustavus J. Simmons, ed., Contemporary Cryptography: The Science of Information Integrity, IEEE Press, 1992.

[7]. Douglas R. Stinson, Cryptography: Theory and Practice, CRC Press, Boca Raton, 1995.

[8]. Bruce Schneier, Applied Cryptography, Second Edition, John Wiley & Sons, New York, 1996.

[9]. Alfred J. Menezes, Paul C. van Oorschot, and Scott A. Vanstone, Handbook of Applied Cryptography, CRC Press, Boca Raton, 1997.

[10]. Ruth M. Davis, "The Data Encryption Standard" Proceedings of Conference on Computer Security and the Data Encryption Standard, National Bureau of Standards, Gaithersburg, MD, Feb. 15, 1977, NBS Special Publication 500-27, pp. 5-9.

[11]. WhitfieldDiffie, "Cryptographic Technology: Fifteen Year Forecast" Reprinted by permission AAAS, 1982 from Secure Communications and Asymmetric Crypto Systems. AAAS Selecte8 Symposia. Editor: C.J. Simmons. Vol. 69, Westview Press, Boulder, Colorado, pp. 38-57.

[12]. C. Boyd. "Modern Data Encryption," Electronics & Communication Engineering Journal, October 1993, pp. 271-278.

[13]. Seung-Jo Han, "The Improved Data Encryption Standard (DES) Algorithm" 1996, pp. 1310-1314.

[14]. A.Kh. AI Jabri, "Secure progressive transmission of compressed images" IEEE Transactions on Consumer Electronics, Vol. 42, No. 3, AUGUST 1996, pp. 504-512 .

[15]. K. Wong, "A single-chip FPGA implementation of the data encryption standard (des) algorithm" IEEE 1998 pp. 827-832 .

[16]. Subbarao V. Wunnava, "Data Encryption Performance and Evaluation Schemes" Proceedings IEEE Southeastcon 2002, pp. 234-238

[17]. Xun Yi," Identity-Based Fault-Tolerant Conference Key Agreement" IEEE transactions on dependable and secure computing, vol. 1, no. 3, July-September 2004, pp. 170-178 .

[18]. M. Backes, "Relating Symbolic and Cryptographic Secrecy" IEEE transactions on dependable and secure computing, vol. 2, no. 2, April-June 2005, pp. 109-123 .

[19]. ElisaBertino, "An Efficient Time-Bound Hierarchical Key Management Scheme for Secure Broadcasting" IEEE transactions on dependable and secure computing, vol. 5, no. 2, April-June 2008, pp. 65-70.

[20]. Clark, A.,"Modern Optimization Algorithms for Cryptanalysis".Proceedings of Second IEEE Australian and New Zealand Conference on Intelligent Information Systems, pp.258-262, 1994.

[21]. Laskari, E. C., Meletiou, G. C., Stamation, Y. C., and Vrahatis, M. N.,"Evolutionary Computation based Crypt-analysis: A first study". Nonlinear Analysis, vol. 63, no.(5- 7), pp. 823-830, 2005.

[22]. R, Vimalathithan, and Valarmathi, M. L.,"Cryptanalysis of S-DES using Genetic Algorithm". International Journal of Recent Trends in Engineering, vol. 2, no. 4, pp.76-79, Nov.2009.

[23]. Shahzad, W., Siddiqui, A. B., and Khan, F. A., "Cryptanalysis of Four-Round DES using Binary Particle Swarm Optimization". Genetic and Evolutionary Computation Conference, pp. 1757-1758, July 8-12, 2009.

[24]. Song, J., Zhang, H., Meng, Q., and Wang, Z.,"Cryptanalysis of Four-Round DES Based on Genetic Algorithm". International Conference on Wireless Communications Networking and Mobile Computing, Issue 21-25, pp. 2326-2329, Sept. 2007.

[25]. Spillman, R., Janssen, M., Nelson, B., and Kepner, M.,"Use of A Genetic Algorithm in the Cryptanalysis of Simple Substitution Ciphers". Cryptologia, vol.17, no.1, pp.

Voltage Unbalance Correction in a Grid Using Inverter

K.Jayakumar[1], N.Sriharish[2], Ch.Rambabu[3]

1 M.Tech Student in Power Electronics, Dept. of EEE at Sri Vasavi Engineering College, Tadepalligudem, A.P, India
2 Assistant Professor, Dept. of EEE at Sri Vasavi Engineering College, Tadepalligudem, A.P, India
3 Professor & HOD, Dept. of EEE at Sri Vasavi Engineering College, Tadepalligudem, A.P, India

Abstract

This paper presents the control of a "voltage unbalance correction in a grid using inverter". The inverters are proposed give additional function to decrease the negative sequence voltage at the point of correction with the utility grid. By using improved multi variable filter, the grid inverter absorbs a small amount of negative sequence current from the grid, and which based up on symmetric sequence voltage decomposition, thereby helping to correct the negative sequence voltage. But the amplitude reduction by each individual inverter system is small as compared to the entire negative sequence component, and these inverter modules can achieve to collect substantial results in the grid. Finally the analyses of the scheme along with the suitable design are presented by using basic circuit diagram and proposed control has been verified by simulation results are shown.

Keywords: PWM inverter, multi variable filter, voltage unbalance, point of correction, negative sequence voltage, distributed generation, grid interfacing etc.

I. Introduction:

In practical three phase power systems, voltage unbalance problems are existing. These problems are mainly caused by single phase and non-linear loads, which are unequally distributed. Therefore these unequal voltage drops are mainly occur across transformers and line impedances. Here the negative sequence voltages are especially troublesome in practical applications. Due to this the zero sequence component are not exist in three wire systems. These voltage unbalance effect is quite serve for electrical machines, power electronic converters and its drives [1]. So to mitigate this voltage unbalance we can go to design power electronic converters for regulating the reactive power [2, 3]. But in underground cables this approach is not suitable because in underground cables the resistance of the cable dominates its inductance. To maintain a balanced voltage at the load terminals, an often used idea is to inject a series voltage [4, 5]. It is straightforward to mitigate the voltage unbalance problem with such converters, but a disadvantage is that they are unused or only lightly loaded when there are no voltage unbalance problems. For dealing with other power quality problems than voltage unbalance, so-called unified power quality conditioners (UPQC) are proposed and continuously improved. However, the UPQC has no energy storage capabilities [6], and should be extended to cope with distributed generation (DG) [7].

Facing the emerging application of distributed generation, power electronics-based grid-interfacing inverters are playing an important role interfacing DGs to the utility grid. In addition to conventional delivery of electricity, ancillary functionality for improvement of power quality problems is being introduced into grid-interfacing inverters [8, 9]. In this paper, it is proposed to integrate voltage unbalance correction into the control of grid-interfacing inverters. This does not require more hardware, since the feedback variables for this control are already available. By controlling the negative-sequence currents, which induce opposite negative-sequence voltage drops on the line impedances, the objective of eliminating negative sequence voltages at the point of connection (PoC) with the grid may be achieved. To investigate the effectiveness of the proposed function, a three-phase four-wire inverter is used to control voltage unbalance correction. The employed inverter operates normally when the utility voltages are balanced, and when unbalanced, performs compensation automatically for negative-sequence voltage, based on utility voltage unbalance factor (VUF) [1]. To this aim, the analysis of negative-sequence current control and high performance detection for symmetrical sequences are introduced in the following. Then, the inverter control scheme and reference signal generation are presented. Finally, the proposed control methods are verified by simulations.

II. Grid-interfacing inverter with integrated voltage unbalance correction:

Fig. (1) shows the structure of a three-phase four-wire grid-interfacing system being connected to the utility grid at the POC through LCL filters. It normally synchronizes with the utility grid and delivers electrical energy to the grid from the DC-bus when pre-regulated distributed sources are connected. The voltage unbalance correction function is added, which intentionally regulates negative sequence currents. Note that, in order to obtain a maximum power factor, most grid-interfacing inverters deliver only positive-sequence currents under either balanced or unbalanced conditions. Therefore, the development of this proposed controller differs from the conventional one, and its design will be presented in the next sections of this paper. In

view of unbalanced situations, a four-leg inverter topology is used as the circuit to eliminate zero-sequence currents. With the theory of symmetric decomposition for three phase systems [10], unbalanced grid voltages can be divided into three groups, namely positive, negative, and zero sequence voltages. Similarly, current quantities can also be separated. By disregarding the mutual coupling between the grid lines in Fig. (1), an equivalent circuit model for each group of sequence components can be derived [11]. The diagram for negative-sequence components is shown in Fig.(2), where the superscript "-" denotes negative sequence. Similarly, the superscript "+" denotes positive sequence. Phasors V_g^- and V_s^- are the negative-sequence voltages of the utility grid and at the PoC, respectively. Current I_s^- is the negative-sequence current equivalent line impedance is represented by *Zg*, the equivalent impedance of the utility grid when the line impedances of the three phases are assumed symmetrical.

Accordingly, a phasor diagram showing the change for negative-sequence fundamental current is drawn in Fig.(3). By changing the amplitude and phase of the negative sequence current I_s^-, the negative-sequence voltage V_s^- can be regulated through the voltage drops across the line impedance. For a given amplitude I_s^-, the voltage changes along the dashed circle and reaches a minimum value at the point *M* where $\theta-$ equals the negative of impedance angle of *Zg*'s. Similarly, zero-sequence voltages at the PoC can be compensated by regulating the zero-sequence currents within the system. This paper only concentrates on the correction of negative-sequence voltages, considering zero-sequence voltages do not exist in case of three wire systems. Of course, zero-sequence voltages can be isolated by transformers when needed. Furthermore, it is noted that measurements of zero-sequence components can be done simply by adding three phase while accurate positive- and negative-sequence components are difficult to be determined. Therefore, zero sequence voltage correction can be trivially added to the control based on the proposed control scheme for negative-sequence voltage correction and is not discussed in this paper.

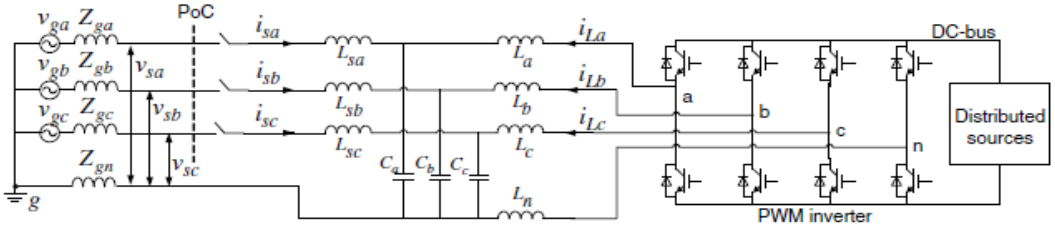


Fig.(1) Three-phase four-wire grid-interfacing four leg inverter at PoC.

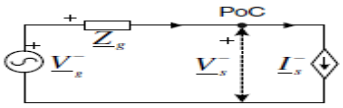


Fig.(2) Negative-sequence equivalent model

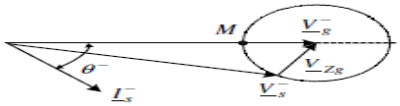


Fig.(3) Phasor diagram of the negative-sequence model.

III. Control scheme:

A. Determination of Negative-sequence Currents

Fig. (3) Illustrates the basic principle of how to correct unbalanced voltage at the PoC with sequence-current control. It is suggested to determine the negative-sequence currents based on the voltage unbalance factor. To assess unbalanced voltages at the PoC, the voltage unbalance factor, *KVUF* is defined as the ratio between the amplitude of the negative-sequence voltage V_s^- and the amplitude of the positive-sequence voltage V_s^+. The following constraint equation is proposed to calculate the desired current amplitude I_s^-:

$$\frac{I_s^-}{I_s^+} = \frac{V_s^-}{V_s^+} = K_{VUF}, \qquad (1)$$

where I_s^+ is the amplitude of the positive-sequence current. Then, the resulting I_s^- is derived based on the ratio of unbalance voltages at the PoC from (1).

However, the voltage unbalance factor at the PoC varies with the controlled negative-sequence currents, because the controller utilizes feed forward measurements of K_{VUF} and operates in a open-loop. Consequently, this strategy may cause the value of K_{VUF} in (1) to vary. To ensure a stable correction, a smooth update method for K_{VUF} is added to the control. The flow chart shown in Fig.(4) illustrates how to derive the final I_s^-. The currently measured quantity is referred to as $K_{VUF}(n)$, and the

previous one is $K_{VUF}(n-1)$. In Fig.(4), the minimum threshold (*Kmin*) of negative-sequence correction is defined according to practical demands, and a coefficient denoted by λ is introduced for smooth regulation when decreasing the output value of K_{VUF}. Note that, for system protection, the current rating of the inverters is always checked before returning I_s^-.

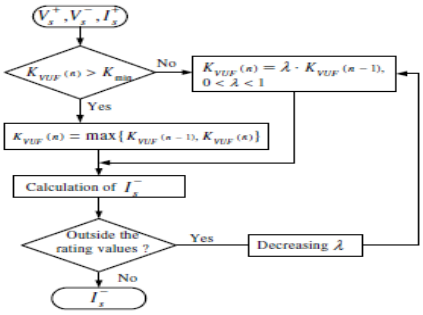


Fig.(4) Flow chart of K_{VUF} determination.

B. Positive- and Negative-Sequence Detection

The factor K_{VUF} is essential to get the amplitude of negative-sequence currents. Thus the separation of sequence voltages is central to get the value of K_{VUF}, as well as to the synchronization with the utility grid. For unbalanced or distorted grid voltages, a multi-variable filter was introduced in [12] for detecting the positive-sequence component in the stationary frame. After modification, this filter is able to directly filter out the fundamental positive and negative-sequence vectors. The following mathematically demonstrates the multi- variable filter for symmetric sequence decomposition.

For unbalanced distorted voltages, the positive- and negative-sequence components are in the $\alpha - \beta$ frame as expressed by

$$v_{\alpha\beta}(t) = v_\alpha(t) + jv_\beta(t)$$

$$= \sum_{k=1}^{\infty}(V_k^+ e^{jk\omega_1 t} + V_k^{o-} e^{-jk\omega_1 t}) \qquad (2)$$

where *k* denotes the harmonic number, $\omega 1$ denotes the fundamental radian frequency, and the superscript symbol "o" denotes conjugate.

Let us look for a filter $G^+(t)$, which can damp all harmonic components of $v_{\alpha\beta}(t)$ but the fundamental positive –sequence component in the stationary frame. That is,

$$v_{\alpha\beta}(t) * G^+(t) = v'_{\alpha\beta}(t), \qquad (3)$$

where the "*" denotes a convolution product, and

$$v'_{\alpha\beta}(t) = V_1^+ e^{jk\omega_1 t} + \sum_{k=3,5,\ldots}^{\infty} U_k^+ e^{jk\omega_1 t} + \sum_{k=1,3,\ldots}^{\infty} U^{o-}_k e^{-jk\omega_1 t} \qquad (4)$$

with $\| U_k^+ \| \ll \| V_1^+ \|$ *and* $\| U^{o-}_k \| \ll \| V_1^- \|$.

Otherwise stated $v'_{\alpha\beta}(t) \approx v^+{}_{\alpha\beta}(t) = V_1^+ e^{j\omega_1 t}$ the fundamental positive-sequence component of $v_{\alpha\beta}(t)$ as defined in (2).

By multiplying $v_{\alpha\beta}(t)$ and $v'_{\alpha\beta}(t)$ with $e^{-j\omega_1 t}$, respectively, which corresponds to a transformation to a positive synchronous rotating frame (PSRF), we obtain from (2) and (3)

$$v_{\alpha\beta}(t)\, e^{-j\omega t} = V_1^+ \sum_{k=2,4,\ldots}^{\infty} (V_k^+ e^{jk\omega_1 t} + V_k^{o-} e^{-jk\omega_1 t}),$$

$$v'_{\alpha\beta}(t)\, e^{-j\omega_1 t} = V_1^+ e^{jk\omega_1 t} + \sum_{k=2,4,\ldots}^{\infty}(U_k^+ e^{jk\omega_1 t} + U^{o-}_k e^{-jk\omega_1 t}) \qquad (5)$$

It can be seen that the fundamental positive-sequence voltage performs as a DC quantity in the PSRF. Therefore, a simple first order filter *H (t)* with

$$L[H(t)] = H(s) = \frac{\omega_b}{s+\omega_b} \tag{6}$$

where ω_b is the corner frequency, is sufficient to get $v'_{\alpha\beta}(t)e^{-j\omega_1 t}$ from $v_{\alpha\beta}(t)e^{-j\omega_1 t}$ under the conditions of (4). This can be expressed with

$$v_{\alpha\beta}(t)\, e^{-j\omega_1 t} * H(t) = v'_{\alpha\beta}(t)\, e^{-j\omega_1 t} \tag{7}$$

or, using Laplace

$$v_{\alpha\beta}(s + j\omega_1)H(s) = v'_{\alpha\beta}(s + j\omega_1) \tag{8}$$

Substituting $s \leftarrow s - j\omega_1$ into (6) and (8), it follows that

$$v_{\alpha\beta}(s).\frac{\omega_b}{s - j\omega_1 + \omega_b} = v'_{\alpha\beta}(s) \tag{9}$$

From (3), we also have

$$v_{\alpha\beta}(s)G^+(s) = v'_{\alpha\beta}(s) \tag{10}$$

Therefore, $G^+(s)$ the filter we are looking for, in the stationary frame should be equal to

$$G^+(s) = H(s - j\omega_1) = \frac{\omega_b}{s - j\omega_1 + \omega_b} \tag{11}$$

By expanding (10) to

$$v'_\alpha(s) + jv'_\beta(s) = \frac{\omega_b}{s - j\omega_1 + \omega_b}[v_\alpha(s) + jv_\beta(s)] \tag{12}$$

the following equations are derived

$$\begin{aligned} v'_\alpha(s) &= \frac{1}{s}[\omega_b\left(v_\alpha(s) - v'_\alpha(s)\right) - \omega_1 v'_\beta(s)] \\ v'_\beta(s) &= \frac{1}{s}[\omega_b\left(v_\beta(s) - v'_\beta(s)\right) - \omega_1 v'_\alpha(s)] \end{aligned} \tag{13}$$

Similarly, the fundamental negative-sequence component follows as

$$v_{\alpha\beta}(t) * G^-(t) = v''_{\alpha\beta}(t) \tag{14}$$

Where, $v''_{\alpha\beta}(t) = V_1^{o+}e^{-j\omega_1 t} + \sum_{k=1,3,\dots}^{\infty} U_k^+ e^{jk\omega_1 t} + \sum_{k=3,5,\dots}^{\infty} U^{o-}_k e^{-jk\omega_1 t}$ (15)

Or $v''_{\alpha\beta}(t) \approx v^-_{\alpha\beta 1}(t)$ similar to (8) and (10), we have

$$\begin{aligned} v_{\alpha\beta}(s - j\omega_1).H(s) &= v''_{\alpha\beta}(s - j\omega_1) \\ v_{\alpha\beta}(s).G^-(s) &= v''_{\alpha\beta}(s) \end{aligned} \tag{16}$$

Where, $$G^-(s) = \frac{\omega_b}{s + j\omega_1 + \omega_b}$$

Correspondingly, the equations below are derived:

$$v_{\alpha}''(s) = \frac{1}{s}[\omega_b\left(v_{\alpha}(s) - v_{\alpha}''(s)\right) - \omega_1 v_{\beta}''(s)]$$

$$v_{\beta}''(s) = \frac{1}{s}[\omega_b\left(v_{\beta}(s) - v_{\beta}''(s)\right) - \omega_1 v_{\alpha}''(s)] \qquad (17)$$

Therefore, the detection for $v_{\alpha 1}^{+}(t) + jv_{\beta 1}^{+}(t)$ and $v_{\alpha 1}^{-}(t) - jv_{\beta 1}^{-}(t)$ are approximately achieved from (13) and (17). These equations can be easily implemented in the $\alpha - \beta$ frame by digital control, without complicated transformation to the SRF and the inverse transformation.

In practical applications, the negative-sequence component is too small to be detected accurately. This is because the input signals involve a large proportion of positive sequence components which are difficult to damp totally. Alternative signals$\hat{v}_{\alpha\beta}(t)$, with

$$\hat{v}_{\alpha\beta}(t) = \hat{v}_{\alpha}(t) + j\hat{v}_{\beta}(t) = v_{\alpha\beta}(t) - v^{+}{}_{\alpha\beta 1}(t) \qquad (18)$$

where the dominant positive-sequence component $v^{+}{}_{\alpha\beta 1}(t) = v_{\alpha 1}^{+}(t) + jv_{\beta 1}^{+}(t)$ is abstracted, and can be used as input signals. This will improve the filtering effect for negative-sequence quantities.

Fig. (5) illustrates the implementation diagram of the multiple-variable filter, where the bandwidth ω_b for the positive- and negative-sequence filter is denoted by ω_{b1} and ω_{b2} respectively (the values can be different and adapted to practical situations). The central frequency ω_1 is set at the fundamental frequency of the grid voltage. In case of grid frequency variations the bandwidth can be increased slightly, or ω_1 can be adaptively updated with the measured fundamental frequency.

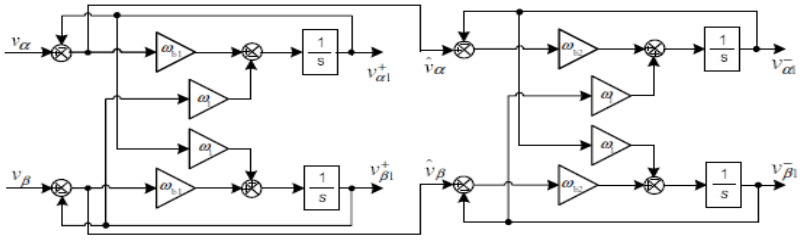


Fig. (5) Implementation diagram of the multi-variable filter

A frequency domain multi-variable filter plot is drawn in Fig. 6, based on (10) and the second equation in (16). Due to unity gain and zero phase-shift of the positive - sequence filter at the central frequency (50Hz), $\hat{v}_{\alpha}(t) + j\hat{v}_{\beta}(t)$ can be directly derived, see Fig. (5).

C. Reference signals generation

Fig. (6) shows the block diagram of the inverter's current reference generator. It consists of the detection of symmetric sequence voltages with a multi-variable filter, the VUF calculation, average power regulation and the signal synthesis. The first two processes have been detailed in the previous two subsections. By utilizing the fundamental positive - and negative-sequence components filtered out by the filter, we can obtain

$$V_{mag}^{+} = \sqrt{v_{\alpha 1}^{+2} + v_{\beta 1}^{+2}}$$

$$V_{mag}^{-} = \sqrt{v_{\alpha 1}^{-2} + v_{\beta 1}^{-2}} \qquad (19)$$

where V_{mag}^{+} and V_{mag}^{-} denote the magnitude of fundamental positive- and negative-sequence voltage, respectively.

Consequently, two groups of per-unit signals can be derived with divisions, that is as shown in Fig. (6). According to the principle described in section II, negative-sequence currents are designed to keep a phase-shift $\theta-$ with the negative-sequence voltage. This phase-shift equals the negative line impedance angle for the maximum correction effect. Its mathematical derivation is

$$i_{s\alpha}^{-} * + j i_{s\beta}^{-} * = (i_{\alpha}^{-} * + j i_{\beta}^{-} *) e^{j\theta^{-}} \qquad (20)$$

The positive-sequence current references are either inphase or in anti-phase with the positive-sequence component of the grid voltage, depending on the desired direction for energy delivery. In this paper, the gain K_{dir} is set -1 in order to deliver energy to the utility grid. In the average power control loop of Fig.(6), the power reference P^* is given, which can be determined according to the application, such as the active power generated by upstream DG or the power demanded by the downstream utility grid. In order to eliminate the effects of double fundamental frequency ripple on the measured average power, the parameters should have a small proportional gain and a big integration time constant. In this work, the gain is chosen as 0.04 and the time constant is 0.02s. The output of the PI controller is used to regulate the amplitudes of the desired currents with the coefficient K_c.

All together, it follows that the current references $i_{s\alpha}^{*}$ and $i_{s\beta}^{*}$ are derived in the stationary frame. This is beneficial for the controller design, since the controller presented in the next section is also designed in the stationary frame. The mathematical manipulations to optimally implement the above digital process are not the subject of this paper, and will be discussed elsewhere.

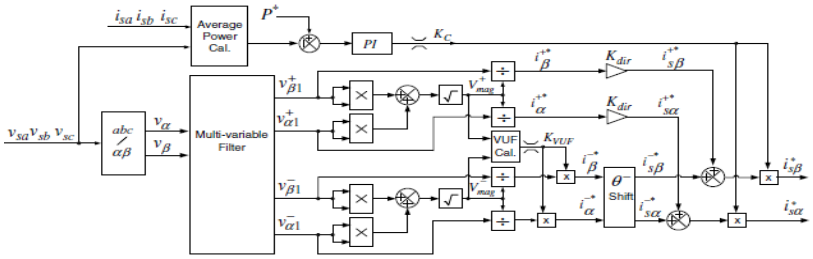


Fig. (6) Current reference generation for the inverter control.

D. Controller for Current Regulation

Fig. (7) shows the controller structure of the grid interfacing inverter. It is constructed by a double-loop current controller, which is an outer control loop with proportional-resonant (PR) controllers for eliminating the zero steady-state error of the delivered currents, and an inner capacitor current control loop with simple proportional controllers to improve stability. Instead of direct sampling, capacitor currents are calculated from the output currents and the inner filter inductor currents. These currents are measured anyway for over-current protection. To eliminate the zero-sequence currents in unbalanced situations, the current reference $i_{s\gamma}^{*}$ should be zero. The control for both positive- and negative-sequence components would be much too complicated and computation-time consuming when conventional PI control with coordinate transformation were used. Therefore, it is preferred to choose a PR controller in the stationary frame. A quasi-proportional-resonant controller with high gain at the fundamental frequency is used,

$$G_i(s) = K_p + \frac{2K_r\omega_{br}s}{s^2 + 2\omega_{br}s + \omega_1^2} \qquad (21)$$

Where K_p are the proportional gain, K_r the resonant gain, and ω_{br} the equivalent bandwidth of the resonant controller. A detailed design for the PR controller has been presented in [13], it is not duplicated here. Through optimizing, the parameters used in the simulation are K_p =0.5, K_r =50, and ω_{br} =20.

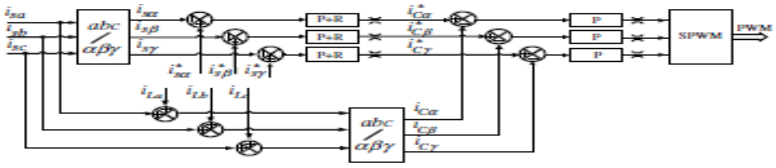
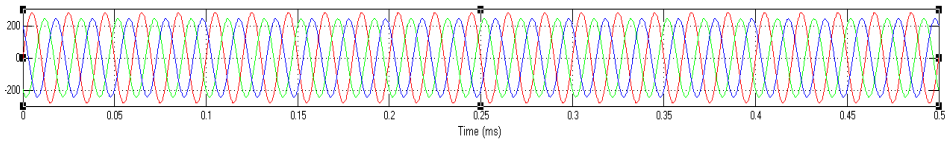


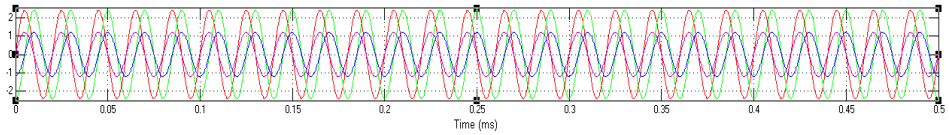
Fig. (7). Structure of the controller for current regulation

VI. Simulation Results:

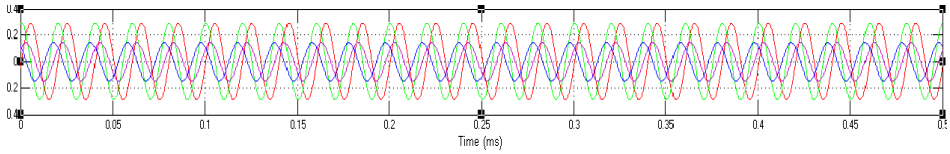
Simulation results from mat lab / simulink are provided to enable the verification of the reference signals generation. The system parameters are shown in the below table. In order to easily observe the effects of negative-sequence correction, we intentionally blown-up the values of the line inductances to the same order as the filter inductors. Therefore, the inductors L_{sa}, L_{sb}, L_{sc} are combined with the line impedances, reducing the LCL structure to an LC one. According to the values of the line impedances in below, we obtain that θ = 45 degrees. For a straightforward test of the effectiveness caused by the negative-sequence voltage correction, only fundamental positive- and negative-sequence components are considered in the grid voltages as given. It should be pointed out that the afore-mentioned control scheme and the multi-variable filter can also be implemented for distorted grid voltages.



(a) unbalanced grid voltages in a-b-c frame



(b) per-unit positive sequence currents $i_{s\alpha}^{+*}$ and $i_{s\beta}^{+*}$ in-phase with the positive-sequence voltage



(c) negative-sequence current $i_{s\alpha}^{-*}$ and $i_{s\beta}^{-*}$ lags the negative-sequence voltage by 45 *degrees* in the $\alpha - \beta$ frame.

Fig. (8). Simulation results of the reference currents generation.

To verify the proposed control method with its integrated correction function, the controller is designed on a Mat lab Simulink. Due to the long computation time of the controller, a sampling frequency of 8 kHz is used. The switching frequency is twice the sampling frequency.

From the above fig.(8) shows the simulation results of reference current generation. And this simulation wave form shows the (a). Unbalanced grid voltages in a-b-c frame, (b).per-unit positive sequence currents $i_{s\alpha}^{+*}$ and $i_{s\beta}^{+*}$ in-phase with the positive-sequence voltage, (c).negative-sequence current $i_{s\alpha}^{-*}$ and $i_{s\beta}^{-*}$ lags the negative-sequence voltage by 45 *degrees* in $\alpha - \beta$ frame.

Fig. (9). show the simulation waveforms of the Grid-interfacing inverter with integrated negative-sequence voltage correction. The plots are the unbalanced grid voltages, the controlled line currents, and the voltages at the PoC, respectively. Using unbalanced grid voltages, the inverter delivers mainly positive-sequence currents to the utility grid and absorbs 10% of the negative-sequence currents. The effectiveness of the multi-variable filter in detecting positive- and negative-sequence

components from unbalanced voltages is shown in Fig.(10). For observing, these simulation waveforms of The RMS value and phase-shift of the positive- and negative-sequence voltages show almost the same results as the calculation results from $a-b-c$ quantities to $\alpha-\beta$ quantities. To observe the negative-sequence voltage correction, the results are illustrated in $\alpha-\beta$ frame by decomposing voltages from the $a-b-c$ frame. As seen in Fig.(11), the amplitude of the negative-sequence voltage at the PoC is reduced, although the decrease is limited to around 10%. Again note that the line impedance parameters have been blown up. In a utility grid, for instance $200\mu H$ line impedance is more realistic, and then the decrease would be around 1% for the same conditions. However, based on multiple modules, the effect of the negative-sequence voltage correction will be more pronounced.

Qualitatively, we can assume that the regulated negative sequence currents by the modules in Fig.(12) can be lumped into a single module, and therefore should behave identical to the single inverter and its results are shown in Fig.(11). And fig.(13) shows the corrected voltage at the PoC when we simply provide more negative-sequence current. It can be seen that the three-phase voltages tend to be balanced. This generally indicates the effectiveness of distributed voltage unbalance correction. However, it must be noted that the proposed method is only an alternative. It is preferable in an unbalanced situation with small voltage deviation, while the conventional methods are suitable for serious situations with large voltage unbalance.

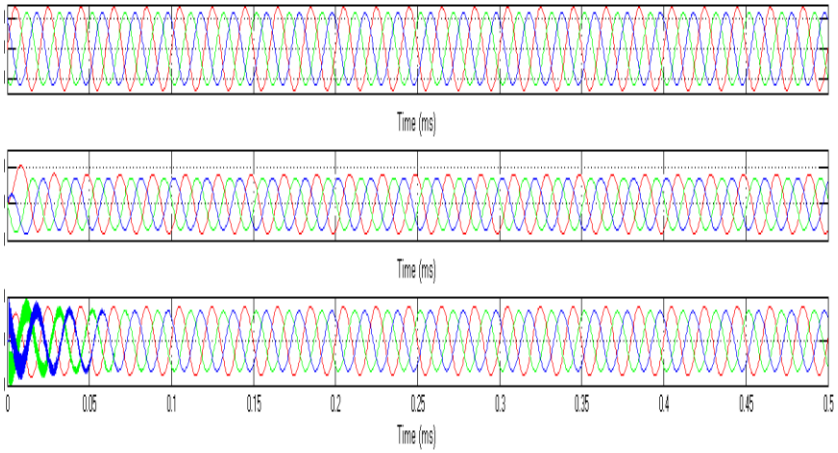


Fig.(9). Simulation results of the grid-interfacing inverter with integrated voltage unbalance correction

(a) Unbalanced grid voltages, (b) Currents delivered by the inverter, (c) Voltages at the PoC.

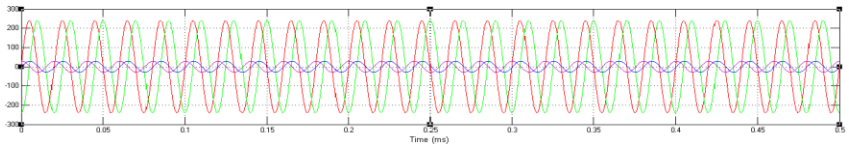


Fig. (10). Simulation waveforms of positive- and negative-sequence voltage detection, where the filtered out fundamental symmetric sequence voltages are derived in $\alpha-\beta$ frame.

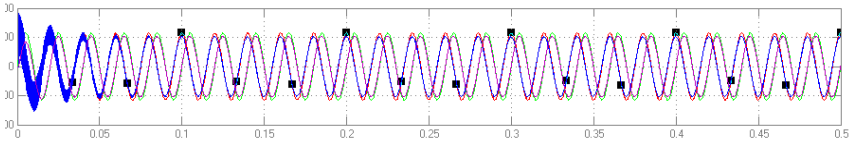


Fig. (11). Simulation results of the negative-sequence voltage correction. The α, β components of the negative-sequence voltage of the PoC $v_{s\alpha}^{-}$, $v_{s\beta}^{-}$ shows a 10% amplitude reduction compared with the negative sequence voltage of the grid $v_{g\alpha}^{-}$, $v_{g\beta}^{-}$.

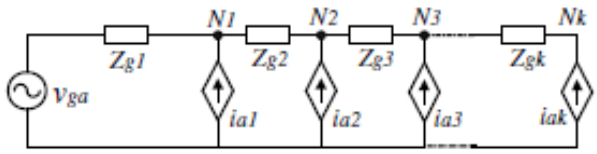


Fig.(12). Per-phase equivalent circuit with multiple modules.

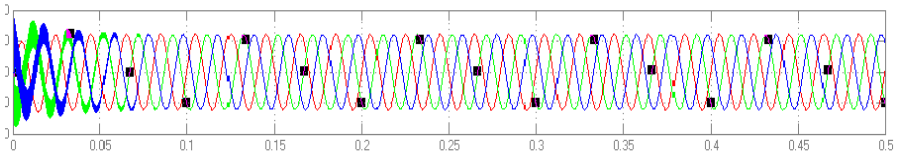


Fig.(13). Simulation waveforms of the negative-sequence voltage correction. The resulting corrected voltages tend to be balanced.

V. System Parameters:

DESCRIPTION	SYMBOL	VALUE
Grid voltage	V_{ga}	198V∟0°
	V_{gb}	171.71V∟125.21°
	V_{gc}	171.71V∟125.21°
Line impedance	Z_{ga},Z_{gb},Z_{gc}	2mH, 0.628Ω
Neutral impedance	Z_{gn}	100uH, 0.03Ω
Filter inductor	L_a,L_b,L_c	2mH, 0.03Ω
	L_n	0.67mH, 0.03Ω
Filter capacitor	C_a,C_b,C_c	5μF
DC-bus	V_{DC}	700V
Switching frequency	f_{sw}	16kHz

Table. (1).

VI. Conclusion:

In this paper, the detailed control of grid-interfacing inverters supporting negative-sequence voltage correction has been presented from basic principle. Based on the voltage unbalance factor and the system's capacity, the inverter absorbs a small amount of negative-sequence current from the grid, thereby correcting the negative sequence voltage. It has been shown that a grid-interfacing inverter, in addition to its normal operation, can help to decrease the negative-sequence voltage at the PoC. By using many of these modules, a substantial improvement is possible. Furthermore, the improved multi-variable filter can filter out positive- and negative-sequence components accurately in case of unbalanced/distorted situations in the stationary frame. The functionality and control scheme are verified by simulation results are shown.

VII. References:

[1] Annette von Jouanne and Basudeb (Ben) Banerjee, "Assessment of voltage unbalance," *IEEE Trans. Power Del.*, vol. 16, no. 4, pp.782-790, Oct. 2001.

[2] Hideaki Fujita, and H. Akagi, "Voltage-regulation performance of a shunt active filter intended for installation on a power distribution system," *IEEE Trans. Power Electron.*, vol 22, no. 3, pp. 1046-1053, May 2007.

[3] Kuang Li, Jinjun Liu, and Zhaoan Wang, and Biao Wei, "Strategies and operating point optimization of STATCOM control for voltage unbalance mitigation in three-phase three-wire systems," *IEEE Trans. Power Del.*, vol. 22, no. 1, pp. 413-422, Jan. 2007.

[4] Kalyan K. Sen, "SSSC-static synchronous series compensator theory, modeling, and application," *IEEE Trans. Power Del.*, vol. 13,

[5] Vijay B. Bhavaraju and Prasad N. Enjeti, "An active line conditioner to balance voltages in a three-phase system," *IEEE Trans. Ind. Applicat.*, vol. 32, no. 2, pp. 287-292, Mar./Apr. 1996

[6] Hideaki Fujita, H. Akagi, "The unified power quality conditioner: the integration of series- and shunt-active filters," *IEEE Trans. Power Electron.*, vol.13, no. 2, pp. 315-322, Mar. 1998.

[7] Dusan Graovac, V. A. Katic, and A. Rufer, " Power quality problems compensation with universal power quality conditioning system," *IEEE Trans. Power Del.* , vol. 22, no. 2, pp. 968-976, Apr. 2007.

[8] Koen J. P. Macken, Koen Vanthournout, Jeroen Van den Keybus, Geert Deconinck, and Ronnie J. M. Belmans, "Distributed Control of Renewable Generation Units With Integrated Active Filter," *IEEE Trans. Power Electron.*, vol. 19, no. 5, pp. 1353-1360, Sep. 2004.

[9] G. Jos, B.-T. Ooi, D. McGillis, F. D. Galiana and R. Marceau, "The potential of distributed generation to provide ancillary services," in *Proc. IEEE Power Eng. Soc. Summer Meeting*, Seattle, WA, July, 2000.

[10] P. M. Andersson, *Analysis of faulted power systems*, New York: IEEE Press, 1995.

[11] Fei Wang, Jorge L. Duate, Marcel A. M. Hendrix, "Weighting function integrated in grid-interfacing converters for unbalanced voltage correction," in *Proc. International Conf. on Renewable Energy and Power Quality (ICREPQ)*, Santander, Spain, 2008.

[12] M. C. Benhabib and S. Saadate, "A new robust experimentally validated phase locked loop for power electroni control," *European Power Electronics and Drives Journal*, vol. 15, no. 3, pp. 36-48, Aug. 2005.

[13] R. Teodorescu, F. Blaabjerg, M. Liserre and P.C. Loh, "Proportional-resonant controllers and filters for grid-connected voltage-source converters," *IEE Proc.-Electr. Power Appl.*, Vol. 153, No. 5, pp. 750-762, September 2006.

VIII. Manuscript of Authors:

K.Jayakumar[1] has received his B.Tech degree in EEE from Gokul Institute of Technology & Sciences, Bobbili in 2009. At present he is pursuing his M.Tech degree with the specialization of power electronics from Sri Vasavi Engineering College, Tadepalligudem, A.P. His areas of interest are power electronics & drives.

N.Sriharish[2] has received the Bachelor of Engineering degree in Electrical & Electronics Engineering from Anna University, in 2005 and Master's degree from JNTU Kakinada in 2010. Currently, he is an Assistant Professor at Sri Vasavi Engineering College, Tadepalligudem, A.P. His interests are in power system, power electronics and FACTS.

Ch.Rambabu[3] has received the Bachelor of Engineering degree in Electrical & Electronics Engineering from Madras University, in 2000 and Master's degree from JNTU Anantapur in 2005. He is a research student of JNTU Kakinada. Currently, he is a Professor & HOD at Sri Vasavi Engineering College, Tadepalligudem, A.P. His interests are in power system control and FACTS.

Hotellings T-square & Principal Component Analysis Approaches to Quality Control Sustainability.

Onwuka, Gerald. I.

Department of Mathematics & Statistics, Kebbi State University of Science and Technology, Aliero

Abstract:

In this paper, Multivariate Techniques are applied to a Production/Manufacturing industry to examine various measurable characteristics of a product since no two items produced are exactly alike. Although, natural or assignable causes of variation possibly must be found in every production process, attempt has been made using this technique to examine correlation and tests for equality of means on collected multivariate data. Principal Component Analysis and Hotelling's T^2 tests were used, with the 3-characteristics measured showing negligible/low correlation with nearly all the correlation coefficients small. Only one variable possess on the average about 70% of the total variation. The result of Hotelling T^2 test shows that the average daily means are not equal and provide estimate of the interval to use. The purpose of this study is to monitor the in-control condition for data on pipes production, which clearly concludes that data for the production process was really obtained from a statistically monitored and controlled process.

Keywords: Quality Control, Multivariate Analysis, PCA & Hotelling T^2.

I Introduction

Statistics technique is applied to a manufacturing industry for the determination of quality raw materials and finished products. It is also applied in modeling manufacturing processes to determine when component could be replaced or the process turned off for turnaround maintenance. Using statistics can help increase not only the quality of the products that are being manufactured but also the quantity. The reason is that manufacturing companies can use the statistics to create a plan of action that will work more efficiently. In order to be able to effectively forecast the future productively with statistics, there would need to be a program setup that can do the following: Forecast production, when there is a stable demand and uncertain demand, Quantify the risk associated within the operations and financial costs. Predict when there will be operational bottleneck, in sufficient time before they occur. Pinpoint when and which specific model will be the cause of uncertainty. Calculate given information to show the statistical outcome. Calculate summary statistics in order to setup sample data.

If there is not a visible quality issue, then there is no way to fix it. When statistics are used to increase the quality of production and products, it is easy to track and make appropriate changes to improve this overall quality. Statistics can also help maintain the quality in the areas of business process, the mechanical and engineering process, in addition to the part production process.

With the possibility of as much up to date, and real time feedback, the quality can be increased almost instantly. In order to setup the right process for statistical tracking and predicting in quality improvement, there would need be a support process and information gathered. This is a step that many manufacturers have had a hard time completing. However, once this is done, the overall payoff of being able to improve quality through statistics has had a huge benefit, increase productivity at a lower cost. And that is what it is all about.

II Back Ground

The quality of the manufactured goods depend on a variety of factors beginning from the quality of raw materials, the process of production, the conditions of the machines and other equipments, the skills of the labour force and the inspection techniques, adopted at every stage of production. The goods can be sold in the market only if they conform to pre-determined quality standards about which the prospective buyers have been briefed. The statistical quality control helps the producer to achieve this objective by keeping the various steps in the production process within statistical control.

III Statistical Quality Control

The field of statistical quality control can be broadly defined as those statistical and engineering methods that are used in measuring, monitoring, controlling, and improving quality. Statistical quality control is a field that dates back to the 1920s. Dr. Walter A. Shewhart of the Bell laboratory was one of the early pioneers of the field. In 1924 he wrote a memorandum showing a modern control chart, one of the basic tools of statistical process control. Harold F.Dodge and Harry G. Romig, two other Bell system employees provided much of the leadership in the development of statistically based sampling and inspection methods. The work of these three men forms much of the basis of the modern field of statistical quality control. World War II saw the widespread introduction of these methods to U.S. industry. Dr W. Edwards Deming and Dr. Joseph M. Juran have been instrumental in spreading statistical quality control methods since World War II. Quality Control is regarded as the most powerful trademark behind market. (Montgomery, 2001)

Statistical quality control is based on the theory of probability and sampling and is extensively used in all types of industries in fact in all repetitive processes; the statistical quality control plays a very important and significant role. (Sharaf Eldin et al, 2006).

2.0 MATERIALS AND METHODS

Data for this study was collected from Tower galvanized products Nigeria Limited Kaduna. The company was incorporated in May 1975 and is managed by CATISA Genera. CATISA is currently managing more than 60 manufacturing operations in Nigeria and more than 400 operations in 55 countries in the world, out of which 65% of these are in Africa. Some of the well-known companies of the group in Nigeria are: Borno aluminum company Ltd. Maiduguri, Queens way aluminum company Ltd. Kaduna., Tower aluminum Nigeria PLC. Ikeja, Midland galvanizing products Ltd. Asaba. among others, the company has pioneered the development of furniture manufacturing industries in the northern states of Nigeria, pipe produced by the company are extensively used in the distribution of water for domestic and industrial use. The company's products enjoy good reputation among other products available in the market. Most of the products are distributed through the facilities available in Kaduna and Kano units. The products include: Head-pan, Cut-sheets, Long-span circular profile roofing (Aluminium, G.I. &Aluzinc) etc The scope of this study covers production data for four(4) different products produced daily by the company. The data consists of the length, circumference, and outside diameter for the four different pipes produced by the company. The pipes are: 1X1" square pipe, 7/8" round pipe, 3" round pipe, 2X2" square pipe,

The data was drawn from sample of size 5 taken from each day's production for 25 days (working days), with the aim of determining the degree of dependency between the various components of each particular product. So as to check the contribution of each component to the total variation (PCA), performing, multivariate test on the data set using Hotelling T^2 approach and to estimate confidence interval using Roy-Bose approach.

This study covers all aspects of production –raw materials, labour, equipment and management. In production industries, data on updated information about the status of the process are collected a long time. These data are frequently served to control charts, in order to see and decide whether production is operating under statistical control or if some special cause has interfered with it. (The Process).Normal operation conditions are set using the control charts, analyze the data from period of normal operations and as long as the process rests within the normal operations conditions, as soon as it moves outside such boundaries, the root of that cause is identified and corrective measures is taken to bring back to normal operation. This is because in any production process, regardless of how well designed, certain amount of inherent or natural variability would always exist.(Montgomery & Lowry, 1995). These "background noise" is the cumulative effect of many small, while a process that is operating with only chance causes of variation present is said to be in statistical control while any variability that are not part of the chance cause pattern are called assignable causes. It may result due to operator errors, improper adjusted machines and sub-standard raw materials. A process that operates in the presence of assignable cause is said to be out of control (Montgomery, 1990).

3.0 RESULTS AND DICUSSION USING THE VARIOUS TECHNIQUES MENTIONED ABOVE

3.1 PRINCIPAL COMPONENT ANALYSIS (PCA)

This is a mathematical procedure which does not require user to specify an underlying statistical model to explain the 'error' structure. (Jackson,1985) In particular, no assumption is made about the probability distribution of the original variables. It is a technique to use to get a 'feel' for a set of data. Hopefully, it assist user to a better understanding of correlation structure and may generate hypothesis regarding the relationship between the variables. Its two main objectives are;

(1) identification of new meaningful variables
(2) reduction of dimensionality of the problem as a prelude to further analysis of the data.

3.1.1 PROCEDURE IN PCA.

(a) Decide if it is worth including all the variables recorded in the original data matrix. And whether any of the variable need to be transformed.
(b) Calculate correlation (or covariance) matrix bearing in mind that a correlation co-efficient should not be calculated for pairs of variables whose relationship is non-linear.
(c) Examine the correlation matrix and observe any natural grouping of variables with 'high' correlation. However, if nearly all the correlation coefficients are "small", there is probably not much point to do principal component analysis because of the results obtained from the analysis (see appendix).. Because the variables i.e diameter, circumference and length of the 4 type of pipe i.e

- 1X1" square pipe
- 7/8" round pipe
- 3" round pipe
- 2X2" square pipe

Used in the analysis have very low correlation coefficient in the various matrix obtained.

These result force the research to the following conclusion on each of the 4 type of pipe used in the analysis. Meaning there was no need to obtain a linear combination of the variables but conduct hypothesis test using Hotellings T^2 and estimate Roy-Bose Confidence interval for each characteristics of each particular type of pipe.(Hotelling, 1947)

(1) From matrix or Table 3.11,3.12,3.13 and 3.14, (see appendix) all the correlation matrix shows a very low relationship between the variables i.e diameter, circumference and length which indicate that there is no further need to conduct principal component analysis.

(2) Also, matrix or table 3.11,3.12,3.13 and 3.14,shows that in each of the type of pipe produced, the diameter contribute higher variation to the total variation in the production process. This is clearly demonstrated by the eigen values, and from the graphs m, n, o and p prove that by the sharp drop from the first eigen value in each graph.

3.2 MULTIVARIATE TESTS

A multivariate statistical tests specifies conditions about the parameters of the population from which sample was drawn. This means that multivariate model considers that the population distribution should have a particular form (e.g a normal distribution) and also involves hypotheses about population parameters.(Morrison,2005) The reliability of the results of multivariate tests depends on the validity of these assumptions:

It is paramount to note that, the most powerful tests are those having more extensive assumptions. The T^2 test, for it to be applied the underlying conditions must be at least 70% satisfied, (Chatfield & Collins, 1980),

1. Observations must be independent (uncorrelated)
2. Observations must be drawn from normally distributed populations
3. Population must have variance-covariance matrix and mean vector.
4. Variables must have been measured in at least interval or ratio scale.

If these conditions are satisfied in the data under analysis, one can choose T^2 or wilks lambda test. When these set of assumptions are met, the tests are most likely to reject H_O when it is false.

3.3 HOTELLING T^2 DISTRIBUTION

Harold Hotelling (1931), propose a multivariate generalization of the student t-distribution and T^2 is used in multivariate hypothesis testing.

If x, x, ---,x_n ~N (μ, □^{2}) with μ and □ unknown, we can test the hypothesis; Ho: $\mu=\mu_b$ using

$t = (\bar{X} - \mu)/□ /\sqrt{n}$

So that $t^2 = (x- \mu_o)^2/(□^2/n)$

$t^2 = n (x- \mu_o)^1 (□^2)^{-1} (x- \mu_o)$.

the generalization for x_1, x_2 --, x_n,~ $N_p (\mu,\Sigma)$ with P≥1, is the Hotelling T^2 statistic

$T^2 = n(x- \mu_o)^1 s^{-1}(x- \mu_o)$

Where mean $\bar{X} = 1/n \Sigma x$, and s^{-1} is the inverse of the variance covariance matrix. n is the sample size. The diagonal elements of S are the variances of the x_i and the off diagonal are the covariance for p variables.

The multivariate techniques should posses' three important properties:

1. They produce a single answer to the question: is the process in control?
2. Has the specified type I error been maintained?.
3. These techniques must take in to account the relationship between the variables.

3.3.1 HOTELLING'S T^2 TEST

Below is the covariance result of the 3`` round pipe data.

	diam	circum	length
diam	0.00119900	0.00025500	-0.00004317
circum	0.00025500	0.01726667	-0.02476333
length	-0.00004317	-0.02476333	0.41491233

Inverse of covariance matrix

843.11	-12.89	0.00
-12.89	63.17	2.58
0.00	2.58	2.58

T^2= 47.3684.

T=14.47.

F(0.05)(3,22)=3.05.

From the result, the null hypothesis is rejected at α= 0.05, that the means are not equal i.e. the mean of diameter, circumference and length for the individual 3``round pipe differs. And that the following intervals should be considered for the acceptance of 3`` round pipes produced.

3.4 CONFIDENCE INTERVAL (Roy – Bose confidence interval) for 3`` round pipe data.

C.I= $\bar{X} \pm K\alpha/2,(n-1)\ S/\sqrt{n}$

$K\alpha/2 = \sqrt{[p(n-1)/(n-p)\ F\alpha,\ ,\ (_{p,\ n-p})]}$

C.I(diameter)= (8.99,9.03).
C.I(circumference)= (30.09,30.25).
C.I(length)= (599.22,600.04).

Also we can use, C.I= $\bar{X} \pm t\alpha/2,(n-1)\ S/\sqrt{n}$

C.I(diameter)= (9,9.02).
C.I(circumference)= (30.12,30.22).
C.I(length)= (599.36,599.9).

Note that this test was done for all 3 types of pipe with interpretation of each before the **(ROY- BOSE C.I)**

4.0 SUMMARY

In this research Quality Control using multivariate techniques tests for independence and equality of means was conducted using multivariate data. From the result of each tests using principal component analysis, the variables are found to be reasonably uncorrelated that the degree of relationship(correlation) is not strong enough. The hypothesis conducted using Hotelling's T^2 , each tests rejected the null hypothesis(H_o : means are the the same).

4.1 CONCLUSION

The multivariate tools used, principal component analysis and Hotelling T^2 are both statistical method of inference. Multivariate methods usually depend upon the assumption of a specific distribution form, for example an approximate multivariate normal distribution. And data for these methods will be in interval or ratio scale. The analysis of the data from pipe Production Company using the multivariate techniques produced a fairly reasonable result. From the analysis using principal component analysis to have a "feel" for the set of data and understanding of their correlation structure, the data set for diameter shows higher variation contribution compared to that for circumference and length. And the correlation that exists from each product variable is "substantial". Meaning nearly all the correlation coefficients are "small", there is probably not much point to further conduct complete principal component analysis. The Hotelling T^2 analysis result for the different product rejected the null hypothesis in favor of the alternative hypothesis (H_o; the means are different). The Roy-Bose confidence interval techniques tend to estimate interval for the production department to consider for the acceptance of a pipe that would be produced when the process is in statistical control with a confidence that items produced reached 95% standard quality. Based on the findings, it is therefore concluded that the process that generates the data is in statistical control.

Numerous quality control related problems are multivariate in nature, and using univariate statistical process control charts is not effective when dealing with multivariate data exhibiting correlated behavior. Therefore, multivariate statistical process control procedure provides reliable result.

APPENDIX

3.2 PRINCIPAL COMPONENT ANALYSIS

Table 3.11 shows the result of the 3`` round pipe data using PCA.

Table 3.11 :

——— **10/10/2012 9:08:34 AM** ———————————————————Welcome to Minitab, press F1 for help.

CORRELATION

	D	C	L
D	1	0.056	-0.002
C	0.056	1	-0.269
L	-0.002	-0.269	1

Covariances: diameter(cm), circumf(cm), length(cm)

	diam	circum	length
diam	0.00119900	0.00025500	-0.00004317
circum	0.00025500	0.01726667	-0.02476333
length	-0.00004317	-0.02476333	0.41491233

Eigen analysis of the Covariance Matrix

Eigenvalue	0.41645	0.01573	0.00119
Proportion	0.961	0.036	0.003
Cumulative	0.961	0.997	1.000
Variable	PC1	PC2	PC3
diameter(cm)	-0.000	-0.017	1.000
circumf(cm)	-0.062	-0.998	-0.017
length(cm)	0.998	-0.062	-0.001

Table 3.12 shows the result of the 7/8`` round pipe data using PCA.

Table 3.12 :

———— **10/10/2012 11:38:11 AM** ————————————————————Welcome to Minitab, press F1 for help.

CORRELATION

	D	C	L
D	1	0.166	0.331
C	0.166	1	0.067
L	0.0331	0.067	1

Covariances: diameter(cm), circumf(cm), length(cm)

	diam	circm	length
diam	0.00258933	0.00133333	0.00872900
circm	0.00133333	0.02500000	0.00550000
length	0.00872900	0.00550000	0.26826933

Eigenanalysis of the Covariance Matrix

Eigenvalue	0.41645	0.01573	0.00119
Proportion	0.961	0.036	0.003
Cumulative	0.961	0.997	1.000
Variable	PC1	PC2	PC3
diameter(cm)	-0.000	-0.017	1.000
circumf(cm)	-0.062	-0.998	-0.017
length(cm)	0.998	-0.062	-0.001

Table 3.13 shows the result of the 2 X 2 round pipe data using PCA.

Table 3.13 :**CORRELATION**

	D	C	L
D	1	0.043	0.074
C	0.043	1	0.42
L	0.074	0.42	1

Covariances: diameter(cm), circumf(cm), length(cm)

	diam	circum	length
diam	0.00251433	0.00419833	0.00105117
circum	0.00419833	0.03793333	0.02292167
length	0.00105117	0.02292167	0.07944733

Eigenanalysis of the Covariance Matrix

Eigenvalue	0.089695	0.028218	0.001982
Proportion	0.748	0.235	0.017
Cumulative	0.748	0.983	1.000
Variable	PC1	PC2	PC3
diameter(cm)	0.031	-0.131	0.991
circumf(cm)	0.407	-0.904	-0.132
length(cm)	0.913	0.407	0.026

Table 3.14 shows the result of the 1 X 1 round pipe data using PCA.

Table 3.14 :**CORRELATION**

	D	C	L
D	1	0.044	-0.283
C	0.044	1	0.034
L	-0.283	0.034	1

Covariances: diameter(cm), circumf(cm), length(cm)

	diam	circum	length
diam	0.00325100	0.00035167	-0.00335833
circum	0.00035167	0.01943333	0.00100000
length	-0.00335833	0.00100000	0.04333333

Eigenanalysis of the Covariance Matrix

Eigenvalue	0.043651	0.019406	0.002960
Proportion	0.661	0.294	0.045
Cumulative	0.661	0.955	1.000

Variable	PC1	PC2	PC3
diameter(cm)	-0.082	0.030	0.996
circumf(cm)	0.040	0.999	-0.026
length(cm)	0.996	-0.038	0.084

From Table 3.11,3.12,3.13 and 3.14,the correlation matrix shows a very low relationship(in appendices Table*) between the variables which indicate that there is no further need to conduct principal component analysis.

Also, Table 3.11,3.12,3.13 and 3.14,shows that in each of the type of pipe produced, the diameter contribute higher variation to the total variation in the production process. This is clearly demonstrated by the eigen values and in the Appendix section, Figures m, n, o and p prove that by the sharp drop from the first eigen value in each graph.

(3)

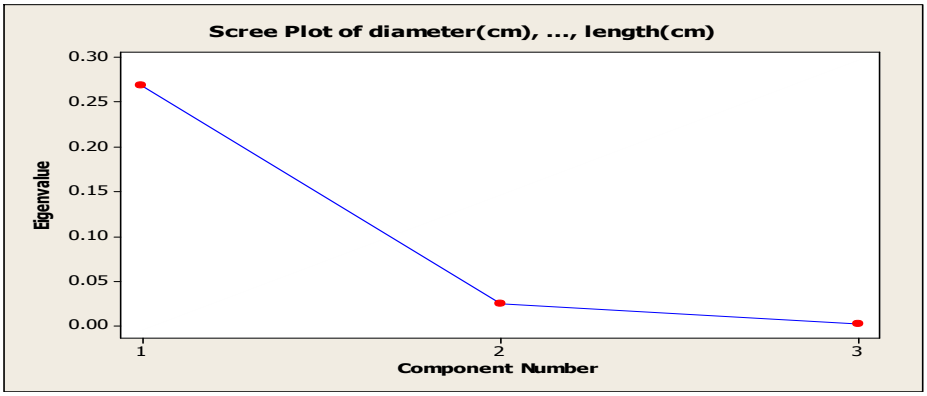


Fig 3-1

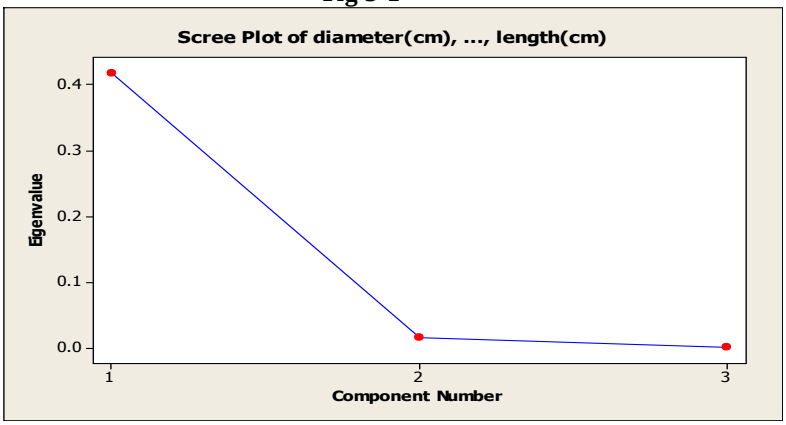


Fig 3.2

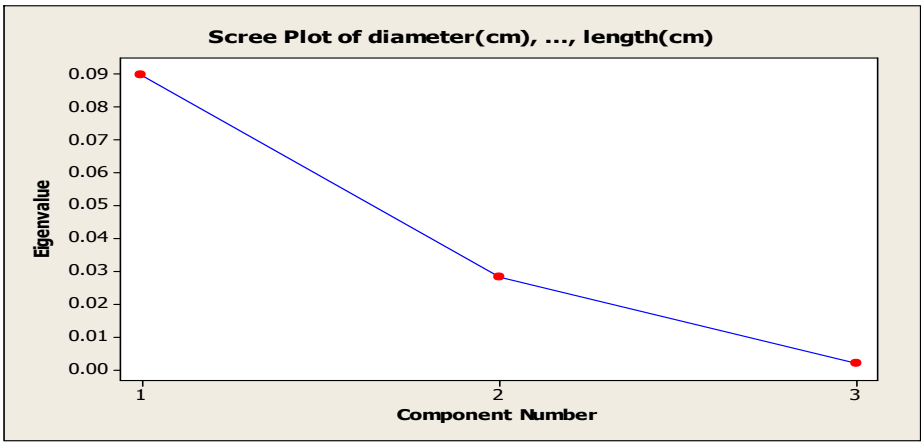
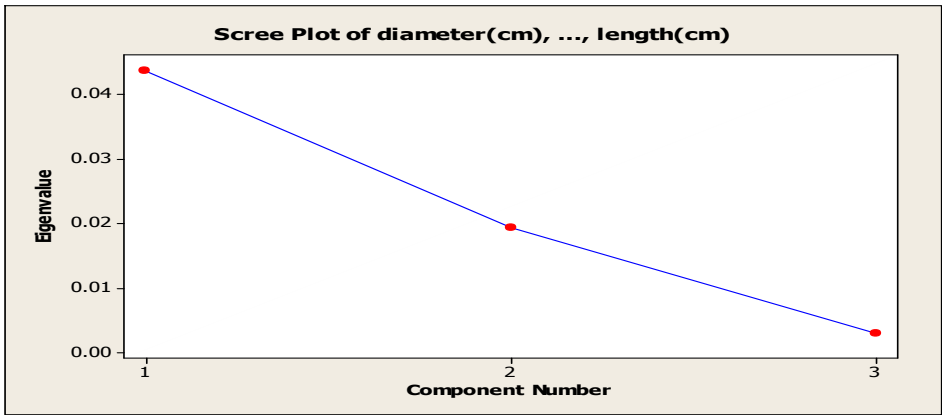


Fig 3.3



Fig, 3.4

References

[1.] Chatfield C. and Collins A.J (1980) Introduction to Multivariate Analysis
[2.] Chapman and Hall, Cambridge
[3.] Hotelling, H. (1931), The Generalisation of Students' ratio. Annals of
[4.] Mathematical Statistics No 2 pp 360-378
[5.] Hotelling, H. (1947), Multivariate Quality Control, in Techniques of Statistical analysis, edited by Eisenhart, Hastay, and Wallis. McGraw- Hill, New York, NY.
[6.] Jackson, J. E. (1985). Principle Component and Factor Analysis: Part 1– principal Component, Journal of Quality Technology, 12, pp. 201- 213.
[7.] Montgomery D, C and Lowry, C.D and (1995), "The performance of
[8.] Control Charts for Monitoring Process Variation," Communication in
[9.] Statistics computation and simulation vol. 2.
[10.] Montgomery, D.C (1990), Statistical Quality Control, 3rd edition John Wiley
[11.] Publish Ltd. New York.
[12.] Montgomery, D.C (2001), Introduction to Statistical Quality Control, 4th
[13.] edition John Wiley , USA
[14.] Morrison, D.F. (2005). Multivariate Statistical Methods. Belmont. CA.
[15.] Brooks /Cole
[16.] Sharaf Eldin, H. J Rashed and M El Khabeery(2006). Statistical Process
[17.] Control Charts applied to Steelmaking Quality improvement. Quality
[18.] Technology and Quantitative management Vol 3.
[19.] Publish Ltd. New York.

Comparative Study and implementation Mixed Level&Mixed Signal Simulation using PSpice and VHDL

[1]G.Ramachandran, [2]N.Manikanda Devarajan [3]T.MuthuManickam,
[4] S.kannan, [5] C. ArunKumarMadhuvappan [6] PM Murali
Assistant Professor, Dept of Electronics& communication Engineering
V.M.K.V Engineering College, Salem Tamilnadu, India -636308

Abstract:
PSpice CMOS & Logic Gates package that is used to analyze and predict the performance of analog and mixed signal circuits. It is very popular especially among Printed Circuit Board (PCB) engineers to verify board level designs. However, PSpice A/D currently lacks the ability to simulate analog components connected to digital circuits that are modeled using Hardware Descriptive Languages (HDLs), such as VHDL and Verilog HDL. Simulation of HDL models in PSpice A/D is necessary to verify mixed signal PCBs where programmable logic devices like Field Programmable Gate Arrays (FPGAs) and Complex Programmable Logic Devices (CPLDs) are connected to discrete analog components. More than 60% of the PCBs that are designed today contain at least one FPGA or CPLD. This paper investigates the possibility of simulating VHDL models in PSpice A/D and also coming future implemented solar cell on various sectors in the trainer kit . A new design methodology and the necessary tools to achieve this goal are presented. The new design methodology will help engineers verify a complete mixed signal design at the board level. This reduces design failures and hence increases reliability. It also reduces the overall time to market. A mixed signal design of softwares,Combinational circuits, analog circuits and Electronic Components are used in the DC motor where efficiency economy and performance are essential .this flexible ,inexpensive circuit eliminates costly PWM devices and complex floating upper rail drives, while delivering efficient motor control and Protection. The application is implemented by following the proposed design methodology.

Keyword: A/D, CMOS, CPLD,D/A, FPGA, SCHEMATIC ,PSPICE ,VHDL

1. Introduction

Over the past few decades, PCBs have brought significant change and advancement to the electronics industry. The overall cost, shape and size of many modern day electronic equipments are dependent on thesize and complexity of the PCB that is utilized. A complex board can contain printed circuitry on both its sides or in many layers, which allows great circuit density and compactness. With the increase in the complexity and reduced time to market, Computer Aided Design (CAD) tools are required to design PCBs,and the need for the CAD tool to support and automate complex design tasks have always remained persistent

The Complexity of the design, necessity to increase design efficiency and to reduce time to design, drives the circuit design methodology and the choice of Electronic Design Automation (EDA) tools. Design methodologies are typically classified into Top down design methodology enables designers to refine an abstract idea progressively as the design process continues. The design process could begin with a very high level behavioral definition of the system and then it can get down to finer details with Register Transfer Level (RTL) and gate level descriptions, as the design progresses. This methodology is more popular with digital circuit designs with the advent of HDLs, Programmable Logic Devices (PLDs) and logic synthesis tools . Traditional or bottom up design methodology allows designers to pick components individually (from a standard set of libraries) and build the design by connecting them appropriately. This methodology is popular in the PCB design flow .Large and complex systems are usually broken into smaller units. These units can be designed using different methodologies. Due to difference in levels of design abstraction, different EDA tools are required to work with different design methodologies. This curtails the ability to verify the functionality of the whole system, which is a potential cause for design failures. Moreover a majority of today's designs are mixed signal circuits, circuits containing different signal domains (eg. analog and digital). A typical example of such a design would be a PCB which has PLDs along with other discrete analog components. In awake of such scenarios, rises a need to have EDA tools that are capable of simulating mixed signal designs as well as designs designed using different methodologies. Such simulations are called mixed level and mixed signal simulation. PSpice A/D, which is very popular among PCB designer, supports mixed signal simulation using traditional design methodology. It however lacks the ability to simulate digital designs modeled using HDLs such as VHDL, Verilog etc.. By enabling simulation of VHDL models in PSpice A/D it is possible to realize a mixed level simulator from a mixed signal simulator. This integrates traditional

designing methodology with top down design methodology and hence helps in verifying the functionality of the whole system and identifying problems much earlier in the design cycles. This paper proposes a new design methodology that will allow simulation of synthesizable VHDL models in PSpice A/D along with discrete analog circuits and provides a low-cost solution for simulating mixed signal designs containing VHDL models for FPGAs and/or CPLDs. However the simulation time of such designs will be directly proportional to the number of gates produced after synthesizing the RTL VHDL.

2. Proposed Design Methodology

A design methodology to achieve total system verification at a board level is presented in figure 1. The complex mixed signal design is divided into two sections, namely analog and digital. While the analog circuits are designed by following the traditional design methodology using PSpice A/D schematic editor, the digital portion in VHDL follows the top down design methodology. Finally, the interfacing software (See Section 3B) abridges the two design methodologies by enabling functional verification of the mixed signal design in PSpice A/D. Verification at PCB level requires the simulation medium to support multiple signal domains (analog and digital) along with different design methodologies (traditional/top down). PSpice A/D is a verification software that already supports mixed signal domains. However, it has limitations in design abstraction levels and most of its digital constructs are available only at gate level. The ability to simulate VHDL models in PSpice will allow verification of designs that are defined using different methodologies. To achieve this goal, the proposed methodology employs a logic synthesis tool (Synplify®), which translates any (RTL) VHDL design into it's equivalent gate level (VHDL) description. Logic synthesis retains the essence and rapidness of top down design methodology and allows engineers to describe designs in a high level abstract and be less concerned about the actual implementation A synthesized VHDL description (gate level netlist) represents the digital system in terms of logic gates and the synthesis tool, Synplify generates a gate level VHDL description that is specific to a technology(FPGA, CPLDs architectures from vendors like Xilinx, Altera, Actel etc.) that was chosen during the logic synthesis process. In order to simulate this gate level VHDL description in PSpice A/D, the resulting netlist needs to be in a format that is understood by the PSpice simulation engine. In other words, the gate level VHDL netlist requires to be translated into a PSpice subcircuit definition. Besides this requirement, simulation of technology-specific gate level VHDL description in PSpice also requires the need for appropriate digital device models within PSpice model libraries. The choice of the technology during logic synthesis determines the ease of translation and the ability to avail or create digital device models in PSpice. Typically, CPLD architectures are relatively simple when compared with FPGAs. CPLDs implement digital circuitry in terms of combinatorial and sequential logic functions. Considering these factors the digital logic described in VHDL is synthesized by targeting at Lattice MACH 111 family of CPLDs. The gate level VHDL description generated by Synplify (synthesis tool) can be translated into a PSpice subcircuit definition and it also contains digital devices that are either currently available or that can be modeled in PSpice A/D. The design flow in this proposed methodology is as follows .The digital logic is described in VHDL and simulated to verify its functionality (top down design methodology)

The RTL VHDL code is synthesized in Synplify using Lattice MACH 111 as the target technology. The gate level VHDL description (after synthesis) is functionally verified for logical equivalency

The gate level netlist is now converted into a PSpice circuit file using the interfacing software which was developed as a part of this research.

The Circuit file is converted into a schematic symbol which can be placed on Cadence OrCAD® Capture (schematic editor) along with other analog components and the complete mixed signal design is verified by simulating in PSpice A/D
The translation of the gate level VHDL netlist into its equivalent PSpice circuit file requires
1. A library of PSpice models for Lattice MACH 111 components.

2. Interfacing software that will utilize components from this library and create a PSpice subcircuit file from the gate level VHDL netlist.

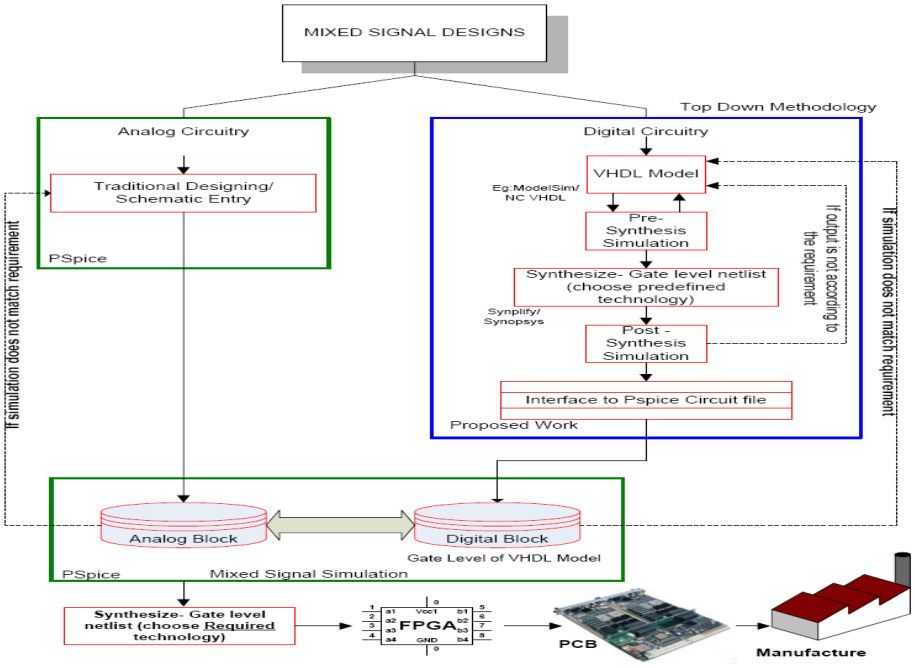


Fig. 1. Proposed design methodology

A. PSpice library of Lattice devices

The gate level VHDL netlist generated by Synplify (synthesis tool) contains components specific to Lattice MACH 111 technology. The following set of combinational and sequential logic elements from the MACH
111 family are utilized by the interfacing software during the translation process.

Combinational logic elements

1) IBUF - Input Buffer
2) OBUF - Output Buffer
3) INV - Logic inverter
4) OR2 - 2 Input logic OR
5) XOR2 - 2 Input logic XOR
6) AND2 - 2 Input logic AND

Sequential logic elements

1) MACHDFF - Reset predominant D flip flop with low preset and reset
2) DFFRH - Reset predominant D Flip flop with preset remaining HIGH all times
3) DFFSH - Reset predominant D Flip flop with reset remaining HIGH all times
4) DFF - Reset predominant D Flip flop The functional behavior of these logic elements are modeled in

PSpice and a library of device models is created for every component that is present in the gate level VHDL netlist.

B. Interfacing software

In the previous section, the creation of PSpice models for Lattice MACH digital devices was discussed. Next, the gate level VHDL netlist needs to be translated into a PSpice subcircuit file. A software program was developed to perform this task. This program will be the interface between the gate level VHDL description and PSpice A/D. Figure 2 presents a flow chart to translate gate level VHDL description into a
PSpice circuit file. The gate level VHDL netlist follows a typical pattern of structured VHDL logic description (Component declaration and definition followed by the main entity and architecture). For every
component defined in the gate level VHDL netlist, there exists an equivalent PSpice model. The interfacing program reads through the gate level VHDL netlist, identifies a Lattice MACH device and replaces it with its equivalent PSpice model in the subcircuit file which it writes simultaneously. The following procedure is followed by the interfacing software

1. Parse through the gate level VHDL netlist and skip until the Main Entity within the file is reached.
2. Within the Main Entity, extract the name of the inputs and outputs. If the inputs/outputs are declared as a BUS, elaborate the bus entries and assign individual net names for each one of the bus entries. PSpice digital device modeling language does not permit BUS declaration.
3. Using the input and output names obtained, define the subcircuit header in the PSpice circuit file by following the PSpice modeling language syntax.
4. Continue to parse the VHDL file and skip until the Main Architecture of the entity is reached.
5. Within the Main Architecture, skip the section where internal signal and component names are declared.
6. Scan the architectural definition and identify the Lattice MACH device that is being "port - mapped"
7. Use "CASE" statements to map the identified component with its equivalent PSpice model.
8. Scan the "port - mapping" definition to identify the input and output net names and assign them to appropriate PSpice model terminals.
9. Loop until the end of architecture section is reached.

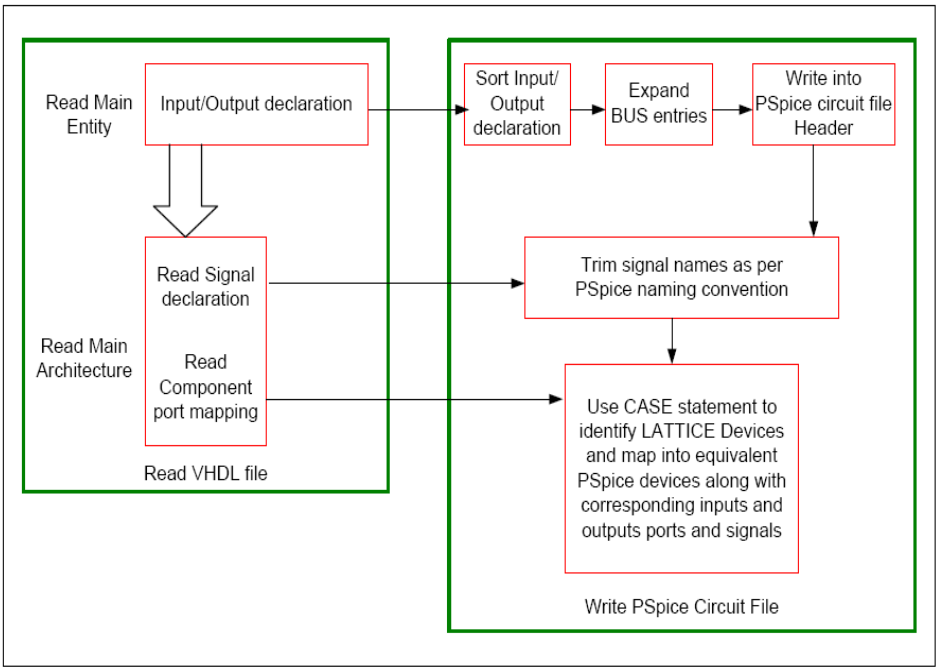


Fig. 2. Flow chart for VHDL-PSpice Conversion program

3. Mixed signal Design:

The ICM7555 and ICM7556 are CMOS RC timers providing significantly improved performance over the standard SE/NE 555/6 and 355 timers, while at the same time being direct replacements for those devices in most applications.

Applications

1. Precision Timing
2. Pulse Generation
3. Sequential Timing
4. Time Delay Generation
5. Pulse Width Modulation
6. Pulse Position Modulation
7. Missing Pulse Detector

Two low cost CMOS ICs manage a 12 VDC ,current limited speed control circuit for DC brush motors .the circuit design uses PWM to chop the effective input voltage to the motor. use of cmos devices gives the benefits of low power ,minimal heat and improved longevity. the overall design is simple ,inexpensive and reliable ,and is useful in application such as embedded .

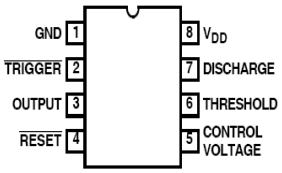


Fig1. ICM7555 (8 LD PDIP, SOIC)

Although the supply current consumed by the ICM7555 and ICM7556 devices is very low, the total system supply current can be high unless the timing components are high impedance. Therefore, use high values for R and low values for C in Figures 2

Circuit concept:

Low –cost DC Motor Speed control with CMOS ICs - A CMOS update of the popular -555 timer device is used because it draws much lower operating currents and then runs cooler and lasts longer in typing operating environments, by adjusting the wiper of the speed –control potentiometer, output-signal duty cycle ,or pulse width ,can be varied from 2% to 98%.operating frequency is fixed at 20khz,to remain in the inaudible range.

Two signals are generated: The PWM signal ,and a direction signal (high=forward ,low=reverse).these signals become inputs to the TC4469, a CMOS quad MOSFET driver .its logic inputs allow proper output chopping and commutation. Similar way using ICs&transistor PWM Driving DC Motor using embedded Technology and its implemented on VLSI Technology(Fig2 &3)

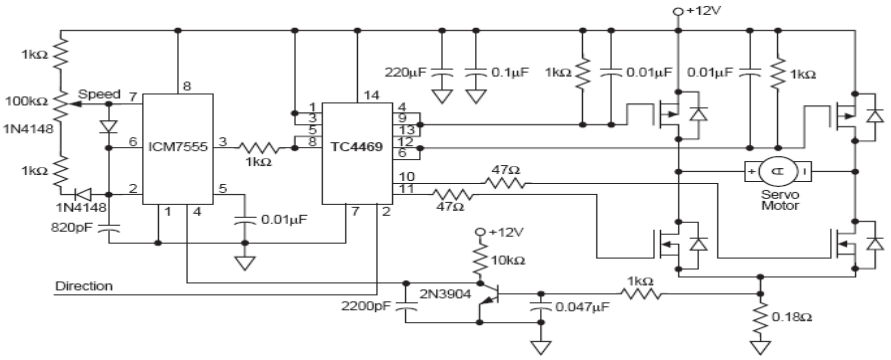


Fig .2. 12VDC Speed control and current limit

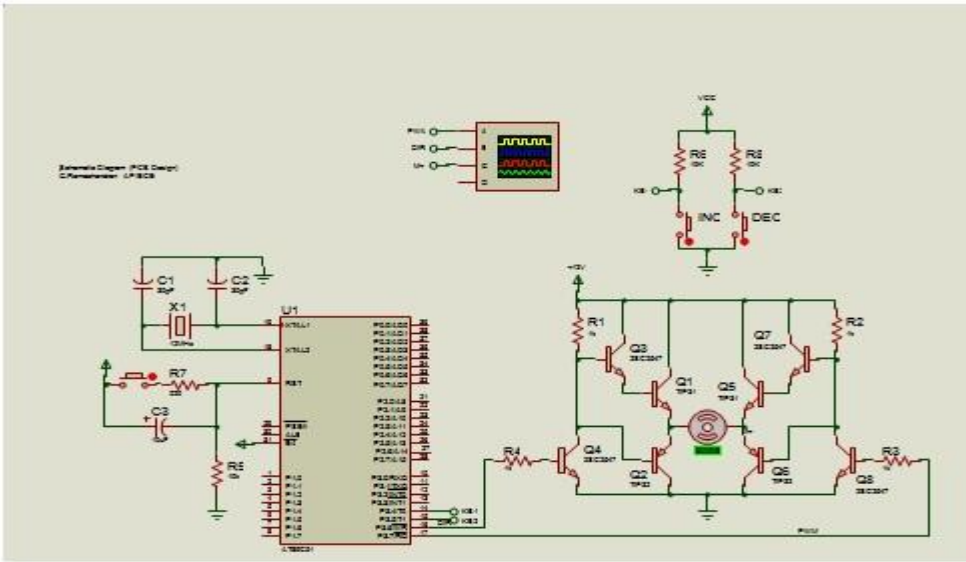


Fig 3. PWM Driving DC Motor

Observe the pulse train applied and motor voltage on the osciloscope.You can change the motor speed using the toggle buttons during simulation This design demonstrates the use of an AT89C51 to control a DC motor using PWM Four Transistors-2SC2547

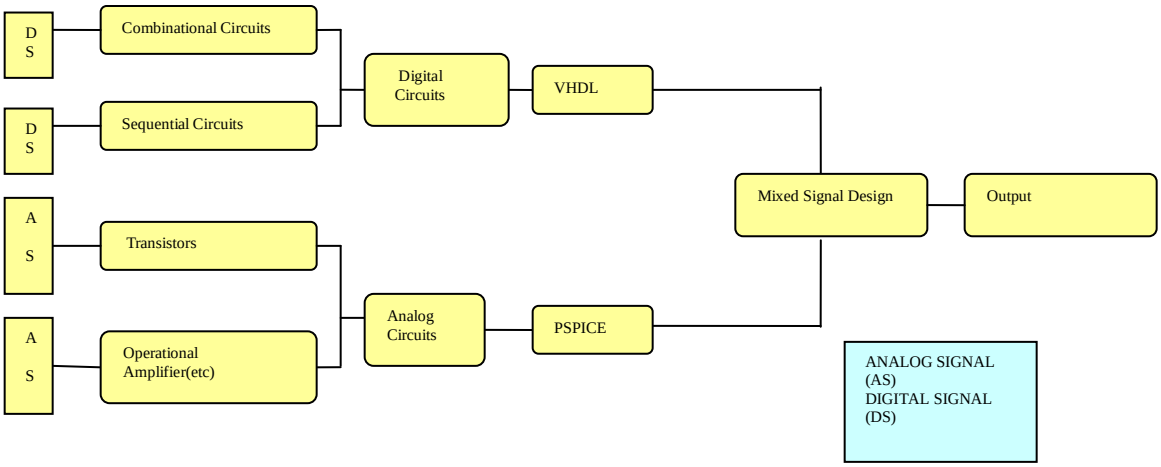


Fig 4 Mixed Level and Mixed Signal Simulation using PSpice and VHDL

3.1Schematic Diagram

3.1.1RTL Schematic

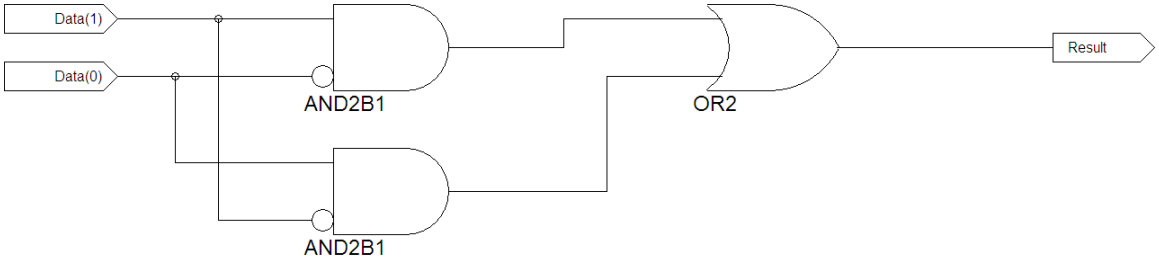


Fig 3.1.1. Basic RTL Schematic

3.1.2Technology Schematic

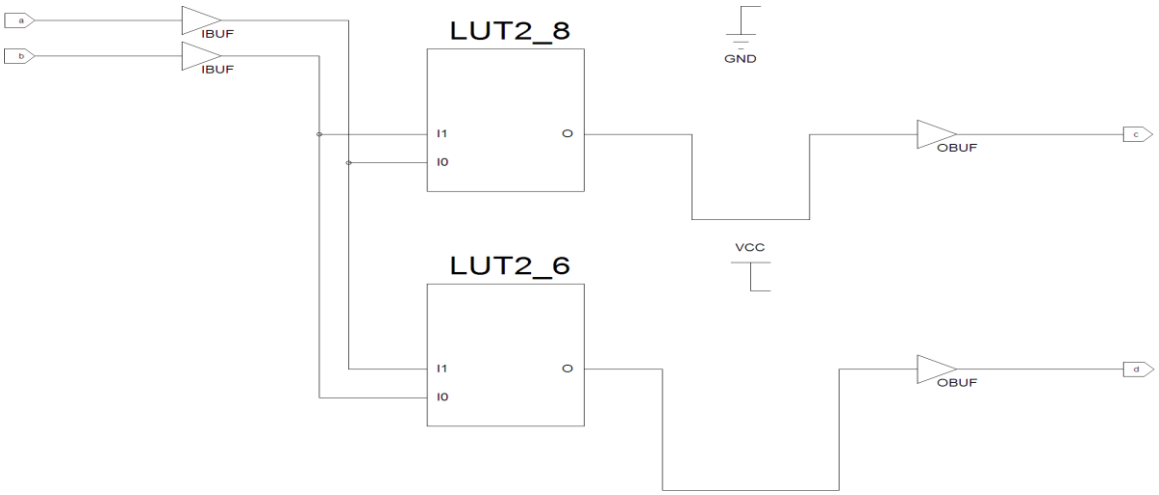


Fig 3.1.2LUT 2 INIT 6 &LUT 2 INIT 8 Technology Schematic

4. Simulation Results:- Mixed level &Mixed signal simulation using VHDL

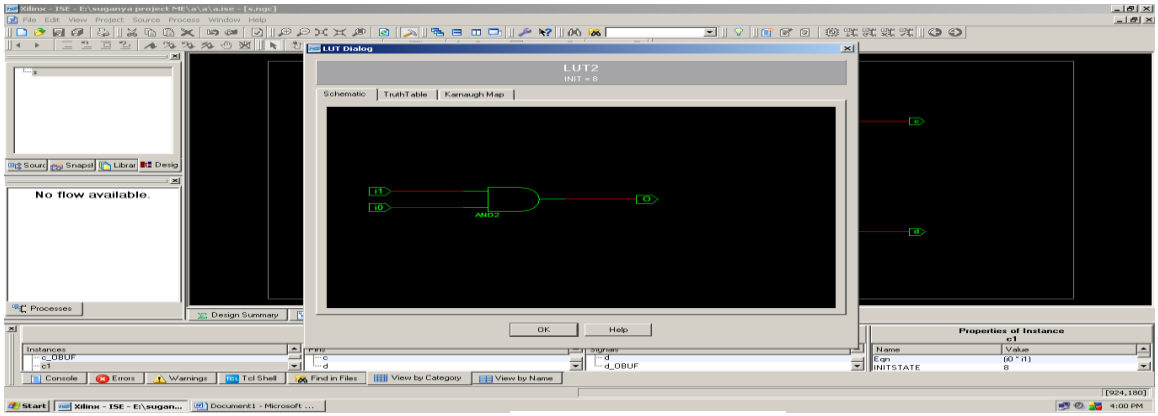


Fig 1. Basic AND, NOT and OR gates

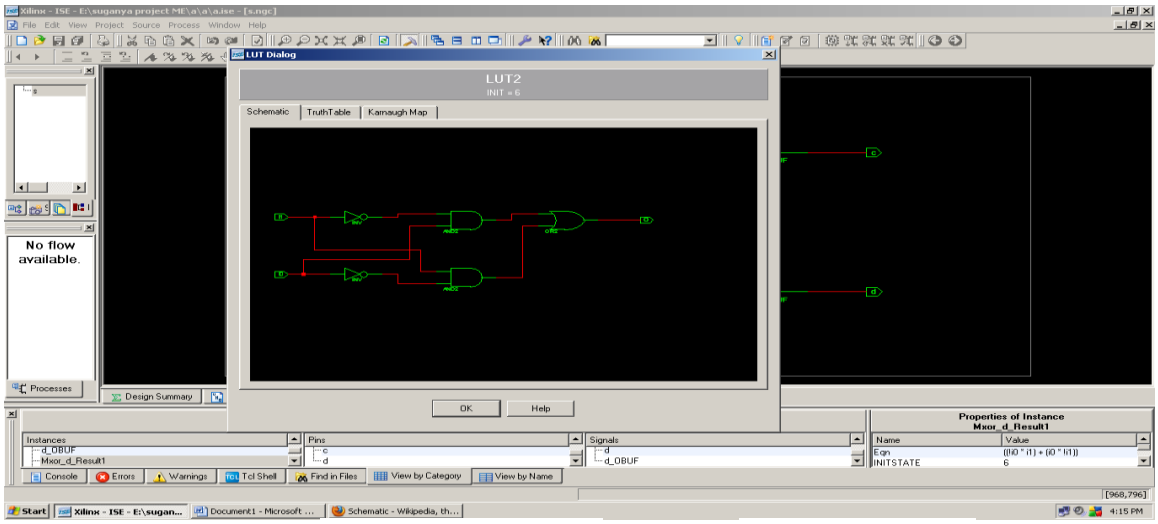


Fig 2. Basic AND, NOT and OR gates Universal ie NAND or NOR gates

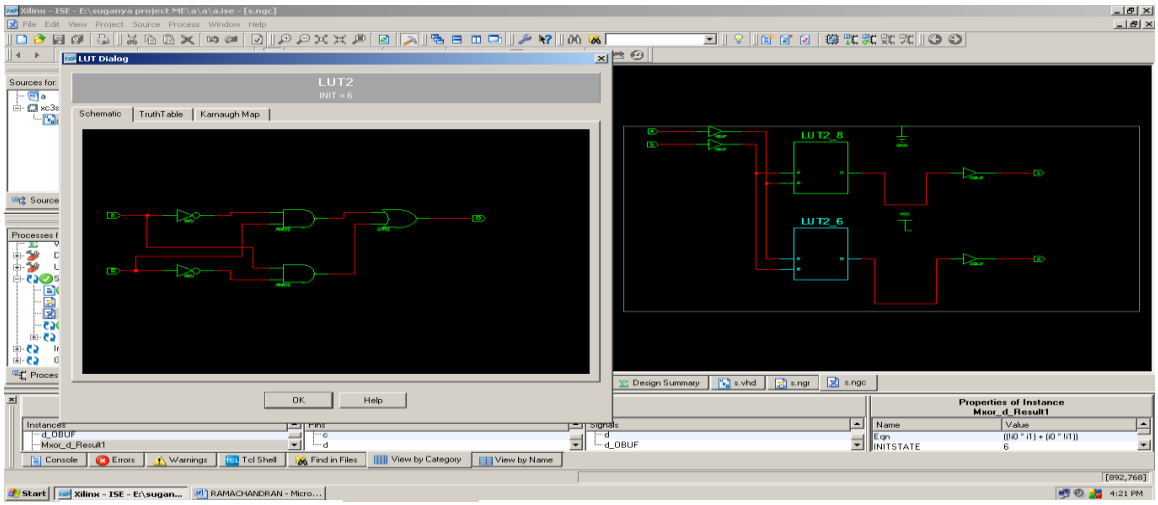


Fig 3. Basic AND, NOT and OR gates Universal ie NAND or NOR gates& Combinational: X-OR GATE or X-NOR Gate

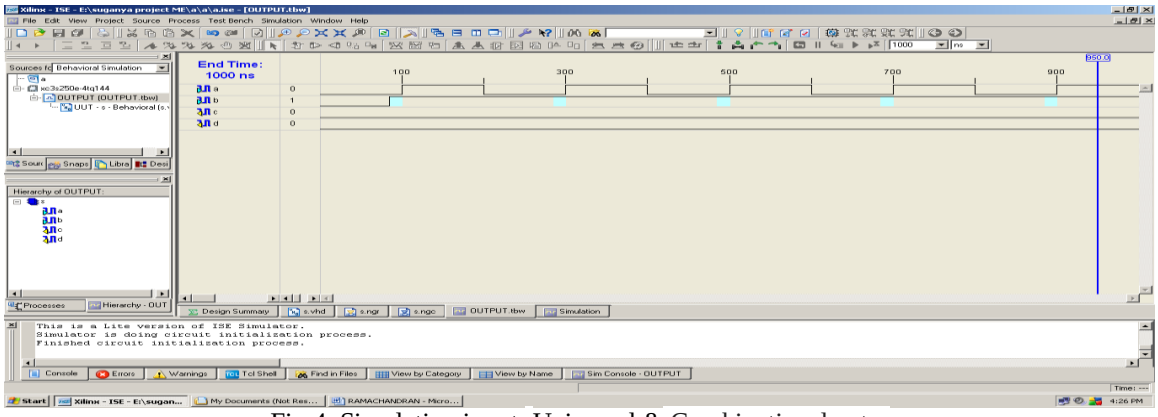


Fig 4. Simulation input: Universal & Combinational gates

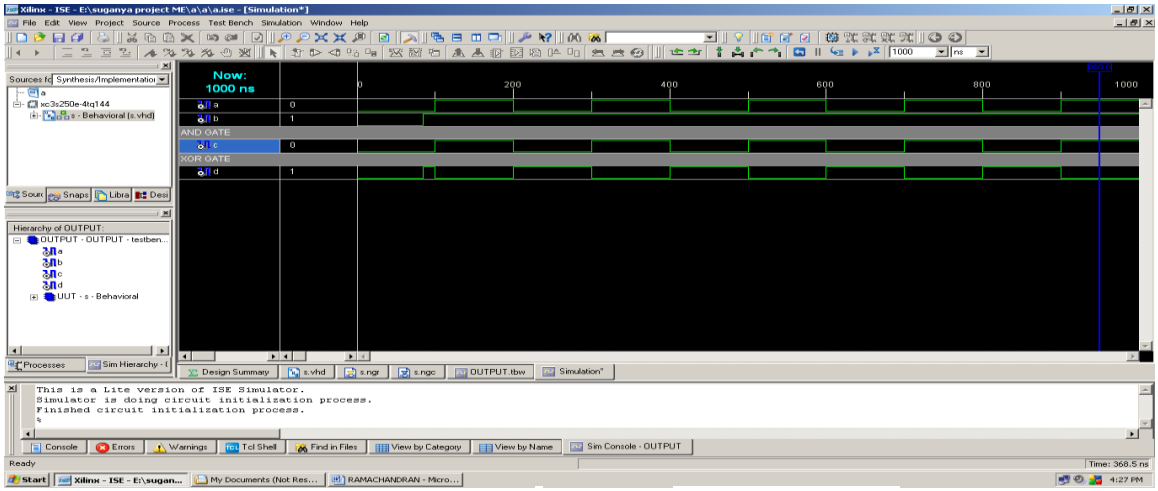


Fig 5. Simulation Output: Universal & Combinational gates

5.Dsh: Schematic: Mixed level &Mixed signal simulation using PSPICE

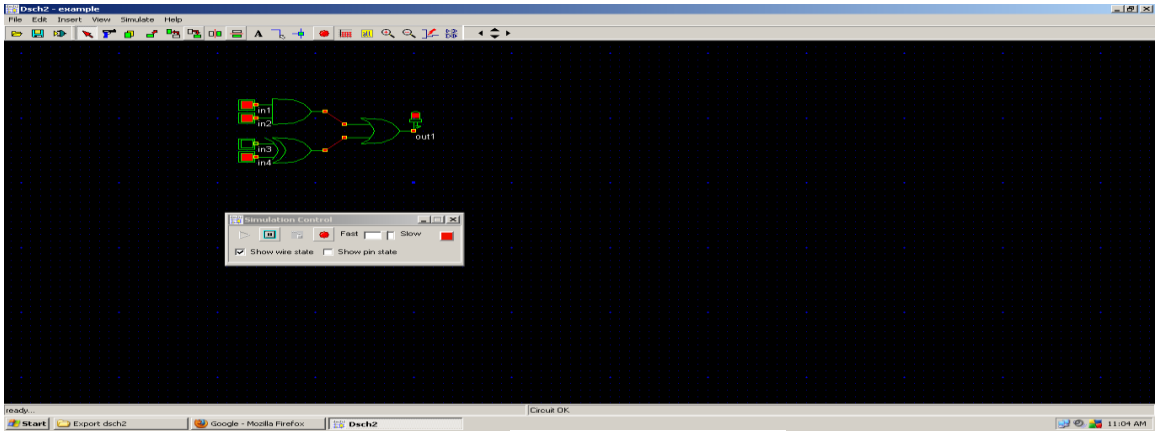


Fig 1. Basic AND, NOT and OR gates

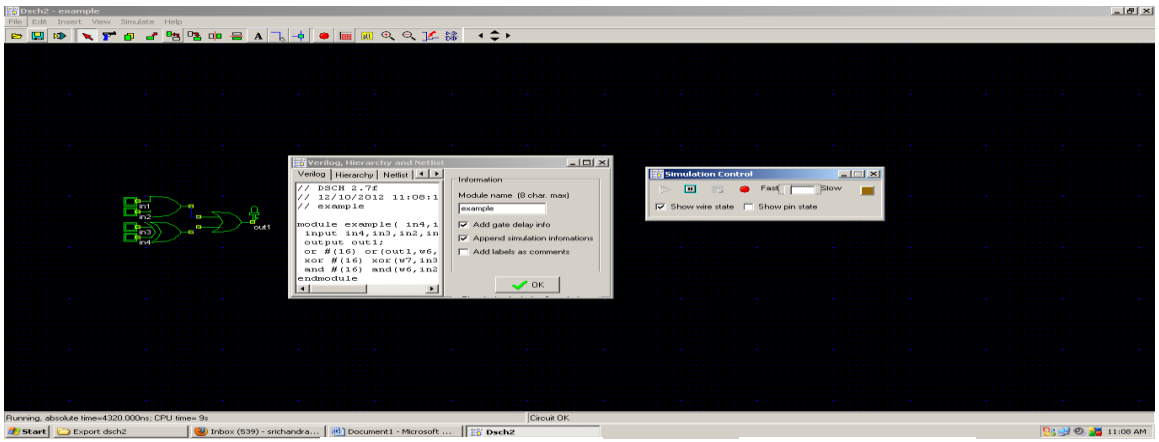


Fig 2. Basic AND, NOT and OR gates Universal ie NAND or NOR gates

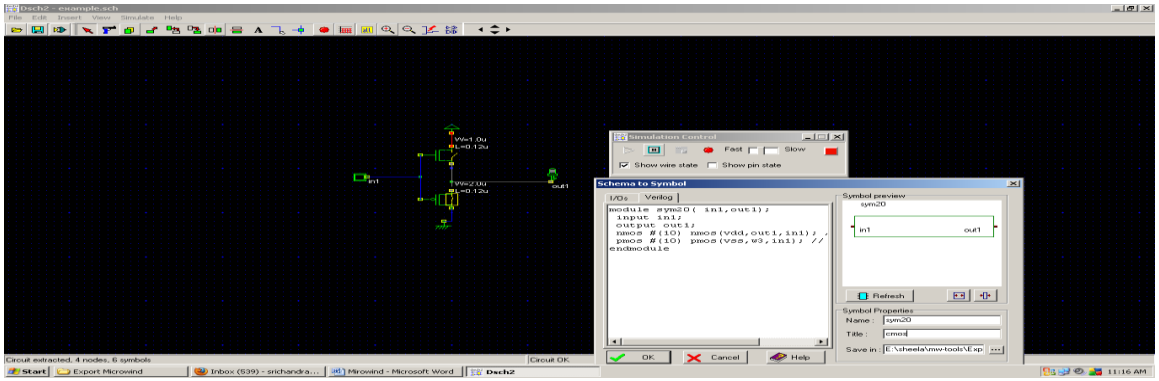
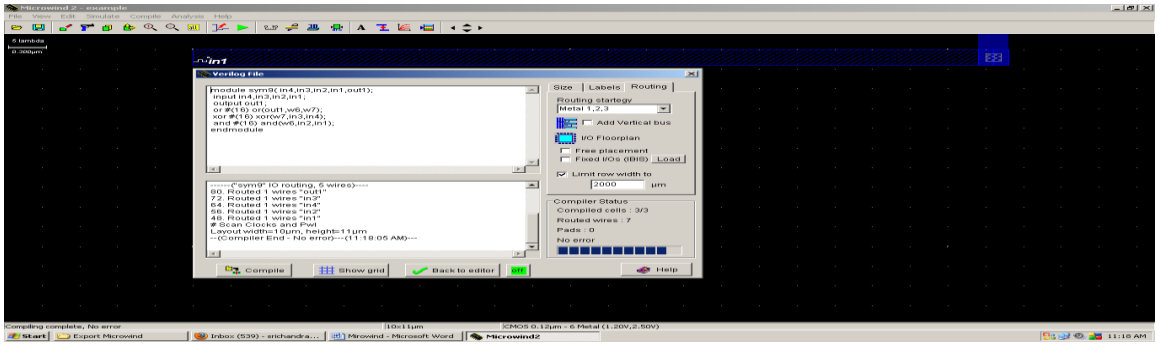


Fig 3. CMOS(Basic AND, NOT and OR gates Universal ie NAND or NOR gates& Combinational: X-OR GATE or X-NOR Gate)



Execution of **CMOS** (Basic AND, NOT and OR gates Universal ie NAND or NOR gates& Combinational: X-OR GATE or X-NOR Gate)

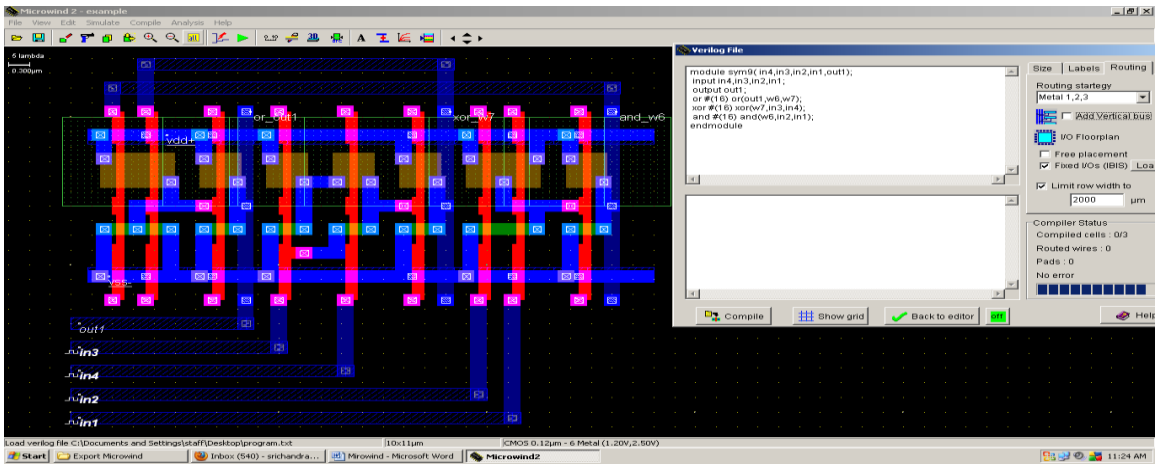


Fig. 3. Block diagram of the design from CMOS Layout

5. Conclusion

Prior to the proposed design methodology, circuit designers had to duplicate the output signals (Johnson counter) from the FPGA with clock generator functions in PSpice. Subjecting the entire design to various input conditions and performing total system verification at the design level was difficult. The new design methodology and the additional tools (*interfacing software and PSpice device library for Lattice devices*) enabled the simulation of the VHDL model in PSpice A/D. It was also possible to precisely model the arrival times of input signals from the VHDL model in PSpice, which in turn produced accurate conversion of DC power to AC. Simulating the entire design at the PCB level allowed engineers to study the behavior of the analog circuits when controlled by the FPGA. The simulation results presented in the previous sections asserts the following

Accurate conversion of the gate level VHDL description into an equivalent PSpice sub circuit file The customized PSpice library of digital devices for the LATTICE MACH 111 technology were modeled accurately

The overall time to run simulation is increased due to the gate level representation of digital logic and the clock interval of the master clock.

Total system verification at the design level was made possible by co-simulating digital circuits modeled in VHDL along with discrete analog components in PSpice. Violation of timing conditions, analog to digital interface problems etc. are some of the typical issues that can be addressed earlier in the design cycle due to
the proposed design methodology.

In this work the possibility of mixed level simulation in PSpice A/D using VHDL was investigated. A new design methodology to simulate synthesizable VHDL models in PSpice A/D was proposed. It combined traditional and top down design methodology in one simulation environment. To achieve this goal additional tools (interfacing software and PSpice device libraries) were required. These tools were developed and their functionality were verified in a practical application (case study). The new design methodology enables verification of mixed signal PCBs that contain PLDs like FPGAs, CPLDs, etc. at the design level and introduces an additional check-point before taking the design to hardware. Overall, it showed an effective way to verify the functionality of the mixed signal design. In examples such as the one studied in this work, different design teams work on different sections of the same design, namely analog and digital. In such situations, where the possibility of error introduced in the design process is high, this proposed methodology provides an easy and simple way to verify the functionality of the entire mixed signal design and provides an excellent chance to identify integration problems much earlier in the design cycle. However, the disadvantage is the increase in overall simulation run time.

6. FUTURE WORK

The following areas were identified for further research , Increased simulation time. Using gate level VHDL netlist can become a bottleneck if the number of gates in the netlist exceeds a few thousand. In order to reduce the simulation time, further investigation on behavioral simulation methods of the VHDL code in PSpice is necessary.

MATLAB Simulink® is a system level simulation tool. It has the ability to simulate electro-mechanical systems and it can integrate with PSpice simulation engine. If these abilities can be combined with the new design methodology then it can provide co simulation between system level, circuit level and HDL based designs.

Improved graphical user interface. The proposed methodology involves repeating a sequence of commands in each of the EDA tools, which can be scripted and automated. In doing so, a lot of details can be hidden from an end user and this would increase the appeal of the solution.

References

[1]. Peter Ashdon. The Designers Guide to VHDL 2nd Edition. Morgan Kaufmann, United States of America, 2002.
[2]. Vaughn Betz, Jonathan Rose, and Alexander Marquardt. Architecture and CAD for Deep Submicron FPGAs. Kluwer Academic, United States of America, 1999.
[3]. Dave Brady and Tom Dewey. Are the benefits of using FPGA's consumed by the obstacles of integrating the FPGAs on Printed Circuit Board? Mentor Graphics White Paper, 2003.
[4]. Cadence. PSpice Reference Guide, 2005. Version 10.5.
[5]. Cadence. PSpice User's Guide, 2005. Version 10.5.
[6]. G.C. Caprini, F. Innocenti, L. Fanucci, S. Ricci, G. Taraschi, P. Terreni, M. Tonarelli, and L. Tosi.
[7]. Embedded system for brushless motor control in space application. MAPLD International Conference, pages 151 – 156, 2004.
[8]. N.L. Eastman. Considerations for Mixed analog/digital PCB design. WESCON/96, pages 297 – 301,1996.
[9]. B Fawcett. Synthesis for FPGAs: an overview. WESCON/94. 'Idea/Microelectronics'. Conference Record, pages 576 – 580, 1994.
[10]. Peter Frey and Radharamanan Radhakrishnan. Parallel mixed technology simulation. Fourteenth Workshop on Parallel and Distributed Simulation, pages 7 – 14, 2000.
[11]. Bashir Al Hashimi. The art of Simulation using Pspice Analog and Digital. CRC Press, United States of America, 1995.
[12]. K. Kundert, H. Chang, D. Jefferies, G. Lamant, E. Malavasi, and F. Sendig. Design of mixed – signal systems - on-chip. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 19:1561 – 1571, 2000.
[13]. Lattice Semiconductor Corporation. Lattice MACH 111 Reference Guide.
[14]. Synplicity. Synplify Pro Reference Guide, 2004. Version 7.5.1.
[15]. G.Ramachandran, T.MuthuManickam, B.SuganyaAbiramavalli,Study And Implementation Of Green Power In Campus Environment International journal of Electronics and Communication Engineering &Technology (IJECET) Pages 325 – 331 2012

Finite Element Simulation of Single Stage Deep Drawing Process for Determining Stress Distribution in Drawn Conical Component

Shishir Anwekar[1], Abhishek Jain[2]

[1, 2] Mechanical Engineering, University Institute of Technology, Barkatullah University, Bhopal, India

Abstract:

To avoid the expensive and difficult experiments, modern manufacturing is based on the analysis of forming processes in numeral atmosphere before the actual production set-up. In the presented work , the single stage deep drawing process of thin walled, mild steel, conical back plate of radial impeller of blowers is approached by means of a finite element analysis. Simulation of the drawing process for determining stress distribution in the drawn component for a particular displacement is explained in the presented work. The distribution of stress in the drawn component is obtained. The study was conducted by using ANSYS12.0, in which, two models have been tested. Both models constructed solely out of axisymmetric, quad 4 node, PLANE 42 elements which have been used to simulate the drawing process of drawing quality mild steel IS2062 grade. The experimental analysis is carried out on two different flat plates having thickness 3 mm and 5 mm from which the conical back plate is manufactured. This study will be beneficial to the tool designer and the manufacturers doing work in this field

Key words: Deep drawing , Finite element simulation, Forming, Manufacturing, Sheet metal, Tool designer,.

1. Introduction

Casting, machining, welding and metal forming are the main methods of manufacturing .Sheet metal forming is an important manufacturing process for producing a large variety of automotive parts and aerospace parts as well as consumer products. Deep drawing is a forming process involving variety of material flow conditions. With the developments in the technology the design of deep drawing is an art than science still today .It depends on the knowledge and experience of the design engineer only. The selection of various parameters is still based on trial and error methods [1],. The use of numerical simulation could contribute towards the development and optimization of the process, leading to considerable economic and technical gains.

Deep drawing is one of the most important sheet metal forming processes in which a 2-d part is shaped into a 3-d part by deep drawing. In deep drawing, a flat blank of sheet metal is shaped by the action of a punch forcing the metal into a die cavity. The application of the finite element method to the numerical simulation of the deep-drawing process has evolved in a significant way in the course of the last few years. Many of the problems associated with numerical simulation of this process have been solved or at least are better understood. Reviewing the various literature available on simulation of drawing, it is understood that most of the research work are focused on drawing of cylindrical product. Relatively a few research work has been done for finite element simulation conical product. The conical shaped product made on hydraulic power press was extensively used in the engineering and day today life.

In the study of Saad Theyyab Faris [1], a numerical procedure was proposed for the design of deep drawing process using finite element method through program code ANSYS 5.4 simplified 2-D axisymmetric model of cylindrical cup are been developed. M. Afteni et al., [2] put forward that the increasing demands for small devices which have multiple applications in automotive industry, in chemical industry but also in medicine leads to new approaches concerning both the simulation and the experimental analysis of the material forming processes. They presented the experimental and numerical studies regarding the micro-deep drawing of Nickle sheets. According to Dr.Sc. Amra Tali ikmi et al,[3] deep drawing is a process for shaping flat sheets into cup-shaped articles without fracture or excessive localized thinning. Their paper describes the use of ABAQUS finite element code in a single stage sheet metal forming simulation on rectangle cup deep drawing. They suggested that the main objective of numerical simulation of the forming process is to reduce the development time of a new product. Eric T. Harpell et al., .[4] modelled various tooling geometries by using finite element analysis. But in their study they worked on aluminium sheet. L.F. Menezes et al. [5] presented a three-dimensional mechanical model for the numerical simulation of the deep-drawing process. The model takes into account the large elastoplastic strains and rotations that occur in the deep-drawing process. According to Abdolhamid Gorji .[6] forming conical parts is one of the complex and difficult fields in sheet-metal forming processes.

Simulation of elastic-plastic behavior of mild steel sheet is carried out with non-linear condition to gain accurate and critical understanding of sheet forming process byAccording to Laxmiputra M Nimbalkar et al. [7]

2. Process Description

During the process a piece of sheet metal is clamped between the die and the blank holder. A force is applied to the blank holder to prevent wrinkling of the sheet and to control the material flow during the deformation. When the punch is pushed in to the die cavity the sheet deforms plastically and thereby it takes the specific shape of the tools.[8] An example of such a deep drawing part is given in figure 1.

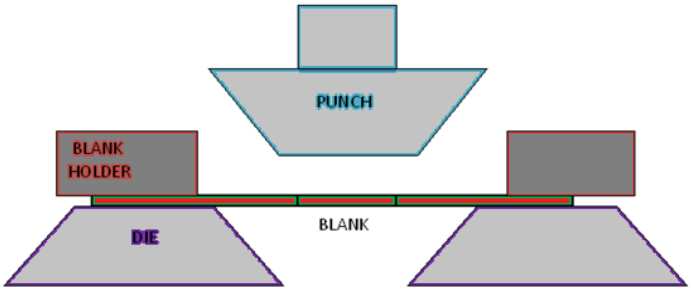


Fig. 1: Set up of a drawing tooling for making conical shaped product from a flat plate

3. Outline Of Finite Element Analysis

Finite element analysis (FEA) has become commonplace in recent years. Numerical solutions to even very complicated stress problems can now be obtained routinely using FEA. Finite element analysis is a simulation technique which evaluates the behaviour of components, equipments and structures for various loading conditions. It is a numerical method , that is used in the analysis of complex mechanical and structural problems. It is used to obtain numerical solutions and can be applied over a wide range of objective functions and over a wide range of loading conditions. The method can be used for analysis of static problems and dynamic problems and can even be used to analyze linear systems and non-linear systems.

A linear system is a type of system where a linear relation-ship exists between the force and deflection and these systems do not take plastic deformation in account. A non-linear system is a type of system that takes plastic deformation into account and can allow testing all the way up to the points of fracture. FEA can be applied to problems that involve heat loss, fluid flow and even electric potential. Due to this wide range of application, the method is very popular. The method is used in the analysis of many complex two dimensional problems, but can also be applied for problems involving three-dimensions. As problems become more complex with finite element analysis, they can take longer to solve. Computers are used to solve complex finite element problems and a wide range of finite element analysis software exists today e.g. LUSAS, ABAQUS, ANSYS etc..

The method works by taking a problem and modeling it into smaller elements through the use of nodes and elements. Physical and geometric properties are assigned to the elements and loads and displacements are applied to the nodes. From that, a finite element analysis is carried out in an attempt to earn numerical values.

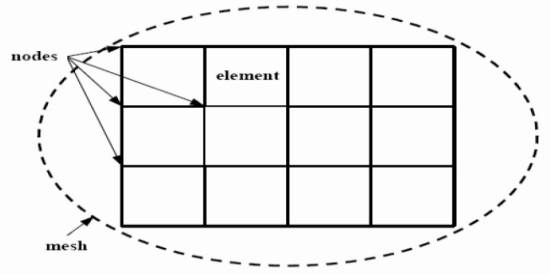


Fig. 2: Finite Element Mesh structure

4. Methodology

Two selected models of back plates of impeller of a blower are investigated both experimentally and by numerical simulation for. The blank from which these back plates are formed have the dimensions as as given in TABLE 1. The material of back plates is mild steel IS 2062 grade. A commercial FE code ANSYS 12.0 structural was used to simulate the deep drawing operation. The thickness of drawn portion is t_f.

The punch is pushed into the die cavity, simultaneously transferring the specific shape of the punch and the die to the sheet metal blank, thus forming a conical cup. The basic shape of punch is frustum of a cone. The back plate is drawn from the blank in a press by the force of the punch. The experimental set up is shown in fig 3.

Table 1 Parameters of experimental set up

Sr. No.	Description	Symbol	Model No.1	Model No.2
1	Top outer diameter of Punch	D_{PT}	450 mm	370 mm
2	Bottom outer diameter of Punch	D_{PB}	251 mm	206 mm
3	Top inner diameter of Die	d_{DT}	455.76	379.34 mm
4	Bottom inner diameter of Die	d_{DB}	256.76	215.34 mm
5	Blank Thickness before drawing	t_i	3.06 mm	5.05 mm
6	Depth of Drawing	h	67.25 mm	80.54 mm
7	Blank outer diameter	D_{BO}	930 mm	540 mm
8	Blank inner diameter	D_{BI}	214 mm	145 mm
9	Punch nose angle	α	30 degree	43 degree

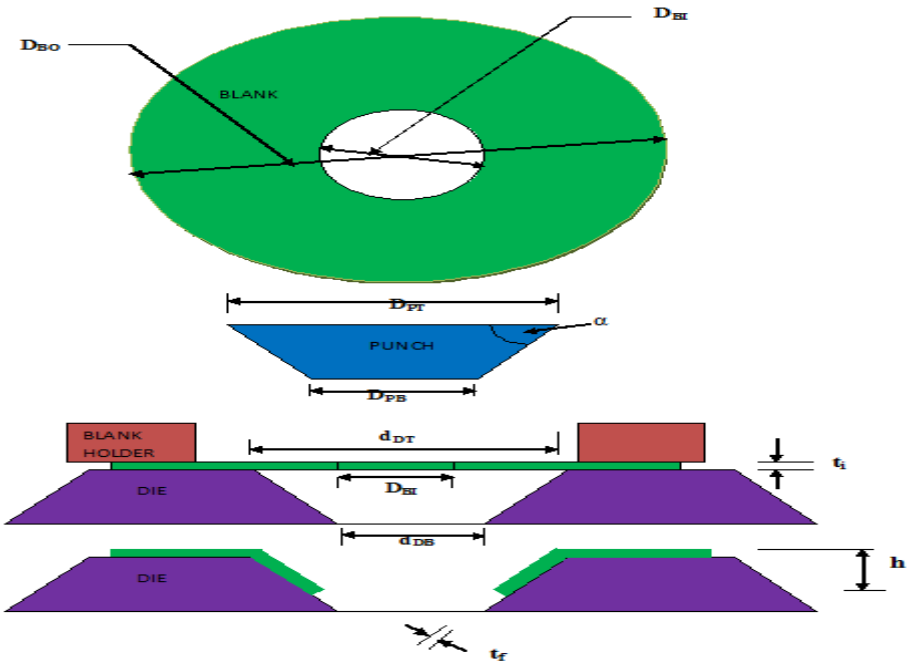


Fig. 3: Diagram of experimental set up

5. Simulation Procedure

5.1 INDUCTION

The stages of this simulation work is summarized sequentially in the following steps :
Preprocessing: (defining the problem)
Define key points/lines/areas/volumes (Solid Modeling)
Define element type and material/geometric properties
Mesh lines/areas/volumes as required
Solution: assigning loads, constraints and solving;
Apply the loads
Specify constraints (translational and rotational)
Finally solve the problem.
Post processing: further processing and viewing of the results;
Lists of nodal displacements and show the deformation
Stress/strain contour diagrams

5.2 SOLID MODELING

It is the process of creating solid models in CAD system. A solid model is defined by volumes, areas, lines, and key points. By using various geometrical data of presented setup as given in table 1 a solid modeling of the same is created in ANSYS.

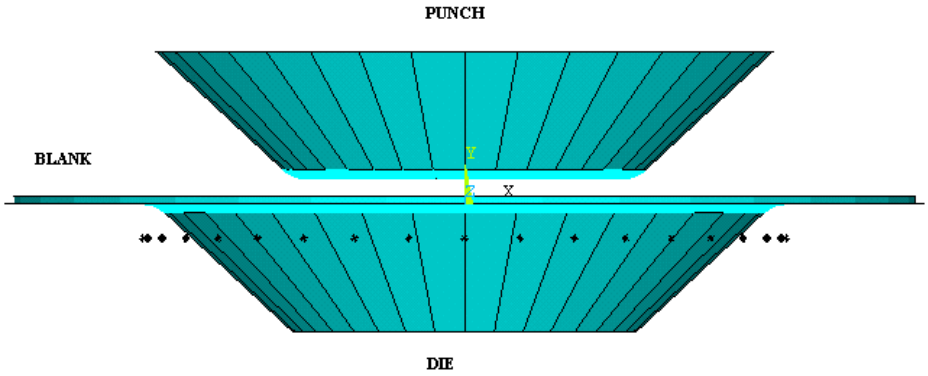


Fig. 4: ANSYS solid modeling of punch, blank and die

The geometry built in ANSYS is shown in fig-4. The structure is divided to ease map meshing of the problem. Work plane options are used to divide the structure. An axisymmetric approach is used to built the geometry. Axisymmetry is the best option to built and analyze deep drawing conical formations.

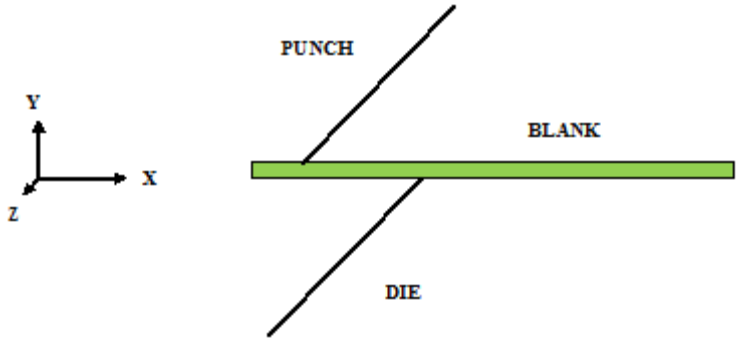


Fig. 5: ANSYS axisymmetric modeling of punch, blank and die

5.3 ELEMENTS

The type of element to be used in the analysis influences the exactness and accuracy of the results to a great extent. Literature review and examination of peer researchers' works show that PLANE42, 2-D elements with axisymmetric behavior have been conveniently used in the numerical analysis of axisymmetric forming process. This element is capable of representing the large deflection effect with plastic capabilities. This element can be used either as a plane element (plane stress or plane strain) or as an axisymmetric element. The element is defined by four nodes having two degrees of freedom at each node: translations in the nodal x and y directions. The element has plasticity, creep, swelling, stress stiffening, large deflection, and large strain capabilities. [10],[11],[12]

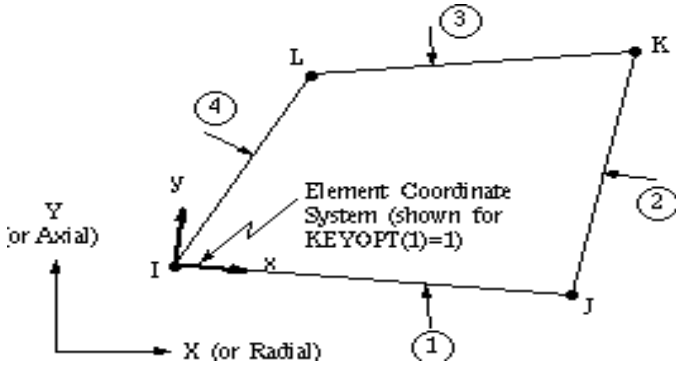


Fig. 6: PLANE 42, 2-D Structural Solid Element

Element size plays an important role throughout a simulation. Element size affects both the computation time and the accuracy of the results. The blank should be meshed finer enough in order to get acceptable results. However, the increase in number of elements results in a drastic increase in computational time. In the presented study in order to achieve good results, size of element is for model number 1 taken as 1.53 mm and for model number 2 it is taken as 2.525 mm.

5.4 MATERIAL PROPERTIES

The material of back plates is mild steel IS 2062 grade. It was selected as a structural, non-linear, isotropic hardening material model in the presented ANSYS simulation. Material properties like yield stress, Modulus of elasticity, Poisson's ratio etc, which are required for fem simulation are obtained from various authentic literature. Tools are assumed as rigid, so there is no need to define material, however the material of punch and die is tool steel.

Table 2: Mechanical Properties of mild steel IS 2062 : source [13]

Sr. No.	Properties	Value
1	Tensile Strength	410 MPa
2	Yield Stress	250 MPa
3	Modulus of elasticity	200 GPa
4	Poisson's Ratio	0.3
5	Friction Coefficient	0.1

5.5 ANSYS MESHING

A quad mapped mesh was generated on all areas apart from the punch/die which is taken as rigid. This was done to achieve a higher number of elements along this line so a solution using contact conditions could be found easier. The figure 16 and 17 shows meshed model of the problem. 4 node PLANE42, 2-D elements is used to mesh the structure. The mapped mesh is good for accurate results as well as for graphical representation which is not proper with free mesh. An expansion option available with ANSYS is used to represent in the three dimensional space. Number of elements in the mesh for model number 1 and 2 are 468 and 158 respectively. Number of nodes for model number 1 and 2 are 1669 and 639 respectively.

Fig. 7: ANSYS meshing for the bank (side view)

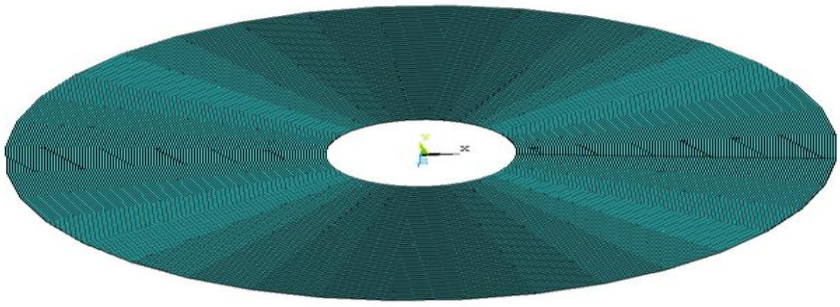


Fig. 8: ANSYS meshing for the bank (top view)

5.6 LOADING & BOUNDARY CONDITIONS

As part of FE analysis, applying loads and constraints i.e. boundary conditions consists of defining which parts from geometrical model moves i.e. defining degree of freedom. Contact surfaces used in the presented work are top blank - bottom punch, bottom blank - top die. In current study movement of blank part is restricted in x- direction. Displacement load of the part of the blank which initially not in contact with die is given in y-direction. Movement of horizontal part of the blank which is on the die is restricted in x- direction as well as in y-direction both. The tools i. e. punch, die and blank holder, in finite element simulation are considered rigid because they are extreme stiff compared to the sheet. For this reason the tool can be presented as a surface only.

6. Solution and Results

Following the successful modeling of geometry, then meshing and correctly applying boundary conditions and loads, a solution was run. The displacement load is applied and problem is executed in the nonlinear domain using material properties specified as in TABLE 2.The solution was a large static displacement analysis. It was carried out with time increments, having the max number of sub steps set as 1000, minimum as 10, and a desired number of 100 specified. The stress distribution was plotted on a contour plot to give a visual indication of stress through the now deformed blank.

The formations of sheet metal along with resulting stress are represented as shown in the following fig. 9 and fig. 10. From the color grid at the bottom of the fig. 9 and fig 10, the values of stress in N/mm^2 at the various points on the drawn component can be obtained. Fig. 9, contour plot for stress for model 1 and Fig. 10, contour plot for stress for model 2 ,shows the distribution of stress in the drawn part of the blank. From the finite element simulation, the region of maximum stress can be identified. Higher stress regions shown in fig 9 and 10.

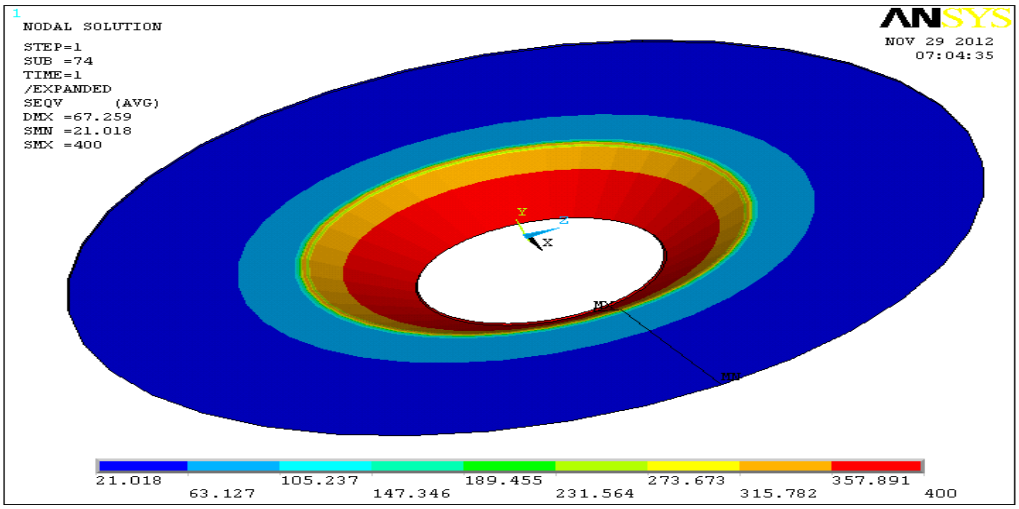


Fig. 9: Contour plot for stress (N/sq mm) for model 1

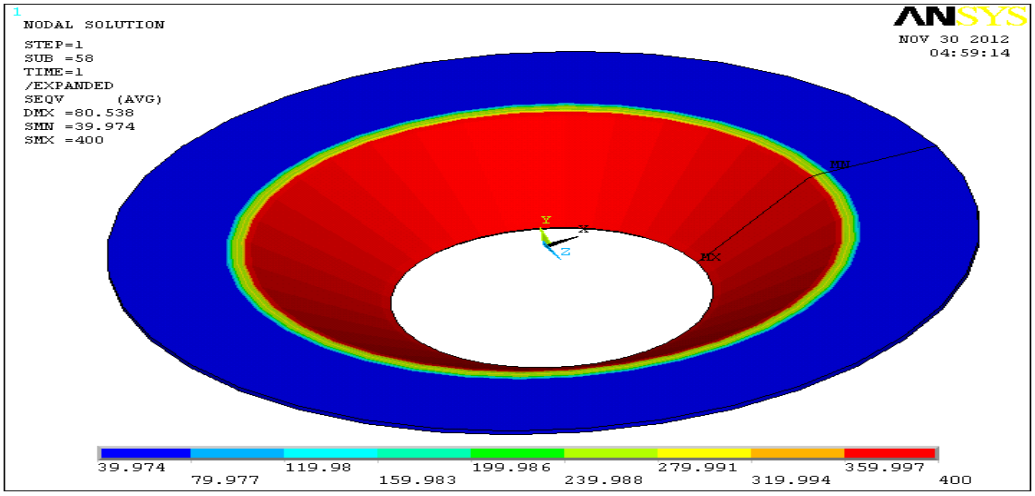


Fig. 10: Contour plot for stress (N/sq mm) for model 2

7. Conclusions

In this paper, a method to simulate a drawing process in ANSYS was explained. In this simulation work the finite element analysis is done on the conical back plate of the impellers Over the course of this work two models were built and tested, these were:

Model No. 1: For M.S. Plate having thickness 3 mm, a 2-D Axisymmetric solid model built using all PLANE 42 elements, 30 degree punch nose angle used.

Model No. 1: For M.S. Plate having thickness 5 mm, a 2-D Axisymmetric solid model built using all PLANE 42 elements, 43 degree punch nose angle used.

Following results can be deduced from this effort:

(1.) With simulation via FEM, designers can estimate field variables such as stress distribution. This information enhances the design capability and knowhow of an experienced process designer and leads to a reduced number of die-tryout tests. We conclude that referring on space state of stress, the location of affected zones were established, which be helpful in punch/ die design and design of other parameters such as punch force, blank holder force etc

(2) The distribution of stress of the drawn component presented.

(3) It is also concluded that ANSYS 12.0 is a very capable finite element software package, which can handle contact, plasticity and large deflection nonlinearities very accurately. Thus, forming processes can easily and accurately be modeled in ANSYS.

(4) It was concluded that, by using FEM, it is possible to produce successful conical shape products.

(5) By using a specialized software, one can be save time and other costs on research work. These simulation and analyses, presented here, suggests that, expensive way to find materials behavior by punch, die and experimental set up can be avoided by using specialized software.

However errors due to material properties may vary from coil to coil and affect the result. Finite element analysis determines an approximate solution of the problem modeled which have errors, these errors generally result from simplifications to the geometry, boundary conditions, material properties, loading and friction and as well as the discretization errors are a consequence of representing a continuum by a finite number of elements.

8. Suggestions For Future Work

Throughout the study it is observed that some of the areas can further be investigated and can be developed. In the presented study, the experiments were performed on the conical shape only. But the behavior of semi-circular, rectangular and triangular shape can be analyzed. In the current study the material is considered isotropic, a study can be developed by considering the anisotropy of the material. Material can be taken other than mild steel. A majority of products of automobile and aircraft industry are manufactured by drawing process. Products of these industries can be taken for analysis work. Investigations on nonsymmetric workpieces can be conducted. Different element models can developed. Remeshing can be applied.

References

[1] Saad Theyyab Faris, Study of the stress and strain distribution during deep drawing and ironing process of metals with a circular profiled die, Diyala Journal of Engineering Sciences, ISSN 1999-8716, Vol. 02, No. 01, June 2009, pp. 80-95

[2] M. Afteni, M. Banu, V. Paunoiu, I. Constantin, Numerical and experimental investigations of the nickel thin sheets micro-deep drawing process, The annals of "dunărea de jos" university of galaţi fascicle v, technologies in machine building,ISSN 1221- 4566, 2011, pp 149

[3] Dr.Sc. Amra Tali ikmi, Muamer Trako, Mladen Karivan, Finite element analysis of deep drawing, 14th International Research/Expert Conference, Trends in the Development of Machinery and Associated Technology, TMT 2010A, September 2010, pp-11-18

[4] Eric T. Harpella,1, Michael J. Worswickb, Mark Finnc, Mukesh Jainc, Pierre Martind,Numerical prediction of the limiting draw ratio for aluminum alloy sheet, ELSEVIER, Journal of Materials Processing Technology 100 (2000), 2000,pp 131-141

[5] L.F. Menezes and C. Teodosiu, Three-dimensional numerical simulation of the deep-drawing process using solid fnite elements, ELSEVIER, Journal of material processing technology 97(2000), 2000,pp 100-106

[6] Abdolhamid Gorji & Hasan Alavi-Hashemi & Mohammad Bakhshi-jooybari & Salman Nourouzi & Seyed Jamal Hosseinipour, Investigation of hydrodynamic deep drawing for conical–cylindrical cups, Int J Adv Manuf Technol (2011) 56:915–927,DOI 10.1007/s00170-011-3263-0, Springer-Verlag London Limited ,2011,pp 915–927

[7] Laxmiputra M Nimbalkar, Sunil Mangshetty, Analyzing the Effect of Blank Holder Shape in Deep Drawing Process Using Fem, International Journal of Engineering Research and Development e-ISSN: 2278-067X, p-ISSN: 2278-800X,Volume 4, Issue 1, October 2012, pp. 23-28.

[8] I.Burchitz, Springback: improvement of its predictability, Project, NIMR project number MC1.02121, Strategic Research programme of the Netherlands Institute for Metals Research, March 2005, pp 1-13

[9] Serhat Yalçin, Analysis and modeling of plastic wrinkling in deep drawing, Thesis, Middle east technical university, September 2010,pp 31-34

[10] Hakim S. Sultan Aljibori, 2009, Finite Element Analysis of Sheet Metal Forming Process, European Journal of Scientific Research, Euro Journals Publishing, Inc. 2009,ISSN 1450-216X Vol.33 No.1, (2009), pp.57-69

[11], Dr. R. Uday Kumar, Finite element anlysis of evaluation of radial stresses in hydro -assisted deep drawing process, International journal of pure and applied research in engineering and technology, IJPRET, Volume 1(2):, ISSN: 2319-507X, 2012, pp 1-10

[12] GAO En-zhi, LI Hong-wei, KOU Hong-chao,CHANG Hui, LI Jin-shan, ZHOU Lian, 2009, Science direct, Transaction of non ferrous metals society of china, 19 (2009),pp 433-437

[13] IS 2062:1999, Indian standard specification code for Mild Steel, Indian standard,Steel for general structural purposes – specification (fifth revision),second reprint January 2002,pp 2-5

The Hardware Implementation of Devices Forming Discrete Signals with Multi-Level Correlation Function

Alexey Smirnov
Professor in the Department of Software, Kirovohrad National Technical University, Kirovohrad, Ukraine

Abstract:
Algebraic approach to study the formation of large assemblies of discrete signals with multi-level correlation function, which is based on the section of circular orbits of group codes, is studied. The number and value of side-lobe levels of the correlation function of the generated sequences, and the power of the assembly signals are determined by remote and structural properties of polynomial rings over finite fields. Proposals for the hardware implementation of devices forming discrete signals with multi-function correlation are developed.

Keywords: ensembles of digital signals, multi-level correlation function

1. Introduction

A promising direction in the development of algebraic methods of the theory of discrete signals is the use of advanced mathematical apparatus of the theory of finite fields and, in particular, the theory of polynomial rings, which allows associating the correlation properties of the sequences, generated from the group and the structural properties of the code sequences [1 – 4]. The studies carried out in this work showed that thrived algebraic approach to the synthesis of discrete signals based on section of circular orbits of the group code allows creating large assemblies of sequences, the correlation properties of which have multi-level structure. The synthesized signals have the most practical interest in multiple access radio control systems [5 – 7]. The use of large assemblies of discrete signals with the improved properties will significantly increase subscriber capacity of radio control systems with code channels division.

Proposals for hardware implementation of devices forming discrete signals with multi-level correlation function are developed in this work. It is shown that the developed solutions allow to generate a sequence with improved correlation and assembly properties and to practically implement the developed in [1 – 4] method of forming of digital signals.

2. Algebraic approach to the formation of discrete sequences with multilevel correlation function

The proposed in [1 – 4] algebraic approach to the formation of large assemblies of discrete signals with multi-level correlation function is based on the section of circular orbits of group codes. The number and value of side-lobe levels of the correlation function of the generated sequences, and the power of signals assemblies are determined by remote and structural properties of polynomial rings over finite fields. Let's briefly examine these provisions constituting the theoretical basis for the formation of discrete signals.

Group code is uniquely determined by leaders (representatives) of its component cyclic orbits. An orbit hereafter refers to the set of code words equivalent to each other with respect to the operation of the cyclic shift. Under the section of the orbits of the group code let's understand the choice of one representative (leader) of each orbit. Distance (correlation) properties of the thus formed set of leaders are determined by remote properties of group codes; herewith the equivalence of cyclic shift operation is absent by definition of orbits section. Let's set this property in the basis of the assembly of discrete signals formation. Sectional diagram of nonzero cyclic orbits of group code is shown in Fig. 1.

Fig. 1 shows the decomposition of the vector space $GF^n(q)$ on the sets of non-intersecting orbit V_ξ, $\xi = 0,\ldots,L$, the group code V representation through the union of a finite number of orbits and the scheme of choosing of orbital leaders – one arbitrary representative from each cyclic subsets V_ξ, $\xi = 0,\ldots,M$ (for convenience the code words $C_{v,u} = \left(c_0^{v,u}, c_1^{v,u}, \ldots c_{n-1}^{v,u}\right)$ are marked by two indices: v – the number of orbit V_v of the code V, $v = 1,\ldots,M$; u – the number of a code word in the orbit $u = 1,\ldots,z_v$, where z_v – the number of code words in the orbit V_v, $z_v \le n-1$).

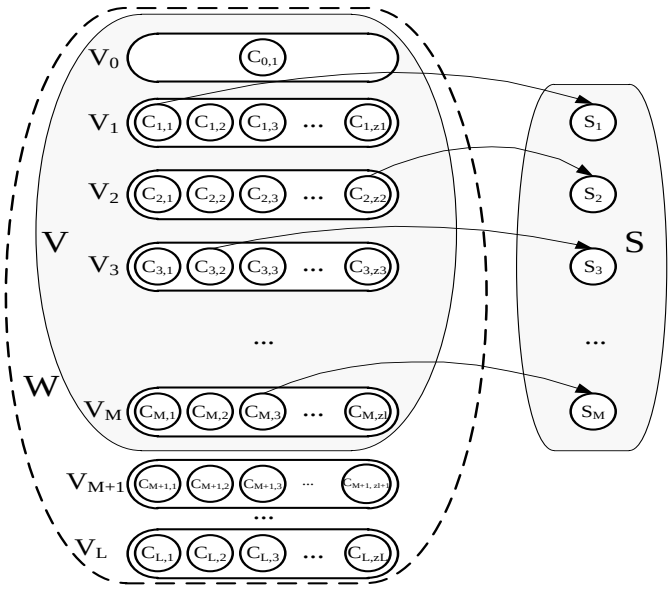


Fig.1. Scheme of nonzero cyclic orbits' section of a group code for the formation of an assembly of discrete signals

Representatives of the orbits of the selected form the set $S=(S_1,S_2,...,S_M)$, where $S_v=C_{v,u}$, $v=1,...,M$, and the selection of an index u with appropriate $C_{v,u}$ is determined by the rule of section of the v-th cyclic group code orbit.

Let's consider the binary case, i.e. restrict ourselves to the properties of the set $S=(S_1,S_2,...,S_M)$, formed by the section of the circular orbits of binary group code. The elements of the formed of discrete sequences (digital signals) $S_v=(s_0^v,s_1^v,..s_{n-1}^v)$ let's define the elements of the selected code words (the leaders of the orbits) as follows: $s_i^v=\begin{cases}1, c_i^{v,u}=1;\\ -1, c_i^{v,u}=0.\end{cases}$

Let's suppose that the considered (n, k, d) code V has a weight spectrum of:

$$\begin{cases} A(0)=1;\\ A(1)=0;\\ A(2)=0;\\ \quad ...\\ A(d-1)=0;\\ A(d);\\ A(d+1);\\ \quad ...\\ A(n). \end{cases} \qquad (1)$$

$w=0,...,n$, where $A(w)$ – the number of code words in the code V with the weight w.

Then the set of digital signals $S=(S_1,S_2,...,S_M)$ formed by the section of cyclic orbits of code V, has correlation and assembly properties, corresponding to the following statement [1 – 4].

Statement

1. Side lobes of the periodic function of auto – (PFAC) and mutual – (PFMC) correlation signals' assembly $S=(S_1,S_2,...,S_M)$ have the following values:

$$PFMC,PFAC=\frac{n-2w}{n}, \qquad (2)$$

for those $w=d,d+1,...,n,$ that $A(w)\neq 0$.

2. For all such $w=d,d+1,...,n$, that $A(w)=0$ the side lobes and PFAC and PFMC will never be $\frac{n-2w}{n}$.

3. The power M of the assembly $S=(S_1,S_2,...,S_M)$ is defined by the number of non-zero orbits of the code V and is bounded below by the expression:

$$M \geq \frac{2^k-1}{n}. \quad (3)$$

The equality holds in case of the maximum period of the sequence of all the orbits forming the code, i.e. if the code is V a set of orbits, formed by sequences of maximum length (m-sequences).

Let's consider the most general case where the binary group (n,k,d) code under $GF(2)$ is given by checking polynomial of:

$$h(x)=f_{i1}(x)f_{i_2}(x)..f_{i_u}(x)=\prod_{s=1}^{m-1}\left(x-\alpha^{i_1(2^s)}\right)\left(x-\alpha^{i_2(2^s)}\right)..\left(x-\alpha^{i_u(2^s)}\right), \quad (4)$$

where, $f_{i_1}(x)$, $f_{i_2}(x)$, ..., $f_{i_u}(x)$ – u arbitrary row of the following minimal polynomial elements $\alpha^{i_1}\in GF(2^m)$, $\alpha^{i_2}\in GF(2^m)$, ..., $\alpha^{i_u}\in GF(2^m)$ respectively, where the order of the elements α^{i_1}, α^{i_2}, ..., α^{i_u} is equal to the order of the multiplicative group of a finite field $GF(2^m)$, $n=2^m-1$, α – a primitive element of the finite field $GF(2^m)$, $n=2^m-1$.

Let's consider, without loss of generality that $i_1=1$. Let's define the check and generator polynomial as follows:

$$h(x)=\prod_{s=0}^{m-1}\left(x-\alpha^{(2^s)}\right)\left(x-\alpha^{i_2(2^s)}\right)..\left(x-\alpha^{i_u(2^s)}\right),$$

$$g(x)=\frac{x^n-1}{h(x)}=\prod_{j\neq 1,i_2,...,i_u}\ \prod_{s=0}^{m_j}\left(x-\alpha^{j(2^s)}\right).$$

Schematically, the process of forming of the check and generator polynomial is shown in Fig. 2. The symbol v stands for the number of the classes of conjugate elements that make up a multiplicative group of a finite field $GF(2^m)$. The first class (elements α^1, α^2, ..., $\alpha^{2^{m_1}}=\alpha^{2^{m-1}}$) contains m conjugacy (which determines the primitive element α). The following classes (elements α^j, α^{2j}, ..., $\alpha^{j2^{m_2}}$) contain m_j conjugacy (m_j divides evenly m) $j\in[1..v]$. For each $j\in[1..v]$ corresponding m_j is defined as the smallest positive integer for which the equality:

$$j=\left(j2^{m_j}\right)\bmod\left(2^m-1\right).$$

If the order of the multiplicative group of a prime number, that is, when:

$$2^m-1=\text{prime number},$$

then:

$$\forall j: m_j=m.$$

A single element of the field $\alpha^0=1$ forms an additional conjugate class of one element.

Fig. 3 shows the corresponding distribution of the elements of a finite field in the polynomials $h(x)$ and $g(x)$. Elements of a finite field of the first u conjugate classes are the roots of the check polynomial $h(x)$. A range of elements of a finite field, which holds the roots of the check polynomial $h(x)$, is determined by the largest value z, for which the condition $\alpha^z=\alpha^{(z)\bmod(2^m-1)}$, is done, that is:

$$z=\max_{s=0,...,m-1}\left\{\left(2^s\right)\bmod\left(2^m-1\right),\left(i_2 2^s\right)\bmod\left(2^m-1\right),...,\left(i_u 2^s\right)\bmod\left(2^m-1\right)\right\}.$$

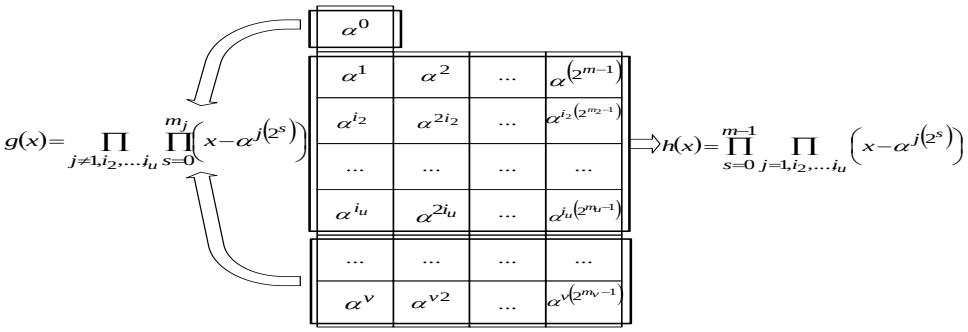


Fig. 2. Scheme of formation of check and generator polynomials of the group code

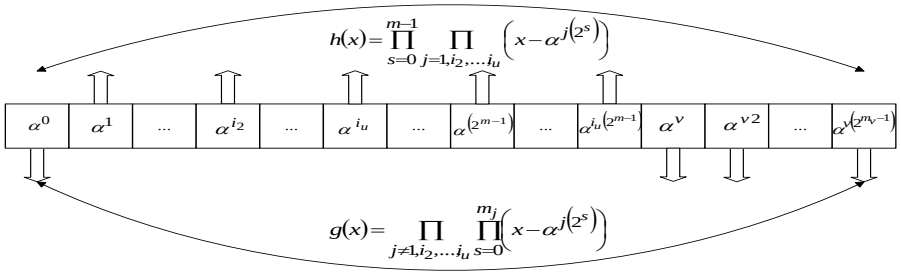


Fig. 3. Distribution of the elements of the finite field to check and generator polynomials of the group code

In general, the roots of polynomials $f_{i_1}(x)$, $f_{i_2}(x)$, ..., $f_{i_u}(x)$ are in the range:

$$\underbrace{\alpha^{i_1},\dots,\alpha^{i_1(2^{m-1})},\dots,\alpha^{i_2},\dots,\alpha^{i_2(2^{m-1})},\dots,\alpha^{i_u},\dots,\alpha^{i_u(2^{m-1})}}_{z\text{ values}},$$

whence:

$$2t=2^m-z-1,$$

and, accordingly:

$$d=2t+1=2^m-z.$$

The corresponding code parameters of a group code are as follows:

$$\left(n=2^m-1, k=zm, d=2^m-z\right). \quad (5)$$

Let's estimate the weight spectrum of the code. Verification polynomial of the code with the parameters (5) contains check polynomials of all the codes, as a cofactor with the check polynomials $h(x)=f_{i1}(x)f_{i_2}(x)..f_{i_y}(x), y\le u$.

It follows that all the code words from the group codes with $h(x)=f_{i1}(x)f_{i_2}(x)..f_{i_y}(x), y\le u$ are code words of considered code with parameters (5), i.e. nonzero components of the weight spectrum are formed by a successive addition (in order of addition of the cofactors in the polynomial $h(x)=f_{i1}(x)f_{i_2}(x)..f_{i_y}(x), y\le u$) of the corresponding pair of elements (for all $y=2,3,\dots,u$):

$$A(z_y)\neq 0,$$

$$A(2^m-z_y)\neq 0,$$

where:

$$z_y=\max_{s=0,\dots,m-1}\left\{\left(2^s\right)\bmod\left(2^m-1\right),\left(i_2 2^s\right)\bmod\left(2^m-1\right),\dots,\left(i_y 2^s\right)\bmod\left(2^m-1\right)\right\}.$$

When $y=1$ we have one non-zero component of the weight spectrum $A(2^{m-1})\neq 0$, which corresponds to $z_y=2^{m-1}$.

The considered in works [1 – 4] cases of building three – and five-level discrete signals correspond to:

$$y = 2:$$

$$z_y = 2^{m-1} + 2^{\frac{m+1}{2}-1},$$

$$A(2^{m-1} + 2^{\frac{m+1}{2}-1}) \neq 0,$$

$$A(2^{m-1} - 2^{\frac{m+1}{2}-1}) \neq 0$$

$$y = 3:$$

$$z_y = 2^{m-1} + 2^{\frac{m+1}{2}},$$

$$A(2^{m-1} + 2^{\frac{m+1}{2}}) \neq 0,$$

$$A(2^{m-1} - 2^{\frac{m+1}{2}}) \neq 0.$$

Thus, the three – and five-level digital signals are a special case of constructing of large assemblies of discrete signals with multi-level correlation functions.

The general expression for estimating the weight range of the group code specified by a check polynomial (4) let's put down as:

$$A(w) = \begin{cases} 1, w = 0; \\ 0, w = 1,\ldots,z_u - 1; \\ \neq 0, w = z_u; \\ \ldots \\ \neq 0, w = z_3 = 2^{m-1} - 2^{\frac{m+1}{2}}; \\ 0, w = z_3 + 1,\ldots,z_2 - 1; \\ \neq 0, w = z_2 = 2^{m-1} - 2^{\frac{m+1}{2}-1}; \\ 0, w = z_2 + 1,\ldots,z_1 - 1; \\ \neq 0, w = z_1 = 2^{m-1}; \\ 0, w = z_1 + 1,\ldots 2^m - z_2 - 1; \\ \neq 0, w = 2^m - z_2 = 2^{m-1} + 2^{\frac{m+1}{2}-1}; \\ 0, w = 2^m - z_2 + 1,\ldots 2^m - z_3 - 1; \\ \neq 0, w = 2^m - z_3 = 2^{m-1} + 2^{\frac{m+1}{2}}; \\ \ldots \\ \neq 0, w = 2^m - z_u; \\ 0, w = w = 2^m - z_u + 1,\ldots 2^m - 1. \end{cases}$$

The corresponding expression on valuing the side-lobe periodic correlation function generally takes the form:

$$\mathrm{PFMC,PFAC}=\begin{cases}\dfrac{2^m-2z_u-1}{2^m-1},\ w=z_u= \\ =\max\limits_{s=0,\dots,m-1}\left\{\left(2^s\right)\bmod\left(2^m-1\right),\left(i_2 2^s\right)\bmod\left(2^m-1\right)\dots,\left(i_u 2^s\right)\bmod\left(2^m-1\right)\right\}; \\ \dots \\ \dfrac{2^m-2z_3-1}{2^m-1}=\dfrac{-1-2^{\frac{m+1}{2}+1}}{2^m-1},\ w=z_3=2^{m-1}-2^{\frac{m+1}{2}}; \\ \dfrac{2^m-2z_2-1}{2^m-1}=\dfrac{-1-2^{\frac{m+1}{2}}}{2^m-1},\ w=z_2=2^{m-1}-2^{\frac{m+1}{2}-1}; \\ \dfrac{2^m-2z_1-1}{2^m-1}=\dfrac{-1}{2^m-1},\ w=z_1=2^{m-1}; \\ \dfrac{2^m-2(2^m-z_2)-1}{2^m-1}=\dfrac{-1+2^{\frac{m+1}{2}}}{2^m-1},\ w=2^m-z_2=2^{m-1}+2^{\frac{m+1}{2}-1}; \\ \dfrac{2^m-2(2^m-z_3)-1}{2^m-1}=\dfrac{-1+2^{\frac{m+1}{2}+1}}{2^m-1},\ w=2^m-z_3=2^{m-1}+2^{\frac{m+1}{2}}; \\ \dots \\ \dfrac{2^m-2(2^m-z_u)-1}{2^m-1},\ w=2^m-z_u= \\ =2^m-\max\limits_{s=0,\dots,m-1}\left\{\left(2^s\right)\bmod\left(2^m-1\right),\left(i_2 2^s\right)\bmod\left(2^m-1\right)\dots,\left(i_u 2^s\right)\bmod\left(2^m-1\right)\right\}.\end{cases} \quad (6)$$

Thus, the generated by the proposed method discrete signals have multilevel features of auto – and mutual-correlation. The values of the lateral emissions take a finite number of values given by the weight properties of a used group code.

Let's estimate the power of the assembly of the formed digital signals. The power of the used code is $2^k=2^{um}$, there are altogether:

$$2^k-1=2^{um}-1$$

of nonzero code words.

If we assume that each code word has the maximum period and each cyclic orbit contains exactly 2^m-1 code words, then the expression for the estimates of the power of the assembly of the formed signals becomes:

$$M=\frac{2^{um}-1}{2^m-1}=2^{(u-1)m}+2^{(u-2)m}+\dots+2^m+1.$$

The analysis of the last expression shows that the use of group codes allows creating large assemblies of discrete signals. Adding of a minimal polynomial as another cofactor in the check polynomial increases the power of the assembly $2^{(u-i)m}$, where $u-i$ – the number of the added minimum polynomials.

3. Development of proposals for the hardware implementation of devices forming discrete signals by a proposed method

The developed method of forming of digital signals allows building large assemblies of weakly correlated binary sequences. Let's consider the possibility of practical formation of large assemblies of weakly correlated discrete signals and constructing the corresponding hardware devices for binary sequences generating.

In case of multi-level sequences, we get the following scheme of formation of discrete signals. (Fig. 4). The device is built through connecting of shift registers to the adder output u. A wiring diagram to make appropriate provision in the register ring feedback shift is selected by the coefficients of primitive polynomials $h_1(x)$, $h_2(x)$, ..., $h_u(x)$ of m degree, respectively. In this case, the length of the binary sequences equals to $n=2^m-1$ and to form them one need to use u shift registers the shift registers with m binary digits. The initial state of the shift registers sets the mode of the formed sequence.

Functions of the feedback shift registers are set by coefficients of primitive polynomials of m degree:

$$h_1(x) = h_{1,0} + h_{1,1}x + h_{1,2}x^2 + \ldots + h_{1,m}x^m = f_{i_1}(x) = \prod_{s=0}^{m-1}\left(x - \alpha^{i_1(2^s)}\right)$$

$$h_2(x) = h_{2,0} + h_{2,1}x + h_{2,2}x^2 + \ldots + h_{2,m}x^m = f_{i_2}(x) = \prod_{s=0}^{m-1}\left(x - \alpha^{i_2(2^s)}\right),$$

$$\ldots$$

$$h_u(x) = h_{u,0} + h_{u,1}x + h_{u,2}x^2 + \ldots + h_{u,m}x^m = f_{i_u}(x) = \prod_{s=0}^{m-1}\left(x - \alpha^{i_u(2^s)}\right)$$

where $f_{i_1}(x),\ f_{i_2}(x), \ldots,\ f_{i_u}(x)$ – minimal polynomial of elements α^{i_1}, α^{i_2}, ..., α^{i_u}, respectively, from the end field $GF(2^m)$, which are defined by their roots $\alpha^{i_1(2^s)}$, $\alpha^{i_2(2^s)}$, ..., $\alpha^{i_u(2^s)}$, $s = 0, 1, \ldots, m-1$.

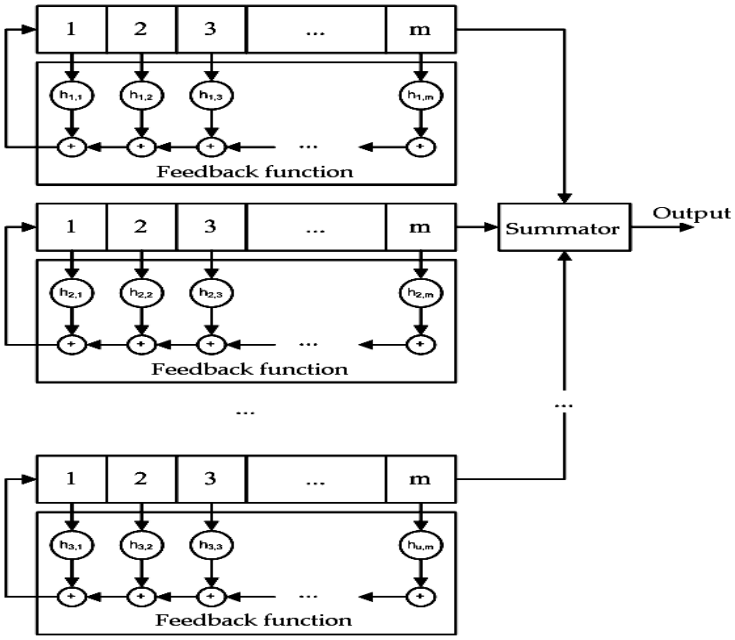


Fig. 4. Block diagram of the formation of discrete signals with multi-level correlation function

The order of elements α^{i_1}, $\alpha^{i_2}, \ldots, \alpha^{i_u}$ equals the order of the multiplicative group of a finite field $GF(2^m)$, α – a primitive element of the finite field $GF(2^m)$.

The device works in the discussed above manner and allows creating:

$$M = \frac{2^{um} - 1}{2^m - 1} = 2^{(u-1)m} + 2^{(u-2)m} + \ldots + 2^m + 1$$

sequences of length $n = 2^m - 1$.

4. Conclusions

Thus, in the course of a conducted research practical suggestions regarding the implementation of the hardware devices forming discrete sequences have been developed.

The designed schemes are implemented computationally by efficient converters, for example, based on circuits with a shift register and an adder (see Fig. 4 – 7). They allow creating large assemblies of discrete signals with improved correlation and assembly properties. Therefore, the developed proposals let to practically implement the developed method of forming discrete signals.

References

[1] Kuznetsov A.A., Smirnov A.A., Sai V. N., Digital Signals with Multi-Level Correlation Function // Radio: Ukr. Interag. Sc. and Eng. Sat – Kharkov: KhNUR. 2011. – Issue 166. ¬ P. 142-152.

[2] Kuznetsov A.A., Smirnov A.A., Sai V. N., Formation of Discrete Signals With Multi-Function Correlation // Information processing systems. – Kh.: KhAFU. – 2011 – Vol. 5(95). – P. 50-60.

[3] Kuznetsov A.A., Use of Complex Discrete Signals for Steganographic Information Security / A.A. Kuznetsov, A.A. Smirnov // International Journal of Engineering Practical Education. – Volume 1, Issue 1. – USA, Indiana: Science and Engineering Publishing Company. – 2012. – P. 21-25.

[4] A.A. Smirnov, Comparative Studies of Methods for the Synthesis of Digital Signals with Special Correlation Properties / A.A. Smirnov, E.V. Meleshko // Abstracts of V International Scientific and Technical Symposium "New Technologies in Telecommunications" (NIICT-Carpathian-2012), Kyiv. 17-21 January 2012 – Kyiv: NIICT. – 2012. – P. 80-81.

[5] Gryanik M.V., Frolov V.I., Technology CDMA – The Future of Mobile Systems in Ukraine. – The World of Communication, 1998, # 3. – P. 40-43.

[6] Naumenko N.I., Stasev J.V., Kuznetsov A.A, Evseev S.P., Theory Of Signal-Code Structures. Kh.: KhAFU, 2008 – 489.

[7] Sklar B., Digital Communication. Theoretical Basis and Practical Application. – M.: Williams, 2003. – 1104p.

Author Name

Alexey Smirnov was born in 1977. Graduated from Kharkiv Military University with a degree in "Automated Control Systems" in 1999.

Candidate of Technical Sciences (PhD). Professor of Department of Software of Kirovohrad National Technical University, Ukraine.

Field of interest: information security and routing issues.

Avoidance of Bottleneck in PCS Network

Sidhi Pandey[1], Alka[2], Pratima Singh[3]

[1, 2, 3] (Computer Science & Engineering, Institute of Technology & Management, India)

Abstract:

This paper contains the concept of Personal Communication Service (PCS) network and schemes to avoid bottleneck in PCS network. One of the key issues in mobile communication is to find the current location of mobile terminal (MT) to deliver the services, which is called as location management. In this paper we will discuss various schemes for improving location management in PCS network and focus on surveying the location management mechanisms in PCS system.

Keywords: De-registration, hand-off, HLR/VLR, Location Management, Mobile terminal, MSC, PCS.

1. Introduction:

Cellular communication has been experiencing a rapid growth in recent years. From its introduction in the early 1980s, cellular communication has been evolved from a costly service with limited availability toward an affordable alternative to wired telephone service. This wide acceptance of cellular communication has led to the development of a new generation of mobile communication network called personal communication services, which can support a large mobile subscriber population while providing various types of services unavailable to traditional cellular system [1].Personal Communication Services networks provide wireless communication services that enable mobile terminals to transfer any form of information between any locations at any time. To support user mobility, the Personal Communication Services networks have to store and maintain location information of mobile stations so that an incoming call can be delivered to the target Mobile Station. The operations on location information consist of location updates and location queries. An update occurs when a Mobile Station changes location. A query occurs when a Mobile Station needs to be located, e.g., to deliver an incoming call to this Mobile Station. The widespread deployment of Personal Communication Services will lead to a tremendous increase in the number of updates and queries to the location database. Thus, a key challenge to location management is to develop efficient database architecture so that the location data can be readily available for signaling such as call setup and routing [1].

2. Location management:

In cellular systems a mobile unit is free to move around within the entire area of coverage. Its movement is random and therefore its geographic allocation is unpredictable. This situation makes it necessary to locate the mobile unit and record its location to Home Location Register and Visitor Location Register when a call has to be delivered to it. Thus, the entire process of the mobility management component of the cellular system is responsible for two tasks:

2.1 Location Management:

It is identification of the current geographical location or current point of attachment of a mobile unit which is required by the Mobile Switching Center to route the call.

2.2 Hand-off:

It is transferring (handing off) the current (active) communication session to the next base station, which seamlessly resumes the session using its own set of channels.

The entire process of location management is a kind of directory management problem where current locations of MU are maintained continuously. Location management involves tracking of Mobile Terminal's location, moving from place to place so as to provide those services timely. Two basic operations in mobility tracking are: location update and paging. Basically, whenever Mobile Terminal moves out of its current LA, its geographical location information is updated to the nearest Base Station. On a call arrival, the network searches the called Mobile Terminal by sending polling signals to the vicinity of last reported location of Mobile Terminal. This searching process is called paging. The total Location Management cost is generally calculated by summing up the cost of location update and paging. Normally, the Location Update costs higher than paging. The network can require more frequent Location Updates, in order to reduce paging cost. Inversely, the network may require rare Location Updates, storing less information about user mobility to reduce computational overhead, but at a higher paging cost. To reduce the total location management cost, it is essential to provide good trade-off among paging and Location Update operations One of the main objectives of efficient location management schemes is to minimize the communication overhead due to updating of Home Location Register. The other related issue is the distribution of Home Location Register to shorten the access path, which is similar to data distribution problem in

distributed database systems. Motivated by these issues, recently a number of innovative location management schemes have appeared in the research world.

The location management approaches can broadly be classified as: Centralized approaches and Distributed approaches. Centralized approach keeps information only on one node in the mobile network. For example, existing location management standards, IS-41 and Global System for Mobile Communication are centralized approaches. Location lookup and update operations are simple in this case but they suffer from severe problems like congestion, central point failure etc. In Distributed approaches user information is distributed among many nodes in the network. This has better stability in comparison to centralized approach but location lookup and update operations are somewhat complex in this case. However, many schemes have come forward to solve this problem [7].

3. Conventional HLR/VLR scheme:

In this section we will describe the existing standard of location management along with the approaches to overcome its drawbacks. The two popular standards used are Global System for Mobile Communication and IS-41. They make use of two types of registers namely, Home Location Register and Visitor Location Register. These two registers are used to store the location information of the mobile terminals [3]. Figure 1 shows the basic architecture under this two-level hierarchy.

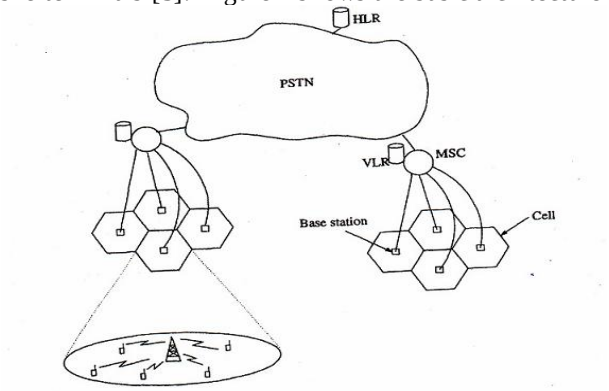


Figure 1: Standard Personal Communication Network architecture

The Home Location Register stores the user profiles of its assigned subscribers. These user profiles contain information such as the type of services subscribed, the quality-of-service requirements and the current location of the mobile terminals. Each Visitor Location Register stores replications of the user profiles of the subscribers currently residing in its associated LA. In order to effectively locate a mobile terminal when a call arrives, each mobile terminal is required to report its location whenever it enters a new LA. This reporting process is called location update. On receiving a location update message, the Mobile Switching Center updates its associated Visitor Location Register and transmits the new location information to the Home Location Register. We call this register update process as location registration. The Home Location Register will acknowledge the Mobile Switching Center for the successful registration and it will also deregister the mobile terminal at the Visitor Location Register of old LA. In order to locate a mobile terminal for call delivery, the Home Location Register is queried to determine the serving MSC of the target mobile terminal. The Home Location Register then sends a message to this Mobile Switching Center which, in turn, will determine the serving base station of the mobile terminal by paging all cells within its associated LA. This location tracking scheme requires the exchange of signaling messages between the Home Location Register and the new and old Mobile Switching Center's whenever the mobile terminal crosses an LA boundary. This may result in significant traffic load to the network especially when the current location of the mobile terminal is far away from its Home Location Register and the mobile terminal is making frequent movements among LA's. Besides, the Home Location Register may experience excessively high database access traffic as it is involved in every location registration and call delivery. This may result in increased connection set up delay during periods of high network utilization [3].

The major steps of the IS-41 location registration scheme are as follows (Fig 2)

Step 1: The mobile terminal moves into a new LA and sends a location update message to the nearby base station.

Step 2: The base station forwards this message to the new serving Mobile Switching Center.

Step 3: The new Mobile Switching Center updates its associated Visitor Location Register, indicating that the mobile terminal is now residing in its services area and sends a location registration message to the Home Location Register.

Step 4: The Home Location Register sends a registration acknowledgement message to the new Mobile Switching Center/ Visitor Location Register together with a copy of the subscriber's user profile.

Step 5: The Home Location Register sends a registration cancellation message to the old Mobile Switching Center/ Visitor Location Register.

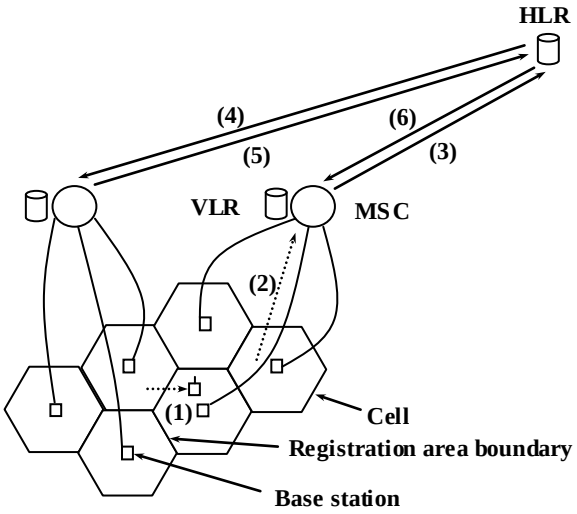


Figure 2: Location registration

Step 6: The old Mobile Switching Center removes the record for the mobile terminal at its associated Visitor Location Register and sends a cancellation acknowledgment message to the Home Location Register.
The IS-41 call delivery scheme is outlined as follows (Fig. 3)

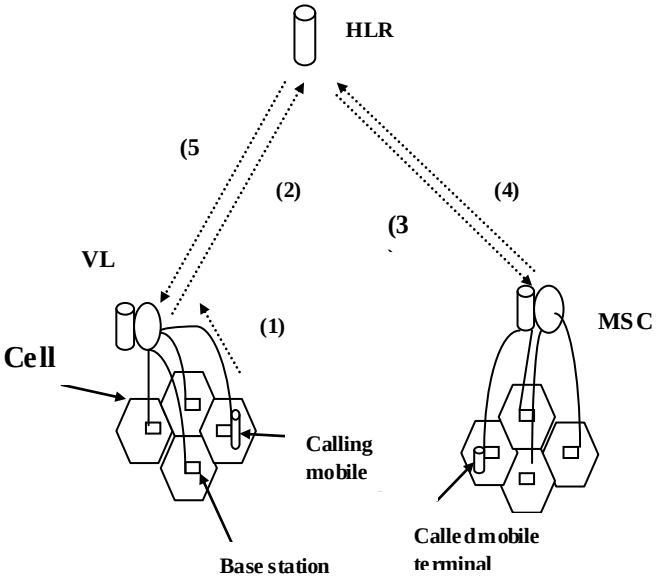


Figure 3: Call delivery

Step 1: The calling mobile terminal sends a call initiation signal to its serving Mobile Switching Center through the nearby base station.
Step 2: The Mobile Switching Center of the calling mobile terminal sends a location request message to the Home Location Register of the mobile terminal.
Step 3: The Home Location Register determines the current serving Mobile Switching Center of the called mobile terminal and sends a route request message to this Mobile Switching Center.
Step 4: The Mobile Switching Center determines the cell location of the called mobile terminal and assigns a temporary location directory number to the called mobile terminal. The Mobile Switching Center then sends this Temporary Location Directory Number to the Home Location Register.
Step 5: The Home Location Register sends the Temporary Location Directory Number to the Mobile Switching Center of the calling mobile terminal. The calling Mobile Switching Center can now set up a connection to the called Mobile Switching Center through the PSTN.

IS-41 and Global System for Mobile standards imply centralized approach, which has some disadvantages: Since every location request as well as location registration is serviced through an Home Location Register, it becomes overloaded. Due to the above reason, traffic on the links leading to the Home Location Register is heavy, which increases time required to establish connection to the mobile host. Any Home Location Register system failure causes all mobiles registered with Home Location Register to be unreachable even though mobile host may be roaming and away from Home Location Register region. Thus Home Location Register is a single point of failure in the network.

4. Modified HLR/VLR Scheme:

In conventional approach when a Mobile Terminal moves from one LA to another LA, which are served by different Visitor Location Register s, for registration of Mobile Terminal at new Visitor Location Register a signal message is transferred to Home Location Register from it, which sends a signal message to old Visitor Location Register to deregister the Mobile Terminal and upon getting an acknowledgement of De-registration from old Visitor Location Register, Home Location Register acknowledges to new Visitor Location Register for registration. This De-registration method is referred to as *explicit de-registration***.** The explicit De-registration scheme may produce significant signaling traffic in the network and require many accesses to the database involved. Due to the increasing number of mobile subscribers, the access rate to the Home Location Register and the Visitor Location Registers is expected to be very high and the databases could possibly become the bottle-neck of the future mobile systems.

Implicit De-registration and timeout/polling deregistration were proposed to reduce signaling traffic and database load due to deregistration. But the comparative study done by Z. Mao, it has been found that by using group deregistration strategy deregistration cost can be reduced significantly. Since this strategy reduces the cost of message transferring as well as database operation which will ultimately reduce total cost of location management. In modified HLR-VLR scheme we try to ignore explicit De-registration message to old Visitor Location Register and its acknowledgement to Home Location Register. Hence when new Visitor Location Register finds a new mobile unit it simply sends a message to Home Location Register which acknowledges the new Visitor Location Register to register it. So signal no. 5 & 6 of fig. 3 can be avoided.

5. Multi HLR Architecture Scheme In PCS Network:

In conventional HLR-VLR scheme, De-registration of a Mobile Terminal from a Visitor Location Register is always explicit. Explicit in the sense that stale entries of Visitor Location Register s are removed with the help of Home Location Register. Actually Home Location Register sends De-registration message to the Visitor Location Register to remove the stale entries when a Mobile Terminal changes its Visitor Location Register. This explicit De-registration increases the total cost by increasing the traffic load. To reduce the traffic load following De-registration strategies were proposed [2].

(A) Distance Based De-registration Scheme.
(B) Time-Based De-registration Scheme.
(C) Polling-Based De-registration Scheme.
(D) Group De-registration Scheme.
(E) Movement-Based De-registration Scheme.

Performance analysis of De-registration strategies in Personal Communication Network shows that the group de-registration scheme is best scheme among time and polling based de-registration schemes.

In the proposed architecture, we have several HLRs zone wise or circle wise instead of a single Home Location Register. It reduces the storage overhead of the Home Location Register. Each Home Location Register can serve more than one Visitor Location Register and each Visitor Location Register can serve more than one RAs. Simply we can say that this architecture contains several conventional HLR-VLR architectures. For each Mobile Terminal we define two types of HLRs: a resident-HLR and a serving-HLR. Resident-HLR is the Home Location Register where Mobile Terminal often resides. While on move, it can enter into the RA being served by another Home Location Register (serving-HLR). When Mobile Terminal will be served by the Home Location Register other than resident- Home Location Register, we will refer it as roaming. In the proposed architecture we define following types of move as:

5.1 Intra-VLR-Resident-HLR Move:

In this type of move, the Mobile Terminal changes its RA and the new RA is still being served by the same Visitor Location Register. The serving Visitor Location Register is being served by the resident-HLR. Now it is obvious that the location update is taking place only at Visitor Location Register not at resident-HLR.

Intra-VLR-Resident-HLR move is shown in fig (4). An Mobile Terminal residing in registration area RA1 moves to another registration area RA2. RA1& RA2 are being served by the same Visitor Location Register, VLR1. The VLR1 is being served by the resident-HLR. Due to movement of Mobile Terminal, this location update is changed at VLR1 not at resident-HLR

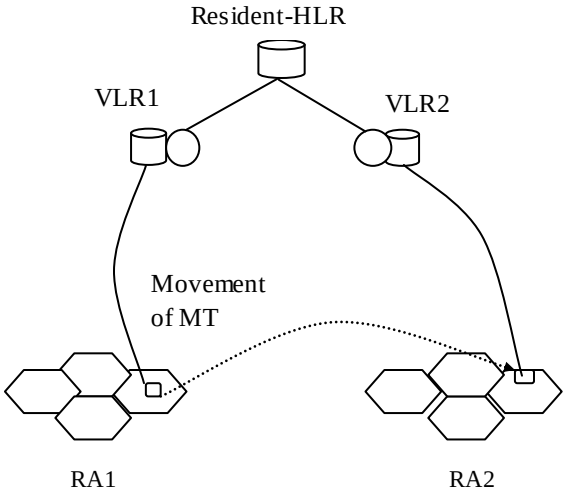


Figure 4: Intra-VLR-Resident-HLR move

5.2 Intra-VLR-Serving-HLR Move:

In this type of move, the Mobile Terminal changes its RA and the new RA is still being served by the same Visitor Location Register. The serving Visitor Location Register is being served by the serving-HLR. Again this information is only updated at Visitor Location Register not at serving-HLR and resident-HLR.

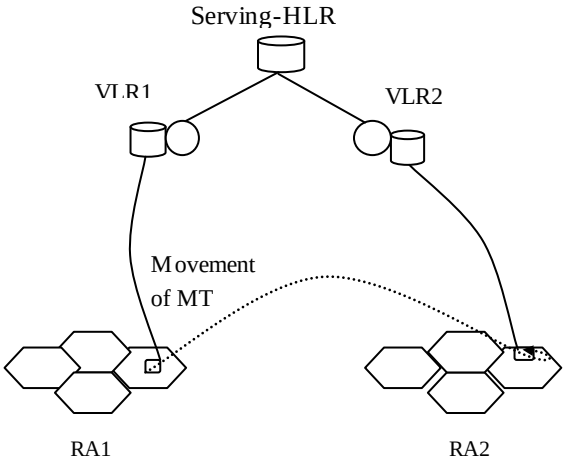


Figure 5: Intra-VLR-Serving-HLR move

Intra-VLR-Serving-HLR move is shown in fig (5). An MT residing in registration area RA1 moves to another registration area RA2. RA1& RA2 are being served by the same Visitor Location Register, VLR1. The VLR1 is being served by the serving-HLR. Due to movement of MT, this location update is changed at VLR1 not at serving-HLR and resident-HLR.

5.3 Inter-VLR-Resident-HLR Move:

In this type of move, the MT changes its RA and the new RA is being served by the new Visitor Location Register. The serving Visitor Location Register is being served by the resident-HLR. Now in this case registration of MT will take place at new Visitor Location Register, de-registration of Mobile Terminal will take place at old Visitor Location Register and finally resident-HLR will update this information in its database.

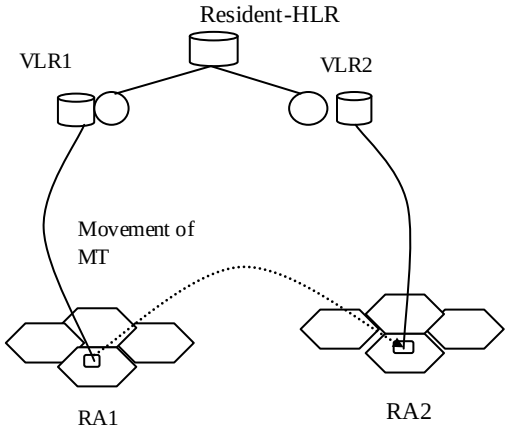


Figure 6: Inter-VLR-Resident-HLR move

Inter-VLR-Resident-HLR move is shown in fig (6). An MT is residing in registration area RA1 which is being served by the VLR, VLR1. This VLR1 is being served by the resident-HLR. On move MT changes its registration area and comes in RA2. This RA2 is being served by the VLR2.Now the registration of Mobile Terminal will take place at VLR2 and De-registration will take place at VLR1. This change in location update will take place on resident-HLR.

5.4 Inter-VLR-Serving-HLR Move:

In this type of move, the MT changes its RA and the new RA is being served by the new Visitor Location Register. The serving VLR is being served by the serving-HLR. Now in this case registration of Mobile Terminal will take place at new VLR, De-registration of Mobile Terminal will take place at old VLR and finally serving-HLR will update this information in its database. However resident-HLR will not be updated. Inter-VLR-Serving-HLR move is shown in fig (7). An MT is residing in registration area RA1 which is being served by the VLR, VLR1. This VLR1 is being served by the serving-HLR. On move Mobile Terminal changes its registration area and comes in RA2. This RA2 is being served by the VLR2.Now the registration of MT will take place at VLR2 and de-registration will take place at VLR1. This change in location update will take place on serving-HLR. Resident-HLR will remain unaffected.

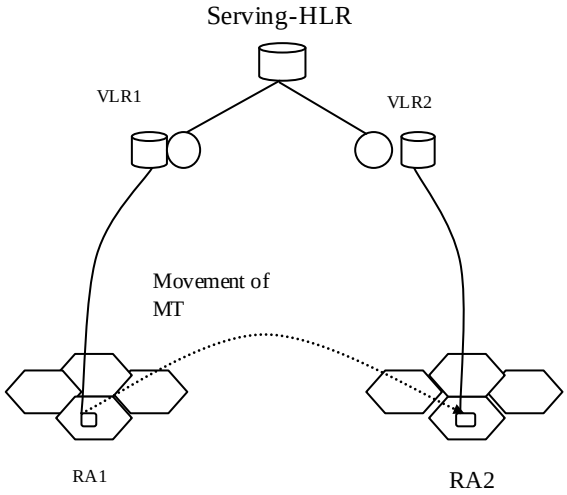


Figure 7: Inter-VLR-Serving-HLR move

5.5 Inter-VLR-Inter-HLR Move:

This type of move occurs into two cases.

1) This type of move occurs when an MT comes back to its resident-HLR from serving-HLR. In this move MT's registration occurs at serving VLR of resident-HLR and this information is sent back to the serving-HLR.

2) On reception of this message serving-HLR de-registers this MT from its database and informs to its VLR (where MT was registered previously) for de-registration.

This case is shown in fig (8). Mobile Terminal is residing in the registration area RA2 of Visitor Location Register2 of serving-HLR. On move Mobile Terminal comes under VLR3 of resident-Home Location Register. In this case both Visitor Location Registers and Home Location Registers change. The Mobile Terminal is now in its resident-HLR. Registration of Mobile Terminal will take place at Visitor Location Register3 and resident-HLR while it's De-registration

will take place at serving-Home Location Register and its corresponding Visitor Location Register where Mobile Terminal was previously resided.

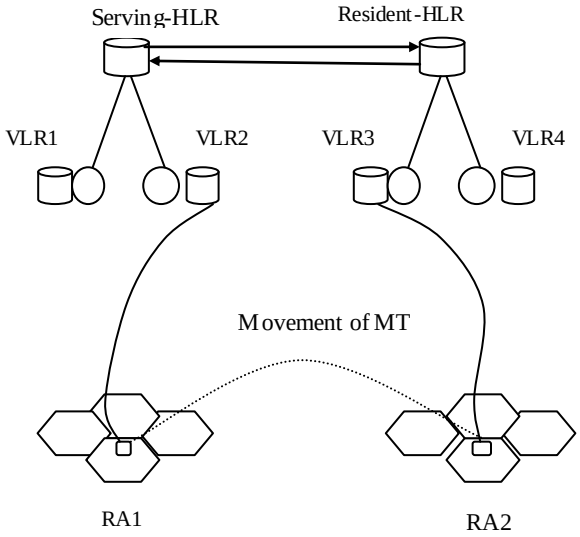


Figure 8: Inter-VLR-Inter-HLR move

3) When a Mobile Terminal moves to another serving-HLR. In this case registration of Mobile Terminal takes place at new serving-Home Location Register and its Visitor Location Register under which Mobile Terminal comes, de-registration occurs at the previous serving- Home Location Register and its associated Visitor Location Register from where Mobile Terminal is coming and resident-Home Location Register is updated.

6. Performance analysis of various location management schemes:

An analytical model to evaluate the performance of the conventional HLR/VLR architecture and a comparison of the same is made with the modified HLR/VLRs. In this analysis, a hierarchical tree of R layers is used, as shown in Fig. 9. The layer R contains the root node and the layer 1 contains the leaf nodes. A database is installed on each node of the tree and the Mobile Terminals are assigned to the leaf nodes. In the HLR/VLRs scheme, the network database, Home Location Register, is situated on the only node of layer R and the Visitor Location Registers are installed on the leaf nodes.

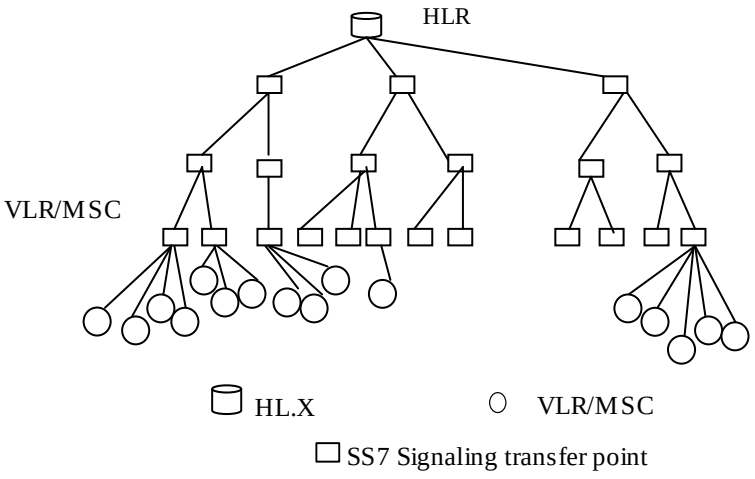


Figure 9: HLR/VLR Architecture

Since total cost consist of location update and location search so total cost is also reduced in modified HLR/VLR scheme. So it can be concluded that proposed modified version of HLR/VLR performs better than the conventional schemes for location management. Also it is observed that the proposed modified HLR/VLR scheme is better than the conventional one as the signaling cost and database updating cost is significantly reduced by using group deregistration strategy for location update only.

Conventional architecture has a single Home Location Register and that's why it suffers from call misrouting and bottleneck during peak load. To remove this, several conventional architectures are group together to form multi Home

Location Register architecture. Now in user profile replication scheme even in high load appropriate information is fetched from Home Location Registers and we significantly minimize the possibility of call misrouting.

7. Conclusion:

In this study, we defined the location management problem and its terminology, reviewed some of the main accomplishments achieved by researchers in the field, and established the fundamental issues that impact this problem. We also attempted to compare the solutions qualitatively according to their effectiveness for different types of updating techniques. That analysis led us to propose a new taxonomy for location management techniques in mobile and wireless environments based on several important factors. In closing, it is worth speculating on the long-term impact of location management issues on mobile and wireless environments and their design. The future for wireless and mobile computing is promising indeed, especially since technological advances continue to support more sophisticated applications for these environments.

References

[1] Guan-Chi Chen and Suh-Yin Lee, "Evaluation of Distributed and Replicated HLR for Location Management in PCS Network," Journal Of Information Science And Engineering 19, 85-101 (2003).
[2] Rajeev R. Kumar Tripathi, Sudhir Agrawal and Swati Tiwari, "Performance Analysis of De-registration Strategy in Personal Communication Network," IJCA, vol.24-no.1, June 2011.
[3] Rajeev R. Kumar Tripathi, Sudhir Agrawal and Swati Tiwari, "Modified HLR-VLR Location Management Scheme in PCS Network," IJCA, vol.6-no.5, Sept 2010.
[4] Sarvpal H. Singh, Sudhir Agrawal and Vijay Kumar "Performance Evaluation of Location Management Techniques in PCS Networks," IJCA, vol.15, no-8, Feb-2011.
[5] S. Mohan and R. Jain, "Two user location strategies for personal communication services," *IEEE Personal Communications*, Vol. 1, 1994.
[6] Haider Safa and Samuel Pierre, "A New Architecture for Improving Location management in PCS Network," Journal of Computer Science1 (2):249-258, 2005.
[7] Sarvpal H. Singh, Sudhir Agrawal and C. Ramakrishna "Location Management optimization schemes for Cellular System," IJCA, vol.4, no-4, july-2010
[8] Rajeev R. Kumar Tripathi, G S Chandel and Ravindra Gupta, "Multi HLR Architecture for improving location management in PCS Network," IJCA, vol.51-no.20, August 2012.

Exploring a Microcontroller Based Hearing Aid with an Output Level Indicator

[1]Aru Okereke Eze , Eng. Dr. Gozie Ihekweaba and [3]Ngwu Rosemary Chinyere
Department of Computer Engineering
Michael Okpara University of Agriculture, Umudike, Umuahia, Abia State, Nigeria

Abstract:
This research explores the design and implementation of a microcontroller based electronic Hearing aid with output level indicator. The incoming sound oscillates the base of a transistor which in turn amplifies the signal at the collector. The output is interfaced to an ADC which converts the signal to digital output. After the conversion, the microcontroller processes the signal and further feeds it in to the DAC for an analog conversion. This output is now fed into the audio level indicator, LM3915, to indicate the level of the audio signal at the ear phone.

Keywords: Microcontroller, Hearing Aid, Digital, Ear, Noise, Electronics.

1. Introduction

Digital hearing aids can do wonders for faded hearing. With the technological advancement in the society, hearing aids can significantly enhance the quality of life for most people with hearing impairment. Therefore, the electronic hearing aid is designed to make sounds louder and therefore easier to hear. Also the design of the circuitry keeps the sound from becoming too loud and helps reduce the effects of background noise.

The two basic types of technology for hearing aids are analog and digital. The first to exist, analog hearing aids process electrical sound in the analog domain; the more recent digital hearing aids process electrical sound in the digital domain. The earliest analog hearing aids simply amplified both speech and noise, and were ordered after testing to determine the particular frequency response needed by the patient. Newer analog hearing aids can be programmed during the fitting process, and some have multiple listening profiles that the patient can select with a button on the hearing aid.[1]. Manufacturers are moving toward their third or fourth generation of digital products. Digital technology is more stable over time. There are fewer components to go wrong and fewer components that are susceptible to moisture and aging changes. This means that the sound you experience on the first day you receive the hearing aid stays consistent until the program is changed.[2] In recent years, there has been an increasing trend toward fitting BTE(Behind The Ear) hearing aids, including receiver-in-canal (RIC) instruments. It is estimated that 51% of the hearing aids fitted in the U.S. are BTE instruments, rather than custom products [3]. In a survey by Kochkin of 2500 hearing aid users, patients reported a desire for hearing aids that do not feedback (85%), fit comfortably (79%), and are less visible (52%) [4]. In this work, we will explore a digital hearing aid with Output level indicator (LM3915) using the microcontroller Atmel 89c52, analog and digital converters.

We preset in section 2 system design, section 3 principle of operation, section 4 user guide, section 5 Conclusion.

2. System Design

The system is designed in such a way that the output is an analog signal which needs to be converted into digital signal for onward interface to a microcontroller. In this work, the ADC was handy to settle the conversion problem. It converts the analog signal to digital signal. After the conversion processes, the microcontroller gives out a digital output which requires conversion back to its analog state. Hence, the output signal can be connected to a microphone to ascertain the nature of the signal.

This analog output is then fed into lm3915, an audio level indicator, to process and give a required graphic result of the signal strength. Apart from analog and digital conversions, the emphasis is also on the individual components of the devices employed and their relevance to the design.

Some of the materials used in this work are listed below:
Audio level indicator (LM3915)
Analog-to-digital converter (ADC0804)
Digital-to-analog converter (DAC0808)
Microcontroller (Atmel 89c52)
Resistors

Capacitors
Transistors
Light Emitting Diodes (LED)
Ear Phones

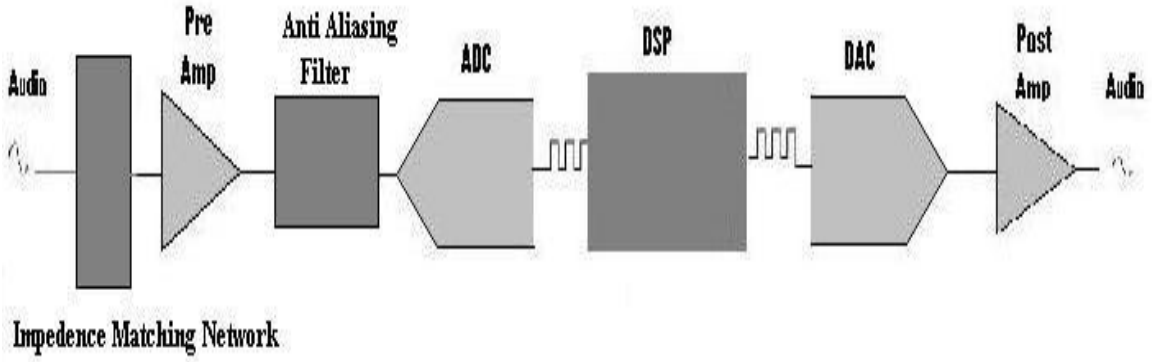


Figure 1: Block Diagram of Hearing Aid

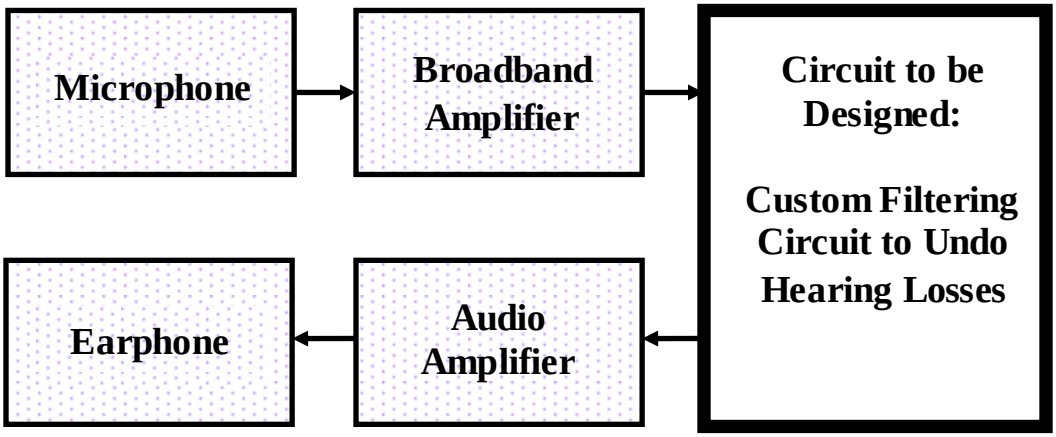


Figure 2. Functional Diagram of Hearing Aid.

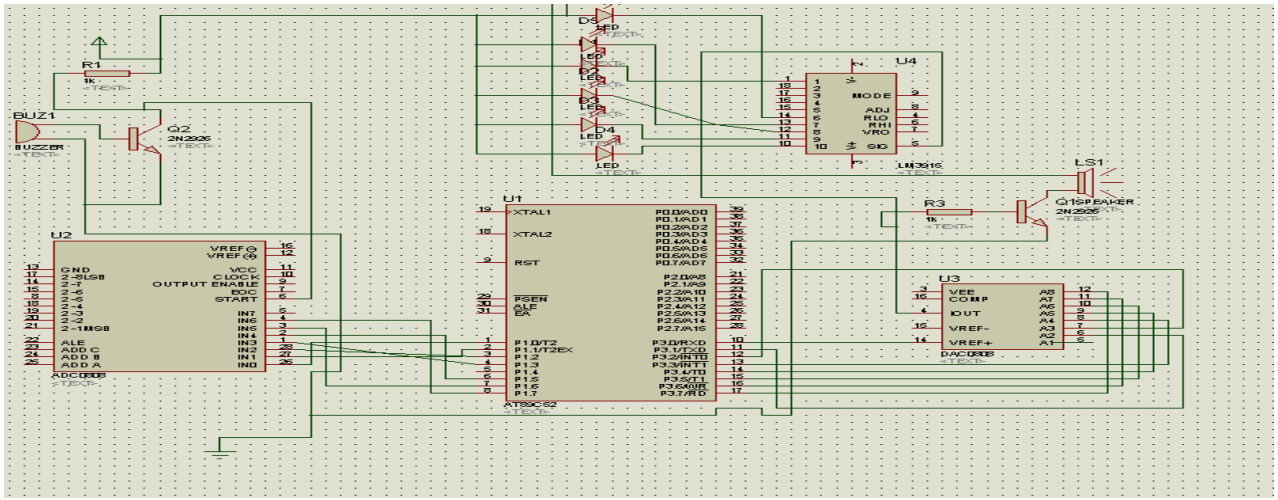


FIG 3. CIRCUIT DIAGRAM OF THE DESIGN

3. Principle Of Operation

The electret mic used in this project is positioned in such a way as to pick noise which is very suitable for hearing aid. The power supply should be 9V so as to pick small signal. The incoming sound oscillates the base of a transistor which in turn amplifies the signal at the collector. The output is interfaced to an ADC which converts the signal to digital output. This is necessary because the microcontroller cannot understand analog signal and therefore cannot process it. After the conversion, the microcontroller processes the signal and further feeds it in to the DAC for an analog conversion.

By so doing, the original signal is replicated at the other end, though in its amplified form. This output is now fed into the audio level indicator, LM3915, to indicate the level of the audio signal at the ear phone. However, this level can be controlled or adjusted according to one's desire, using a potentiometer. One end of the earphone is plugged into the jack on the system and the other end in the ear.

The subsystems of this work were interconnected or integrated to work as a system. The integration of this system is necessary so as to tailor each unit of the system to perform a targeted result. The transducer was interconnected to analog-to- digital converter to be able convert the analog quantity of the ambient temperature to a digital output. The ADC in turn was interfaced to a microcontroller for processing and giving its digital output to a DAC to reconvert to an analog signal. With all these interconnectivity, the different subsystems can now be said to be working as a system, in that they can interact with each other to produce a definite result.

USER GUIDE

To use the microcontroller based hearing aid, Power ON the switch and the red LED comes on. Plug in the Ear phones and Listen. The audio level is indicated by the five LEDs on the box. The loudest level is indicated by the yellow LED. Power OFF when not in use to avoid running the battery down. If the audio level LEDs glow continuously or do not correspond to the level of sound, then turn OFF the switch and ON again. If it persists, it could be as a result of either low battery or microcontroller lock up.

4. Conclusion

No matter what you do for a living, impaired hearing will affect your job performance. This is the information age and one of the primary ways to receive information is through hearing. The microcontroller based digital hearing aid can go a long way to achieve in people, increased enjoyment of social activities, improved ability to use the telephone, greater enjoyment of television and music, improved relationships and understanding of speech, increased ability to hear environmental sounds, increased self esteem and greater confidence. The design of the hardware using available components was achieved based on the principles of operation of individual electronic devices. The software program entails writing of programs that will drive the other subsystems to perform the desired operation.

References

[1]. Rickets T.A. (2011) Digital Hearing Aids: "Current state of the art" retrieved September 04, 2012 from http://www.asha.org/public/hearing/treatment/digital_aid.htm

[2]. Imran khan(2012) Advanced Hearing Aid Technology Retrieved September 20 2012, from http://www.eldercareresourcespittsburgh.com/advanced-hearing-aid-technology/

[3]. Alexandria V (2006) Hearing Industries Association.(HIA) special survey results on BTEs, directional and telecoil use. HIA journal 11, 15-25

Aru, Okereke Eze is a lecturer in the Department of Computer Engineering, Michael Okpara University of Agriculture, Umuahia, Abia State, Nigeria. His research

Interests include Computer Hardware design and maintenance, digital systems design using microcontrollers and other computer related subjects.

Ngwu, Chinyere Rosemary is a student in the Department of Computer Engineering, Michael Okpara University of Agriculture, Umuahia, Abia State, Nigeria. Her research interests include Computer Hardware design and maintenance, digital systems design using microcontrollers, etc.

Eng. Dr. Gozie Ihekweaba is a senior lecturer in the Department of Computer Engineering, Michael Okpara University of Agriculture, Umuahia, Abia State, Nigeria. His research Interests include Computer Hardware design and maintenance, digital systems design, Digital Electronics and other computer related subjects

Object-Oriented Full Function Point Analysis: An Empirical Validation

[1,]Sheeba Praveen, [2,]Dr. Rizwan Beg
[1,]Dept. CSE,Integral University Lucknow ,U.P
[2,] Dept .CSE & IT,Integral university Lucknow,U.P.

Abstract

this research work focuses on validation work of my proposed work to determine the functional size of real time application at early stage. This paper will describe how to calculate the functional size of real time system using my proposed model that is Object Oriented Size Estimation Model for Real Time Application [1]. Here in this paper I am validating my proposed model with the help of Real Time System. I am taking Human Emotion detection which is real time software and applying OOFFPA on it. OOFFPA metrics of my proposed model will calculate the size in terms of Function Point of this HED software and after calculating the size I compare this size with the size which is calculated by other metrics. Comparisons will prove that OOFFP metrics of my proposed model is best for size measurement of real time system.

Keywords: OOFFPA, OPA, FPA, OO, TFPCP, FFP, OOSE, UML, OOFP, SC, AS, AG, GN/SP, MX, OOMC, OOTC, OOCC, OOUCG, OORCG, OOCECE, OOCECX, OOCICR, OOCICW, OOEI, OOEO, OOEQ, OOILF, OOEIF, OOFC.

1. INTRODUCTION

This research is focuses on the cost estimation of Real Time System and for that I choose "Human Emotion Detection (HED)" which is a Real time system. This case study will help to validate my proposed work. In this case study I am using OOFFP metrics and Object oriented development procedures to estimate the size and cost of HED real time software.The HED can be applied in robotics in AI. The Robot can detect the human emotion and act accordingly the situation. This code can be interfaced with the help of embedded system using some programming language.
The camera will act as a sensor which captures the image of human and analyses the image on the basis of his/her facial expression. It detect the emotion i.e. Happy, Sad, Strange, Normal, Surprise etc. means act accordingly the situation.
Scenarios of HED are following:

1. Take an image as an input
2. Then apply skin color segmentation on an image
3. Find the largest connected region
4. Check the probability to become a face of the largest connected region
5. Convert the RGB image into Binary image
6. To detect that the image is human face or not
7. Separate the Eye and Lips from the image
8. Convert the Eye and Lips into Binary image
9. Apply the Bezier Curve on Eyes and Lips
10. Detect the human Emotion From the image

2. OBJECT ORIENTED DEVELOPMENT LIFE CYCLE

A. Object Oriented Design

In this research I follow the object oriented development life cycle to estimate size and cost of project [].

1. Object modeling Techniques
2. Dynamic modeling techniques
3. Functional modeling techniques

All three object oriented design techniques and associated models (D) are used in the proposed Size estimation model. These models help to estimate the size of the project. So In this research I am applying Management Function count and Control function count on all three models of object oriented design.

4.2.1 Unified Modeling Language

The Unified Modeling Language (UML) defines a large number of different diagrams. They are divided into following three categories: Static structure diagrams, Behavior diagrams and Implementation diagrams. In order to calculate the function point from the above diagrams, we use the sequence diagrams and class diagrams. Because these diagrams includes the information about all functions and data manipulated in the system. All these UML diagram are explained in my proposed paper" Full Functional size measurement Model applied to UML-based Real Time Application" [1].

3. PROPOSED RULES FOR OOFFPA FUNCTION POINT

Aim is to calculate the Unadjusted Function Point. Here I am proposing the following five steps to apply OOFFPA to the requirements/design specifications (class diagrams and sequence diagrams) based on the UML.

Step1 Determine the Type of Function Point Count
Step2 Identify the Counting Boundary
Step3 A. Count Data Function Types
Step4 Count Transactional Function Types
Step5 Count new Control Data Function Type
Step6 Count new Control Transactional function Type
Step7 Calculate Unadjusted Function Point Count
Step8 Determine value adjustment factor
Step9 Calculate Adjusted Function point

B. **Count Data Function Types**

a. Object Oriented ILF Complexity and Contribution (OOILF):

The OOILF steps are as follows:
1. Rate the OOILF complexity.
2. Translate the complexity to unadjusted function points.
3. Calculate the OOILF contribution to the total unadjusted function point count.

1. Rate Ooilf Complexity:

Rate the complexity of the OOILF using the following complexity matrix.

RER/DET	1 to 19 DET	20 to 50 DET	51 or more DET
1 RET	Low	Low	Average
2 to 5 RETs	Low	Average	High
6or more RETs	Average	High	High

Table1. OOILF

2. Translate OOILFs :

The following table translates the external inputs' functional complexity to unadjusted function points. Low=7, Average=10, High =15.

3. Calculate Ooilf Contribution:

The following table shows the total contribution for the OOILF function type.

Type	ILF	SR	ILF OOFP	SR OOFP	Total OOFP
SC	20	7 (L)	20*7	140+60	200
AS	14	7 (L)	14*7	98+60	158
AG	14	7 (L)	14*7	98+60	158
GN	7	10 (A)	7*10	70+60	130
MX	7	10 (A)	7*10	70+60	130

Table2. OOILF

b. Object Oriented EIF Complexity and Contribution (OOEIF):

The OOEIF steps are as follows:
1. Rate the OOEIF complexity.
2. Translate the complexity to unadjusted function points.
3. Calculate the OOEIF contribution to the total unadjusted function point count.

1. Rate Ooeif Complexity:

Rate the complexity of the OOEIF using the following complexity matrix.

RET/DET	1 to 19 DET	20 to 50 DET	51 or more DET
1 RET	Low	Low	Average
2 to 5 RETs	Low	Average	High
6or more RETs	Average	High	High

Table3. OOEIF

2. Translate Ooilfs :

The following table translates the external inputs' functional complexity to unadjusted function points. Low=5, Average=7, High =10.

3. Calculate Ooeif Contribution:

The following table shows the total contribution for the OOEIF function type. As there are 12 concrete methods in the model, service requests contribute 12 * 5 = 60 OOFPs (SR OOFP). The Value 7 is rated as Low and it is weighted 4.

Type	EIF	SR	ILF OOFP	SR OOFP	Total OOFP
SC	8	5 (L)	8*5	40+60	100
AS	7	5 (L)	7*5	35+60	95
AG	7	5 (L)	7*5	35+60	95
GN	5	10 (A)	5*10	50+60	110
MX	4	10 (H)	4*10	40+60	100

Table4. OOEIFs

C. **Count Transactional Function Types**

a. Object Oriented External Inputs Complexity and Contribution (OOEI):

The OOEI steps are as follows:

1. Rate the OOEI complexity.
2. Translate the complexity to unadjusted function points.
3. Calculate the OOEI contribution to the total unadjusted function point count.

1. Rate Ooei Complexity:

Rate the complexity of the OOEI using the following complexity matrix.

RER/DET	1 to 5 DET	6 to 19 DET	20 or more DET
0 to1 FTR	Low	Low	Average
2 to 3 FTRs	Low	Average	High
4 or more FTRs	Average	High	High

Table5. OOEIs Complexity Rate

2. Translate OOEIs :

The following table translates the external inputs' functional complexity to unadjusted function points. Low=3, Average=4, High =6. The following table shows the functional complexity for each OOEI.

OOEI	FTRs	DETs	Functional Complexity
Browse Image	1	4	L
Apply Skin colour	1	3	L
Largest connected Region	2	10	A
RGB Image	2	8	A
Binary Image	1	4	L
Image Lip	0	2	L
Image Eye	0	2	L
Image Eyebrow	0	2	L
Bezier Curve	1	3	H

Table6.OOEIs

3. Calculate OOEI Contribution:

The following table shows the total contribution for the OOEI function type.

Total no. of OEI Function Type	Functional Complexity	Total Complexity	Total function type
8	5(L)	5*3	15
	2(A)	2*4	8
	1(H)	1*6	6
	Total	+	29

Table7. Total OOEIs

b. Object Oriented External Output Complexity and Contribution (OOEO):

The OOEO steps are as follows:

1. Rate the OOEO complexity.

2. Translate the complexity to unadjusted function points.
3. Calculate the OOEO contribution to the total unadjusted function Point count.

1. Rate Ooeo Complexity:

Rate the complexity of the OOEO using the following complexity matrix.

RER/DET	1 to 5 DET	6 to 19 DET	20 or more DET
0 to1 FTR	Low	Low	Average
2 to 3 FTRs	Low	Average	High
4 or more FTRs	Average	High	High

Table8. OOEOs Complexity Rate

2. Translate Ooeos :

The following table translates the external inputs' functional complexity to unadjusted function points. Low=4, Average=5, High =7.

The following table shows the functional complexity for each OOEO.

OOEO	FTRs	DETs	Functional Complexity
Valid Face	0	6	H
Connected	2	4	A
Happy emotion	1	3	L
Fear Emotion	1	5	L
Surprise Emotion	1	7	L
Sadness Emotion	1	4	L
Anger Emotion	1	6	L
Disgust Emotion	1	2	L

Table9.OOEOs

3. Calculate Ooeo Contribution:

The following table shows the total contribution for the OOEO function type.

Total no. of OOEO Function Type	Functional Complexity	Total Complexity	Total function type
6	4(L)	5*4	16
	2(A)	2*5	10
	0(H)	1*7	0
	Total	+	26

Table10. Total OOEIs

c. Object Oriented External Inquiries Complexity And Contribution (OOEQ)

The OOEQ steps are as follows:

1. Rate the OOEQ complexity.
2. Translate the complexity to unadjusted function points.
3. Calculate the OOEQ contribution to the total unadjusted function point count.

1. Rate Ooeq Complexity:

Rate the complexity of the OOEQ using the following complexity matrix.

RER/DET	0 to 5 DET	6 to 19 DET	20 or more DET
0 to1 FTR	Low	Low	Average
2 to 3 FTRs	Low	Average	High
4 or more FTRs	Average	High	High

Table11. OOEQs Complexity Rate

2. Translate Ooeqs :

The following table translates the external inputs' functional complexity to unadjusted function points. Low=3, Average=4, High =6.

OOEQ	FTRs	DETs	Functional Complexity
Fear samples	1	2	L
Surprise samples	1	1	L
Sadness samples	1	4	L
Anger samples	1	6	L
Disgust samples	1	4	L
Happy samples	1	3	L

Table12.OOEQs

3. Calculate Ooeqs Contribution:

The following table shows the total contribution for the EI function type.

Total no. of OEI Function Type	Functional Complexity	Total Complexity	Total function type
5	5(L)	5*3	15
	0(A)	0*4	8
	0(H)	0*6	6
	Total	+	29

Table13. Total OOEQs Complexity Rate

The following table shows the total contribution for the OOEI, OOEO and OOEQ function type

Function Type	Total Function	Low	Average	High	Total function type
OOEI	9	6*3	2*4	1*6	32
OOEO	8	6*4	1*5	1*7	26
OOEQ	6	6*3	0*4	0*6	18
+					76

Table14.Total Function type

D. **Count new Control Data Function And Transactional Function Type**

a. Object Oriented Full Function Points (OOFFP) new Control and Transactional Function Count

OOFFP new function types	**Description**	**No. of sub-processes**
OOUCG	Data updated by the application	8
OORCG	Data not updated by the application	10
OOECE	Incoming external message	1
OOECX	Outgoing external message	1
OOICR	Referred attribute in an elementary action	7
OOICW	Update attribute in an elementary action	15
	Total	42

Table15.Total new control and Transactional Function count.

Low=2, Average=3, High =5(for NCDFC).
Low=3, Average=4, High =6(for NTFC)

Function Type	Total function	Low	Average	High	Total function type
OOUCG	8	6*2	2*3	0*5	18
OORCG	10	6*2	3*3	1*5	26
OOECE	1	1*3	0*4	0*6	3
OOECX	1	0*3	1*4	0*6	4
OOICR	7	3*3	3*4	1*6	27
OOICW	14	7*3	7*4	0*6	49
					127

Table16.Total New function Type

E. **Calculate UnAdjusted Function Point Count (UFP)**

The following table shows the contribution of the application functionality to the unadjusted function point count.

Function Type	Total functional Complexity
OODFC	1341
OOTFC	76
NCDFC & NTFC	42
Total	1459

Table17. Total UFP of OOFFP

F. **Procedures to Determine the VAF**

The following steps outline the procedures to determine the value adjustment factor (chapter3).

1. Evaluate each of the 14 general system characteristics on a scale from zero to five to determine the degree of influence (DI).

2. Add the degrees of influence for all 14 general system characteristics to produce the total degree of influence (TDI).

3. Insert the TDI into the following equation to produce the value adjustment factor.

VAF = (TDI * 0.01) + 0.65

For example, the following value adjustment factor is calculated if there are three degrees of influence for each of the 14 GSC descriptions

(3*14).

VAF = (42 * 0.01) + 0.65

VAF = 1.07

G. **Calculate adjusted Function point count(AFP)**

Using the complexity and contribution counts for this example, the development project count is shown below. The value adjustment factor (VAF) for this example is 1.07.

AFP = UFP * VAF

AFP = 1459 * 1.07

AFP = 1561.13 or 1561

H. **Assumptions & Results**

Past date indicate that one FP translate into 60 times of code (if an OOP language is to be used) LOCs = 60 * 1561 = 93660 (approximately) Past project have found an average of 3 errors per function point during analysis and design reviews and 4 errors per function point during unit and integration testing. Thus, possible number of errors in analysis and design reviews should be 3*1561 i.e. 4683. At the time of testing

Possible number of errors should be 4*1561 i.e. 6244. Thus total possible number of errors should be 10927.

4. Verification of Results

After implementation it was found that lines of code are 94181, which is more than calculated LOCs (on the basis of FPs in analysis phase) by a value of 1561.

Errors found at the time of analysis and design reviews are 4683 and errors found at the time of testing are 6244. Thus total errors found are 10927 which is more than calculated by a value of 1561.

5.1 COMPARATIVE STUDY OF HED ON DIFFERENT METRICS

Metrics	Count number	Error-proness
LOC	94181	10927
FP	1850	7259
OOFFP	1561	5463.5

Table17. Metrics comparison

5. CONCLUSION

This above table shows that the OOFFP metrics is best for size measurement of real time system. By using this size we can calculate the Productivity and quality of real time system. So OOFFPA helps to increase the performance MIS as well as real time software.

References

[1] Full Functional size measurement Model applied to UML-based Real Time Application is accepted in International Conference on Recent Trends in Control Communication and Computer Technology(RTCCT-2012), Paper ID: RTCCCT-29SEP12-030

URL http://www.interscience.ac.in/Chennai/RTCCCT/index.html

[2] Abran, A., Desharnais, J.-M., Maya, M., St-Pierre, D., & Bourque, P. (1998). Design of Functional Size Measurement for Real-Time Software. Montréal, Université du Québec à Montréal [www document]..

URL http://www.lrgl.uqam.ca/publications/pdf/407.pdf

[3] Bohem, R. (1997). Function Point FAQ. Metuchen,USA, Software Composition Technologies, Inc URL http://ourworld.compuserve.com/homepage/softcom/

[4] Desharnais, J.-M. & Morris, P. (1996). Validation Process in Software Engineering: an Example with Function Points. In Forum on Software Engineering Standards (SES'96), Montreal [www document]..

URL http://www.lrgl.uqam.ca/publications/pdf/104.pdf

[5] Introduction to Function Point Analysis (1998). GIFPA, Issue 2, summer [www document]. URL http://www.gifpa.co.uk/news/News2Web.pdf

[6] International Standards Organisation (ISO) (1991). Information Technology ± Software Product Evaluation ± Quality Characteristics and Guidelines for their Use (ISO/IEC IS 9126). Geneve: ISO/IEC.

[7] Software Metrics - why bother?.(1998). GIFPA, Issue 1, spring

URL http://www.gifpa.co.uk/news/Issue1_ed2.pdf

[8] St-Pierre, D., Maya, M., Abran, A., Desharnais, J.-M. & Oligny, S. (1997a). Full Function Points: Counting Practices Manual. Montréal, Université du Québec à Montréal [www document].

URL http://www.lrgl.uqam.ca/publications/pdf/267.pdf

[9] St-Pierre, D., Maya, M., Abran, A., Desharnais, J.-M. & Oligny, S. (1997b). Measuring the functional size of real-time software. Montréal, Université du Québec à Montréal [www document].

URL http://www.lrgl.uqam.ca/publications/pdf/330.pdf

[10] Fenton, N. E. & Pfleeger, S. L. (1996). Software Metrics - A Rigorous & Practical Approach. London: International Thomson Computer Press.

[11] Functional Size Measurement for Real-Time Software. Montréal, Université du Québec à J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[12] An Empirical Validation of Software Quality Metric Suites on Open Source Software for Fault-Proneness Prediction in Object Oriented System I. S. Jacobs and C. P. Bean, "Fine particles, thin films and

Parametric Analysis of Four Wheel Vehicle Using Adams/Car

Jadav Chetan S.[1], Patel Priyal R.[2]

[1] Assistant Professor at Shri S'ad Vidya Mandal Institute of Technology, Bharuch-392001, Gujarat, India.
[2] PG Student at Shri S'ad Vidya Mandal Institute of Technology, Bharuch-392001, Gujarat, India.

Abstract:

Inspiring from the Multibody dynamics this paper has been carried out for estimating the dynamics of vehicle in motion. As there is more and more importance is given to the handling performance to the vehicle for its comfort, certain parameters like roll, yaw, pitch and, side slip angle etc., are to be studied, analyzed, controlled and adversely changed for increasing the overall performance and vehicle behavior. Here the focus is intended on the roll, pitch and yaw phenomena of the vehicle while changing the lane tack and its effects are simulated at different speeds. Results show certain variation and its effect. Hence this tool for multibody dynamics proves more and more efficient for such conditions.

Keywords: ADAMS/CAR, Lateral force, Multibody Dynamics, Pitch, Roll, Tire slip angle and, Yaw.

1. Introduction:

The focus of the paper is to estimate vehicle response in terms of roll ,yaw and pitch effects in front-wheel steered, rear-wheel driven four wheeled vehicles in real time. It gives a measure of the lateral forces produced at the tire-road contact patches while cornering and lane change during vehicle motion which make the vehicle turn. Physically it represents the twist in the treads of the tires and body frame. It is very difficult if not possible to directly measure the yaw rate angles of the chassis in vehicle, hence indirect methods have to be applied to estimate them. Knowledge of side slip angle is a required for advanced vehicle control systems like braking control, stability control, security actuators and for validating vehicle simulators. These controllers increase the safety of the vehicle, and make the response more predictable. The knowledge of vehicle dynamics can also be applied to decrease road damage caused by vehicles. The wheel hub is coupled to the vehicle through the suspension and steering mechanisms. Thus the steering and suspension mechanisms affect each other's behaviour. While cornering, changes will be induced in variables associated with both the mechanisms. The project endeavours to analyse these changes to estimate yaw rate rolling and pitch angle based on chassis deflection information. This work has been performed to meet the objectives:

1. To develop a method of estimating roll, yaw and, pitch in front-wheel steered four wheel vehicles.
2. To instrument a vehicle and conduct on road tests to validate the estimation method.

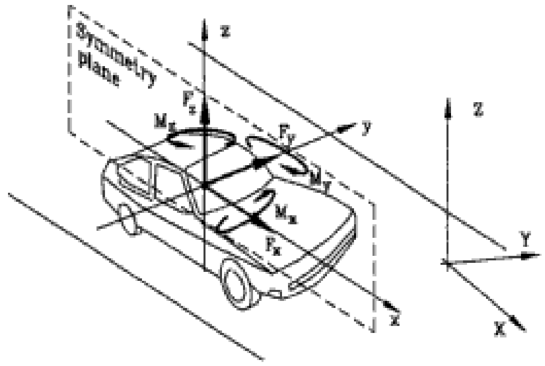
A model is constructed in ADAMS/CAR to simulate test conditions and predict the results for the tests to be conducted. The predictions of this model are verified with experimental results from literature. An open-loop estimator that uses a three degree of freedom vehicle model is used to estimate results using real time experimental data from tests conducted. We conclude that a three degree-of-freedom model is a good start for use in estimation techniques of vehicle dynamics and performance.

2. Vehicle dynamics:

The basic information required to understand vehicle architecture as well as dynamic behaviour of front wheel steered, four wheel vehicles is its mathematical formulation. It also explains how lateral forces are generated while a vehicle negotiates a turn, and their relation with the slip angle. Four wheel vehicle model is discussed in detail. This model is used to build observers (estimators) for estimating vehicle handling behaviour. A brief explanation of currently employed methods of estimation lateral forces and slip angles is given below. Also, the system consists of the dependent mechanism: steering and suspension system. The system shown here consists of rack and pinion steering mechanism and Double Wishbone Suspension.

The vehicle coordinate system shown in the Figure is explained below:

- Linear motion along x direction is known as longitudinal motion.
- Rotational motion about x axis is known as roll.
- Linear motion along y direction is known as lateral or transverse motion.
- Rotational motion about y axis is known as pitch.
- Linear motion along z direction is known as vertical motion.
- Rotational motion about z axis is known as yaw.



3. Vehicle Model Formulation For Parameters Identification:

For simulating the lateral dynamics of the vehicle a four wheel 3 Degree of freedom model is used containing lateral velocity (V), yaw rate (r) and, roll angle (ϕ).The input of the model is the steering angle on the front tires. Also the continuum mass of the vehicle is modelled by three lumped masses which are front and rear unsprung masses (M_{uf}, M_{ur}) and sprung mass (M_s) so the entire vehicle mass is:

$$M_t=M_s + M_{uf} +M_{ur}. \tag{1}$$

In order to derive the equation of motion a moving reference frame is attached to the vehicle with its origin at the centre of gravity as shown in Figure. Since the coordinate system is attached to the vehicle the inertia properties of the vehicle will remain constant Also as the result of symmetry assumption all the products of inertia are ignored .The state variables are assumed to be lateral velocity yaw rate and roll angle.

Using the above assumptions the equations describing the motion are:

$$M_t(\dot{V}+ru) + M_s h_s\ddot{\phi} = F_{yfr}+F_{yfl}+F_{yrr}+F_{yfl}, \tag{2}$$

$$I_{xx}\ddot{\phi} + M_s h_s(\dot{V}+ru) = L_s, \tag{3}$$

Where:

$$L_s = M_s g h_s\phi - K_\phi\phi - C_\phi\dot{\phi}, \tag{4}$$

Where, Ms is the sprung mass, which is the mass supported by the vehicle suspension, (I_{xx})is the sprung mass moment of inertia about longitudinal axis (x), (I_{zz})is the moment of inertia of the entire vehicle about vertical axis (z), K_ϕ and C_ϕ are roll stiffness and roll damping coefficient of suspensions respectively (h_s)is the vertical distance of CG from the roll axis, a and b are the distances of the front and rear axles from CG ,u is the longitudinal speed of the vehicle which is constant in vehicle maneuvers and F_{yfr}, F_{yfl}, F_{yrr}, F_{yrl} are the tire cornering forces of front right front left rear right and rear left respectively.

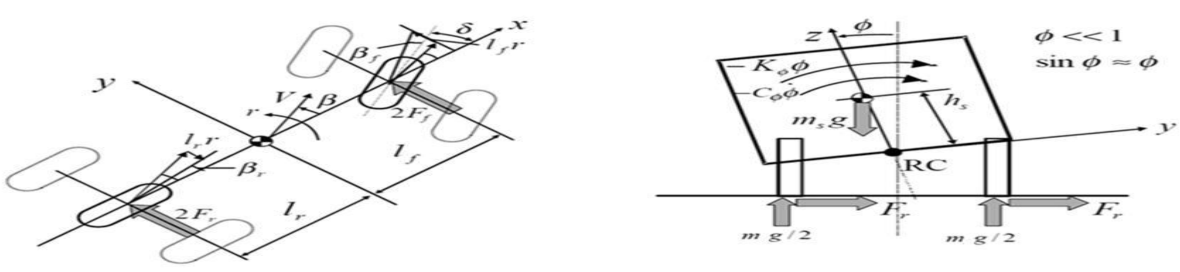


Figure 2 .Vehicle Model

The cornering force of a tire is mainly dependent on the slip angle vertical load longitudinal slip and camber angle of that tire. In this paper a tire model in which the cornering force of the tire is a function of the cornering stiffness vertical load slip angle and longitudinal slip has been used and the effects of camber angle and aligning moments are ignored. In order to compute tire forces the slip angles of the tire should be calculated as below:

$$\alpha_{fr} = \delta - \tan^{-1}\left(\frac{V+ar}{u-t_{f_2}^{\cdot}}\right) \quad (5)$$

$$\alpha_{fl} = \delta - \tan^{-1}\left(\frac{V+ar}{u+t_{f_2}^{\cdot}}\right) \quad (6)$$

$$\alpha_{rr} = \tan^{-1}\left(\frac{br-V}{u-t_{r_2}^{\cdot}}\right) \quad (7)$$

$$\alpha_{rl} = \tan^{-1}\left(\frac{br-V}{u+t_{r_2}^{\cdot}}\right) \quad (8)$$

Where δ is the steer angle as the input of the model and t_f , t_r are the front and rear tread widths of the vehicle respectively.

3.1 Lateral load for its effects:

Also, for obtaining the lateral load transfer some equations to describe vertical forces on each tire have been written. In this concept lateral load transfer is assumed to be the result of three phenomena which are body roll, roll centre height and unsprung mass.

Lateral load transfer, due to body roll, is an follows:

$$F_{f1} = \left(\frac{k_f h_s M_s}{K_\emptyset t_f}\right).\left(a_y \cos\emptyset + g \sin\emptyset\right) \quad (9)$$

$$F_{r1} = \left(\frac{k_r h_s M_s}{K_\emptyset t_r}\right).\left(a_y \cos\emptyset + g \sin\emptyset\right) \quad (10)$$

Where k_f and k_r are the front and rear roll stiffness and a_y is the lateral acceleration Which is given as:

$$a_y = \dot{V} + ru + \frac{M_s h_s \ddot{\emptyset}}{M_t} \quad (11)$$

Lateral load transfer, due to roll center height, is as follows:

$$F_{f2} = \frac{M_s b h_f a_y}{t_f(a+b)} \quad (12)$$

$$F_{r2} = \frac{M_s a h_r a_y}{t_r(a+b)} \quad (13)$$

Where h_f and h_r are the front and rear roll center heights respectively.

And lateral load transfer,due to unsprung masses, is :

$$F_{f3} = M_{uf} a_y \frac{n_f}{t_f} \quad (14)$$

$$F_{r3} = M_{ur} a_y \frac{n_r}{t_r} \quad (15)$$

The vertical load on each tire can be described by the following equations :

$$F_{zfr} = \frac{W_f}{2} - F_{f1} - F_{f2} - F_{f3} \quad (16)$$

$$F_{zfl} = \frac{W_f}{2} + F_{f1} + F_{f2} + F_{f3} \quad (17)$$

$$F_{zrr} = \frac{W_r}{2} - F_{r1} - F_{r2} - F_{r3} \quad (18)$$

$$F_{zrl} = \frac{W_r}{2} + F_{r1} + F_{r2} + F_{r3} \quad (19)$$

Where W_f and W_r are the static load distribution on the front and rear axles and can be computed by the following equations:

$$W_f = M_t g \frac{b}{a+b} \quad (20)$$

$$W_r = M_t g \frac{a}{a+b} \quad (21)$$

3.2 Lateral Force and Tire Slip Angle:

While cornering, a vehicle undergoes lateral acceleration. As the tires provide the only contact of the vehicle with the road, they must develop forces which result in this lateral acceleration. When a steering input is given, the successive treads of the tires that come in contact with the road are displaced laterally with respect to the treads already in contact with the road. Thus an angle is created between the angle of heading and the direction of travel of the tire. This angle is known as the tire slip angle which gives an estimate of twist of the treads of the tire. It can also be defined as the ratio of the lateral and forward velocities of the wheel. The twisted treads try to get back to their original positions, thus producing the force required for lateral acceleration. This force is known as the Lateral Force (F_y) or the Cornering Force. At a given load, the cornering force grows with slip angle. At low slip angles (5 degrees or less) the relationship is linear. In this region, cornering force is often described as $F_y = C_\alpha \alpha$. The proportionality constant C_α is known as cornering stiffness and is defined as the slope of the curve for F_y versus α at $\alpha = 0$.

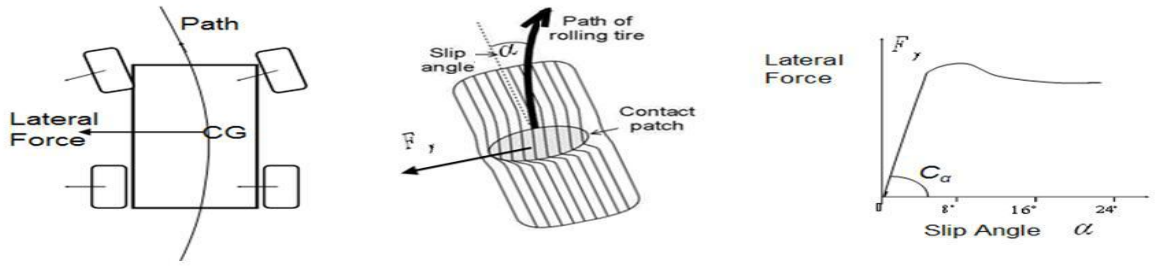
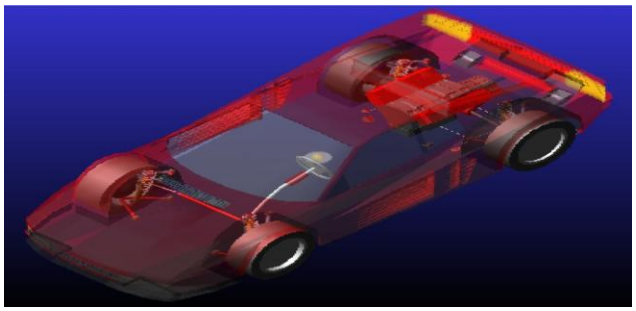


Figure 3. Three Degree of Freedom Automobile Model

Since the stability control system has the ability to affect the vehicle's attitude and motion, a function normally reserved for the driver, it needs to accurately interpret what the driver intends for the vehicle motion in order to provide added directional control (within physical limitations) as a driver's aid. Responsiveness, consistency, and smoothness are essential for a driver's confidence and comfort with the system. These are the guiding principles for our development. A driver typically expresses directional intent through the steering wheel. The angular position of the steering wheel is the first measure of driver intent.

4. Simulation RESULTS:

The below shown figure indicates the model which is to be simulated for the estimation of vehicle dynamics and its handling. The model shown below is ADAMS/CAR generated model:



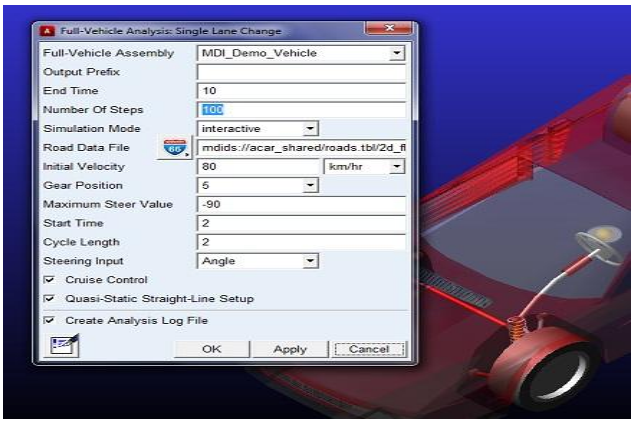


Figure 4. ADAMS/CAR and Full Vehicle Analysis

With reference to the car model here full vehicle analysis is to be carried out. Certain graphs are discussed here.The graph shown here is angular v/s time for the rolling motion of vehicle chassis at speed of 40, 80 and, 120 km/hr.

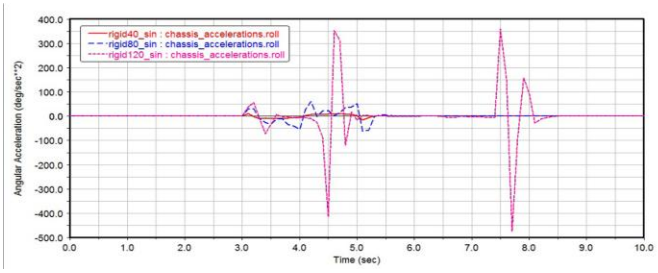


Figure 5. Angular acceleration v/s time for Roll

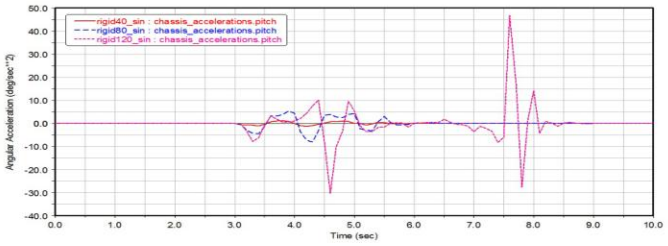


Figure 6. Angular acceleration v/s time for Pitch

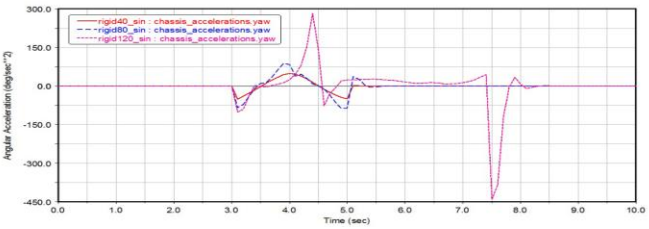


Figure 7.Angular acceleration v/s time for Yaw

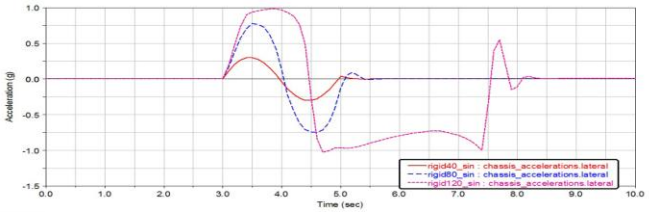


Figure 8.Angular acceleration v/s time for lateral acceleration

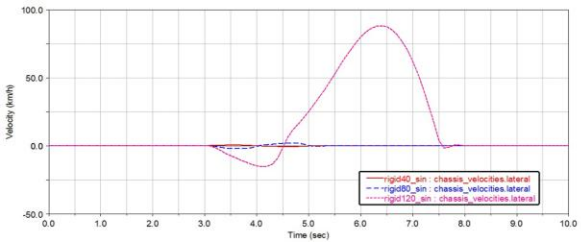


Figure 9.Chassis velocity v/s time for lateral

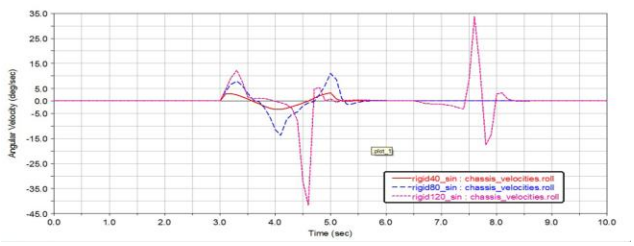
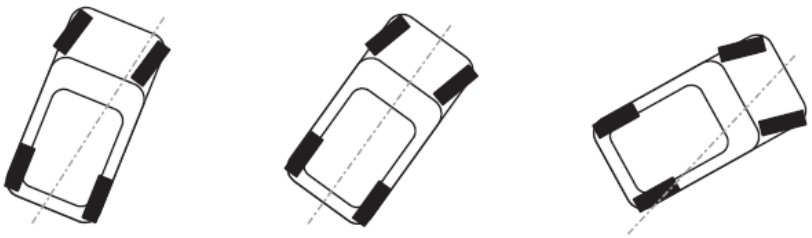


Figure 10.Angular Velocity v/s time for Roll

Here the lateral effect on vehicle handling is shown over here and its effects to yaw, roll and, pitch is mentioned:In scenario (a), yaw rate gain is reduced further than lateral acceleration gain. In order to accommodate the changes in both lateral acceleration and yaw rate, the radius of the path must increase and so the vehicle has a period of adjustment to a new, wider line in the curve. Most drivers notice this and instinctively reduce vehicle speed to restore the desired path over the ground. If uncompensated, it leads to a vehicle departing the course (road, track, etc.) in an attitude that is basically forwards. This is by far the most common behaviour for road vehicles.In scenario (b), lateral acceleration and yaw rate gain change in some connected manner and the vehicle will maintain course although it might need some modification to steering input. Excess speed for a curve will lead to the vehicle running wide but with no sense of 'turning out of the curve'. Such a vehicle generally feels benign although the progressive departure can mean it is unnoticed by inattentive drivers.In scenario (c), lateral acceleration gain reduces more than yaw rate gain. This leads to an 'over-rotation' of the vehicle when viewed in plan. Depending on the severity of the mismatch, the change may lead to a spin out of the curve. From inside the vehicle there is a pronounced sense of the rear end of the vehicle departing first but objectively the vehicle may not actually oversteer in the classical sense.it may simply move 'towards neutrality'. Vehicles that preserve yaw rate gain as they lose linearity are widely regarded as fun to drive and sporty.



4.1 Yaw, Rate and angular Acceleration

The vehicle is driven at known speeds, on a pre- marked course. Hence, by knowing the trajectory of the curve and speed of the vehicle, the yaw rate and lateral acceleration is calculated.

5. Conclusion

This paper aimed at estimating the dynamic parameters like roll, yaw, pitch and side slip angle in front-wheel steered, rear-wheel driven four wheeled vehicles. Side slip angle cannot be measured directly; hence estimation has been built to calculate side slip angle from measurable variables. A model was constructed in ADAMS/CAR to predict this variable for a candidate set of test conditions. Since the predictions of the ADAMS/CAR model were found to be in reasonable agreement with the experimental results reported in literature, it can be conclude that the three degree-of-freedom model provides good prediction capabilities for estimation of vehicle dynamics.

Reference

[1] M R Bolhasani and S Azadi, "Parameter Estimation of Vehicle Handling Model Using Genetic Algorithm" Scientia Iranica Vol.11 No. 1&2, pp 121-127.
[2] Marcelo parado, Argemiro costa, "Bus handling validation and analysis using ADAMS/CAR" ADAMS user conference 2007.
[3] Mohan M N and S R Sankapal, "Study of handling characteristics of a passenger car through Multibody Simulation". SASTech vol. 5 no. 1 April 2005.
[4] S Hegazy, H Rahnejat and K Hussain, "Multi-body dynamics in full-vehicle handling analysis" Proc Instn Mech Engrs Vol 213 Part K..
[5] John Grogg, Qinghui Yuan and Jae Lew, "Dynamic Modeling of Torque-Biasing Devices for Vehicle Yaw Control" SAE Automotive Dynamics, Stability & Controls Conference and Exhibition Novi, Michigan February 14-16, 2006.
[6] Michael Blundell and Damian Harty, "The Multibody Systems Approach to Vehicle Dynamics, Elsevier Butterworth-Heinemann 2004.

Study of Genetic Algorithm for Process Scheduling in Distributed Systems

Usha Barad

1,PG Student of Computer Engineering, Merchant Engineering College, Gujarat, India

Abstract : This paper presents and evaluates a new method for process scheduling in distributed systems. The problem of process scheduling in distributed system is one of the important and challenging area of research in computer engineering. Scheduling in distributed operating system has an important role in overall system performance. Process scheduling in distributed system can be defined as allocating processes to processor so that total execution time will be minimized, utilization of processors will be maximized and load balancing will be maximized. Genetic algorithm is one of the widely used techniques for constrain optimization. The scheduling in distributed systems is known as an NP-complete problem even in the best conditions, and methods based on heuristic search have been proposed to obtain optimal and suboptimal solutions. In this, paper using the power of genetic algorithms. We solve this problem considering load balancing efficiently. We evaluate the performance and efficiency of the proposed algorithm using simulation result.

Keyword: Distributed system, scheduling, Genetic algorithm, load balancing

I. Introduction

Scheduling in distributed operating systems is a critical factor in overall system efficiency. A Distributed Computing system (DCS) is comprised of a set of Computers (Processors) connected to each other by communication networks. Process scheduling in a distributed operating system can be stated as allocating processes to processors so that total execution time will be minimized, utilization of processors will be maximized, and load balancing will be maximized. The computational complicated process cannot be executed on the computing machine in an accepted interval time. Therefore, they must be divided into small sub-process. Process scheduling in distributed system is done in two phases: in first phase processes are distributed on computers and in second processes execution order on each processor must be determined [1]. Several methods have been proposed to solve scheduling problem in DCS. The methods used to solve scheduling problem in distributed computing system can be classified into three categories graph theory based approaches , mathematical models based methods and heuristic techniques [1]. Heuristic algorithm can be classified into three categories iterative improvement algorithms [2], the probabilistic optimization algorithms and constructive heuristics. Heuristic can obtain sub optimal solution in ordinary situations and optimal solution in particulars. The first phase of process scheduling in a distributed system is process distribution on computer. The critical aspects of this phase are load balancing. Recently created processes may be overloaded heavily while the others are under loaded or idle. The main objectives of load balancing are to speared load on processors equally, maximizing processors utilization and minimizing total execution time [4]. The second phase of process scheduling in distributed computing system is process execution ordering on each processor. Genetic algorithm used for this phase. Genetic algorithm is guided random search method which mimics the principles of evolution and natural genetics . Genetic algorithms search optimal solution from entire solution space. In dynamic load balancing, processes must be dynamically allocated to processors in arrival time and obtain a near optimal schedule, therefore the execution of the dynamic load balancing algorithm should not take long to arrive at a decision to make rapid task assignments.

A GA starts with a generation of individuals, which are encoded as strings known as chromosomes. A chromosome corresponds to a solution to the problem. A certain fitness function is used to evaluate the fitness of each individual. Good individuals survive after selection according to the fitness of individuals. Then the survived individuals reproduce offspring through crossover and mutation operators. GA-based algorithms have emerged as powerful tools to solve NP-complete constrained optimization problems, such as traveling salesman problem, job-shop scheduling and flow-shop scheduling, machine learning, VLSI technology, genetic synthesis and etc.In this paper using the power of genetic algorithms we solve this problem considering load balancing efficiently. The proposed algorithm maps each schedule with a chromosome that shows the execution order of all existing processes on processors. The fittest chromosomes are selected to reproduce offspring; chromosomes which their corresponding schedules have less total execution time, better load-balance and processor utilization. We assume that the distributed system is non-uniform and non-preemptive, that is, the processors may be different, and a processor completes current process before executing a new one. The load-balancing mechanism used in this paper only schedule processes without process migration and is centralized.

2. Model and Problem Definition

2.1. System Model

The system used for simulation is loosely coupled non-uniform system, all task are non-pre-emptive and no process migration are assumed. The process scheduling problem considered in this paper is based on the deterministic model. The most important purposes of distributed system are assigned as providing an appropriate and efficient environment for sharing resource, having an acceptable speed and high reliability and availability.Network topology, processors speed, communication channels speed and so on. Since we study a deterministic model, a distributed system with *m* processors, *m* > 1 should be modeled as follows:P= {p1, p2, p3,..., pm} is the set of processors in the distributed system. Each processor can only execute one process at each moment, a processor completes current process before executing a new one, and a process cannot be moved to another processor during execution. G is an m × m matrix, where the element guv 1≤ u,v ≤ m of G, is the communication delay rate between Pu and Pv. H is an m × m matrix, where the element huv 1≤ u, v ≤m of H, is the time required to transmit a unit of data from Pu and Pv. It is obvious that huu=0 and ruu=0.T = {t1, t2, t3,...,tn} is the set of processes to execution. E is an n × m matrix, where the element eij 1 ≤ i ≤ n, 1 ≤ j ≤ m of E, is the execution time of process ti on processor pj. In distributed systems the execution time of an individual process ti on all processors is equal.F is a linear matrix, where the element fi 1≤ i ≤ n of F, is the target processor that is selected for process ti to be executed on.C is a linear matrix, where the element ci i ≤ i ≤ n of C, is the processor that the process ti is presented on just now.

The problem of process scheduling is to assign for each process ti Є T a processor fi Є P so that total execution time will be minimized, utilization of processors will be maximized, and load balancing will be maximized.

The processor load for each processor is the sum of process execution times allocated to that process [1].

$$Load\ (p_i) = \sum_{j=1}^{\substack{No.of\ allocated\\ processes\ on\\ processor\ i}} a_{j,i} + \sum_{k=1}^{\substack{No.of\ new\ assigned\\ processes\ on\\ processor\ i}} a_{k,i} \quad \ldots\ldots (1)$$

The length or *max span* of schedule T is the maximal finishing time of all the processes or maximum load. Also, communication cost (CC) to spread recently created processes on processors must be computed:

Maxspan (T) = max (Load (p_i)) for each 1≤i≤ Number of processors … (2)

$$CC(T) = \sum_{i=1}^{nNo.of\ new\ processes} (r_{c_i f_i} + h_{c_i f_i} X d_i) \quad \ldots\ldots (3)$$

The processor utilization for each processor is obtained by dividing the sum of processing times by scheduling length. The average of process utilization is obtained by dividing the sum of all utilizations by number of processors [2].

$$U(p_i) = {Load\ (p_i)}/{maxspan} \quad \ldots\ldots\ldots (4)$$

$$AvgU = {(\sum_{i=1}^{No.of\ processors} U(p_i))}/{No.of\ processors} \quad \ldots\ldots (5)$$

2.2. Genetic algorithm

Genetic algorithm is guided random search algorithm based on the principles of evolution and natural genetics. It combines the exploitation of the past results with the exploration of new areas of the search space. Genetic algorithms, as powerful and broadly applicable stochastic search and optimization techniques, are the most widely known types of evolutionary computation methods today. In general, a genetic algorithm has five basic components as follows [3]:

1. An encoding method that is a genetic representation (genotype) of solutions to the program.
2. A way to create an initial population of individuals (chromosomes).
3. An evaluation function, rating solutions in terms of their fitness, and a selection mechanism.
4. The genetic operators (crossover and mutation) that alter the genetic composition of offspring during reproduction.
5. Values for the parameters of genetic algorithm.

Genetic algorithm maintains a population of candidate solutions that evolves over time and ultimately converges. Individuals in the population are represented with chromosomes. Each individual is numeric fitness value that measures how well this solution solves the problem. Genetic algorithm contains three operators. The selection operator selects the fittest individuals of the current population to serve as parents of the next generation. The crossover operation chooses randomly a pair of individuals and exchanges some part of the information. The mutation operator takes an individual randomly and alters it. As natural genetics, the probability of crossover is usually high, the population evolves iteratively (in the genetic algorithm terminology, through generation) in order to improve the fitness of its individuals The structure of genetic algorithm is a loop composed of a selection, followed by a sequence of crossovers and mutations. Probabilities of crossover and mutation are constants and fixed in the beginning. Finally, genetic algorithm is executed until some termination condition achieved, such as the number of iterations, execution time, execution time, result stability, etc.

3. FRAMEWORK FOR GENETIC ALGORITHM

3.1 String representation

The genetic representation of individuals is called genotype. The main criteria in selecting the string representation for the search node is that the new string generated from the application of genetic operator must represents legal search node for the problem. Every process is present and appears only once in the schedule. The string representation used in this paper is an array of n × m digits, where n is the number of processes and m shows the processor that the process is assigned to. Process index shows the order of execution on that processor.

3.2 Initial population

A genetic algorithm starts with a set of individuals called initial population. Most GA-Based algorithms generate initial population randomly. Here, each solution *i* is generated as follows: one of the unscheduled processes is randomly selected, and then assigned to one of the processors. The important point is the processors are selected circularly, it means that they are selected respectively form first to last and then come back to first. This operation is repeated until all of processes have been assigned. An initial population with size of *POPSIZE* is generated by repeating this method.

3.3 Fitness function

The fitness function is essentially the objective function for the problem. It provides a mean to evaluate the search nodes and also controls the reproduction process. For the process scheduling in distributed system problem, we can consider factor such as load-balancing, processors utilization etc. We take into account this objective in following equation. Fitness function of schedule T is

$$f(T) = \frac{Average\ utilization}{Scheduling\ length}$$

This equation shows that a fitter solution has less scheduling length and higher processor utilization.

3.4 Reproduction

The reproduction process is typically based on the fitness values of the strings. The principal is that string with higher fitness value will have higher chance of surviving to the next generation. We calculate fitness of the entire individual in population, picking the best individual of them and then form a group. We can make a slight modification to basic reproduction operation by always passing the best string in the current generation to the next generation. This modification will increase the performance of genetic algorithm.

3.5 Crossover

Crossover is generally used to exchange portions between strings. The crossover operation picks top two strings from best group. Random task Ti from one of these two strings picks and put on the new string with care of precedence relation. If Ti task at same position then Ti is coping on the same place for new string. If we randomly choose two same parents (A=B) and if one of parents e.g. A the best string we use operator mutation on the second B string. Otherwise we mutate the first set and child generate randomly.

3.6 Mutation

Crossover operation of genetic algorithms (GAs) cannot generate quite different offspring from their parents because the acquired information is used to crossover the chromosomes. An alternate operator, mutation, can search new areas in contrast to the crossover.Mutation is used to change the genes in a chromosome. Mutation replaces the value of a gene with a new value from defined domain for that gene.

3.7 Termination condition

We can apply multiple choices for termination condition: maximum number of generation, equal fitness for fittest selected chromosomes in respective iterations.

```
Genetic Algorithm
{
Randomly create an initial population
Assign a fitness value to each individual
Form Group of best individual
WHILE NOT termination criteria DO
{
Assign a priority value to the individual in group
Choose two best individual from the group;
Crossover surviving individuals;
Mutation child;
Recorded best individual in group and eliminate worst one
in group.
}
}
```

Genetic Algorithm Steps

4 . Conclusion

The paper makes an analysis of existing genetic algorithm and their parameters.Scheduling in distributed operating systems has a significant role in overall system performance and throughput. This algorithm considers multi objectives in its solution evaluation and solves the scheduling problem in a way that simultaneously minimizes scheduling length and communication cost, and maximizes average processor utilization and load-balance. Most of existing approaches tend to focus on one of the objectives. Experimental results prove that our proposed algorithm tend to focus on more objective simultaneously and optimize them.

References

[1] Dr Apurva Shah And Vinay Hansora "A Modified Genetic Algorithm For Process Scheduling In Distributed System" IJCA Special Issue On "Artificial Intelligence Techniques - Novel Approaches & Practical Applications"AIT, 2011

[2] M. Nikravan And M.H.Kashani " A Genetic Algorithm For Process Scheduling In Distributed Operating Systems Considering Load Balancing", European Conference On Modelling And Simulation.

[3] C.C.Shen, & W.H.Tsai, "A Graph Matching Approach To Optimal Task Assignment In Distributed Computing Using A Minimax Criterion", IEEE Trans. On Computers, 34(3), 1985, 197-203.

[4] G.L.Park, "Performance Evaluation Of A List Scheduling Algorithm In Distributed Memory Multiprocessor Systems", International Journal Of Future Generation Computer Systems 20, 2004, 249-256

[5] L.M.Schmitt, "Fundamental Study Theory Of Genetic Algorithms" , International Journal Of Modelling And Simulation Theoretical Computer Science 259, 2001, 1 – 61.

[6] W.Yao, J.Yao, & B.Li, "Main Sequences Genetic Scheduling For Multiprocessor Systems Using Task Duplication", International Journal Of Microprocessors And Microsystems, 28, 2004, 85-94.

[7] A.T. Haghighat, K. Faez, M. Dehghan, A. Mowlaei, & Y. Ghahremani, "Multicast Routing With Multiple Constraints In High-Speed networks based on genetic algorithms" , *In ICCC 2002 Conf.*, India, 2002, 243–249.

Parameter Optimization of Tube Hydroforming

Edina Karabegović[1], Miran Brezočnik[2]

[1] University of Bihać, Faculty of Technical Engineering of Bihać,
[2] University of Maribor, Faculty of Mechanical Engineering, Maribor

Abstract:

Tube hydroforming is mostly applied in automotive industry. In this respect, necessity for the procedure improvement of fluid forming is constant. One of the reasons of its improvement is the procedure performance in optimal conditions. The process parameters have the direct influence on quality and optimal of forming procedure. This paper provides an example of the fluid pressure optimization in T-shape tube hydroforming. Three types of material have been analysed, with three wall thickness and three course levels of axial printers. For the optimization, the evolutional method with applied genetic algorithm (GA) was utilized. The application of GA is significant in solving of many problems in engineering practice. The simplicity and adaptability of the genetic algorithm to the engineering problem results with the increasing volume of applications in a research work. In this paper we investigated interactions of the internal parameters of the T tube hydroforming process, towards achieving the GA model for the optimal internal pressure, necessary for hydroforming.

Keywords: Hydroforming, tube, modelling, optimization, parameter,genetic algorithm, T-shape,

1. Introduction

A significant development of unconventional procedure processing has brought up by the market demands for the rapid part changes in automotive industry, whereas economic and market demands couldn't be satisfied with the conventional method of automotive part manufacturing. The demand that lays for nowadays products is a very high level of quality which implies product manufacturing, with lesser number of constituent elements and lesser quantity of materials (thin-walled elements). The automotive manufacturers are faced with this problem, where apart from functional, the securing and ergonomic product demands lays. It justifies the increasing application of hydroforming in automotive and aviation industry. Hydroforming in that area has showed as satisfying, as besides technological process advantages (possibilities for amplification, narrowing, tube calibration), their advantage is significant in different material application, large dimension parts and complex shape. One of the procedure improving methods is the parameter optimization of the tube hydroforming process. From previous researches, it is observed that the parameter process analyses of element tube hydroforming have been carried out by many other researchers. Investigations were related on modelling and the parameter optimization and considered as the most effective on element tube forming. For modelling and optimization, there are applied analytic methods, numerical methods, evolutional methods, artificial neural networks, finite neural method and alike. A group of researchers, Giuseppe Ingaro and the others (2009) have been investigating and optimizing the internal pressure functioning and printer-counter in Y string hydroforming process in the surging zone, considering the tube wall thickness change. The applied methods are numerical and experimental. Researchers Nader Abedrabo and the others (2009) are analysing and optimizing the internal fluid pressure and axial shift in symmetrical tube amplification, using the finite element method and genetic algorithm. Researchers Zhan Yong and the others (2009) apply FE simulation and GS method for the internal pressure optimization and axial shift in element tube hydroforming. The results of all investigations have influence on the improvement and development of the hydroforming process. This paper provides an example of the internal fluid pressure optimisation in T-shape tube hydroforming, with three wall thickness levels (s=1-2-3mm), three materials ($\sigma_{0.2}$= 164-290-412 N/mm^2) and three axial shifts (10-15-20mm). The genetic algorithm method has been applied (1-3,10 - 15).

2. Tube Hydroforming Process

Almost all recent technological processes are based on utilization of the unconventional procedure of metal processing with plastic forming. It is significant that about 10% of manufacturing in automotive industry includes fluid plastic forming of a tube. The demands are corresponding to the techno-economic justification and hydroforming process utilization. Some of the reasons of the utilization: (4, 7, 8, 10):satisfactory quality of the obtained tube elements,a single tube part fabrication, the welding utilization reducing, possibility of different shape tube forming and wall thickness,abrication possibility of the high repetition percentage products.

Some of the examples of element tube forming are given in Figure 1.

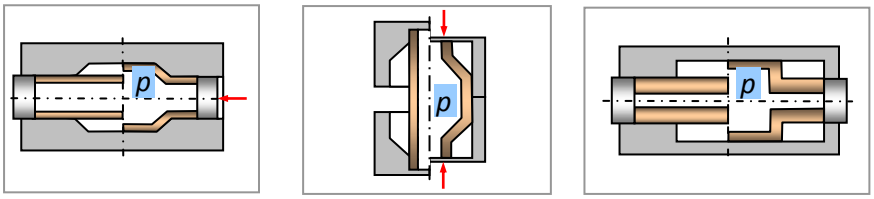


Figure 1: Tube elements obtained by hydroforming

2. 1. T tube Hydroforming

In T tube forming, without drain narrowing, the deformation of the drain terminate in the initial forming stage, while rising of the drain height is obtained by plastic tube deformation, i.e. material inflow from the tube to the drain.

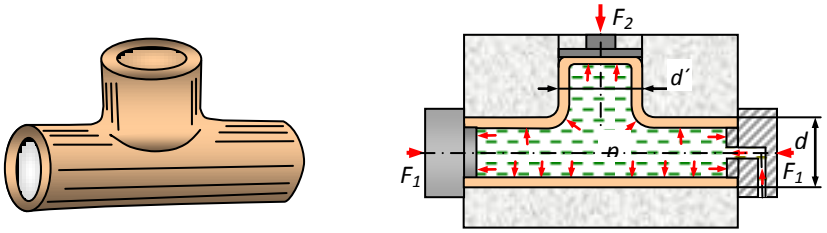
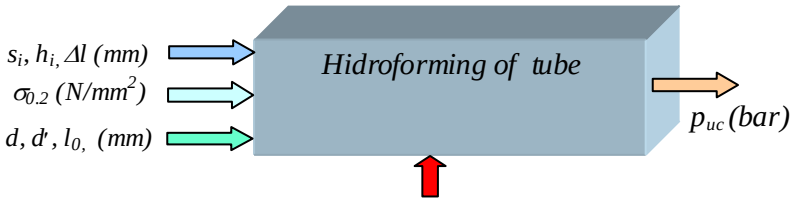


Figure 2: T –shape hydroforming tube

A successful tube forming doesn't depend only on dimensions and contact stress distribution, i.e. the stress-strain states in certain zones of plastic forming part. It also depends on the requisite forces for the operation achievement and the other parameters like: material characteristics, tribological conditions and geometry setup. With defining of the block scheme in process of the input/output sizes, the mutual dependence of the influential parameters of the hydroforming process is recognized (4, 7, 8. 10).



Legend:

- pressure of the cylinder, $p_c\ (N/mm^2)$
- stress of the flow, $\sigma_{0.2}\ (N/mm^2)$
- diameter of the tube, $d_i\ (mm)$
- diameter of the bulge, $d'\ (mm)$
- height of the bulge, $h_1 (mm)$
- wall thickness, $s_i\ (mm)$
- the shift, $\Delta l\ (mm)$
- length of the tube, $l_0\ (mm)$

C = const.

- a device(machine tool)
- a tool,
- the fluid,
- the outer diameter d_v=*const*.
- an abrasion.

Figure 3: Inner/outer scheme block of tube hydroforming process sizes

The internal fluid tube pressure (p_{uc}) via T-shape tube is obtained and represents a cost function for modelling and optimisation of parameter forming process.

2. 2. Tube hydroforming device

Researches of tube hydroforming have been done on the tube hydroforming machine in the laboratory of the Faculty of Technical Engineering of Bihać, Figure 4.

Figure 4: Experimental analysis device of tube hydroforming process

The working pressure, obtained during the forming process has reached the valuation to 3200 (bar).

2. 3. Measuring equipment for the internal fluid pressure

The measuring amplifier device, Spider 8 from Hottinger Saldwin Messtechnick (HBM), Germany (10), has been utilized for the internal fluid pressure measurement of T-shape tube hydroforming.

Figure 5: Measuring amplifier device, Spider 8, P3MB – Fluid pressure sensor in the tube

The device contains eight independent measuring channels, where various sensors can be connected, based on principle of the electronic dimension change. The fluid pressure sensor P3MB, differs with small dimensions, and has been fitted at the fluid output of the multiplier (10).

2.4. Hydroforming tubes

For the analysis there has been used tubes made of three types of material (aluminium alloy, brass and steel), with the outer diameter d_v=20 (*mm*) and length *lo*= 80 (*mm*).

Figure 6: Initial tube shapes for the aluminium alloy, brass and steel

In Table 1 there are the basic material characteristics of which the initial tube shapes have been made.

Table 1. Properties of Materials

TYPES OF MATERIALS								
Aluminium alloy AlMgSi0.5			*Brass Cu63Zn*			*Steel Ck10 (soft annealed)*		
Mechanical Properties		*Chemical properties*	*Mechanical properties*		*Chemical properties*	*Mechanical Properties*		*Chemical properties*
$\sigma_{0.2}$	σ_m	*Al 98.5%*	$\sigma_{0.2}$	σ_m	*Cu 63%*	$\sigma_{0.2}$	σ_m	*C=0.10%*
164 N/mm²	*175 N/mm²*	*Mg≤1 Si=0.5*	*164 N/mm²*	*164 N/mm²*	*Zn 37%*	*164 N/mm²*	*164 N/mm²*	*Si≤0.10% Mn≤0.30% P≤0.035% S≤0.035%*

2. 5. Genetic algorithm modelling

Modelling represents a mathematical law description of the parameter process change in certain time and space. In engineering practice problem solving, the genetic algorithm, with its simplicity and repeatability has given solutions for better problem solving. The genetic algorithm (GA) is adjusted to the evolutional organism previsions, and as a modelling shape is necessary for solution improving of cost functions. The classic GA works with the coded variables. The GA seeks the solution by the dot population (not only by one dot). The linear model has been utilized for the modelling of the first line (general shape) for the triple-factor interaction:

$$y = b_0 + b_1 s_0 + b_2 \sigma_{0.2} + b_3 \Delta l + b_4 s_0 \sigma_{0.2} + b_5 \sigma_{0.2} \Delta l + b_6 s_0 \Delta l + b_7 s_0 \sigma_{0.2} \Delta l \quad (1)$$

The modelling aim is to optimize coefficients b_0, b1, ...,b7, in the equation 1.

A single organism coding has been derived (9):

$$\left(\overbrace{(b_0, b_1, \ldots b_7)}^{O_1}, \overbrace{(b_0, b_1, \ldots b_7)}^{O_2}, \ldots, \overbrace{b_0, b_1, \ldots b_7}^{O_m} \right) \quad (2)$$

The population P(t) in a generating time t (Figure 5), is an organism set (O_1, O_2,…., O_m), and the single organism is the dot (solution) in a multidimensional area. A single gene is a coordinate of the dot (3, 5, 6, 9). The initial stochastic generated population in our concrete example consists of the m organisms. Each organism is made of eight genes (see equation 1):

$$O_i = b_0, b_1, \ldots, b_7, \quad (3)$$

Where (i) is a single organism index, i.e. the mathematical model considering the equation 1. The coefficient evolution of the model from equation 1, i.e. overall mode, is carried out in a way given in the pseudocode, Figure 7.

Evolutionary algorithm
$t := 0$
Generate the initial random population of the solution $P_{(t)}$
evaluate $P_{(t)}$
repet
edit $P_{(t)} \rightarrow P_{(t+1)}$
evaluate $P_{(t+1)}$
$t := t + 1$
Till the criteria for the evaluation abruption is not achieved

Figure 7: Pseudo code evolutionary algorithm

In the evaluation phase of population $P_{(t)}$, in the generation time t, an absolute deviation $D_{(i,t)}$ of the individual model (organism) *i* for all measurements has to be calculated:

$$D_{(i,t)} = \sum_{j=1}^{n} \left| E_{(j)} - P_{(i,j)} \right| \quad (4)$$

Where:

- $E_{(j)}$-experimental values for j measuring
- $P_{(i,j)}$ prediction value of individual model *i*, for j measuring
- n- a maximum number of measurements

The equation (4) is a raw adaptive measure. The aim of the optimization by GA is to achieve the equation solution (4) with the lack of deviation. But the smallest given absolute solution value doesn't indicate that the smallest percentage model deviation is achieved. Therefore, the average percentage deviation, for all measurements and for individual model *i* was introduced:

$$\Delta(i) = \frac{D(_{i,t})}{\left| E_{(f)} \right| n} \cdot 100\% \quad (5)$$

The equation (5) was not utilized as an adequate measure for the evaluation population, but is utilized in order to obtain the best organism in the population after the settled optimisation.

3. The Measurement Results

3. 1. Experimental results

The table shows the values of obtained experimental results of the internal fluid pressure (p_{uc}) in the tube, upon T-string hydroforming.

Table 2: Experimental measurement results of the internal fluid pressure

Number of exp. Nj	Input values of parameters			Output values
	s_0	$\sigma_{0.2}$	Δl	p_{uc}
	mm	N/mm^2	mm	bar
1	1	164	10	432
2	3	164	10	1220
3	1	412	10	1120
4	3	412	10	2940
5	1	164	20	477
6	3	164	20	1298
7	1	412	20	1210
8	3	412	20	3090
9	2	290	15	1603.75

The diagram of experimental values for the fluid pressure of T tube forming, made of steel, wall thickness s_0=2 mm, obtained by measuring on Spider 8 device.

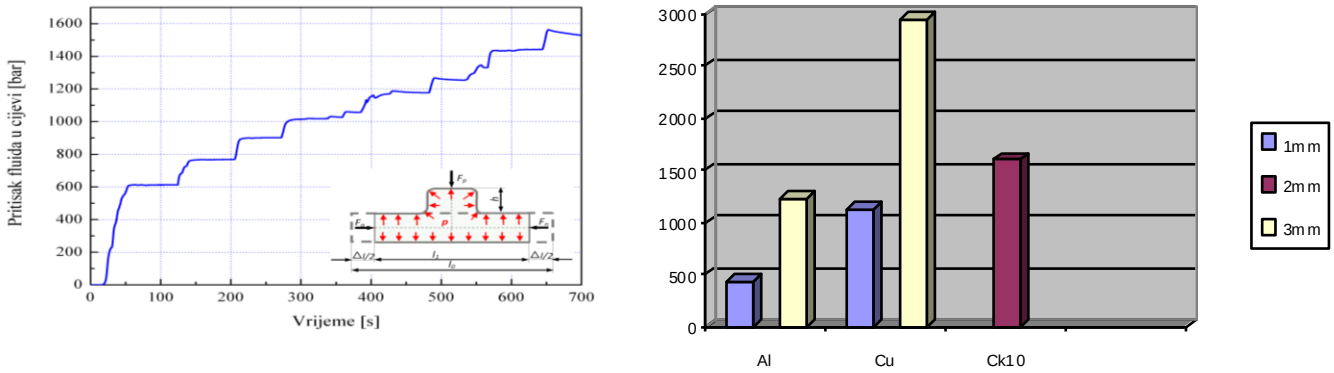


Figure 8: Fluid pressure in the tube by experimental measurement,fluid pressure in the tube

The fluid pressure in the tube for three types of material has been given in Figure 8.

3. 2. GA model of fluid pressure in the tube

The analysis of the experimental values of the fluid pressure in the tube, with the GA application has been obtained. The population carried the 100 organisms, and the number of the system motions for GA was 10. The maximum number of the generations was 1000.

The best mathematical model obtained by GA for the fluid pressure in the tube has a shape:

$$p_{uc} = -81,1683 + 43,9205 s_0 + 0,559208 \sigma_{0.2} + 0,59908 \Delta l + 2,02987 s_0 \sigma_{0.2} + 0,0132628 \sigma_{0.2} \Delta l + 0,860965 s_0 \Delta l + 0,00514024 s_0 \sigma_{0.2} \Delta l \quad (6)$$

The figure 9 shows the best evolution solution (i.e. the best mathematical model for the fluid pressure in the tube), over the generations. The solution improvement in the initial phase was rapidly, and after the generation of 100, relatively small solution improving has been perceived.

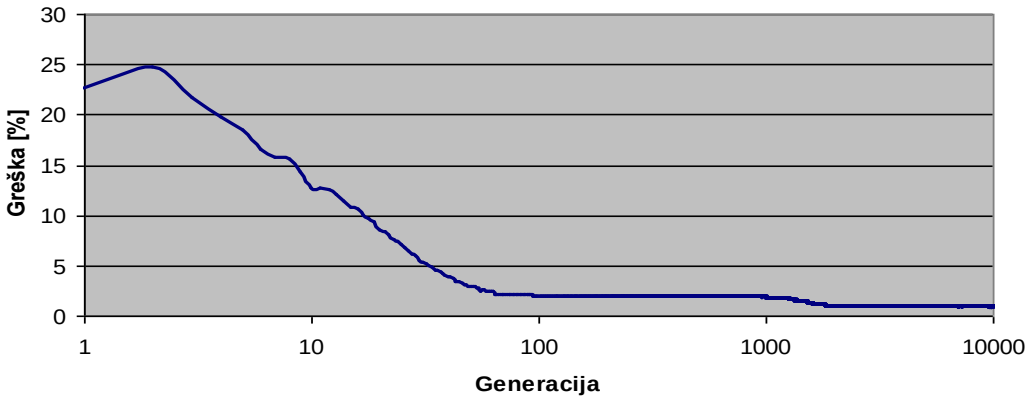


Figure 9: Improvement of the best generation solutions

3. 3. Comparative results of the experimental values by the GA model

Experimental researches of the T tube hydroforming process parameters, the values for the defined outer parameter are obtained, i.e. tube fluid pressure (p_{uc}). These values are compared with the values (p_{uc}) obtained by the GA model, Table 3.

Table 3: Comparative data of the experimental and obtained GA model values

Number of exp. Nj	*Input values of parameters*			*Output values*	*GA model*
	s_0	$\sigma_{2,0}$	Δl	p_{uc}	p_{uc}
	mm	N/mm^2	*mm*	*bar*	*bar*
1	1	164	10	432	432,1424
2	3	164	10	1220	1219,8601
3	1	412	10	1120	1119,8733
4	3	412	10	2940	2939,9021
5	1	164	20	477	476,9239
6	3	164	20	1298	1298,7208
7	1	412	20	1210	1210,2943
8	3	412	20	3090	3089,8979
9	2	290	15	1603.750	1483,3960

The mean percentage deviation within experimental and prediction values, obtained by the GA model is Δi = 0.85148%.

4. Conclusion

Plastic forming with the fluid application has been known since the last century, and researches in this area are significant for the process improvement. With the process parameter optimization of the plastic forming, the techno-economic justification of the process is achieved.With the mathematical modelling and optimization of the experimental values for the fluid pressure in the tube, with the applied genetic algorithm (GA), the mathematical model equation is obtained, with the percentage deviation Δi = 0.85148%. The mathematical model for the fluid pressure in the tube refers to the optimal work area, describing the solution regression for the derived experiment. The solutions in this area ensure the forming efficacy, with the optimal internal pressure for the generated dimension lines: $s_0 = 1mm \div 3mm$, $\sigma_{0.2} = 164 \div 412 \;\; N/mm^2$ i $\Delta l = 10mm \div 20mm$.

References

[1] Jurković, M.: Matematičko modeliranje inženjerskih procesa i sistema, Mašinski fakultet Bihać, 1999.

[2] Jurković, M., Tufekčić, Dž.: Tehnološki procesi projektiranje i modeliranje, Univerzitet u Tuzli, Mašinski fakultet, Tuzla, 2000.

[3] Brezočnik, M.: Uporaba genetskega programiranja v inteligentnih proizvodnih sistema, Fakultet za strojništvo, Maribor, Slovenija, Tiskarna Tehniških Fakultet, Maribor, 2000.

[4] Karabegović, E., Jurković, M.: The Analytical Models of hydraulic Forming and Discussion, 8thInternational Research/Expert Conference, Trends in theDevelopment of Machinery and Associated Technology, TMT 2004, Neum, 2004., p.119-122.

[5] Brezočnik, M. and Kovačić, M.: Integrated Genetic Programming and Genetic Algorithm Approach to Predict Surface Roughness, Materials and Manufacturing Processes Vol.18. N_0. 3, p. 473-489, 2003.

[6] Brezočnik, M., Gusel, L. and Kovačić, M.: Comparison Between Genetic Algorithm and Genetic Programming Approach for Modeling the Stress Distribution, Materials and Manufacturing Processes Vol.20. N_0. 3, p. 497-508, 2005.

[7] Karabegović, E., Jurković, M., Jurković, Z: Theoretical Analysis of Hidroforming of Tubes, Manufacturing Engineering, vol. 5, N_0. 3/2007, p.58-62.

[8] Jurković, M., Karabegović, E., Jurković, Z., Mahmić, M.: Theoretical Analysis of the Tube Hydroforming Process Parameters and a Suggestion for Experimenta, 11thInternational Research/Expert Conference, Trends in theDevelopment of Machinery and Associated Technology, TMT 2007, Hammamet, Tunisia, 2007. p.83-86.

[9] Brezočnik, M., Balić, J.: System for Discovering and Optimizing of Mathematical Moldels Using GeneticProgramming and Genetic Algorithms, 8th International DAAAM Symposium: Intelligent Manufacturing & Automation, Dubrovnik, Croatia, 1997, p. 37-38.

[10] Karabegović, E.: Modeliranje i optimizacija parametara procesa hidrooblikovanja tankostijenih cijevnih elemenata, Doktorska disertacija, Tehnički fakultet Bihać, jul 2009.

[11] Altan, T. and Jirathearanat S.: Successful tube hydroforming:Watching parameters, accurately simulating the process yield good results, TPJ-The Tube&Pipe Journal, 2001.

[12] Rosa Di Lorenzo, Giuseppe Ingarao, Francisco Chinesta: Integration of gradient based and response surface methods to develop a cascade optimisation strategy for Y-shaped tube hydroforming process design, Advances in Engineering Software, Volume 41 , Issue 2 (February 2010), ISSN:0965-9978, p. 336-348.

[13] Giuseppe Ingarao, Rosa Di Lorenzo, Fabrizio Micari: Internal pressure and counterpunch action design in Y-shaped tube hydroforming processes: A multi-objective optimisation approach, Computers and Structures, Volume 87 , Issue 9-10 (May 2009), ISSN:0045-7949, p. 591-602.

[14] Zhang Yong i dr.: Optimization for Loading Paths of Tube Hydroforming Using a Hybrid Method, Materials and manufacturing processes, ISSN 1042-6914, 2009, vol. 24, n°4-6, p. 700-708.

[15] Nader Abedrabbo i dr.: Optimization of a Tube Hydroforming Process,Red Cedar Technology, AB 2027, Rev.04.09.

Development Of Embedded Ethernet Drivers For Arm9

[1,]T.Satyanarayna [2,] S.Latha(Associate Proffesor)

, [1,2,]Dept of Electronics & Communication Engineering, Aurora Technological & Research Institute,
. Jawaharlal Nehru Technological University.

Abstract

with the widely application of ARM technique, building the embedded operating system based on ARM processor has been a hot point of research. In this paper, the design of network device driver in Linux operating system based on ARM920T processor is implemented on the S3C2410- S development platform made in Beijing universal pioneering technology. Focused on discussing implementation principle of embedded Linux network drivers and detailed to analysis of the frame structure of the program code.

Keywords-ARM9 processor; embedded linux; network device driver; CS8900A

1. Introduction

Now the people more and more like open-source software. As a powerful and stable open-source operating system, Linux is acclaimed by the thousands of computer expert and amateur. In the embedded field, Linux can be cured in dozens of megabytes of memory chips or SCM after small cutting. So that it can be used in a specific context of embedded Linux. Strong network support functions of Linux achieved support for multiple protocols including TCP / IP, and it meets the demand for embedded systems application networking for the 21st century. Therefore, when developing and debugging embedded systems, network interface almost become indispensable module.

2. INTRODUCTION OF LINUX NETWORK DEVICE DRIVER

Linux network device driver is an important part of Linux network application. The whole Linux Network driver follows the common interface. For each network interface, it uses a device data structure. Generally, the network device is a physical device, such as Ethernet card.Network driver must solve two problems: first, not all network device driver based on Linux kernel have control equipment; second, Ethernet device in the system is always called / dev/eth0, dev/eth1 etc., regardless of the underlying device driver. When initialization routines of each network device are called, the driver will return a status, whichindicating whether it is orientation to an instance of the driven controller. If the driver does not find any device, then the entries pointed to the device lists by the 'dev_base' will be deleted. If the driver can find a device, then the rest of the device data structure is filled by this device information and the address of support function in network device driver.

3. Architecture Of Linux Network Device Driver

Shown in Fig. 1, the architecture of Linux network driver can be divided into four levels. Linux kernel source code provided the network device interface and the code above level. Therefore the main work which transplanting specific network hardware drivers is to complete the corresponding code of the device driver function layer. According to the specific bottom hardware features, Structure variable of network device interface 'struct net_device' type is defined and corresponding operation function and interrupt handling program are implemented.

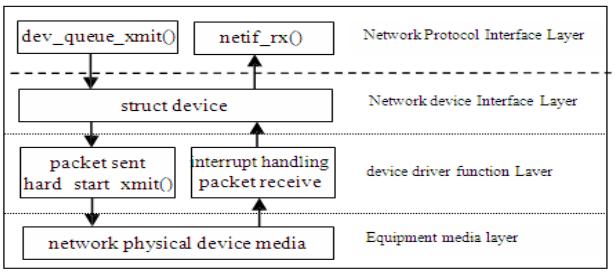


Figure 1.Architecture of Linux network driver

4. Ethernet Controller Chip Cs8900a Cs8900a

is a 16-bit Ethernet controller produced byCIRRUS LOGIC, embedded on-chip RAM, 10VASE-T transceiver filter and direct ISA bus interface. The salient feature of the chip is flexible to use, and it can dynamically adjust according to needs for its physical layer interface, data transfer mode and work mode, and it can adapt to different application environment through setting internal register.

CS8900A can be operated in memory mode and I / O

mode. When CS8900A is configured to Memory Mode operation, its internal registers and frame buffer are mappedto a serial 4KB host memory block, the host can directly access CS8900A's internal registers and frame buffer through the block.

5. Design And TheImplementation Principle Of Network Driver

Linux network system can complete data transfer between all levels through the socket buffer sk_buff, data structure sk_buff is each protocol data processing object. sk_buff is the media of exchange data between driver and network. When the driver sends data to the network, data source and data length must be obtained. Data must be saved in sk_buff also after the driver received data from thenetwork. Thus upper layer protocol can process it.For the actual development of Ethernet driver, corresponding template program in the kernel source tree can be consulted, focused on understanding the implementation principle of network driver and program structural framework. Then the code is rewritten to develop specific hardware, and achieve the appropriate operation function. Through transplanting and preparing the embedded CS8900A network card driver on embedded development board (S3C2410 processor) to show the implementation principles of network driver.

A. Initialization Function

Initialization of network equipment is completed mainly by initialization function which is referred by init function pointer in device data structure. After the kernel load the network driver module, initialization process will be called. First it is necessary to detect whether network physical device exist, which is completed by detecting the physical device hardware characteristic. Then the resource equipment needed is configured, such as interrupts. Next the device'device' data structure is constructed. The relevant variable in the device is initialized by detected data, and finally thedevice is registered to the Linux kernel and applies memory space.In this instance, the initialization function is "_init cs8900a_s3c2410（void）".
In the network device driver, the device 'device' data structure is dev_cs89x0, which is defined as follows.
#ifdef MODULE static struct net_device dev_cs89x0={
'''',
0, 0, 0, 0,0, 00, 0, 0, NULL,NULL};The function 'cs89x0_probe' detect the existence of the network physical device, but device initialization is completed by twofunctions 'cs89x0_probe' and'cs89x0_probe1' together. In 'cs89x0_probe1', the function pointers of 'device' data structure are filled.After completion of filling pointer, to register by'register_netdev' (struct net_device * dev) function. Two function

register_netdev' and 'unregister_netdev' are defined in file 'net_init.c '.Since there is 'init' function, there should also be'cleanup' function, because they are essential function of each driver. The 'cleanup' function run when module is unloaded, which complete mainly the work of resource release. Such as cancel device registration, free memory, release the port, etc. All in all, it is some action contrary to init.The function of Cancellation of network device registration is 'unregister_netdevice' defined in file / net / core / dev.c, this function is called in 'unregister_netdev'.

B. Open Function

When system response 'ifconfig' command, a network interface will be opened (closed). The 'ifconfig' command given addressto interface by callingioctl' (SIOCSIFADDR). Response of SIOCSIFADDR is accomplished by the kernel, and device-independent. Then, the 'ifconfig' command set IFF_UP bit of dev-> flag to open the device by calling 'ioctl' (SIOCSIFFLAGS). The device's open method is called through the above called. In the network device driver, Open function is called when network device is activated, that device status becomes from down to up. So a lot of initialization work can be done here. In open function, operation on the register uses two
functions: 'readreg' and 'writereg'. 'readreg' function is used to read the register contents, 'writereg' function is used to write registers, the code is as follows:

```
inline int readreg(struct net_ device *dev,int portno){
outw(portno,dev->base_addr+ADD_PORT);
return inw(dev->base_addr+DATA_PORT);
}inline void writereg(struct net_ device
*dev,int portno,int value){ outw(portno,dev->base_addr+ADD_PORT); outw(value,dev-
>base_addr+DATA_PORT);
}
```

C. Close Function

Close function (net_close) releases resources to reduce system burden. It is called when the device status becomes from up to down. In addition, if the driver is loaded as a module, the macro MOD_DEC_USE_COUNT need to call in close also, to reduce the frequency of equipment cited in order to uninstall the driver.

D. Send Function

Sending and receiving of data packets are two key processes achieving Linux network driver, good or bad of the two processes affected directly the driver's overall running quality. First of all, when the network device driver is loaded, device is initialized by init function pointer in the device domain calling network device initialization function. If the operation is successful, device is opened by open function pointer in the device domain calling network device open function. Next, hardware packet header information is set up by building hardware packet header function pointer hard_header in the device domain. Finally, the packet sent is completed through protocol interface layer function dev_queue_xmit calling function pointer hard_start_xmit in the device domain. If sent successfully, sk_buff is released in hard_start_xmit method, and return 0 (send success). If the device can not be process temporarily, such as the hardwareis busy, and then return 1. At this time if the dev-> tbusy is set to non-zero, the system regards that the hardware is busy, it will not re-send until dev-> tbusy is set to 0. tbusy's setting zero task is completed by interrupt generally. Hardware interrupt at the end of sent, at this time tbusy can be set to 0, then call mark_bh () to notify system to re-sent. In the case of sending unsuccessful, dev-> tbusy can not be set to non- zero, this system will try to resent continuously. If hard_start_xmit is not sent successful, then sk_buff cannot be release. The data which is sent in Sk_buff already contains the frame head of hardware requirement. Hardware frame head need not be fill in send method, data can be sent directly to the hardware. sk_buff is locked, and it is assured that other programs will not access it.

E. Receive Function

Receive function is different from receive data, network interface does not provide receive function pointer similar to net_receive_packet, because that network device receive data is achieved through interrupts. Upon receiving the information, an interrupt is generated, the driver will apply a sk_buff (skb) in interrupt handler, and the data read from hardware is placed to applied buffer. Next, some information is filled in sk_buff. skb-> dev = dev judges protocol type ofreceived frame, and fills in the skb-> protocol (multi- protocol support). Pointer skb-> mac.raw point to the hardware data and then discard the hardware frame head (skb_pull). Also skb-> pkt_type is set, indicating the type of data link layer. If the data packet is obtained, then net_rx() is implemented, net_rx () is a subroutines of data reception, it is called by interrupt service routine. Last the data is sent to the protocol layer by calling netif_rx (). netif_rx () put data into the processing queue and then return, the real processing will be implemented after interrupted return. Thus interrupt time is reduced. After netif_rx () is called, the driver cannot be saved in data buffer skb. In the protocol layer, flows control of receiving data packets is divided into two levels: first, netif_rx () function is limited the number of data frames from the physical layer to protocol layer. Secondly, each socket has a queue, which limits the number of data frames from the protocol layer to the socket layer. In transmission, the dev-> tx_queue_len
parameter in driver limits the length of the queue.

F. Interrupt Hhandler

In the open function, the interrupt is applied, the interrupt handler is net_interrupt. The preparation of writing this function is to understand the interrupt process of network control chip. For the CS8900A chip, this process can explain the flow chart using fig. 2.

First kernel need to read ISQ (Interrupt Status Queue)
value, ISQ event has the following 5 types:

```
#define ISQ_RECEIVER_EVENT          0x04
#define ISQ_TRANSMITTER_EVENT       0x08
#define ISQ_BUFFER_EVENT            0x0c
#define ISQ_RX_MISS_EVENT           0x10
#define ISQ_TX_COL_EVENT            0x12
```

After receiving a data packet (RxEvent), to deliver
function net_rx() to process. Interrupt handler parameter is set as follows:

```
Static void net_interrupt(int irq， void *dev_id， structpt_regs *regs
```

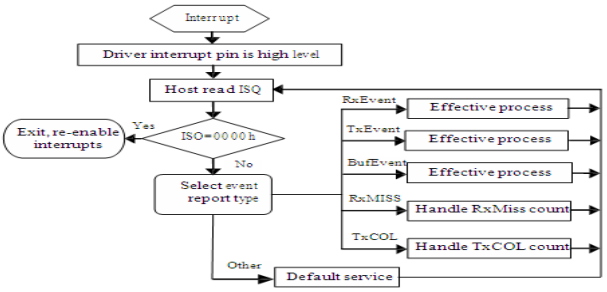


Figure 2.Interrupt handling flowchart

6. DRIVER TEST

Driver is only provided interface between kernel and user. To use the device, not only to install the driver, but also must write an application to use it.

After the network driver is designed, the kernel module can be compiled, and the custom kernel module as part of the system source code is compiled a new system.

A. Configuring The Kernel

In Linux2.6 kernel, want to compile the kernel module, first it is necessary to configure and construct properly kernel source tree in / usr / src, that is, to extract the need kernel source code to / usr / src /, and to use the command 'make menuconfig'or 'make gconfig' configure the kernel under the main directory of kernel source (Here is / usr/src/linux-2.6.18.3), then using the 'make all' to compile the kernel completely.

B. Take CS8900A NIC driver for example, introduce how to compile network device driver into the kernel.

- To create a new directory cs8900a under the directory 'drivers' of system source code tree;
- To copy cs8900a.c and cs8900a.h to drivers/cs890a directory;
- To compile 'Makefile' file under drivers/cs890a directory;

```
#Makefile for CS8900A network Driver
obj-$(CONFIG_DRIVER_CS8900A) +=cs8900a.o
```

- To compile 'Kconfig' file under drivers/cs890a directory;

```
#Just for CS8900A network device
menu "CS8900A network device support"
config DRIVER_CS8900A tristate "CS8900A support" This is a network driver module for CS8900A.
endmenu
```

- To add a line in front of'endmenu' statement of 'Kconfig' file In the 'driver' directory source "drivers/cs8900a/Kconfig"

In main directory of kernel source code tree, "CS8900A network device support [y / n /?]" can be found in the option 'Device Drivers' through command 'make menuconfig' or 'make gconfig' , if select "y", the kernel will provide support for network driver.

To recompile the kernel can be obtain kernel of supported CS8900A card, then download the kernel to the development board. By configuring network parameters, the behavior of the network card driver can be test.

7. CONCLUSION

In recent years, Internet has a rapid development. Remote monitoring and remote maintenance become very easy after embedded system access internet, so the network of embedded systems is very important. Embedded system achieved internet access; its prerequisite is that system software has TCP/IP protocol support. Fortunately, Linux kernel provides support for multiple protocols including TCP/IP. This paper takes network chip CS8900A for example, and introduces the key process of implementing network driver.

8. Acknowledgment

This research was supported by the Open Project Program of Key Laboratory of Intelligent Manufacture of Hunan Province (Xiangtan University), China (No.2009IM03).

References

[1] Q.Sun, Developing Detail Explain of Embedded Linux Application, Beijing, Posts & Telecom Press, July 2006.
[2] T.Z.Sun and W.J.Yuan, embedded design and Linux driver development guide, Beijing, Publishing House of Electronics Industry, October 2009.
[3] M.liu, Embedded system interface design and Linux driver development, Beijing, Beihang University Press, May 2006.
[4] L.G.Zhou, Example of ARM Embedded Linux System Building and Driver Development, Beijing, BeihangUniversity Press, 2006.
[5] C.Qian, R.H.Xu and Q.R.Wang, "Device Driver Development Based on Linux Operating System", Control & Automation,2004,(09).
[6] Q.N.Cao, B.Zhao and K.Y.Meng, "Design and realization of embedded linux network communication system based on ARM9 platform ", Journal of Northwest University(Natural Science Edition), 2009,(1).
[7] J.Zhao,X.Q.Ding, "Development and Implementation Principle of Network Driver Based on Embedded Linux", Microcomputer Information, 2008,(17).
[8] F.J.Li and W.D.Jin, "Research and Implementation of Network Driver in Embedded Linux", Modern Electronic Technique, 2005(16)
[9] W.Su, "Design of Linux Network Device Driver" Financial Computer of Huanan, 2005(06).
[10] D.Cao and K.Wang, "Research of Network Device Driver based on Linux", Computer Knowledge and Technology, 2005(21).
[11] Q.Wu, SH.H.Zhou and ZH.X.Ma, "Development of Linux Network Driver Based on USB Device", Microcomputer Information, 2007,(02). V12-448

Temperature Control of Shell and Tube Heat Exchanger by Using Intelligent Controllers-Case Study

[1,] Mr.P.Sivakumar, [2,]Dr.D.Prabhakaran , [3,]Dr.T.Kannadasan

[1,]Chemical Engineering Department, M.Tech Scholar Coimbatore Institute of technology Anna university, chennai.
[2,] Chemical Engineering Department, Associate Professor Coimbatore Institute of technology Anna university, chennai.
[3,]Chemical Engineering Department, Professor and Head Coimbatore Institute of technology Anna university, chennai.

Abstract

Temperature control of the shell and tube heat exchanger is characteristics of nonlinear, time varying and time lag. Since the temperature control with conventional PID controller cannot meet a wide range of precision temperature control requirement, the temperature control system of the shell and tube heat exchanger by combining fuzzy and PID control methods was designed in this paper. The simulation and experiments are carried out; making a comparison with conventional PID control showing that fuzzy PID strategy can efficiently improve the performance of the shell and tube heat exchanger.

Keywords: Control algorithm, Fuzzy logic, PID Controller, Tuning

1. Introduction

In many industrial process and operations Heat exchanger is one of the simplest and important unit [1] for the transfer of thermal energy. There are different types of heat exchangers used in industries; the shell and tube heat exchanger system being most common. The main purpose of exchanger is to maintain specific temperature conditions, which is achieved by controlling the exit temperature of one of the fluids (mainly hot fluid) in response to variations of the operating conditions. The temperature control of heat exchanger is nonlinear, time varying and time delay system. For these situations, nonlinear control strategies can provide significant improvements over PID control [2], [12]. Control of temperature using PID controllers, compared to other methods, is more effective and economical. The heat exchangers need to respond to highly non linear features and work well under different operating points. In order to achieve a wide range of high accurate temperature, neuro-fuzzy control and PID control methods were combined. The main design is to assume neuro-fuzzy reasoning control methods according to different error 'e' and error change 'ec' to get self –tuning PID parameter based on conventional PID controller. The simulation of the controller was accomplished carrying out experiments in an actual heat exchanger system.

2. Temperature Control System

2.1 Principal of temperature control in shell and tube heat exchanger

The control of temperature in a shell-and-tube heat exchanger is demonstrated in figure 1, with cold water flowing on the tube side and steam on the shell side [5] where steam condenses and heats the water in the tubes. The controlled variable is the tube-side outlet temperature, and the manipulated variable is the steam flow rate on the shell-side.

$$M_c C_p (T_{in} - T_{out}) = M_s A \qquad (1)$$

where, M_c, M_s, C_p, T_{in} and T_{out} refer to cold water flow rate, steam flow rate, specific heat of water, inlet water temperature, and outlet water temperature respectively.

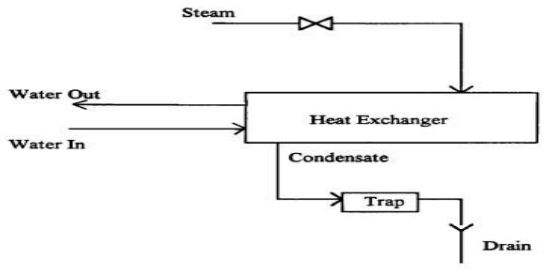


Fig 1: Shell and Tube heat exchanger

The dynamics of the process are complex because of various nonlinearities introduced into the system. The installed valve characteristic of the steam may not be linear [3]. Dead-time depends on the steam and water flow rates, the location and the method of installation of the temperature-measuring devices. To take into account the non-linearity and the dead-time, gain scheduling features and dead-time compensators have to be added. Also, the process is subjected to various external

disturbances such as pressure fluctuations in the steam header, disturbances in the water line, and changes in the inlet steam enthalpy and so on.

2.2 Mathematical Model of heat exchanger

The total heat in the heat exchanger system can be expressed as equation 2. [5].

$$Q_f = Q_s + \sum_{i=1}^{n} C_i \rho_i V_i dT_i \qquad (2)$$

where, Q_f, Q_s, C, ρ, V and dT refer to total system heat productivity, total system heat dissipating capacity, specific heat capacity, heat transfer medium density, volume, and temperature variation.

Total system heat dissipating capacity Q_s is given by equation 3.

$$Q_s = \sum_{i=1}^{n} k_i A_i (T_{in} - T_{out}) \qquad (3)$$

where, k_i and A_i refer to heat transfer coefficient, heat transfer area of the heat exchanger system.

The heat exchanger equations can be expressed as in equation 4. [12]

$$C_w \rho_w q_w (T_{wo} - T_{wi}) d\tau = C_f \rho_f q_f (T_{fo} - T_{fi}) d\tau \qquad (4)$$

where, the subscripts w and f refer to cold and hot water of the heat exchanger system. Therefore considering all above equations, the differential equation of the shell and tube heat exchanger is shown in equation 5.

$$\frac{dT}{d\tau} + FT = N(x - \tau) \qquad (5)$$

where,

$$F = \frac{k_i A_i}{C_o \rho_o V} \qquad (6)$$

The transfer function of controlled object can be derived from equation 5, it is described as the first-order with pure time delay, expressed as equation 7.

$$G(s) = \frac{K}{1 + Ts} e^{-\tau s} \qquad (7)$$

where, T, K and τ refer to time constant, system gain, and delay time.

3. Intelligent Controllers

3.1 PID Control of Shell and Tube Heat Exchanger

PID controller is the most popular controller used because it is easy to operate and very robust. Implementation of the latest PID controller is based on a digital design. These digital PID include many algorithms to improve their performance, such as anti wind-up, auto-tuning, adaptive, fuzzy fine-tuning and Neural Networks with the basic operations remaining the same. The performance specifications such as rise time, overshoot, settling time and error; steady state can be improved by tuning value of parameters K_p, K_i and K_d of the PID controller. The output

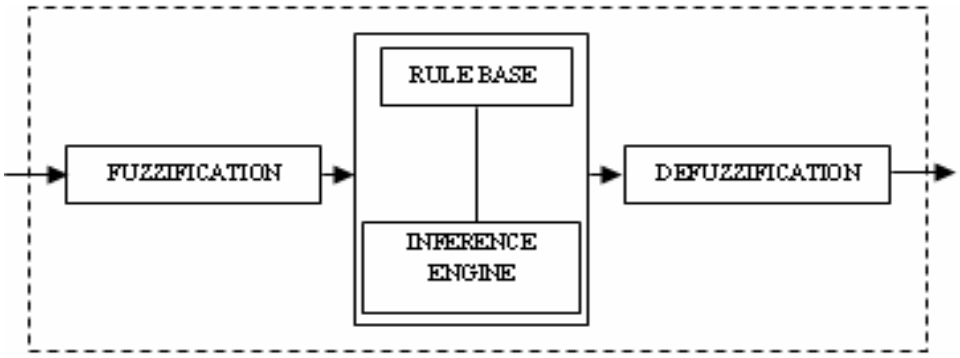


Fig 2: Main Parts of Fuzzy PID Control

is mathematically represented as equation 8 and 9.

$$y(t) = K_P[e(t) + T_d \frac{de(t)}{dt} + \frac{1}{T}\int_0^t e(t)dt] \quad (8)$$

$$y(t) = K_P e(t) + K_P T_d \frac{de(t)}{dt} + \frac{K_P}{T}\int_0^t e(t)dt] \quad (9)$$

3.2 Structure and Parts of Self tuning Fuzzy Controller

Fuzzy logic controller as shown in Figure 2 consists of main four parts fuzzification, rule base, inference engine and de-fuzzification [6], [8]. Fuzzy PID Self-tuning Control takes error "e" and Change-in-error "ec" as the input of Fuzzy PID controller. Using fuzzy control rules on-line, PID parameters "K_p", "K_i", "K_d" are amended, which constitute a self-tuning fuzzy PID controller, the principle of which is shown in Figure 3. The language variable values of error "e" and the error rate of change "ec" is (NB, NM, NS, ZO, PS, PM, PB) seven fuzzy values. And then setting up the suitable fuzzy control table for Kp, Ki, K_d three parameters tuning separately[8], [9].According to the fuzzy rules table, appropriate vague and ambiguous methods is to be selected to make dynamic tuning for K_p, K_i, K_d.

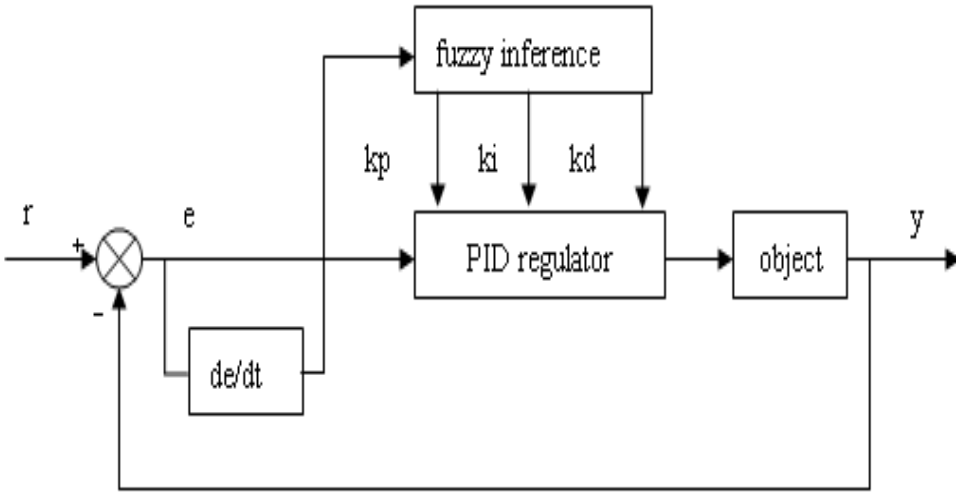


Fig 3: Structure of Fuzzy PID Control

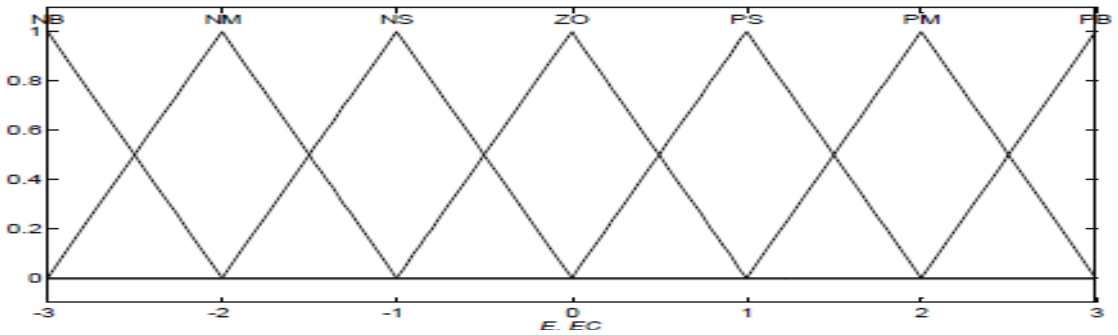


Fig 4: Input membership function for e and ec

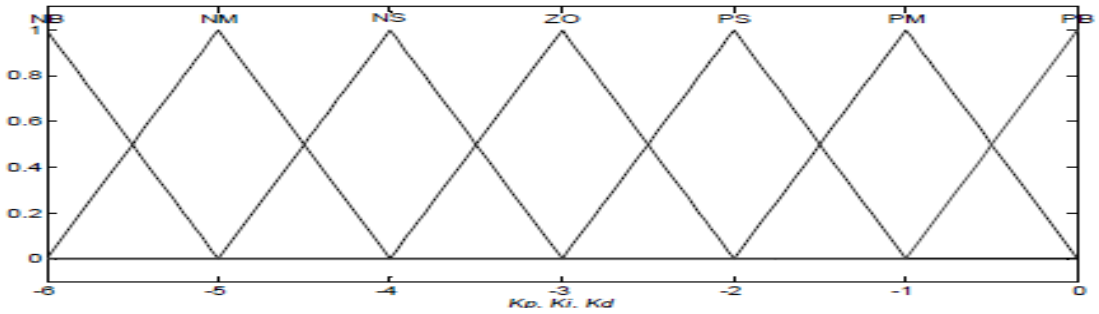


Fig 5: Output membership function for K_p, K_i, K_d

The fuzzy
controller takes two inputs (e, and error change ec) and three outputs (K_p, K_i, K_d). When the error is large, it is controlled according to the characteristics of PID control where the output value automatically closes to the given value. When the error becomes smaller to a certain extent, the fuzzy control takes effect. The input error, error change and output

membership functions use triangular functions, which are shown in figure 4 and 5.Larger K_p is chosen to speed up the system response speed. At the same time, in order to avoid the probable differential super-saturation, smaller K_i is chosen. In order to avoid large overshoot, the integral is limited by setting K_d is zero.

TABLE 1
THE FUZZY CONTROL RULE FOR K_p

ec / Δk_p / e	NB	NM	NS	ZO	PS	PM	PB
NB	PB	PB	PM	PM	PS	ZO	ZO
NM	PB	PB	PM	PS	PS	ZO	NS
NS	PM	PM	PM	PS	ZO	NS	NS
ZO	PM	PM	PS	ZO	NS	NM	NM
PS	PS	PS	ZO	NS	NS	NM	NM
PM	PS	ZO	NS	NM	NM	NM	NB
PB	ZO	ZO	NM	NM	NM	NB	NB

TABLE 2
THE FUZZY CONTROL RULE FOR K_i

ec / Δk_i / e	NB	NM	NS	ZO	PS	PM	PB
NB	NB	NB	NM	NM	NS	ZO	ZO
NM	NB	NB	NM	NS	NS	ZO	ZO
NS	NB	NM	NS	NS	ZO	PS	PS
ZO	NM	NM	NS	ZO	PS	PM	PM
PS	NM	NS	ZO	PS	PS	PM	PB
PM	ZO	ZO	PS	PS	PM	PB	PB
PB	ZO	ZO	PS	PM	PM	PB	PB

THE FUZZY CONTROL RULE FOR K_d

ec / Δk_d / e	NB	NM	NS	ZO	PS	PM	PB
NB	PS	NS	NB	NB	NB	NM	PS
NM	PS	NS	NB	NM	NM	NS	ZO
NS	ZO	NS	NM	NM	NS	NS	ZO
ZO	ZO	NS	NS	NS	NS	NS	ZO
PS	ZO	ZO	ZO	ZO	ZO	ZO	ZO
PM	PB	NS	PS	PS	PS	PS	PB
PB	PB	PM	PM	PM	PS	PS	PB

In order to make the overshoot of the system respond relatively small and to ensure the system response speed, K_p is set smaller, and K_i and K_d values are chosen respectively. In order to make the system have better steady state, K_p and K_i are set larger, and to avoid oscillations near the set point, K_d is set properly. When ec is small, K_d is set middle, and when ec is

large, K_d is set small. According to the given rules, the control rule table of PID parameters can be obtained and the control rules for K_p, K_d and K_i is listed in table I, II and III.

Defuzzification

In this paper, the weighted average method to the fuzzy evaluation is used to get the precise control value with formula as shown in equation 10.

ec Δk_i e	NB	NM	NS	ZO	PS	PM	PB
NB	NB	NB	NM	NM	NS	ZO	ZO
NM	NB	NB	NM	NS	NS	ZO	ZO
NS	NB	NM	NS	NS	ZO	PS	PS
ZO	NM	NM	NS	ZO	PS	PM	PM
PS	NM	NS	ZO	PS	PS	PM	PB
PM	ZO	ZO	PS	PS	PM	PB	PB
PB	ZO	ZO	PS	PM	PM	PB	PB

ec Δk_i e	NB	NM	NS	ZO	PS	PM	PB
NB	NB	NB	NM	NM	NS	ZO	ZO
NM	NB	NB	NM	NS	NS	ZO	ZO
NS	NB	NM	NS	NS	ZO	PS	PS
ZO	NM	NM	NS	ZO	PS	PM	PM
PS	NM	NS	ZO	PS	PS	PM	PB
PM	ZO	ZO	PS	PS	PM	PB	PB
PB	ZO	ZO	PS	PM	PM	PB	PB

ec Δk_i e	NB	NM	NS	ZO	PS	PM	PB
NB	NB	NB	NM	NM	NS	ZO	ZO
NM	NB	NB	NM	NS	NS	ZO	ZO
NS	NB	NM	NS	NS	ZO	PS	PS
ZO	NM	NM	NS	ZO	PS	PM	PM
PS	NM	NS	ZO	PS	PS	PM	PB
PM	ZO	ZO	PS	PS	PM	PB	PB
PB	ZO	ZO	PS	PM	PM	PB	PB

Where u_i is the fuzzy set values, $\mu(u_i)$ is membership degree of fuzzy values and u_0 is evaluation result. After the three parameters are adjusted by the fuzzy controller, the output control parameters are calculated from the equation 9.

3.3 Neuro- Fuzzy Controller

In the field of artificial intelligence, neuro-fuzzy refers to combinations of artificial neural networks and fuzzy logic. Neuro-fuzzy hybridization results in a hybrid intelligent system by combining the human-like reasoning style of fuzzy systems with the learning and connection structure of neural networks. The drawbacks are the complexity and the darkness of their structures. Industries use the PID technique since it is a crisp control. The self tuning of the P, I, D parameters are quite difficult and the resultant control is with overshoot and with large time constants. To avoid this a combination of Neuro-Fuzzy PID controllers are used for controlling the temperature in the process.

4. Experimental Results And Discussions

4.1 Step Response for the Intelligent Controllers

Comparison between conventional PID, fuzzy PID and Neuro-Fuzzy PID controller's temperature control was performed. According to the analysis fuzzy controller is designed in MATLAB and the fuzzy self-tuning PID control system model is designed by SIMULINK. The step response of the Fuzzy control and conventional control is shown in Figure 7,8.

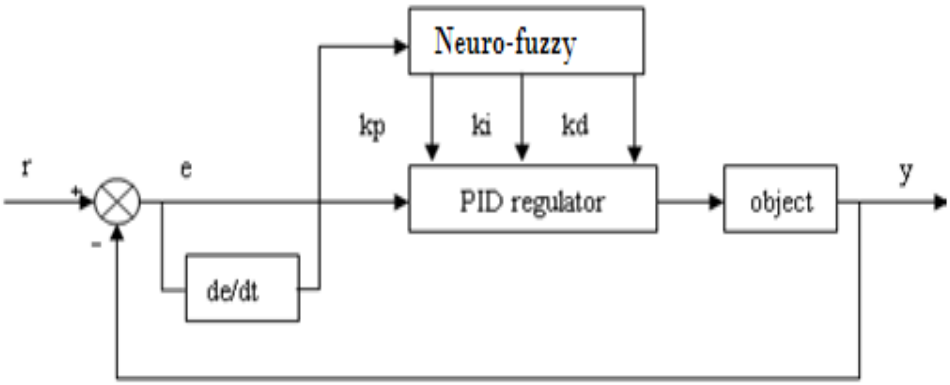


Fig 6: Structure of Neuro-Fuzzy PID Control

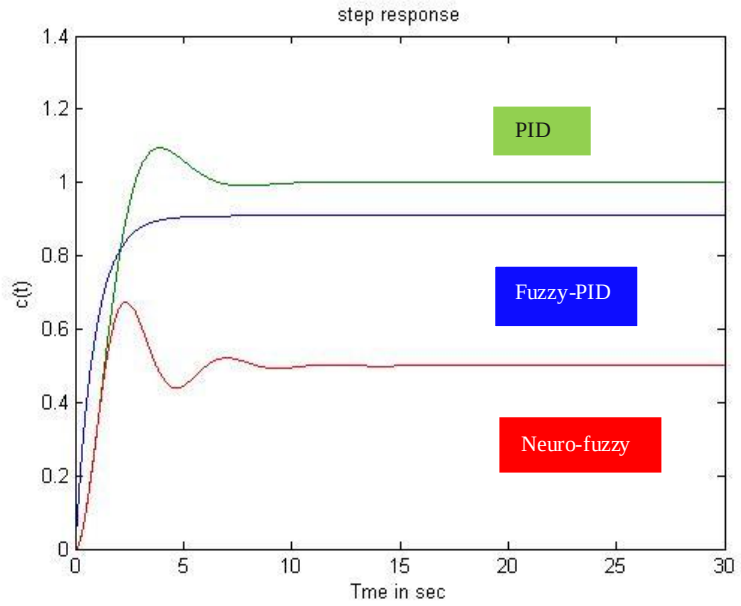


Fig 7: Step Response for the Intelligent Controllers

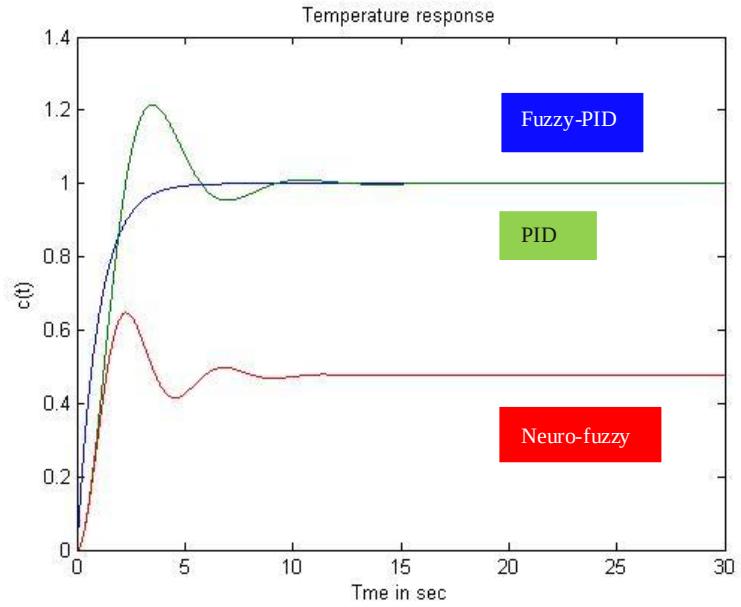


Fig 8: Response of Temperature control

The initial tuning of the PID controller is accomplished based on the Ziegler-Nicholes method, and the gain coefficients are K_p=1.1, K_i=0.003 and K_d=52. Fuzzy PID Controller has a small overshoot, fast response and the steady state error is less than 1%.

4.2 Response of the Temperature control system

The temperature control experiment is conducted in the actual Shell and tube heat exchanger system. The target outlet temperature of heat exchanger is 60˚C and figure 7 shows the response of heat exchanger.The results suggest that neither the settling time nor control accuracy is satisfied enough when conventional PID controller is used in shell and tube heat exchanger. The steady state error of the PID controller is greater than fuzzy PID. The experimental results shows that fuzzy self-tuning PID control has better dynamic response and steady state error characteristics.

4.3 Comparison of various parameters

Table IV shows the comparison of various parameters from the above graph for the various types of Controllers.

TABLE 4
COMPARISON OF PARAMETERS

Parameters	PID	Fuzzy PID	Neuro Fuzzy PID
Rise Time	1.8	2.6	3.2
Peak over shoot	0.68	0.82	1.2
Settling Time	14	5.8	8.6
Peak Time	2.6	4.1	4.2

5. CONCLUSION

In this paper design of a temperature control of a shell and tube heat exchanger based on Neuro-fuzzy PID control was discussed by comparing it with PID and Fuzzy PID. The analysis fuzzy controller is designed in MATLAB and the fuzzy self-tuning PID control system model is designed by SIMULINK. The results suggested that self-tuning parameter fuzzy PID controller has a smaller system overshoot, faster response and less steady state error thereby making it stronger than conventional PID controller. It was thus concluded that fuzzy self-tuning PID control has better dynamic response and steady state error characteristics. The control rule table of PID parameters was obtained and the control rules for K_p, K_d and K_i was tabulated. The actual system used obtains a good control effect and can satisfy the requirements of the temperature control system of the shell and tube heat exchanger.

6. References

[1] H. Thal-Larsen, 1960, Dynamics of heat exchangers and their models, ASME J. Basic Eng., pp. 489 – 504.
[2] M.Chidambaram, and Y.S.N. Malleswara rao, 1992,Nonlinear Controllers for heat exchangers, J. Proc. Cont., vol 2(1), pp 17-21.
[3] A.W.Alsop, and T.F.Edgar, 1998, Non-linear heat exchanger control through the use of partially linearised control variables, Chein. Eng. Commn.,vol.75, pp 155 – 170.
[4] Wang Li-Xin, 1994,Adaptive fuzzy systems and control: Design and stability analysis [M], Prentice-Hall Englewood chiffs NewJersey.
[5] H. Yamashita, R. Izumi, S. Yamaguchi, 1978, Analysis of the dynamic characteristics of cross-flow heat exchangers with both fluids unmixed, Bull. JSME, vol. 21 (153), pp 479-485.
[6] Hassan B.Kazemian, 2001,Development of an intelligent Fuzzy Controller, IEEE International Fuzzy Systems Conference. Vol.1, , pp. 517–520.
[7] Mann G. K.I., Hu B.G.Gosine R.G, 1999, Analysis of direct fuzzy PID controller structures, IEEE Trans on Systems, Man and Cybernetics. Vo1.29, pp. 371-388.
[8] M. Sugeno, 1985, Industrial applications of fuzzy control, Amsterdam, The Netherlands: Elsevier.
[9] Li, H.X. & H. Gatland, 1996,Conventional fuzzy logic control and its enhancement, IEEE Transactions on Systems, Man and Cybernetics, 26(10), pp.791-797.
[10] Jean-Jacques E. Slotine, Weiping Li, 1991,applied nonlinear control, Pretice-hall, Inc., London, pp. 278-282.
[11] George K. I. Mann, Bao-Gang Hu, and Raymond G. Gosine, 2001,Two-level tuning of fuzzy PID controllers, IEEE transactions on systems, man, and cybernetics-part B: cybernetics, Vol. 31, No.,pp. 263-269.
[12] K. M. Hangos, J. Bokor, and G. Szederkényi, 2004, Analysis and control of nonlinear process control systems, Advanced Textbooks in Control and Signal Processing, 1st Edition, Ch. 4, Springer-Verlag London limited, pp. 55-61.

Performance Evaluation of Routing Protocols in MANETs under Wormhole Attack

[1,]Pardeep Kaur, [2,] Deepak Aggarwal
[1,]M. Tech Student , [2,] Assistant Professor
[1,2],Department of CSE & IT, BBSBEC, Fatehgarh Sahib, Punjab, India

Abstract

Mobile Ad-Hoc Network is a group of wireless mobile nodes connected to each-other without any central administrator. Nodes can move from one place to another in the network or may leave or join the network at any time. Due to this the topology of the network changes rapidly. So the routing protocols are required that can adopt the frequent changes in the network topology. Due to the absence of central administrator the MANETs are vulnerable to attacks. In this paper comparison of reactive protocols i.e AODV and DYMO has been done under three types of wormhole attack. Performance is measured with metrics like Packet Delivery Ratio, Average End-to-End Delay, Throughput and Jitter by varying the number of nodes.

Keywords-AODV, DYMO, MANET, Wormhole

1. Introduction

Mobile Ad-Hoc Network is a group of wireless mobile nodes connected to each-other without any central administrator. The nodes can leave or join the network at any time. Nodes act as routers that relay packets generated by other nodes to their destination [Jeroen Hoebeke et al., 2006]. Due to the movement of nodes the topology of the network changes rapidly. The nodes which are near to each other or within each other's radio range can communicate directly. But nodes which are far away they use intermediate nodes to send data. MANETs has advantages like Simple, cheap and fast setup of networks, more robust concerning failure of single component due to decentralized structure because of these they are used in many applications like wireless sensor networks, rescue operations, sports events and conferences etc.

2. Routing Protocols

Proactive protocols are also known as table driven protocols. In these protocols each node maintains a route in their routing table to all the destination nodes in the network. Due to that, routes are discovered for every mobile node of the network, without any request for communication by the hosts [Gurjinder Kaur et al., 2011]. The routing tables are updated periodically or when a change occurs in the network topology. Some of proactive protocols are DSDV, OLSR and STAR. Reactive protocols are also known as on-demand routing protocols. In these protocols a route is only discovered when source node want to send data to the destination node. Source node broadcast a route request message to find a route to the destination. Some of the reactive routing protocols are DSR, AODV and DYMO. Due to the random movement of nodes, the network topology becomes unpredictable and changes rapidly. In order to find the most adaptive and efficient routing protocols for dynamic MANET topologies, the behavior of routing protocols need to be analyzed at varying node speeds, network size, number of traffic nodes and node density [Fahim Maan et al., 2010].AODV and DYMO routing protocols are used in simulation.

2.1 AODV

Ad-hoc on–demand distance vector is a reactive routing protocol. This property implies that it requests a route when it needs one and the nodes which do not want to take part in active communication, need not to maintain routing tables. AODV uses the sequence number to find fresh routes. AODV has two basic operations: route discovery and route maintenance. AODV uses RREQ, RREP and RERR messages to find and maintain the routes.In route discovery , when a source node desire a route to the destination node for which it does not have a route, it broadcast a route request (RREQ) message in the network. RREQ message contains source IP address, destination IP address, sequence number, hop count and broadcast ID. A neighbor receiving a RREQ may send route reply (RREP), if it is either the destination or if it has unexpired route to the destination. When destination node send a route reply (RREP) message to the source node, a forward path is formed. Now source node will send the data through this path.In route maintenance, when a link breakage in an active route is detected, the node notifies this link breakage by sending a route error (RERR) message to the source node [Dong-Won Kum et al., 2010] . The source node will reinitiate the route discovery process if it still has data to send.

2.2 DYMO

DYMO is a successor of AODV. It is a combination of AODV and DSR routing protocols. Similar to AODV, DYMO has two basic operations, route discovery and route maintenance. In route discovery, the source node broadcast a RREQ message throughout the network to find the destination node. During this process, each intermediate node records a route to the source node and rebroadcast the RREQ after appending its own address. This is called the path

accumulation function. When the destination node receives the RREQ, it responds with RREP to the source node. Each intermediate node that receives the RREP records a route to the destination node. When the source node receives RREP message, the route is established between the source node and the destination node. As path accumulation function can reduce the routing overhead, although the packet size of the routing packet is increased [Dong-Won Kum et al., 2010]. When a link breaks, the source of the packet is notified. RERR message is sent to inform the source node.

3. Wormhole Attack

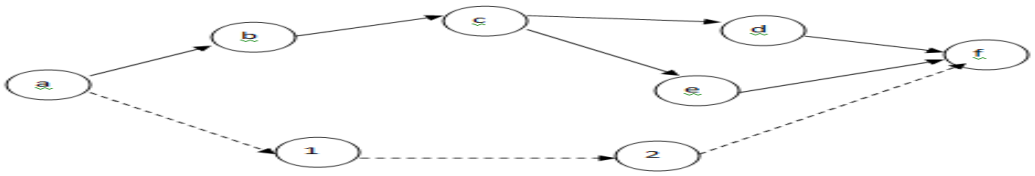


Fig 1. Wormhole attack

Wormhole is a severe type of attack, where two attackers are connected to each other through high speed off-channel link. In this wormhole node receives the packet at one location and send it to other wormhole node through high speed off-channel link. The worst can happen that nodes can be in dilemma that they are close to the destination even though they are at far distance.

Three types of wormhole attack are:

1. All Pass: In this wormhole nodes will pass all the packets irrespective of their size.
2. All Drop: In this all the packets are dropped by wormhole nodes.
3 Threshold: Wormhole drops all the packets size greater than or equal to the threshold value.

4. Simulation And Results

The Qualnet 5.2 simulator is used for simulation. The MAC protocol IEEE 802.11 was used with a data rate of 2 Mbps.

Table 1. Simulation Parameters

Parameter	Value
Terrain Size	1500m×1500 m
No. of Nodes	25/50/75/100
No. of wormhole nodes	4/8/12/16
Traffic Type	CBR
No. of CBR links	5
Mobility Model	Random Waypoint
Routing Protocols	AODV, DYMO
MAC	802.11
Packet Size	512 bytes
Speed	0-10m/s
Pause Time	10 sec
Simulation Time	400 sec
Attack Type	Wormhole

4.1 Performance Metrics

Performane Metrics used to measure the performance are:

4.1.1 Packet Delivery Ratio: Packet delivery ratio is calculated by dividing the number of packets received by the destination through the number of packets originated by source.

4.1.2 Average End-to-End Delay: Average end-to-end delay is the average time it takes a data packet to reach to destination in seconds. It is calculated by subtracting "time at which first packet was transmitted by source" from "time at which first data packet arrived to destination.

4.1.3 Throughput: It is defined as total number of delivered data packets divided by the total duration of simulation time.

4.1.4 Jitter: Jitter is the variation in the time between packets arriving, caused by network congestion, and route changes.

Results without wormhole attack

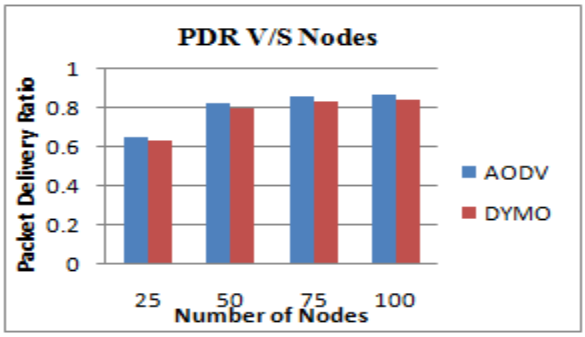


Fig 2: PDR without wormhole

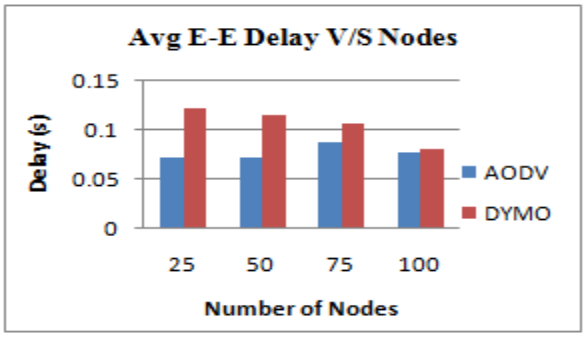


Fig 3 Avg. E-to-E Delay without wormhole

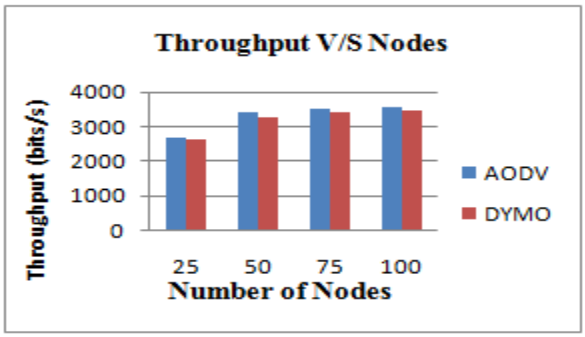


Fig. 4 Throughput without wormhole

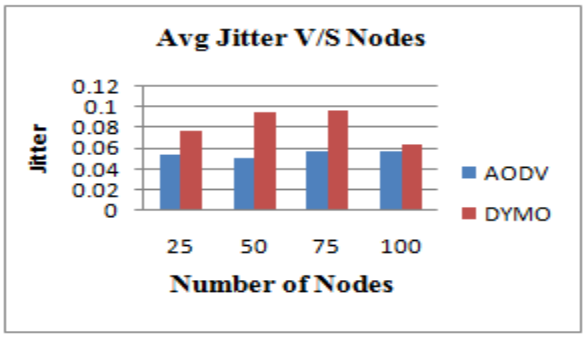


Fig. 5 Avg. jitter without wormhole

Results with Wormhole

- **All Pass**

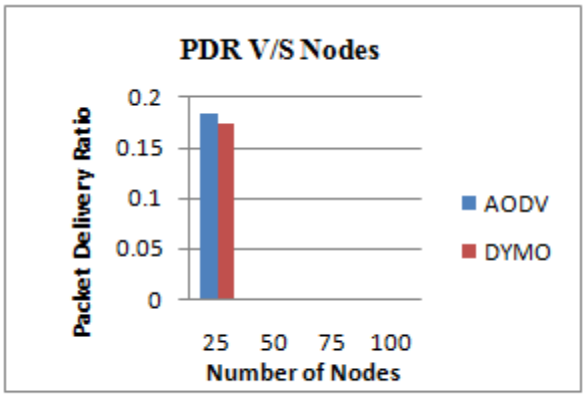


Fig. 6 PDR for All Pass mode

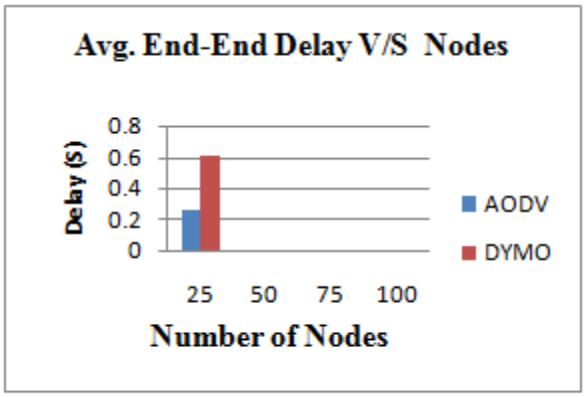


Fig. 7 Avg. E-to-E Delay for All Pass mode

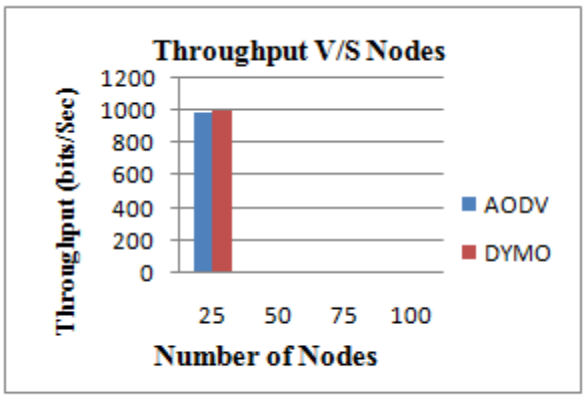


Fig. 8 Throughput for All Pass mode

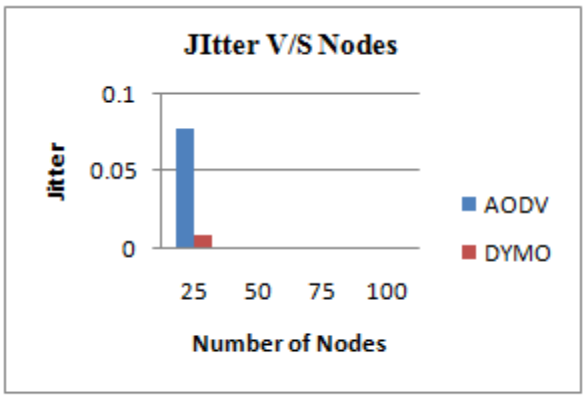


Fig. 9 Avg. Jitter for All Pass mode

➢ **ALL Drop** In all drop, all the packets that are sent by the sender to the receiver are dropped by the wormhole nodes.

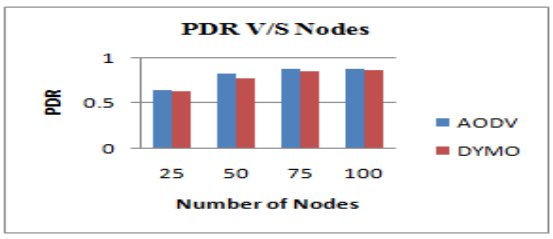


Fig. 10 PDR for All drop mode

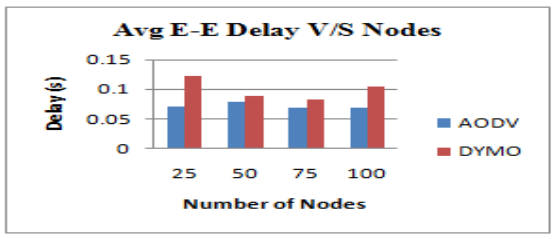


Fig. 11 Avg E-to-E Delay for All Drop mode

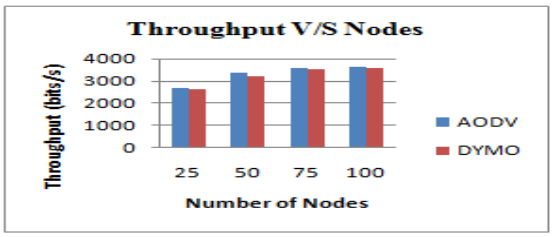


Fig. 12 Throughput for All Drop mode

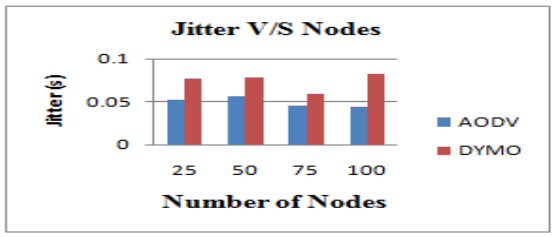


Fig. 13 Avg Jitter for All Drop mode

➢ **Threshold (150 Bytes)** In this case, wormhole drops all the packets which are above 150 bytes in size.

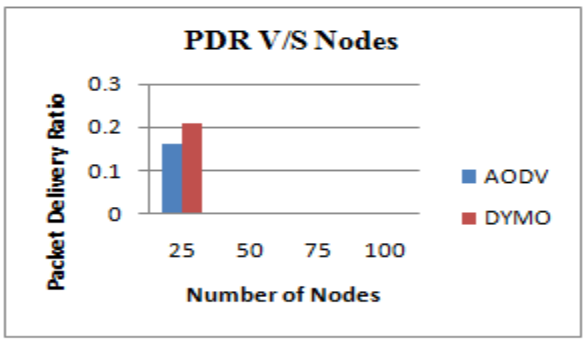


Fig. 14 PDR for Threshold mode

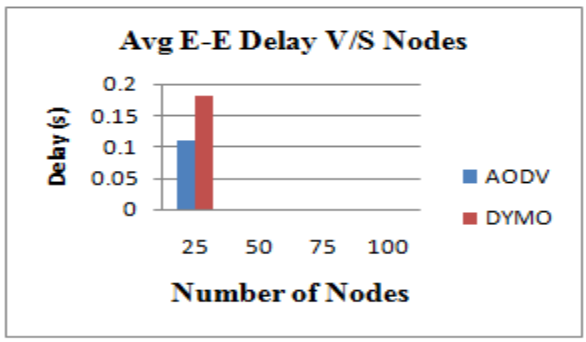


Fig. 15 Avg. E-to-E Delay for Threshold mode

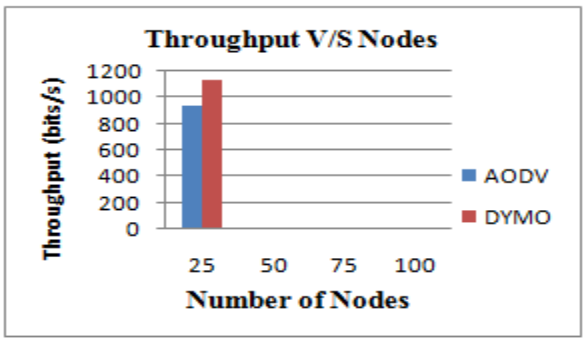


Fig. 16 Throughput for Threshold mode

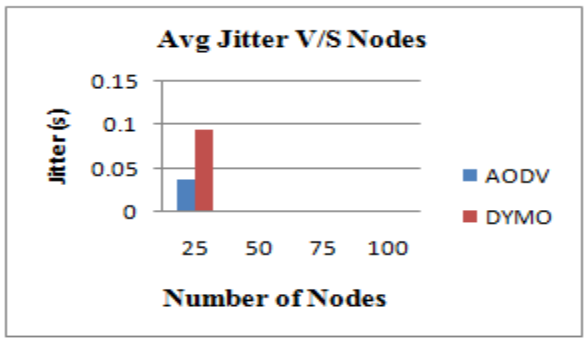


Fig. 17 Jitter for Threshold mode

5. Conlusion And Future Work

From simulation results it is concluded that AODV perform better than DYMO without wormhole attack. But under the wormhole attack the performance of bothe the protocols decreased. But still AODV has better performance than DYMO. All Pass and Threshold has effected the performance greatly. At 25 nodes it has shown some results but as the number of nodes increased the wormhole nodes came into existance and decreased the performance completely. All drop has less affect on the performance as it drop the route request packets and routes are established through other nodes in the network.In future work performance can be measured with different performance metrics like routing overhead by varying pause time and speed. Security mechanism to prevent from these types of attacks can also be developed.

References

[1] Priti Garg, Asma Tuteja. Comparative Performance Analysis of Two Ad-hoc Routing Protocols, International Conference on Network and Electronics Engineering, IPCSIT vol.11,IACSIT Press, Singapore, 2011.

[2] Pravin Ghosekar, Girish Katkar, Dr. Pradip Ghorpade. Mobile Ad Hoc Networking: Imperatives and Challenges, IJCA Special Issue on "Mobile Ad-hoc Networks", 2010.

[3] Jeroen Hoebeke, Ingird Moerman, Bart Dhoebt and Piet Demeester, , An Overview of Mobile Ad-hoc Networking: Applications and Challenges, 2006.

[4] S. Suresh kumar, T. V.P Sundararajan and A Shanmugam. Performance Comparison of three types of wormhole attack in Mobile Adhoc Network, proceedings of the international conference on information science and applications, Chennai, India, 2010.

[5] Gurjinder Kaur, V. K Jain and Yogesh Chaba. Wormhole attacks: Performance Evaluation of On Demand Routing Protocols in Mobile Ad-hoc Networks, world conference on information and communication technologies,pp. 1155-1158, 2011.

[6] Dong-Won Kum, Jin-Su Park, You-Ze Cho and Byoung-Yoon Cheon. Performance Evaluation of AODV and DYMO Routing Protocols in MANET, IEEE CCNC, 2010.

[7] Fahim Maan, Nauman Mazhar. MANET Routing Protocols vs Mobility Models: A Performance Evaluation,ICUFN, 2010.

[8] MIAO Quan-xing, XU Lei. DYMO Routing Protocol Research and Simulation Based on NS2 International Conference on Computer Application and System Modeling, 2010.

[9] Rashid Sheikh, Mahakal Singh Chandel, durgesh Kumar Mishra. Security issues in MANET: A Review, IEEE, 2010.

Fundamental Theorem Of Algebra A Study

Dr Mushtaq Ahmad Shah
Department of Mathematics CMJ University Shilliong 79300

Abstract: This paper is a study of the Fundamental Theorem of Algebra which states that every polynomial equation of degree n has exactly n zeroes. It gives a historical account of the theorem in different periods; by different mathematicians it also includes the contribution of different countries. In addition to this I present different proofs of Fundamental Theorem of Algebra by using different techniques which is actually the main theme behind the paper.

Keywords:- Polynomials, zeroes, analytic, bounded, constant, Maximum Modulus,

1. Introduction

When we speak of the early history of algebra, first of all it is necessary to consider the meaning of the term. If by algebra we mean the science which allows us to solve the equation $ax^2 + bx + c = 0$, expressed in these symbols, then the history begins in the 17th Century ; if we remove the restrictions as to these particular signs and allow for other and less convenient symbols. we might properly begins the history in the 3rd century ; if we allow for the solution of the above equation by geometric methods, without algebraic symbols of any kind, we might say that the algebra begins with the Alexandrian School or a little earlier; and if we say we should class as algebra any problem that we should now solve by algebra (even through it was at first solved by mere guessing or by some cumbersome arithmetic process), then the science was known about I800 B.C,, and probably still earlier. It is first proposed to give a brief survey of the development of algebra, recalling the names of those who helped to set the problems that were later solved by the aid of equation, as well as those who assisted in establishing the science itself. These names have been mentioned in Volume 1 and some of them will be referred to when we consider the development of the special topics of algebra and their application to the solution of the elementary problems. It should be borne in mind that most ancient writers outside Greece included in their mathematics works a wide range of subjects. Ahmes (c.1550 B.C.), for example, combines his algebra with arithmetic and mensuration, and even shows some evidence that trigonometry was making a feeble start. There was no distinct treatise on algebra before the time of Diophantus (c.275). There are only four Hindu writers on algebra whose are particularly noteworthy. These are Aryabhata, whose Aryabha-tiyam(c.510) included problems in series, permutation , and linear and quadratic equations; Brahmagupta, whose Brahmasid-dhanta(c.628) contains a satisfactory rule for the solving the quadratic, and whose problems included the subjects treated by Aryabhata : Mahavira whose Ganita-Sari Sangraha (c.850) contains a large number of problems involving series, radicals, and equations; and Bhaskara, whose Bija Ganita (c.1150) contains nine chapters and extended the work through quadratic equations.

Algebra in the modern sense can hardly be said to have existed in the golden age of Greek mathematics. The Greeks of the classical period could solve many algebraic problems of considerable difficulty, but the solutions were all geometric. Hippocrates (c.460 B.C.), for example, assumed a construction which is equivalent to solving the equation

$$x^2 + \sqrt{\frac{3}{2}} \cdot ax = a^2.$$

With Diophantus (c.275) there first enters an algebraic symbolism worthy of the name, and also a series of purely algebraic problems treated by analytic methods. Many of his equations being indeterminate, equation of this type are often called Diophantine Equations. His was the first work devoted chiefly to algebra, and on his account he is often, and with must justice, called the father of the science. The algebraists of special prominence among the Arabs and Persians were Mohammed ibn Musa, al- Khowarizmi, whose al- jabr w'al muqabalah(c.825) gave the name to the science and contained the first systematic treatment of the general subject as distinct from the theory of numbers; Almahani (c.860), whose name will be mentioned in connection with the cubic; Abu kamil (c.900), who drew extensively from al- khawarizmi and from whom Fibonacci (1202) drew in turn; al-Karkhi(c.1020), whose Fakhri contains various problems which still form part of the general stock material of algebra; and Omar Khanyyam (c.1100), whose algebra was the best that the Persian writers produced.

Most of the medieval Western scholars who helped in the progress of algebra were translators from the Arabic. Among these were Johannes Hispalensis (c.1140), who may have translated al-Khowarizmi's algebra; Gherardo of Cremona (c.1150), to whom is also attributed a translation of the same work; Adelard of Bath (c.1120), who probably translated an asronomical work of al-Khowarizmi, and who certainly helped to make this writer Known; and Robert of Chester, whose translation of al-Khoarizmi's algebra is now available in English.

The first epoch-making algebra to appear in print was the Ars Magna of Cardan (1545). This was devoted primarily to the solutions of algebraic equations. It contained the solutions of the cubic and biquadratic equations, made use of complex numbers, and in general may be said to have been the first step towards modern algebra.The next great work on algebra to appear in print was the General Trattato of Tartaglia (1556-1560), although his side of the controversy with Cardan over the solution of the cubic equation had already been given in his qvesitised inventioni diverse (1546). The first noteworthy attempt to write algebra in England was made by Robert Recorde, who Whetstone of Witte (1557) was an excellent textbook for its first time. The next important contribution was Masterson's incomplete treatise of 1592-1595, but the work was not up to standard set by Recorde. The first Italian textbook to bear the title of algebra was Bombelli's work of 1572. In this book the material is arranged with some attention to the teaching of the subject.

By this time elementary algebra was fairly well perfected and it only remained to develop a good symbolism. Every real polynomial can be expressed as the product of real linear and real quadratic factors.Early studies of equations by al-Khwarizmi (c 800) only allowed positive real roots and the Fundamental Theorem of Algebra was not relevant. Cardan was the first to realise that one could work with quantities more general than the real numbers. This discovery was made in the course of studying a formula which gave the roots of a cubic equation. The formula when applied to the equation x3 = 15x + 4 gave an answer involving √-121 yet Cardan knew that the equation had x = 4 as a solution. He was able to manipulate with his 'complex numbers' to obtain the right answer yet he in no way understood his own mathematics. Bombelli, in his Algebra, published in 1572, was to produce a proper set of rules for manipulating these 'complex numbers'. Descartes in 1637 says that one can 'imagine' for every equation of degree n, n roots but these imagined roots do not correspond to any real quantity Viète gave equations of degree n with n roots but the first claim that there are always n solutions was made by a Flemish mathematician Albert Girard in 1629 in L'invention en algebre However he does not assert that solutions are of the form a + bi, a, b real, so allows the possibility that solutions come from a larger number field than C. In fact this was to become the whole problem of the Fundamental Theorem of Algebra for many years since mathematicians accepted Albert Girard's assertion as self-evident. They believed that a polynomial equation of degree n must have n roots, the problem was, they believed, to show that these roots were of the form a + bi, a, b real Now Harriot knew that a polynomial which vanishes at t has a root x - t but this did not become well known until stated by Descartes in 1637 in La geometrie, so Albert Girard did not have much of the background to understand the problem properly.

A 'proof' that the Fundamental Theorem of Algebra was false was given by Leibniz in 1702 when he asserted that $x^4 + t^4$ could never be written as a product of two real quadratic factors. His mistake came in not realizing that √i could be written in the form a + bi, a, b real. Euler, in a 1742 correspondence with Nicolaus(II) Bernoulli and Goldbach, showed that the Leibniz counter example was false. D'Alembert in 1746 made the first serious attempt at a proof of the Fundamental Theorem of Algebra . For a polynomial f(x) he takes a real b, c so that f (b) = c. Now he shows that there are complex numbers z1 and w_1 so that

$$|z_1| < |c|, |w_1| < |c|.$$

He then iterates the process to converge on a zero of f. His proof has several weaknesses. Firstly, he uses a lemma without proof which was proved in 1851 by Puiseau, but whose proof uses the Fundamental Theorem of Algebra Secondly, he did not have the necessary knowledge to use a compactness argument to give the final convergence. Despite this, the ideas in this proof are important. Euler was soon able to prove that every real polynomial of degree n, $n \leq 6$ had exactly n complex roots. In 1749 he attempted a proof of the general case, so he tried to prove the Fundamental Theorem of Algebra for Real Polynomials: Every polynomial of the nth degree with real coefficients has precisely n zeros in C. His proof in Recherches sur les racines imaginaires des équations is based on decomposing a monic polynomial of degree 2^n into the product of two monic polynomials of degree $m = 2^{n-1}$. Then since an arbitrary polynomial can be converted to a monic polynomial by multiplying by ax^k for some k the theorem would follow by iterating the decomposition. Now Euler knew a fact which went back to Cardan in Ars Magna, or earlier, that a transformation could be applied to remove the second largest degree term of a polynomial. Hence he assumed that

$$x^{2m} + Ax^{2m-2} + Bx^{2m-3} + \ldots = (x^m + tx^{m-1} + gx^{m-2} + \ldots)(x^m - tx^{m-1} + hx^{m-2} + \ldots)$$

and then multiplied up and compared coefficients. This Euler claimed led to g, h, ... being rational functions of A, B, ..., t. All this was carried out in detail for n = 4, but the general case is only a sketch. In 1772 Lagrange raised objections to Euler's proof. He objected that Euler's rational functions could lead to 0/0. Lagrange used his knowledge of permutations of roots to fill all the gaps in Euler's proof except that he was still assuming that the

polynomial equation of degree n must have n roots of some kind so he could work with them and deduce properties, like eventually that they had the form a + bi, a, b real. Laplace, in 1795, tried to prove the Fundamental Theorem of Algebra using a completely different approach using the discriminant of a polynomial. His proof was very elegant and its only 'problem' was that again the existence of roots was assumed Gauss is usually credited with the first proof of the Fundamental Theorem of Algebra . In his doctoral thesis of 1799 he presented his first proof and also his objections to the other proofs. He is undoubtedly the first to spot the fundamental flaw in the earlier proofs, to which we have referred many times above, namely the fact that they were assuming the existence of roots and then trying to deduce properties of them. Of Euler's proof Gauss says ... if one carries out operations with these impossible roots, as though they really existed, and says for example, the sum of all roots of the equation

$$x^m + ax^{m-1} + bx^{m-2} + \ldots = 0$$

is equal to –a even though some of them may be impossible (which really means: even if some are non-existent and therefore missing), then I can only say that I thoroughly disapprove of this type of argument. Gauss himself does not claim to give the first proper proof. He merely calls his proof new but says, for example of d'Alembert's proof, that despite his objections a rigorous proof could be constructed on the same basis. Gauss's proof of 1799 is topological in nature and has some rather serious gaps. It does not meet our present day standards required for a rigorous proof. In 1814 the Swiss accountant Jean Robert Argand published a proof of the Fundamental Theorem of Algebra which may be the simplest of all the proofs. His proof is based on d'Alembert's 1746 idea. Argand had already sketched the idea in a paper published two years earlier Essai sur une manière de représenter les quantitiés imaginaires dans les constructions géometriques. In this paper he interpreted i as a rotation of the plane through 90° so giving rise to the Argand plane or Argand diagram as a geometrical representation of complex numbers. Now in the later paper Réflexions sur la nouvelle théorie d'analyse Argand simplifies d'Alembert's idea using a general theorem on the existence of a minimum of a continuous function.

In 1820 Cauchy was to devote a whole chapter of Cours d'analyse to Argand's proof (although it will come as no surprise to anyone who has studied Cauchy's work to learn that he fails to mention Argand !) This proof only fails to be rigorous because the general concept of a lower bound had not been developed at that time. The Argand proof was to attain fame when it was given by Chrystal in his Algebra textbook in 1886. Chrystal's book was very influential.Two years after Argand's proof appeared Gauss published in 1816 a second proof of the Fundamental Theorem of Algebra . Gauss uses Euler's approach but instead of operating with roots which may not exist, Gauss operates with indeterminates. This proof is complete and correct. A third proof by Gauss also in 1816 is, like the first, topological in nature. Gauss introduced in 1831 the term 'complex number'. The term ‘conjugate’ had been introduced by Cauchy in 1821. Gauss's criticisms of the Lagrange-Laplace proofs did not seem to find immediate favour in France. Lagrange's 1808 2nd Edition of his treatise on equations makes no mention of Gauss's new proof or criticisms. Even the 1828 Edition, edited by Poinsot, still expresses complete satisfaction with the Lagrange- Laplace proofs and no mention of the Gauss criticisms. In 1849 (on the 50th anniversary of his first proof!) Gauss produced the first proof that a polynomial equation of degree n with complex coefficients has n complex roots. The proof is similar to the first proof given by Gauss. However little since it is straightforward to deduce the result for complex coefficients from the result about polynomials with real coefficients.

It is worth noting that despite Gauss's insistence that one could not assume the existence of roots which were then to be proved reals he did believe, as did everyone at that time, that there existed a whole hierarchy of imaginary quantities of which complex numbers were the simplest. Gauss called them a shadow of shadows. t was in searching for such generalisations of the complex numbers that Hamilton discovered the quaternions around 1843, but of course the quaternions are not a commutative system. The first proof that the only commutative algebraic field containing R was given by Weierstrass in his lectures of 1863. It was published in Hankel's book Theorie der complexen Zahlensysteme.Of course the proofs described above all become valid once one has the modern result that there is a splitting field for every polynomial. Frobenius, at the celebrations in Basle for the bicentenary of Euler's birth said Euler gave the most algebraic of the proofs of the existence of the roots of an equation, the one which is based on the proposition that every real equation of odd degree has a real root. I regard it as unjust to ascribe this proof exclusively to Gauss, who merely added the finishing touches. The Argand proof is only an existence proof and it does not in any way allow the roots to be constructed. Weierstrass noted in 1859 made a start towards a constructive proof but it was not until 1940 that a constructive variant of the Argand proof was given by Hellmuth Kneser. This proof was further simplified in 1981 by Martin Kneser, Hellmuth Kneser's son. In this dissertation we shall use various analytical approaches to prove the theorem All proofs below involve some analysis, at the very least the concept of continuity of real or complex functions. Some also use differentiable or even analytic functions.

This fact has led some to remark that the Fundamental Theorem of Algebra is neither fundamental, nor a theorem of algebra.

Some proofs of the theorem only prove that any non-constant polynomial with real coefficients has some complex root. This is enough to establish the theorem in the general case because, given a non-constant polynomial $p(z)$ with complex coefficients, the polynomial

$$q(z) = p(z)\overline{p(\overline{z})}$$

has only real coefficients and, if z is a zero of $q(z)$, then either z or its conjugate is a root of $p(z)$.

Different Proofs of the theorem: Statement of Fundamental theorem of algebra Every polynomial equation of degree n has exactly n zeroes An expression of the form

$P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$

Where $a_0, a_1 \ldots a_{n-1}, a_n \neq 0$

are real or complex numbers and p(x) is called a polynomial equation of degree n and the equation $p(x) = 0$ is called a polynomial equation of degree n .

By a zero of the polynomial x or a root of the equation $x = 0$, we mean a value of p(x) such that $p(x) = 0$.

First proof: For the proof of the Theorem we must know the following Theorem known as Liouville's Theorem.

STATEMENT: If a function f(z) is analytic for all finite values of z and is bounded, then f(z) is constant. "or"

If f is regular in whole z-plane and if $|f(z)| < K$ for all z then f(z) is constant.

PROOF: Let a & b be arbitrary distinct points in z-plane and let C be a large circle with center $z = 0$ and radius R such that C encloses a & b. The equation of C is

$|z| = R$ so that $z = Re^{i\theta}$

$dz = iRe^{i\theta}$

$|dz| = R\,d\theta$

f(z) is bounded for all z $\Rightarrow |f(z)| \leq$ m for all z where $m > 0$

By Cauchy integral formula

$$f(a) = \frac{1}{2\pi i}\int_c \frac{f(z)}{z-a}d(z)$$

$$f(b) = \frac{1}{2\pi i}\int_c \frac{f(z)}{z-b}d(z)$$

$$f(a) - f(b) = \frac{1}{2\pi i}\int_c \left(\frac{1}{z-a} - \frac{1}{z-b}\right) f(Z)d(Z)$$

$$f(a) - f(b) = \frac{a-b}{2\pi}\int_c \frac{f(z)d(z)}{(z-a)(z-b)}$$

$$|f(a) - f(b)| \leq \frac{|a-b|}{2\pi}\int_c \frac{|f(z)||d(z)|}{(|z|-|a|)(|z|-|b|)}$$

$$|f(a) - f(b)| \leq \frac{M|a-b|.2\pi R}{2\pi(R-|a|)(R-|b|)}$$

$$|f(a) - f(b)| \leq \frac{MR|a-b|}{(R-|a|)(R-|b|)} \to 0 \text{ as } R \to \infty$$

• • • f(a) –f(b) = 0

"or" f(a) = f(b)

Showing there by f(z) is constant

[$\int_c |dz|$ = circumference of circle, $C = 2\pi R$]

The liouville's Theorem is one of the most outstanding Theorems in Complex Analysis which has no counter part in Real Analysis. In fact the Theorem does not hold for real function.

2. Proof Of The Fundamental Theorem Of The Algebra

We shall prove it by contradiction suppose $p(z) \neq 0$ for any value of z. Then

f(z) = $\frac{1}{p(z)} = \frac{1}{a_0 + a_1 z + \ldots + a_n z^n}$

f(z) = $\frac{1}{z^n\left(\frac{a_0}{z^n} + \frac{a_1}{z^{n-1}} + \frac{a_2}{z^{n-2}} + \ldots + a_n\right)} \to 0 as z \to \infty$

$\therefore$ for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$|f(z)| < \varepsilon$ when $|z| < \delta$

Since f(z) is continuous in the bounded closed domain $|z| \leq \delta$ therefore f(z) is bounded in the closed domain $|z| \leq \delta$, so there exists a positive number k such that

$|f(z)| < k$ for $|z| \leq \delta$

If M = Max(ε , k) , Then we have

$|f(z)| = \left| \frac{1}{p(z)} \right| < M$ for every z

Hence by Liouville's Theorem f(z) is constant. This gives a contradiction. Since p(z) is not constant for n =1, 2, 3, and $a_n \neq 0$ Thus p(z) must be zero for at least one variable of z.

i.e. p(z) = 0 must have at least one root say α_1 then we have p(α_1) = 0

Now p(z) = p(z) – p(α_1)

i.e. $P(z) = (a_0 + a_1 z + a_2 z^2 + \ldots + a_n z^n)$

$-(a_0 + a_1 \alpha_1 + a_2 \alpha_1^2 + \ldots + a_n \alpha_1^n)$

"or" $p(z) = a_1(z - \alpha_1) + a_2(z^2 - \alpha_1^2) + \ldots + a_n(z^n - \alpha_1^n)$

i.e. $P(z) = (z - \alpha_1)\, p_1(z)$

where $p_1(z)$ is a polynomial of degree n-1. Again $p_1(z) = 0$ must have at least one root say α_2 (α_2 may be equal to α_1) proceeding as above we have

$P(z) = (z - \alpha_1)(z - \alpha_2)p_2(z)$

where $p_2(z)$ is a polynomial of degree n – 2 continuing in this way we see that p(z) = 0 has exactly n roots.

Second proof : For the second proof of the theorem we must know the following theorem known as Rouche's Theorem

Statement: If f(z) and g(z) are analytic inside and on a simple closed curve C and $|g(z)| < |f(z)|$ on C ,
Then f(z) and f(z) + g(z) both have the same number of zeros inside C

Proof : Suppose f(z) and g(z) are analytic inside and on a simple closed curve C and

$|g(z)| < |f(z)|$ on C

Firstly we shall prove that neither f(z) nor f(z) + g(z) has zero on C
If f(z) has a zero at z = a on C then f(a) = 0

But $|g(z)| < |f(z)|$ on C

which is absurd Again if f(z) +g(z) has a zero at z = a on C
Then f(a) + g(a) = 0 so that

$$f(a) = -g(a)$$

"or"

$$|g(a)| = |f(a)|$$

Again we get a contradiction , Thus neither f(z) nor f(z) + g(z) has a zero on C
Let N_1 and N_2 be number of the zeros of f and f + g respectively inside C.
we know that f and f + g both are analytic within and on C and have no poles inside C.
Therefore, by usual formula.

$$\frac{1}{2\pi i}\int_c \frac{f'}{f}dz = N - P$$

gives $$\frac{1}{2\pi i}\int_c \frac{f'}{f}dz = N_1$$

and

$$\frac{1}{2\pi i}\int_c \frac{f'+g'}{f+g}dz = N_2$$

subtracting we get

$$\frac{1}{2\pi i}\int_c \left(\frac{f'+g'}{f+g} - \frac{f'}{f}\right)dz = N_2 - N_1 \longrightarrow \text{(I)}$$

Taking $\frac{g}{f} = \phi$ so that g = ϕ f

$$|g| < |f| \Longrightarrow \left|\frac{g}{f}\right| < 1 \Longrightarrow |\phi| < 1$$

Also $$\frac{f'+g'}{f+g}$$

$$= \frac{f' + f'\phi + \phi' f}{f + \phi f}$$

$$= \frac{f'(1+\phi) + \phi' f}{f(1+\phi)}$$

"or" $$\frac{f'+g'}{f+g} - \frac{f'}{f} = \frac{\phi'}{1+\phi}$$

Using in (I) we get

$$N_2 - N_1 = \frac{1}{2\pi i}\int_c \frac{\phi'}{1+\phi}dz$$

$$N_2 - N_1 = \frac{1}{2\pi i}\int_c \phi'(1+\phi)^{-1}dz \longrightarrow \text{(II)}$$

Since we have seen that $|\phi| < 1$ and so binomial expansion of $(1+\phi)^{-1}$ is possible and binomial expansion thus obtained is uniformly convergent and hence term by term integration is permissible, Hence

$$\int_c \phi'(1+\phi)^{-1}dz = \int_c \phi'[1-\phi+\phi^2-\phi^3+\ldots]dz$$

$$\int_c \phi'(1+\phi)^{-1}dz = \int_c \phi'dz - \int_c \phi'\phi dz + \int_c \phi'\phi^2 dz + \ldots$$

$$\longrightarrow \text{(III)}$$

The function f and g both are analytic within and on C and $f(z) \neq 0$ for any point on C. Hence $\frac{g}{f} = \phi$ is analytic and non–zero for any point on C. Therefore ϕ and it's all derivatives are analytic

∴ By caucly's integral theorem, each integral on R.H.S of (3) vanishes consequently.

$$\int_c \phi'(1+\phi)'dz = 0$$

In this event (2) takes the form

$N_2 - N_1 = 0$ or $N_1 - N_2$

3. Proof of Fundamental Theorem of Algebra

Consider the polynomial

$a_0 + a_1z + a_2z^2 + \ldots + a_nz^n$

such that $a_n \neq 0$

Take $f(z) = a_nz^n$

$g(z) = a_0 + a_1z + \ldots + a_{n-1}z^{n-1}$

Let C be a circle $|z| = r$ where $r > 1$.

Then

$|g(z)| \leq |a_0| + |a_1|r + |a_2|r^2 + \ldots + |a_{n-1}|r^{n-1}$

$|g(z)| \leq |a_0|r^{n-1} + |a_1|r^{n-1} + |a_2|r^{n-1} \ldots + |a_{n-1}|r^{n-1}$

$|g(z)| \leq [|a_0| + |a_1| + |a_2| + \ldots + |a_{n-1}|]r^{n-1}$

But $|f(z)| = |a_nz^n| = |a_n|r^n$

$$\therefore \left|\frac{g(z)}{f(z)}\right| \leq \frac{[|a_0| + |a_1| + |a_2| + \ldots + |a_{n-1}|]r^{n-1}}{|a_n|r^n}$$

$$\left|\frac{g(z)}{f(z)}\right| = \frac{|a_0| + |a_1| + |a_2| + \ldots + |a_{n-1}|}{|a_n|r}$$

Now if $|g(z)| < |f(z)|$ so that $\left|\frac{g(z)}{f(z)}\right| < 1$, then

$$\frac{|a_0| + |a_1| + |a_2| + \ldots + |a_{n-1}|}{|a_n| r} < 1$$

This

$$\Rightarrow r > \frac{|a_0| + |a_1| + |a_2| + \ldots + |a_{n-1}|}{|a_n|}$$

Since r is arbitrary and hence by choosing r large enough, the last condition can be satisfied so that $|g(z)| < |f(z)|$. Now applying Rouche's theorem, we find that the given polynomial $f(z) + g(z)$ has the same numbers of zeros as $f(z)$ But $f(z) = a_n z^n$ has exactly n zeros all located at $z = 0$. Consequently $f(z) + g(z)$ has exactly n zeros. Consequently the given Polynomial has already n zeros.

Third Proof : For the proof we must know the following theorem known as Maximum Modulus principle.

Statement: Suppose f(z) is analytic within and on a simple closed countor C and f(z) is not constant . Then $|f(z)|$ reaches its maximum value on C (and not inside C), that is to say , if M is the maximum value of $|f(z)|$ on C, then $|f(z)| < M$ for every z inside C.

PROOF: We prove this theorem by the method of contradiction, Analyicity of f(z) declares that f(z) is continuous within and on C. Consequently $|f(z)|$ attains its maximum value M at the same point within or on C. we want to show that $|f(z)|$ attains the value M at a point lying on the boundary of C (and not inside C) . Suppose, if possible, this value is not attained on the boundary of C but is attained at a point z = a within C so that

$\max |f(z)| = |f(a)| = M$ .....(1)

and $|f(z)| \leq M \ \forall$ Z within C.....(2)

Describe a circle Γ with a as center lying within C. Now f(z) is not constant and its continuity implies the existence of a point $z = b$ inside Γ such that $|f(b)| < M$

Let $\varepsilon > 0$ be such that $|f(b)| = M - \varepsilon$

Again $|f(z)|$ is continous at $z = b$

and so $\left|[|f(z)| - |f(b)|]\right| < \frac{\varepsilon}{2}$

Whenever $|z - b| < \delta$

Since

$|[|f(z)| - |f(b)|]| \geq |f(z)| - |f(b)|$

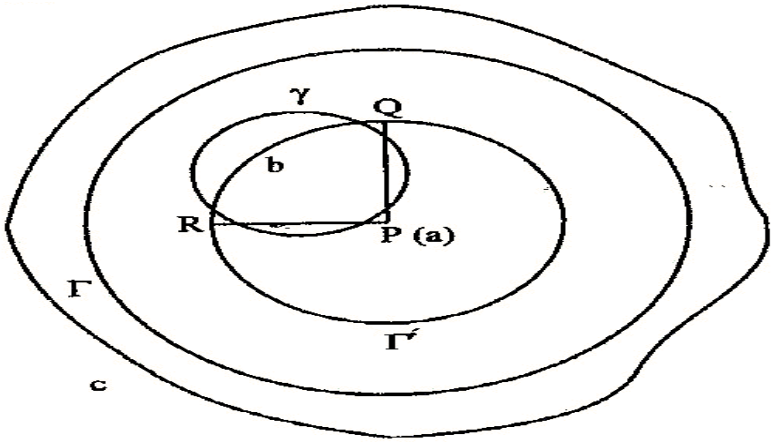
"or" $|f(z)| - |f(b)| \leq \left|[|f(z)| - |f(b)|]\right| < \frac{\varepsilon}{2}$

"or" $|f(z)| - |f(b)| < \frac{\varepsilon}{2}$

"or" $|f(z)| < |f(b)| + \frac{\varepsilon}{2}$

$= M - \in + \frac{\varepsilon}{2} = M - \frac{\varepsilon}{2}$

"or" $|f(z)| < M - \frac{\varepsilon}{2} \ \forall$ z s.t $|z - b| < \delta$ ...(3)



We draw a circle γ with center at b and radius δ. Then (3) shows that

$|f(z)| < M - \frac{\varepsilon}{2} \quad \forall$ Z inside γ .....(4)

Again we draw another circle Γ́ with center at a and radius $|b - a| = r$

By Cauchy's Integral formula .

$$f(a) = \frac{1}{2\pi i}\int \frac{f(z)}{z-a}dz$$

on Γ́, $z - a = re^{i\theta}$

$$f(a) = \frac{1}{2\pi i}\int_0^{2\pi} f(a+re^{i\theta})\frac{rie^{i\theta}d\theta}{rie^{i\theta}}$$

If we measure θ in anti-clock wise direction & if

< QPR = α then

$$f(a) = \frac{1}{2\pi}\left[\int_0^{\alpha} + \int_{\alpha}^{2\pi}\right] f(a+re^{i\theta})d\theta$$

$$\therefore |f(a)| \le \frac{1}{2\pi}\int_0^{\alpha}|f(a)+re^{i\theta}|d\theta + \frac{1}{2\pi}\int_{\alpha}^{2\pi}|f(a+re^{i\theta})|d\theta$$

$$|f(a)| < \frac{1}{2\pi}\int_0^{\alpha}\left(M - \frac{\in}{2}\right)d\theta + \frac{1}{2\pi}\int_{\alpha}^{2\pi} M d\theta$$

$$f(a) = \left(M - \frac{\in}{2}\right)\frac{\alpha}{2\pi} + M\frac{(2\pi-\alpha)}{2\pi} = M - \frac{\alpha\in}{4\pi}$$

Then $M = |f(a)| < M - \frac{\alpha\in}{4\pi}$

"or" $M < M - \frac{\alpha\in}{4\pi}$. A contradiction

For M cannot be less than $M - \frac{\alpha\in}{4\pi}$

Hence the Required results follow.

4. Minimum Modulus Principle

STATEMENT:- Suppose f(z) is analytic within and on a closed contour C and Let $f(z) \neq 0$ inside C suppose further that f(z) is not constant . Then $|f(z)|$ attains its minimum value at a point on the boundary of "C" that is to say, if M is the minimum value of $|f(z)|$ inside and on C. Then $|f(z)| > m \ \forall$ z inside C

PROOF: f(z) is analytic within and on C and $f(z) \neq 0$ inside C. It follows that $\frac{1}{f(z)}$ is analytic within C. By Maximum Modulus principal $\frac{1}{|f(z)|}$ attains its Maximum value on the boundary of C. So that $|f(z)|$ attains its Minimum value on the boundary of C. Hence the theorem

5. Proof Of The Fundamental Theorem Via Maximum

Modulus Principle Proof: Assume p(z) is non-constant and never zero. $\exists$ M such that $|p(z)| \geq |a_0| \neq 0$ if $|z| > M$. Since |p (z)| is continuous, it achieves its minimum on a closed interval. Let z_0 be the value in the circle of radius M where p (z) takes its minimum value.

So $|p(z_0)| \geq |p(z)|$ for all $z \in$ C, and in particular

$|p(z_0)| \geq |p(0)| = |a_0|$.

Translate the polynomial.

Let $p(z) = p((z - z_0) + z_0)$;

Let $p(z) = Q(z - z_0)$.

Note the minimum of Q occurs at

$z = 0$: $|Q(0)| \geq |Q(z)|$ for all $z \in$ C.

$Q(z) = c_0 + c_j z^j + \cdots + c_n z^n$,

Where j is such that cj is the first coefficient (after c_0) that is Non-zero. I must show Q(0) = 0 Note if $c_0 = 0$, we are done.

We may rewrite such that

$$Q(z) = c_0 + c_j z^j + z^{j+1} R(z)$$

We will extract roots.

Let

$$re^{i\theta} = -\frac{c_0}{c_j}$$

Further, Let

$$z_1 = r^{\frac{1}{j}} + e^{\frac{i\theta}{j}}$$

So, $$c_j z^j = -c_0$$

Let $\in > 0$ be a small real number. Then

$$Q(\in z_1) = c_0 + c_j \in^j z_1^{\,j} + \in^{j+1} z_1^{\,j+1} R(\in z_1)$$

$$|Q(\in z_1)| \leq |c_0 + c_j \in^j z_j^{\,j}| + \in^{j+1} |z_1|^{j+1} R(\in z_1)$$

$$|c_0| - \in^j |c_0| + \in^{j+1} |z_1|^{j+1} N,$$

Where N chosen such that $N > |R(\in z_1)|$, and $\in$ is chosen so that

$$\in^{j+1} |z_1|^{j+1} < \in^{j} |c_0|$$

Thus,

$$|Q(\in z_1)| < |c_0|,$$

But this was supposed to be our minimum. Thus , a contradiction. Hence proved

6. Proof of the Fundemental Theorem via Radius of convergence

We now prove the Fundamental theorem of Algebra: As always, p(z) is a non constant polynomial. Consider

$$f(z) = \frac{1}{p(z)} = b_0 + b_1 z + \ldots$$

and

$$p(z) = a_n z^n + \ldots + a_0, a_0 \neq 0$$

Lemma. $\exists c, r \in C$ such that $|b_k| > cr^k$ for infinitely many k.

Now, 1= p(z)f(z). Thus, $a_0 b_0 = 1$. This is our basic step .Assume we have some coefficient such that $|b_k| > cr^k$. We claim we can always find another .Suppose there are no more .Then the coefficient of z^{k+n} in p(z)f(z) is

$$a_0 b_{k+n} + a_1 b_{k+n-1} + \ldots + a_n b_k = 0$$

Thus, as we have $|b_j| > cr^j$ in this range ,we have the coefficient satisfies

$$|a_0| r^n + |a_1| r^{n+1} + \ldots + |a_{n-1}| r \leq |a_n|$$

f $$r \leq \min\left\{1, \frac{|a_n|}{|a_0| + \ldots + |a_{n-1}|}\right\}$$

This will give that

$$|b_k| = \frac{|a_0 b_{k+n} + \ldots + a_{n-1} b_{k+1}|}{|a_n|}$$

$$|b_k| \leq \frac{|a_0 b_{k+n}| + \ldots + |a_{n-1} b_{k+1}|}{|a_n|} \leq cr^k$$

for sufficiently small.

Let z = $\frac{1}{r}$, Then

$$|b_k z^k| = \frac{|b_k|}{r^k} > c$$

This is true for infinitely many k, hence the power series diverges, contradicting the assumption that function is analytic and its power series converges everywhere.

7. Proof Of The Fundamental Theorem Via Picard's Theorem

Statement: If there are two distinct points that are not in the image of an entire function p(z) (ie. $\exists z_1 \neq z_2$ such that for all

$$z \in C, p(z) \neq z_1 or z_2),$$

then p(z) is constant.

We now prove the Fundamental Theorem of Algebra;

Let p(z) be a non-constant polynomial, and assume p(z) is never 0.

Claim:- If p(z) is as above , p(z) does not take on one of the variable $\frac{1}{k}$ for $k \in N$

Proof: Assume not. Thus, $\exists z_k \in C$ such that $p(z_k) = \frac{1}{k}$. If we take a circle D centered at the origin with sufficiently large radius, then $|p(z)| > 1$ for all z outside D. Thus each $z_1 \in D$, we have a convergent subsequence. Thus we have $z_{n_i} \to z'$ .but

$$p(z') = \lim_{n_i \to \infty} p(z_{n_i}) = 0 .$$

Thus there must be some k such that $p(z) \neq \frac{1}{k}$. Since there are two distinct values not in range of p, by picard's Theorem it is now constant. This contradicts our assumption that p(z) is non-constant. Therefore, $p(z_0) = 0$ for some z_0 .

Remark:- One can use a finite or countable versions of picards. Rather than missing just two points, we can modify the above to work if Picard instead stated that if we miss finitely many (or even countably) points, we are constant. Just look at the method above, gives $\frac{1}{k_1}$. We can then find another larger one, say $\frac{1}{k_2}$ .And so on. We can even get uncountably many such points by looking at numbers such as $\frac{\pi}{2}$ (using now the transcendence of C is 1).

8. Proof Of The Fundamental Theorem Via Cauchy's Integral Theorem

Statement:- Let f(z) be analytic inside and on the boundary of some region C. Then

$$\int_{\partial C} f(z) = 0$$

We now prove the Fundamental Theorem of Algebra.

Proof: Without loss of generality let p(z) be a non- constant polynomial and assume p(z) = 0. For $z \in R$ .assume p(z) $\in R$ (Otherwise, consider $p(z)\overline{p}(z)$.

Therefore, p(z) doesn't change signs for $z \in R$, or by the Intermediate Value Theorem it would have a zero.

$$\int_0^{2\pi} \frac{d\theta}{p(2\cos\theta)} \neq 0$$

This follows from our assumption that p(z) is of constant signs for real argument, bounded above from 0. This integral equals the contour integral

$$\frac{1}{i}\int_{|z|=1} \frac{dz}{zp(z+z^{-1})} = \frac{1}{i}\int_{|z|=1} \frac{z^{n-1}}{Q(z)}$$

If z $\neq 0$, Q(z) $\neq 0$

If z =0, then Q(z) $\neq 0$ Since

$P(z + z^{-1}) = a_n(z + z^{-1}) + \ldots$

$z^n p(z + z^{-1}) = z^n(\ldots a_n z^{-n}) + \ldots$

$= a_n + z(\ldots)$

Thus, $Q(z) = a_n$, which is non-zero . Hence, $Q(z) \neq 0$,

Consequently $\frac{z^{n-1}}{Q(z)}$ is analytic. By the Cauchy Integral

Formula $\frac{1}{i}\int_{|z|=1} \frac{z^{n-1}}{Q(z)} \neq 0.$ Thus, a contradiction!

9. The Fundamental Theorem Of Algebra

Our object is to prove that every non constant polynomial f(z) in one variable z over the complex numbers C has a root, i.e. that is a complex number r in C such that f(r) = 0. Suppose that

$f(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0$

where n is at least 1, $a_n \neq 0$ and the coefficients a_i are fixed complex numbers. The idea of the proof is as follows: we first show that as |z| approaches infinity, |f(z)| approaches infinity as well. Since |f(z)| is a continuous function of z, it follows that it has an absolute minimum. We shall prove that this minimum must be zero, which establishes the theorem. Complex polynomial at a point where it does not vanish to decrease by moving along a line segment in a suitably chosen direction.
We first review some relevant facts from calculus.

10. Properties Of Real Numbers And Continuous Functions

Lemma 1. Every sequence of real numbers has a monotone (nondecreasing or nonincreasing) subsequence. Proof. If the sequence has some term which occurs infinitely many times this is clear. Otherwise, we may choose a subsequence in which all the terms are distinct and work with that. Hence, assume that all terms are distinct. Call an element "good" if it is bigger than all the terms that follow it. If there are infinitely many good terms we are done: they will form a decreasing subsequence. If there are only finitely many pick any term beyond the last of them. It is not good, so pick a term after it that is bigger. That is not good, so pick a term after it that is bigger. Continuing in this way (officially, by mathematical induction) we get a strictly increasing subsequence. That proves the theorem

lemma 2. A bounded monotone sequence of real numbers converges.

Proof. This is sometimes taken as an axiom about the reals. What is given here is an intuitive justification. We assume the sequence is non-decreasing: for the other case, take the negatives. The boundedness forces the integer parts of the terms in the sequence to stabilize, since a bounded montone sequence of integers is eventually constant. Thus, eventually, the terms have the form $m + f_n$ where m is a fixed integer, 0 is less than or equal to $f_n < 1$, and the f_n are non decreasing. The first digits of the f_n (after the decimal point) don't decrease and eventually stabilize: call the stable value a_1. For the f_n which begin with $.a_1$ ... the second digits increase and eventually stabilize: call the stable value a_2. Continuing in this fashion we construct a number $f = .a_1\ a_2 \ldots a_h \ldots$ with the property that eventually all the f_n agree with it to as many decimal places as we like. It follows that f_n approaches f as n approaches infinity, and that the original sequence converges to m + f. That proves the theorem

Lemma 3. A bounded sequence of real numbers has a convergent subsequence. If the sequence is in the closed interval [a,b], so is the limit.

Proof. Using Fact 1, it has a monotone subsequence, and this is also bounded. By Fact 2, the monotone subsequence converges. It is easy to check that if all terms are at most b (respectively, at least a) then so is the limit. That proves the theorem

Lemma 4. A sequence of points inside a closed rectangle (say a is less than or equal to x is less than or equal to b, c is less than or equal to y is less than or equal to d) has a convergent subsequence. The limit is also in the rectangle.

Proof. Using Fact 3 we can pick out a subsequence such that the x-coordinates converge to apoint in [a,b]. Applying lemma 3 again, from this subsequence we can pick out a further subsequence such that the y-coordinates converge to a point in [c,d]. The x-coordinates of the points in this last subsequence still converge to a point in [a,b]. That proves the theorem

Lemma 5. A continuous real-valued function f defined on a closed rectangle in the plane is bounded and takes on an absolute minimum and an absolute maximum value.

Proof. We prove the result for the maximum: for the other case consider -f. For each integer n = 1, 2, 3, ... divide the rectangle into n^2 congruent sub rectangles by drawing n-1 equally spaced vertical lines and the same number of equally spaced horizontal lines. Choose a point in each of these sub rectangles (call these the special points at step n) and evaluate f at these points. From among these choose a point where f is biggest: call it P_n. The sequence $P_1, P_2, P_3, \ldots$ has a convergent subsequence: call it $Q_1, Q_2, Q_3, \ldots$, where

$Q_k = P_{\{n_k\}}$ Let Q be the limit of the sequence Q_k. It will suffice to show that f(Q) is bigger than or equal to f(P) for every point P in the rectangle. If not choose P in the rectangle such that f(P) > f(Q). For each k let P'_k be a special point at step n_k in a sub rectangle (among the $(n_k)^2$) that contains P. It follows that P'_k approaches P as k approaches infinity, since

both sides of the sub rectangles approach zero as k approaches infinity. For every k, $f(Q_k)$ is at least $f(P'_k)$, by the choice of Q_k. Taking the limit as k approaches infinity gives a contradiction, since $f(Q_k)$ approaches f(Q) and, by the continuity of f, $f(P'_k)$ approaches f(P) as k approaches infinity. That proves the theorem. The result is valid for a continuous real- valued function on any closed bounded set in R^2 or R^n, where a set is closed if whenever a sequence of points in the set converges, the limit point is in the set.

Lemma 6. Let f be a continuous real- valued function on the plane such that f(x,y) approaches infinity as (x,y) approaches infinity. (This means that given any real number B, no matter how large, there is a real number $m > 0$ such that if $x^2 + y^2$ is at least m then f(x,y) is at least B.) Then f takes on an absolute minimum value at a point in the plane.}

Proof. Let B = f(0,0). Choose $m > 0$ such that if $x^2 + y^2$ is at least m then f(x,y) is at least B. Choose a rectangle that contains the circle of radius m centered at the origin. Pick Q in the rectangle so that the minimum value of f on the rectangle occurs at Q. Since (0,0) is in the rectangle f(Q) is at most B. Since outside the rectangle all values of f are at least B, the value of f at Q is a minimum for the whole plane, not just the rectangle. That proves the theorem

Lemma 7. Let g be a continuous function of one real variable which takes on the values c and d on a certain interval. Then g takes on every value r between c and d on that interval. Proof. Let g(a) = c and g(b) = d. We may assume without loss of generality that a < b. Replacing g by g - r we may assume that g is positive at one endpoint, negative at the other, and never vanishes. We shall get a contradiction. We shall construct a sequence of intervals $I_1 = [a,b]$, I_2, ..., I_n, ... such that $I_{[n+1]}$ is contained in I_n for each n, g has values of opposite sign at the end points of every I_n, and I_n has length $\frac{b-a}{2^{\{n-1\}}}$. In fact if I = $[a_n,b_n]$ has already been constructed and M is the midpoint of I_n, then g(M) has opposite sign from at least one of the numbers $g(a_n)$, $g(b_n)$, and so we can choose one of $[a_n, M]$ or $[M, b_n]$ for $I_{\{n+1\}}$. The a_n are non decreasing, the b_n are non increasing, and $a_n < b_n$ for all n. It follows that both sequences have limits. But $b_n - a_n$ approaches 0 as n approaches infinity, so the two limits are the same. Call the limit h. Since a_n approaches h, $g(a_n)$ approaches g(h). Similarly, $g(b_n)$ approaches g(h). Since $g(a_n)$ and $g(b_n)$ have opposite signs for all n, this can only happen if g(h) = 0. That proves the theorem

Remark:- Consider a polynomial f(x) with real coefficients of odd degree. Then lemma 7 implies that f has at least one real root. To see this, we may assume that f has a positive leading coefficient (otherwise, replace f by -f). It is then easy to see that f(x) approaches +infinity as x approaches + infinity while f(x) approaches -infinity as x approaches - infinity. Since f(x) takes on both positive and negative values, lemma 7 implies that f takes on the value zero.
We want to note that if u, u' are complex numbers then

$$|u + u'| \leq |u| + |u'|.$$

To see this note that, since both sides are non-negative, it suffices to prove this after squaring both sides, i.e. to show that $|u + u'|^2 \leq |u|^2 + 2|uu'| + |u'|^2$. Now, it is easy to see that for any complex number v,

$$|v|^2 = v(\overline{v}),$$

$\overline{v}$ where denotes the complex conjugate of v. Using this the inequality above is equivalent to

$$(u + u')\overline{(u + u')} \leq uu' + 2|uu'| + u'(\overline{u'}).$$

Multiplying out, and canceling the terms which occur on both sides, yields the equivalent inequality

$$u\overline{u'} + \overline{u}u' \leq 2|uu'| = 2|u||u'| = 2|u\overline{u}| = 2|u|\left|\overline{u'}\right|$$

Let w = u(u')-. Then $\overline{w} = (\overline{u})(\overline{u'}) = (\overline{u})u'$.

Thus, what we want to show is that $w + \overline{w} \leq 2\left|\overline{w}\right|$.

If w = a + bi where a, b are real this becomes the assertion that $2a \leq 2\{a^2+b^2\}^{1/2}$

or $a \leq \sqrt{a^2 + b^2}$,

which is clear. Moreover, equality holds if and only if a is non negative and b is zero, i.e., if and only if $w = u(\overline{u'})$ is a non-negative real number.

We also get that $|u \pm u'| \geq |u| - |u'|$: replacing u' by if necessary we can assume the sign is -, and we already know that $|u| \leq |u-u'| + |u'|$, which is equivalent. Finally, we want to justify carefully why, when n is a positive integer, every complex number has an n^{th} root. First note that the fact holds for non-negative real numbers r using lamma 6 applied to the function F:R to R given by $F(x) = x^{n-r}$: the function is non-positive at x = 0 and positive for all sufficiently large x, and so takes on the value 0. We can now construct an n^{th} root for $r(\cos t + i \sin t)$,namely

$$r^{1/n} \cos\left(\frac{t}{n}\right) + i \sin\left(\frac{t}{n}\right),$$

Using De-Moivre's formula.

11. Proof of the fundamental theorem of algebra

Let $f(z) = a_n z^n + ... + a_0$, where the a^i are in C, $n > 0$, and a_n is not 0. If we let $z = x + yi$ with x and y varying in R, we can think of f(z) as a C-valued function of x and y. In fact if we multiply out and collect terms we get a formula

$f(z) = P(x,y) + iQ(x,y)$

where P and Q are polynomials in two real variables x, y. We can therefore think of

$|f(x+yi)| = (P(x,y)^2 + Q(x,y)^2)^{1/2}$

as a continuous function of two real variables. We want to apply lamma 6 to conclude that it takes on an absolute minimum. Thus, we need to understand what happens as (x,y) approaches infinity. But we have

$$|f(z)| = |z^n| + |a_n| \left| 1 + b_{\frac{n-1}{z}} + b_{\frac{n-2}{z^2}} + ... + b_{\frac{0}{z^n}} \right|,$$

where $b_i = a_i/a_n$ for $0 \leq i \leq n-1$. Now

$$\left| 1 + \left(b_{\frac{\{n-1\}}{z}} + b_{\frac{\{n-2\}}{z^2}} + ... + b_{\frac{\{0\}}{z^n}} \right) \right| \quad ...(A)$$

$$> |1| - \left| \left(b_{\frac{\{n-1\}}{z}} + b_{\frac{\{n-2\}}{z^2}} + b_{\frac{0}{z^n}} \right) \right|$$

The term that we are subtracting on the right is at most

$$\left| b_{\frac{\{n-1\}}{|z|}} \right| + \left| b_{\frac{\{n-2\}}{|z^2|}} \right| + ... + \left| b_{\frac{0}{|z^n|}} \right|,$$

and this approaches 0 at |z| approaches infinity. Hence, for all sufficiently large |z|, the quantity on the left in the inequality labeled (A) is at least 1/2, and so |f(z)| is at least $|z|^n \left| a_{\frac{n}{\frac{1}{2}}} \right|$ (1/2) for large |z|. Thus, |f(z)| approaches infinity as |z| approaches infinity. This implies, by lamma 6, that we can choose a point z = r = a + bi where |f(z)| is an absolute minimum. The rest of the argument is devoted to showing that f(r) must be zero. We assume otherwise and get a contradiction. To simplify calculations we are going to make several changes of variable. Simplification 1. Let g(z) = f(z+r), which is also a polynomial in z of degree n. g (resp. |g|) takes on the same set of values as f (resp. |f|). But |g| is minimum at z = 0, where the value is |f(0+r)| = |f(r)|. Thus, we may assume without loss of generality that |f| is minimum at z = 0. (We change back the notation and don't refer to g any more.) We are now assuming that $a_0 = f(0)$ is not 0. Let $a = |a_0|$. Simplification 2. Replace f by $(1/a_0)f$. All values of f are divided by a_0. All values of |f| are divided by a. The new function still has its minimum absolute value at z = 0. But now the minimum is 1. We still write f for the function. Thus, we can assume that f(0) is 1 (this means that $a_0 = 1$) and that 1 is the minimum of |f|. Simplification 3. We know that a_n is not 0. Let k be the least positive integer such that a_k is not 0. (It might be 1 or n.) We can write

$$f(z) = 1 + a_k z^k + ... + a_n z^n$$

We next observe that if we replace f(z) by f(cz) where c is a fixed nonzero complex number the set of values of f (and of |f|) does not change, nor does the constant term, and 0 = c(0) stays fixed. The new f has all the same properties as the old: its

absolute value is still minimum at z = 0. It makes life simplest if we choose c to be a k^{th} root of $(-1/a_k)$. The new f we get when we substitute cz for z then turns out to be $1 - z^k + a'_{k+1} z^{k+1} + ... + a'_n z^n$. Thus, there is no loss of generality in assuming that $a_k = -1$. Therefore, we may assume that

$$f(z) = 1 - z^k + a_{k+1} z^{k+1} + ... + a_n z^n.$$

If n = k then $f(z) = 1 - z^n$ and we are done, since f(1) = 0. Assume from here on that k is less than n. The main point. We are now ready to finish the proof. All we need to do is show with $f(z) = 1 - z^k + a_{k+1}z^{k+1} + ... + a_n z^n$ that the minimum absolute value of f is less than 1, contradicting the situation we have managed to set up by assuming that f(r) is not 0. We shall show that the absolute value is indeed less than one when z is a small positive real number (as we move in the complex plane away from the origin along the positive real axis or x-axis, the absolute value of f(z) drops from 1.) To see this assume that z is a positive real number with $0 < z < 1$. Note that 1-z is then positive. We can then write

$$|f(z)| = |1 - z^k + a_{k+1}z^{k+1} + ... + a_n z^n|$$

$$\leq |1-z^k| + |a_{k+1}z^k + ... + a_n z^n|$$

$$= 1 - z^k + |a_{k+1}z^{k+1} + ... + a_n z^n|$$

$$\leq 1 - z^k + |a_{k+1}|z^{k+1} + . . . + |a_n|z^n$$

(keep in mind that z is a positive real number)

$$= 1 - z^k (1 - w_z), \text{ where}$$

$$w_z = |a_{k+1}|z + ... + |a_n|z^{n-k}.$$

When z approaches 0 through positive values so does w_z. Hence, when z is a small positive real number, so is $z(1-w_z)$, and so for z a small positive real number we have that

$$0 < 1 - z^k(1-w_z) < 1.$$

Since $|f(z)| < 1 - z^k(1-w_z)$

it follows that |f(z)| takes on values smaller than 1, and so |f(z)| is not minimum at z = 0 after all. This is a contradiction. It follows that the minimum value of |f(z)| must be 0, and so f has a root .That proves the theorem.

12. Fundemental Theorem of Algebra via Fermat's Last Theorem

For the proof of the theorem we must know the following lemmas

Lemma 1:If an algebraic equation f(x) has a root α, then f(x) can be divided by x-α without a remainder and the degree of the result f'(x) is less than the degree of f(x).

Proof:
Let $f(x) = x^n + a_1 x^{n-1} + . . . + a_{n-1}x + a_n$
Let α be a root such that f(α) = 0
Now, if we divide the polynomial by (x-α), we get the following
$f(x)/(x - \alpha) = f_1(x) + R/(x-\alpha)$

where R is a constant and $f_1(x)$ is a polynomial with order n-1.
Multiplying both sides with x-α gives us:
$f(x) = (x - \alpha)f_1(x) + R$
Now, if we substitute α for x we get:
f(α) = 0 which means that the constant in the equation is 0 so R = 0. That proves the lemma.

Theorem: Fundamental Theorem of Algebra

For any polynomial equation of order n, there exist n roots r_i such that:
$xn + a_1 x^{n-1} + ... + a_{n-1} x + a_n = (x - r_1)(x - r_2) . . . (x - r_n)$
Proof: Let $f(x) = x_n + a_1 x^{n-1} + ... + a_{n-1}x + an$
We know that f(x) has at least one solution α_1.
Using Lemma 1 above, we know that:
$f(x)/(x - \alpha_1) = f'(x)$ where deg f'(x) = n-1.

So that we have:
$f(x) = (x - \alpha_1)f'(x)$
But we know $f'(x)$ has at least one solution α_2
$f'(x)/(x-\alpha_2) = f''(x)$ where deg $f'(x) = n-2$. And
$f(x) = (x - \alpha_1)(x - \alpha_2)f''(x)$
Eventually we get to the point where the degree of
$fn(x)=1$ .In this case, $f_n(x) = x - \alpha_n$. This establishes that there are n roots for a given equation $f(x)$ where the degree is n. Putting this all together gives us:
$f(x) = (x - \alpha_1)(x - \alpha_2). \ . \ . (x - \alpha_n)$
Now, since $f(x)=0$ only when one of the values $\alpha_i=x$, we see that the n roots α_i are the only solutions. So, we have proven that each equation is equal to n roots.One important point to remember is that the n roots are not necessarily distinct. That is, it is possible that $\alpha_i = \alpha_j$ where $i \neq j$. That proves the theorem

Fundamental theorem of Algebra due to Cauchy
We will prove
$f(z) = z^n + a_1z^{n-1} + a_2z^{n-1} + . \ . \ . + a_n = 0$
where a_i are complex numbers $n \geq 1$ has a complex root.

Proof:- let $a_n \neq 0$ denote $z = x + iy$ x, y real Then the function

$g(x,y) = |f(z)| = |f(x + iy)|$

Is defined and continuous in R^2
Let $c = \sum_{j=1}^{n} |a_j|$ it is +ve using the triangle inequality
We make the estimation

$$|f(z)| = |z|^n \left|1 + \frac{a_1}{z} + \frac{a2}{z^2} + \ldots + \frac{|a_n|}{z^n}\right|$$

$$|f(z)| \geq \left(1 - \frac{|a_1|}{|z|} - \frac{|a_2|}{|z|} + \ldots + \frac{|a_n|}{|z|^n}\right)$$

$$|f(z)| = |z|^n \left(1 - \frac{c}{|z|}\right) \geq \frac{1}{2}|z|^n$$

Being true for $|z| > \max(1, 2c)$ denote $r : \max\left(1, 2c, \sqrt[n]{2|a_n|}\right)$ conside the disk $x^2 + y^2 \leq r$ Because it is compact the function g(x,y) attains at a point $(x_0\ y_0)$ of the disk its absolute minimum value (infimum) in the disk if $|z| > r$ we have

$$g(x, y) = |f(z)| \geq \frac{1}{2}|z|^n > \frac{1}{2}r^n \geq \frac{1}{2}\left(\sqrt[n]{2|a_n|}\right) = |a_n| > 0$$

Thus

$$g(x, y) \leq g(0,0) = |a_n| < |f(z)| \ for |z| > r$$

Hence $g(x_0, y_0)$ is the absolute minimum of g(x,y) in the whole complex plane
we show that $g(x_0\ y_0) = 0$ therefore we make the antitheses that $g(x_0y_0) > 0$

Denote $z_o = x_o + iy_o = z_o + UB$

$$f(z) = f(z_o + u) = b_n + b_{n-1}U + \ldots + b_1U^{n-1} + U^n$$

Then $b_n = f(z_0) \neq 0$ by the antithesis

More over denote

$$c_j = \frac{b_j}{b_n}(i = 12\ldots n), c_0 = \frac{1}{b_n}$$

And assume that $c_{n-1} = c_{n-2} = \ldots = c_{n-k+1} = 0$
But $c_{n-k} \neq 0$ then we may write

$$f(z) = b_n(1 + c_{n-k} + c_{n-k}u + c_{n-k-1}u^{k+1} + \ldots + c_0u^k)$$

If $c_{n-k}u^k = p(\cos\alpha + i\sin\alpha)\ \&\ u = e^{(\cos\phi + \sin\phi)}$

Then

$$c_{n-k}u^k = pe^{n(\cos\alpha + k\phi) + i\sin(\alpha + k\phi)}$$

By Demorvie's identity choosing

$$\rho \leq 1\ \&\ \phi = \frac{\overline{\pi\alpha}}{k}$$ we get

$c_{n-k}u^k = -pe^k$ & can make the estimate

$$\left|c_{n-k-1}u^{k+1} + \ldots + c_0u^n\right| \leq \left|c_{n-k-1}\right|e^{k+1} + \ldots + \left|c_0\right|c^n$$

$$\leq \left(\left|c_{n-k-1}\right| + \ldots + \left|c_0\right|e^{k+1} = \mathrm{Re}^{k+1}\right)$$

Where R is constant

Let now $\rho = \min\left(1, \sqrt[k]{\frac{i}{p} \cdot \frac{p}{2R}}\right)$ we obtain

$$|f(z)| = |b_n||1 - pe^k + h(u)|$$

$$|f(z)| = |b_n||1 - pe^k + h(u)| +$$

$$f(z) \leq |b_n||1 - pe^k + \mathrm{Re}^{k+1}|$$

$$\leq b_n\left|1 - e^k\left(p - R\frac{p}{2R}\right)\right|$$

$$\leq \frac{|b_n|}{2} < |bn| = |f(z_0)|$$

Which results is impossible since $|f(z)|$ was absolute minimum. Consequently the antithesis is wrong & the proof is settled.

13. Another Proof Of Fundamental Theorem Of Algebra

The proof depends on the following four lemmas

Lemma1: Any odd-degree real polynomial must have a real root.
Proof:

We know from intermediate value theorem, suppose $p(x) \in R[x]$ with degree p(x) = 2k +1 and suppose the leading coefficient $a_n > 0$
(the proof is almost identical if $a_n > 0$). Then

$P(x) = a_n x^n$ + (lower terms)
And n is odd . Then,

(1) $\lim_{x\to\infty} p(x) = \lim_{x\to\infty} a_n x^n = \infty$ since $a_n > 0$.

(2) $\lim_{x\to-\infty} p(x) = \lim_{x\to-\infty} a_n x^n = -\infty$
since $a_n > 0$ and n is odd.
From (1) , p(x) gets arbitrarily large positively , so there exists an x_1 with $p(x_1) > 0$. Similarly, from (2) their exits an x_2 with $p(x_2) < 0$ A real polynomial is a continuous real – valued function for all $x \in R$, Since $p(x_1)\, p(x_2) < 0$, it follows by the intermediate value theorem that their exits an x_3 , between x_1 and x_2, such that $p(x_3) = 0$.

Lemma 2: Any degree two complex polynomials must have a complex root.

Proof: We know from consequence of the quadratic formula and of the fact that any complex number has a square root.
If $p(x) = ax^2 + bx + c$, $a \neq 0$, then the roots formally are

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

From DeMoivere's theorem every complex number has a square root, hence x_1 , x_2 exists in C. They of course may be same if $b^2 - 4ac = 0$.

Lemma 3: If every non-constant real polynomial has a complex root, then every non-constant complex polynomial has a complex root.
Proof: According to concept of the conjugate of complex polynomial
Let $P(x) \in R[x]$ and suppose that every non-constant real polynomial has at least one complex root. Let $H(x) = P(x)\overline{P}(x)$. from previous lemma $H(x) \in R[x]$. By supposition there exists a $z_0 \in C$ with $H(z_0) = 0$. then

$$P(z_0)\overline{P}(z_0) = 0,$$

And since C has no zero divisors, either

$$P(z_0) = 0 \text{ or } \overline{P}(z_0) = 0.$$

then previous lemmas

$$\overline{P}(z_0) = \overline{\overline{P}(\overline{z_0})} = P(\overline{z_0}) = 0.$$

Therefore, $\overline{z_0}$ is root of p(x).

Lemma 4:-
Any non-constant real polynomial has a complex root.

Proof: Let $f(x) = a_0 + a_1 x + \ldots + a_n x^n \in R[x]$
with $n \geq 1$, $a_n \neq 0$. The proof is an induction on the degree n of f(x). Suppose $n = 2^m q$ where q is odd. We do the induction on m. if m = 0 then f(x) has odd degree and the theorem is true from lemma 1. Assume then that theorem is true for all degrees $d = 2^k q'$ where $k > m$ and q' is odd . Now assume that the degree of f(x) is $n = 2^m q$. Suppose F' is the splitting field for f(x) over R in which the roots are $\alpha_1, \ldots, \alpha_n$. We show that at least one of these roots must be in C. (In fact, all are in C but to prove the lemma we need only show at least one.)
Let $h \in Z$ and from the polynomial

$$H(x) = \prod_{i<j} (x - (\alpha_i + \alpha j + h\alpha_i \alpha_j))$$

This is in F'[x] we chose pairs of roots $\{\alpha_i, \alpha_j\}$, so the number of such pairs is the number of ways of choosing two elements out of $n = 2^m q$ elements. This is given by

$$\frac{(2^m q)(2^m q - 1)}{2} = 2^m q(2^m q - 1) = 2^{m-1} q'$$

With q'odd. Therefore, the degree of H(x) is $2^{m-1}q'$.

H(x) is a symmetric polynomial in the root $\alpha_1,...,\alpha_n$. Since $\alpha_1,...,\alpha_n$ are the roots of a real polynomial, from lemma 3 any polynomial in the splitting field symmetric field symmetric in these roots must be a real polynomial.
Therefore, $H(x) \in R[x]$ with degree $2^{m-1}q'$.By the inductive hypothesis, then, H(x) must have a complex root. This implies that there exists a pair $\{\alpha_i, \alpha_j\}$ with

$$\alpha_i + \alpha_j + h\alpha_i \alpha j \in C$$

Since h was an arbitrary integer , for any integer h_1 there must exists such a pair $\{\alpha_i, \alpha_j\}$ with

$$\alpha_i + \alpha_j + h\alpha_i \alpha j \in C$$

Now let h_1 vary over the integers. Since there are only finitely many such pairs $\{\alpha_i, \alpha_j\}$, it follows that there must be least two different integers h_1 , h_2 such that

$$z_1 = \alpha_i + \alpha_j + h_1\alpha_i\alpha_j \in C$$

and

$$z_2 = \alpha_i + \alpha_j + h_2\alpha_i\alpha_j \in C$$

Then $z_1 - z_2 = (h_1 - h_2)\alpha_i\alpha_j \in C$

and since $h_1, h_2 \in Z \subset C$ it follows that $\alpha_i\alpha_j \in C$. But then $h_1\alpha_i\alpha_j \in C$, from which it follows that $\alpha_i\alpha_j \in C$.

Then,

$$P(x) = (x - \alpha_i)(x - \alpha_j) = x^2 - (\alpha_i + \alpha_j)x + \alpha_i\alpha_j \in C[x]$$

However, P(x) is then a degree–two complex polynomial and so from lemma 2 its roots are complex. Therefore, $\alpha_i\alpha_j \in C$ and therefore f(x) has a complex root.

It is now easy to give a proof of the Fundamental Theorem of Algebra. From lemma 4 every non constant real polynomial has a complex root. From lemma 3 if every non constant real polynomial has a complex root, then every non- constant complex polynomial has a complex root providing the Fundamental Theorem

REFERNCES

[1]. Open Mappings and the Fundamental Theorem of Algebra R. L. Thompson Mathematics Magazine, Vol. 43, No. 1.
[2]. (Jan., 1970), pp. 39-40.
[3]. What! Another Note Just on The Fundemental Theorem of Algebra ? R. M. Redheffer The American Mathematical
[4]. Monthly, Vol. 71, No. 2. (Feb., 1964), pp. 180-185
[5]. Another Proof of the Fundamental Theorem of Algebra Daniel J. Velleman Mathematics Magazine, Vol.70, No. 3. (Jun.,1997), pp.282-293.
[6]. [Euler and the Fundamental Theorem of Algebra William Dunham The College Mathematics Journal, Vol.22, No. 4 (Sep., 1991), pp.282-293.
[7]. [An Elementary Constructive Proof of the Fundamental Theorem of Algebra P.C. Rosenbloom The American Mathematical Monthly, Vol. 52, No. 10. (Dec., 1945), pp.562-570.
[8]. Proof of the Fundamental Theorem of Algebra J.L. Brenner; R.C. Lyndon The American Mathematical Monthly, Vol. 88, No.4. (Apr., 1981), pp.253-256.

[9]. [Yet Another Proof of the Fundamental Theorem of Algebra R.P. Boas,Jr. The American Mathematical Monthly, Vol. 71, No. 2. (Feb., 1964), p.180.
[10]. The Fundamental Theorem of Algebra Raymond Redheffer The American Mathematical Monthly, Vol. 64,
[11]. No. 8 (Oct.,1957), pp.582-585.
[12]. A Proof of the Fundamental Theorem of Algebra R.P.Boas, Jr. The American Mathematical Monthly,
[13]. Vol. 42, No. 8.(Oct.,1935), pp.501-502
[14]. An Easy Proof of the Fundamental Theorem of Algebra Charles Fefferman The American Mathematical
[15]. Monthly, Vol. 74, No. 7. (Aug.- Sep., 1967), pp.854-855[11] Eulers: The Masster of Us All William Dunham
[16]. The Mathematica Association of America Dolciani Mathematical Expositions No.22, 1999.
[17]. The Fundamental Theorem of Algebra Benjamin Fine Gerhard Rosenberger Springer New York
[18]. Mathematical Analysis S.C Malik, Savita Arora New Age International Publication
[19]. Functions of a Complex Variable Goyal & Gupta Pragati Prakajhan Publication
[20]. A.R. Schep, A Simple Complex Analysis and an Advanced Calculus Proof of the FundamentalTheorem
[21]. of Algebra, Am Math Monthly 116, Jan 2009,67-67
[22]. David Antin's translation of Heinrich Dorrie's 100 Great Problems of Elementary
[23]. A. R. Schep, A Simple Complex Analysis and an Advanced Calculus Proof of the Fundamental Theorem of
[24]. Algebra, Am Math Monthly 116, Jan 2009, 67-68.
[25]. Heinrich Dorrie, 100 Great Problems of Elementary Mathematics.
[26]. Complex Analysis Kunihiko Kodaira Cambridge Studies in advanced mathematics
[27]. Elementary Theory of Analytic Functions of one or Several Complex Variables Henri Cartan
[28]. Complex variables An introduction Carlos A. Berenstein Roger gar Springer-Verlag
[29]. Fundamental Theorems of Operator Algebra V.1 Elementary Theorem V.2 Advanced Theorem Kadison- Richard.
[30]. Fundamental Theorem of Algebra Fine and Gerhard
[31]. Fundamentals of College algebraic Miller
[32]. Fundamental concept of higher Algebra Albert.
[33]. Fundamental concept of Algebra Meserve.
[34]. Fundamental Concept of Algebra Chevalley. C
[35]. Fundemental Theorey of One Complex Variable Green and Krantz
[36]. Function of complex variable Franklin
[37]. Function of Complex Variable with application Phillips vig Theory & Technique carrier range.
[38]. Function of a Complex variable and some of their applications Fuchs and Shabat Smitnov & Lebeder
[39]. Function of complex variable Goursat E.
[40]. Function of complex variable Conway J.B
[41]. Function of one complex variable Macrobart TM
[42]. Foundation of the Theorem of Algebras Volume I & Volume II Hankoock. Harris
[43]. Complex Analysis Freitag Busam
[44]. Function of Several variables lang.

Promoting a culture of health and safety at work in cement plants

Taleb Mounia[1], Chaib Rachid[2], Chetouani Yahyia[3]

[1]Electromechanical mining engineering department, Larbi Tebbessi University, Tébessa, Algeria
[2] Transport engineering and environment laboratory, Mentouri University, Constantine, Algeria
[3] Chemical engineering department, Rouen University, France

Abstract:

Safety is a priority of any industrial activity. It is a positive cultural element that allows other improvements in the factory. An administration that does not attain to manage safety is not in a position to manage other functions. However, as work accidents and occupational diseases have an enormous impact on the health of workers and considerable economic and social impacts. In addition, with the increasing complexity of industrial tissue and with the rapidity that the techniques develop in the big factories, risks assessment becomes a crucial and strategic answer to preserve workers health and safety on the one hand and to maintaining a qualified labour on the other hand. These are data, among others, that have triggered the alarm signal and impose the necessity of an increased safety in the factories. Therefore, a priori assessment of these risks and the implementation of a prevention approach within a factory is required to become one of the main drivers of progress. Hence, for some employers, employees and their representatives, health and safety at work do not mean so much. In addition, with the permanent evolution of work, even its risks, it becomes increasingly insufficient to establish general safety rules of, relying solely upon standards and regulations to comply **[1]**, but move to awareness, information, training and motivation of staff on the role of health and safety at work, steps previously required for the implementation of a prevention, even to a mitigation measures relevant and effective. That allows to define a general policy of prevention and to bring to successful management of industrial risk within the entity. Hence, it has become essential to give all staff a real sense of safety that will predict and act in very affective way; objective of this work.

This article presents a technique of analysis to better understand the dynamic of the policy in terms of health and safety at work established in the cement plants.

Keywords: health and safety at work, promote health, evaluation, accidents, prevention, sensitization.

1. Introduction

The policy in terms of health and safety at work is not only a matter of laws and regulations. These are essential and must be strictly applied at each company, even at each workplace. However, with labour market rapidly changing, all the work accidents cannot be avoided. Henceforth, with the intensity of human activity, its ever-increasing pace and the permanent intervention of man, the risk is always present. Moreover, there are health and security at work that do not mean so much. But to get a measurable improvement in working conditions and reduction of occupational accidents and occupational diseases, it is necessary to unite them at a range of other instruments such as information, sensitization, training, social dialogue, good practices, social responsibility of companies, economical incitements and integration into a process of continuous improvement at the HSW. The daily drama of work accidents calls an energetic reaction of all the concerned actors, as much as level of factories state and workers. Today, the programs are more focused on the emergence of new risks related to changes in the world of work and society. Now, the safety and health at work took the community strategies form. Despite the progress in this field, the current community strategy aims to reduce by 25% work accidents across the UE union by 2012.To attain this objective, all actors are invited to take action at all levels-European, national, local and workplace.

Therefore, taking into account health and security at work becomes a national objective to achieve in order to preserve industrial performance. However, an industrial company needs to rely upon all its strength to not only improve performance, but to be the best in its field, keeping in that any change in culture takes time and must consider the particularity of the factory [2]. Thus, achieving to a continuous reduction, sustainable and homogeneous work accidents and occupational diseases is an imposed necessity. That is why it is urgent to deal with the problem from all sides in relation with the HSW. Thus, everything has to be done to avoid any accident that breaks or handicaps a human life; objective of this work

2. Causes of accidents at work

The development and implementation of coherent national strategies are a major concept and a fundamental pillar of the community strategy objectives. The passage to new strategies based on the information is a global phenomenon. In 2006, the ILO adopted its promotional framework for safety and health at work and the WHO adopted a global action plan 2008-2017 on health workers.

In addition, many countries do not belong to the UE have developed SHW strategies more clear and comprehensive efforts and set priorities for safety and health at work, case of our country. However, it is very difficult to convince the employers and the decision makers that improving working conditions can be profitable **[3**] and that the improvement of health and safety at work can generate an enormous economical benefits, not only for the factories, but also for society as a whole **[3,4]**.

That is why, to engage in a process of a continuous improvement in a subject of health and safety at work, the evaluation is required [**5**]. Causes of accidents recorded in **07** cement plants are presented as follows and often linked to:

- Lack of experience and motivation;
- The high turnover of workers at the workplace;
- Ephemeral labor relationships;
- Complex situation on workplace;
- Lack of communication, information and awareness;
- Insufficient practices;
- Insufficient safety of a machine;
- Some careless;
- High pace of work;
- Stress;
- Total absence of preventive safety strategy;
- Lack of clear strategy on health and safety at work;
- Lack of awareness of responsibilities;
- Lack of means and resources to deploy and to support solutions in health and safety at work;
- Insufficient involvement of employees and of general direction;
- Questioning of the existing organizational processes;
- An underestimation of the consequences less-evaluation of occupational injury (WA and OD);
- Weak awareness of the perspectives offered of health and safety at work;
- Lack of human resources, financial resources and information to manage health and safety at work;
- Burn-in presence or interior resignation **[6]**, in all sectors: body be present at work, but absent mentally or physically ill;
- Some people insist on working even if they are not alright and even when a doctor recommends a few days off;
- Lack of hygiene, organization and work atmosphere;
- Potential benefits of better management of health and safety at work are often unrecognized.

3. The work objective

Conscious of the importance of industrial safety control for the future of the cement industry in our country and to promote a culture of health and safety at work, it is hoped to initiate a deep collective reflection on the entire issues raised by technological risks. The majority of our cement plants have very limited human resources, financial resources and information to manage health security in their community. In addition, they are often difficult to join by organizations that may offer support on the one hand and taking awareness of the problem of ill-being at work by the public opinion across mediated cases on the other hand. These catastrophes have emerged the ill-being at work as a society fact and a stake of public debate. Now, the best way to save is to avoid accidents. How? By supporting health and safety at work in the factory with workers concert. And if there are accidents, the best way to minimize expenses is to ensure that workers can quickly return and in a sustainable way to work. For that, the factory should identify and assess all potential risks associated with the occupation **[7]**. This is why our goal is to promote a culture of health and safety at work, even to energize a policy to attain as quickly as possible a remarkable decrease and a measurable improvement in working conditions (reduction of occupational accidents and occupational diseases), primarily in the sectors most exposed: the trend towards zero accident and thus empower and establish a policy health and safety at each cement plant. The problem is primarily cultural: a collective awareness that too often depends on crisis situations is needed. Therefore, this paper orients the reader to progressively implementing a policy of prevention, based mainly on awareness, information, motivation and regulation. It should allow factories managers, employees and their representatives **[8]** to:

- Attain the better control of occupational risks;
- Respect the physical integrity of each, with a proper evaluation and implementation of preventive measures, taking into accounts the work organization and its technical and human components;
- Help respecting the legislation in effect;
- Improve working conditions and the employee's welfare;
- Engage in a process of continuous improvement in health and safety at work;
- Widely, improve work conditions in a continuous and sustainable development and a sustainable culture of prevention at the entity.

The reasons of this ambition are multiples:

- ✓ The well-being of employees, health and security at work are foremost priorities to protect workers. This objective goes beyond all others and explain alone the urgency of mobilization;
- ✓ The life and health of workers and their welfare at work also bring more value to society as a whole and at each factory in particular;
- ✓ Indeed, in a secondary way, economically it is also a matter of economic justice, equality and fair lawful competition. A healthy workplace can improve public health in general as well as productivity and competitiveness of factories.

Finally, issues of health and safety at work cost very expensive to workers, to factories and to society as a whole. More good practical working conditions enhance workers productivity and improve the quality of goods and services..

4. Methodology

The analysis of event data can be extremely useful for understanding where, how, and when workplace accidents occur. It can also help the factory employees to determine priority actions and the best time to intervene. In other words, it contributes to the process of building a strategy for managing performance in health and safety at work. Performance required in this area and economic issues associated will lead safety services to be interested in work accidents, their consequences al stakes bring security services to become interested in works accidents, and organizations they result. These concerns are expressed and treated through the envisaged concepts and methods. This article presents an analysis technique to better understand the dynamics of health policy and safety at work established within cement plants. This is a census of the different causes of work accidents listed in 07 cement plants. These causes have been formulated over a questionnaire form filled up at least by 100 workers and their representatives in each cement plant. An initial study at the Elma Labiod cement plant gave the following results, table 1.

Table 1. Personal reactions on the accidents causes in the entity.

Number of causes	Yes	No
1	96	4
2	83	17
3	56	44
4	68	32
5	85	15
6	68	32
7	70	30
8	96	4
9	70	30
10	94	6
11	80	20
12	67	33
13	73	27
14	84	16
15	68	32
16	51	49
17	41	59
18	62	38
19	79	21
20	78	22
21	81	19
22	72	28
23	70	30
24	14	86

Notice: According to the result of survey carried out at the cement factory of Elma Labiod, it can be noticed that the factory is far from the regulations and standards in health and safety at work and that much remains to be done to develop a risk culture for all and ensure a level of acceptable safety. Now, well-being at work is everyone's business. Health and safety at work should not be restricted to a few companies at the forefront of willing social progress or whose leaders are particularly sensitive to the issue. In close collaboration; the public authority and all field players must perform the daily work of information and explanation of risk prevention measures.

These can only be successful if everyone is widely and continuously involved. Therefore, it has become essential to give all staff a real sense of safety that will anticipate and act very effectively. To carry out this campaign, a methodology of work consists of 7 stages is proposed, namely:

1. Consciousness and sensitization of the industrial safety concept;
2. Return of experience;
3. Draw up a map of accidents;

4. Release priority actions of the appropriate prevention;
5. Seek appropriate solutions by using the Ishikawa diagram (7M);
6. Establish an action plan for implementation;
7. Evaluation of recommended results.

5. Conclusion

Safety is a priority of any industrial activity. It is a positive cultural element that allows further improvements in the factory. Hence, the well-being at work is everyone's business [9]. That is why, if a better protection of employees is to be attained, a change of perspective is needed, who should know what is happening and can happen to organize their lives and have good reflexes in case of emergency. Once, the responsibilities are clarified and understood by all, many improvements would be so easier to achieve. In fact, it is to increase transparency and develop a common language and common vocabulary for professionals and their representatives on matters relating to risk all partners and establish and disseminate this communication tools. As a result, many university officials, teachers and industrial representatives have understood the need necessity and the opportunity to teach general safety issues and with its specific rules.

References

[1] R. Chaib, M. Benidir and I. Verzea, Risks evaluation in a company: Case of engines and tractors factory, World Journal of Engineering. Vol.3, N°.4 (2006), Pp 88-94. Sun Light Publishing, Canada. ISSN 1706-5284.

[2] Jean Bufferne, La TPM : un système de production, Techno méca, Technologie 155, Avril 2008, Pp24-31.

[3] Jos C.M. Mossink, Pourquoi et comment procédés à des évaluations économiques au niveau de l'entreprise ? Série protection de la santé des travailleurs N°2, Bibliothèque de l'OMS : Publié sous la direction de Deborak Imed Nelson, 2004, ISSN 1729-3502.

[4] Rachid Chaïb, Ion Verzea, Mohamed Benidir and Ahmed Bellaouar; for sustainable development as regards prevention, safety and health at work in a company; *international symposium in knitting and apparel – iska 2010, iasi, november 19 – 20, 2010.*

[5] ACFCI, Guide pour la mise en place par étapes d'un système de management de la santé et de la sécurité au travail, I.S.B.N. 2-85723-466-X, Dépôt légal : Octobre 2007 ;

[6] Monique Legault Faucher, Vous avez dit présentéisme ? Prévention au travail Hiver 2009, Pp 34-38.

[7] S. Auduberteau et K. Ganino; La prévention des risques professionnels : hygiène et sécurité au travail, collection "les diagnostics de l'emploi territorial" hors série N°5, Direction Recrutement et Protection Sociale, octobre 2003.

[8] Bellilet. G et Benyounes. M, *Contribution à l'amélioration de la sécurité industrielle dans les entreprises*, PFE (DEUA) université de Constantine, Département de génie mécanique, juillet 2006, 136p.

[9] Aour. S, Laacheche. A et Meloul. N, *Contribution à l'optimisation de la sécurité industrielle dans l'entreprise*, PFE (DEUA) université de Constantine, Département de génie mécanique, juillet 2007, 128p.

Annexes

Relative questionnaire of work accidents in a cement plant: Elma Labiod cement plant, Tebessa, Algeria.

1- Lack of experience and motivation ;
2- Important rotation of workers and work ;
3- Ephemeral work relationships;
4- Complex situation on work site ;
5- Lack of communication, information and sensitization ;
6- Insufficient practices;
7- Insufficient safety of machine ;
8- Certain rashness ;
9- Well work pace;
10- au stress,
11- Total absence of safety prevention strategy ;
12- Lack of clear strategy in a subject of health and security at work;
13- Lack of responsability conscience ;
14- Insufficient means and resources to display and live the solutions of health and security at work;
15- Insufficient implication of workers and general direction ;
16- A discount in question of existing organizational processes;
17- Consequences less-evaluation of d'une lésion Professional injury (AT and MP) ;
18- Weak conscience of given perspectives of health and security at work;
19- Few human resources, financial means and information to manage health and security at work;
20- Burn-in Presence or interior resignation at level of all activity sectors : be body present at work, but absent minded or be sick physically;
21- Some people insist to work even when it does not work, even when a doctor recommended some rest days ;
22- Lack of hygiene, organization and work atmosphere;
23- Potential advantages often unrecognized of better management of health and security at work ;
24- Other causes.

Note: This questionnaire has to be filled up at least by 70% of workers in each cement plant.

Performance Evaluation of Energy Traffic In Ipv6 Networks

Dharam Vir[1], S.K.Agarwal[2], S.A.Imam3

[1] Head of Section, Department of Electronics Engg., YMCA University of Science & Technology, Faridabad India,
[2] Professor, Department of Electronics Engineering, YMCA University of Science & Technology, Faridabad India,
[3] Assistant Professor, Department of Electronics & Communication Engineering, Jamia Millia Islamia, New Delhi.

Abstract:

In this paper, we present a study of energy traffic based simulative and analytical methods in IPv6 networks. This research examine to find out which MANET routing protocol performs better in the case of TCP/IP (Application and Physical layer) under congested IPv6 networks. We investigates & undertakes simulation based study of Ad-hoc routing protocols in wireless sensor Network. We compare the five MANET routing protocols AODV, DYMO, Olsrv2 Niigata, OLSR Inria and RIPng with varying network nodes and fixed random waypoint mobility model using QualNet 5.0.1 Simulator. The metrics used for performance evaluation in TCP/IP application layer are Throughput, Average Jitter, End-to End delay, Total packets received / efficiency. In addition, the energy traffic model in the physical layer we simulate Total energy consumed in transmit mode, Total energy consumed in received mode and Total energy consumed in ideal mode in Ipv6 networks. The simulation has been carried out using QualNet 5.0.1 which is scalable network simulator. Finally results obtained by scrutinized from different scenarios to provide qualitative evaluation of the protocols.

Keywords: AODV, DYMO, Energy Traffic, IPv6, Olsrv2 Niigata, OLSR Inria, RIPng , QualNet 5.0.1

1. Introduction

A MANET [1] [2] [9] consists of only mobile nodes with wireless interfaces and provides wireless lattice connectivity among them. Each node can communicate with each other directly when the two nodes are in transmission range. When the two nodes are not in transmission range, the MANET routing protocol automatically selects the next hop node to the destination node. We can introduce IPv6; the next generation internet protocol was developed as a successor to IPv4 to increase the scalability of the internet. The IPv6 protocol was developed to solve the IPv4 address exhaustion problem, so it expands the IP address space from 32 to 128 bit. Also IPv6 increases the Minimum Transmission Unit (MTU) requirement from 576 to 1,280 bytes considering the growth in link bandwidth [10] [13]. IPv6 was developed by the IETF to overcome the inadequacy of IPv4. The 128 bit address space of IPv6 is beyond anyone's imagination. According to Beijnum (2006) it is, "340,282,366,920,938,463,463,374,607,431,768,211,456" for IPv6 while there is only "4,294,967,296" possible addresses for IPv4. IPv6 was designed not only to increase the address space, but also includes unique benefits such as scalability, security, simple routing capability, easier configuration "plug and play", support for real-time data and improved mobility support. IPv6 has full support for IPSec, and IPv6 is more secure when compared to IPv4. The processing of an IPv6 packet will be more efficient than an IPv4 packet. However, that is not the only enhancement that comes with IPv6. Following is an outline of some efficiency enhancements that IPv6 brings [4]:

- IPv6 header has a fixed length
- IPv6 header is optimized for processing up to 64 bits at a time (32 in IPv4)
- IPv4 header checksum that is calculated every time a packet passes a router was removed from IPv6
- Routers are no longer required to fragment oversized packets; they can simply signal the source to send smaller packets
- All broadcasts for discovery functions were replaced by multicasts.

1.1. RANDOM WAYPOIINT MOBILITY MODEL:

Mobility models are used for simulation purposes when new network protocols are evaluated [3] [9]. The Random waypoint model is a random mobility model used to describe the movement of mobile users, and how their location changes with time. It is one of the most popular mobility model to evaluate Mobile ad hoc network (MANET) routing protocols, because of its simplicity and wide availability. Using this model, the mobile nodes move randomly and freely without any restriction i.e. the destination, direction and speed of all chosen randomly and independently of all other nodes.

1.2. Energy Traffic Model:

The Battery power consumption of the mobile devices depends on the operating mode of its wireless network interfaces. Considering a broadcast transmission between the nodes of the active network, then wireless interfaces can be assumed to be in any of the following operating modes: [6] [11] [12]

- Transmit: source to destination node packet transmitting,
- Receive: source to destination nodes packets received,
- Idle: the node is ready to transmit or receive packets,

Sleep: it is the low power consumption mode state when a node cannot transmit or receive until woken up. The rest of the paper is organized as follows; in section 2, MANET Routing Protocols and their detail steps to design and implementing a network model using QualNet. Section 3 Mobility and Energy Traffic, QualNet designed scenario discussed in section 4. and also describes how the statistics in QualNet was collected. Section 5 describes the simulation results followed by section 4. Finally section 5 concludes the research work with possible future work.

2. Manet Routing Protocols:

The routing of the information is the most challenging task due to the inherent characteristics of the wireless sensor networks like dense deployment, mobility of nodes and energy constraint. The major issues related to this are: maximizing network lifetime, minimum latency, resource awareness, topological changes, location awareness and scalability. We are taking five routing protocols such as AODV, DYMO, OLSRv2-Niigata, OLSR-Inria and RIPng for our simulation and evaluation comparison [14].

2.1. Ad-Hoc On Demand Distance Vector Protocol (AODV):

The Ad hoc On Demand Distance Vector (AODV) [7] [8] is a routing protocol which is designed for ad hoc mobile networks. AODV is capable of both multicast as well as unicast routing. It builds and maintains routes between source nodes to desired destination nodes. AODV consists of a routing table which contains next hop information with sequence number.

The protocol consists of two processes:

(i) Route discovery
(ii) Route maintenance

In route discovery process a source node broadcasts a route request (RREQ) packet across the network. While this Route Request packet propagates in the network, a reverse route to the source is established along the way. RREQ packet contains the source node's IP address, current sequence number, broadcast ID and the most recent sequence number for the destination of which the source node is aware. When this packet reaches the destination (or a node having route to the destination), a Route Reply packet is sent, in unicast, to the source node using this reverse path [9] [5]. The maintenance of routes is done only for the dynamic routes. A destination node after receiving the RREQ may send a route reply (RREP) reverse to the source node. The source node receives the RREP, and begins to forward data packets to the destination. A route is considered active as long as there are data packets intermittently travelling from the source node to the destination node along that path. Once the source stops sending data packets, the links will time out and ultimately be deleted from the intermediate node routing tables. In route maintenance process if a link breaks occurs while the route is active; the node upstream of the breaking link propagates a route error (RERR) message to the source node to inform it of the now unreachable destinations. After receiving the RERR, if the source node still requests the route, it can reinitiate route discovery.

2.2. Dynamic MANET On-demand (DYMO):

The DYMO [2] [6] [14] routing protocol enables reactive multihop unicast routing between source node to participating destination nodes. The working of DYMO is similar to AODV with small modification. The protocol also consists of route discovery and route maintenance process. During route discovery, the source node initiates broadcasting of a Route Request (RREQ) throughout the network to find a route to the destination nodes. During this hop by-hop dissemination process, each intermediate node records a route to the source nodes. When the destination node receives the RREQ, it responds with a Route Reply (RREP) sent hop-by-hop (multihop) toward the source node. Each intermediate node that receives the RREP creates a route to the target, and then the RREP is unicast hop-by- hop toward the source. When the source node receives the RREP, routes have been established between the source node and destination node In route maintenance process this protocol made two operations. In order to shield routes in use, node extends route life times upon successfully forwarding a packet. In order to reply to changes in the network topology, DYMO routers examine links over which traffic is flowing. When a data packet is received and a route for the destination node is not known or the route is broken down, then the DYMO source router is notified. A Route Error (RERR) is sent toward the source to indicate the current route to a particular destination is invalid or missing. When the source receives the RERR, it deletes the route, than the source node later receives a packet for forwarding to the same inference, it will need to perform route discovery once more for that destination.

2.3. Optimised Link State Routing Protocol (OLSR Inria)

The Optimised Link Sate Routing Protocol (OLSR Inria) [7] [13] [14] is supports the large, dense mobile networks, with high nodal mobility and topological changes. It uses periodic messages to update the topological information of the network among the respective nodes. It uses the concept of multi-point relays to calculate the route towards any source to destination in the network. The multi-point relays provide the optimal routes, and due the pro-active nature of the protocol based on link state algorithm. OLSR Inria is an optimization over a pure link state protocol as it squeezes the size of information send in the messages, and reduces the number of retransmissions. It provides optimal routes in terms of number of hops. OLSR Inria is particularly suitable for large and dense networks [12]. The functioning of the OLSR Inria protocol is based on periodically diffusing a topology control packet in the network. In OLSR Inria each node uses the most recent information to route a packet. Each node in the network selects a set of nodes in its neighborhood, which retransmits its packets. This set of selected neighbor nodes is called the multipoint relays (MPR) of that node. The neighbors that do not belong to MPR set read and process the packet but do not retransmit the broadcast packet received form node. For this purpose each node maintains a set of its neighbors, which are called the MPR Selectors of that node.

2.4. Optimized Link State Routing protocol v_2 Niigata (OLSRv2 Niigata)

OLSRv2-Niigata also supports the QualNet simulator [8]. But two features have not been yet implemented; OLSR packet fragmentation, and multiple addresses and multiple interfaces handling.

2.5. Routing Information Protocol next generation (RIPng)

RIPng is a proactive Interior Gateway Protocol based on the distance-vector algorithm [15]. RIPng is intended for use within the IPv6-based Internet. As it is a distance-vector routing protocol, it forms routing tables by exchanging routing table information with each router. There are two types of updates. One is a Regular update, which is periodically sent and contains the whole routing table information. The other is a Triggered update, which is sent when a router's routing table changes and contains only those routing entities which have been modified. When a router receives a packet, it updates its routing table and if its routing table has changed, it sends a triggered update to its neighbor router.

3. Simulation Scenarios:

We have using the QualNet 5.0.1 simulator for our analytical evaluation. In our simulation model, nodes are placed randomly within a 1500m x 1500m physical terrain area so that the average node degree for 10-100 nodes is respectively. In this scenario wireless connection of varying network size (100 nodes) for MANET is used for analytical comparison performance of routing protocol AODV, DYMO, OLSRv2-Niigata, OLSR-Inria and RIPng over it data traffic of Constant Bit Rate (CBR) is applied between source and destination. The nodes are placed randomly over the region of 1500m x 1500m. The network of size 100 nodes. The Qualnet5.0.1 simulator network simulator is used to analyze the parametric performance of all protocols defined above. We choose a square area in order to allow nodes to move more freely with equal node density. We have tested five different routing protocols and no. of different scenarios characterized by different network conditions. Each data point in the simulation graphs represent an average value obtained from 10 randomized simulation runs. The basic scenarios parameters are listed in table 1. The table 1 parameters implementing in the simulator then analyze the performance of AODV, DYMO, OLSRv2-Niigata, OLSR-Inria and RIPng routing protocols. The animated simulations of network size 100 are shown in Figure 1.The performance is analyzed with varying network size keeping energy traffic load and random way point mobility constant. The metrics are used to study the protocols Average Jitter, Throughput , Average End to End delay, percentage efficiency of total Packet received , Energy consumed in transmit mode, Energy consumed in receive mode, and Energy consumed in Ideal mode. The results are shown in from Figure 2 to Figure 8. We evaluate the performances metrics in Application and Physical layers of designed scenarios. The performance matrices are given below:

- Throughput
- Average Jitter
- End-to-End Delay
- Total Packet Received / Efficiency
- Energy Consumed in Transmit mode
- Energy Consumed in Receive Mode
- Energy Consumed in Idle Mode

Table 1. Simulation Parameters for Energy Based Performance Analysis of AODV, DYMO, OLSRv2-Niigata, OLSR-Inria and RIPng Routing Protocols

Simulator Parameters	
Mac Type	IEEE 802.11
Protocols under studied	AODV, DYMO, OLSRv2-Niigata, OLSR-Inria, RIPng
Transmission range	600m
Node movement model	Random way point, 0-5m/s, pause time 0s
Traffic type	CBR
Antenna	Omni directional
Node Speed	10m/s, 20m/s, 50m/s, 100m/s
Propagation model	Two Ray Ground
Channel Frequency	2.4 GHz
Network Protocols	IPv6
Scenario Parameters	
Number of nodes	10 to 100
Topology area	1500x1500
Packet size	512
Item to send	100
Simulation time	30 Seconds
Battery Charge Monitoring Interval	60 Sec.
Full Battery Capacity	1200 (mA,h)
Performance Matrices in Application Layer	Average Jitter, End to End Delay, Throughput, Total Packet received
Performance Matrices in Physical Layer	Energy consumed (in mjules) in transmit mode Energy consumed (in mjules) in received mode Energy Consumed (in mjules) in ideal mode
Energy model Parameters	
Energy Model	Mica motes
Energy Supply Voltage	6.5 Volt
Transmit Circuitry Power Consumption	100.0 mW

Receive Circuitry Power Consumption	130.0 mW
Idle Circuitry Power Consumption	120.0 mW
Sleep Circuitry Power Consumption	0.0 mW
Energy Model Specifications	
Initial Energy (Joules)	15
Transmission Power (Watt)	1.4
Receiving Power (Watt)	1.0
Idle Power (Watt)	0.0

Scenario designed for AODV, DYMO, OLSRv2-Niigata, OLSR-Inria, RIPng with Varying Network Size (10-100). The parameters of Table 1 deployed in QualNet simulator assigned in 10-100 nodes sources are randomly distributed over a 1500m x 1500m area. The maximum speed varies from 1 to 21 m/s. Pause time is set at 5 sec. Consequently, most nodes move at all times.

3.1. Snapshot of Simulation

The simulations of energy traffic model were performed using QualNet Simulator 5.0.1, the traffic sources are CBR (continuous bit rate). The source-destination pairs are multiplying randomly over the network. During the simulation, each node starts its journey from a source node to destination node. This process repeats throughout the simulation, causing continuous changes in the topology of the underlying network. Fig.1 Shows the running simulation of snapshot when we applying CBR (1 - 20) nodes and DYMO routing protocol.

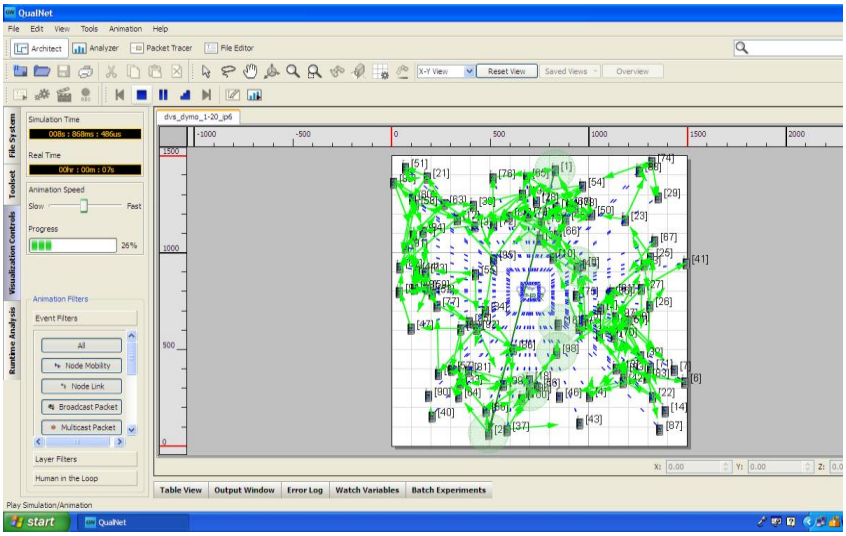


Figure1. Snapshot of QualNet Animator in action for applying DYMO protocol using 100 nodes.

4. Results & Analysis

4.1. Analysis and impact of Throughput (Bits/s):

The throughput of the protocols can be defined as successful average rate of data packets received at its destination among the packets sent by the source. Throughput of all protocols decreases when the size of network increases. The throughput is measured in bits per second (bit/s or bps). For better system performance the number of bits per sec must be high.

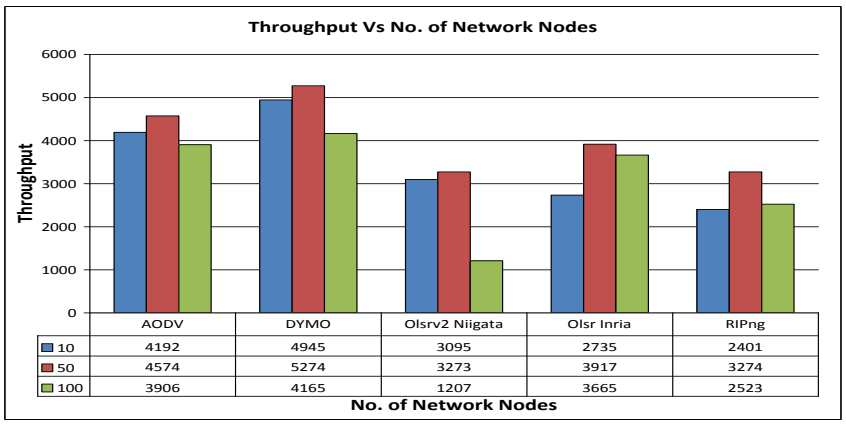


Figure.2 shows the impact variation of throughput for various routing protocols which considered for Ipv6 as parameter in application layer.

It has observed in Fig.2 that the throughput of DYMO is better than AODV & OLSRv2-Niigata, OLSR-Inria and RIPng whereas the performance of DYMO is better than others.

- DYMO, AODV, Olsr Inria, Olsrv2 Niigata, And Ripng Are Having Minor Degradation.
- By Observation The Throughput Is Maximum For DYMO Which Is Respectively By AODV, Olsr Inria, Olsrv2 Niigata, And RIPng for Ipv6. RIPng gives the minimum throughput for Ipv6 network.

4.2. Analysis And Impact Of Average Jitter (S):

The discrepancy in Jitter which is caused due to obstruction by network, timing drift, route changes, topology change etc. in a network. Low value of jitter provides the better performance of any protocol. This includes all possible delays caused by buffering during route discovery.

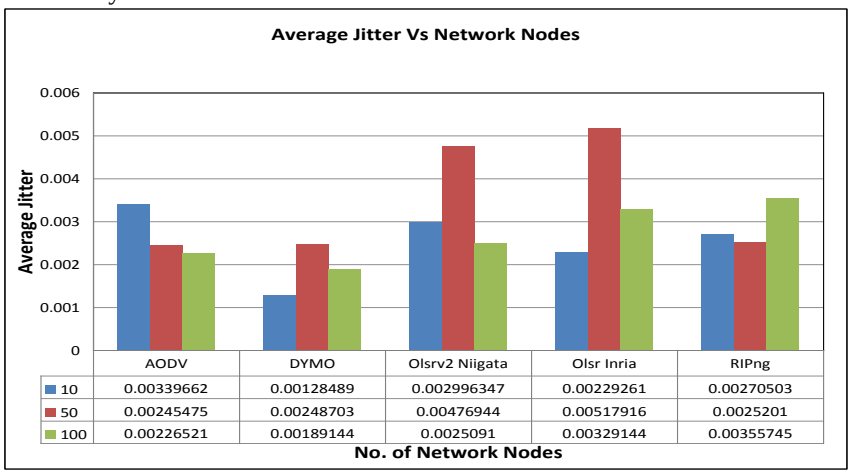


Figure 3. Shows the impact variation of average jitter for various routing protocols which considered for Ipv6 as parameter in application layer.

- DYMO shows the constant least jitter when mobility is restricted to only 60 nodes.
- By observation the Jitter is maximum for Olsrv2 Niigata which is followed by AODV, RIPng, Olsr Inria and Olsrv2 Niigata and DYMO. DYMO gives the minimum jitter for Ipv6 network.
- Olsr Inria gives an average amount of jitter.

4.3. Analysis of Average End-to-End Delay (AE2ED):

The successful data packet delivered and divides that sum by the number of successfully received data packets. The average time taken in delivery of data packets from source to destination nodes.

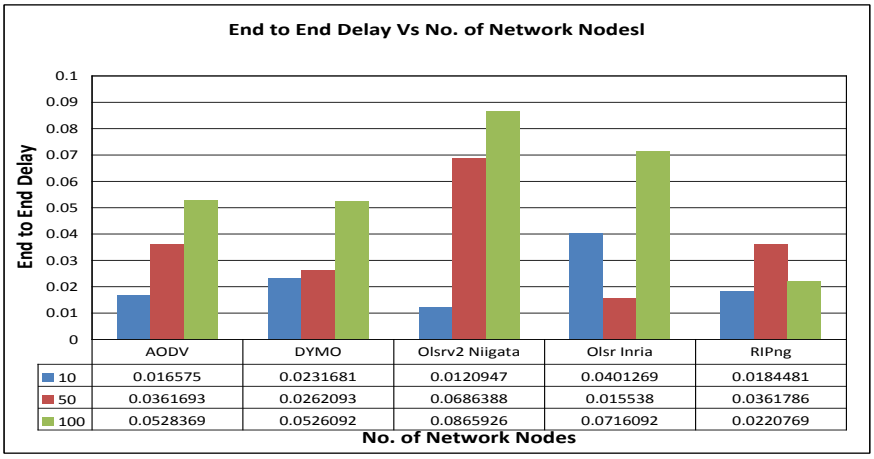


Figure 4. Shows impact variation of Average End to End Delay for various routing protocol as parameter Ipv6 network.

- By observation the Average End to End Delay is maximum for Olsrv2 Niigata which is followed by Olsr inria, AODV, DYMO then RIPng. RIPng gives the minimum average End to End delay for Ipv6 energy model.

4.4. Total Packet Received/Efficiency :

Ratio between the data packets received from to the destination and those generated by CBR sources. This evaluates the ability of the protocol to discover routes and its efficiency.

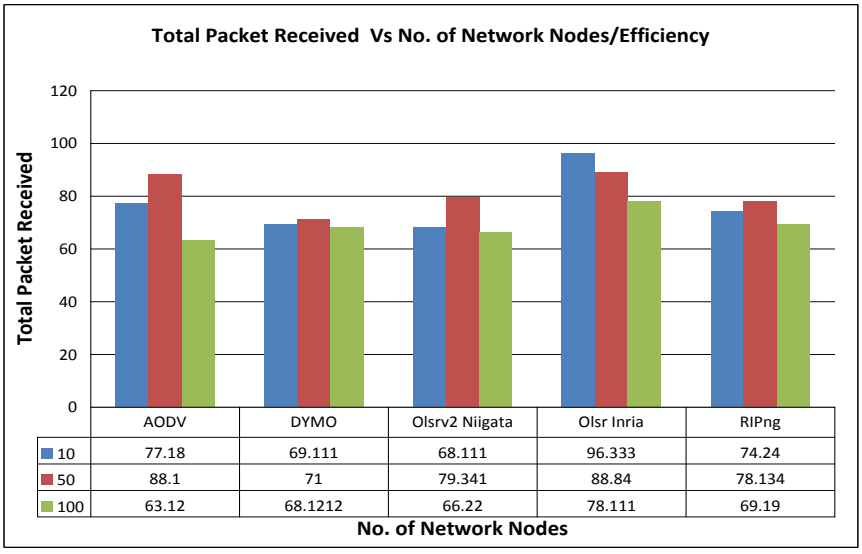


Figure 5: Comparison of Routing protocol with varying network size in effect to Total Packet Received in Application Layer

- By observation of Fig.5 the Total Packet Received in Ipv6 is maximum for Olsr inria which is followed by Olsrv2 Niigata, AODV, DYMO then RIPng. RIPng protocol received the minimum packets for Ipv6 in application layer.

4.5. Analysis And Impact Of Energy Consumed In Transmit Mode:

The mobility, efficiency, scalability, response time of nodes, lifetime of nodes, and effective sampling frequency, all these parameters of the MANET depend upon the energy. In case of power failure the network goes down break therefore energy is required for maintaining the individual health of the nodes in the network, during transmission of data as well receiving the packets.

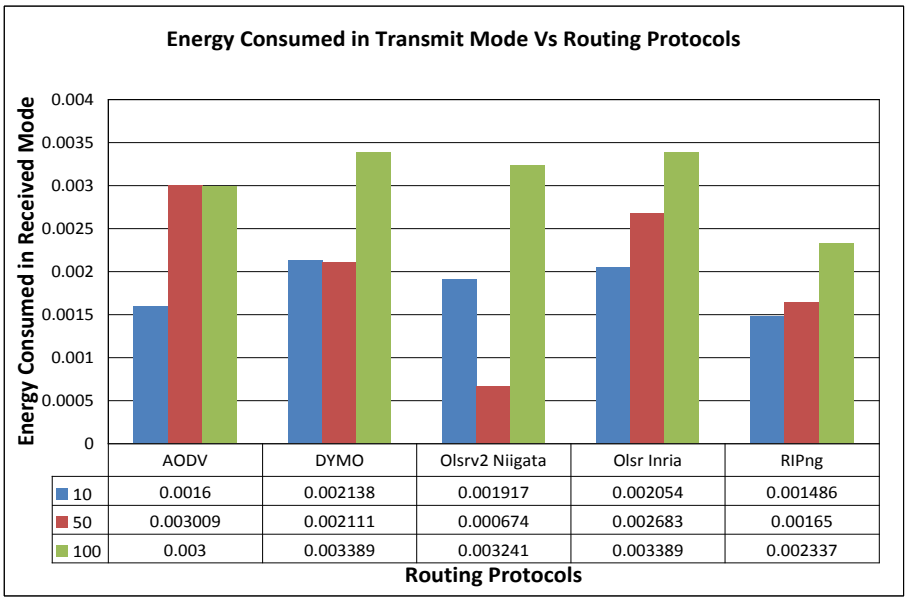


Figure 6. Shows the impact variation of Energy consumption in transmit mode with different routing protocols

Fig. 6 shows the total energy consumed (Joules) by all the nodes while varying the number of nodes in the network connection by (10-100). The routing packet is increased which impacts that energy consumption also increased of all protocols in Ipv6 network. AODV performed better than all other protocols due to route cache.

- By observation from graph the maximum energy consumes by AODV, followed by DYMO, Olsr inria , Olsrv2 Niigata and RIPng. RIPng consumes the minimum power in transmit mode for Ipv6 networks.

4.6. Analysis and impact of energy consumed in receive mode:

The mobile ad-hoc network routing protocol efficiency depends upon the energy of network. If more power failure then efficiency of network goes down therefore energy consumption in received mode is required for maintaining the efficiency of the nodes in the network, during transmission of data as well receiving the packets.

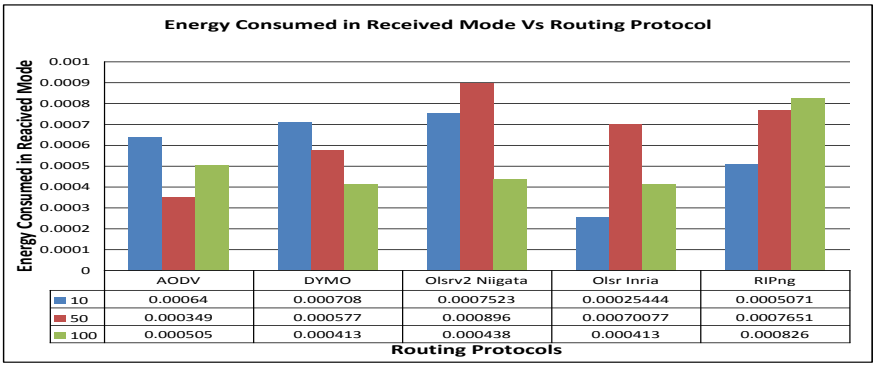


Figure .7 shows the impact variation of Energy consumption in receive mode with different routing protocols.

- By observation from graph the maximum energy received by AODV which is followed by DYMO, Olsr inria, RIPng than Olsrv2 Niigata in Ipv6 network.

4.7. Analysis and impact of energy consumed in ideal mode:

The energy consumption in idle mode that there is maximum consumption in AODV followed by Olsrv2 Niigata, DYMO, Olsr inria than RIPng.

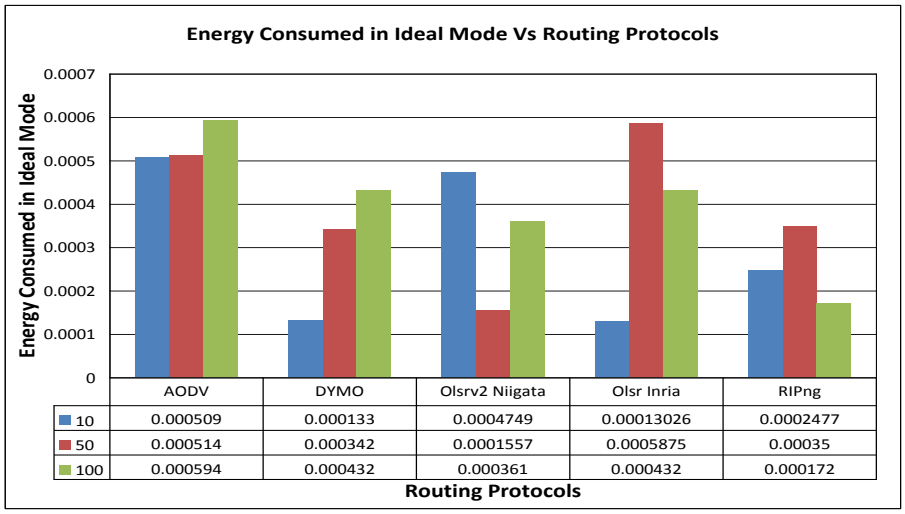


Figure 8. Shows the impact variation of Energy consumption in ideal mode with different routing protocols.

- By observation we are considering the energy consumed in idle mode AODV consumed more and RIPng consumes very less in idle mode but in the case of Olsrv2 Niigata , it is consumes in between DYMO and Olsr inria in Ipv6 network.

5. Conclusion

In this paper we have made a comparison between five different types of routing protocols in Ipv6 network i.e., AODV, DYMO, Olsrv2 Niigata, Olsr inria and RIPng. These results of comparison are very much useful for researcher to be implemented in professional purposes. We are observed that route maintenance and route construction mechanisms have much effect on protocol performance in Ipv6 network. The above graphical simulation results showed that the OLSR inria throughput is almost the same as the OLSRv2 Niigata packet throughput. Both take a different path as if the network topology is same Ipv6 network. We simulate and analyzed energy model comparison and impact shown in above graphs. As far as we can conclude, the performance of DYMO and Olsr inria was promising in almost all scenarios but with a high end-to-end delay varying between (10 to 50) nodes. AODV was the third best performing protocol but resulted to be more sensitive than the others to network size and traffic load. AODV performance is not much affected by mobility. Olsrv2 Niigata is the route maintenance mechanism does not locally repair the broken links which results in initiating another route discovery, which introduces extra delays with more routing overhead. We can conclude that Olsr inria is more reliable and more adaptable to changing network conditions in Ipv6 network. As mobility increases, the average end-to-end delay decreases. For future work we can next perform using QualNet simulator taking all above Manet routing protocols AODV, DYMO, Olsrv2 Niigata, Olsr inria and RIPng using Dual IP (Ipv4 and Ipv6) taking all performance matrices parameters same.

References

[1] Boukerche A., 2004. Performance Evaluation of Routing Protocols for Ad Hoc Wireless Networks Mobile Networks and Applications, Vol. 9, pp 333 - 42

[2] C. E. Perkins, Performance Comparison of Two On-Demand Routing Protocols for Ad Hoc Networks. In: IEEE personal communication. 2001

[3] Md. Arafatur Rahman, Farhat Anwar, Jannatul Naeem and Md. Sharif Minhazul Abedin, "Simulation Based Performance Comparison of Routing Protocol on Mobile Ad-hoc Network (Proactive, Reactive and Hybrid)", International Conference on Computer and Communication Engineering (ICCCE 2010), 11-13 May 2010, Kuala Lumpur, Malaysia.

[4] C.Siva Rammurty and B.S. Manoj , "Ad hoc wireless networks architectures and protocols" ISBN 978-81-317-0688-6, 2011.

[5] C.E. Perkins, and E. M. R., 1999. Ad hoc on-demand distance vector routing(AODV). In: Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications, New Orleans, LA, pp 90-100.

[6] Ha Duyen Trung, W. B., and Phan Minh Duc, 2007. Performance evaluation and comparison of different ad hoc routing protocols Computer Communications, Vol 30, pp 2478-96.

[7] Alexander Klein," Performance Comparison and Evaluation of AODV, OLSR, and SBR in Mobile Ad-Hoc Networks" PP 571-575, 2008 IEEE

[8] Hong Jiang, and J. J. G-L-A., 2001. Performance Comparison of Three Routing Protocols for Ad Hoc Networks. In: Proc. of IEEE Tenth International Conference on Computer Communications and Networks (ICCCN), pp 547 - 54.

[9] J. Broch. et al, 1998 A performance comparison of multi-hop wireless ad hoc network routing protocols. In: Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking, Dallas, Texas, United States, pp 85 - 97.
[10] T.Clausen, C.Dearlove, J.Dean, C.Adjih, "Generalized MANET Packet/Message Format," draft-ietf-manet-packetbb-02, internet draft, July, 2006.
[11] P. Johansson, T. L., and N. Hedman, 1999. Scenario-based Performance Analysis of Routing Protocols for Mobile Adhoc Networks. In: International Conference on Mobile Computing and Networking, Proceedings of the 5th annual ACM/IEEE, Seattle, Washington, United States, pp 195 - 206.
[12] Samir R. Das, C. E. P., and Elizabeth M. Royer, 2000. Performance Comparison of Two on-Demand Routing Protocols for Ad Hoc Networks. In: In Proceedings of the IEEE Conference on Computer Communications (INFOCOM), Tel Aviv, Israel.
[13] T.Clausen, "Optimized Link State Routing Protocol version 2,"draft-ietf-manet-olsrv2-02, internet draft, Jun. 2006.
[14] I.Chakeres, C.Perkins, "Dynamic MANET On-demand (DYMO) Routing," draft-ietf-nanet-dymo-05, June, 2006.
[15] The Qualnet 5.0.2 simulator tools online available www.scalable-networks.com.

Motion Blur Image Fusion Using Discrete Wavelate Transformation

Er. Shabina Sayed

Department Of Information Technology, MHSS COE, Mumbai, India

Abstract

The methodology for implementing a image fusion system using deconvolution and discrete wavelate transformation is proposed in this papers. This project proposes a method to remove the motion blur present in the image taken from any cameras. The blurred image is restored using Blind de-convolution method with N=20 number of iteration and DWT using averaging, maximum likelihood and window based method.the comparision result of both the method prove that image restoration using dwt gives better result than image restoration using deconvolution.

Keywords: multisensory system,pyramid transform,discrete wavelet transform,Motion blur,blind deconvolution,

1. Introduction

With the recent rapid developments in the field of sensing technologies multisensory systems[1,2] have become a reality in a growing number of fields such as remote sensing, medical imaging, machine vision and the military applications for which they were first developed. The result of the use of these techniques is a great increase of the amount of data available. Image fusion provides an effective way of reducing this increasing volume of information while at the same time extracting all the useful information from the source images. Multi-sensor images often have different geometric representations, which have to be transformed to a common representation for fusion. This representation should retain the best resolution of either sensor. A prerequisite for successful in image fusion is the alignment of multi-sensor images. Multi-sensor registration is also affected by the differences in the sensor images.However, image fusion does not necessarily imply multi-sensor sources, there are interesting applications for both single-sensor and multi-sensor image fusion, as it will be shown in this paper.The primitive fusion schemes perform the fusion right on the source images.One of the simplest of these image fusion methods just takes the pixel-by-pixel gray level average of the source images. This simplistic approach often has serious side effects such as reducing the contrast. With the introduction of pyramid transform in mid-80's[3], some sophisticated approaches began to emerge. People found that it would be better to perform the fusion in the transform domain. Pyramid transform appears to be very useful for this purpose. The basic idea is to construct the pyramid transform of the fused image from the pyramid transforms of the source images, and then the fused image is obtained by taking inverse pyramid transform. Here are some major advantages of pyramid transform:

- It can provide information on the sharp contrast changes, and human visual system is especially sensitive to these sharp contrast changes.
- It can provide both spatial and frequency domain localization There are many transformations which can be used but Basically this paper makes the contribution of the two important transformation .

2. Discrete Wavelet Transformation(DWT)

The wavelet transform[4,7], originally developed in the mid 80's, is a signal analysis tool that provides a multi-resolution decomposition of an image in a bi orthogonal basis and results in a non-redundant image representation. These bases are called wavelets, and they are functions generated from one single function, called mother wavelet, by dilations and translations. Although this is not a new idea, what makes this transformation more suitable than other transformations such as the Fourier Transform or the Discrete Cosine Transform, is the ability of representing signal features in both time and frequency domain.Fig.1 shows an implementation of the discrete wavelet transform. In this filter bank, the input signal goes through two one-dimensional digital filters. One of them, H_0, performs a high pass filtering operation and the other H_1 a low pass one. Each filtering operation is followed by sub sampling by a factor of 2. Then, the signal is reconstructed by first up sampling, then filtering and summing the sub bands.

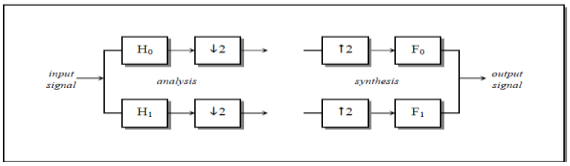


Figure 1.two channel filter bank

The synthesis filters F_0 and F_1 must be specially adapted to the analysis filters H_0 and H_1 to achieve perfect reconstruction [3]. By considering the z-transfer function of the 2-chanel filter bank shown in Fig.1 it is easy to obtain the relationship that those filters need to satisfy. After analysis, the two subbands are:

$$\frac{1}{2}\left[H_{0(z^{1/2})}X\left(z^{1/2}\right)+H_{0(-z^{1/2})}X\left(-z^{1/2}\right)\right] \quad (1)$$

$$\frac{1}{2}\left[H_{0(z^{1/2})}X\left(z^{1/2}\right)+H_{0(-z^{1/2})}X\left(-z^{1/2}\right)\right] \quad (2)$$

Then, the filter bank combines the channels to get $\hat{x}(n)$. In the z-domain this is $\hat{X}(z)$. Half of the terms involve $X(z)$ and half involve $X(-z)$.

$$\hat{X}(z)=\frac{1}{2}[F_0(z)H_0(z)+F_1(z)H_1(z)]X(z)+$$
$$\frac{1}{2}[F_0(z)H_0(-z)+F_1(z)H_1(-z)]X(-z) \quad (3)$$

There are two factors to eliminate: aliasing and distortion. For alias cancellation choose:

$$F_0(z)=H_1(-z)$$
$$F_1(z)=H_0(-z) \quad (4)$$

The distortion must be reduced to a delay term, to achieve this Smith and Barnwell suggested [8]:

$$H_1(z)=-z^{-N}H_0(-z^{-1}) \quad (5)$$

With these restrictions the final filtering equation is

$$X(Z)=\frac{1}{2}z^{-N}[H_0(z)H_0(z^{-1})+H_0(z^{-1})H_0(z)]X(z) \quad (6)$$

Fig.2 represents one step in a multiscale pyramid decomposition of an image [3]. The algorithm applies a one-dimensional high and low pass filtering step to the rows and columns separately in the input image. The inverse transform filter bank structure is represented in Figure 4.

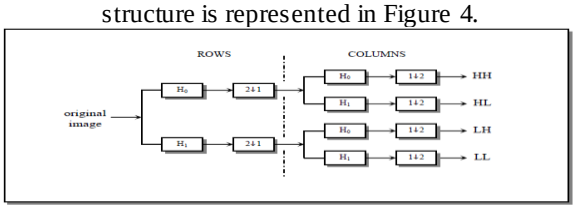


Figure 3 filter bank structure of DWT analysis

Successive application of this decomposition to the LL sub band gives rise to pyramid deco mposition where the sub images correspond to different resolution levels and orientations as exemplified in Fig.5. Some images decomposed with the wavelet transform are shown in figure.6

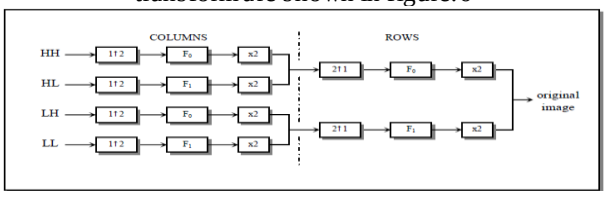


Figure 4 filter bank structure of the reverse DWT synthesis

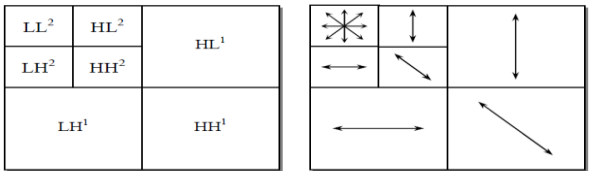


Figure 5 Image decomposition. Each sub band has a natural orientation

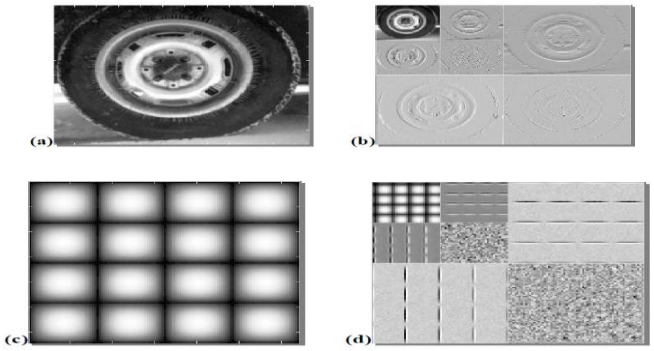


Figure 6 Images (a) and (c) shows original images and (b) and (d) their wavelet decomposition

3. Blind deconvolution

3.1_Proposed Method

Restoration techniques [1,5] are oriented toward modeling the degradation and applying the inverse process in order to recover the original image. The image gets blurred due to the degradation. Blur is of many types but for this paper motion blur is considered.

3.2_Block Diagram

Fig. 7 shows the block diagram of the proposed method[1]. In the original image noise is added to get blurred noisy image. To remove motion blur, the blurred image is restored using restoration algorithms. Finally the filtered Images are fused using image fusion method to get the fused image.

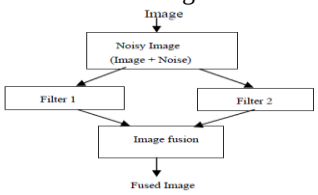


Figure 7 Block diagram

3.3 Noise model

Generally the noise is modeled as zero mean white Gaussian additive noise. But here we have modeled noise as sum of the multiplicative noise and additive Gaussian noise as

$v(x, y) = f(x, y)*\sigma1(x, y)+ \sigma2(x, y)$ (7)

Where $\sigma1(x, y)$ is the multiplicative noise and $\sigma2(x, y)$ is the additive noise.

3.4 Image restoration

In order to remove motion blur, various image restoration algorithms have been proposed. Blind deconvolution adopts regularized iteration to restore the degraded image. But it requires large computational complexity. For this reason, the work proposes the implementation of wiener filter to reduce the computational complexity with better acceptable restoration results of image restoration method.

3.5 Point Spread Function (PSF)

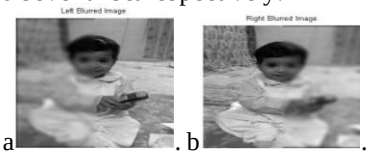
The General form of motion blur function [6] is given as follows,

$$h(x, y) = \begin{cases} \frac{1}{L}, & \text{if } \sqrt{x^2 + y^2}\frac{L}{2}, \frac{x}{2} = -\tan(\varnothing) \\ 0, & otherwise \end{cases} \quad (7)$$

As seen that motion blur function depends on two parameters: motion length (L) and motion direction (ϕ).

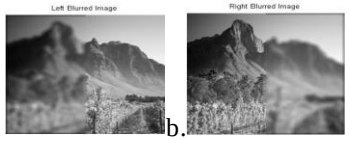
4. Fusion Techniques

In this research, two fusion approaches have been developed. These two approaches are the blind deconvolution & the discrete Wavelet Transform. These two methods, were selected for being the most representative approaches, especially the approach based on the Wavelet Transform has became the most relevant fusion domain in the last years. This section describes the technique specifications, their implementation and presents experimental results and conclusions for each of them. It is important to note that all the source images used in this research were already correctly aligned on a pixel-by-pixel basis, a prerequisite for successful image fusion. The fusion techniques have been tested with four sets of images, which represent different possible applications where fusion can be performed. The first set of images, Figure 10 called 'kid' represent the situation where, due to the limited depth-of-focus of optical lenses in some cameras, it is not possible to get an image which is in focus everywhere. The second set of images, Figure 11 corresponds to navigation . In this case, a millimeter wave sensor is used in combination with a visual image. An example of fusion applied to medicine is represented in the third set of images, Figure 12. One of the images was captured using a nuclear magnetic resonance (MR) and the other using a computed tomography (CT).Remote sensing is another typical application for image fusion. The fourth set of images in figure 13 illustrates the captures of two bands of a multispectral scanner. The purposes of them are navigation applications and surveillance applications for the sixth and seventh set respectively.



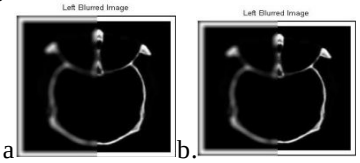
a. b.

Figure 10 Set 1 (a) focus on left, Image 2 (b) focus on right



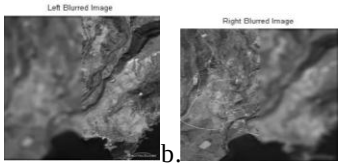
a. b.

Figure 11 Set 2 (a) focus on left, Image



a b.

Figure 12 Set 3. Image 1 (a) focus on left, Image 2 (b) focus on right

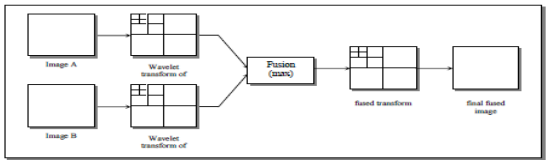


a. b.

Figure 13 Set 4. Image 1 (a) and Image 2 (b) multispectral scanner

4.1 Technique

An alternative to fusion using pyramid based multi resolution representations is fusion in the wavelet transform domain. As mentioned in *section II.* The wavelet transform decomposes the image into low-high, high-low, high-high spatial frequency bands at different scales and the low-low band at the coarsest scale. The L-L band contains the average image information whereas the other bands contain directional information due to spatial orientation. Higher absolute values of waveletcoefficients in the high bands correspond to salient features such as edges or



lines.

Figure 14.wavelet fusion scheme

Wavelet transform is first performed on each source images, and then a fusion decision map is generated based on a set of fusion rules as shown in figure 14. The fused wavelet coefficient map can be constructed from the wavelet coefficients of the

source images according to the fusion decision map. Finally the fused image is obtained by performing the inverse wavelet transform. From the above diagram, we can see that the fusion rules are playing a very important role during the fusion process. Here are some frequently used fusion rules in the previous work as shown in figure 15.

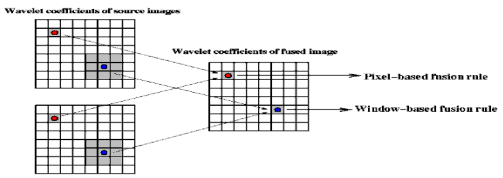


Figure 15.Frequently used fusion rules

When constructing each wavelet coefficient for the fused image. We will have to determine which source image describes this coefficient better. This information will be kept in the fusion decision map. The fusion decision map has the same size as the original image. Each value is the index of the source image which may be more informative on the corresponding wavelet coefficient. Thus, we will actually make decision on each coefficient. There are two frequently used methods in the previous research. In order to make the decision on one of the coefficients of the fused image, one way is to consider the corresponding coefficients in the source images as illustrated by the red pixels. This is called pixel-based fusion rule. The other way is to consider not only the corresponding coefficients, but also their close neighbors, say a 3x3 or 5x5 windows, as illustrated by the blue and shadowing pixels. This is called window-based fusion rules. This method considered the fact that there usually has high correlation among neighboring pixels. In this research, it has been thought that objects carry the information of interest, each pixel or small neighboring pixels are just one part of an object. Thus, we proposed a region-based fusion scheme. When there is a need to make the decision on each coefficient, it consider not only the corresponding coefficients and their closing neighborhood, but also the regions the coefficients are in. It is observe that the regions represent the objects of interest.

5. Experimental Results

This section demonstrate some experimental results of the proposed method.The results are compared on the basis of the RMSE, for different type of images (kids,medical,satellite) image restored using blind deconvolution (N = 20 iterations) .[J,PSF] = deconvblind(I, INITPSF) deconvolves image I using the maximum likelihood algorithm, returning both the deblurred image J and a restored point-spread function PSF.. The results of blind deconvolution are as shown in Table 1.

Table1.RMSE of the images restore using blind deconvolution(N=20).

Sr.no	Type of image	Blurred noisy image (rmse)	Blind deconvolution for N=20 (rmsek)
1.	Medical	69.9514	77.1854
2.	Kids	45.8178	47.7941
3.	Satellite	46.7137	47.8174

The Results of the second experiment of Wavelet based Fusion are shown in Table 2 (Fused image_1) and Table 3(Fusedimage_2), Table 4 (Fused image_3).where Fused image_1 is the image fusion using DWT by pixel averaging method. Fused image_2 is the image fusion using DWT by maximum likelihood method. Fused image_3 is the image fusion using DWT by window based method.

Table 2 RMSE of the images fusion using averaging pixel method

Sr.no	Type of image	Blurred noisy image (rmse blur)	Image fusion using DWT (rmsek)
1.	Medical	39.4644	62.4161
2.	Kids	26.8725	32.4585
3.	Satellite	38.1441	39.0473

Table 3 RMSE of the images fusion using maximum likelihood method .

Sr.no	Type of image	Blurred noisy image (rmse blur)	Image fusion using DWT (rmsek)
1.	Medical	39.4644	78.8207
2.	Kids	26.8725	51.4416
3.	Satellite	38.1441	50.8605

Table 4 RMSE of the images fusion using window based method .

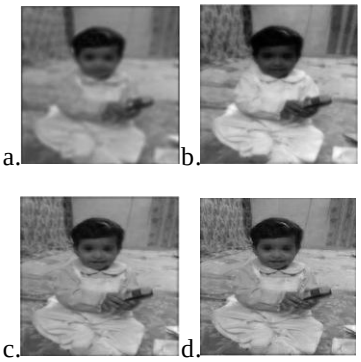


Figure 16.a) kid image a)image fusion using deconvolution. b)image fusion using pixel averaging method. c)image fusion using maximum likelihood method.d)image fusion using window based method

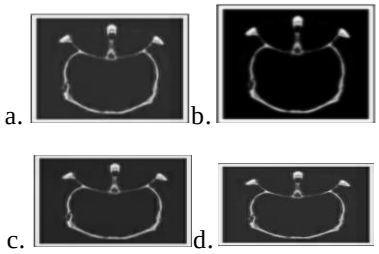


Figure 17 medical image a) image fusion using deconvolution.b)image fusion using pixel averaging method.c)image fusion using maximum likelihood method.d)image fusion using window based method

Sr.no	Type of image	Blurred noisy image (rmse blur)	Image fusion using DWT (rmsek)
1.	Medical	39.4644	30.3697
2.	Kids	26.8725	23.6879
3.	Satellite	38.5179	39.8529

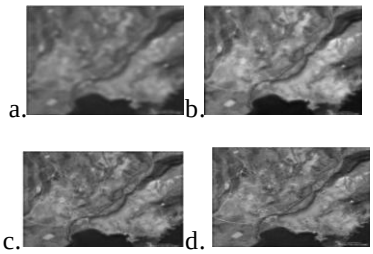


Figure 18. satellite image a) image fusion using deconvolution.b)image fusion using pixel averaging method.c)image fusion using maximum likelihood method.d)image fusion using window based method

6. Conclusion

This research proposes a method to remove the motion blur present in the image taken from any cameras. The blurred image is restored using Blind de-convolution method with N=20 number of iteration and DWT using averaging, maximum likelihood and window based method. The result based on deconvolution does not improve the image quality drastically. If we compare the rmse of blurred image and fused image we can see that fused image rmse is higher than blurred image rmse. It is also proved by performing visual comparison among all the fused image. There is still significant difference between blurred image and fused image of fusion using pixel average and maximum likelihood approach. Medical image rmse is significantly higher than blurred image in all the methods except window based method. The primitive fusion schemes like pixel averaging and maximum pixel perform the fusion right on the source images. These methods often have serious side effects such as reducing the contrast of the image as a whole. But these methods do prove good for certain particular cases wherein the input images have an overall high brightness and high contrast. Further window based method compared for fusion, and it gave the best results. If computationally it's performance is compared it's rmse of fused image for all type of image is minimum, also the fused image quality is improved because the fused image rmse is lower than blurred image rmse. We can do further satisfaction by visual comparison of all the fused images. The challenge is to design a method that exhibits the most appropriate compromise among computational complexity, reliability, robustness to noise, and portability for a given application.

7. References

[1] Hui Li; Manjunath, B.S.; Mitra, S.K" Multi_sensor image fusion using the wavelettransform" Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference, 1994 , Page(s): 51 - 55 vol.1 .

[2] Implementation and Comparative Study of Image Fusion Algorithms " International Journal of Computer Applications (0975 – 8887)Volume 9– No.2, November 2010"

[3] Gihong Qu, Dali Zhang and Pingfan Yan, "Medical image fusion by wavelet transform modulus maxima," Optics Express, vol. 9, No. 4 pp.184-190,Aug. 2001

[4] S.Mallat "An Improved Image Denoising Method Based on Wavelet Thresholding"journal of signal and information processing PP.109-116 DOI: 10.4236/jsip.2012.31014

[5] Zhu Shu-Long, "Image fusion using wavelet transform,"Symposium on Geospatial Theory, Processing and Applications, Ottawa 2002.

[6] Hongbo Wu; Yanqiu Xing "Pixel-based Image Fusion Using Wavelet Transform for SPOT and ETM+ Image" Publication Year: 2010 , Page(s): 936 - 940

[7] Y. Xia, and M. S. Kamel "Novel Cooperative Neural fusion Algorithms for Image Restoration, Image Fusion", Feb 2007.