محور الهندسة

Production of 68Ga from 68Zn at Low Energy Medical Cyclotron

Iman Tarik Al-Alawy [a,] *
Hamza Abed Al-Kadhim Mezher[b]
[a] Department of Physics, Collage of Science, University of Al-Mustansiriyah, Iraq
[b] Department of Physics, Collage of Science, University of Karbala, Iraq
* Email: profimantarik@gmail.com

Abstract

In this work 68Ga was produced via the 68Zn(p,n)68Ga nuclear reaction. Since 68Ga is an important positron emitting radionuclide for positron emission tomography. 68Zn target was irradiated with a proton beam. The production yield achieved was 6124MBq/μA (165.51Ci/μA) at 12MeV with 0.012% impurity of 67Ga. Stopping power and calculated yields for induced protons to produce Gallium from 68Zinc isotope in the energy range from threshold energy up to 85 MeV proton energy have been calculated, for 68Zn(p,n)68Ga, 58Zn(p,2n)67Ga, and 68Zn(p,3n)66Ga reactions. Complete energy range starting from threshold energy for each reaction have been analyzed statistically and the adopted cross sections were reproduced in fine steps of incident proton energy in 0.01 MeV intervals with their corresponding errors. The stopping power according to Zeigler formula was used in order to obtain the cross sections and calculated yield for each reaction based on the complete spectrum of cross sections.

Keywords: induced proton, Zinc-68 target, production yields, Gallium-68**,** cyclotron

1. Introduction

Among the Gallium radioisotopes 68Ga is of special interest to nuclear medicine as it can be employed both for medical imaging (via Positron Emission Tomography: PET) and for targeted radio-immunotherapy of tumors. Its relative short half-life ($T_{1/2}$ = 67.71m; decay scheme: EC(100%), Q-value(2921.1keV), β+(829.5keV, 88.91%, 0.7334MBq/μA) [1]. The benefit of its half-life that it is short enough to limit the patient's exposure during these studies [2] Several research centers have already investigated its production routes via proton induced nuclear reactions on highly enriched 68Zn (stable isotopic with abundance 18.45%, Q-value(-5462.1keV) solid target [1]. Nowadays, many nuclear reaction processes are used in practice

providing the purest form of both radionuclide and chemical in very high yield. Due to the relative high price of the enriched materials, however, alternative production methods were studied in recent years. As a result of these detailed studies, proton induced reactions on highly enriched 68Zn target have been suggested as candidates for this purpose. The investigated 68Zn(p,n)68Ga, 68Zn(p,2n)67Ga, 68Zn(p,3n)66Ga nuclear reactions were follows in the present work. Compilation and evaluation of these production routes of Gallium was already the subject of an IAEA's Coordinated Research Project [3]. Furthermore, Alves et al. [4] have also published a comprehensive evaluation of these production related cross section data. Additionally, the limited number of higher energy proton accelerators in research centers could also limit the widespread application of the proton ways. Surveying the information on cross section data of all Gallium producing from 68Zn+p reactions, it was found one process that could be useful even at low energies, namely the 68Zn(p,n) 68Ga (EC(Q = 2921.1keV)) reaction [5]. The excitation function of this reaction was studied by different authors in the past, published in IAEA libraries. To evaluate the practical production circumstances of 68Ga (i.e. production energy range, activation yield, contamination level, etc.) at a biomedical cyclotron via the 68Zn(p,n)68Ga reaction, it is important to have regular cross section databases of all 68Zn+p reactions that form 66,67,68Ga radioisotopes below 18 MeV. From the point of view of 68Ga, the presence of 66Ga($T_{1/2}$=9.49h, EC(Q=5175keV), 0.97MBq/μA) and 67Ga($T_{1/2}$= 3.2617day, EC(Q=1001.3keV), 4.16E-7MBq/μA) [1] at EOB means the major radio-Gallium contaminations. The presence of 67Ga at EOB may limit the length of the application period of the compounds. These reactions seem to be well measured and can be used for calculating the thick target yields of the above reactions with the required precision. We discuss here in detail the actual production possibility of 68Ga via irradiations with three different energy proton beams.

The cross sections evaluation for calculated yields for induced proton on Zinc to produce Gallium, Zinc target elements are calculated according to the available International Atomic Energy Agency (IAEA) libraries and other experimental published data. The stopping power depends on the type and energy of the incident particle and on the properties of the materials it passes. In passing through matter, fast charged particles ionize the atoms or molecules which they encounter. The yield for a target having any thickness can be defined as the ratio of the number of nuclei formed in the nuclear reaction to the number of particles incident on the target. Thick target yield is defined for a fixed macroscopic energy loss, E_{in}-E_{out}, in a thick target. Integral yield is defined for a finite energy loss down to the threshold of the reaction, E_{in}-E_{th}. The recommended cross sections

discussed in the present work and the target stopping powers of Ziegler [6,7] and SRIM program (2003) were used to Evaluation the calculated yields for a target of significant thickness. The cross sections of calculated yields for induced proton on Zinc to produce Copper published by different authors [8-25] in the energy range 3.766– 85.0MeV. Adopted values have been calculated, the cross sections were reproduced in fine steps of incident proton energy in 0.01 MeV intervals with the corresponding errors. In this study the stopping power have been calculated using SRIM program and Ziegler formulae [6,7] corpuscle to three regions based on the velocity of the incident proton (V). The calculated adopted cross sections for these reactions have been evaluated and systematic behaviors of calculated yields with proton energy and target numbers (Z) have been observed throughout the studied isotopes.

2. Theretical part

a. Stopping Power

Incident protons with certain energy will lose all their energies in a definite distance in a medium before it stopped completely. The mechanism for the stopping power of ions penetrating condensed matter depends on the charge and velocity of the incident corpuscle and the nature of the matter, for that reason one can be compilation the energy loss of the charge corpuscle to three regions (high, intermediate and low energy). The behavior of ions in each region can be explained as the following [26]:

i. The high energy region

This region can occurs when the velocity of the incident corpuscle (V) is ($V \geq 2V_0Z_1$) where (Z_1) is the atomic number of ion and (V_0) represents the Bohr velocity ($V_0 = 2.18\times10^6$ m/s) and this is about the velocity of the conduction electrons in solid. Ions with velocity below (V_0) have adiabatic collisions with target electrons and hence small stopping power. The stopping power increases with decreasing ion-velocity [27].

The electronic stopping power (S_e) is to prevail with Bethe (1933) equation applies in this region [28]:

$$-\frac{dE}{dx} = NS_e = \frac{4\pi K^2 e^4 Z_1^2}{mV^2} NZ_2 \left[\ln\left(\frac{2mV^2}{I}\right) - \ln\left(1-\beta^2\right) - \beta^2 \right] \quad \ldots(1)$$

Where N is the atomic density of the medium [$N=N_a(\rho/A)$], N_a is the Avogadro's number ($N_a=6.022\times10^{23}$mole^{-1}), ρ is the density of matter, A is the mass number, e,m are the charge and mass of the electron respectively, Z_1, Z_2: are the atomic numbers of ion and target respectively, β

is the ratio between incident corpuscle velocity and the velocity of light, I is the mean ionization and excitation potential, K is the coulomb constant $K = \frac{1}{4\pi\varepsilon_0} = 8.99\times10^{9}\mathrm{Nm^2C^{-2}}$.

ii. The Intermediate Energy Region

The intermediate energy region occurs when the velocity (V) of the incident corpuscle is in the range ($2V_0Z_1>V\geq V_0Z_1^{2/3}$); it includes the maximum stopping power. In this region the effect of effective charge is clear and that is because of loss its energy which is mean decrease of corpuscle velocity and charge Z_1 decreased too, and that because of loss or acquire electrons and there will be elastic collision with the nuclei of atoms occur. Thus equation (1) was modified, and its express electronic stopping power as Bethe-Bloch (1933) [28].

$$-\frac{dE}{dx} = NS_e = \frac{4\pi K^2 e^4 Z_1^2}{mV^2} NZ_2 L \qquad \ldots (2)$$

Where L is the stopping atomic number and depends on the velocity of incident corpuscle and the medium of the target.

$L= L_0 + Z_1 L_1 + Z_1^2 L_2$ … (3)

Where $L_0 = \ln(2wv^{1/2} / I) - C/Z_2$ …(4)

C/Z_2 is the shell correction, $Z_1 L_1$ is the Barkas effect correction from the polarization, $Z_1^2 L_2$ is Bloch-correction to transform from quantum to classical form.

iii. The Low energy region

It occurs when the incident corpuscle velocity (V) ($V<V_0Z_1^{2/3}$) in this region, to calculate the cross section for electronic stopping on the Thomas-Fermi potential as a function of velocity. The equation for this region is given by [29,30]:

$$S_e = 8\pi e^2 a_0 \frac{Z_1^{7/6} Z_2}{Z^{2/3}} \left(\frac{V}{V_0}\right) \qquad \ldots (5)$$

Where $Z^{2/3} = Z_1^{2/3} + Z_2^{2/3}$ … (6)

and a_0 represents the Bohr radius, $a_0 = \frac{h^2}{me^2} = 5.29\times10^{-11}\ \mathring{A}$ …(7)

In the scope of this work, the electronic stopping powers were programmed and using the empirical formulae given by Ziegler as follows [31]:

1- Energy range (1-10)×10^{-3} MeV

$$-\frac{dE}{dx} = A_1 E^{1/2} \quad \ldots (8)$$

2- Energy range (10-999)×10^{-3} MeV

$$\left(-\frac{dE}{dx}\right)^{-1} = \left(-\frac{dE}{dx}\right)_{Low}^{-1} + \left(-\frac{dE}{dx}\right)_{High}^{-1} \quad \ldots (9)$$

$$\left(-\frac{dE}{dx}\right)_{Low} = A_2 E^{0.45} \quad \ldots (10)$$

$$\left(-\frac{dE}{dx}\right)_{High} = \left(\frac{A_3}{E}\right)\ln\left[1+\left(\frac{A_4}{E}\right)+A_5 E\right] \quad \ldots (11)$$

3- Energy range (1000-100.000)×10^{-3} MeV

$$\left(-\frac{dE}{dx}\right) = \left(\frac{A_6}{\beta^2}\right)\left[\ln\left(\frac{A_7\beta^2}{1-\beta^2}\right) - \beta^2 - \sum_{i=0}^{4} A_{i+8}(\ln E)^i\right] \quad \ldots (12)$$

Where E is the proton energy in (MeV), A_i are the coefficients given by Ziegler [31,32] , and β is the ratio between incident corpuscle velocity and the velocity of light.

b. Calculated Yield

The Yield of calculated detected per incident particle, Y, for an ideal, thin, and uniform target and mono-energetic particles beam of incident energy E_b is given by [33].

$$Y = (nt)\sigma(E_b)\varepsilon(E_b) \quad \ldots (13)$$

Where n is the number of target atoms per unit volume, t is the target thickness, σ is the reaction cross section, and ε is the proton-detection efficiency. For target which is not infinitesimally thin, the beam loses energy as it passes through the target, and the Yield is then given by [34,35].

$$Y = \int_{E_{thr}}^{E_b} \frac{n\sigma(E)\varepsilon(E)f dE}{-\frac{dE}{dx}(E)} \quad \ldots (14)$$

In which (E_{thr} =E_b-ΔE). Where E_{thr} is the reaction threshold energy, ΔE is the energy loss of the beam in the target, f is the number of target atoms in each target molecule, $-\frac{dE}{dx}(E)$ is the stopping power of the medium as a function of the beam energy.

If the target is sufficiently thick, and there exist one atom per each molecule (i.e., $f=1$) and taking the efficiency $\varepsilon(E)=1$, then the resulting calculated yield is called the thick-target yield which is given by [23]:

$$Y(E_b)=\int_{E_{thr}}^{E_b}\frac{n\sigma(E)dE}{-(dE/dx)} \quad \ldots (15)$$

The thick target yield can be calculated from the following expression [36,37]:

$$Y=\frac{N_A H}{Mzq_e}\int_{E_{in}}^{E_{out}}\left(\frac{dE}{dx}\right)^{-1}\sigma(E)dE \quad \ldots (16)$$

where N_A is Avogadro's number, z is the projectile's atomic number and qe is the electron's charge. The integral is carried out over the beam's energy range, from the incoming energy E_{in} to the outgoing energy E_{out}, and the integrand is the inverse of the medium's stopping power multiplied by the cross-section for the corresponding value of the energy. The typical units for the thick target yield in radioisotopes production practice is the $MBq/\mu A$.

3. Data Reduction and Analysis

Method used to obtain the adopted cross sections is as the following:

a. The sets of experimental cross sections data with their corresponding errors were collected for different authors with different energy intervals. The data was re-arranged with energy interval 0.01MeV.

b. The normalization for the statistical distribution of cross sections errors to the corresponding cross section values for each author has been done.

c. The interpolation of the nearest data for each energy interval as a function of cross sections and their corresponding errors, have been done using Matlab-8.0.

d. The interpolated values were calculated to obtain the adopted cross section which is based on the weighted average calculation according to the following expressions [38]:

$$\sigma_{w.a.} = \frac{\sum_{i=1}^{n} \frac{\sigma_i}{(\Delta\sigma_i)^2}}{\sum_{i=1}^{n} \frac{1}{(\Delta\sigma_i)^2}} \quad \ldots (17)$$

Where the standard deviation error is:

$$S.D. = \frac{1}{\sqrt{\sum_{i=1}^{N} \frac{1}{(\Delta\sigma_i)^2}}} \quad \ldots (18)$$

Where σ_i: is the cross section value.

$\Delta\sigma_{i:}$ is the corresponding error for each cross section value.

4. Results and Discussion

The production of 68Ge via nuclear reaction has studied in which three potential nuclear reaction routes were presented and discussed. 68Ge can be produced via the 68Zn(p,n)68Ga; 68Zn(p,2n)67Ga or 68Zn(p,3n)66Ga nuclear reactions. However, the nuclear reaction consists in proton bombardment of enriched solid Zinc target nuclear reactions. The results required lower energies to provide significantly higher yields. Table 1 presents theoretical production yields obtained from three distinct nuclear reactions using solid enriched Zinc target. The available measuring data of proton energies and their associated cross sections were collected from the International Atomic Energy Agency (IAEA), using EXFOR library [39]. The available were evaluated in the present work in order to calculate the adopted cross sections using adopt.m program, which is written in the present work using Matlab-8.0. The adopted evaluated cross sections are plotted as a function of incident proton energy starting from threshold energy for each reaction. The results of calculated yields are discussed as follows:

1. $^{68}_{30}Zn_{38}(p,n)^{68}_{31}Ga_{37}$ Reaction

The cross sections data from EXFOR library measured by Mcgee et al.(1970) [8], Hille et al. (1972) [9], Barrandon et al.(1975) [10], Kotelnikova et al.(1980) [11], Esat et al. (1981) [12], Tarkanyi et al.(1990) [13], Levkovskij (1991) [14], Hermanne et al.(1991) [15], Vinogradov et al.(1993) [16] and Zhuravlev et al. (1995) [17] have been plotted, interpolated, and recalculated in steps of 0.01MeV from threshold energy 3.766MeV up to 85MeV of the incident proton energy in order to obtain the adopted cross sections of this reaction. The

results of adopted cross sections and the experimental results for the authors mentioned above are shown in figure 1.

2. $^{68}_{30}Zn_{38}(p,2n)^{67}_{31}Ga_{36}$ **Reaction**

The cross sections data from EXFOR library measured by Mcgee et al.(1970) [8], Litte and Lagunas-Solar (1983) [18], Tarkanyi et al.(1989) [19], Levkovskij (1991) [14], Hermanne et al. (1991) [15], Szelecsenyi et al.(1994) [20], Hermanne (1997) [21], and Szelecsenyi et al. (1998) [22] have been plotted, interpolated, and recalculated in steps of 0.01MeV from threshold energy 3.766MeV up to 85MeV of the incident proton energy in order to obtain the adopted cross sections of this reaction. The results of adopted cross sections and the experimental results for the authors mentioned above are shown in figure 1.

3. $^{68}_{30}Zn_{38}(p,3n)^{66}_{31}Ga_{35}$ **Reaction**

The cross sections data from EXFOR library measured by Mcgee et al. (1970) [8], Levkovskij (1991) [14], Hermanne et al.(1991) [15], Szelecsenyi et al.(1994) [20], Hermanne (1997) [21], Szelecsenyi et al.(1998) [22], Hermanne et al.(1999) [23], Stoll et al.(2002) [24], and Szelecsenyi et al.(2005) [25] have been plotted, interpolated, and recalculated in steps of 0.01MeV from threshold energy 3.766MeV up to 85MeV of the incident proton energy in order to obtain the adopted cross sections of this reaction. The results of adopted cross sections and the experimental results for the authors mentioned above are shown in figure 1.

Table 1. Proton-induced nuclear reactions occur in 68Zn enriched Zinc and theoretical productions yield for irradiation with protons.

Zinc nuclide (abundance in natZn)	Abundance % [1]	Nuclear reaction	Half-life [1]	Decay [1]			Theo. Yield (MBq/μA)	Exp. Yield (MBq/μA) at EOB	Theo. Yield (atom * 1.0E-9) (pw)
				β^+ (MeV)	EC (%)	γ- rays			
68Zn (18.8% [1])	99.5	68Zn(p, n)68Ga	67.8 min	1.899 (88.9%)	11.1	511(178%) 1077(3.24%) 1883(0.142%)	599.0 [4] 657.1(pw)	5032[40] 6124 (pw)	3094 at 12MeV
		68Zn(p, 2n)67Ga	3.26 days		100	93.3(70.6%) 184.6(21.3%) 300.2(16.67%)	23.2 [4] 34.4 (pw)	3.94E-2[40] 7.34E-1(pw)	8956 at 25MeV
		68Zn(p, 3n)66Ga	9.49 h	4.153 (56%)	44	511(114%) 833(5.9%) 1039(37%)	9.31E-03[4] 11.6E-3(pw)	------	5485 at 38MeV

						2190(5.3%) 2751(22.7%)			

In analyzing for induced proton on Zinc to produce, the discrete nuclear states that are populated in ordinary decays have discrete separations, widths, and lifetimes. Thus, to calculate the cross sections at a given incident proton energy of a nuclear state, it is very unlikely that the overlap of the energy distributions of two different states could cause confusion as to the stationary state resulting from the decay. The states are overlap and strongly mixed, these states do not have distinctly observable wave functions because the widths of unstable states are small compared with their separation the states are distinct and observable. Thus the instability of the compound nucleus, results in an uncertainty in the energy of these states.

The stopping power of Zinc target element induced by proton-particles has been calculated in the present work using two methods:

1- SRIM (2003), as experimental results where SRIM is a program build by Ziegler empirical formulae [41].
2- Ziegler empirical formulae as a theoretical calculation results. Using Ziegler coefficients [31,32].

For these calculations, the (stop.m) program has been written in Matlab-8.0 for this purpose. The calculated yields for induced proton on Zinc to produce Gallium are very important quantity as well as the cross sections in analyzing problems of diagnosis, physical therapy, and medicine treatments as the following:

Gallium-68 radionuclide finds significant widely applications used in nuclear medical field. The production possibility in different energy region via different nuclear reactions is clarified in the following discussion:

1- For diagnosing tumors, Gallium-68 tracers are being in many neuroendocrine tumor studies in human.
2- Gallium-68 decays principally by positron emission and it used in conjunction with positron emission tomography (PET) scanners for imaging various organs and their physiological functions. This is an economical source in hospitals.
3- Gallium-68 finds significant application in assessment to detect blood-brain barrier defect to image tumor.

Therefore, the calculated yield for Zinc target was calculated. The main aim of this study is to increase calculated yields for induced proton on Zinc to produce Gallium with low impurity by decreasing the energy

of proton beams to avoid producing different targets. Figure 2 illustrates the stopping power and calculated yields for induced proton on 68Zn to produce 68Ga. In figure 2, the calculated yields seem to depend strongly on the structure of the individual nucleus, the incident proton energy, and stopping power of the target element. In general, the stopping power behavior decreases with the increased calculated yields. This behavior was agreed with the reference [32,33]. It is clear from the calculated results shown in these figures that for the calculated yield values for 20-100% abundance target element, and follow the trend in the asymmetry parameter of proton excess (N-Z)/A so that by increasing this parameter the maximum calculated yields will be decrease. This increment may be attributed to the fact that by decrease the numbers of neutrons the outer shells are populated by an excess calculated which increases the occurrence probability for induced proton on Zinc to produce Gallium.

As generators are expensive and deliver limited activities of 68Ga, the latter can also be produced directly. Among several possible nuclear reactions the irradiation of Zn by protons is the preferred nuclear reaction as it leads to a larger production yield and uses protons is the simplest of all cyclotron projectiles [42]. Since several radioisotopes of gallium with longer half-lives are also produced when irradiating natural Zn, it is necessary to bombard enriched 68Zn. However, the irradiation of 68Zn by protons also leads to the production of the undesired long-lived 67Ga radioisotopic impurity through the 68Zn(p,2n)67Ga reaction for protons with energy higher than 12MeV, and the 66Ga impurity through the 68Zn(p,3n)66Ga reaction for protons with energy higher than 20MeV. This compels us to select the energy beam wisely to achieve high purity radionuclide and be suitable for human use. Therefore, producing 68Ga by using the solid 68Zn target can produce yields of about 6124MBq/μA (165.51Ci/μA) at 12MeV. However, such large activities of 68Ga are achieved at the expense of several important technical difficulties. This process requires significant optimization studies and even additional quality control prior irradiation in order to guarantee that it is suitable for production and to avoid product contamination from the backing layer.

5. Conclusions

In the case of incident proton on Zinc for the production of Gallium, a comparison of all experimental and theoretical results showed that cross section theory was successful in reproducing most experimental data. The recommended excitation functions and calculated integral yields help to optimize the energy range for each nuclear reaction for the production of Gallium seem to be the most useful.

Eventually 68 Zn(p,n)68Ga nuclear reaction requires lower energies to provide significantly higher yields. It provides the use of the enriched material and mainly because facilities with 66Ga and 67Ga at low energy protons are scarce.

We have examined and recommended the excitation function of the 68Zn(*p*,n)68GA, 68Zn(p,2n)67Ga, and 68Zn(p,3n)66Ga nuclear reactions up to 85 MeV. Our new data showed good agreement with the updated values of the only data set available in the literature. Practical production possibilities up 12MeV were investigated in detail for (p,n) reaction for highly enriched 68Zn target with no impurity of 66Ga and 0.012% 67Ga. It could be concluded that the above reaction can be employed for practical purposes at biomedical cyclotrons, however. The available yields are not only large at 12 MeV, but it has no background of 66Ga and a very low contamination of 67Ga. The production was at high energy of irradiation 12MeV which could provide enough activities 6124MBq/μA (165.51Ci/μA).

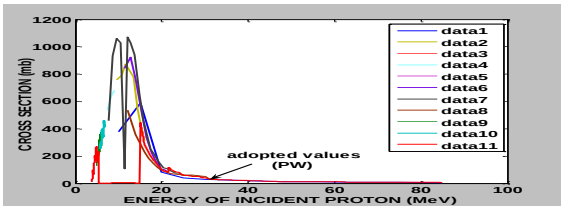


Figure 1 Excitation functions of 68Zn(p,n)68Ga reaction.
Data 1:Ref. No.[8]; Data 2:Ref. No.[9]; Data 3:Ref. No.[10];

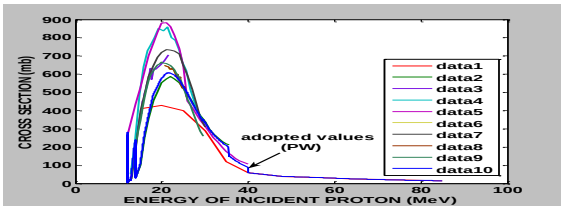


Figure 2 Excitation functions of 68Zn(p,2n)67Ga reaction.
Data 1:Ref. No.[8]; Data 2:Ref. No.[18]; Data 3:Ref. No.[19];

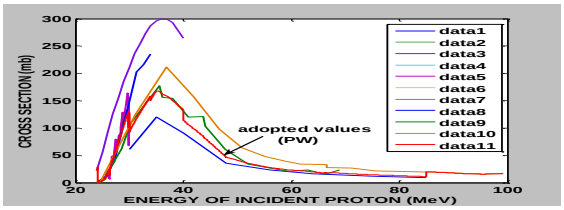


Figure 3 Excitation functions of 68Zn(p,3n)66Ga reaction.
Data 1:Ref. No.[9]; Data 2:Ref. No.[14]; Data 3:Ref. No.[15];

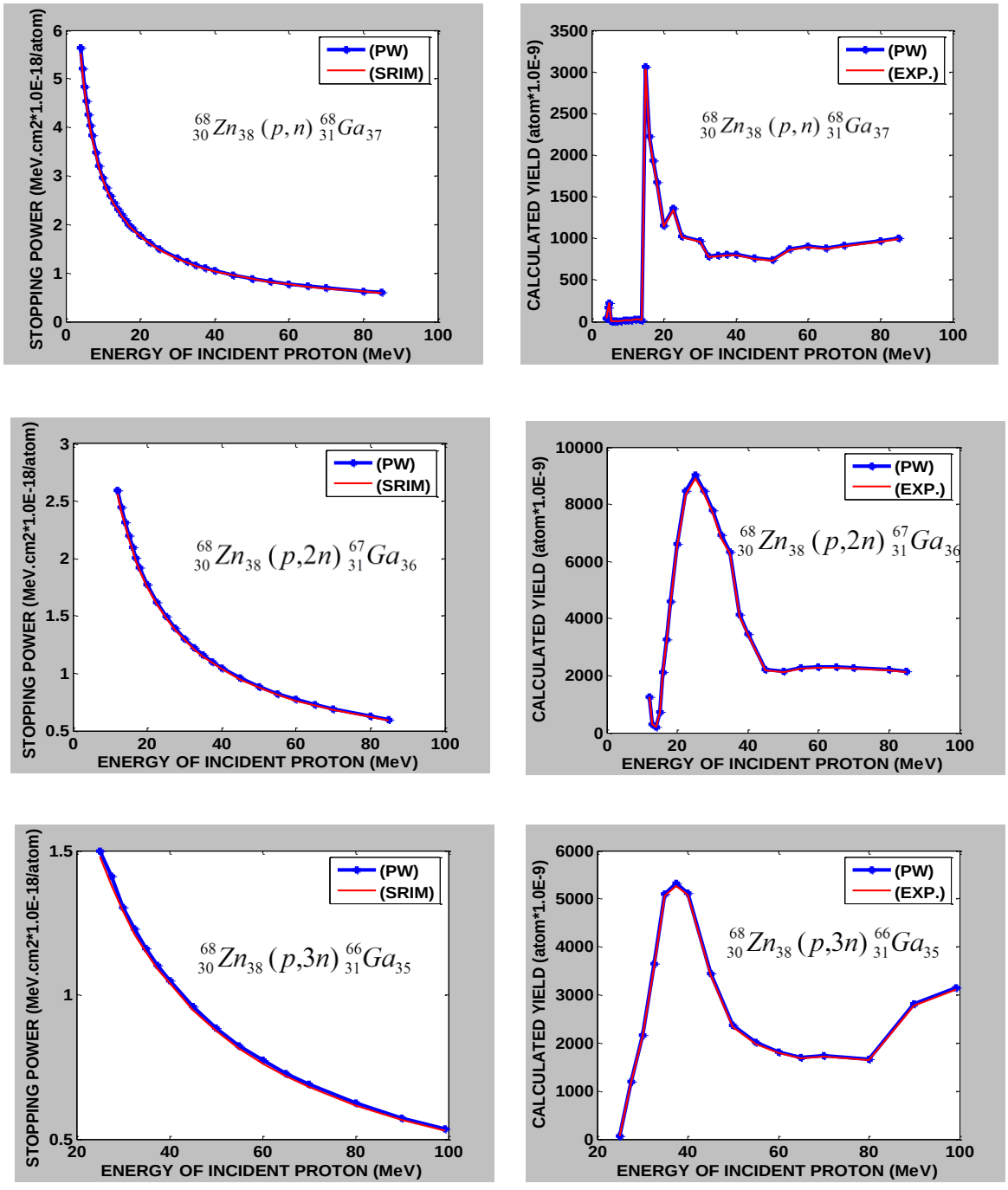


Figure 2 Left side; calculated and SRIM (2003) stopping power verses incident proton energy on 68Zn. Right side; calculated and experimental yield based on the adopted cross section of incident proton on 68Zn reaction.

Acknowledgments

The authors wish to express their deep thanks and gratitude to the International Atomic Energy Agency (IAEA) to allow access to the experimental nuclear data libraries and made available.

References

1. NuDat 2.7 database. Available from: http://www.nndc.bnl.gov/ nudat2.
2. Carroll, V., Demoin, D. W., Hoffman, T. J., Jurisson, S. S., Inorganic chemistry in nuclear imaging and radiotherapy: current and future directions, Radiochim. Acta **100**, 653–667 (2012).
3. IAEA Technical Reports Series No. 473: Nuclear Data for the Production of Therapeutic Radionuclides, (Technical Editors: S. M. Qaim, F. Tarkanyi, and R. Capote) ISBN 978-92-0-115010-3 (2011).
4. Alves V. H. P., Do Carmo S. J. C., Neves A. C. B., and Silva M., Production of copper-64 and gallium-68 with a medical cyclotron using liquid targets, Modern Physics Letters A, **32**(17), 1740013-(1-21) (2017). DOI: 10.1142/S0217732317400132
5. Experimental Nuclear Reaction Data (EXFOR) (Database Version of February 26, 2013). Available from: www-nds.iaea.org/exfor.
6. Ziegler J.F., Stopping and Ranges Elements, Helium Pergamum Press, Oxford, **4** (1977).
7. SRIM (2003) program, experimental results from the original work by: J.F.Ziegler, The Stopping and Range of Ions in Matter, **2-6**, Pergamon Press (1977-1985).
8. Mcgee T., Rao C.L., Saha G.B. and Yaffe L., Nuclear Interactions of $^{45}_{21}Sc_{24}$ and $^{68}_{30}Zn_{38}$ with Protons of Medium Energy, Nuclear Physics, Section A, **150**, 11 (1970).
9. Hille M., Hille P., Uhl M., and Weisz W., Excitation Functions of (p,n) and (α,n) Reactions on $_{28}Ni$, $_{29}Cu$ and $_{30}Zn$, Nuclear Physics, Section A, **198**, 625 (1972).
10. Barrandon J. N., Debrun J. L., Kohn A., and Spear R. H., A Study of the Main Radioisotopes Obtained by Irradiation of $_{22}Ti$, $_{23}V$, $_{24}Cr$, $_{26}Fe$, $_{28}Ni$, $_{29}Cu$ and $_{30}Zn$ with Protons from 0 to 20MeV, Nuclear Instrument and Methods in Physics Res., **127**, 269 (1975).
11. Kotelnikova G. V., Lovchikova G. N., Salnikov O. A., Simakov S. P., Trufanov A.M.and Fetisov N.I., The Investigation of Neutron Energy Spectra For $^{68}_{30}Zn_{38}(p,n)^{68}_{31}Ga_{37}$ Reaction, Fiz.-Energy Institut, Obninsk Reports, No.1141, P.1 (1980).
12. Esat M. T., Spear R. H., Zyskind J. L., Shapiro M. H., Fowler W. A., and Davidson J. M., Test of global Hauser-Feshbach calculations for proton-induced reactions on $^{68}_{30}Zn_{38}$, Physical Review, Part C, Nuclear Physics, **23**, 1822 (1981).
13. Tarkanyi F., Szelecsenyi F., Kovacs Z., and Sudar S., Excitation functions of proton induced nuclear reactions on enriched $^{66}_{30}Zn_{36}$, $^{67}_{30}Zn_{37}$ and $^{68}_{30}Zn_{38}$ production of $^{67}_{31}Ga_{36}$ and $^{66}_{31}Ga_{35}$, Radiochimica Acta, **50**, 19 (1990).
14. Levkovskij V.N., Activation cross section nuclides of average masses (A=40-100) by protons and alpha-particles with average energies (E=10-50 MEV), Act. Cs. By Protons and Alphas, Moscow (1991).
15. Hermanne A., Walravens N., and Cicchelli O., Optimization of isotope production by cross section determination, Journal Nuclear Data for Science and Technology, **6**(2), 19-24 (1991).

16. Vinogradov V. M., Zhuravlev Yu. Yu., Zarubin P. P., Kolozhvari A. A., Sergeev V. O., and Sitnikova I. V., Excitation functions of (p,n) reactions on zinc isotopes in the range of E(p) from 4.9 to 5.9MeV, Bull.Russian Academy of Sciences - Physics, **57**, 906 (1993).
17. Zhuravlev Yu. Yu., Zarubin P. P., Zeic Yu. V., Kolozhvari A. A., and Chelgunov I. V., Excitation functions of (p,n) reactions on nuclei of isotopes Zn from E(p)=5.6 To 6.8 MeV, Izv. Rossiiskoi Akademii Nauk, Ser. Fiz., **59**, 118 (1995).
18. Little E., and Lagunas-solar C., Cyclotron production of $^{67}_{31}Ga_{36}$ Cross sections and thick-target yields for the $^{67}_{30}Zn_{37}(p,n)$ and $^{68}_{30}Zn_{38}(p,2n)$ reactions, Applied Radiation and Isotopes, **34**, 631 (1983).
19. Tarkanyi F., Szelecsenyi F., Kovacs Z., and Sudar S., Excitation functions of proton induced nuclear reactions on enriched $^{66}_{30}Zn_{36}$, $^{67}_{30}Zn_{37}$ and $^{68}_{30}Zn_{38}$ production of $^{67}_{31}Ga_{36}$ and $^{66}_{31}Ga_{35}$, Radiochimica Acta, **50**, 19 (1990).
20. Szelecsenyi F., Boothe T. E., Tavano E., and Plitnikas M. E., New cross section data for $^{66,67,68}_{30}Zn+p$ reactions up to 26 MeV, Conf. on Nucl. Data for Sci. and Techn., Gatlinburg, P.393 (1994).
21. Hermanne A., Evaluated cross section and thick target yield data of $_{30}Zn+P$ processes for practical applications, Nuclear Physics, Section A, **4**(8), 261(1997).
22. Szelecsenyi F., Boothe T. E. , Takacs S., Tarkanyi F., and Tavano E., Evaluated cross section and thick target yield data bases of $_{30}Zn+P$ processes for practical applications, Applied Radiation and Isotopes, **49**, 1005 (1998).
23. Hermanne A., Szelecsenyi F., Sonck M., Takacs S., Tarkanyi F., and Vanden Winkel P., New Cross Section Data on $^{68}_{30}Zn_{38}(p,2n)^{67}_{31}Ga_{36}$ and $^{nat}_{30}Zn(p,xn)^{67}_{31}Ga_{36}$ Nuclear Reactions for the Development of a Reference Data Base, Journal of Radioanalytical and Nuclear Chemistry, **240**, 623 (1999).
24. Stoll T., Kastleiner S., Shubin Yu. N., Coenen H. H., and Qaim S. M., Excitation functions of proton induced reactions on $^{68}_{30}Zn_{38}$ from threshold up to 71 MeV, with specific reference to the production of $^{67}_{29}Cu_{38}$, Radiochimica Acta, **90**, 309 (2002).
25. Szelecsenyi F., Steyn G. F., Kovacs Z.,Vander Walt T. N., Suzuki K., Okada K. and Mukai K., New cross-section data for the $^{66}_{30}Zn_{36}(p,n)^{66}_{31}Ga_{35}$, $^{68}_{30}Zn_{38}(p,3n)^{66}_{31}Ga_{35}$, $^{nat}_{30}Zn(p,x)^{66}_{31}Ga_{35}$, $^{68}_{30}Zn_{38}(p,2n)^{67}_{31}Ga_{36}$ and $^{nat}_{30}Zn(p,x)^{67}_{31}Ga_{36}$ nuclear reactions up to 100 MeV, Nucl. Instrum. Methods in Physics Res., Sec. B, **234**(4), 375 (2005).
26. Lindhard H. H., and Schardff M., Hydrogen Stopping Powers and Ranges in all Elements, Phys. Rev, **124**, 128 (1961).
27. Powers D. and Olson H. J., Evaluated cross section and thick target yield data bases of Zn+p processes for practical applications, Phys. Rev, **73**, 2271(1980).
28. Beth H. A., New cross sections and inter comparison of proton monitor reactions on Gallium, Bloch F., Ann. Phys., **16**, 285 (1933).
29. Lindhard H. H., and Winther A., limiting factor for the progress of radionuclide based diagnostics and therapy, Mat. Fys. Medd. Dan. Vid. Selsk., **34**, 264 (1964).
30. Ashely J. C., Ritchi R. H., and Brant W., In vitro and in vivo evaluation of copper 64-octreotide conjugates, Phys. Rev., B5, 2329 (1972).
31. Ziegler J. F., Handbook of Stopping Cross-Sections for Energetic Ions in all Elements, Pergamon Press, Oxford, **5** (1980).

32. Andersen H. H., and Ziegler J. F., Hydrogen Stooping Powers and Ranges in all Elements, Pergamon Press, Oxford, **3** (1977).
33. Nukulin V. Ya., and Polukhin S. N., Saturation of the Neutron Yield from Mergajoule Plasma Focus Facilities, Journal of Plasma Physics, **33**(4), 304 (2007).
34. Becturts K. H. and Wirtz K., Neutron Physics, Springer (1964).
35. Norman E. B., Chupp T. E., Lesko K. T., and Schwalbac P., Differential neutron production cross sections for 800-MeV protons, Nucl. Phys. A, **390**, 561 (1982).
36. Vora M. M., Schlyer D. J., Van den Winkel P., Ruth T. J., Cyclotron produced radionuclides: Principles and Practice Technical Report, International Atomic Energy
37. Francisco José Cerqueira Alves. Estudo da viabilidade de produção de carbono-10 em ciclotrões de baixa energia para utilização em PET. PhD thesis, Universidade de Coimbra (2002).
38. Knole G. F., Radiation Detection and Measurement, John Wiley and Sons, 90-92 (2000).
39. Experimental Nuclear Reaction Data (EXFOR).https://www-nds.iaea.org/exfor/exfor.htm.
40. Sadeghi M., Kakavand T., Rajabifar S., Mokhtari L., and Nezhad A. R., Cyclotron production of 68Ga via proton-induced reaction on 68Zn target, Nukleonika, **54**(1), 25−28 (2009).
41. Sabet M., Rowshanfarzad P., Jalilian A. R., Ensaf M. R., and Rajamand A. A., Production and quality control of $^{66}_{31}Ga_{35}$ radionuclide, Nukleonika Journal , **51**(3), 147 (2006).
42. Szelecsenyi F., Kovacs Z., Nagatsu K., Fukumura K., Suzuki K., and Mukai K., Radiochim. Acta, **100**, 5 (2012).

Spectral Efficiency Improvement of the Optical Communication Systems

Thanaa Hussein Abd[*1], Ali Abdul Khadhum Ruhaima[*2] and Dunya Mohee Haider[*3]
*Corresponding author: eng_thanaa@yahoo.com (Thanaa H.A)
*[1,2] Al-Nisour University Collage
[*3]Madent AlelmUninersity Collage

Abstract

In this paper, different modulation techniques with the polarization-interleaving design are examine and compared to improve the spectral efficiency of the high capacity Wavelength Division Multiplexing (WDM) systems.

The system is simulated with channel at a bit rate of 40 Gb/s to evaluate the performance with no Stimulated Raman Scattering (SRS), Cross Phase Modulation (XPM) and Four Wave Mixing (FWM) effects.

In order to evaluate the modulation formats performance, Q factor verses lunched power curves and the eye diagrams are presented. A reference Q factor to evaluate the performance of the formats is taken as 17 dB, which corresponds to a BER of 1×10^{-12}. The communication system is simulated by the (Optisystem v 7.0) simulation package.

Keywords: Wavelength Division Multiplexing (WDM) systems Stimulated Raman Scattering (SRS), Cross Phase Modulation (XPM) and Four Wave Mixing (FWM)

1. Introduction

In order to meet the ever-increasing demand in telecommunication capacity, fiber optic communication systems have been evolving dramatically over the past decade. The fiber optic communication traffic growth has been at a rate of about 2 dB per year, representing a traffic increase of a factor of 100 in 10 years. The capacity increase in fiber optic communication systems has been achieved mainly by deploying more fiber links, populating more wavelength channels per fiber link through dense wavelength-division-multiplexing (DWDM), and increasing the data rate per wavelength channel [1].

Transmission through optical fiber cables plays an important role in the communication networks of nowadays. The main motivation for this is the enormous potential bandwidth of optical fiber (>100 THz), which contains several orders of magnitude larger than the bandwidth of copper media like coaxial cables or twisted wire pairs [2].

Fig.(1) compares the most widely used radio frequency (RF) and optical communication technologies for wireline and wireless communications, based on this metric. Fig.(1) shows clearly that optical communication systems can support Tb/s capacities over many thousand kilometers, which makes them the ideal technology base for high-capacity wireline networking [3].

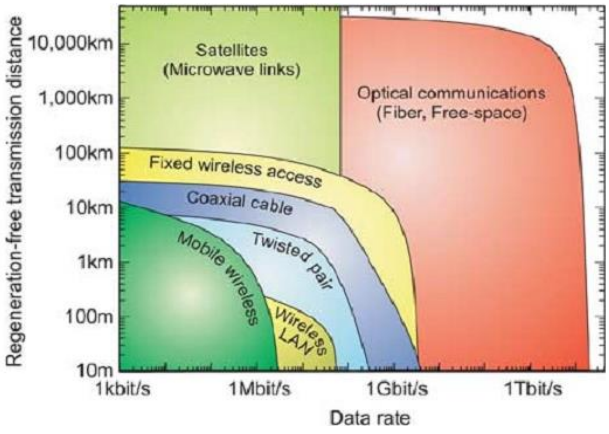


Figure. 1 Regeneration- free transmission distance versus data rate for various wireless and wireline communication technologies [3].

However, this project is aimed to improve the spectral efficiency of the optical fiber communication systems by using various amplitude and phase modulation formats. The performance and robustness of the following modulation formats against channel impairments are tested using (Optisystem v.70) software package. The tested modulation formats are:

1. Non-Return to Zero On-Off Keying (NRZ-OOK).
2. Return to Zero On-Off Keying (33% RZ-OOK and 67% RZ-OOK).
3. Return to Zero Duobinary (RZ-DUO).
4. Return to Zero Alternate Mark Inversion (RZ-AMI).
5. Non-Return to Zero Differential Phase Shift Keying (NRZ-DPSK).
6. Return to Zero Differential Phase Shift Keying (33% RZ-DPSK and 67% RZ-DPSK).
7. Non-Return to Zero Differential Quadrature Phase Shift Keying (NRZ-DPSK).

This paper is organized as follows: In section II, the evaluation of optical fiber communication systems is presented. Section III, shows the system capacity and spectral efficiency. The performance analysis of the new proposed system is presented, and finally, conclusions are given in section IV.

2. Evolution of optical fiber communication

The evolution of optical fiber communication systems can be divided into several generations. The first generation of lightwave systems operated near 0.8 μm and used Gallium Arsenide (GaAs) semiconductor lasers. After several field trials during the years 1977-79, such systems became available commercially in 1980. They operated at a bit rate of 45 Mb/s and allowed repeater spacing of up to 10 km. The larger repeater spacing compared with 1-km spacing of coaxial systems was an important motivation for system designers because it decreased the installation and maintenance costs associated with each repeater.

It was clear during the 1970s that the repeater spacing could be increased considerably by operating the lightwave system in the wavelength region near 1.3 μm, where fiber loss is below 1 dB/km. Furthermore, optical fibers exhibit minimum dispersion in this wavelength region.

This realization led to a worldwide effort for the development of Indium Gallium Arsenide Phosphide (InGaAsP) semiconductor lasers and detectors operating near 1.3 μm [4].

The second generation of fiber-optic communication systems became available in the early 1980s, but the bit rate of early systems was limited to below 100 Mb/s because of dispersion in multimode fibers. This limitation was overcomed by the use of single-mode fibers. In 1987, the second-generation lightwave systems, operating at bit rates of up to 1.7 Gb/s with a repeater spacing of about 50 km, were commercially available. The repeater spacing of the second-generation lightwave systems was limited by the fiber losses at the operating wavelength of 1.3 μm (typically 0.5 dB/km). Losses of silica fibers became minimum near 1.55 μm. Indeed, a 0.2 dB/km loss was realized in 1979 in this spectral region.

However, the introduction of third generation lightwave systems operating at 1.55 μm was considerably delayed by large fiber dispersion near 1.55 μm. Conventional InGaAsP semiconductor lasers could not be used because of pulse spreading occurring as a result of simultaneous oscillation of several longitudinal modes. The dispersion problem can be overcomed either by using dispersion-shifted fibers designed to have minimum dispersion near 1.55 μm or by limiting the laser spectrum to a single longitudinal mode. Both approaches were followed during the 1980s [4].

The third generation lightwave systems operating at 2.5 Gb/s became available commercially in 1990. Such systems are capable of operating at a bit rate of up to 10 Gb/s. The best performance is achieved by using dispersion-shifted fibers combined with lasers oscillating in a single longitudinal mode. A drawback of the third-generation 1.55 μm systems is that the signal is regenerated periodically by using electronic repeaters spaced apart typically by 60-70 km. The spacing of the repeaters can be increased by making use of a homodyne or heterodyne detection scheme because its use improves the receiver sensitivity. Such systems are referred to as coherent lightwave systems. Coherent systems were under worldwide development during the 1980s, and their potential benefits were demonstrated in many system experiments. However, commercial introduction of such systems was postponed with the advent of fiber amplifiers in 1989.

The fourth generation of lightwave systems makes use of optical amplification for increasing the repeater spacing and of wavelength-division multiplexing (WDM) for increasing the bit rate. In most WDM systems, fiber losses are compensated periodically by using Erbium-Doped Fiber Amplifiers (EDFA) spaced 60-80 km apart. Such amplifiers were developed after 1985 and became available commercially in 1990. This performance indicated that an amplifier-based, all-optical, submarine transmission system was feasible for intercontinental communication. By 1996, not only transmission over 11,300 km at a bit rate of 5 Gb/s had been demonstrated by using actual submarine cables, but commercial transatlantic and transpacific cable systems also became available. Since then, a large number of submarine lightwave systems have been deployed world widely [4].

3. System Capacity and Spectral Efficiency

The use of WDM can increase the system capacity because it transmits multiple bit streams over the same fiber simultaneously. When N channels at bit rates B_1, B_2, … B_N are transmitted simultaneously over a fiber of length L, the total bit rate of the WDM link becomes:

$$B_T = B_1 + B_2 + \cdots B_N \qquad (1.1)$$

For equal bit rates, the system capacity is enhanced by a factor of N. The most relevant design parameters for a WDM system are the number of channels N, the bit rate B at which each channel operates, and the frequency spacing $\Delta\nu_{ch}$ between two neighboring channels.

The product NB denotes the system capacity and the product $N\Delta\nu_{ch}$ represents the total bandwidth occupied by a WDM system [4].

The spectral efficiency (η_s) is defined as [4]:

$$\eta_s = \frac{B}{\Delta v_{ch}} \qquad (1.2)$$

Where B is the single-channel bit rate and Δv_{ch} is the channel spacing. The units of the spectral efficiency is (b/s/Hz). The throughput of a Dense Wavelength Division Multiplexing (DWDM) transmission system can be increased in many ways by using a wider optical bandwidth, by increasing spectral efficiency, or by some combination of the two. Utilizing a wider bandwidth typically requires additional amplifiers and other optical components, therefore; raising spectral efficiency is often the more economical alternative [5].

4. 1×40 Gb/s System Design and Description

The performance of the single channel system where there are no SRS, XPM or FWM effects analyzed. In this section, a description of the single channel system is presented. The system is simulated at 40 Gb/s data rate and the number of samples per bit is 64.

A. The structure of the transmitter

The transmitter section is shown in fig.(2):

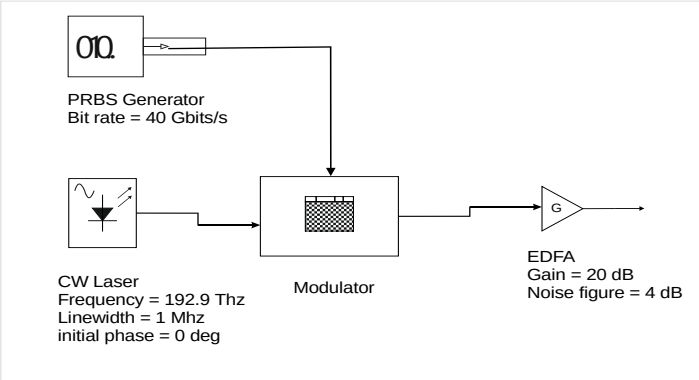


Figure. 2 The transmitter section of the simulated single channel system.

A laser source is used, which has a narrow beam width. The parameters of the simulated CW laser are shown in table (1).

Table (1) Simulation parameters of the CW laser.

Parameter	Value
Frequency	192.9 THz
Power	(-5) dBm to 10 dBm
Linewidth	1 MHz
Initial phase	0

The optical signal is then applied to the modulator, which is used to modulate the electrical data with the optical carrier generated by the CW laser.

The Mach-Zehnder modulator is an intensity modulator based on an interferometric principle. In order to decrease the chirp induced in the optical signal to increase the transmission distance, a chirp free MZM structure is used in the modulation process. The value of the switching bias voltage is taken as 4 V for all the modulation formats. The value of the bias voltage of the two-modulator arms differs from modulation format to another.

After the modulation of the electrical data with the optical signal, an EDFA amplifier is used as a booster amplifier to compensate the losses incurred in the transmitter and then boosts the signal in the optical fiber channel. The EDFA is operated in the power control mode with a gain of 20 dB and a noise figure of 4 dB.

B. Optical fiber channel

Then the signal is lunched in the optical fiber channel. An ITU-T G.652 optical fiber is used. In order to decrease the pulse broadening resulting from the chromatic dispersion effects, a dispersion compensating fiber (DCF) is used in the simulation of the fiber channel.

The simulation parameters of the optical fibers channel are listed in table (2).

Table (2) Simulation parameters of the SMF, the DCF, and the system losses.

ITU-T G.652 Fiber, 60 km span length	
α (dB/km)	0.2
Dispersion parameter D (ps/(nm.km))	17

Dispersion slope S (ps/(km.nm2))	0.075
Effective area (μm2)	70
DGD parameter (ps/$\sqrt{km}$)	0.1
n2 (m2/W)	2.6e (-20)
SBS Threshold	-7 dBm
SRS Threshold	32 dBm
DCF parameters	
α (dB/km)	0.5
Dispersion parameter D (ps/(nm.km))	-85
Dispersion slope S (ps/(km.nm2))	-0.3
Effective area (μm2)	22
DGD parameter (ps/$\sqrt{km}$)	0.1
n2 (m2/W)	2.6e (-20)
SBS Threshold	-8 dBm
SRS Threshold	31 dBm
System losses	
Connecter loss	0.25 dB
Splicer loss	0.15 dB
Insertion loss	4 dB
Number of connectors	4
Number of splicers	1
Number of multiplexers	3
Margin loss	6
Total system loss	34.15B

c. The structure of the receiver

After the propagation in the optical fiber channel, the signal is then applied to the receiver to detect it and convert it into an electrical signal. In general, there are two connection types of the optical receiver depending on the type of the modulation format used in the simulation. For the amplitude and the pseudo

multilevel modulation formats, a direct detection scheme is used in the detection process. The block diagram of the direct detection scheme is shown in fig.(3).

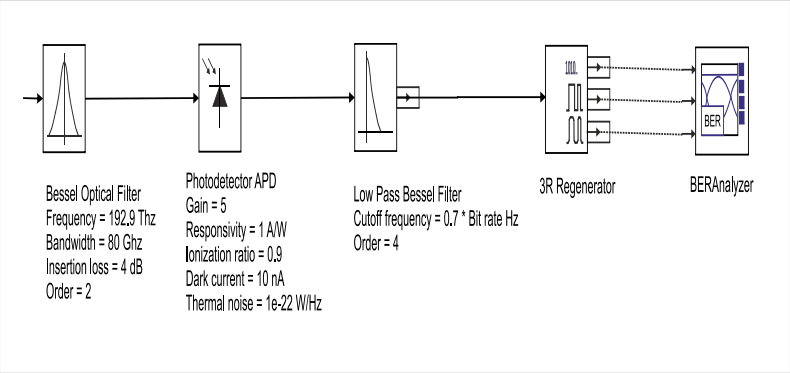
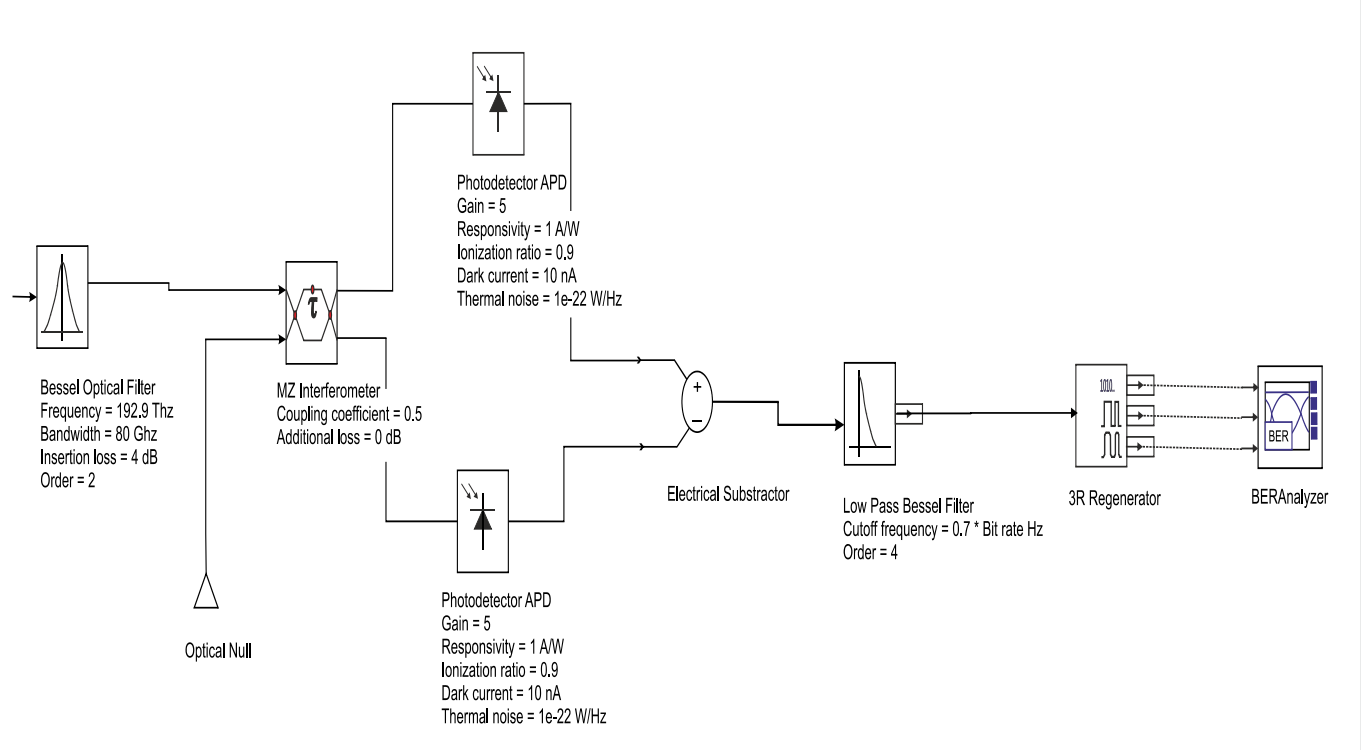


Fig.(3) Block diagram of the direct detection scheme.

An optical band pass Bessel filter of a second order with a center frequency of 192.9 THz and 80 GHz bandwidth is used in front of the photodetector to decrease the noise and the crosstalk in the received signal. Then an Avalanche Photodiode (APD) is use in to convert the optical signal into electrical signal. The simulation parameters of the direct detection scheme are listed in table (3).

Table (3) Simulation parameters of the receiver system diagram.

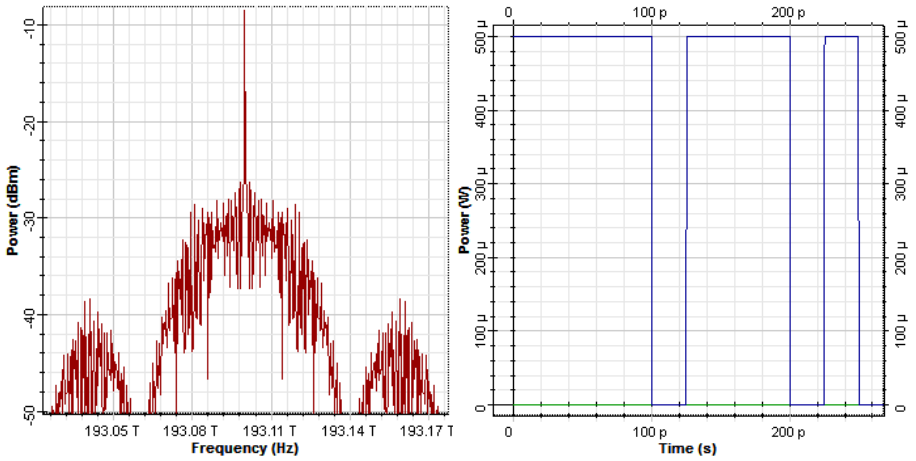
Parameter	Value
Optical band pass filter	2rd order Bessel filter
Center frequency	192.9
Optical bandwidth	80 GHz
Insertion loss	4 dB
Photodetector type	APD
Gain	5
Responsivity	1 A/W
Ionization rate	0.9
Dark current	10 nA
Thermal noise	1e-22 W/Hz
Electrical low pass filter	4th order Bessel filter
Cutoff frequency	28 GHz



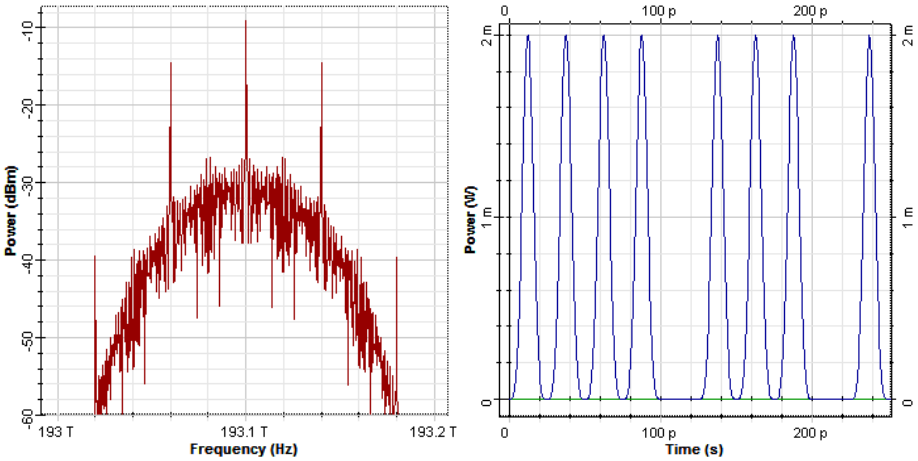
For the phase modulation formats, a balanced detector is used to correlate each bit with a delayed version of it is used in the detection process. Its block diagram is shown in fig.(4).The received signal is first applied to an optical band pass Bessel filter which is the same as the one used in the direct detection scheme. Then an MZI is used to split the received signal into two equal parts. Two APDs and then low pass Bessel optical filter with the same simulation parameters of the one used in the direct detection scheme are used in the balanced scheme. This scheme is used to make a correlation between every bit and a 1-bit delayed version of it.

5. Performance Analysis

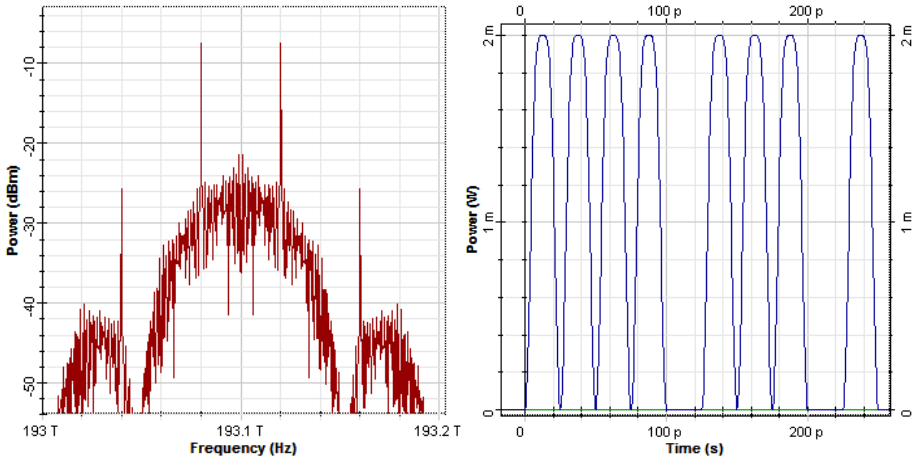
Fig.(5) shows the optical spectrum and waveform of the various modulation formats considered in the simulations at 40 Gb/s. Table (4) shows the (SE) of the modulation formats considered. It shows that in general DQPSK formats have the highest values of the (SE) among the compared modulation formats because of their reduced bandwidth.



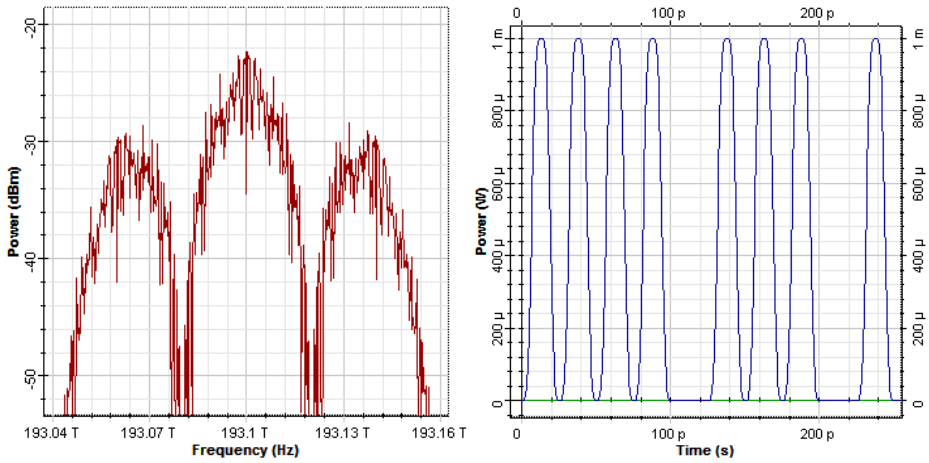
NRZ-OOK



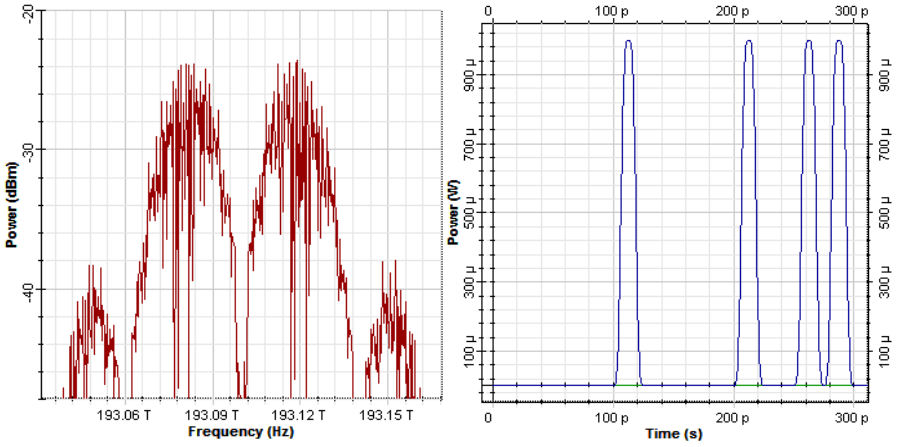
33% RZ-OOK



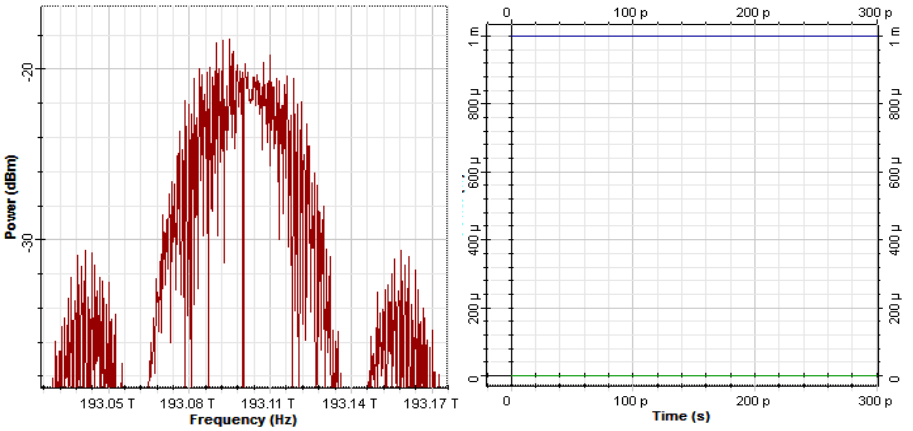
67% RZ-OOK



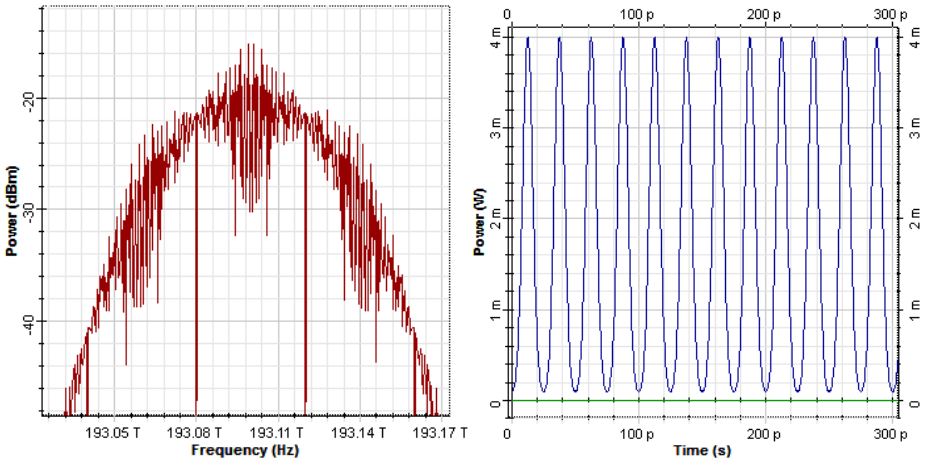
RZ-DUO



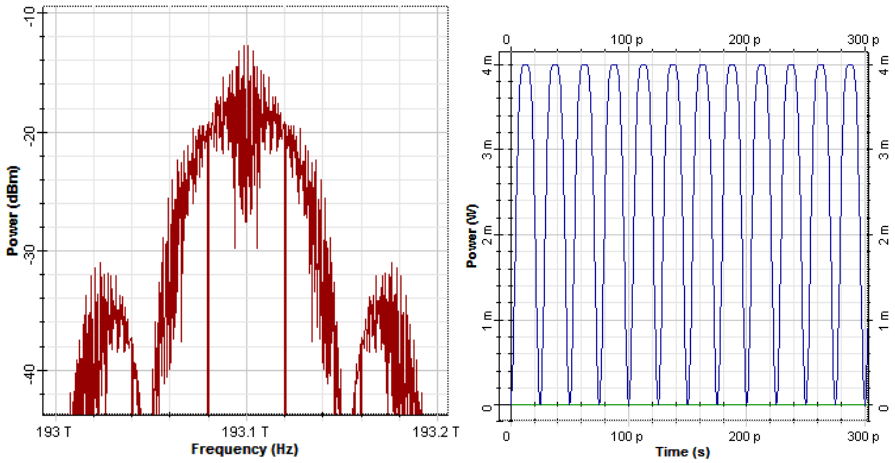
RZ-AMI



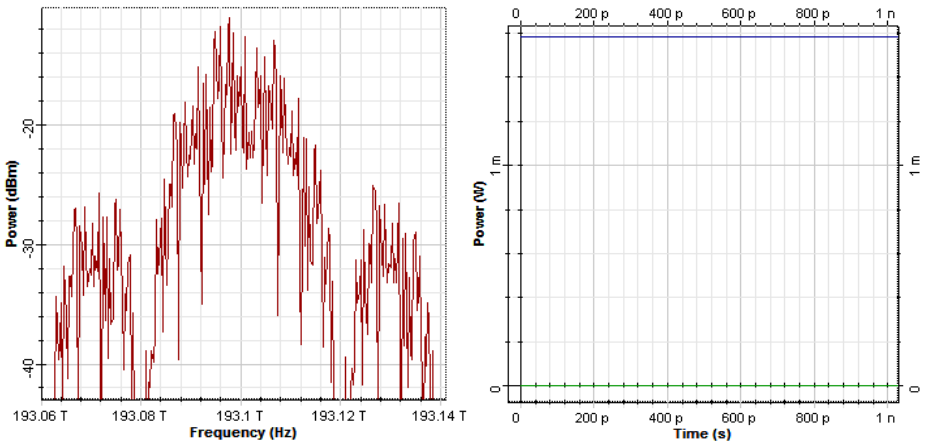
NRZ-DPSK



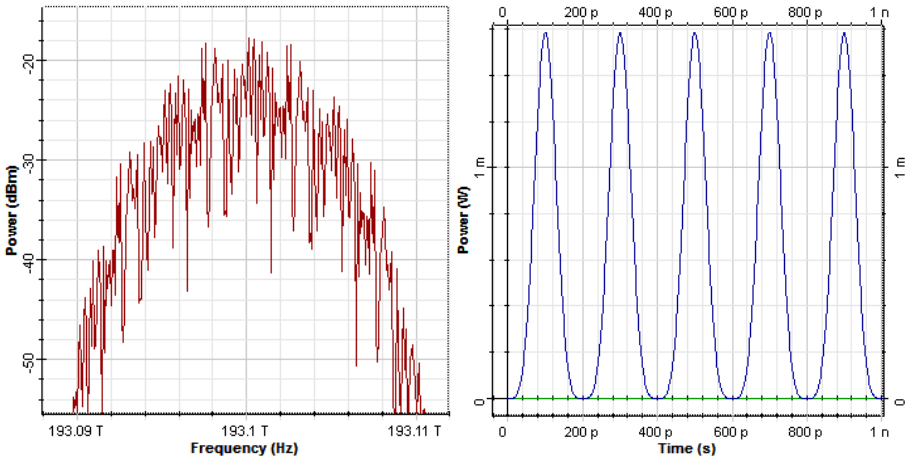
33% RZ-DPSK



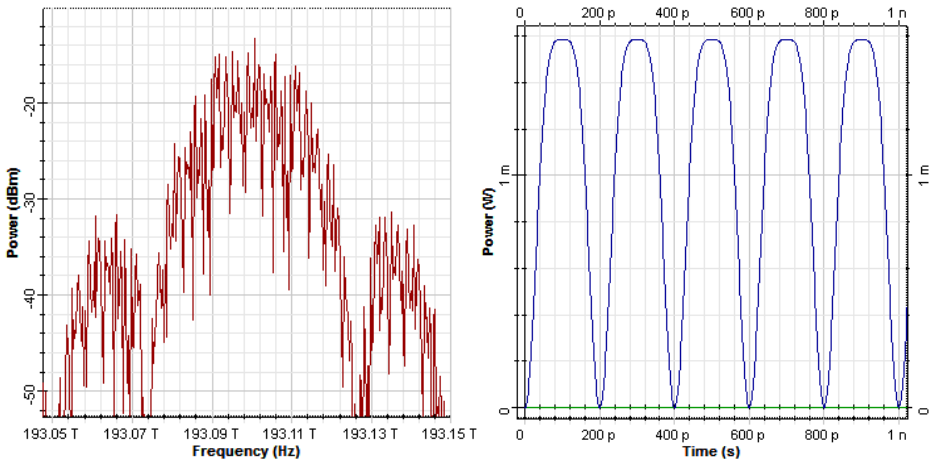
67% RZ-DPSK



NRZ-DQPSK



33% RZ-DQPSK



67% RZ-DQPSK

Fig.(5) Optical Spectrum and waveform of the simulated modulation formats at 40 Gb

Table (4) Spectral Efficiency (SE) of the simulated modulation formats.

Modulation Format	SE (b/s/Hz)	Modulation Format	SE (b/s/Hz)
NRZ-OOK	0.5	33% RZ-DPSK	0.25
33% RZ-OOK	0.27	67% RZ-DPSK	0.4
67% RZ-OOK	0.33	NRZ-DQPSK	1
RZ-DUO	0.33	33% RZ-DQPSK	0.5
RZ-AMI	0.5	67% RZ-DQPSK	0.8
NRZ-DPSK	0.5		

Fig.(6) Optical eye diagrams of the simulated modulation formats at the input (to the left) and at the output of the single channel (to the right).

A complete description of the eye diagram is presented in appendix D. Now a comparison of the modulation formats at the single channel system is presented. Fig.(7) shows a comparison of the amplitude modulation formats for the single channel system at five spans. The results are summarized in table (5).

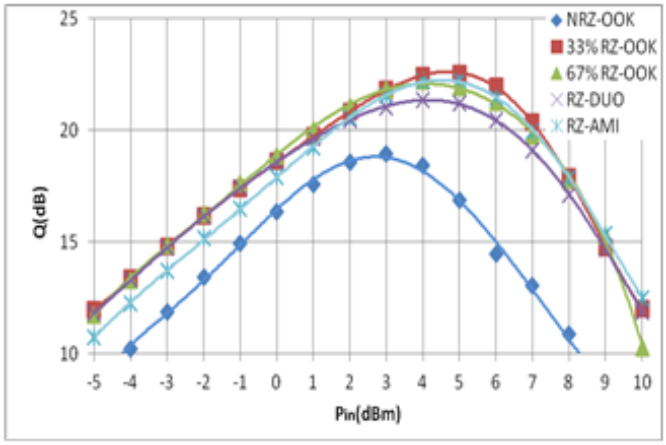


Fig.(7) Comparison between amplitude modulation formats for single channel system.

Table (5) Maximum Q factor, Power range and the spectral efficiency of amplitude modulation formats at single channel system.

Modulation format	Max. Q (dB)	P (dBm)	SE (b/s/Hz)
NRZ-OOK	18.6	0.5-4	0.5
33% RZ-OOK	22.6	(-1.4)-8.4	0.27
67% RZ-OOK	22	(-1.4)-8.4	0.33
RZ-DUO	21.4	(-1.4)-8.4	0.33
RZ-AMI	22.2	(-0.6)-8.4	0.5

Fig (8) shows a comparison of the phase modulation formats for the single channel system at five spans. The results are summarized in table (6).

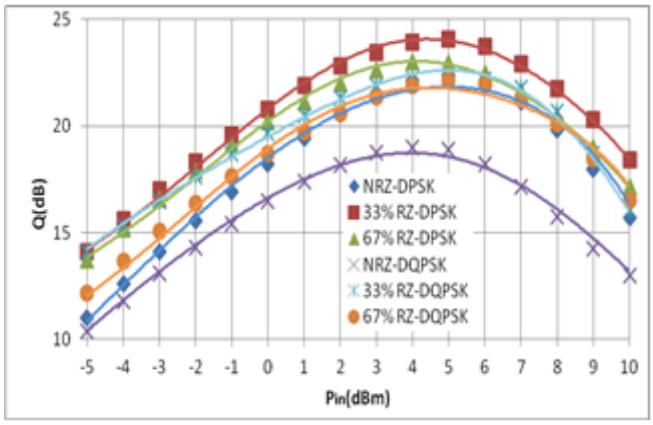


Fig.(8) Comparison between phase modulation formats for single channel system.

Table (6) Maximum Q factor, Power range and the spectral efficiency of phase modulation formats at single channel system.

Modulation format	Max. Q (dB)	P (dBm)	SE (b/s/Hz)
NRZ-DPSK	21.8	(-1)-9.8	0.5
33% RZ-DPSK	24	(-2.8)-…	0.25
67% RZ-DPSK	23	(-2.4)-10	0.4
NRZ-DQPSK	18.8	0.8-7.2	1
33% RZ-DQPSK	22.6	(-2.6)-9.6	0.5
67% RZ-DQPSK	21.8	(-1.4)-10	0.8

Fig.(9) shows a comparison between the NRZ modulation formats in single channel system and five spans.

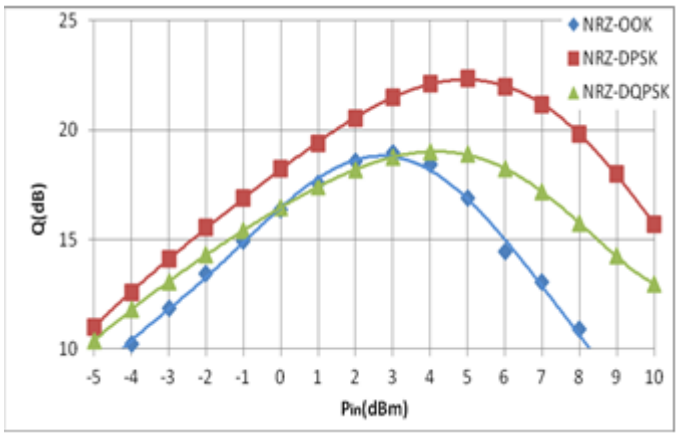


Fig.(9) Comparison between NRZ modulation formats for single channel system.

Fig.(10) shows a comparison between the 33% RZ and the pseudo multilevel modulation formats in single channel system at five spans.

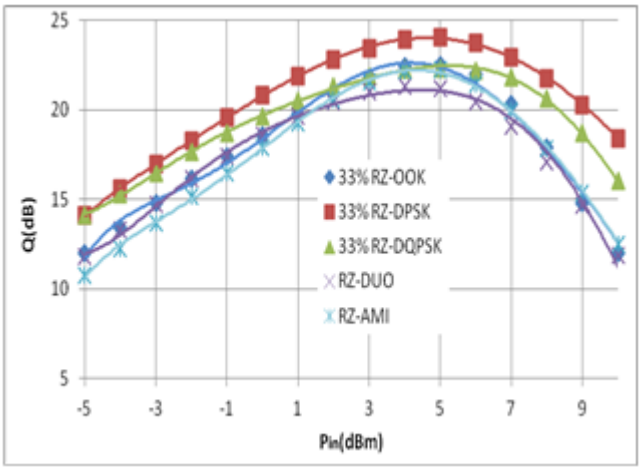


Fig.(10) Comparison between 33% RZ and the pseudo multilevel modulation formats at single channel system.

Fig (11) shows a comparison between the 67% RZ modulation formats in single channel system at five spans.

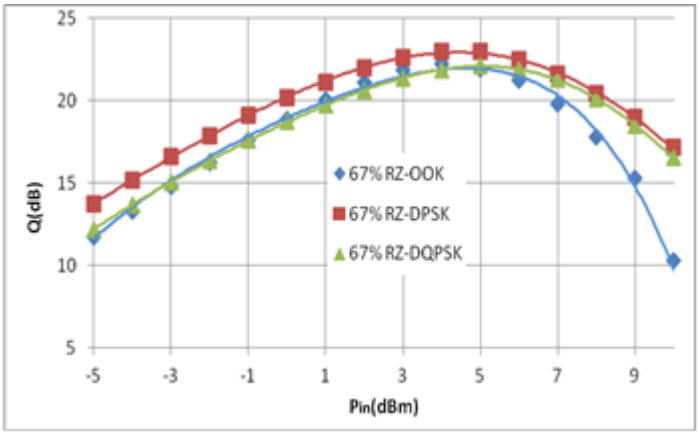


Fig.(11) Comparison between 67% RZ modulation formats at single channel system.

From the obtained results, it is shown that 33% RZ-OOK, 67% RZ-OOK and RZ-AMI nearly have the same Q factor. Among these formats, RZ-AMI gives the maximum (SE) of 0.5 (b/s/Hz) and the optimum performance because of the decreased bandwidth and the π phase shift between the alternate 1-bit. Although NRZ-OOK format has a (SE) of 0.5 (b/s/Hz) it has the worst performance. Its reduced bandwidth comes at the cost of the increased chirp in the signal.

So, that DPSK formats perform better than DQPSK formats. This is because the Euclidean distance for DPSK formats larger than that of the DQPSK formats, which makes them perform better with respect to the noise and they require less OSNR.

As well, the 33% RZ-DPSK gives the optimal performance as it gives the highest Q factor and power range. A disadvantage of 33% RZ-DPSK is its small value of the (SE) with only 0.25 b/s/Hz.

The performance of the 67% RZ-DPSK has a higher Q factor than 67% RZ-DQPSK with an (SE) of 0.4 b/s/Hz, but after 6 dBm, its performance is nearly the same as that of 67% RZ-DQPSK with only 1 dB difference because of the nonlinear effects that begin to rise at the high power region.

Conclusion

The conclusion can be summarized as following points:

1. The eye diagrams shown in the last chapter shows the effectiveness of using the polarization interleaving in reducing the nonlinear effects and the crosstalk especially at low channel spacing.
2. The modulation formats with the highest spectral efficiency has the least value of Q factor. For example, by looking at table (2.8), it is shown that among the DQPSK formats, NRZ-DQPSK has the highest (SE) of 1 b/s/Hz followed by 67% RZ-DQPSK with 0.8 b/s/Hz and then 33% RZ-DQPSK with 0.5 b/s/Hz while the maximum Q factor is in a reversed order.
3. RZ formats are found to give a better performance than their corresponding NRZ formats due to their reduced pulse duration, which leads to increase their robustness to CD and the nonlinear effects but with a reduced spectral efficiency because of their wider spectrum.
4. DPSK formats are found to be highly affected by decreasing the channel spacing from 100 GHz to 50 GHz. For example, the optimum format between them is 33% RZ-DPSK has a maximum Q factor of 24 dB in the single channel system with a power range from (-2.8) dBm and continues above 20 dB. At 100 GHz channel spacing, its Q factor is decreased to 19.7 dB with a power range from 8.1 dBm to 16.5 dBm while at 50 GHz, its Q factor is decreased to 15 dB and does not reached the 17 dB (BER = 1×10^{-12}). In addition, DPSK formats have a small values of the (SE) as shown in table (2-8).

Reference

[1]I. B. Djordjevic, B. Vasic, J. Rorison, "Multi-Weight Unipolar Codes for Spectral-Amplitude-Coding Optical CDMA Systems," IEEE Communications Letters, vol. 8, pp. 259-261, April 2004.

[2] J. A. Salehi, "Code division multiple access technique in optical fiber networks- Part I: Fundamentals principles,'' IEEE Trans. Commun., vol. 37, pp.824-833, Aug. 1989.

[3] E.D.J. Smith, R.J. Blaikie, D.P. Taylor, "Performance enhancement of spectral amplitude-coding optical CDMA using pulse-position modulation," IEEE Trans. Commun. 46 (1998) 1176–1184.

[4] Z. Wei, H. Ghafouri-Shiraz, Codes for Spectral-amplitude-coding optical CDMA systems, J. Lightwave Technol, vol. 50, August 2002, pp. 1209-1212.

[5] Z. Wei, H. Ghafouri-Shiraz, Codes for Spectral-amplitude-coding optical CDMA systems, J. Lightwave Technol, vol. 50, August 2002, pp. 1209-1212.

[6] S.A.Aljunid, ,M.Ismail, A.R.Ramli, Borhanuddin M Ali, and Mohamad Khazani Abdullah, "A NewFamily of Optical Code Sequences for Spectral-Amplitude-Coding Optical CDMA Systems" IEEE Photonics Technology Letters, Vol. 16, No. 10, October 2004.

[7] Hillal Adnan Fadhil, S. A. Aljunid, R. B. Ahmed, "Effect of random diagonal code link of an OCDMA scheme for high-speed access networks," J. Optical Fiber Technology.15 (2009) 237-241.

DEVELOPMENT MODEL OF BALANCING SUBCHANNELS BETWEEN SUBSCRIBER STATIONS IN WIMAX NETWORK

Dr. Haider D. Al-Janabi [1]
dr.haideraljanabi@gmail.com
+9647731186547

Dr. Hussam D. Al-Janabi [2]
en.hussam2@gmail.com
+9647722569770

Bashar B. Qas Elias [3]
basharbahaa1989@yahoo.com
+9647702501023

Dr. Aymen M.K. Al-Dulaimi [4]
dr.haideraljanabi@gmail.com
+9647735173736

[1,3] *Al Nisour University College - Department of Computer Techniques Engineering*

[2] *Al-Mustafa University College - Department of Computer Techniques Engineering*

[4] *Al-Farahidi University College - Department of Communications Technology Engineering*

Abstract

In this paper, the Developed Model of balancing allocated subchannels between subscriber stations is presented. According to the requirement of the Quality of Service (QoS) included the solutions of the unbalancing allocated of the subchannels between the subscriber stations SS, the Orthogonal Frequency Division Multiplexing Access (OFDMA) techniques is proposed. Two simulation designs based on the Full Usage Subchannel (FUSC) mode which is the one mode of the OFDMA consist of 8 and 16 subchannels as example, to investigate and offer the guaranteed for providing throughput for subscriber stations. Matlab version r2014b is used as a software simulation.

Keywords: WiMAX, Subchannel, OFDMA, Scheduling blocks, resource balancing.

الخلاصة

في هذا البحث، يتم عرض نموذج مطور لموازنة القنوات الفرعية المخصصة بين المستخدمين. ووفقاً لمتطلبات جودة الخدمة شملت حلول عدم توازن القنوات الفرعية بين محطات المشتركين، يقترح تقنيات النفاذ المتعدد بالتقسيم التعامدي للتردد. و يتكون تصميم المحاكاة على أسلوب الاستخدام الكلي للقنوات الفرعية، وهو احد الاساليب المستخدمة في توزيع القنوات الفرعية، سبيل المثال قناة مكونة من 8 و 16 قناة فرعية، للتحقيق وتقديم ضمان توفير القدرة الإنتاجية لمحطات المشتركين يستخدم برنامج الماتلاب نسخة 2014 بمثابة برنامج للمحاكاة.

1. Introduction

Presently, the latest info communication services in terms of improving their mobility and availability is directly related to further deployment of wireless telecommunication technologies. However, the limiting factor in this case is low performance provided by wireless technologies in comparison with the wired solutions.

WiMAX (Worldwide Interoperability for Microwave Access) technology, which surely takes one of the leading places in the market of wireless solutions, the main resource unit is a subchannel which is result of splitting the original frequency downlink. The number of such channels is entirely determined by the channel bandwidth (1.25 ÷ 20 MHz) and selected operating mode and may vary from 2 to 32 for the DL FUSC mode, and from 3 to 60 for DL PUSC mode. It should be noted that the level of satisfaction of the Quality of Service requirement requests of a particular user station (Subscriber Station, SS) is determined by the number of subchannels allocated and modulation and coding scheme (Modulation and Coding Scheme, MCS) used [1].

In this regard, there is an actual problem associated with improvement of frequency and time resource allocation methods in WiMAX technology. Currently there are known several approaches to optimization the process of frequency and time resource allocation in WiMAX technology [1-4]. Special attention should be paid to solutions based on balancing the number of subchannels and slots [2-4] allocated to particular subscriber station within generated bursts.

1. Model for Subchannels Balancing in Wireless Network of IEEE 802.16 Standard

In the method of balancing the number of subchannels allocated to subscriber station it is assumed that there are known the following inputs [5-6]:

- Bandwidth of used frequency channel is from the range of 1.25 MHz to 20 MHz; selected mode of usage of the subchannels (FUSC, PUSC, OPUSC, OFUSC,and TUSC);
- $\boldsymbol{N}$ is the total number of the SSs in the network;
- $\boldsymbol{K}$ is the number of subchannels used depending on the selected channel bandwidth (for ex-ample, using the mode of DownLink FUSC the number of subchannels can take values 2, 8, 16, 32; while for DownLink PUSC they are 3, 15, 30, 60);
- $R^{n}{}_{req}$ is the required transmission rate for service of the n–th SS (Mbps) ;
- $R^{n,k}$ is the bandwidth of k -th subchannel allocated to the n -th SS.

In the WiMAX technology the duration of frame can vary and take values equal to 2, 2.5, 4, 5, 8, 10, 12.5, and 20 ms. Taking into account that the useful part of the symbol has a fixed duration Tb= 89,6

μs, the number of symbols in frame will take values 19, 24, 39, 49, 79, 99, 124, 198 according to the indicated size of frame [5].

Moreover, between the symbols there is a guard interval Tg , which can take four values concerning the length of the useful part of symbol (Table 1).

Table 1

Guard Interval Values	$T_b/4$	$T_b/8$	$T_b/16$	$T_b/32$
T_g (μs)	22.4	11.2	5.6	2.8
OFDMA Symbol Duration, T_b+T_g (μs)	112	100.8	95.2	92.4

The capacity of the k -th subchannel allocated to the n-th SS ($R^{n,k}$) represents the number of transmitted bits per time unit (second) and can be calculated according to the formula [5, 6]:

$$R^{n,k}=\frac{R_c^{n,k}K_b^{n,k}K_s(1-BLER)}{T_b+T_g+T_{RTG}+T_{TRG}}, \qquad (1)$$

Where $R_c^{n,k}$ is the speed of the code used at coding of a signal of the n-th SS; $K_b^{n,k}$ is the bit load of the symbol of the *n*-th SS; K_s is the number of subcarriers for the data transmission in one subchannel (for the Downlink FUSC submode K_s=48 and for Downlink PUSC K_s=24); T_{RTG}= 105 μs is the duration of the interval of switching from receiving to transmission (receive/transmit transition gap, RTG); T_{TRG}=60 μs is the duration of the interval of switching from transmission to receiving (transmit/receive transition gap, TRG); *BLER* is the probability of the block error obtained at the expense of the Hybrid Automatic Repeat Request mechanism (HARQ) [1].

While solving a problem of subchannels allocation within the framework of the represented model it is necessary to provide calculation of the control variable (X_n^k), defining the order of subchannels allocation. According to the physics of solved problem the following limitation should be over the control variables:

$$x_n^k \in \{0,1\},\ (n=\overline{1,N},\ k=\overline{1,K}), \qquad (2)$$

$$x_n^k=\begin{cases}1, \text{ if } k-\text{th subchannel allocated to the } n-\text{th SS};\\ 0, \text{ otherwise}.\end{cases}$$

Total number of control variables depends on amount of subscriber stations in the network and used subchannels respectively defined by the expression *N*. *K*. When calculating the required variables (X_n^k), it is necessary to meet a number of important conditions limitations.

Condition of fixing one subchannel only for one subscriber station is defined according to the expression:

$$\sum_{n=1}^{N} x_n^k \le 1,\ (k=\overline{1,K}). \qquad (3)$$

Condition of scheduling the transmission rate for the *n-th* subscriber station on the *k-th* subchannel not exceeding the capacity of subchannel is defined by the expression

$$\sum_{k=1}^{K} R^{n,k} x_n^k \ge R_{req}^n \delta_n, \qquad (4)$$

$$\delta_n = \begin{cases} 1, \text{ if for } n-\text{th SS service guarantee necessary;} \\ 0, \text{ otherwise.} \end{cases}$$

For optimal balancing the number of subchannels allocated to each SS, the system introduced additional conditions limitations to the control variables X_n^k :

$$\frac{\sum_{k=1}^{K} R^{n,k} x_n^k}{R_{req}^n} \ge \beta,\ (n=\overline{1,N}) \qquad (5)$$

Where β is a control variable too, characterizing lower bound of satisfaction level of QoS requirements to access rate. In general β =0.

To improve the quality of service in WiMAX network in solving the problem of balancing the number of subchannels allocated to SS it is needed to maximize the lower bound meeting QoS requirements to access rate, i.e.

$$\beta \to \max. \qquad (6)$$

Thus, the model of balancing the number of subchannels allocated to subscriber station in WiMAX network based on solution of optimization problem associated with maximizing the lower level allocated bandwidth to each subscriber station (6) according to its QoS requirements for ac-cess rate. As the constraints stated in solving the optimization problem are conditions (1)-(5). For-mulated optimization problem belongs to class of mixed-integer linear programming, because some variables of (6) are Boolean, balancing variable (6) is a positive real variable, and objective function (6) and constraints (2)-(5) are linear.

2. Research of balancing process the subchannels number with guarantee of access rate for some subscriber stations (channel bandwidth 5 MHz, number of subscriber stations is 3)

For the second case, when guaranties of access rate provided just for some subscriber stations, considered the same input data as for the first. But guaranties provided only to first subscriber station:

$\delta_1 = 1$, $\delta_2 = 0$, $\delta_3 = 0$. Then, on Fig. 6 shown how adaptively changed the lower bound of satisfaction level of QoS requirements to access rate depending on the growth requirements of the first station to access rate ($R_{req1} = 1 \div 3$, step 0.18 Mbps). As it can be seen from Fig. 5, with $R_{req1} \leq 2$ Mbps each SS also received the required QoS level, because ($\beta \geq 1$).

With growth of R_{req1} presented method due to limited volume of available channel resource adaptively reduced level of quality of service for each station in dependence on their requirements for access rate (Fig. 1). However, the first station always got required level of service for access rate (Fig. 2), and for second (Fig. 3) and third (Fig. 4) stations as expected QoS requirements were not satisfied with $R_{req1} > 2$.

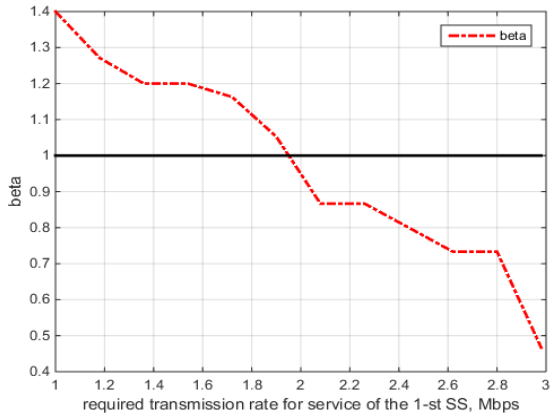


Fig. 1

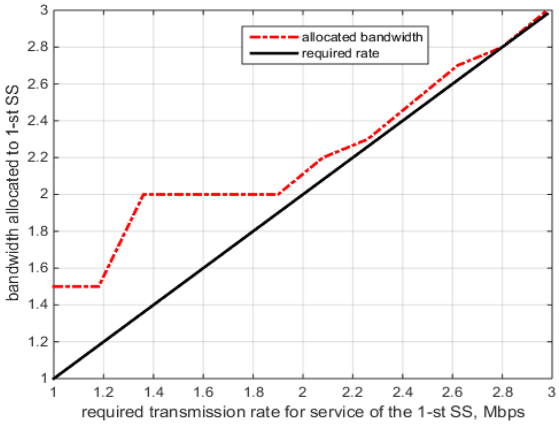


Fig. 2

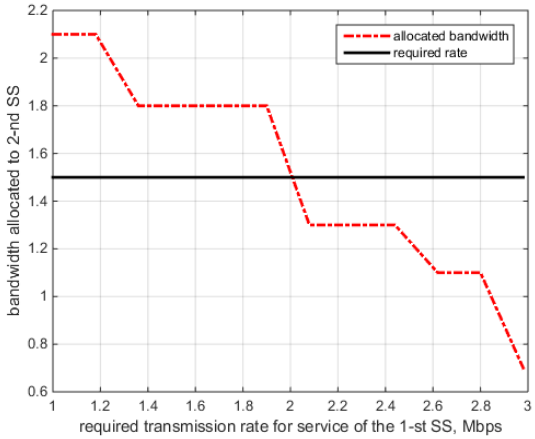


Fig. 3

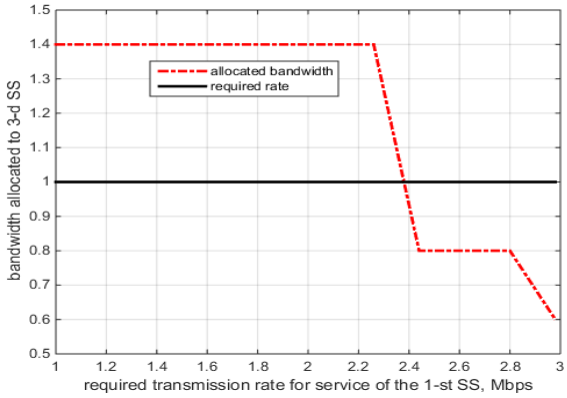


Fig. 4

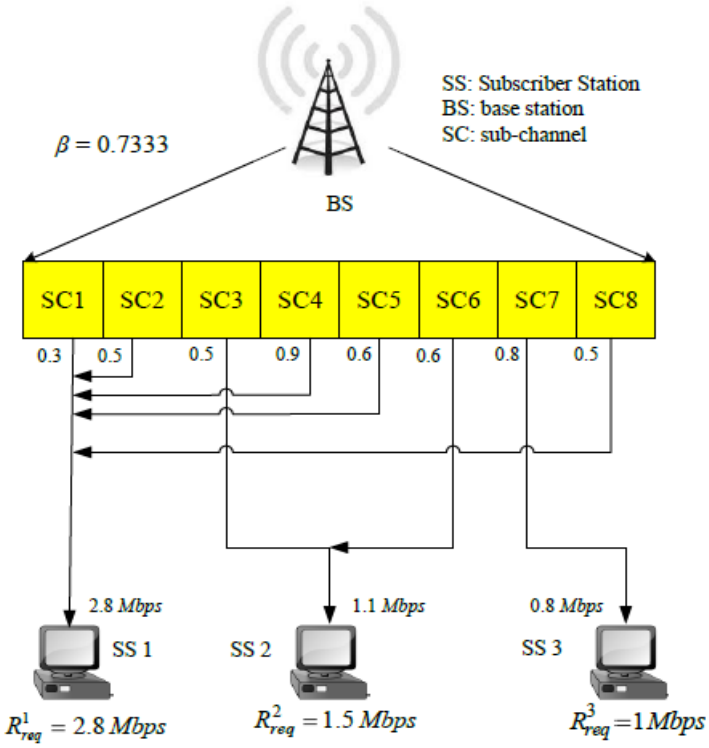


Fig. 5

Fig. 5 shows an example of solving the problem of subchannel allocation between the three SSs during the next initial data: R_{req1} = 2.8 Mbps, R_{req2} = 1.5 Mbps, R_{req3} =1 Mbps (fig. 6); service disciplines are with guaranties for 1-st SS, and best effort service for 2-nd and 3-d SSs. Within the whole solution it was provided a balanced allocation (β = 0.7333) of eight subchannels between the three SSs, while the first station provides guaranties with respect to the increasing demands of Bandwidth (fig. 2).

The results of numerous calculations showed that the cause of these problems is the fact that with a relatively small number of subchannels, which are provided the discrete allocated bandwidth for each SS, ensuring the ideal fair balancing of resources is difficult. This is especially true for the case where the number of subscriber stations or QoS-demands grow, and channel bandwidth and number of formed subchannels remain unchanged or even reduced.

3 Research of Balancing Process the Subchannels Number without Guarantee of Access Rate (channel bandwidth 10 MHz, number of subscriber stations is 3)

Within research considered the following input data:

- number of subscriber stations is 3 (N=3);
- number of available subchannels is 16 (K =16);

- access rates required for each station: R_{req1}= 1÷7 Mbps (step 0.45 Mbps), R_{req2} = 2.5 Mbps, R_{req3}= 3 Mbps; matrix of bandwidths:

$$R\left|R_{n,k}\right| = \begin{Vmatrix} 0.3 & 0.5 & 0.2 & 0.9 & 0.6 & 0.3 & 0.4 & 0.5 & 0.3 & 0.5 & 0.2 & 0.9 & 0.6 & 0.3 & 0.4 & 0.5 \\ 0.5 & 0.7 & 0.5 & 0.8 & 0.1 & 0.6 & 0.5 & 0.3 & 0.5 & 0.7 & 0.5 & 0.8 & 0.1 & 0.6 & 0.5 & 0.3 \\ 0.6 & 0.3 & 0.6 & 0.5 & 0.3 & 0.2 & 0.8 & 0.1 & 0.6 & 0.3 & 0.6 & 0.5 & 0.3 & 0.2 & 0.8 & 0.1 \end{Vmatrix}.$$

As an example, consider the case when the process of balancing the number of subchannels to subscriber stations makes guaranties for access rate: δ_2= 1, n = 1,3. In analogy with Fig. 1- 4 on Fig. 11 ÷ 14 it was shown the effect on the process of balancing channel resource use with increasing QoS demands on the bandwidth, for example, of the first subscriber station. As was shown on Fig. 6-9, with $R_{req1} \geq 4.3$ all stations have received a satisfactory quality of service (β =1), while with further increase of QoS demands of the first station available network resources are allocated in accordance with the requirements of the stations to their allocated downlink bandwidth. Increasing the amount of available for allocation frequency subchannels (8 to 13), while maintaining the same number of subscriber stations (N= 3), it is possible to achieve almost ideal balancing of the frequency resource (Fig. 6) which in the limit should be a linear dependence [5,6].

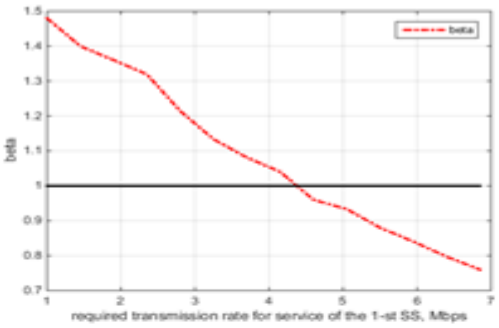


Fig. 6

Fig. 7

Fig. 8

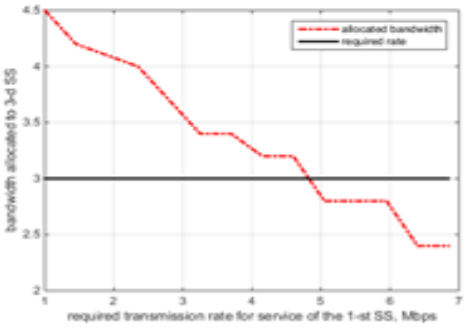


Fig. 9

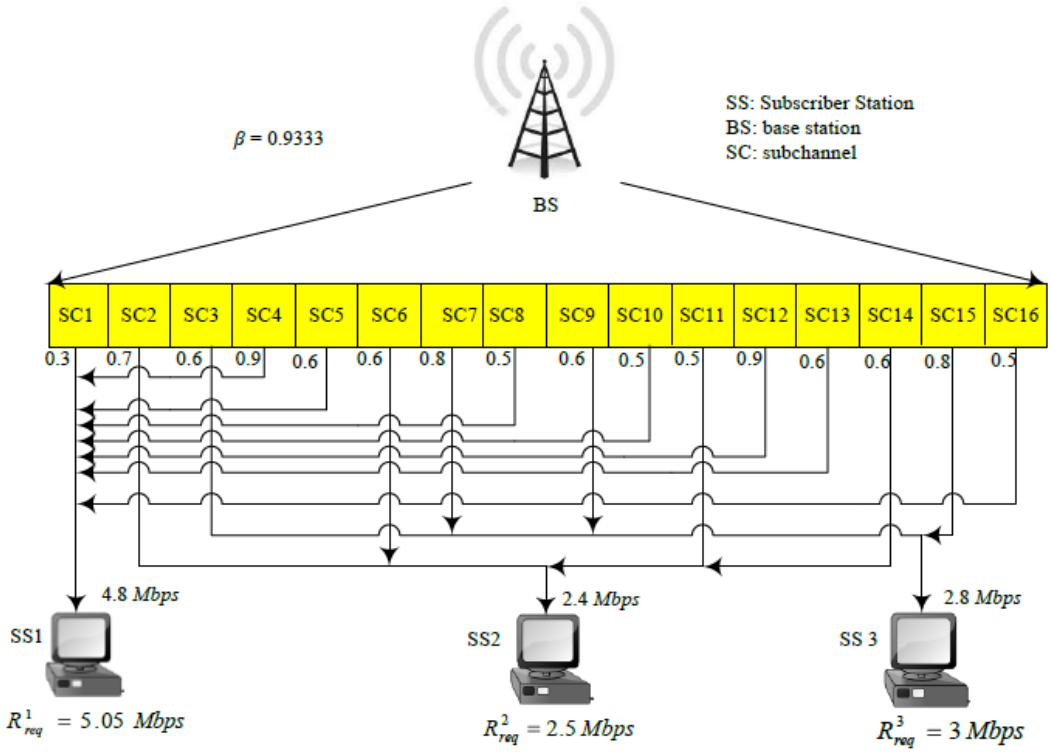


Fig. 10

On the Fig. 10 it was shown an example of solving the problem of subchannels allocation between the three SSs according to following initial data: R_{req1}= 5.05 Mbps, R_{req2} = 2.5 Mbps, $Rreq_3$= 3 Mbps (fig. 6); service discipline for this example is best effort service for all SSs. Within the obtained solution provided balanced allocation (β = 0.9333) of sixteen subchannels between three SSs according to their QoS requirements on bandwidth.

Conclusion

The tasks of ensuring the quality of service in WiMAX networks are very important with the basic QoS indicator access rate. Depending on the type of provided service the required access rate can differentiate and, in some cases, even be guaranteed.

Research of proposed model confirmed the adequacy of the model (1)-(5) and effectiveness of solutions as a whole in terms of providing different types of service level (with and without guaranties) to subscriber stations. Results showed that in balancing the number of subchannels

allocated to SSs with a deficit of network resource adaptive reduction of provided level of quality of service performed dependently on the type of possible guaranties and requirements regarding the access rate. Furthermore, increasing the amount available for allocation frequency subchannels (from 8 to 16), while maintaining the same number of subscriber stations ($N = 3$), it is possible to achieve almost ideal balancing of the frequency resource which in the limit should be a linear dependence.

References:

[1] Puranik S.R., Vijayalakshmi M., Kulkarni L. A Survey and Analysis on Scheduling Algorithms in IEEE 802.16e (WiMAX) Standard // International Journal of Computer Applications. 2013. P. 1-10.

[2] Garkusha S., Andrushko Yu., Lemeshko O. Analysis Results of WIMAX Dowlink Traffic Management Model In Congestion Conditions // Proceedings of World Telecommunications Congress 2014 (WTC 2014). 2014. P. 1-4.

[3] S. Garkusha, H. Al-Janabi, A. Al-Dulaimi Results of Development of Sub-Channels Scalable Scheduling Model in WiMAX Network / // Proceedings of First International IEEE Conference «Problems of Infocommunications. Science and Technology» (PICS&T-2014). – Kharkiv, Ukraine: 14-17 October, 2014. – P. 49-53.

[4] Lemeshko A.V., Al-Janabi H.D., Al-Dulaimi A.M.K. Model progress of subchannel distribution in WIMAX in antennas system // X Anniversary International Conference On Antenna Theory And Techniques, ICATT'2015. 2015. P. 276-278.

[5] G Sergiy, HD Al-Janabi, AMK Al-Dulaimi Model of the balanced distribution of subchannels in the mesh-network using WiMAX technology// Electronics and Nanotechnology (ELNANO), 2014 IEEE 34th International Conference, 2014.

[6] Garkusha S. Model Of The Balanced Distribution Of Subchannels in The Mesh-Network Using WiMAX Technology / S. Garkusha, H.D. Al-janabi, A.M.K. Al-Dulaimi // 2014 IEEE 34th International Conference on Electronics and Nanotechnology (ELNANO) Kyiv, Ukraine, 15-18, April 2014. – Kyiv Polytechnic Institute. – P. 421-425.

The use of watermarking technique in controlling the security and privacy of the iris images

استخدام تقنية العلامة المائية في السيطرة الأمنية والحفاظ على الخصوصية لصور قزحية العين

Dr. Rasha T. Mohammed, & Dr. Abdulsalam M. Saeed
Alrasheed University College, Department of Computer Techniques Engineering, Baghdad, Iraq, e-mail: rashathabit@yahoo.com

Abstract

The biometric identification systems are increasingly used because of their advantages in secure access control and prevention of fraud. Exchanging and storing biometric information (such as iris images) needs protection against attacks (i.e, theft, deleterious modification ...). One of the methods that are widely used to protect digital images is image watermarking. This paper presents an iris image watermarking technique that can be used to ensure the safety of the images in addition to data hiding. The proposed watermarking algorithm hides the watermark and other important information into the iris image without affecting the quality of the iris region and thus the watermarked image can be used for the recognition and identification purposes.

Keywords: Biometric watermarking, reversible watermarking, discrete wavelet transform, histogram shifting.

الخلاصة

هناك زيادة مضطردة لاستخدام منظومات التمييز البايومترية، وذلك لميزتها في السيطرة على الايصال الآمن ومنع التلاعب. ان تبادل وخزن المعلومات البايومترية (مثل قزحية العين) بحاجة للحماية ضد الهجمات الخارجية (كالسرقة او التغيرات المؤذية ...). إحدى الطرق التي تستخدم بشكل واسع لحماية الصور الرقمية هي صورة العلامة المائية. تعرض هذه الورقة تقنية العلامة المائية لصورة قزحية العين، التي يمكن استخدامها لضمان سلامة هذه الصور بالإضافة إلى إخفاء البيانات. إن خوارزمية العلامة المائية المقترحة تقوم بإخفاء العلامة المائية وغيرها من المعلومات الهامة في صورة القزحية دون التأثير على جودة منطقة القزحية، وهكذا صورة العلامة المائية يمكن استخدامها لأغراض التمييز وتحديد الهوية.

1- Introduction

Biometric identification systems are increasingly used nowadays which depend on collecting the personal traits data and saving that data for authentication purposes. There are different personal traits such as the features of the face, fingerprint, voice, and iris images [1]-[4]. The biometric systems provide considerable avail in comparison with the traditional

authentication techniques, however, protecting the biometric data form unauthorized modifications is a critical issue. The receiver of the biometric data should guarantee that the data have been received from an authorized person [5].

Iris-based biometric systems [6] are highly recommended because of their security and ease of use [7]. storing iris images or exchanging them through unsecured channel such as internet will increase the threat of modifying the images or stealing the important information related to these images by unauthorized persons. Therefore, iris images and their related information should be saved and exchanged in a secure manner. One of the techniques that have been applied to protect this kind of images is image watermarking.

Image watermarking is a technique used to protect images either by hiding watermarks (such as text, logo, voice, …) in the original image or by hiding the original image in a cover media. A number of watermarking schemes are available to protect iris images such as the schemes in [5], [7]-[10]. In [5], the iris image has been protected by transforming the image using discrete cosine transform (DCT) and embedding a demographic text in the DCT coefficients. The watermark embedding method depends on interchanging three middle band DCT coefficients pairs. In [7], the iris image template has been extracted using DCT and converted to a binary code. Another image has been used as host image which is transformed using discrete wavelet transform (DWT) and the singular values of the DWT coefficients are modified to carry the binary code. In [8], two iris image watermarking scenarios have been investigated to study the effect of watermarking on iris image recognition. In the first scenario the iris template is extracted and embedded in a cover image (i.e., iris watermarks); in the second scenario the iris image watermarking is applied by embedding owner's information in the image. The results of this work proved that the second scenario (i.e., iris image watermarking) performs better than the first scenario in terms of recognition performance, however, the watermarking process produce distortion in the watermarked iris image. In [9], the iris image template has been embedded in a cover face image using watermarking technique based on parallel-bit stream method. The bits embedding process is performed by applying the XOR operation between the bit and the pixel value of the image. In [10], the iris image template has been extracted and the Bose–Chaudhuri–Hocquenghem codes (BCH) have been calculated. The BCH codes are embedded in the singular values of DCT coefficients for a cover image.

As explained previously, the watermarking techniques in [5], [7]-[10] are performed either by calculating the iris templates and using these templates as watermarks or by watermarking the iris image using owner's watermarks. These techniques have been successfully applied to protect iris images, however, they suffer from some limitations. In the case of using the iris image template as a watermark the recognition process will be affected especially when the image undergoes some attacks. In the case of watermarking the iris image the performance is better in terms of recognition but this process will introduce some distortions in the iris image. Therefore, an iris image watermarking technique that can recover the original iris image after extracting the embedded watermark is required. This process can be obtained using reversible image watermarking technique.

This paper presents a reversible iris image watermarking technique that can embed a watermark and other important information in the image using histogram shifting process of the discrete wavelet transform coefficients. At the watermark-extraction side, the original iris image can be recovered without any loss and thus the watermarking process will not produce distortions in the iris image and therefore will not affect the image recognition process. The rest of the paper is organized as follows: section 2, presents the proposed technique for iris image watermarking; section 3, explains the experiments that have been conducted and their results; and section 4, contains the conclusions.

2- The proposed technique for iris image watermarking

This section contains a demonstration of the proposed watermarking technique for iris images. The first subsection contains a general description of the proposed technique, the second subsection contains the details of the watermark embedding and extraction algorithms, and the third subsection contains the flow charts for the abovementioned algorithms.

2.1 General description of the proposed technique

As explained in section 1, the previous iris image watermarking techniques suffer from the problem of distortions that are generated in the iris image after the watermark embedding process which may affect the iris image recognition process and thus may lead to incorrect recognition. To solve this problem, we suggested the use of reversible (i.e., lossless) image watermarking technique in order to ensure the integrity of iris image by recovering the original iris image after the watermark extraction process. To ensure the security while saving and exchanging iris images and their related information (such as the personal information about

the owner of the iris image), the information is embedded in the iris image and thus it will not be a clear text to a reader (only authorized person can extract and read that information).

2.2 The proposed algorithms

The proposed technique for iris image watermarking includes two algorithms that are the watermark embedding and the watermark extraction algorithms that are explained in the following subsections.

A) Watermark embedding algorithm

In the watermark embedding process, the secret information (i.e., text watermark with some other important data) are inserted in iris images to provide security when saving these images and to ensure the integrity when exchanging them. This process can be summarized in the following steps.

Step 1: Read the original iris image and the text file of the secret information. The text file contains a text watermark and some other important data such as the personal information about the owner of the iris image.

Step 2: Convert the text file to a binary sequence (*Bseq*) and read the length (i.e., the total number of bits) of the resulting binary sequence (*Lseq*).

Step 3: Read the size of the iris image (*M*×*N pixels*) and choose the side length of the image block (*Bs*).

Step 4: Divide the iris image into non-overlapping blocks of size (*Bs*×*Bs pixels*) and apply wavelet transform (*wname*) to obtain the wavelet coefficients. The resultant coefficients are divided into four subbands (i.e., *CA*, *CH*, *CV*, and *CD*) as shown in Figure (1).

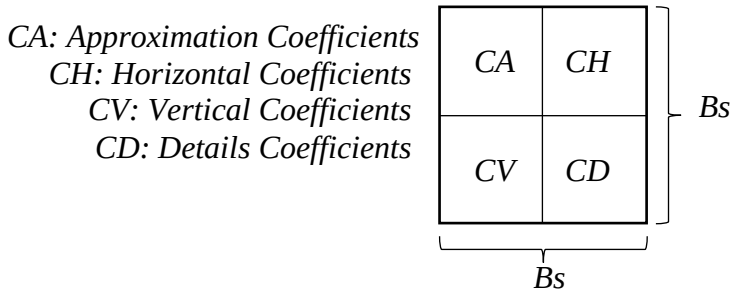


Figure (1): Wavelet coefficients for an image block of size (*Bs*×*Bs*).

Step 5: choose the horizontal and vertical subbands (i.e., *CH* and *CV*) to carry the binary bits (*Bseq*). The binary bit embedding algorithm can be explained as follows:

Input: binary bit (*b*), horizontal subband (*CH*), and vertical subband (*CV*).

Output: Watermarked *CH* (*WCH*), and watermarked *CV* (*WCV*).

1. Calculate the average of the coefficients in *CH* and *CV* (*AvCH*, and *AvCV*).
2. Calculate the differences values (*d1= AvCH - AvCV*) and (*d2= AvCV - AvCH*).
3. Set a threshold value (*Thr*) and shift the histogram of the subbands using the following rules:

   If (*d1< Thr*) and *b=1* then increase *CH* and decrease *CV* by the value *((Thr-d1)/2).*

   If (*d2 < Thr*) and *b=0* then decrease *CH* and increase *CV* by the value *((Thr-d2)/2).*

   Else keep the block unchanged.
4. Save *d1* and *d2* as side information which are necessary for the extraction process. The resultant *CH* and *CV* subbands after modification process are the watermarked subbands (*WCH* and *WCV*).

Step 6: Replace the original *CH* and *CV* subbands with (*WCH* and *WCV*) respectively, then apply inverse wavelet transform to obtain the watermarked block.

Step 7: Repeat the binary bit embedding process until finish the binary sequence (*Bseq*).

Step 8: Replace the original blocks with the watermarked blocks to obtain the watermarked iris image.

B) Watermark extraction algorithm

In the watermark extraction process, the secret information is extracted from the iris image and the original iris image is recovered to ensure its safety from any distortions and thus the watermarking process will not affect the iris image recognition. This process can be summarized in the following steps.

Step 1: Read the watermarked iris image and its side information that have been generated during the watermark embedding process.

Step 2: Read the size of the watermarked iris image ($M \times N$) and divide the image into non-overlapping blocks of size ($Bs \times Bs$ *pixels*).

Step 3: Apply wavelet transform (*wname)* to obtain the wavelet coefficients (i.e., *CA*, *WCH*, *WCV*, and *CD*).

Step 4: Choose the horizontal and vertical subbands (i.e., *WCH* and *WCV*) to extract the binary bits (*Bseq)*. The binary bit extraction and block recovery algorithm can be explained as follows:

Input: *WCH* and *WCV*.

Output: binary bit (*b*), recovered block.

1. Calculate the average of the coefficients in *WCH* and *WCV* (*AvWCH*, and *AvWCV*).
2. Extract the binary bit and recover the original *CH* and *CV* subbands by applying the following rules:

   If (*AvWCH* > *AvWCV*) then set the extracted bit (i.e., *b=1*); decrease *WCH* and increase *WCV* by the value *((Thr-d1)/2)*.

   If (*AvWCH* < *AvWCV*) then reset the extracted bit (i.e., *b=0*); increase *WCH* and decrease *WCV* by the value *((Thr-d2)/2)*.

Step 5: Replace *WCH* and *WCV* with the recovered *CH* and *CV* subbands, then apply inverse wavelet transform to obtain the recovered block.

Step 6: Repeat the binary bit extraction and block recovery procedure until finish extracting the binary sequence (*Bseq*).

Step 7: Replace the watermarked blocks with the recovered blocks to obtain the original iris image.

2.3 Flow charts for the proposed algorithms

The steps of the watermark embedding and extraction algorithms can be summarized in the flow charts shown in Figures (2) and (3). Figure (2) shows the flow chart for the watermark embedding process and Figure (3) shows the flow chart for the watermark extraction process.

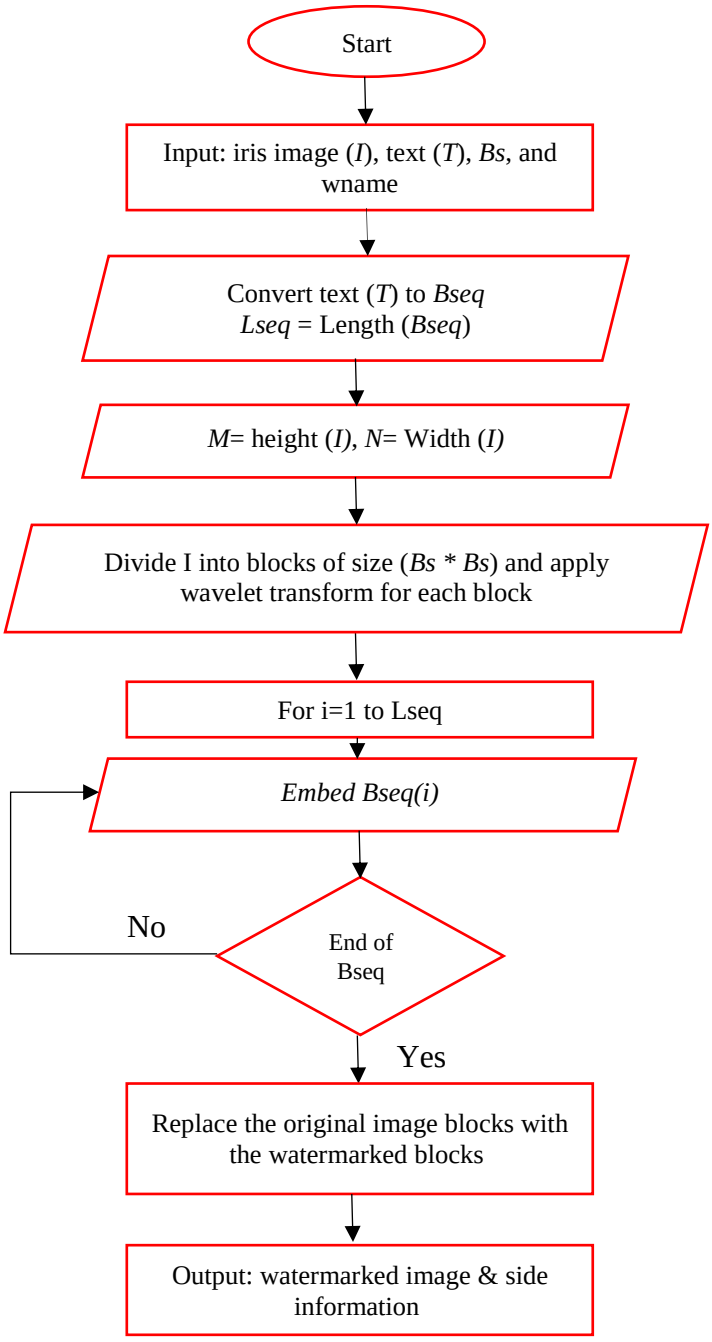


Figure (2): The proposed watermark embedding algorithm.

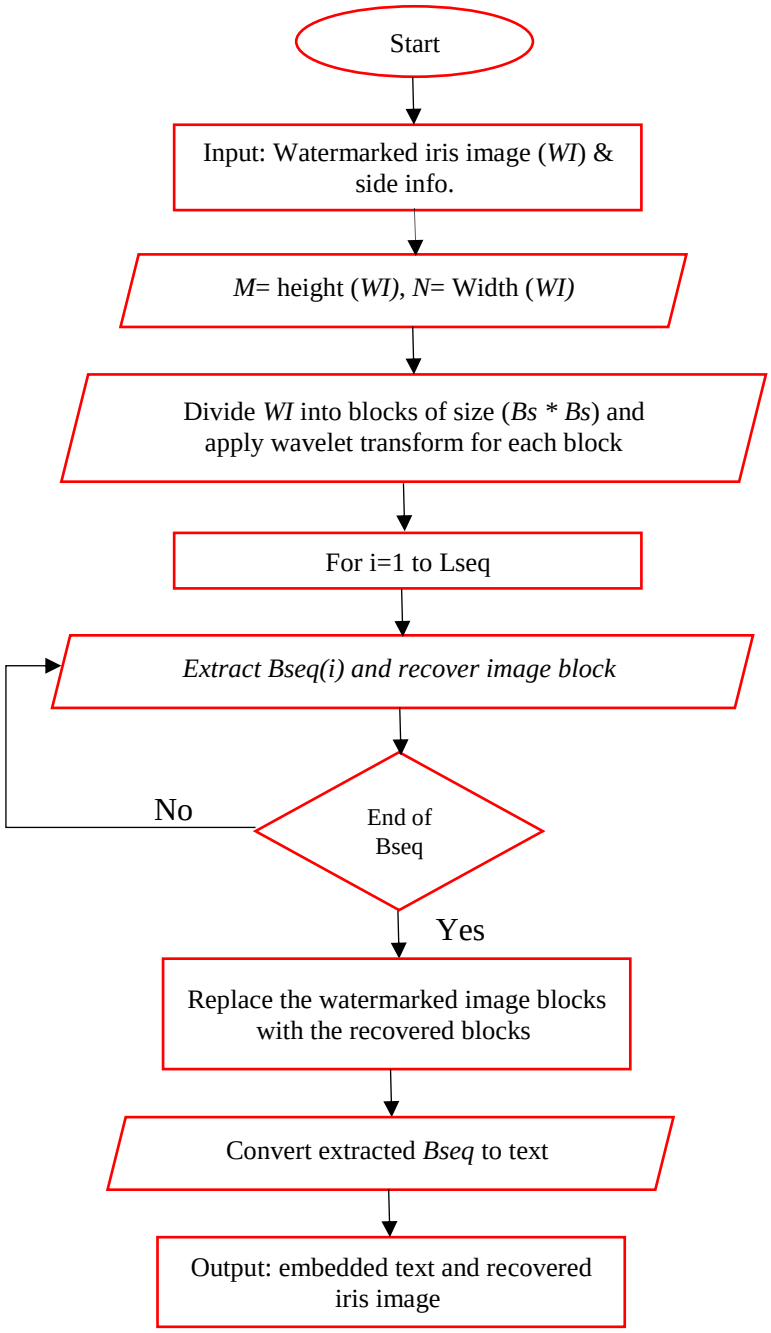


Figure (3): The proposed watermark extraction algorithm.

3- Experiments and Results

This section introduces the experiments that have been conducted to evaluate the performance of the proposed iris image watermarking technique. Two factors in the proposed technique are studied to obtain the best performance. The first factor is the wavelet family type (*wname*) and the second factor is the size of image block (*Bs*×*Bs*). In these experiments, we

used ten different iris images from IIT Delhi iris database [11] as shown in Figure (4). A sample text file (shown in Figure (5)) has been used in the experiments which contains a text watermark (Al-Rasheed University College) and sample personal information about the owner of the iris image.

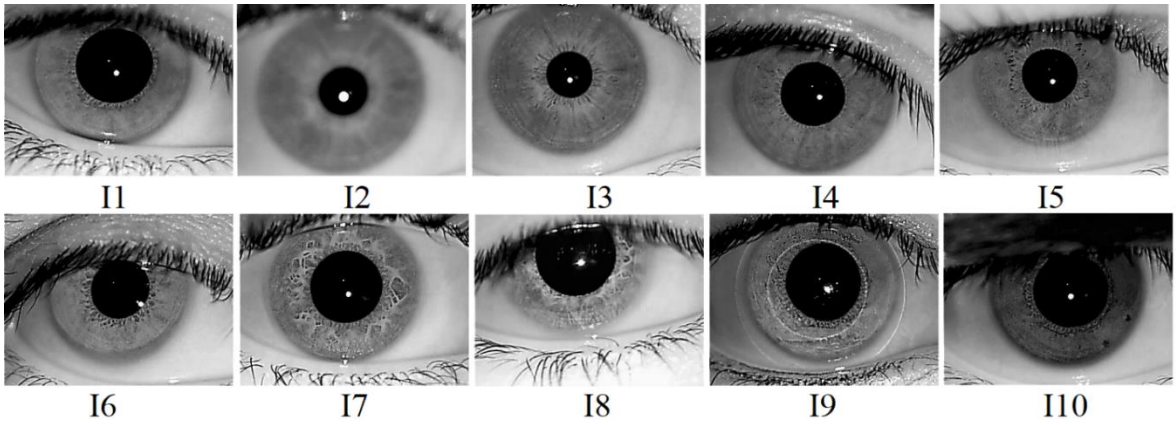


Figure (4): Test iris images [11].

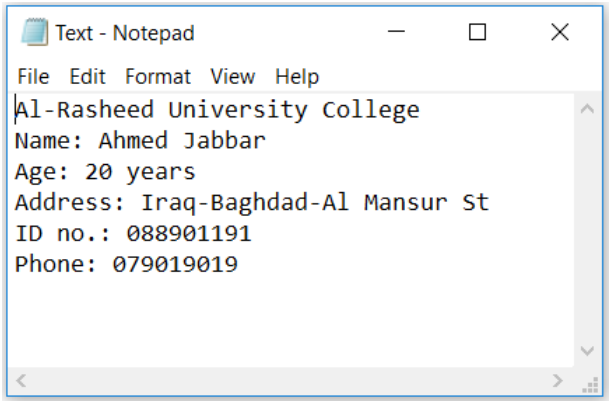


Figure (5): Sample text file.

3.1 Wavelet family type (wname) experiments

In these experiments, 16 different wavelet families have been tested to check their effect on the following:

1- The extracted binary sequence (Bseq) by calculating the bit error rate (BER). When the watermark extracted correctly the BER=0.

2- The reversibility of iris images by calculating the images similarity (IR). When the iris image recovered correctly the IR=1.

3- The visual quality of the watermarked images by calculating the peak signal to noise ratio (PSNR). The higher the PSNR the better the visual quality of the watermarked iris image.

The experiments have been conducted at Bs=8 for the 16 different types of (wname); In Table (1) and Table (2), samples of the experimental results are presented. Figure (6) shows the average BER for ten test images at each (wname), Figure (7) shows the average IR for ten test images at each (wname), and Figure (8) shows the average PSNR for ten test images at each (wname).

Table (1): Experimental Results for wname= 'haar'

Iris Image	PSNR(dB)	IR	BER
I1.gif	37.84798	1	0
I2.gif	38.55284	1	0
I3.gif	38.28459	1	0
I4.gif	37.81775	1	0
I5.gif	37.96416	1	0
I6.gif	38.32527	1	0
I7.gif	37.91788	1	0
I8.gif	36.87863	1	0
I9.gif	37.99909	1	0
I10.gif	38.95256	1	0
Average	38.05407	1	0

Table (2): Experimental Results for wname= 'coif1'

Iris Image	PSNR(dB)	IR	BER
I1.gif	39.6321886	1	0.031021898
I2.gif	39.57092191	1	0.002737226
I3.gif	39.50981607	0	0.009124088
I4.gif	39.50573964	1	0.026459854
I5.gif	39.47032362	0	0.020985401
I6.gif	39.4421254	1	0.028284672
I7.gif	39.29384924	1	0.048357664
I8.gif	39.15950324	1	0.04379562
I9.gif	39.65246724	0	0.030109489
I10.gif	39.91069863	1	0.006386861
Average	39.51476336	0.7	0.024726277

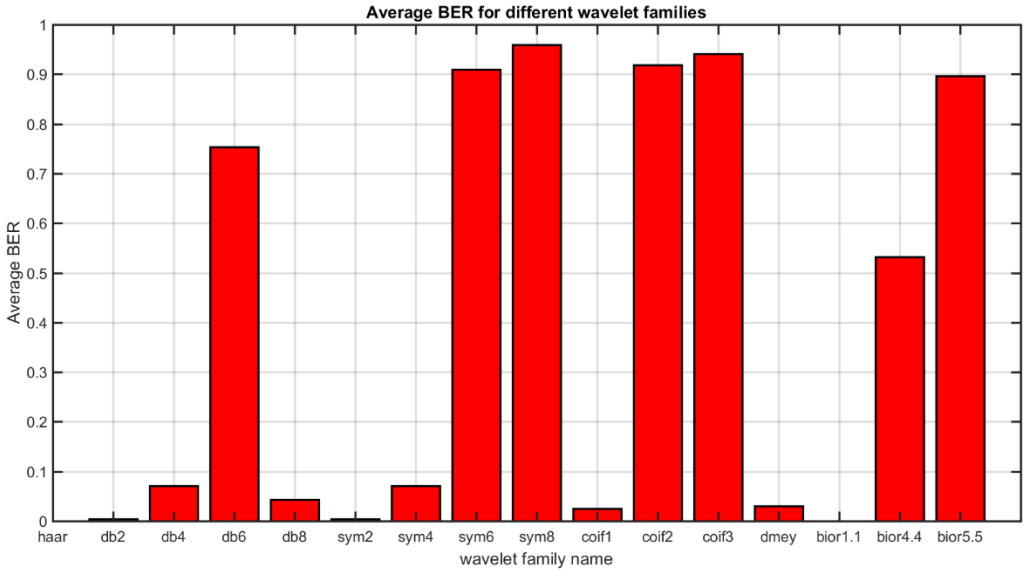


Figure (6): Average BER for test images at different wavelet families

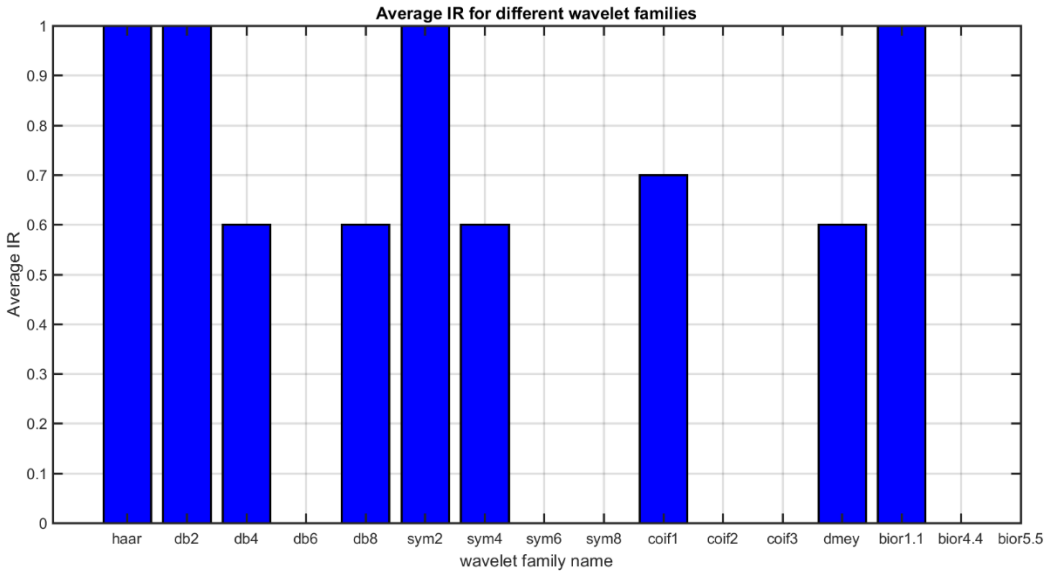


Figure (7): Average IR for test images at different wavelet families

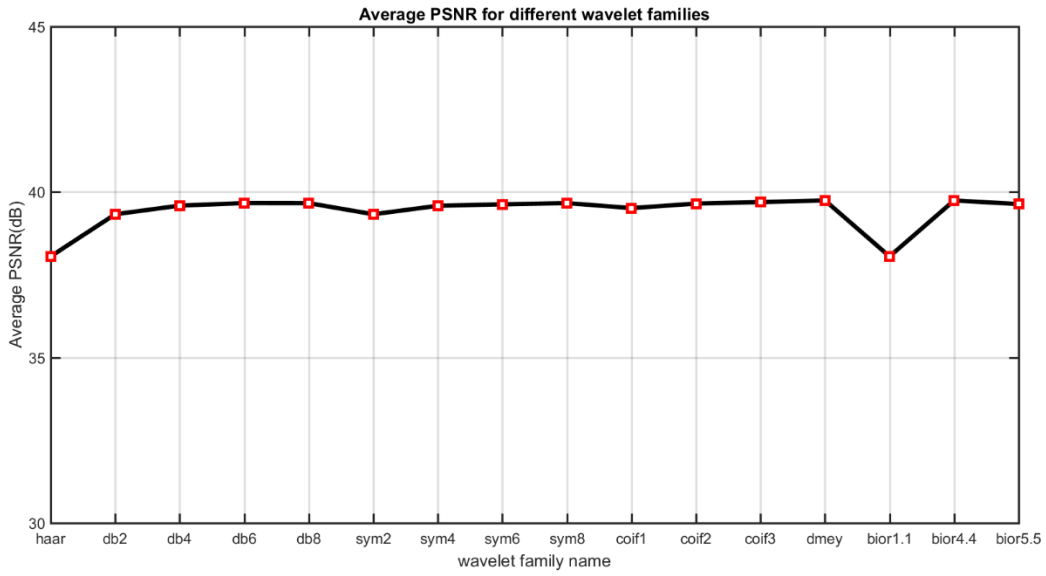


Figure (8): Average PSNR for test images at different wavelet families

The results proved that the wavelet families 'haar' and 'bior1.1' obtained the best results in terms of BER and IR at the cost of lower visual quality in comparison with other wavelet families.

In the other wavelet families, the watermark cannot be extracted correctly and the image in some cases is not completely reversible. Therefore, the wname 'haar' and 'bior1.1' are the best choices for the proposed technique.

3.2 Block size effect experiments

The experiments in this section have been conducted to evaluate the performance of the proposed technique at different block size ($Bs \times Bs$). The wname = 'haar' and different block sizes have been applied (i.e., 2×2, 4×4, and 8×8); the results are shown in Table (3), Table (4), and Table (5), respectively. As illustrated by the experimental results, the best block size for the proposed technique is 8×8.

Table (3): Experimental Results for block size (2×2)

Iris Image	PSNR(dB)	IR	BER
I1.gif	47.4789175257647	0	0
I2.gif	48.7654829278841	0	0
I3.gif	42.4183265375851	0	0
I4.gif	48.674169121663	1	0
I5.gif	49.9280875578606	1	0
I6.gif	41.8425123892404	1	0
I7.gif	42.2970592262592	0	0

I8.gif	44.6881604440366	1	0
I9.gif	40.5870813197162	0	0
I10.gif	50.5780116360988	0	0

Table (4): Experimental Results for block size (4×4)

Iris Image	PSNR(dB)	IR	BER
I1.gif	43.1847469856641	0	0
I2.gif	44.3276310317094	0	0
I3.gif	41.4362303928099	1	0
I4.gif	42.0259764189976	0	0
I5.gif	41.7522855754857	1	0
I6.gif	41.3632008834701	1	0
I7.gif	39.6010505391117	1	0
I8.gif	39.8753930119007	0	0
I9.gif	39.5720565355737	0	0
I10.gif	45.1186379437444	1	0

Table (5): Experimental Results for block size (8×8)

Iris Image	PSNR(dB)	IR	BER
I1.gif	37.8479777220166	1	0
I2.gif	38.5528358016813	1	0
I3.gif	38.284587121697	1	0
I4.gif	37.8177501171335	1	0
I5.gif	37.9641583254588	1	0
I6.gif	38.3252724524041	1	0
I7.gif	37.9178767065788	1	0
I8.gif	36.8786264574888	1	0
I9.gif	37.9990891339057	1	0
I10.gif	38.9525564252083	1	0

4- Conclusions

In this paper, a reversible iris image watermarking technique has been presented to ensure the safety and integrity of iris images when saving and exchanging them. The proposed technique depends on dividing the image into non-overlapping blocks and transforming the blocks using discrete wavelet transform. The watermark embedding process depends on shifting the histogram of the horizontal and vertical subbands of the wavelet coefficients. The experimental results proved that the wavelet family types ‘haar’ and ‘bior1.1’ obtained the best

results and the best block size is (8×8). Since the iris image can be recovered correctly after the watermark extraction process, its recognition process will not be affected.

References

[1] D. N. Parmar, and B. B. Mehta, "***Face Recognition Methods & Applications***," Int. J. Computer Technology & Applications (IJCTA), Vol 4 (1), pp. 84-86, 2013.
[2] M. R. Islam, M. S. Sayeed, and A. Samraj, "***Biometric Template Protection Using Watermarking with Hidden Password Encryption***," *2008 International Symposium on Information Technology*, Kuala Lumpur, pp. 1-8, 2008.
[3] H. N. M. Shah, M. Z. Ab Rashid, M. F. Abdollah, M. N. Kamarudin, C. K. Lin and Z. Kamis, "***Biometric Voice Recognition in Security System***," Indian Journal of Science and Technology, Vol 7(2), pp. 104–112, February 2014.
[4] K. W. Bowyer, K. P. Hollingsworth, and P. J. Flynn, "***A Survey of Iris Biometrics Research: 2008–2010,"*** M.J. Burge and K.W. Bowyer (eds.), Handbook of Iris Recognition, Advances in Computer Vision and Pattern Recognition, 2013.
[5] M. A. M. Abdullah, S. S. Dlay, W. L. Woo, "***Securing Iris Images with a Robust Watermarking Algorithm based on Discrete Cosine Transform***," In: 10th International Conference on Computer Vision Theory and Applications (VISAPP 2015).
[6] K. W. Bowyer, K. Hollingsworth, P. J. Flynn, "***Image understanding for iris biometrics: A survey***," Computer Vision and Image Understanding, Vol (110), pp. 281–307, 2008.
[7] C. J. Anoopa, "***IRIS BIOMETRIC WATERMARKING USING SINGULAR VALUE DECOMPOSITION AND WAVELET BASED TRANSFORM***," Proceedings of 1st IRF International Conference, Coimbatore, 9th March-2014.
[8] J. Dong and T. Tan, "***Effects of watermarking on iris recognition performance***," 2008 10th International Conference on Control, Automation, Robotics and Vision, Hanoi, pp. 1156-1161, 2008.
[9] E. Mostafa, M. Mansour, and Heba Saad, "***Parallel-Bit Stream for Securing Iris Recognition***," IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 3, No 2, May 2012.
[10] J. Lu, T. Qu, and H. R. Karimi, "***Novel Iris Biometric Watermarking Based on Singular Value Decomposition and Discrete Cosine Transform***," Hindawi Publishing Corporation Mathematical Problems in Engineering Volume 2014, Article ID 926170, 6 pages.
[11] IIT Delhi Iris Database version 1.0, (Accessed on April 2018).
http://web.iitd.ac.in/~biometrics/Database_Iris.htm

Implementation of OSPF protocol in network designed in Iraq

Sura kh. Ibrahim & Zainab Taha Jebur
Department of computer tech. engineering
Al-Nisour College University
Sura.eng@nuc.edu.iq
Zainab.t.eng@nuc.edu.iq

Abstract:

OSPF is an important internal gateway protocol which used in internet layer of TCP/IP model and in network layer in OSI model. OSPF is used in application of current networks like ad hoc wireless and mobile communication systems, in ISP, enterprise networks in schools, hospitals…etc. it also used in control plane of SDN. OSPF will control on decision of which have the shortest path of network branches depending on their weights. Principle work of OSPF protocol determine the shortest path by checking link state of network. Also, principle of this protocol, it divides the same network into areas (sub networks).
In this paper, OSPF network has one area called (area 0), which has 6- Routers distributed among cities on Iraq's map. MATLAB (2017b) and GNS-3 (2.1.3) softwares are used to calculate theoretical and experimental results respectively. The results proved that the best path which has the low weights and running time. These results also proved that if the shortest path is down, it can be used the other link (back up) in the network.

Keywordss: ospf, GNS-3, Matlab, SDN.

الخلاصة

ospf هو بروتوكول بوابة الداخلية الهامة التي تستخدم في طبقة الإنترنت من نموذج tcp/ip وطبقة الشبكة في نموذج osi. يستخدم ospf في تطبيق الشبكات الحالية مثل أنظمة الاتصالات اللاسلكية واللاسلكية المتنقلة، و في شبكات المؤسسات في المدارس والمستشفيات ... الخ. كما أنها تستخدم في الشركات المجهزة للنت وفي وحدة السيطرة في شبكات ال SDN. حيث يسيطر هذا البرتوكول على قرار اختيار المسارات التي لديها أقصر مسار في فروع الشبكة اعتمادا على أوزانها. مبدأ عمل بروتوكول ospf تحديد أقصر الطرق عن طريق التحقق من حالة وصلة الشبكة. أيضا، مبدأ هذا البروتوكول فإنه يقسم الشبكة نفسها الى مناطق (الشبكات الفرعية).
في هذا البحث شبكة ospf لديها منطقة واحدة تسمى (المنطقة صفر) ، والتي لديها 6- من الموجهات موزعة بين المدن على خريطة العراق. ((MATLAB (2017b) and GNS-3 (2.1.3) برامج تستخدم لحساب النتائج النظرية والعملية على التوالي. أثبتت النتائج أن أفضل مسار له الأوزان المنخفضة ووقت التشغيل. وأثبتت هذه النتائج أيضا أنه إذا كان أقصر مسار متوقف ، فإنه يمكن استخدام الارتباط الآخر (النسخ الاحتياطي) في الشبكة.

1. Introduction

With the expansion of the existing wireless nets and the emergence of new applications that require a real time communication with low cost, higher efficiency and less complexity; routing protocols become one of the most important decisions in the design of these networks [1].

OSPF is a dynamic routing protocol. It is a link-state routing protocol and is a part of the interior gateway protocols group. OSPF keeps track of the complete network topology and all the routers and connections within that network [2]. Each router will calculate a mathematical data structure called "shortest path tree"that describes the shortest path to the destination address. This is where OSPF gets its name, it will try to open the shortest path first [3].

OSPF routing protocol is very important protocol to consider when setting up routing instructions on the network. As OSPF gives the routers the ability to learn the most optimal (shortest) paths, it can definitely speed up data transmission from source to destination [4].

Dijkestra algorithm is described as greedy algorithm. It is a graph search algorithm that solves the single-source shortest path problem [5]. It can also be used for finding weights of the shortest path, from a single vertex to a single destination vertex. For example, if the vertices of the graph represent driving distances between pairs of cities connected by a direct road, dijkestra algorithm can be used to find the shortest path first is widely used in network routing protocols. Most notably OSPF (open shortest path first) [6].

In this paper**,** OSPF network has one area called (area 0), which has 6- Routers distributed among cities on Iraq's map. MATLAB (2017b) and GNS-3 (2.1.3) softwares are used to calculate theoretical and experimental results respectively. The results proved that the best and shortest path which has the low weights and low time complexity. These results also proved that if the shortest path is down, it can be used the other link as a backup in the network in spite of it has high weights.

Design of OSPF area network:

OSPF network is designed in some of cities in Iraq. It consists of 6-vertices (routers) and 6-edges which is connected between the vertices collected in one area called (area 0). The 6-routers are distributed among cities on Iraq's map. The first single source router begin from Arbil city which directly connected with neighbor router exists in Baghdad city with weight of 20. The router of Baghdad connected to two paths with router neighbors exist in Wasit city by weight of 11 and in Anbar city by weight of 20. The two paths that come from routers of Wasit and Anbar cities collected in neighbor router exists in Najaf city with weights

of 21 and 30 respectively. The router of Najaf city is directly connected to a single destination router of Basrah city with weight of 22.

Design of OSPF area network is shown in figure (1.1);

Fig.(1) OSPF area net in Iraq's map.

Weight means cost of route path from one path to another in network.

Theoretical analysis and calculations:

Dijkestra algorithm used in calculations to find the shortest path of OSPF area design network [7,8].

Input: a weighted graph *G=(V,E)*,

when *V* ; no. of the vertices (routers) in network.

E; no. of the edges (paths) in network.

Output: the path between two given routers ***s, t*** that minimizes the total weight (cost).

when ***s***; the source router in network.

t; the destination router of network.

A graph of network must be connected to calculate the shortest path as shown in the following table;

Name of paths (vertices)	**Edges(s,t)**	**weights**
Arbil → baghdad	(1,2)	20
Baghdad → Wasit	(2,4)	11
Baghdad →Anbar	(2,3)	20
Wasit → Najaf	(4,5)	21
Anbar → Najaf	(3,5)	30
Najaf → Basrah	(5,6)	22

The shortest path from Arbil to Basrah is come from Baghdad through Wasit to Najaf because the weight from Arbil to Basrah through Wasit equal to 74, but from Arbil to Basrah through Anbar equal to 92.

Simulation Results for theoretical calculation:

By using MATLAB software; when OSPF protocol is implemented The OSPF network for routers of cities in Iraq is drawn as shown in figure (2);

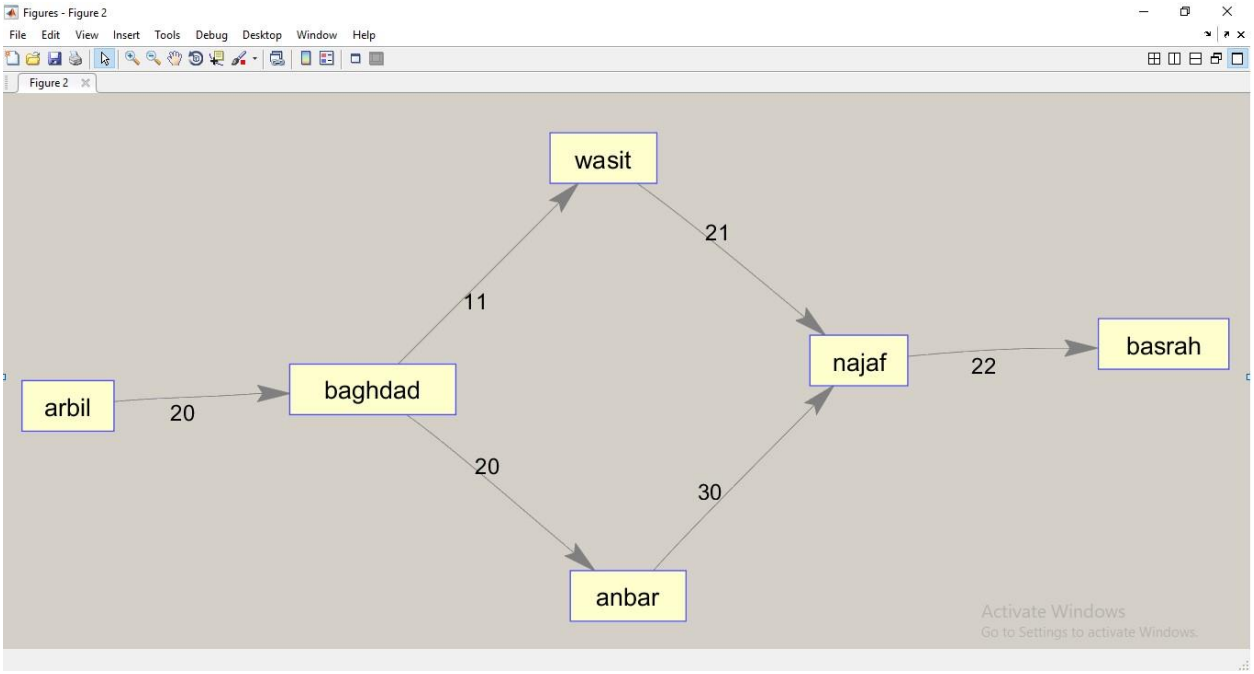


Fig (2) Shortest path of OSPF area net.

The shortest path of OSPF network for routers after implementation is drawn in figure (3);

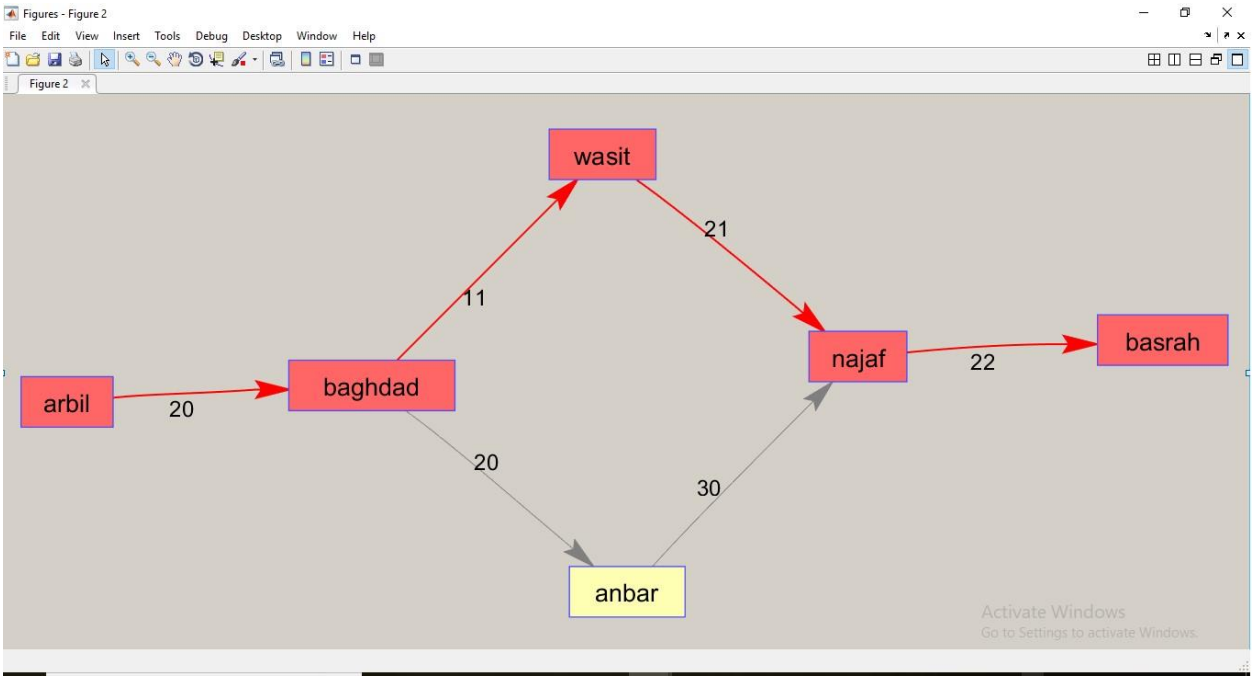


Fig.(2) Shortest path of OSPF area net.

A graph of the best path indicates in red color.

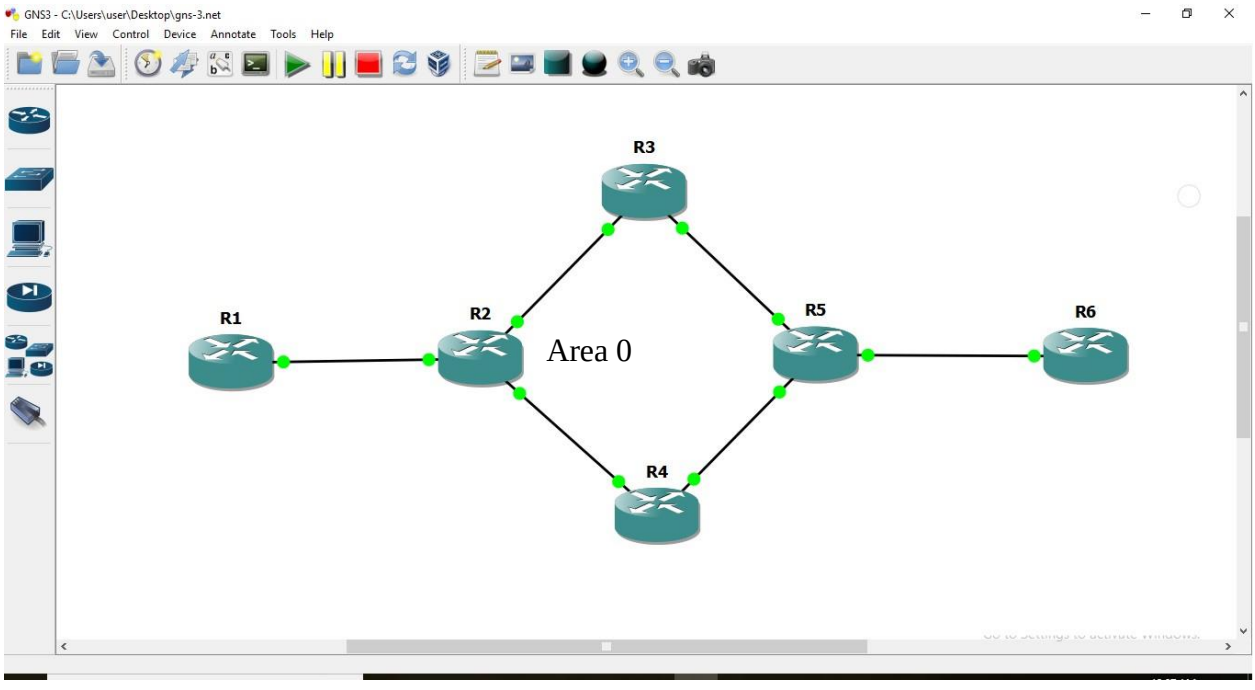
The command that used in MATLAB program to find and draw the shortest path is

[dist,path,pred]=graphshortestpath(cm,1,6)

Emulation Results for practical measurements:

To emulate design the network after implementation OSPF protocol, it used GNS-3 software;

The OSPF network (area 0) for routers of cities in Iraq is drawn as shown in figure (4);



(a) OSPF net /area 0 .

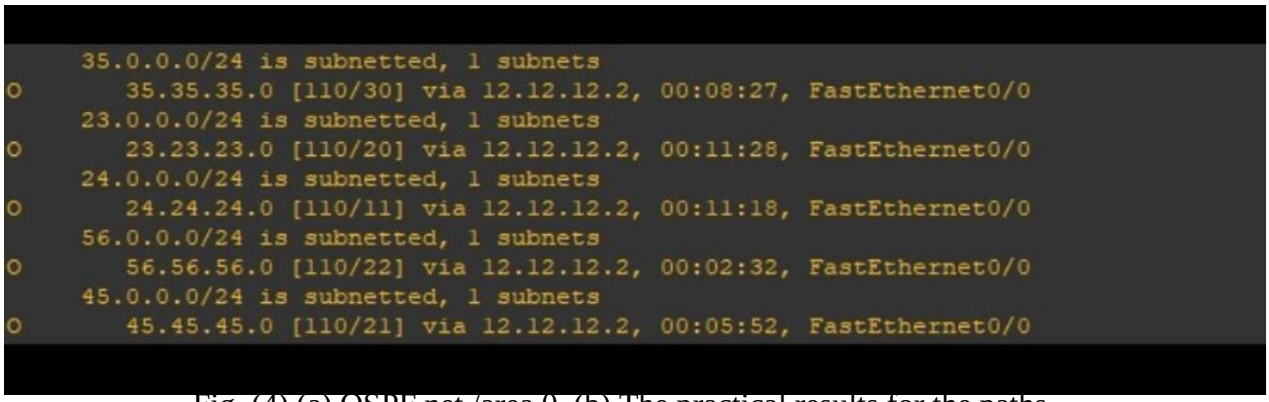


Fig. (4) (a) OSPF net /area 0. (b) The practical results for the paths

Emulation results show the IP addresses of networks that represent the paths from Arbil router to Basrah router. IP addresses paths are illustrate in the following table;

Routers	**Names**	**IP address**	**Weights**
R1→R2	Arbil→ Baghdad	12.12.12.0	20
R2→R3	Baghdad→ Anbar	23.23.23.0	20
R2→R4	Baghdad→ Wasit	24.24.24.0	11
R3→R5	Anbar→ Najaf	35.35.35.0	30
R4→R5	Wasit →Najaf	45.45.45.0	21
R5→R6	Najaf→ Basrah	56.56.56.0	22

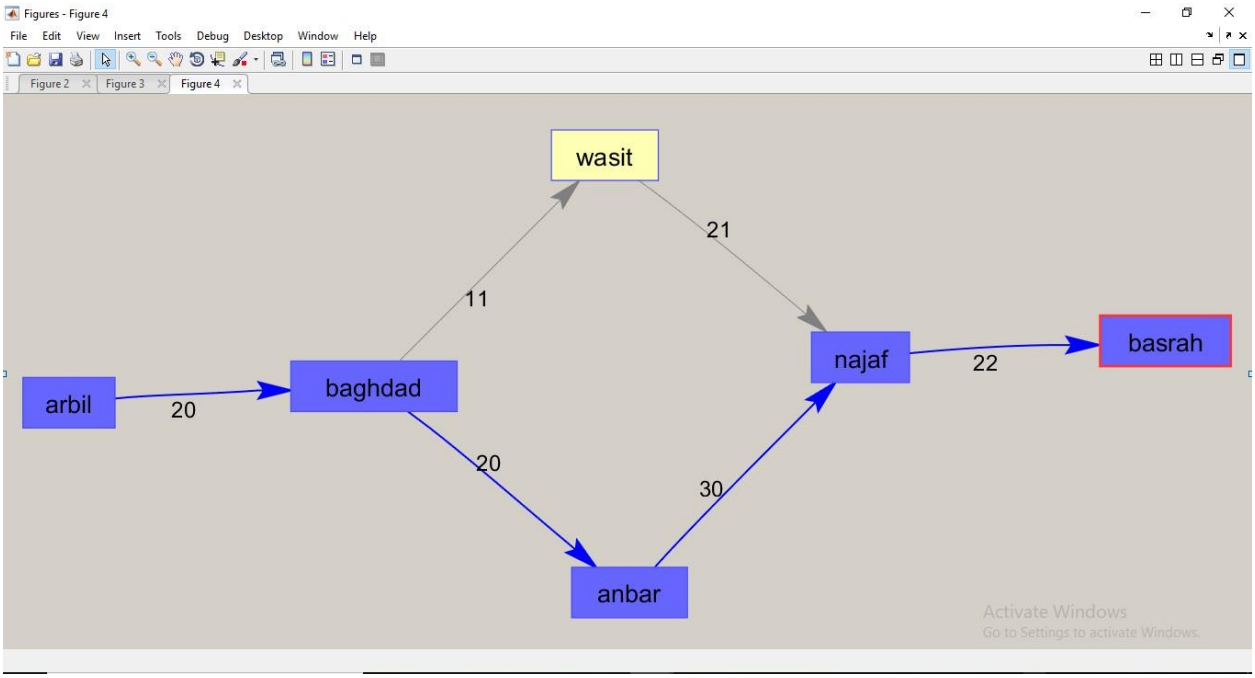
From table; it indicates the following ;

The shortest path comes from (R1-R2), (R2-R4), (R4-R5) to (R5-R6), when the weights are 20 + 11 + 21 +22 = 74. 74 is a no. of the total weights for the shortest path.

The total weights for the other path are 20 + 20 + 30 + 22 = 92.

Then; the shortest path has low weights rather than the other path.

If one of edges of the shortest path is down, it can be used the other path which is called a back up path as shown below

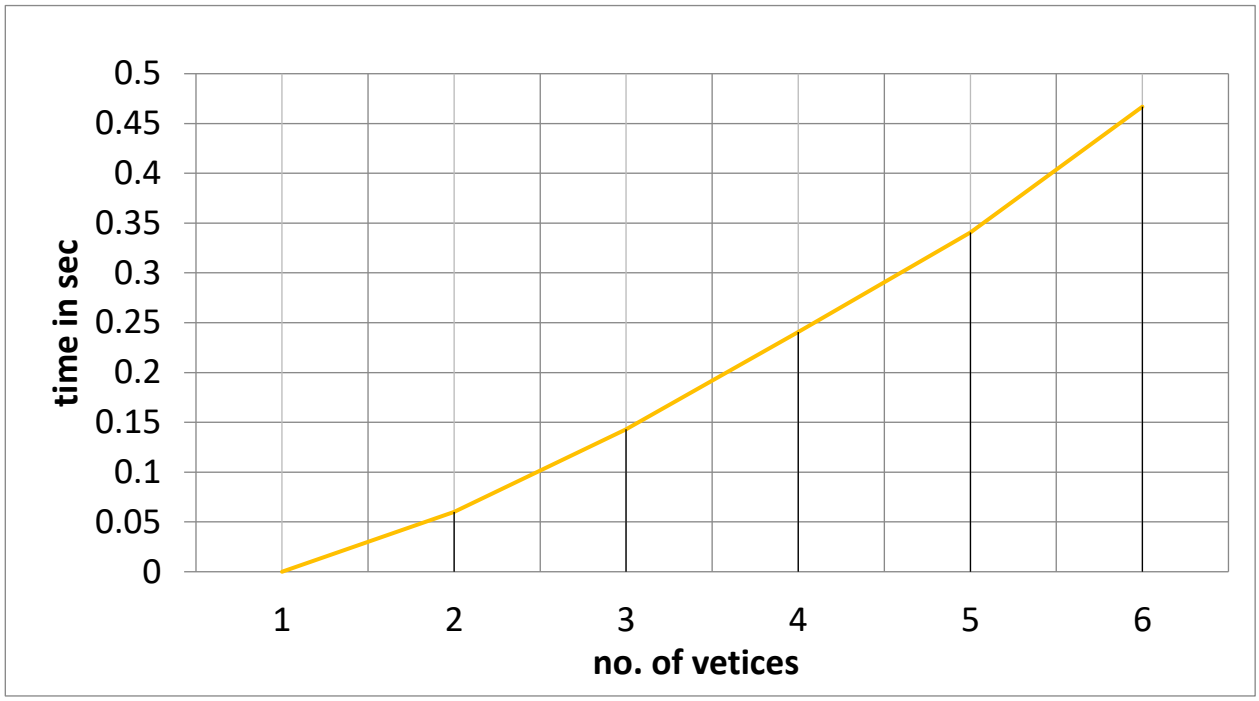



```
     35.0.0.0/24 is subnetted, 1 subnets
O       35.35.35.0 [110/20] via 45.45.45.5, 02:12:16, FastEthernet0/1
     23.0.0.0/24 is subnetted, 1 subnets
O       23.23.23.0 [110/30] via 45.45.45.5, 00:03:46, FastEthernet0/1
     56.0.0.0/24 is subnetted, 1 subnets
O       56.56.56.0 [110/11] via 45.45.45.5, 02:12:16, FastEthernet0/1
     12.0.0.0/24 is subnetted, 1 subnets
O       12.12.12.0 [110/20] via 24.24.24.2, 02:12:26, FastEthernet0/0
```

Figure(6); backup path using ip address.

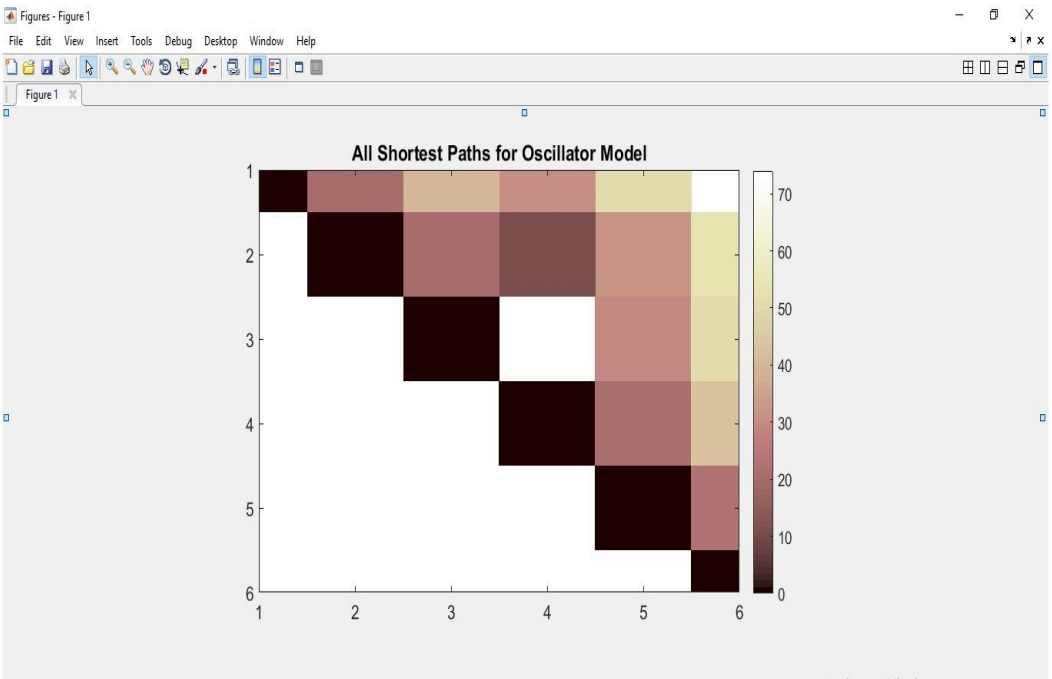
The running time for the shortest path has 19.02 msec., but the running time of the other path has 21.87 msec. so the shortest path has low running time compared with the other path. Theoretically when use (e log v),the running time is

Shown below;



Figure(7), The running time for the shortest path

The full scanning for the routers in Iraq country shows the shortest path between arbil and basrah as in figure (8);



Figure(8); the full scanning for routers in Iraq.

Conclusion:

Theoretical analysis and practical measurements indicate a good results to choose the best path for Iraq network when implementing OSPF protocol. The results show that the best path from Arbil to Basrah roads through Baghadad ,Wasit and Najaf roads with low weight of 74 compare with the other path that come from Baghdad, Anbar and Najaf roads with weight of 92. The best path has also low running time with 19.02 msec. compare with the other path with running time of 21.87 msec. one of feature of OSPF protocol, if one of edges in the shortest path is down, it can be used the other path as a back up in network.

Then, the desired paths of networks which that have low weights and low running time that make their CPU routers is intensive.

References:

[1] Shubhi, Prashant shukla " Comparative analysis of distance vector routing and link state protocols " , IJIRCCE, VOL. 3, Issue 10, October, 2015.

[2] Bhawana; " OSPF Implementation and applications in MATLAB " , IJTRE, VOL. 3, Issue 10, June , 2016.

[3] Uma Agarwal, Vipin Gupta , " Network routing algorithm using genetic algorithm and compare with route guidance algorithm " , IJSRET , ISSN: 2278-0882 , IEERET-2014 conference proceeding , 3-4 Novmber , 2014.

[4] Cisco book; "IP routing :OSPF configuration guide, Cisco IOS XE release 3s" ,iro-xe-3s-book< https://www.Cisco.com.

[5] Phani Raj Tadimety, " OSPF : a network routing protocol" , ISBN: 978-1-4842-1422-4, book< https://www.apress.com, 2015.

[6] Wikipedia , " Open shortest path first " , https://en.m.wikipedia.org/wiki/open-shortestpathfirst .

[7] Jaehyun Park , " Shortest path algorithm " , CS 97SI, Stanford university, June 29 , 2015.

[8]Melissa Yan , " Dijkestra algorithm ", http://math.mit.edu/~rothvoss/18.304.3PM/Presentationsl1-Melissa.pdf.

Electronic Learning System to Improve the Education Quality Using Of Satellite Communication Techniques

Ali Abdul Hussein Shyaa
Medical Technical Institute
Middle Technical University, Baghdad
E-mail: shyaa.ali1@gmail.com

Abstract

Facilitating the process of learning in real life by employing of available communication techniques (mostly by internet) is termed as E-learning. With time advance, education of today developed with more publicity so then, students from remote venues are getting access to their course of interest even though far once of courses, with such capability of education, students start searching the interested courses and teaching style as well, that raised a new concern to be applied into classical methods of educations. Rural and remote regions can be provided with quality education through E-learning, the same is proposed in this article to connect student with remote lecturers and tutors effectively.

Keywords*: E-learning, Satellite, GSM, Education, IoT, Orbit, CEC, UGC.*

Introduction

Learning is defined in [1] as skills and information that leads to facilitate people's lives and given by means of study. Furthermore, all knowledge is a mother of skills and it can be gained by multiple ways such as practical experience or teaching through a knowledgeable person; that is termed to the classical means of learning. Form the other hand, E-learning may term as gaining of knowledge in smarter way with more ease of information and knowledge delivery irrespective of location and time [2]. Many applications are used with E-learning such that: virtual classrooms, web learning, computer and mobile based leaning; another definition can be stated which is the information delivery by utilizing digital electronic and communication technologies such that internet, satellites, CD-ROM and interactive television [3]. All electronic learning systems (E-learning) are basically constructing of three functional components such content (things to be taught), service to be delivered to candidates and technology to be used for ensuring of those services. Students can be provoked for knowledge gaining by providing all interested education means such that virtual classrooms, quizzes, etc.

all those facilities can be accessed distantly, the same had been developed when internet and communication technologies have been developed.

Countries like India, have developed interested contributions in terms of distance education and electronic learning, EDUSAT is a unique kind satellite which being used for E-learning purposed and manufactured by ISRO (Indian space research organization) [4]. This satellite is connecting of large number of instates which are widely involved in electronic learning.

Since electronic learning is facilitating the efforts of its candidates to travel for attending classical lectures, it is well known as flexible and far less cost. Many advantages are obvious for this type of education system which are summarized as:

- Cost effective: if we compared the electronic learning with traditional learning, the first is lesser cost.
- Rapid and speedy: study that given in [5] has proved that, electronic learning is fifty present faster that classical learning, the study made one student course completion and feedbacks.
- Gap less communication learning: it reduces the problem of multiple instructors of the same subject as in here, every instructor is having own study material and own teaching methods.
- Mobility education: access is more ease and location independent [6].
- Time independent: candidate can join any time and less time constrains are provided by electronic learning.
- Upgradable: system content can be upgraded easily and quickly.

Students can be grouped together from different locality and share their knowledge and thoughts through electronic forums and discussion groups within E-learning system, such concept will help to construct a reliable community of students with portentous outcomes.

Developing process of electronic learning system at any county is strongly dependable on internet; so as to we have to study the internet users and population of internet in that particular region of interest. Statistical and published researched shown that internet users and their population is widely increased by the end of 2012.

For developing of E-learning system at middle-east region, studies shows a good population of internet users within this area so that efficient system can be built.

i. E-learning system

Indian satellite for education purposes which is designed and used by some large education institutes like IGNOU and NIOS will be detailed in this section.

1. EDUSAT

This family of satellite was designed on late time of 2014 as first educational server satellite from its kind in India. The mail agenda of this satellite was to provide a interactive distance education across the country; this technology is reflected the contribution of Indian space organization in education and their commitment to deliver education to those bodies on remote regions hence, they are maximizing the number of learner. [7] is listing the conceptual structure of this satellite; video and voice from live lectures and offline lectures are now possible to be broadcasting throughout all parts of country with help of this satellite which is placed in 74 degree east longitude of Geostationary orbit. Table 2 in below is summarizing the technical specifications of this satellite.

Table 1: EDUSAT design specifications.

Subject	Details
Mass to be lifted of	1920 Kilo gram
Dimensions	2.4x1.65x1.15 (in meters)
Power supply	Solar panels 2044 w
Life of service	84 month

The link budget of this satellite can be listed as below [7]:

- Six upper extended: 37 dB-watt coverage (national) with C-band transponders.
- Five lower bands: 55dB- watt for beam spot coverage which generates isotropic radiation of power at Ku-band.
- One deacon band: to control the uplink power and gaining as accurate as possible positioning of ground antenna.
- One lower band: at Ku-bands for national coverage at 50 dB-watt.

Figure 1, is giving the EDUSAT satellite architecture.

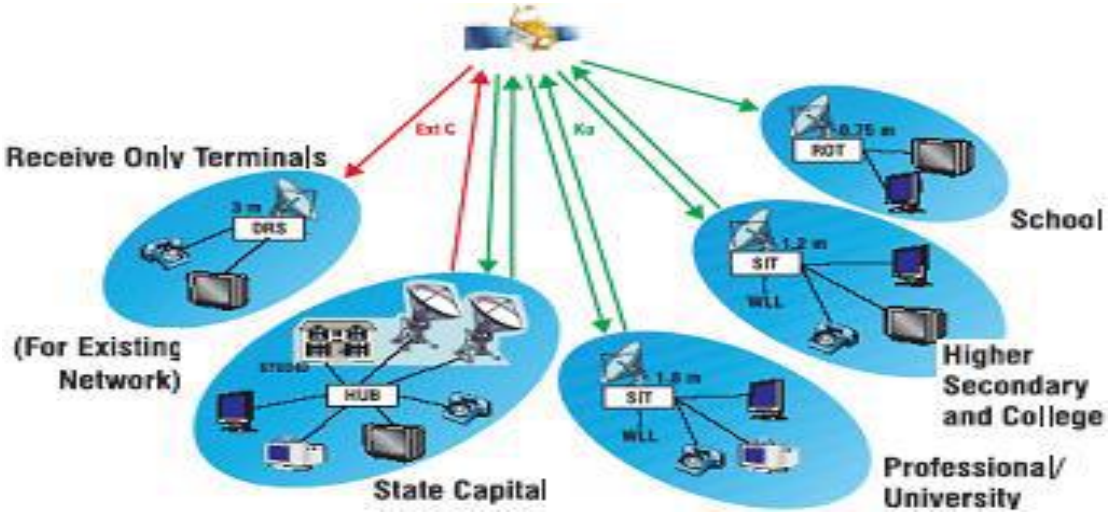


Figure 1: conceptual structure of EDUSAT.

Engineering education facilities are ensured to be delivered by EDUSAT across India, this setup is essentially helped to connect large number of facilities and students from different engineering colleges. The task of this satellite is facilitated by Indian space research organization (ISRO) by allotting a division for EDUSAT in space. At any particular university, the Grants Commission (UGC) are in turn communicating the subject syllabus by special band (channel) called consortium educational communication (CEC) [8]. Toper universities in India are using the services of EDUSAT such as IIT Chennai, IGNOU, IIT Kharagput and many others.

2. INSAT

This is multi-purpose satellite orbit that owned by Indian space research organization (ISRO), it is known as Indian national satellite system, main agenda of this satellite is to enhance communication services and broadcasting including education services. This satellite is developed jointly DoT, DoS and IRaD; and considered as largest satellite system in pacific Asia region. Table 3 is listing the series of INSAT which are recently launched and in service.

Table 2: INSAT satellite family.

SN.	Version
1	INSAT-2E
2	INSAT-3B
3	INSAT-3A
4	INSAT-3E

5	INSAT-3C
6	GSAT-2
7	KALPANA-1
8	EDUSAT
9	INSAT-4A
10	EDUSAT-1
11	GSAT-12
12	GSAT-8

Services like disaster alerting, weather forecasting are also provided by this satellite; 175 transponders are provided by his system on C band, EX-C band and Ku band making it a multipurpose satellite. By this more than 9.5 million people are able to watch television broadcast also satellite mobile services are being provided by the same. As stated in [9], telemedicine network is also supported by this system which is capable to link data of 160 hospital across the country.

ii. E-learning Case studies

1. IGNOU

As practice of electronic learning systems, Indira Gandhi National Open University is one good case that deploying an E-learning system. This organization had established around 30 years ago and serving around two million local students today; from the other hand, students from thirty countries are also being provided with education from this university. Very large network of regional centres and study schools are affiliated under this organization; advance learning methods and effective technologies of distance learning are being used within this university. In the domain of E-learning systems, IGNOU has started developing the environments for e-learners as follow:

- PAN AFRICAN Electronic Network and SAKSHAT: are developed by IGNOU as one of top portals of education.
- Electronic GyanKosh: is national digitized repository.

Those systems are used as platform for live virtual education, library and information by this university.

2. NIOS

The national institute of open schooling which is known formerly as NOS national open schooling is enrolling a 1.6 million student today. This organization is recognized as largest schooling institute in the world; NIOS was established on early 1979 by central board of secondary education (SBSE). This organization is declared globally as sustainable learner school of centric education and training by open learning and distant education as well. Following are some facts about NIOS:

- Largest open and distance schooling system in the world and in service since 1979 with 2235000 enrolled students since 1990.
- More than 20000 students are enrolling annually in vocational education and more than 250000 are enrolling in other open courses.
- NIOS is getting in touch with their leaner by more than 900 centres of vocational around the country and in abroad also.

NIOS is drawn essential participation in education development by providing a self-instructional print of study material, personal contact program supported by CD-ROM, moreover, practical sessions of training are also provided. NIOS learning programs are taking place also by help of radio and TV broadcasting.

Proposed system

Employing of satellite communication to create E-learning system can fill up the gap between urban large infrastructure learning organizations and others rural educational organization of less or few resource and infrastructure and hence good outcomes from all prospection could be yielded. The proposed Electronic learning system is comprised of three main units as in hereafter.

A. Mobile web applications: centralized educational services such as courses, tutorials and sessions can be accessed by students and teachers as well with help of smart phone applications of web facilities. These terminals i.e. mobile phones can be used by professors to login with their own authentication identity and then all of system services will be possible for them such as posting an assignment, taking the attendance report or even starting a quiz. Another service such as sending of feedback or results to the parents by SMS or email can be embedded with this application. From the other hand, students can login as well with their allotted authentication information and

participating the work through this application, they can also download any given materials by their teachers and attending the quizzes online, the mark can be given as well according the their performance. This system is expected to increase the interest of students to learn and hence developing the public ability.

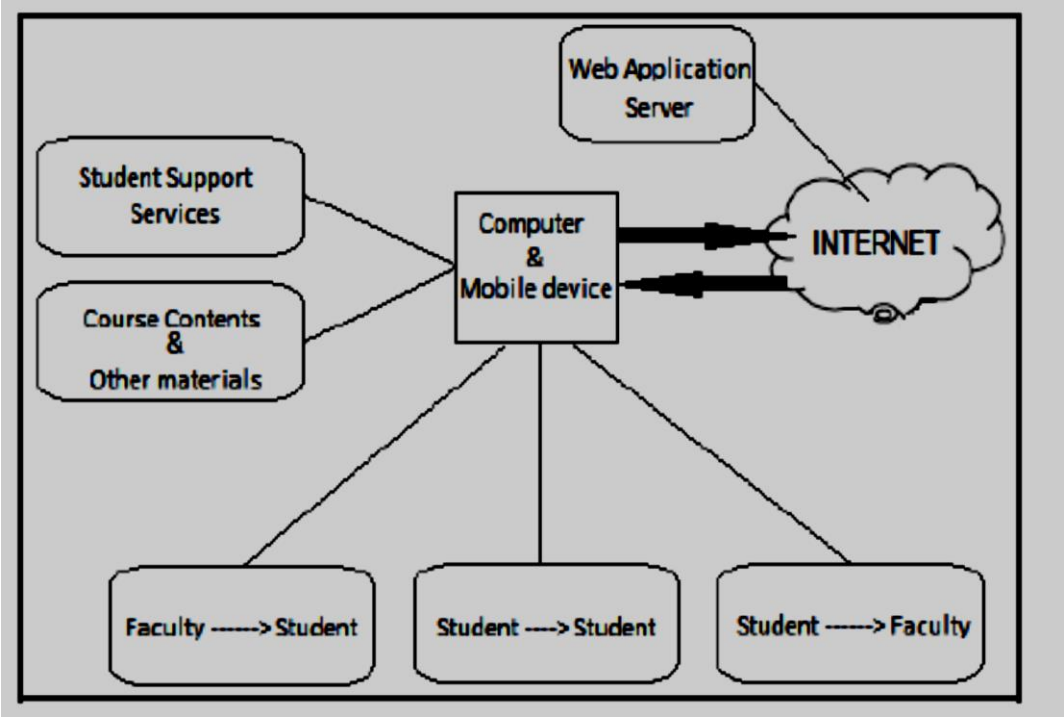


Figure 2: Mobile web application structure.

The figure above is detailing the structure of mobile applications that based on web and expected to provide the mentioned services in the above section. Apart from the mobile unit, user may login into the centralized system by using personal computers or laptops;

B. Channels of communication: it can be wire or wireless means of transmission, channel is related to the medium that used to share the data between the centralized system and candidate. Our proposed system is depending basically on educational satellite such as INSAT so, the channel of interest in our case will be satellite channel that contains of two beams: the uplink which is used to transmit the data from ground station to the satellite and downlink to receive the information by any terminal from satellite.

C. Ground station: this unit is performing the multimedia broadcasting such as video and voice; transition of live lectures or recorded once to the satellite in orbits and also

performing the control signalling to ensure successful delivery of information between any pair of communication. One point to be considered is that wider coverage is assured, that can be achieved by using a satellite station with multiple bands to cover as larger as possible geographical area.

Practical model

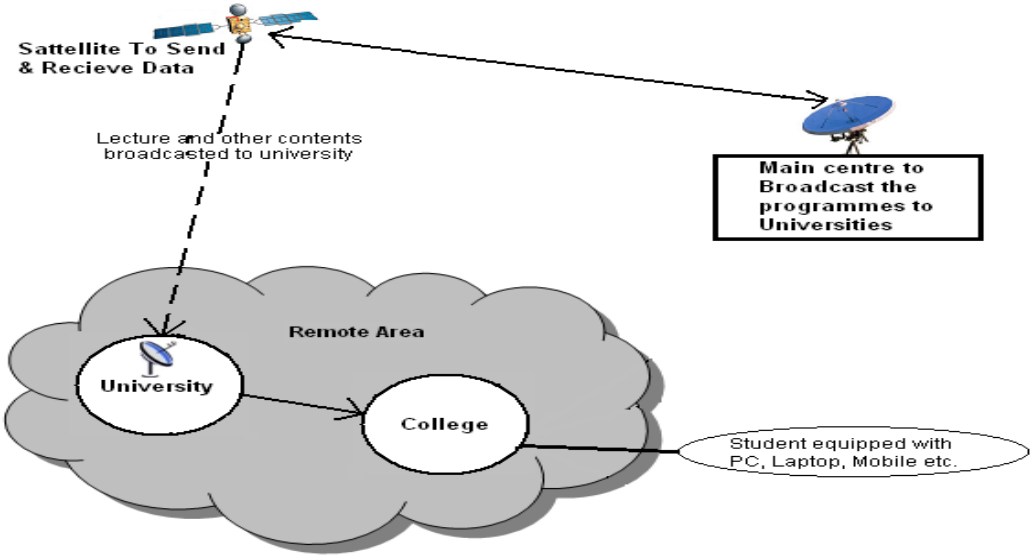


Figure 3: practical model of proposed system.

In our model, lectures or any other educational material can be receded or live broadcasted to the satellite station which is in turn rerouting this information to the registered receivers (universities) in different coverage regions, those receivers will be applying this material on the registered website of every college. A web application can be used over smart phone to access those courses or it can be done directly by web browsers from any terminal computer to access college website. Concept of interactive classrooms will provide the facilities of remotely attending of live lectures and communicating with the subject export by means of voice, the proposed work of us involves a web application with login facility where student, teachers can be login and the last can remark any student as absent or present, the event activity can be displayed also of this virtual classroom. Subject expert can send email to any parents.

Conclusion

In this article, we proposed electronic learning system underlying by satellite to improve education quality and increase the student ability of leaning. Satellite can help to deliver live

and interactive information for larger geographical areas. Three functional components are depended in our system; the data is being accessed by web application or personal computers.

References

[1]. Tina Klomsri and Matti Tedre, "E-learning Opens Door to the Global Community: Novice Users' Experiences of E-learning in a Somali University", MERLOT Journal of Online Learning and Teaching, Vol. 11, No. 2, June 2015.
[2]. http://www.elearningconsulting.com. Last accessed on 05/01/2012
[3]. E-learning Through Satellite, Information Paper, International Journal of Recent Trends in Engineering, Vol 1, No. 2, May 2009
[4]. https://www.isro.gov.in/publications-0 Last accessed on 10/04/2018
[5]. http://www.satelliteone.com/supportfiles/Spot_Beam_Short.pdf. Last accessed on 02/01/2018
[6]. Mamta Garg , Manoj Kumar Jindal ," EduSat- E-learning Through Satellite –Reaching the Unreached", International Journal of Recent Trends in Engineering, Vol 1, No. 2, May 2009.
[7]. Dr. S. Arulchelvan," EDUSAT Networks in Imparting Efficient Teaching-Learning Solutions", International Journal of Education and Psychological Research (IJEPR) ISSN: 2279-0179 Volume 2, Issue 2, pp: 23-32, April 2013.
[8]. Yazmi USA LLC," Operations management of INSAT satellites",ResearchGate
[9]. Sofia Tsekeridou and Thanassis Tiropanis, "Satellite-enabled educational services specification and requirements analysis based on user feedback", Int. J. Knowledge and Learning, Vol. 4, Nos. 2/3, 2008.

New Algorithm for solving the ODE with unbounded control operator

Ali Abdul Kadhum Ruhaima [(1)] & Dunya Mohee Haider[(2)]
Ali.abd.eng@nuc.edu.iq Al Nisour College University; E-mail: (1)
(2) Madent Alelm College Unversity ; Email: dunyamoheehaydee@gmail.com

Abstract

In this paper, a general frame work and for approximating solution ofinfinite dimensional linear quadratic optimal control problem via simegroup approach was presented.

A new algorithm for solving such problems was presented and applied on some examples that were solved using this new algorithm to reach the approximate solutions.

Keywords: ODE, linear quadratic, simgroup, operator, Banach space, Riccati

الخلاصة

في هذا البحث تم التطرق لتقريب الحل الامثل لمسائل السيطرة المثلى الخطية التربيعية اللامنتهية البعد باستخدام طريقة شبه الزمرة. تم تهيئة خوارزمية جديدة لحل تلك المسائل للوصول للحل التقريبي له.

1. Introduction

The theory of one parameter semigroups of linear operators on Banach space started earlier, acquired its core in 1948 with Hille-Yosida generation theorem, and attained its first apex with the 1957 edition of" semigroups and functional Analysis" by E.Hille and Philips. In the 1070's and 1980's, the theory reached a certain state of perfection.

1. Some basic definitions

Definition 1.1 [Pazy, 78]

Let X be a Banach space. A one parameter family $S(t): R^{+} \to \pounds(x)$ $0 < t < \infty$ of bounded linear operators from X into X is a simegroup of bounded linear operators on X if

i) $S(0) = I$, where I denote the Identity operator on X

ii) $S(t+s) = S(t).S(s)$ for every $t, s \geq 0$

Furthermore, the linear operator A defined on the domain

$$.D(A) = \left\{x \in X: \lim_{t\to 0}\left(\frac{S(t)x-x}{t}\right) exist\right\}$$

By $$Ax = \lim_{t\to 0}\left(\frac{S(t)x-x}{t}\right), x \in D(A)$$

is defined to be the infinitesimal generators of semigroup $S(t)$.

Definition 1.2 [Pazy, 78]

Let X be a Banach space, and let $S(t)$ be a bounded operators on X. $S(t)$

$\lim_{t\to 0}\|S(i)-I\|_{L(x)}=0$

Definition 1.3 [Balakrishnan, 76]

Let H be a Hilbert space. The infinitesimal generator A is called dissipative if $R(<Ax,x>_H)\leq 0$ for all $x\in D(A)$.

2. **Infinite Dimensional Linear Quadratic Regulator Problem** [Klaus, 00]

Consider the dynamic system

$\dot{x}=A_0x(t)+Bu(t)$

$x(0)$ is given , let H be real Hilbert space

$A: D(A_0)\subseteq H\to H$ is infinitesimal generators of C_0-semigroup $S(t)$ and A_0 is closed linear operator with dense domain such that for all $x\in D(A_0)$.

$A_0x=\lim_{t\to 0}\frac{S(t)-x}{t}$

$S(t)$: strongly continuous semigroup of operator A_0

B: bounded linear operator$H_u\to H$, where H_u stand for Hilbert spaces for control input (space) and $u(.)$ Is an element of $L_2\big((0,T);H_u\big)$.

The optimal control theory (Known as "the linear quadratic regulator problem (LQR)" is to find u in $L_2\big((0,T);H_u\big)$ so as minimize the cost functional:

$$q(u)=\int_0^T\left\{<Qz(t),z(t)>+\int_0^T<u(t),u(t)>\right\}dt$$

Q: Linear bounded nonnegative, definite operator for the Hilbert space$H$ to itself.

2. Theorem [Vurgin, 04]

Suppose A_0 is the infinitesimal generator of strongly continuous simpgroup $S(t)$ over a Hilbert space H, and suppose $\pi(t)$ for each t, $0\leq t\leq T$, is a linear bounded operator on H into H strongly continuous in $0\leq t\leq T$. Then the equation

$\frac{d}{dt}<x(t),y>=<x(t),\big(A_0+\pi(t)\big)'y>\quad ;y\in D(A_0)\quad x_0(0)$ is given

has a unique continuous solution $x(t)$, $0\leq t\leq T$ and

$$x(t)=\Phi(t,s)x(t);\quad t\geq s\geq 0;$$

3.1 Optimal control problem in infinite time

In this section we introduce an optimal control problem with infinite time.

Let H and U be Hilbert spaces endowed with inner product $<\cdot,\cdot>_H$ and $<\cdot,\cdot>_U$, respectively $\left\{< u,v >_H= \int_a^b u(x)v(x)dx\right\}$

Let $Q: H \to H$ and $J: U \to U$be bounded linear operators with the following properties:

i) the operator J is a positive defiite operators on U

ii) the operator Q is a positive definite operators on H

Consider the following optimal control problem (LQR) :

Find $u \in L_2\big((0,\infty); U\big)$ such that u minimize

$\Im(z_0,u) = \int_0^\infty \{< Qz(t), z(t) >_H + < Ju(t), u(t) >_U\}dt$

Subject to

$\frac{dz(t)}{dt} = A_0 z(t) + Bu(t)$ on H

$z(0) = z_0$

Where A_0 and B are defined.

4 . Computational algorithm for solving infinite dimensional LQR problem.

Step 1 (main problem)

Let H and U be Hilbert spaces endowed with inner product $<\cdot,\cdot>_H$ and $<\cdot,\cdot>_U$, respectively $\left\{< u,v >_H= \int_a^b u(x)v(x)dx\right\}$

Let $Q: H \to H$ and $J: U \to U$be bounded linear operators with the following properties:

i) the operator J is a positive definite operators on U

ii) the operator Q is a positive definite operator on H

Consider the following optimal control problem (LQR) :

Find $u \in L_2\big((0,\infty); U\big)$ such that u minimize

$\Im(z_0,u) = \int_0^\infty \{< Qz(t), z(t) >_H + < Ju(t), u(t) >_U\}dt$

Subject to

$\frac{dz(t)}{dt} = A_0 z(t) + Bu(t)$ on H

$z(0) = z_0$

Where A_0 and B , Q and $J$are defined as follow

$A_0: D(A_0) \subseteq H \to H$, the closed linear operatorA_0 with dense domain infinitesimal generators of C_0-semigroup $S(t)$

B: bounded linear operator $H_u \to H$ where $H_u \equiv L_2((0,\infty); U)$

Step 2 (Generating of semigroup)

Let V be a subspace of H such that $D(A_0) \subseteq V \subseteq H = H^* \subseteq V^* \subseteq D(A_0{}')$. Where * stand for the dual space. Let $S(t)$ be a C_0 −semigroup defined on $V$with infinitesimal generator A_0.

Step 3 (choosing the finite dimensional subspaces)

Finite dimensional subspace should be selected such that $V_N \subseteq V^*$. Since V_N is finite dimensional, it is span of a finite number of basis functions in V.

Step 3.1 (choosing the finite dimensional operators)

Define suitable operators that approximate A, B, Q and J of main problem, Let A_N, B_N, Q_N and J_N the approximate operators to A, B, Q and J respectively can be selected such that

1. The operator $A_N: V_N \to V_N$ generate a C_0-semigroup on V_N
2. The operator $B_N: U \to V_N$ is bounded.
3. The operator $Q_N: V_N \to V_N$ is bounded and semidefinite.
4. The operator $J_N: U \to U$ is bounded and positive definite.

Step 3.2 (Define the orthogonal projection)

Let P_N denote the orthogonal projection as$P_N: H \to V_N$. By restricting the bilinear form $a(.,.)$ toV_N, we have

$A_N: V_N \to V_N$ Such that $a(x,y) =< A_N x, y >$ for all n$x, y \in V_N$.

To define the operators $B_N, Q_N \; and \; J_N$, we utilize the projection operator P_N as

$B_N: U \to V_N$ Such that$B_N = P_N B$.

$Q_N: V_N \to V_N$ Such that$Q_N = P_N Q$.

Step 3.3 (matrix approximation of linear operators)

One of most important finite approximation is the matrix approximation of suitable linear operators, be the basis of V_N defined a $N \times N$ matrix to be:

$[M_N] = \left[< b_i, b_j >_H\right]_{i,j=1}^{N}$; Where the notation [,] is standing of matrix representation, and $<.,.>$ is standing for a suitable inner product. Notice that $[M_N]$ is a symmetric

matrix. The matrix representation for approximate operators A_N, A_N', B_N, Q_N can be done as follow:

Find the suitable basis $b_i \ , i = 1, \dots, N$ from the subspaceV_N, M_N-inner product on R^N by

$$< \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix}, \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix} >= [a_1 \quad \cdots \quad a_N][M_N]^T \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix} \text{ , for all } \quad \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix}, \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix} \in R^N$$

1. The matrix representation of the operatorA_N with respect to the basis vector $\{b_i\}_{i=1}^N$ and M_N-inner product of the matrix $[A_N]$ defined by

$$< A_N \sum_{i=1}^N a_i b_i, \sum_{j=1}^N c_j d_j > =< [A_N] \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix}, \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix} >_{N \times N} \text{ , for all } \quad \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix}, \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix} \in R^N$$

2. The matrix representation of $[A_N]'$ is given by

$$[A_N]' = [M_N]^{-1}[A_N]^T[M_N]$$

3. The matrix representation of the operatorQ_N with respect to the basis vector $\{b_i\}_{i=1}^N$ and M_N-inner product of the matrix $[Q_N]$ defined by

$$< Q_N \sum_{i=1}^N a_i b_i, \sum_{j=1}^N c_j d_j > =< [Q_N] \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix}, \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix} >_{N \times N} \text{ , for all } \quad \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix}, \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix} \in R^N$$

4. Note that U is finite dimensional. Thus the operator $B: U \to H$ is an operator of finite rank, this implies that there exist vector $f_i \in H, i = 1, \cdots, m$ such that

$$B \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} = \sum_{i=1}^m f_i u_i \qquad \forall \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} \in R^m$$

In this case $[B_N] \in R^{m,n}$

The matrix representation of B_N is the matrix given by

$$< B_N \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}, \sum_{j=1}^m c_j b_j > =< [B_N] \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}, \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix} >_{M \times N} \text{ , for all } \quad \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} \in R^m, \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix} \in R^N$$

Furthermore

$$[B_N]' = [B_N]^T[M_N]$$

Step 4 (finite approximate of LQR problem)

Consider the above problem where$U = R^m$. The main problem is then approximated as follows:

Find $u_N \in L^2((0,\infty);U)$ so that

u_N Minimizes the following cost functional

$$\Im(z_{0,N}, u_N(.)) = \int_0^\infty \{< Q_N z_N(t), z_N(t) >_H + < J_N u_N(t), u_N(t) >_U\} dt$$

Subject the following dynamical control system

$$\frac{dz_N(t)}{dt} = A_N z_N(t) + B_N u_N(t) \quad \text{on } H$$

$$z_N(0) = z_{0,N}$$

Where A_N, B_N are defined in step 3

So in matrix representation in step 3 for our problem we write the matrix linear quadratic regular problem $([LQR_N])$

Find $[u_N(t)] \in L^2((0,\infty);R^N)$ such that $[u_N]$ minimizes

$$\Im([z_{0,N}], [u_N(.)]) = \int_0^\infty \{< [Q_N][z_N(t)], z_N(t) >_{M_N} + < [J_N][u_N(t)], [u_N(t)] >_U\} dt$$

Subject to

$$\frac{[dz_N(t)]}{dt} = [A_N][z_N(t)] + [B_N][u_N(t)] \quad \text{on } R^N$$

Where $[A_N], [B_N]$, $[Q_N]$, $[J_N]$ are defined in step 3.3

Step 5 (finding the algebraic Riccati equation (ARE))

Theorem (4.1) [Gibson, 81]

Consider the above optimal control problem where the operators are satisfying the following conditions:

The operator J is positive definite operator on U. The operator Q is positive semidefinite operator on H. The closed linear operator A_0 with dense domain is infinitesimal generators of C_0-semigroup $S(t)$ and the operator B is bounded linear operator. Suppose the semigroup $S(t)$ is exponentially stable. Then there is a unique control $u_0(t)$ minimizing which has the form

$$u_0(t) = -B'\pi(t)z_0(t)$$

Where

$$\pi(t)z_0(t) = \int_t^\tau S(s-t)Jz_0(s)ds$$

Where $z_0(t)$ is the unique continuous solution of

$< z_0(t), y > =< z_0(0), y > + \int_0^t < z_0(s), (A_0 + \pi)'y > ds, \quad y \in D(A_0')$.

π is bounded linear operator mapping $(\pi: H \to H)$ and characterized as the unique solution of

$0 =< Qz, y > +< \pi z, A_0 y > +< A_0 z, \pi y > -< B\pi z, J^{-1}B'y >$

The Algebraic Riccati equation for the system ( ) is defined as follows

$< Qz, y > +< \pi z, A_0 y > +< A_0 z, \pi y > -< B\pi z, J^{-1}B'y >= 0$

$< (Q + A_0\pi + \pi A_0 - \pi B J^{-1}B'\pi)z, y >= 0$

But z and y does not equal to zero. Hence

$(Q + A_0'\pi + \pi A_0 - \pi B J^{-1}B'\pi) = 0$

Or $A_0'\pi + \pi A_0 - \pi B J^{-1}B'\pi + Q = 0$ (this equation called Algebraic Riccati Equation (ARE)

Then the optimal control is $u(t) = -J^{-1}B'\pi z(t)$ and hence is a solution of feedback

$\dot{z}_N = A_N z_N + B_N u_N$, on using the semigroup

$\dot{z} = Az + B(-B'\pi z) = Az - \pi z$

$= Az - \int_t^\tau S(s-t)'Jz(s)ds$

On using the theorem ( ) on the system we get the following ARE

$A_N'\pi_N + \pi_N A_N - \pi_N B_N J_N^{-1} B_N'\pi_N + Q_N = 0$

On using the hypotheses of the theorem ( ) (LQR_N) is transformed into solving the following matrix equation

$[M_N]^{-1}[A_N]^T[M_N][\pi_N] + [\pi_N][A_N] - [\pi_N][B_N][J_N]^{-1}[B_N]^T[M_N][\pi_N] + [Q_N] = 0$

Step 6 (finding the Optimal Control operator)

Define the bounded linear operator $K: H \to U$ where K defined as $K = J^{-1}B'\pi$, such that $A_0 - BK$ generates an exponentially stable semigroup on H. We get the optimal control:

$u(t) = -J^{-1}B'\pi z(t) \quad 0 \le t \le \infty$

If the system ( ) – ( ) satisfy the hypotheses of theorem ( ). The optimal control u_N is defined as

$u_N(t) = -K_N z_N(t) \equiv -J_N^{-1}B_N'\pi_N z(t)$

The optimal control $u = -J^{-1}B'\pi z$ can be approximated by

$u_N = -J_N^{-1} B_N{}' \pi_N z \equiv -K_N z_N$

The optimal control for $[(LQR_N])$is become:

$$[u_N(t)] = -[K_N][z_N(t)] \equiv -[J_N^{-1}][B_N]'[\pi_N][z(t)]$$
$$= -[J_N^{-1}][B_N]^T[M_N][\pi_N][\pi_N][z(t)]$$

Let $[P_N] = [M_N][\pi_N]$

$[A_N]^T[P_N] + [P_N][A_N] - [P_N][B_N][J_N^{-1}][B_N]^T[P_N] + [M_N][Q_N] = 0$

Then the optimal control ( ) can be rewritten as follows

$$[u_N(t)] = -[J_N^{-1}][B_N]^T[P_N][\pi_N][z_N(t)]$$

Where $[\pi_N] = [M_N]^{-1}[P_N]$

Convergence of the Riccati and feedback optimal control operator

In this section we provide the frame work for proving of the Riccati $\pi_N \to \pi$

Let $H = L^2(0,1)$, $V = H_0^2(0,1)$ $and\ U = R^m$

Let A, B, Q and J satisfy the hypotheses in the theorem(4.1) and consider the following problem (#)

$\frac{dy}{dt} = Ay(t) + Bu(t)$ $on\ H$

$y(0) = y_0$

Together with associated performance measure

$\Im(y_0, u) = \int_0^\infty \{< Qy(t), y(t) >_H + < Ju(t), u(t) >_U\} dt$

Consider the following optimal control problem(#) minimize $\Im(y_0, u)$ over $u \in L^2\big((0,\infty); U\big)$ subject to ( ) ( ). The problem (#) is approximated by a sequence of approximate problem. Let $V_N, N = 1,2,\cdots$ be a sequence of finite dimensional subspace of $V \subseteq H$ and $P_N$be the orthogonal projection operators $P_N : H \to V_N$ under the H inner product. Let $S_N(t)$be the seq. of C_0 −semigroup defined on V_N with infinitesimal generator $A_N \in \mathcal{L}(V_N)$ and $S_N(t)$' denote the respective adjoint semigroup. Additionally, let $B_N \in \mathcal{L}(U, V_N)$ and $Q \in \mathcal{L}(V_N)$

The following approximate optimal control problem ($\#^N$) has been considered as follows

Consider the problem ($\#^N$): minimize the following objective function

$\bar{\Im}\big(y_{0,N}, u_N\big) = \int_0^\infty \{< Q_N y_N(t), y_N(t) >_{V_N} + < J_N u_N(t), u_N(t) >_U\} dt$

Over $u_N \in L^2\big((0,\infty); U\big)$ subject to

$\frac{dy_N}{dt} = A_N y_N(t) + B_N u_N(t)$ $on\ H$

$y_N(0) = P_N y_0$

We shall say that the function $u_N \in L^2((0,\infty);U)$ is admissible control for initial state $y_N(0)$ if $\Im_N(y_N(0),u_N)$ is finite. Bancks and Kunish theorem (Vurgin) establishes sufficiency for the convergences of $\pi_N \to \pi$ using the following assumptions:-

H1) for each $y_N(0) \in V_N$, there exists a control $u_N \in L^2((0,\infty);U)$ for $(\#^N)$, and any admissible control for $(\#^N)$ derives the state to zero asymptotically.

H2) i) for each$z \in H$; $S_N(t)P_N z \to S(t)z$, with convergence uniform in t on bounded subsets on$[0,\infty)$.

ii) for each $z \in H$; $S'_N(t)P_N z \to S'(t)z$, with convergence uniform in t on bounded subsets on $[0,\infty)$.

iii) for each $u \in U$; $B_N P_N u \to Bu$

and for each $u \in U$; $B'_N P_N u \to B'u$

iv) each $z \in H$; $Q_N P_N z \to Qz$

C1) each $z \in V$; the exist an element $z \in V_N$ such that $\varepsilon(N) \to 0$ as$N \to \infty$.

C2) the pair (A,B) is stabilizable

C3) There exist a bounded operator $K: H \to U$ the stabilizes (A,B). Furthermore, there exist an integer N_0 such that for $N \geq N_0$, the order (A_N,B_N) is stabilizable by K.

Theorem (3) (Banks-Kunich) [Banks, 84]

Suppose (H1) and (H2) and (C1)-(C3) holds and suppose that $J > 0, Q > 0, Q_N > 0$. Then there exit an unique Riccati operator π and π_N associated with (#) and$(\#^N)$ on H and V_N respectively such that if P_N denote the orthogonal projection intoV_N then

$\pi_N P_N z \to \pi z \qquad \forall z \in H$

$S_N(t)P_N z \to S(t)z \qquad \forall z \in H$

$u_N(t) \to u(t) \qquad \forall u \in U$

With these statement holding uniformly in t on compact subsets of $[0,\infty)$. Here, $S_N(t)$ and $S(t)$ are the semigroup generated by $A_N - P_N B_N J^{-1} B_N^{-1} P_N \pi_N$ and $A - BJ^{-1}B'\pi$ and $u(t)$ and $u_N(t)$ are the optimal feedback for (#) and$(\#^N)$ respectively and π_N is positive definite Riccati operator that solve ARE

$A'_N \pi_N + \pi_N A_N - \pi_N B_N J_N^{-1} B'_N \pi_N + Q = 0$

Now when we construct (LQR_N) , then they will be the feedback gains $K_N = J_N^{-1}B'_N P_N \pi_N$ converge to$K = J^{-1}B'\pi$, the resulting feedback gain for (LQR), specially we want $\|K - K_N\|_{L(H,V)} \to 0$

Remark

1. $\pi_N P_N z \to \pi z \qquad \forall z \in H$
2. $S_N(t)P_N z \to S(t)z \qquad \forall z \in H$
3. $u_N(t) \to u(t) \qquad \forall u \in U$
4. $A_N z \to Az \qquad \forall z \in H$
5. $K_N \to K$

The approximate optimal control is $u_N = -K_N z_N = -J_N^{-1}B'_N \pi_N z_N$

The approximate optimal solution is

$\dot{z}_N = A_N z_N + B_N u_N \qquad ; z(0) = z_0$

$\dot{z}_N = A_N z_N - B_N K_N z_N$

$\dot{z}_N = (A_N - B_N K_N) z_N$

Hence $z_N \to z$

and

$\Im(z_{0,N}, u_N(.)) \to \Im(z_0, u_N(.))$

Example: simple problem

Let H and U be a Hilbert spaces endowed with inner product $<\cdot,\cdot>_H$ and $<\cdot,\cdot>_U$, respectively. Let $Q: H \to H$ and $J: U \to U$ be a bounded linear operator with the following properties

i) The operator J is a positive definite operator on.

ii) The operator Q is a positive semidefinite operator on.

Consider the following optimal control problem(LQR):

Find $u_N \in L^2((0,\infty); U)$ such that u minimize

$\Im(z_0, u) = \int_0^\infty \{< Qz(t), z(t) >_H + < Ju(t), u(t) >_U\} dt$

Subject to

$\frac{d}{dt}\begin{bmatrix} v(t) \\ w(t) \end{bmatrix} = \begin{bmatrix} \mu & R \\ 0 & -\mu \end{bmatrix}\begin{bmatrix} v(t) \\ w(t) \end{bmatrix} - \begin{bmatrix} p \\ q \end{bmatrix} u(t)$

$z(0) = z_0 \qquad$ In H.

Where $A,\ B,\ Q$ and J are defined as follow:

$A: D(A) \subseteq H \to H$; The closed linear closed operator A with dense domain is infinitesimal generator of C_0-semigroup$S(t)$.

$B: H_u \to H$ Bounded linear operator; where H_u instead of$L^2\big((0,\infty);U\big)$ where $R > 0$, $\mu > 0$ and p, q are not both zero.

If we define $A = \begin{bmatrix} \mu & R \\ 0 & -\mu \end{bmatrix}$ and $B = \begin{bmatrix} p \\ q \end{bmatrix}$ then the linear system can be written as

$$\frac{d}{dt}\begin{bmatrix} v(t) \\ w(t) \end{bmatrix} = A.\begin{bmatrix} v(t) \\ w(t) \end{bmatrix} - B.u(t)$$

Where $p \neq 0$ and $q \neq 0$ and $R > 0$, $\mu > 0$

So we use the theorem (3)

Then the positive definite solution π of ARE

$A'\pi + \pi A - \pi\pi BB'\pi + Q = 0$

Is given by

$$\pi(\mu, p, R) = \begin{bmatrix} \frac{\mu+\sqrt{\mu^2+p^2}}{p^2} & \frac{R}{p^2} \\ \frac{R}{p^2} & \frac{R^2-p^2}{2\mu p^2} \end{bmatrix}$$

As R become large, the matrix A become more non normal

Moreover $p \geq\geq 1$ $\lim_{R\to\infty^+} \|\pi(\mu, p, R)\| = \infty$

Thus, the ARE can become ill-condition for large value of R.

However when $R \to 0$ the matrix A become normal (self-adjoint) and when $R = 0$

$$\pi(\mu, p, 0) = \begin{bmatrix} \frac{\mu+\sqrt{\mu^2+p^2}}{p^2} & 0 \\ 0 & \frac{R^2-p^2}{2\mu p^2} \end{bmatrix}$$

References:

[1 Balakrishnan, 76] A.V. Balakrishnan, “Applied Functional Analysis”, 1976

[Banks, 84] Banks and Kunisch, “The Linear regulator problem for parabolic system “, SIAM J., control and optimization, 1984.

[Fucik, 80] Svatopulk Fucik, “Solvability of Nonlinear Equations and boundary value problem”, D. Reidel Publishing Company, Holland, 1980

[Gibson, 81] J.S. Gibson “The Riccati integral equations for optimal control prblems in Hilbert spaces “, SIAM J. , control and optimization , 1981.

[Klaus, 00] Klaus Engel, “one parameter semigroups for linear evolution equations”, 2000.

[Pazzy, 78] Ammon Pazzy, “Semigroups of linear operators and applications to partial differential equations “, Springer -Verlag, New-York 1978.

[Vugrin, 04] Eric D. Vurgin , “ On approximation and optimal control of non-normal distributed parameter system”, PhD Thesis, Balacksburg, Virginia, USA, 2004.

Breast cancer detection using photoacoustic imaging

Luma W. Jameel
Alnisour University College (Private Sector)
Email: luma.eng@nuc.edu.iq

Abstract

This work is about a tumor detection using photoacoustic technique, the photoacoustic wave's generation and detection in tissue phantom (liver in chicken breast) which used to represent the real tissue. The one dimensional photoacoustic signal was generated using Nd:YAG laser, detected using piezoelectric detector and displayed on a storage oscilloscope, the experimental work is done at Different depth, thickness and the distance to the detector for several phantoms and different laser energy 250,400 and 500 mJ.

Keyword : Breast cancer, photoacoustic imaging, PA

الخلاصة

هذا العمل هو الكشف عن الاورام باستخدام تقنية التصوير الضوئي الصوتي، ان توليد الموجه الضوئية الصوتية وكشف عنها في الخليا باستخدام شي تجريبي التي تستخدم لتمثيل الخلايا الحقيقة.الاشارة احادية البعد تولدة باستخدام Nd:YAG ليزر وتكشف باستخدام كاشف البيزو وتعرض باستخدام عارض الموجات، التجربة العملية تمت باستخدام عمق وعرض ومسافة مختلفة وطاقات مختلفة.

Introduction

Medical imaging is very important for medical diagnostics and research. There are so many medical imaging techniques, and some of them have weaknesses. For instances, x-ray computerized tomography is limited by the accumulation of ionizing radiation which are harmful to human body, ultrasound imaging is limited by its poor contrast and pure optical imaging techniques are unable to effectively visualize the structures several centimeters deep in the tissue due to the strong scattering of biological tissue etc[1].

A sound or stress wave is produced because of the thermoelastic expansion by absorption of short EM pulse that is induced by a slight temperature rise, typically in millikelvin range, as a result PA signal is emitted. The excited PA signal is locally determined by the EM absorption

and scattering properties, the thermal properties, including the thermal diffusivity, thermal expansion coefficient, and the elastic properties of the sample [2].

The experimental work using a phantom to simulate a real tissue.

Theory

EM energy in the optical (from visible (400–700 nm) to near-IR (700–1100 nm)) and RF regions is used for PA excitation in soft tissues because the waves in these regions are nonionizing, safe for human and can provide the high contrast and enough penetration depths [1].

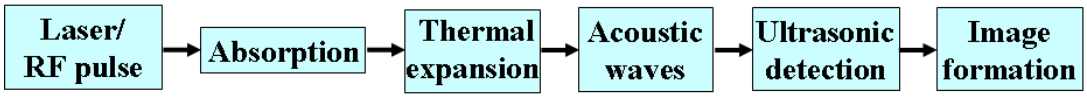


Fig1: Schematic illustration of PA imaging.

The optical absorption in biological tissues can be due to endogenous molecules such as Hb or melanin, or exogenously delivered contrast agents (exogenous Hb, for example). Blood usually has orders of magnitude larger absorption than surrounding tissues, so there is sufficient endogenous contrast for PA imaging to visualize blood vessels, for imaging *in vivo* subcutaneous vasculature (the arrangement or the distribution of blood vessels in an organ or body part) in small animals and human, monitoring tumor or cancer, mapping blood oxygenation, imaging functional brain, detecting skin melanoma (a malignant tumor of melanocytes which are found in skin, also in the bowel and the eye), monitoring of vascular damage during tumor photodynamic therapy [1].

Laser Tissue Interaction

When the laser light strikes a tissue surface, it can be reflected and refracted, scattered, absorbed or transmitted. The fractional intensity that goes into these different processes depends on the optical properties of the tissue like it is reflectivity, scattering and absorption coefficients, particle size, as well as the laser parameters like wavelength, energy, pulse duration, operation mode and output spectral profile [3].

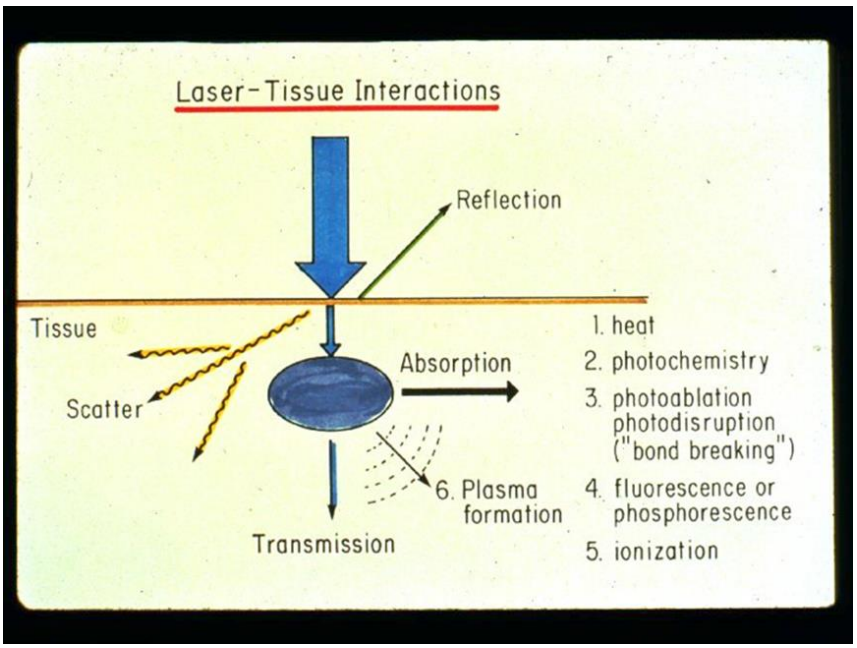


Fig2: Laser-tissue interaction [4].

Two laws are frequently applied; they describe the effect of either the thickness or concentration on absorption, respectively. They are commonly Lambert's law and Beer's law (Beer-Lambert's law) and are expressed by:

$$I_{(z)} = I_{\circ} ex\, p(-\mu_a z) \qquad (1)$$

$$I_{(z)} = I_{\circ} \exp(-kcz) \qquad (2)$$

Where, z is the optical axis, I_z is the intensity at a distance z, I is the incident intensity, μ_a is the absorption coefficient of medium, c is the concentration of the absorbing agent and k depends on the internal parameter other than concentration.

Photoacoustic Tissue Interaction

EM-pulse excited pressure acts as an acoustic source and initiates further acoustic wave propagation in three-dimensional (3D) space. In the low-megahertz frequency range, ultrasound in soft tissues has the properties of low scattering and deep penetration. The total attenuation results from the combined losses due to both absorption and scattering, while the scatter component accounts for about 10%–15% of the total attenuation. The attenuation of all

tissues is temperature and frequency dependent. The frequency dependency of ultrasonic attenuation can be represented by the expression μ =*afb*, where μ is the ultrasonic attenuation coefficient, *a* and *b* are constants, and *f* is the frequency of ultrasound. A mean value of ultrasound attenuation equals ~0.6 dB cm−1 MHz−1 for soft tissues [2].

Acoustic Impedance is the resistance to travel that a sound beam encounters as it passes through a medium, such as human tissue. Just as velocity is dependent on density, so is acoustic impedance. The acoustic impedance is directly proportional to tissue density.

In human soft tissue, acoustic impedance is defined as the product of the density and velocity.

$Z = \rho V$ (3)

Where Z: acoustic impedance, ρ: density of medium, V: velocity of sound. Slight difference in acoustic impedance between two tissues create an interface that will cause a portion of the ultrasound wave to be reflected [5].

This report is concerned with the thermoelestic production of sound in an optically absorbing medium. Is essence, thermoelestic sound is produced by the transient heating of a restricted volume by light energy. This condition is referred to as "stress confinement" and it requires that the optical irradiance, often done with a laser, occurs over a very short amount of time. More specifically, the laser pulse duration must be less than the amount of time needed for acoustic energy to propagate out of the area of absorption.

If the condition for stress confinement is met, and if the fluid is stationary with isotropic acoustic properties, the wave equation for acoustic pressure,

P, generated within the laser-irradiated volume is [6]:

$$\nabla^2 P(r,t) - \frac{1}{v^2}\frac{\partial^{"}}{\partial t^2}\, P(r,t) = \frac{\beta}{c_p}\frac{\partial}{\partial t} H(r,t) \qquad (4)$$

Where H(r,t) is the heating function defied as the thermal energy deposited by the EM radiation per time per volume (υ-the speed of sound, β is the isobaric volume expansion coefficient, and c_p is the specific heat capacity). In general P and H depend on the positive r=(x,y,z) and time t.

Instantaneous laser irradiation (denoted by the impulse function, $\delta(t)$, the heating function can be modeled as:

$H(r,t) = H(r)\,\delta(t)$ (5)

Where: H(r) is the volumetric heat density, $\delta(t)$ is optical absorption depths.

Experimental work

The work will done using Q-Switch Nd-YAG laser consists of many parts and systems such as: light route system, power supply, control system and cooling system. The light route system is installed into the hand piece, while the other systems are installed into the machine box of power supply. Nd:YAG laser with power supply, control panel and hand piece are shown in figure 3.

This type of laser is used in medical applications for removing eyebrow and tattoos as is can penetrate into deep layers of the skin with wavelength 1064 nm, so it is easily absorbed by blue, black and green color pigment; therefore, it is very suitable to treat the pigment disease in the dermal layer.

While for wavelength 532 nm it can penetrate into superficial layers of the skin; therefore, it is very easily absorbed by the brown, red and deep-brown color pigment. For example, eliminating freckles, pigment spots and other light-color tattoos. Technical Specifications of a Q-Switched Nd:YAG laser are illustrated in table 1 [7].

Table 1: Nd:YAG laser Technical Specifications

Laser type	Q switch Nd: YAG laser
Laser wavelength	1064nm/ 532nm
Maximum pulse energy	1000 mJ
Pulse width	10 ns
Frequency	1-6 Hz
Display	Button Control Screen
Cooling method	Inner loop water cooling circuit with air cooled external heat exchanger
Power supply	AC220V/ 5A/ 50Hz

Fig3: Nd:YAG laser

Piezoelectric detector (LDT1-028K) can be used for detecting physical phenomena such as vibration or impact, so that it was chosen for detecting the photoacoustic wave which was resulted from Nd:YAG laser after passing through the phantom. The piezo film element is laminated by a sheet of polyester (Mylar), and produces a useable electrical signal output when forces are applied to the sensing area. The dual wire lead attached to the sensor allows a circuit or monitoring device to process the signal as shown in figure 4 [8].

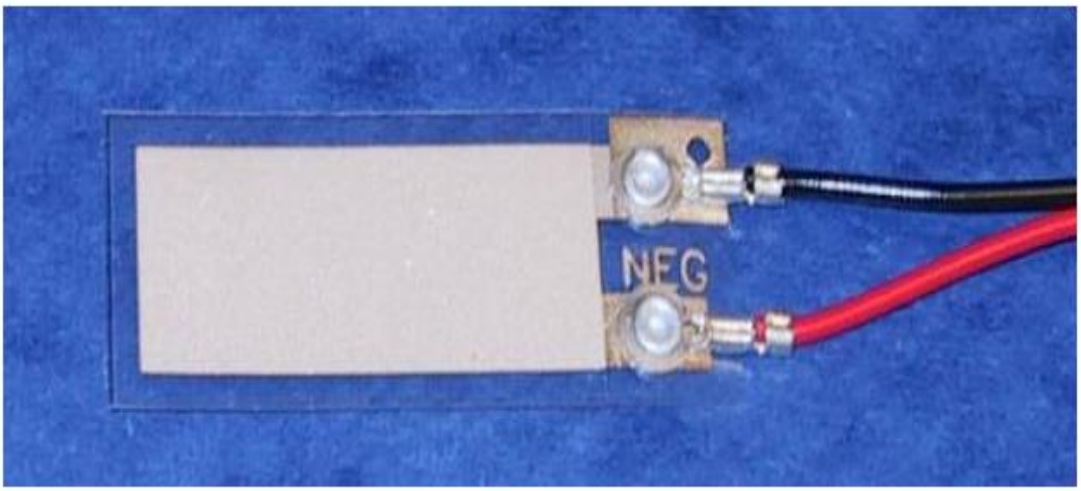


Fig 4: Piezoelectric Sensor (LDT1-028K) with lead attachment.

The Phantom

To demonstrate the ability of the photoacoustic imaging and provide one dimensional measurements of the object, several phantoms have been prepared of well-defined absorption

structure. A sandwich of chicken breast (test sample simulating normal tissue) and piece of liver (test sample simulating blood vessel or tumor).

The liver was embedded within the chicken breast with different thicknesses and depths as listed in table 2.

Table 2: Different depth, thickness and the distance to the detector for several phantoms

Phantom	Depth (cm)	Thickness (cm)	Distance to detector (cm)
No.1	0.2	0.3	0.8
No.2	1	0.4	1.2
No.3	0.5	0.65	0.8
No.4	0.8	0.8	0.3
No.5	0.7	0.9	1.1
No.6	0.8	1.3	0.6

The distance between the phantom and the detector was varied with changing the thickness and the depth of the liver. The phantom which is used in the present work is shown in figure 5.

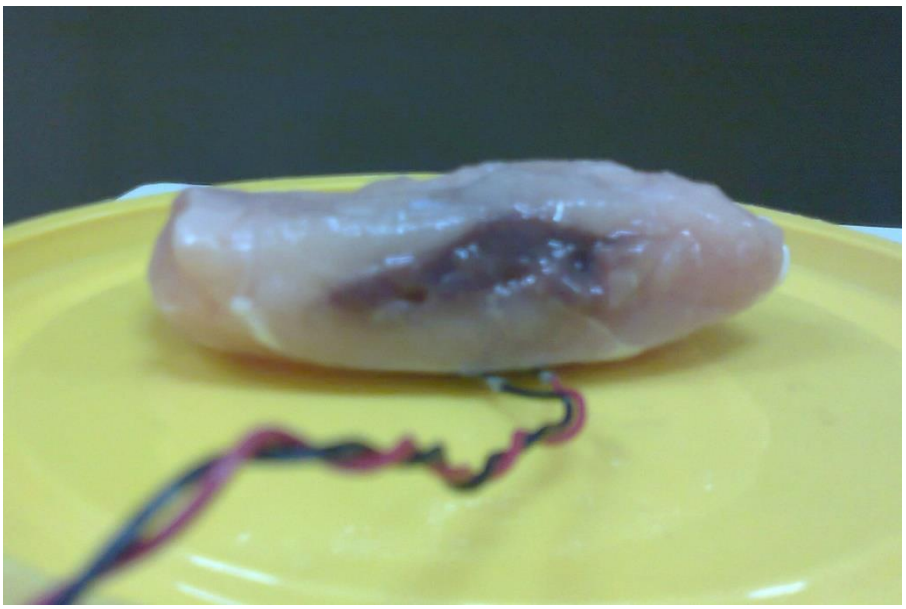


Figure 5: Phantom with liver inside it.

Storage Oscilloscope

Storage oscilloscope type PDS5022S (Portable Color Digital Storage Oscilloscope) has been used in this experiment to save the signal detected from the piezoelectric detector.

Photoacoustic Experimental Arrangement

In this arrangement, Nd:YAG laser with wavelength 1064 nm and different energies of 250, 400 and 500mJ was used to provide a photoacoustic wave which is consisted when the laser radiation is incident perpendicular to the bulk phantom and detector. The generated wave was detected using piezoelectric detector and the signal was displayed on the screen of the storage oscilloscope

3 Experimental Results

This part presents the experimental results of the phantom shown in table 2 signals collected from the PA experiments performed on the phantoms explained in chapter3. These signals are time-domain voltage amplitude detected by the piezoelectric transducer that represents the acoustic pressure waves induced by the phantom.

Figure 6, 7 and 8 show a photoacoustic voltage signal with time of phantom no.1 which has 0.2 cm depth of target, 0.3 cm thickness of target and 0.8 cm distance to the detector at 500, 400 and 250mJ pulse energy respectively.

The peak amplitude of positive photoacoustic signal is 180mV, 150mV and 110mV respectively, the thickness measured from the signal is 0.4125 cm, 0.45cm and 0.4875cm.

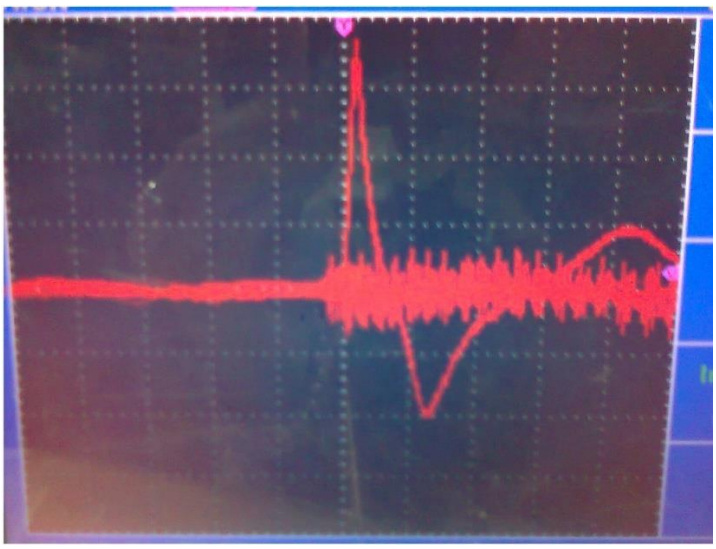


Fig 6: Voltage waveform resulting experimentally at 500mJ pulse energy of phantom no.1.

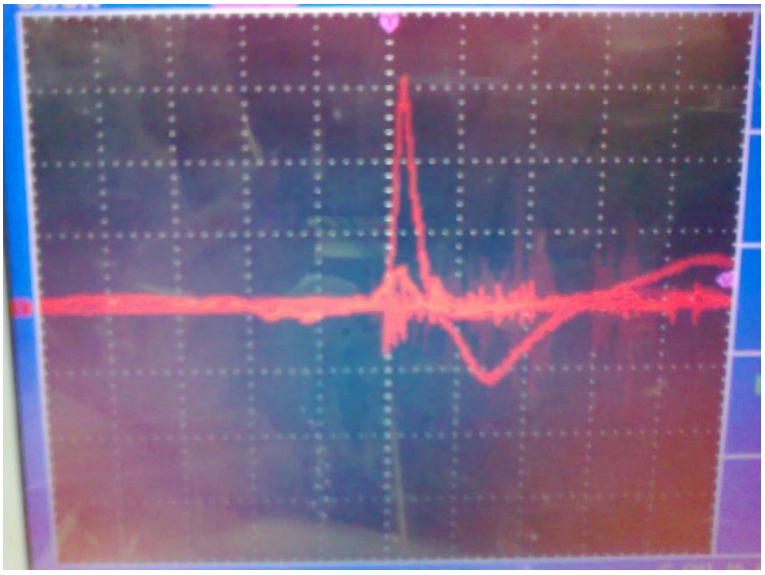


Fig 7: Voltage waveform resulting experimentally at 400mJ pulse energy of phantom no1.

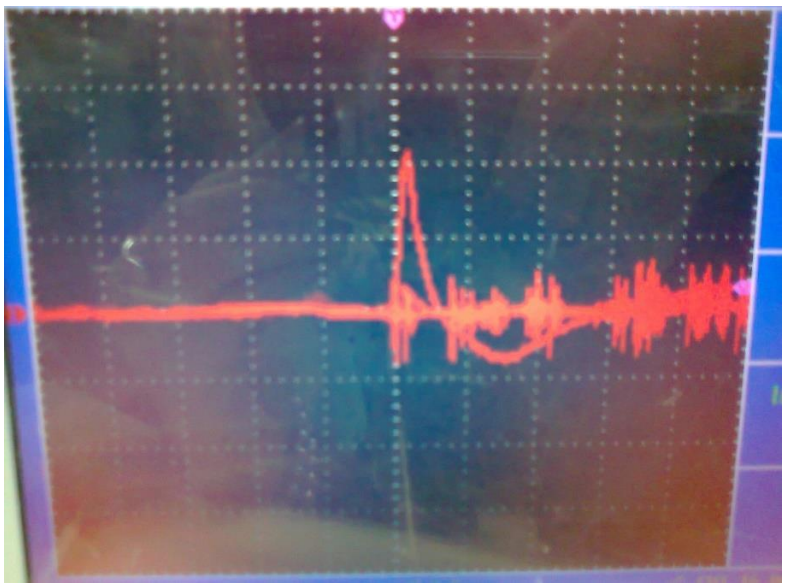


Fig 8: Voltage waveform resulting experimentally at 250mJ pulse energy of phantom no.1.

Figure 9, 10 and 11 show a photoacoustic voltage signal with time of phantom no.2 which has 1 cm depth of target, 0.4 cm thickness of target and 1.2 cm distance to the detector at 500, 400 and 250mJ pulse energy respectively.

The peak amplitude of positive photoacoustic signal is 200mV, 170mV and 160mV respectively, the thickness measured from the signal is 0.4875 cm and 0.56cm.

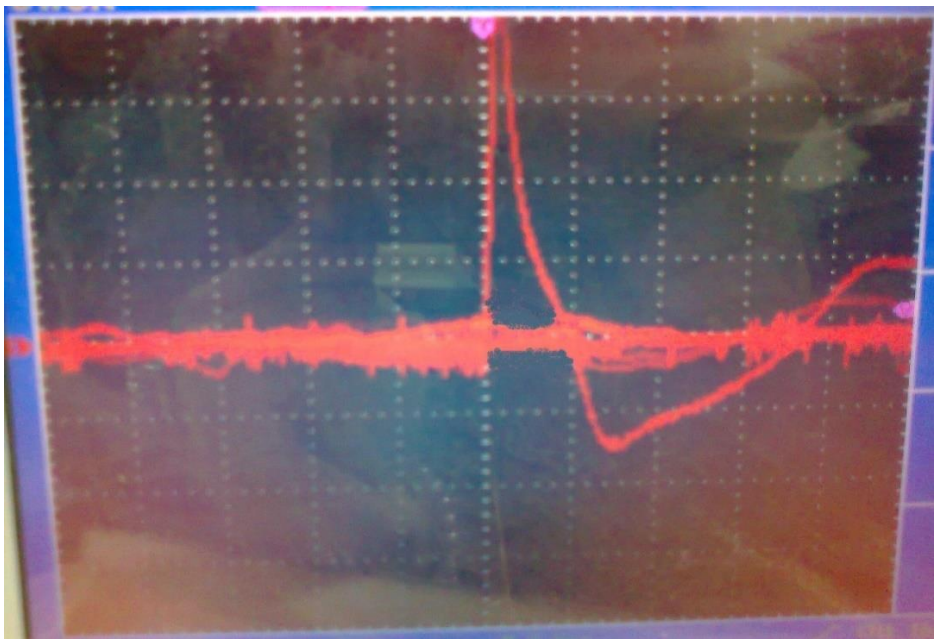


Fig 9: Voltage waveform resulting experimentally at 500mJ pulse energy of phantom no.2.

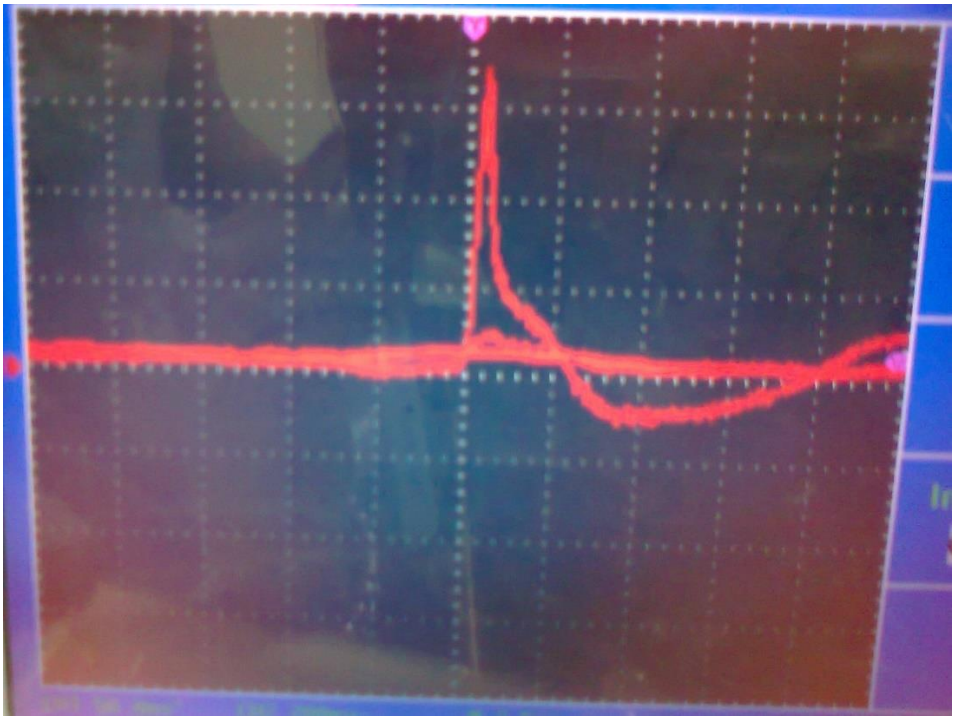


Fig 10: Voltage waveform resulting experimentally at 400mJ pulse energy.

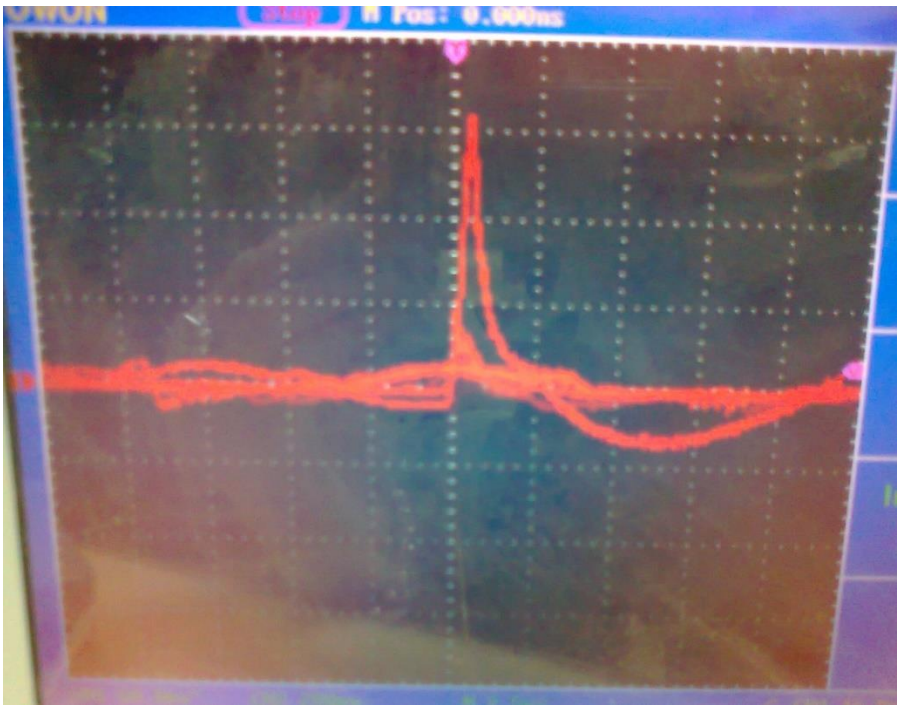


Fig 11: Voltage waveform resulting experimentally at 250mJ pulse energy.

Figure 12, 13 and 14 show a photoacoustic voltage signal with time of phantom no.3 which has 0.5 cm depth of target, 0.65 cm diameter of target and 0.8 cm distance to the detector at 500, 400 and 250mJ pulse energy respectively.

The peak amplitude of positive photoacoustic signal is 200mV, 110mV and 90mV respectively, the thickness measured from the signal is 0.45 cm, 0.675cm and 0.45cm.

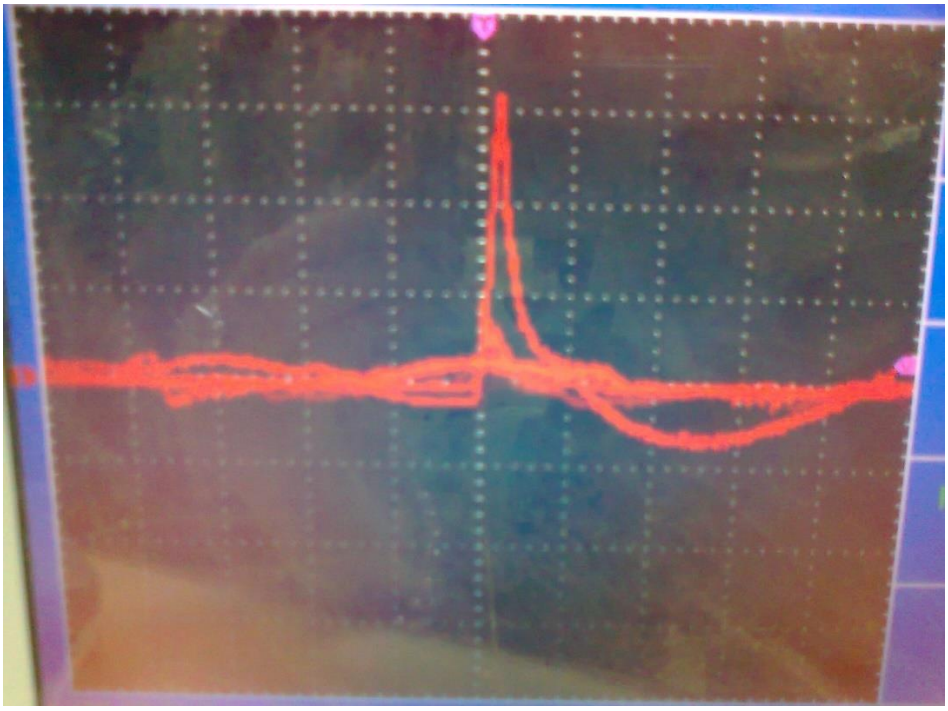


Fig 12: Voltage waveform resulting experimentally at 500mJ pulse energy.

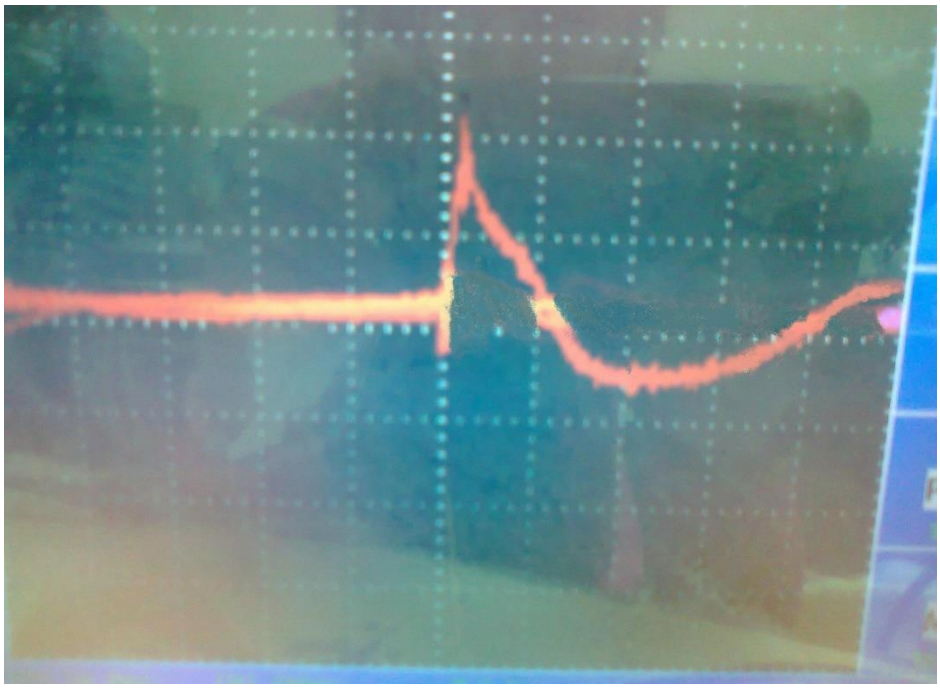


Fig 13: Voltage waveform resulting experimentally at 400mJ pulse energy.

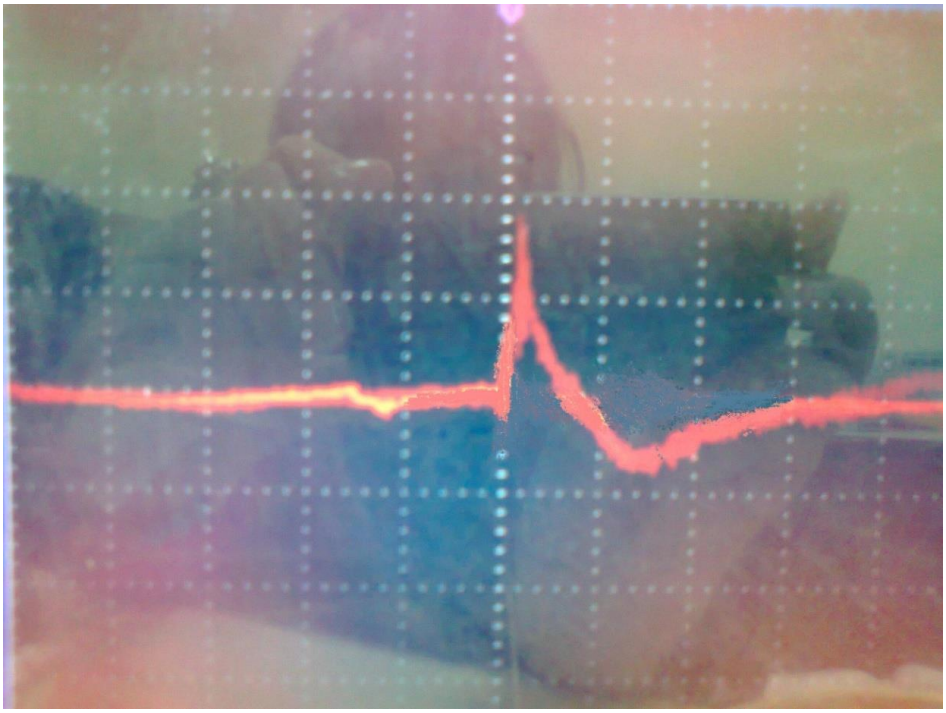


Fig 13: Voltage waveform resulting experimentally at 250mJ pulse energy.

Figure 14, 15 and 16 show a photoacoustic voltage signal with time of phantom no.4 which has 0.8 cm depth of target, 0.8 cm thickness of target and 0.3 cm distance to the detector at 500, 400 and 250mJ pulse energy respectively.

The peak amplitude of positive photoacoustic signal is 120mV, 90mV and 70mV respectively, the thickness measured from the signal is 0.675 cm, 0.405cm and 0.45cm.

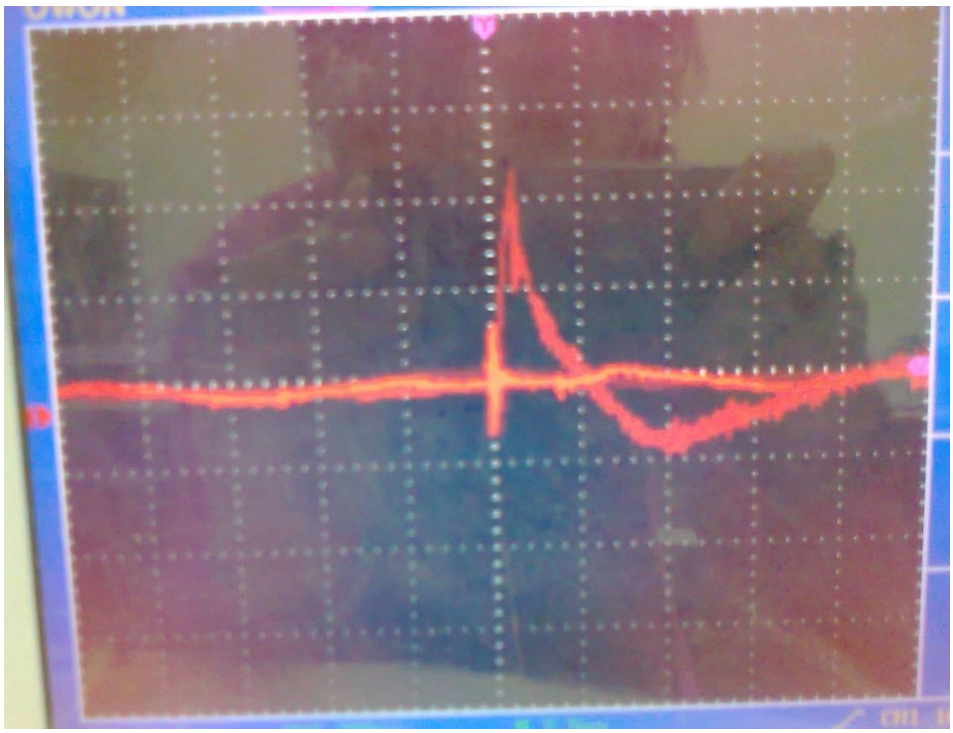


Fig 14: Voltage waveform resulting experimentally at 500mJ pulse energy.

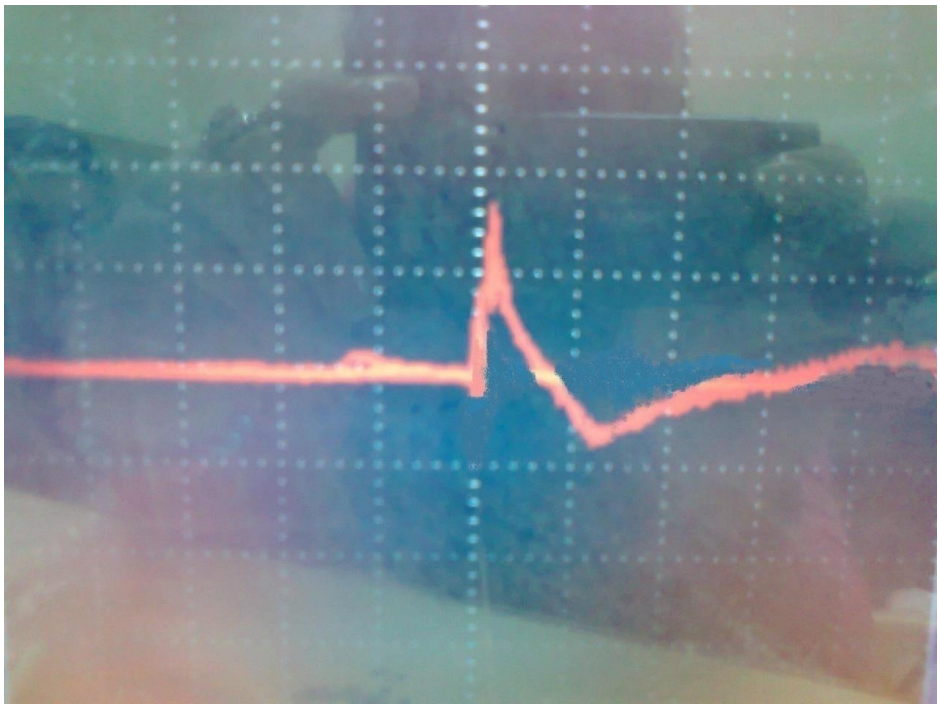


Fig 15: Voltage waveform resulting experimentally at 400mJ pulse energy.

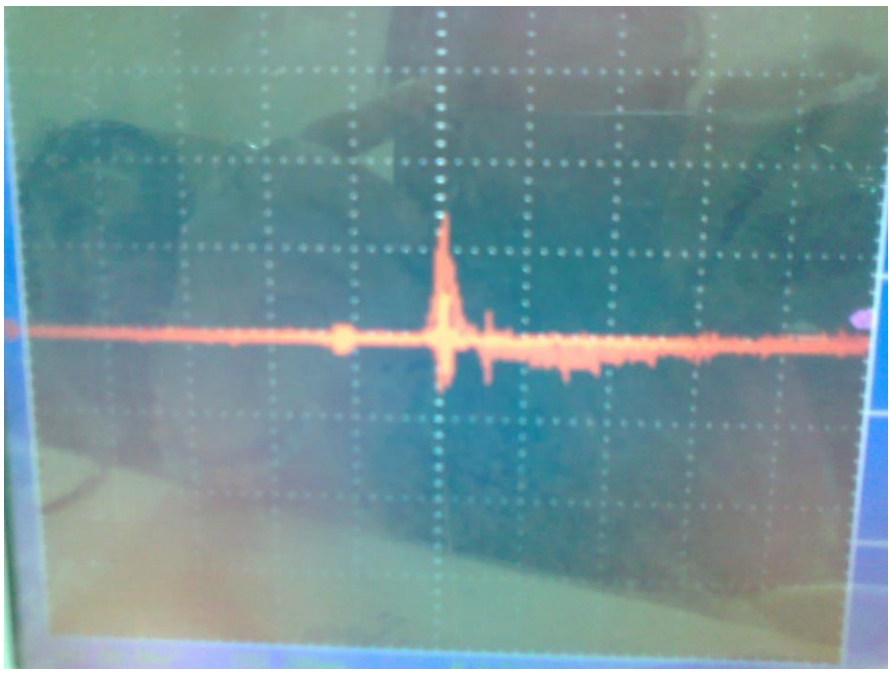


Fig 16: Voltage waveform resulting experimentally at 250mJ pulse energy.

Figure 17 and 18 shows a photoacoustic voltage signal with time of phantom no.5 which has 0.7cm depth of target, 0.9 cm diameter of target and 1.1 cm distance to the detector at 500 and 400mJ pulse energy respectively.

The peak amplitude of positive photoacoustic signal is 250mV and 180mV respectively, the thickness measured from the signal is 0.375 cm and 0.4125cm.

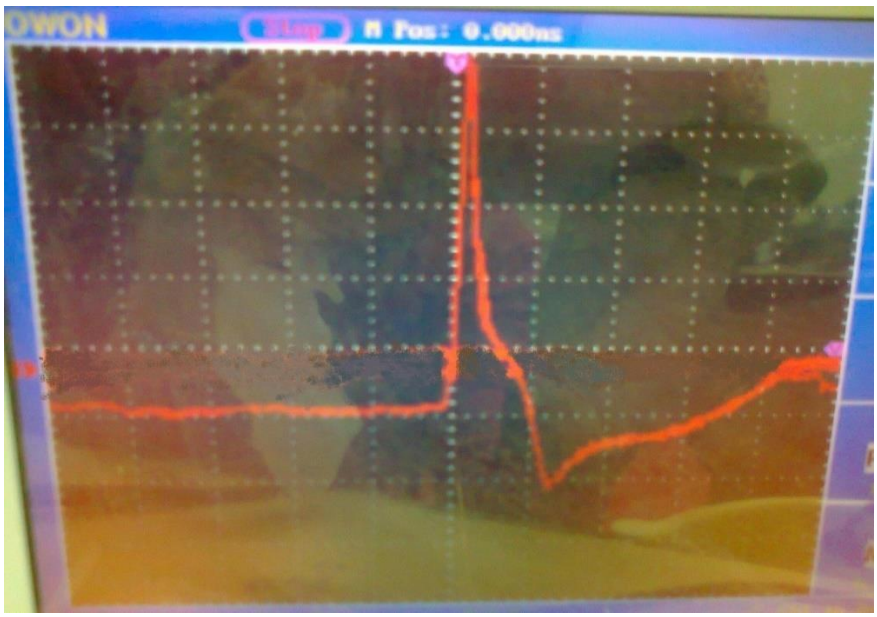


Fig17: Voltage waveform resulting experimentally at 500mJ pulse energy.

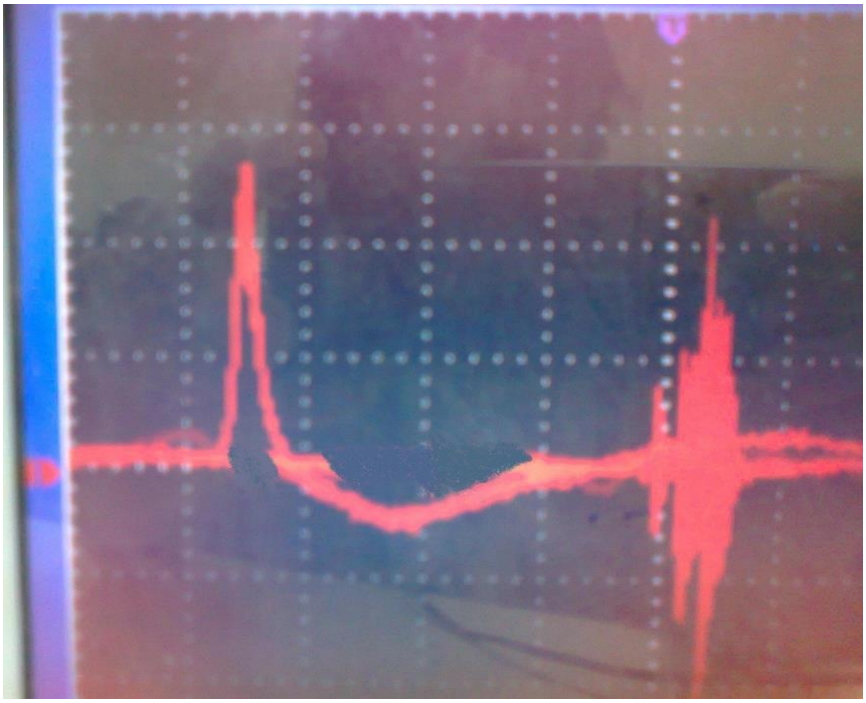


Fig18: Voltage waveform resulting experimentally at 400mJ pulse energy.

Figures 19, 20 and 21 show a photoacoustic voltage signal with time of phantom no.6 which has 0.8 cm depth of target, 1.3 cm thickness of target and 0.6 cm distance to the detector at 500, 400 and 250mJ pulse energy respectively.

The peak amplitude of positive photoacoustic signal is 170mV, 140mV and 120mV respectively, the thickness measured from the signal is 0.5625 cm, 0.5625cm and 0.525cm.

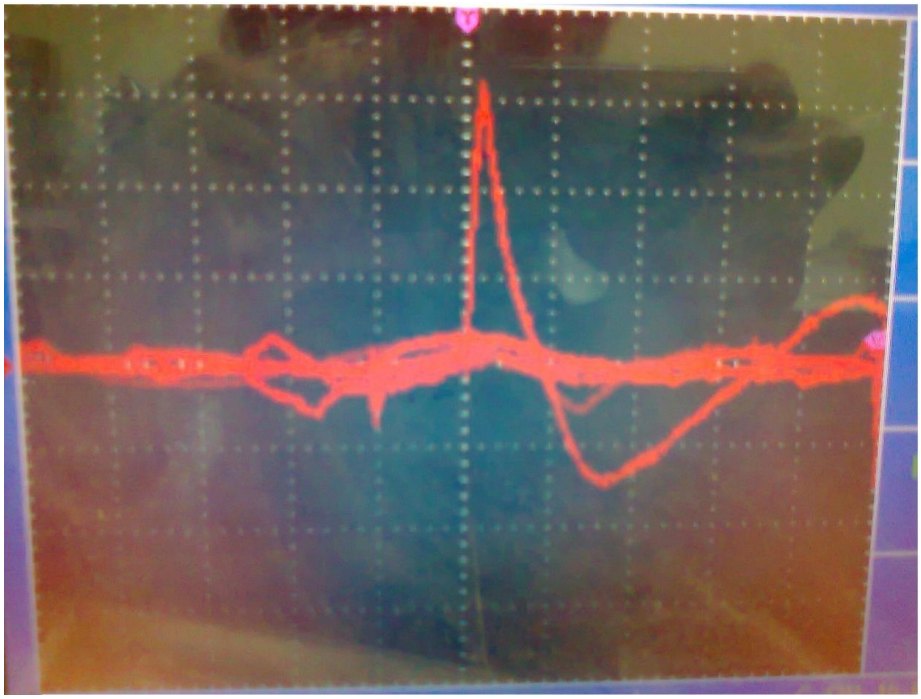


Fig 19: Voltage waveform resulting experimentally at 500 mJ pulse energy.

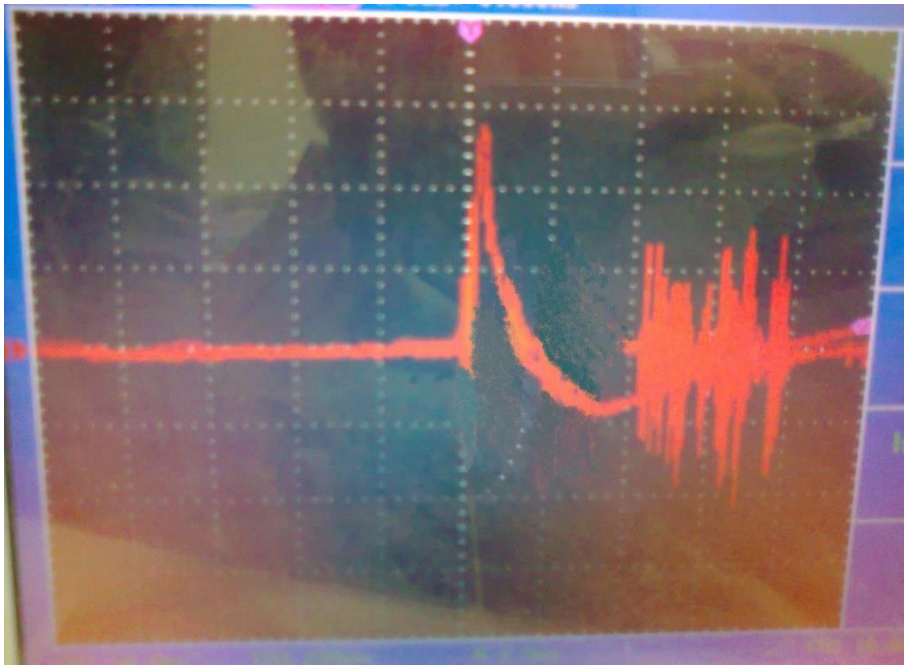


Fig 20: Voltage waveform resulting experimentally at 400mJ pulse energy.

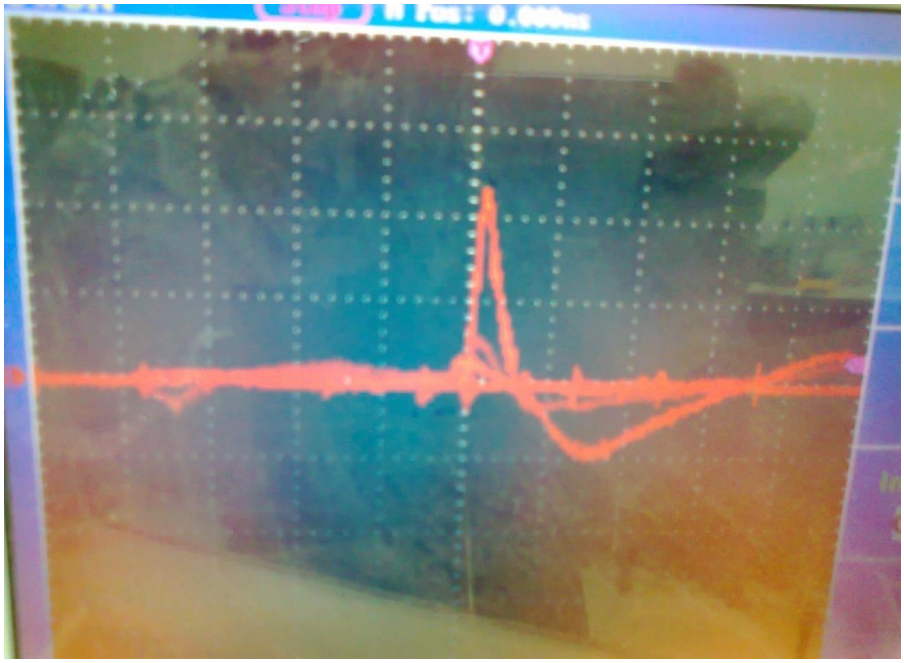


Fig 21: Voltage waveform resulting experimentally at 250mJ pulse energy.

The amplitude of photoacoustic signal is linearly proportional to the laser pulse energy so the amplitude of signal has reduced when the pulse energy has decreased as shown in the results above.

The positive peak of the experimental results represents the absorption properties of the target and its amplitude is much larger than the negative peak which represents the acoustic wave generated from a short laser pulse which is must be shorter than the stress confinement time.

The second small positive peak in the figures represents reflection of acoustic wave by the walls of sample and it can be considered as noise. There are many noise sources which may affect the photoacoustic signal, the vibration of Nd:YAG laser, the noise from a digital oscilloscope and the noise comes from the coupling of the wire between the piezoelectric detector and the oscilloscope.

The photoacoustic signal amplitude as a function of depth can be used to measure the light attenuation, the effective attenuation coefficient of chicken breast tissue was fitted to 1.2 cm^{-1} based on Beer-Lambert's law. This light attenuation in chicken breast muscle is similar to that in some human tissues.

A long wavelength allows light to penetrate deeply with less attenuation, but it is not absorbed as great a degree as light with a shorter wavelength.

Conclusions

The acoustic signal in the experimental results has a larger amplitude for the positive peaks than the negative peaks. The time between the maximum positive peak and minimum negative peak is measured as the one-dimensional size of the liver piece. No consistent correlation has been done between the actual thickness and the measured from the experiment results for many reasons as that could be related to use a slab sample not a spherical one which is difficult to build and coupling in the breast chicken tissue, the coupling of the liver piece with the surrounding tissue, the coupling of the piezoelectric detector with the phantom, noise from the surrounding tissue (breast chicken) and other noise from laser, detector and oscilloscope.

Reference

[1] Y. SU, C. Hofer, and S. S. Gambhir, "A Photoacoustic tomography system or imaging of biological tissues", J.Phys. D:Appl, Phys. 38 2640-2644.
[2] M. Xu and L. V. Wang, "Photoacoustic Imaging in biomedicine", review of scientific instruments, 2007.
[3] M.M.Jawad, S. T,.AbdlQader, A.A.Zadian, B.B.Zadia, A.W. Naji, "An Overview of Laser Principle, Laser-Tissue Interaction Mechanisms and Laser Safety Precautions for Medical Laser Users",International Journal of Pharmacology 7(2); 149-160, 2011.
[4] Internet web. www.amyshah.com
[5] J. Beun, "Physical Principles of General and Vascular Sonography"
[6] G.J.Diebold, T.Sun and M.I.Khan, "Photoacosutic Waveforms Generated by Fluid Bodies", Springer-Verlag, Berlin, Heidlberg, Pp.263-269, 1992.
[7]"Nd:YAG Laser" HuefieTongda technology-Diamond-288 pattern EPLS, www.iplsupply.com/3a-YAG-laser.html
[8] "Piezo Film Sensor Technical Manual" Measurement Specialties, Inc., www.meas-spec.com

Modern Methods and Technology in University Engineering Education Using Computer Programs

Nabeh Natic Abdullatif Alderoubi[1], Humam Kareem Jalghaf[2], Salman H. Omran[3], Ahmed Sadeq Yousif[2], and Laith Jaafer Habeeb[2]

[1]United States / Southeast Community College - Computer Aided Design and Drafting

[2]University of Technology / Mechanical Engineering Department

[3]Ministry of Higher Education and Scientific Research

Abstract

In the Mechanical Engineering Department / University of Technology, it was completed the final preparations and procedures for teaching, evaluation and examination without the use of any paper either in the lectures or in the exams for CAD and CAE using the SolidWorks and ANSYS programs for the third and fourth stages respectively. This type of learning was called Paperless Learning. The computer is only used in teaching and evaluation of students and conducting all the term and final exams for the purpose of involving the student in the application of the material interactively and thus applying the theoretical material practically on the computer, which enables the student to complete the simulation to solve all engineering issues using the computer. The results show that students have shown interest in these materials. The use of the paperless learning method enables students to develop professional skills in computer design and engineering software, as well as enhance their capabilities in other programs and computers to suit the requirements of the labor market. This method can be applied to any other material that contains programming or computer use such as computer engineering, science or any computer-aided material.

الخلاصة

تم في قسم الهندسة الميكانيكية / الجامعة التكنولوجية الانتهاء من التحضيرات والاجراءات النهائية للتدريس والتقييم (CAD) والامتحان بدون استخدام اي ورقة سواءاً في المحاضرات او في الامتحانات لمادتي التصميم بإستخدام الحاسوب للمرحلتين الثالثة والرابعة على ANSYS والـ SolidWorks بإستخدام برنامجي الـ (CAE) والهندسة بإستخدام الحاسوب اي يتم استخدام الحاسوب (Paperless Learning). التوالي. حيث أُطلق على هذا النوع من التعلّم اسم التعلم بدون ورق فقط في التدريس وتقييم الطلاب واجراء كافة الامتحانات الفصلية والنهائية لغرض اشراك الطالب في تطبيق المادة بشكل

تفاعلي وبالتالي تطبيق المادة النظرية عملياً على الحاسبة ممّا يُمكّن الطالب من انجاز عملية المحاكاة لحل كافة المسائل الهندسية بإستخدام الحاسبة. وتظهر النتائج أن الطلبة قد اظهروا اهتمامهم بهذه المواد وكذلك فإن استخدام طريقة التعلم بدون ورق تمكن الطلاب من تطوير المهارات المهنية في برمجيات التصميم والهندسة بإستخدام الحاسوب، فضلاً عن تعزيز امكانياتهم في البرامج واجهزة الحاسبة الأخرى وبما يلائم متطلبات سوق العمل. يمكن تطبيق هذه الطريقة على أي مواد أخرى تحتوي على البرمجة أو استخدام الكمبيوتر مثل مواد هندسة الحاسوب أو العلوم أو أي مادة بمساعدة الحاسوب.

1. Introduction

Technology has become an important educational tool, and the competition among technology companies to provide advanced materials and learning devices has assisted university professors and lecturers in enhancing the educational process. To date, the world is undergoing enormous changes in various fields, including in education in which the conventional methods of learning and teaching are going through a transitional stage. For example, the US, the UK, and several European countries have already made huge advancements in the field of e-learning [1]. Several educational institutions in the Middle East, such as Isra University, Petra University, and Hashimate University in Jordan, are also offering e-learning classes [2]. Evidently, many calls have been made to pursue learning without using papers and without the actual presence of teachers and physical libraries. Consequently, these calls make us rethink about the teaching methods which use, as well as enriching these methods, by using novel and beneficial resources for a promising and improved learning process. This endeavor may result in immense developments and benefits for learners, as well as the environment by reducing the use of papers (thereby preventing deforestation) [3]. Evidently, the ease of using computers and printers in printing various types of documents has resulted in the massive use of paper. Thus, creating a paperless environment requires electronic documentation through a word processing document or a digital image, and submitting or uploading it directly either in its original form or in the previously printed form [4]. Meanwhile, education development is one of the important goals of developed countries [1] with a particular focus on developing a well-educated generation. Thus, these countries utilize their human and financial resources in prioritizing and pursuing human development through high-quality education. Consequently, achieving high standards in education necessitates altering the conventional learning process through interactive, learning-based methods using personal computers (PCs) or tablets [5, 6]. At the University of Technology, a novel paperless learning

(PLL) method is introduced to replace the conventional method of learning. The PLL method is applied successfully in the CAD and CAE classes of third and fourth year students, respectively.

2. Reasons

Through teaching more than one subject during the past 20 years, it was noticed that the level of the student began to decrease and their interest in the curricula materials decreased and their understanding of the scientific subjects became weak and did not rise to the required level because of many things including the general situation of the country and the abnormal conditions that passed in Iraq as well as some old teaching methods, education, training or learning, which are incompatible with the great scientific development that has occurred in the basic education sectors and university education. Thus, it has noticed that the level of student understanding, the extent of its application and its efficiency in the use of computer, software and computer aided engineering materials has become very low. Therefore the group - a group of professors - and with the help of the department and the university to develop a five-year plan to modify the curricula in these materials as well as the mechanism and methods of learning and the preparation of new lectures in line with the global evolution of these materials and arranged a lot of logistics to achieve this purpose that is increasing the interest of the student to study more and give the scientific material to him better and create the appropriate atmosphere for him to create again.

3. Details of the Method

The details of paperless learning are explained as follows:

1- The staff of teachers is fixed for each material and decide by the senior of the material at a rate of one for every eight students to answer questions and queries of students and training them to use engineering programs. As teaching these materials using computers requires more than one teacher of the division because the teaching of CAD and CAE requires the involvement of a group of teachers to provide the opportunity to answer questions to students equally based on the total number of students in the division, ensuring the access of the information and knowledge to all students at an equal time depending on the type of lecture delivered, whether class or tutorial or quiz, etc., as well as their

discussion in assignments and projects and give them grades and evaluation in each lecture. Where the teacher of the material presents the lecture on the smart board using the computer or (tablet) and then answer the questions of students and discussed with the help of staff of professors with specialization where the distribution of students in the form of groups and each group of students have one professor.

2- The duties of each teacher are determined in advance. The senior will give the lecture in the first half hour of the lecture. All the teachers in the next hour will answer the students' questions and solve all the problems in their weekly classroom duties or computer problems related to the various programs used by the student during the course - SolidWorks, ANSYS, image or text files browsing programs - or problems in the hardware of the computers. The last half hour of the lecture is intended to receive the weekly homework and evaluate it on the computer and give the degree at same time after reviewing the files and date, as the student is documenting everything that was done on the computer in a manner of film recorded all the steps of the exercise and save in a movie file to eliminate the cheating phenomenon and cloning duties among students.
3- The duties of teaching staff are not limited to answer and discuss students only but include modification and correction of the program or project presented by the student, whether H.W. or seminar or project interactively on the computer to solve the problems faced by the student during his/her work and assess his/her performance.
4- Communicate through the media of social communication -Facebook-: The notifications and updates on the grades, absences and quick answer to the inquiries of students in their homes are completing their duties or preparation for the next lecture will be on Facebook where the staff of the department open a special account of these two materials, it is similar to that in the western universities of the Black Board.
5- Each student saves the homework and classwork on his/her computer in the laboratory, which has its number in the list of attendance list, on the hard disk (D) in a folder bearing the number and the name of the student and be within the folder of the stage and specialization and the division in English, and this is the context in the quiz and examination too.
6- Homework & Classwork, called laboratory activity, is collected and divided on the number of duties and delivered every week. The delay is not accepted except with a formal excuse.

Without an excuse, the compensated duty is assessed from 50% instead of 100%. It has an additional duty set by the supervising professor and is usually from previous years' questions or chooses duties specially prepared for the late ones.

7- Students are taught how to use the computer screen by dividing it into more than one window, starting with two main windows one for the main program and one for the homework or text file. The student uses one of these two windows to insert the picture of the lecture or exam question (instead of the paper) at the same time, the question is solved using the second window of the program on the same screen. A large number of windows can be used if the question or exercise contains many parts, such as a 10-part machine.

8- Exams and sudden examinations (quizzes) are done in the same context for continuous evaluation and the final exam is done on computers only, and in the case of secret exam, student saves solution files which include design and drawing files or design, drawing and simulation files, in the same form of duties and in the same method of saving except that the student information (name and division. etc.) can only be on the main folder without the subfolders and save the name of the Project or Exam sub files. Where the examination committee save these answers and copping it from computers to external hard and then the members of the examination committee give a code for each student and change the name of the main folder and then give these confidential answers to the professors of the material to check the answers and deliver grades to the committee again and then members examination committee return to the secret code and restore the student's degree.

9- The grades are presented on the Facebook site of the two materials. The objections are received on the grade and the students' observations and the objections on the grade are recorded. The senior study the solution and the degree immediately after the exams and at the end of the first lecture after the exam - in the case of the exam is not secret - or after handing over the files of the objectors by examinations committee after exams - in case of secret exam.

10- It is customary to grades and assessment of classwork or homework or projects putted by the teacher to the student and be final, but one thing unique to this updated method also is that the student puts the degree for himself/herself and all according to his/her effort. In the event that the student wishes to obtain a high degree and did not receive it because of the delay in the delivery of the duty or absence without a legitimate excuse or handing us

a duty containing shortages or was unable to answer the questions of teaching staff fully or for other reasons, he can improve or maximize the degree by donning additional one or more duty at any time or during the holidays to improve the degree and this opportunity will be granted to all students within the conditions determined by the teaching staff. In this way it is easy to ensure that the student does not waste his/her extra time or repeated holidays, but rather learns and improves himself in the use of programs and improve his/her grades. And there are many exercises and duties prepared for this purpose in advance and be more difficult than the original duty - as mentioned - and the degree to be a few for each additional exercise in order to depend on the student to deliver duties (classwork/homework) or understand the material on time, and that delay or compelling reason is an exceptional circumstance and not a general case.

11- For the final projects of both materials, the student chooses between taking one or more of the projects that are presented - in the first lecture at the beginning of the year - and be from global sites in these disciplines or to search for a project in which a more complex engineering problem on the internet to ensure that the degree is higher in evaluation or devising a problem and finding solutions to it. In these projects, the group encourage students to solve more realistic issues whose subjects are the same engineering subjects taught by students (mechanics, materials, fluids, heat transfer, vibrations, etc.) and they think about them and find solutions, thus stimulating the student to search for information and harness to solve the problem by retrieving all the information gained during his/her study.

4. Results and Discussion

1- Some of the most important results are that the students have shown their interest and passion in these subjects and some of them are coming to the teaching staff and ask them to give the subject of the forth year so they can solve the duties and learn the material for the coming year in the summer vacation. It have also noticed a remarkable improvement in the level of students and their understanding of many things and topics as they not only study them but apply them and research in the details that increase their capabilities and thus review and see the results of science and the work of animation for engineering cases

are closer to understanding and more entrenched in the imagination of the student form listening or paper solution of equations.

2- The staff develop the lectures always according to what was obtained and observed during the previous year or the reactions of students (feedback) and thus the curriculum is a constant development even during the one year to accommodate the development of these materials in the developed western universities.

3- Things that have also been observed for this type of teaching - through our experience as teachers and teaching more than a material - that the control of teaching on the lecture is much easier and that students rioters - if they find - cannot talk about things that do not belong to the lecture or do not pay attention to teaching or lecture or preoccupied with other things do not belong to the material, because any busyness for the student, even for a limited seconds causes him to lose certain information or important step and this in turn reduces the performance and thus the degree of classwork. It has find that the students are very integrated with the lecture and for the whole two hours and this fact did not find previously and even some students pre-teaching to solve some things and this confirms that students are studying the material before entering the new lecture because the group give students all the lectures at the beginning of the academic year.

4- The results obtained also that students learn to use the computer and solve problems (Hardware & Software) better than they were in the past because of the need to install and use the computer and its general and engineering programs.

5- Most students deliver tasks and projects quickly and ahead of time due to them because the project that delivers quickly is a higher degree, and even some students began to pursue the staff of the subjects in the squares and corridors and during their exit after the end of working hours to ask to solve some of the problems facing them in the design and simulation engineering issues, some of which are somewhat complex and above the level of primary studies students, but insist on delivering such projects to ensure a high degree.

6- Also, one of the important results that the students cannot cheat even one of them looked at the computer of the other for many logical reasons, including: the different scientific possibilities of students and their thinking where to start and any part of the system and any orders and any view begin drawing and design and simulation, speed of use the

computer , mouse, keyboard and even the speed of computers to execute different commands because the computers are slightly different from each other and so on.

7- Figures (1) and (2) show some of the work of the students in the use of software design using computer and computer programming as a final project image without mentioning the exact scientific details or parts constituent of these systems and projects. Figure (1) shows the student outputs in CAD software (SolidWorks). Figure (2) shows the student outputs in CAE software (ANSYS Workbench).

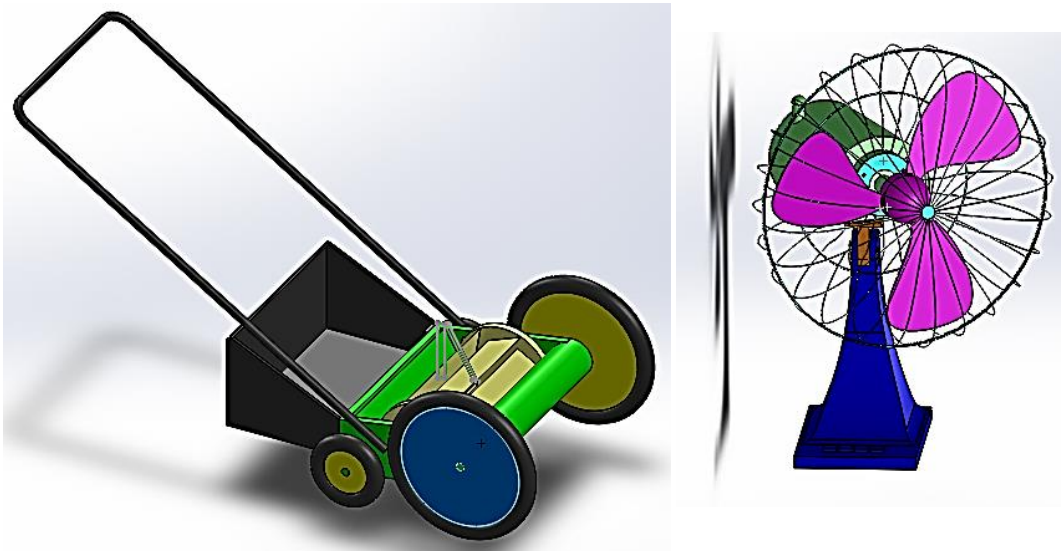


Figure (1) student outputs in CAD software (SolidWorks).

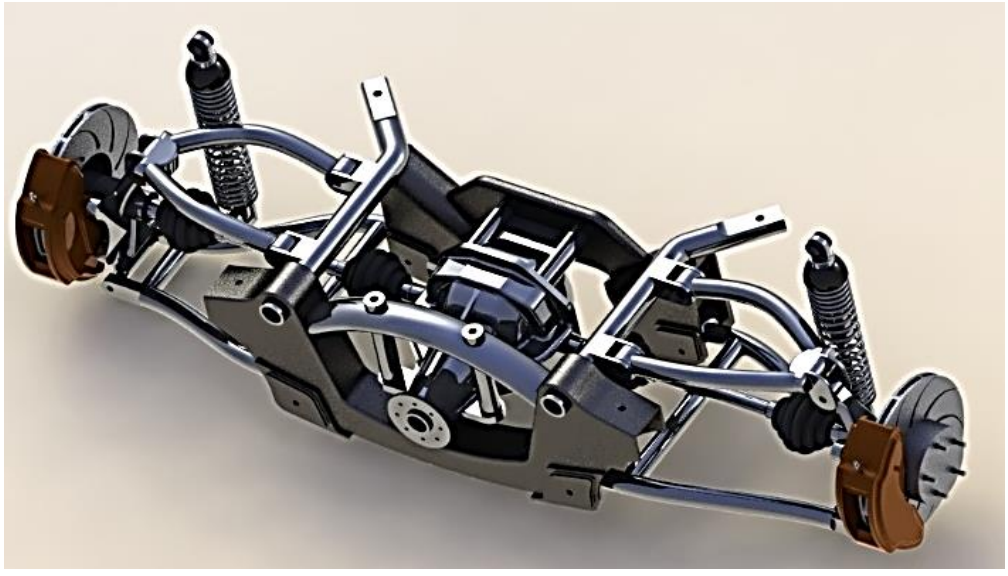
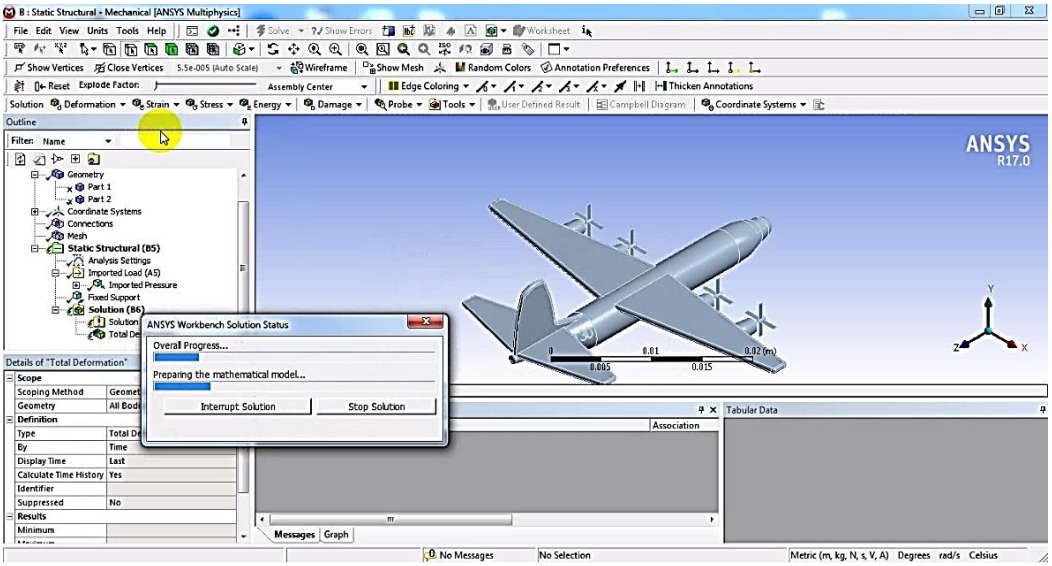


Figure (1) Continue.



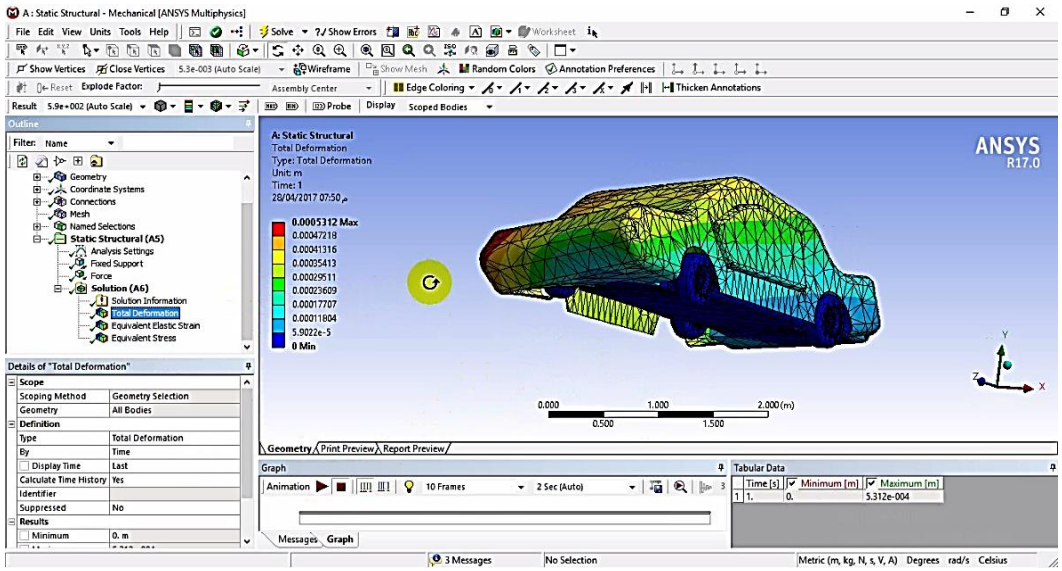


Figure (2) student outputs in CAE software (ANSYS Workbench).

5. Student survey

The following table and chart present a student questionnaire assessment that lists students' answers as percentages (yes, neutral, and no) of selected evaluation elements for a random sample. The survey shows that students tend to be positive in evaluating

materials and professors and understand lectures greatly and are disciplined to a large extent during the lecture.

Table (1): The views of the mechanical engineering student's questionnaire on the curriculum and teaching methods in CAD.

No.	Item	Yes %	Neutral %	No %
1.	The student is provided with a class syllabus (content, goals, references, and grade distribution).	100	0	0
2.	The materials are beneficial and meaningful.	94	6	0
3.	The professor attends the lectures during the specified lecture time.	96	4	0
4.	The professor marks the absent students for each class.	94	6	0
5.	The professor encourages the students to practice college traditions and directs them to engage in various activities.	78	18	4
6.	The professor has various scientific approaches in evaluating the students.	84	4	12
7.	The professor has the ability to manage the class and encourage students to learn the practical side of the scientific materials.	88	4	8
8.	The students can substantially understand the scientific materials.	76	18	6
9.	The content of the class materials is acceptable.	90	10	0
10.	Practical applications and follow-ups are connected with the class materials.	84	16	0

11.	The lectures start and end according to the academic calendar.	88	12	0
12.	The professor provides illustrative examples during the lecture sessions.	84	12	4
13.	The professor's voice can be heard and clear.	66	20	14
14.	The teaching language used is understandable.	82	18	0
15.	Students are following up with the professor.	72	28	0
16.	The professor agrees to meet students outside the lecture time.	80	20	0
17.	The professor assigns homework and tests outside the lecture time.	80	16	4
18.	The professor returns homework and tests to students after grading them.	72	12	16
19.	The professor compensates for missed lectures in case of his/her absences.	72	18	10
20.	The professor uses a variety of teaching methods to deliver the scientific materials to the students.	60	30	10
21.	Textbooks and references are available in the library.	56	24	20
22.	References are available outside the library.	64	20	16
23.	The students use the library to read lecture references.	44	12	44
24.	The students use the library to read the material thumbnails only.	52	20	28
25.	The class professor is highly qualified to teach this subject.	74	26	0

26.	The professor is committed to teaching the class subject.	96	4	0
27.	The professor has acceptable and positive personality, decent looking, and cooperate with others.	92	8	0
28.	The professor exudes honesty and integrity inside the campus.	94	6	0

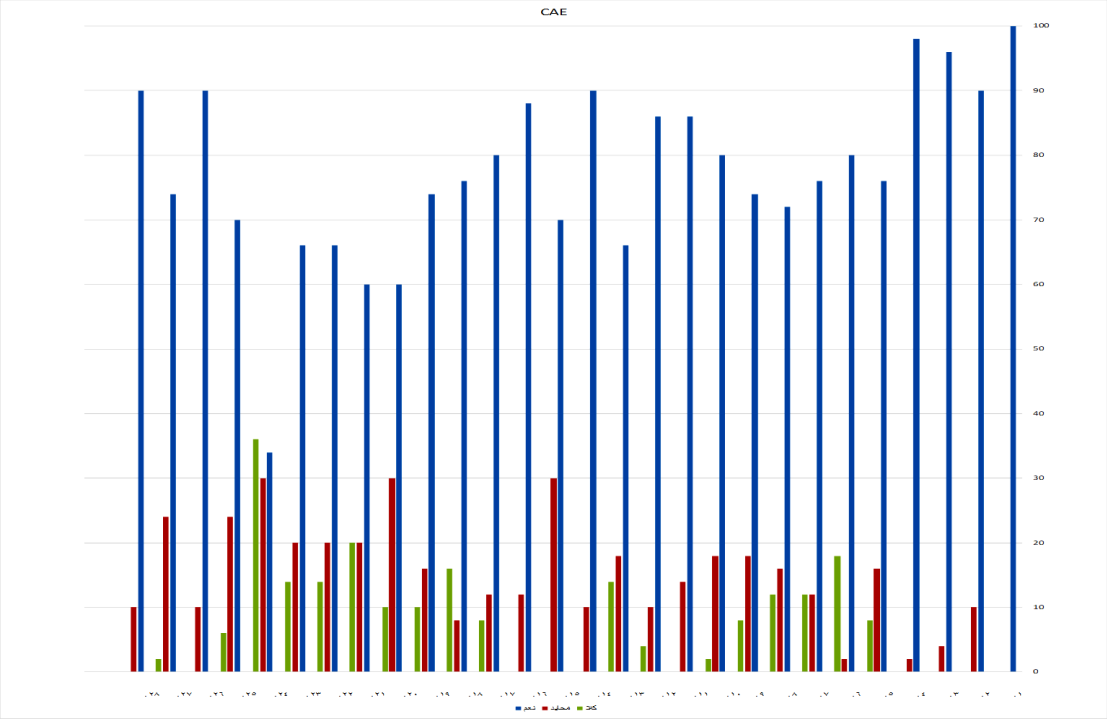


Figure (3) The views of the mechanical engineering students' questionnaire on the curricula and methods of teaching in CAE.

6. Conclusions

The following are the main conclusions drawn from this novel approach:

1- Students showed interest in materials that are of a distinctive nature and also the procedures that followed.

2- Using the paperless learning method enables students to develop professional skills in computer design and engineering software, as well as enhance their capabilities in other

programs and computers to suit the requirements of the labor market. It has been noted that companies and the labor market need students who are proficient in computers, and simulation, and this has made it clear to some students after graduation that the preference to appoint graduates who are good at computer and design.

3- This method can be applied to any other materials that contain programming or computer use such as computer engineering, science or any computer-aided material.

4- Students learn how to connect many scientific topics interactively.

5- In the survey, students tend to be positive in the evaluation of materials and teachers, as this behavior was not observed in other material as far as it knows.

References

[1] Michelle Ramim and Yair Levy, "Securing e-learning systems: A case of insider cyber attacks and novice IT management in a small university", Journal of Cases on Information Technology (JCIT), 2006. 8(4): p. 24-34.

[2] Shadi Aljawarneh, Zahraa Muhsin, Ayman Nsour, Faisal Alkhateeb, Eslam AlMaghayreh, "E-learning Tools and Technologies in Education: A Perspective", E-learning, 2010.

[3] Jason Tan, Mihailo Veljovic, Jessica Tom, Jordan Stewart, "Paper or Paperless? A Point-Counterpoint Student's Perspective", UBC PSSJ, 2014. 1(1).

[4] Reaz, M., Hussain, S. and Shajib Khadem, "Multimedia university: A paperless environment to take the challenges for the 21st century", AACE Journal, 2007. 15(3): p. 289-314.

[5] Peter Osmon, "Paperless classrooms: a networked Tablet PC in front of every child", Proceedings of the British Society for Research into Learning Mathematics, 2011. 31(2): p. 55-60.

[6] Rossman, Parker, "The Emerging Worldwide Electronic University: Information Age Global Higher Education". Contributions to the Study of Education, Number 57, Greenwood Publishing Group, 88 Post Road West, Box 5007, Westport, CT 06881, 1992.

Design an Acceleration Fuzzy Logic Controller (AFLC) for Speed DC Servo Motor

Assist. Prof. Muna Hadi Saleh
Baghdad University/Engineering Collage/Electrical Engineering Department
drmuna4@yahoo.com

Abstract:

Fuzzy logic is well suited to the control of non-linear, time varying and /or any system for which it is difficult to obtain an exact mathematical model. To design a good fuzzy controller, through understanding of the desired process behavior is required. Many papers have designed fuzzy algorithm by using the deviation from the set point and its rate of change. In this research the variables can be replaced to a change in rate change in error term (CCE) and the effect of acceleration on the system performance has been considered. The quantitative and qualitative study of Acceleration Fuzzy Logic Controller (AFLC) is made. The result showed there is an efficient method for speed control of a separately excited D.C motor by using this new controller.

Keywords: Generalized Modus Ponens, Acceleration Fuzzy Logic Controller (AFLC), DC Servo motor.

تصميم مسيطر المنطق المضبب المعجل للسيطرة على سرعة محرك مؤازر

أ.م.د. منى هادي صالح
جامعة بغداد/ كلية الهندسة / قسم الهندسة الكهربائية

الخلاصة

يعتبر المنطق الضبابي مناسبا جدا للسيطرة على المنظومات الغير خطية والتي يصعب الحصول على انموذج رياضي لها. من اجل تصميم جيد لمسيطر مضبب لتصميم وحدة تحكم غامض جيدة ، من خلال فهم سلوك العملية المطلوبة هو مطلوب. لغرض تصميم مسيطر جيد يجب فهم تصف المنظومة المراد السيطرة عليها. عدة بحوث سابقة صممت المسيطر المضبب بأستخدام متغيري الخطأ ومعدل التغيير في الخطأ. في هذا البحث تم تصميم مسيطر مضبب وان متغيرات هذا المسيطر هي الخطأ والتغير في معدل التغير في الخطأ (التعجيل). كما تم دراسة تأثير التعجيل على اداء المنظومة. اوضحت النتائج ان اضافة هذا المسيطر الجديد على محرك مؤازر اعطى طريقة كفوءة في السيطرة على سرعته.

1. Introduction

Modern control theory has success in area where the system is well defined, but it has failed to cope with the practicalities of many industrial process despite the development of the huge body of mathematical knowledge technique from the field of artificial intelligent may be usefully employed to control a complex, non-linear dynamic plant. Although such plants may be difficult to control manually, it may be possible to control them by means of a suitable heuristic program. The effectiveness of such programs has been demonstrated in chess playing and theorem proving etc. These programs may be very complex and hence difficult to construct, and may also take a long time to evaluate decisions. Thus, they have not often been applied to control algorithmic plant [1].

Although in theory it should be possible to do so. Fuzzy control algorithm is used to implement linguistically expressed heuristic control policy directly to automating those processes which are complex and poorly-defined where modeling difficulties and lack of suitable measurements, non-linear process. The problem of control and decision-making in fuzzy environment presents one of the most fundamental and challenging issues measure-our in ability to design and analysis of man-machine systems, [2].

In a large measurer to design such machines systems from a fundamental difference between human intelligence on the one hand, and machine intelligence on the other. The difference lies in the ability of the human brain-inability which present day digital computers do not possess to think and reason in imprecise, non quantitative, fuzzy terms. It is this ability that makes it possible for humans to decipher sloppy hand writing understand distorted speech, and focus on that information which is relevant to decision [3, 4].

It is the lack of this ability that makes even the most sophisticated large scale computers incapable of communicating with humans in natural-rather than artificially-constructed-languages. The first implementation of a rule-based controller was to control a laboratory steam engine, which was a two inputs-two outputs process [5]. In most previous studies a consistent control the authors applied fuzzy logic techniques to design a fuzzy controller with two inputs and one output. In this research the acceleration fuzzy controller designed by including acceleration as a new variable replacing the ordinary term (velocity) or the rate of error. The

robustness of this controller was demonstrated using a digital simulation to study its performance in response to various changes in the process parameters [6, 7].

2. Related Works

The field of application of fuzzy logic in designing controllers for industrial plants is widely. It should be noted that fuzzy logic is a minor an aspects of the whole field of fuzzy mathematics and perhaps little related to the majority of the rest of the literature on the subject. Vague logical statements can be used to construct computational algorithms which may be used to drive inferences from vague data, this method is useful in the treatment of complex humanistic systems. This method could be applied to "hard" systems such as industrial plant controllers. In such cases where a linguistic control protocol can be obtained from a skilled operator of a plant, fuzzy logic, can be used to synthesis this protocol.

In order to realize precise speed control of driving motor, an adaptive fuzzy Proportional–Integral–Derivative (PID)control strategy for motors was established based on the existing PID control theory. The motor speed control model is built by simplifying the parameters of a brushless DC motor used the Sim Power Systems toolbox in MATLAB/Simulink environment, which involves the simulation of motor speed control including low speed, high speed, and road bump situations in city traffic environment. Simulink environment, which involves the simulation of motor [8].

A comparative study on DC motor using different types of controllers was presented with optimization techniques and literature review on speed control of DC motor [9]. Shreekant1 Dr. M. S. Aspalli used PWM for speed control of a DC motor is generated using ATMEGA16 microcontroller. DC motor is driven by a driver IC IR2100. The H-bridge MOSFET Chopper is driven by a driver IC IR2100 through PWM signal generated by the microcontroller. A practical one and high feasible according to economic point of view and accuracy and it had been tested on 0.5HP DC motor. [10].

The position control scheme of DC servo motor drive had been presented. The Fuzzy Logic Controller (FLC) is developed for controlling the position of DC servo motor drive. The DC servo motors are highly preferred because of high power rating and speed of the motor. The fuzzy input variable (error) has seven membership functions, the fuzzy input variable

(change of error) has five membership functions, and the fuzzy output variable represented by applied voltage has five membership functions [11].

3. Fuzzy Logic

During the past dozen years, the seminal work of Zadeh has stirred widespread interest in the modeling of human reasoning through the use of fuzzy logic, leading to many important contributions from a word wide community of scholars and scientists In growing numbers-investigating a wide variety of fields ranging from psychology, philosophy and economics to natural sciences and engineering-are exploring this new path to the understanding of human reasoning and cognition, and developing novel methods for dealing with systems that are too complex for analysis by conventional quantitative techniques [11].

Fuzzy logic is a logic that is much closer in spirit to human thought and language than the conventional logical systems. But, the goal of science and objective knowledge is to construct models that are closer and closer approximations to reality, [12]. Considering that human thinking is articulated by an internal language that is implicit and an external language that is explicit in nature. General form of a fuzzy rule is "IF X is A THEN Y is B", where X and Y are fuzzy variables and A, B are fuzzy sets. The inference, based on the fuzzy rule-has the same form of fuzzy rule, Zadeh proposed a compositional rule of inference, also called the Generalized Modus Ponens (GMP) [13].

4. Mathematical Modeling of D. C. Servo Motor System [14, 15]

A DC motor is used in a control system where an appreciable amount of shaft power is required. The DC motors are either armature-controlled with fixed field, or field-controlled with fixed armature current. DC motors used in instrument employ a fixed permanent-magnet field, and the control signal is applied to the armature terminals. Figure (1) represented the servo motor model. Let's consider:-

$E_a(t)$ =Input voltage

$i_a(t)$ =Armature current

R_a = Armature resistance

L_a = Armature inductance

$E_b(t)$ =Back e.m.f

T_m = Developed Torque

ω_m =Motor angular velocity

J = Motor moment of inertia

B =Viscous friction coefficient

K_b=Back e.m.f constant

K_T=Torque constant

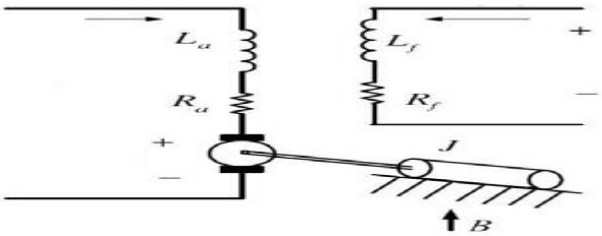


Figure (1): DC servo motor model
Here, the differential equation of armature circuit is:

$$E_a(t) = R_a.i_a(t) + L_a.\frac{di_a(t)}{} + E_b(t) \quad \text{.....(1)}$$

The Torque equation is:

$$T_m(t) = J.\frac{d\omega_m(t)}{} + B.\omega_m(t) \quad \text{..... (2)}$$

The torque developed by motor is proportional to the product of the armature current and field current i.e.

$$T_m(t) = K_f.i_f.i_a \quad \text{..... (3)}$$

Where, K_f is constant. In armature – controlled D.C. motor the field current (i_f) is kept constant i.e.

$$T_m = K_T.i_a \quad \text{..... (4)}$$

Where, $K_T = K_f.i_f$ is torque constant. The back e.m.f. of motor is proportional to the speed i.e.

$$E_b(t) = K_b.\omega_m \quad \text{.....(5)}$$

Where, K_b is back e.m.f. constant. In order to create the block diagram of system initial conditions are zero and Laplace transform is implemented to the equations. i.e.

$$E_a(s) = R_a.I_a(s) + sL_a.I_a(s) + E_b(s)\ I_a(s) = Ea(s) - E_b(s) \quad .....(6)$$

$$sL_a + R_a\ T_m(s) = sJ.\omega_m(s) + B.\omega_m(s)\ \omega_m(s) = T_m(s) \quad .....(7)$$

$$sJ + B\ T_m(s) = K_T.I_a(s) \quad .....(8)$$

$$E_b(s) = K_b.\ \omega_m(s) \quad .....(9)$$

The block diagram of DC servo motor is shown in Figure (2).

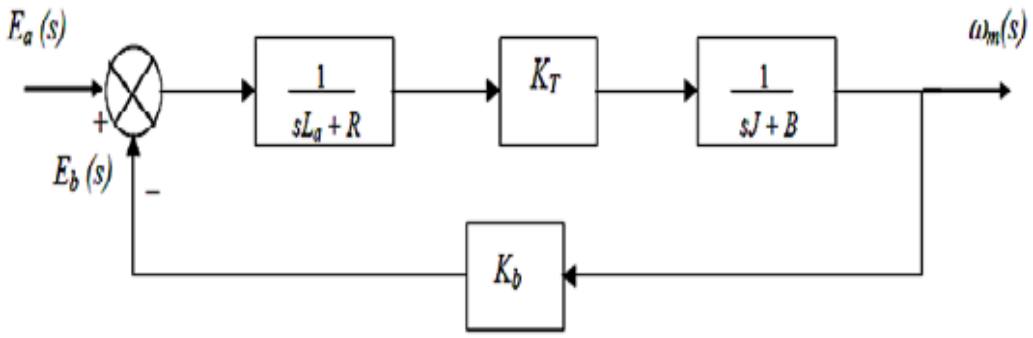


Figure (2): Block diagram of separately excited DC servo motor

The motor used in this experiment is a 25V D.C. motor with no load speed of 4050 rpm. The parameter values are

R-resistance 1 Ω

L-inductance 29.79 mH

J-moment of inertia 0.01 $kg.m^2$

K_{t_b}-torque constant 0.052 Nm/A

K-electromotive force constant 0.1 V/rad/s

B-viscous friction coefficient 0.004 N.m/rad/s.

5. The Proposed Design of an Acceleration Fuzzy Logic Controller (AFLC)

There is to-date no generalized method for the formulation of fuzzy control strategies, and design remains an ad hoc trial-and-error exercise. Recently, the relay feedback test has received a great deal of attention in low-order modeling and controllers tuning. The purpose of any plant controller is first to relate the state a variables to action variables i.e., to periodically look at the values of the state variables and from the expressed relationships to compute the value of the action variable. Now the controller of a physical system need not itself be physical but may be purely logical. Furthermore, where the known relationships are vague and qualitative a acceleration fuzzy logic controller is constructed to implement the known heuristics. Thus, in such a controller the variables are equated to non-fuzzy universes giving the possible range of measurement or action magnitudes. These variables, however, take on linguistic values which are expressed as fuzzy subsets of the universes. Figure (3) gives the block diagram of control system after applying the acceleration fuzzy controller.

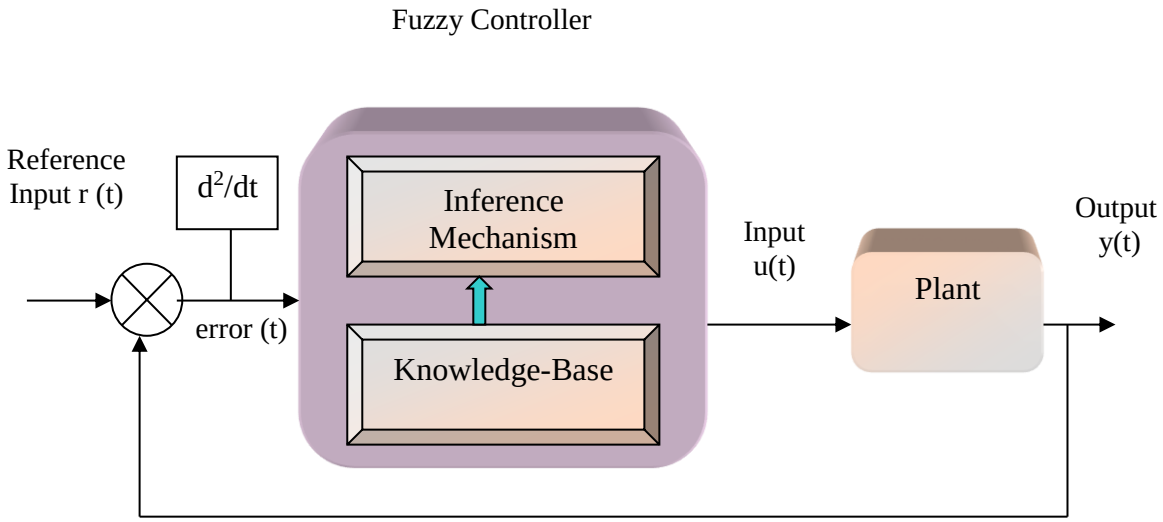


Figure (3): Control system with fuzzy logic controller

Fuzzy controller is in fact a collection of linguistic rules which describes the control policy. These rules are represented by IF and THEN statements such as **: " If E is PL AND CCE is NL then Action PucL "**Where E is error, CCE the change in rate change of error and **PucL** (Positive Large) is the control action (UC) to the process. These rules are expressed as fuzzy sets of the respective universe of discourse by corresponding membership functions, as

shown in Figure(4), triangle shape. Then the fuzzy set definition for (E), (CCE) and control action (CV) are five fuzzy sets are:

Operator Variables	Fuzzy Variables
Positive Large	PL
Negative large	NL
Zero	Z
Positive Small	PS
Negative Small	NS

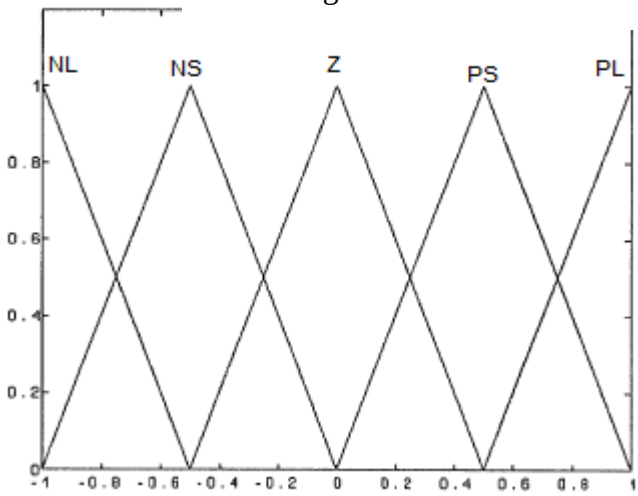


Figure (4): Five Fuzzy Sets (Triangles Shape)

While Table (1) shows the 25-rules used in the design of acceleration fuzzy controller.

Table (1) 25-rules used in the design of acceleration fuzzy controller.

E / CCE	NL	NS	Z	PS	PL
NL	PucL	PucL	PucL	PucS	Zuc
NS	PucS	PucS	PucS	Zuc	NucL
Z	PucS	PucS	Zuc	NucS	NucL
PS	PucS	Zuc	NucS	NucL	NucS
PL	Zuc	NucS	NucL	NucL	NucL

The complete design procedure for this controller is described as follows:-

1) Calculate the error (E) and its rate of change (CE) from the set point R and measured output (C), then calculate the change in change of error (CCE), at instant (i);

$$E(i) = R - C(i) \quad \text{----------} \quad (10)$$

$$CE(i) = E(i) - E(i-1) \quad \text{----------} \quad (11)$$

$$CCE(i) = CE(i) - CE(i-1) \quad \text{-------} \quad (12)$$

2) Convert the real numbers of E and CCE to their fuzzy variables by mapping the corresponding membership values.

3) Find the degree of fulfillment (DOF) for all the rules in acceleration fuzzy controller, the truth values of each rule are obtained by ANDing E(i) and CCE(i) e.g.

Rule (1) = $E1 \cap CCE_1$

Rule (2) = $E2 \cap CCE_2$

And so on for Rule (i) = $Ei \cap CCEi$.

4) By another ANDing operation between DOF degree of fulfillment control action of subset (C_n) element which calculated by

Rule(1) = $DOF_1 \cap U_1$

Rule(2) = $DOF_2 \cap U_2$

Rule(n) = $DOF_n \cap U_n$

5) The overall fuzzy output C_n is calculated as the union of output control action related to each rule.

6) In final stage, calculate the scalar control action Uc, by using center of gravity method as bellow: -

$$Uc = \frac{\sum_{n=1}^{N} I * W_n'(x)_n}{\sum_{n=1}^{N} W_n'(x)_n} \qquad \forall \ (x < 0.5 \qquad) \quad \text{-------} \quad (13)$$

Where W_n is the weight represents the elements (membership) of the net control action vector and I represents value on interval n.

6. Controller Evaluation

The fuzzy control algorithm is used to implement linguistically expressed heuristic control policies directly. The fuzzy algorithm have used in the majority of control applications. Control system has taken the form of continuous systems. Since digital implementations of the fuzzy logic controllers have by necessity been discrete system. Using acceleration as an input to fuzzy controllers the controller become more adaptive that is mean acceleration term an estimate, the previous behavior of the system better than rate in error change (velocity).

This capability gives the more powerful to fuzzy controller. The results illustrate the powerful, adaptively and excellent tracking to follow the output response of the system when the system have sudden change in step input. For simplicity the rules are evaluated at regular intervals. The choice of sampling interval depends on the process being controlled and should be selected so that at cheats twice the largest frequency in continuous signal as Shannon sampling theorem said. Figure (5) shows the simulated results of applying the acceleration fuzzy controller to the digital plant (2/S(S+1)) with unit step input. Eq. (14) is first decomposed into a first-order model, a time delay element with a fuzzy logic controller and predictor are designed separately for the two processes. Consider a normalized first-order process with the delay removed and having unit gain and unit time constant

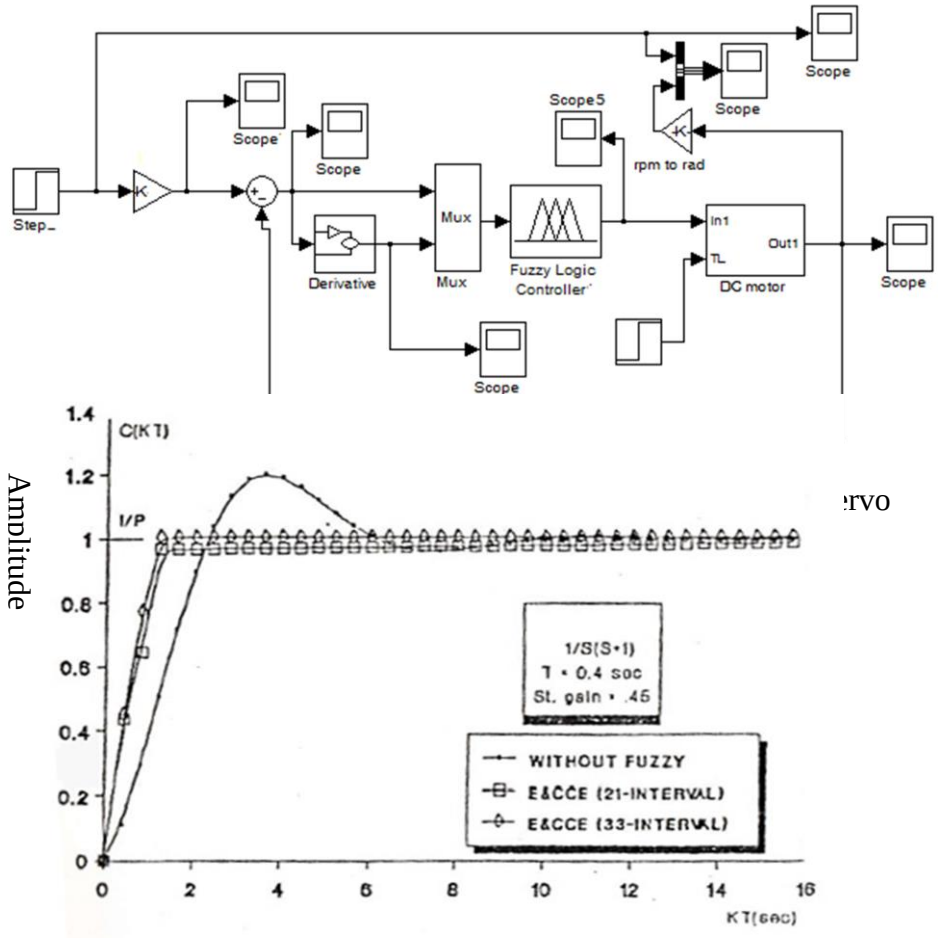
$$G(s) = \frac{1}{s+1} \qquad \dots\dots(14)$$

System G(s) is then converted to a sampled data process with sampling interval equal to (5 sec) of the system time constant such that the sampled data process can capture the original system characteristic. With a sampling interval equal to (5sec) of the time constant of the process, the sampled data process obtained from the z-transform of Eq. (15) with a zero order hold is given by;

$$G(z) = \frac{1 - e^{-0.2}}{z - e^{-0.2}} \qquad \dots..(15)$$

The effect of different interval of universe of discourse has been shown in both Figures (6, 7). One of the most important advantages for use acceleration instead of velocity (rate of change) is the good estimation for the future change in the system's response, which give the controller a good resolution and ability to track the signal carefully in step by step in order to

examine this ability of acceleration fuzzy controller. In this research the new acceleration fuzzy controller gives a very good response when insert this controller with unstable system.



ervo

Figure (6): Acceleration fuzzy controller with Ts-0.4 sec, and different time intervals.

-8-

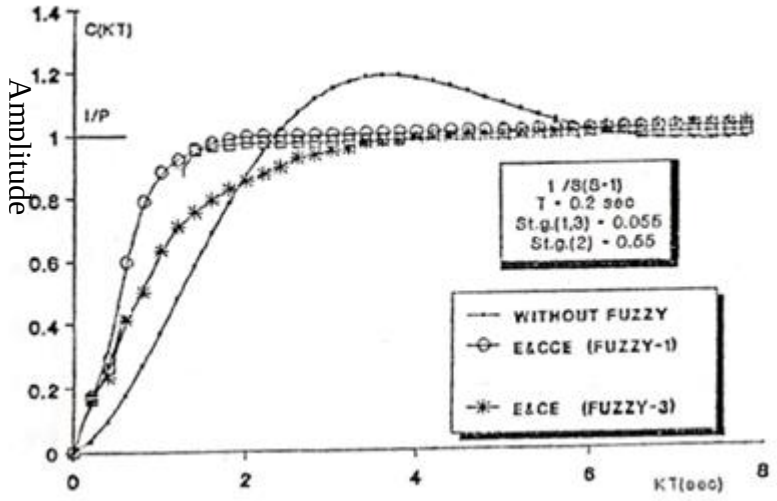


Figure (7): Acceleration fuzzy controller with Ts-0.2 sec, and different time intervals.

7. Conclusion

In practice, acceleration fuzzy controllers have shown themselves to be useful in those domains where a good process model was not available, but human operators could perform adequately. The installation of acceleration fuzzy controller rules has for some of the process industries been the first time when a result of the consistent, automatically applied control strategy was feasible. As a result of the consistency, the management of the plant has been able to observe its long term behavior with only natural plant disturbances to cope with, freed from the variations created by the different control strategies of the shift operators. As a result of consistency of approach, two benefits may be obtained.

1) The organization benefits from a more consistent product, providing a marketing edge over their competitors.

2) The consistency of process input means that experiments may be performed showing up behavior of the process that was previously masked by the variations in performance.

Reference

[1] R.M. Tong, "A control Engineering Review of Fuzzy Systems," Automatic control, 2009.

[2] E.H. Mamdani, J. J. Qstergaard and E. Lembessis, "Use_Of_Fuzzy_Logic_For_Implementing_Rule-Based Control of Industrial Process," Times/studies in the management sciences, 20(2004).
[3]Juan Luis Castro, Miguel Delgado, Carlos Javier Mantas, "A New Approach for the Execution and Adjustment of a Fuzzy Algorithm," Fuzzy Sets and Systems 121, 2010.
[4] Solmaz Ghazanfar Aharia, Nader Ghaffari-Nasabb, Ahmad Makuib, and Seyed Hassan Ghodsypoura "A Portfolio Selection Using Fuzzy Analytic Hierarchy Process: A Case Study of Iranian Pharmaceutical Industry," International Journal of Industrial Engineering Computations, 2011.
[5] L.A.Zadeh, E.sanchez, "Approximate Reasoning in Intelligent Systems Decision and Control," Pergamon Books Ltd. 2006.
[6] P. Magrez and P. Smets, "Fuzzy Modus Ponens: A New Model Suitable For Applications In Knowledge-Based Systems", Int. J. of intelligent systems,
Vol. 4, No. 2, summer 2008.
[7]Yogesh Garge, Prof. S. M. Chaudhari, "PID and Fuzzy Control of BLDC Motor," International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Issue 4, April 2017.
[8] Wanmin Lia,b, Menglu Gub, Lulu Weib, " Speed Control Simulation of the Electric Vehicle Driving Motor," International journal of performability engineering, Vol. 13,no. 7, 2017.

[9]Tarei1, Satish Arora, ,"Comparison of Speed Control of DC Motor using PID Controller and Various Optimization Techniques a Review," Samikshya, IJESC,2017.
[10]Shreekant1 Dr. M. S. Aspalli,"Chopper based Speed Control of DC Motor using ATMEGA16 Microcontroller," International Journal for Scientific Research and Development| Vol. 4, Issue 12, PP.2321-0613, 2017.
[11]Zain-Aldeen S. A.Rahman," Design a Fuzzy Logic Controller for Controlling Position of D.C. Motor," International Journal of Computer Engineering In Research Trends, Volume 4, Issue 7, pp. 285-289, July-2017.
[12] A. Prithiviraj, K. Krishnamoorthy, B. Vinothini, "Fuzzy Logic Based Decision Making Algorithm to Optimize the Handover Performance in HetNets Circuits and Systems," Vol.7, pp.3756-3777, 2016.
[13]A. Robles, E. Latrille, M. V. Ruano and J. P. Steyer, "A fuzzy-logic-based controller for methane production in anaerobic fixed-film reactors," Journal of Environmental Technology, Vol. No. 38, 2017.
[14] I.J. Nagrath and M. Gopal, "Control Systems Engineering," New Age Publication, 2003.
[15] Katsuhiko Ogata, "Modern Control Engineering," Fifth Edition, Copyright, Pearson Education, Inc., publishing as Prentice Hall, 2010.

Experimental study on the effect of the acceptance angle of the compound parabolic concentrator on the performance of the solar collector

Dr. Talib Z. Farge[1] , Dr. Abd UlJabbar Owaid[2] & Asst. Lec. Mina A. Nsaif[3]

1. Assistant Professor, Electromechanical Engineering Department, University of Technology, Iraq drtalebzf@yahoo.com

2. Co Supervisor, Electromechanical Engineering Department, University of Technology, Iraq aljabbara.2007@yahoo.com

3. Assistant Lecturer, Physical education and sports science Department, Al Nisour University College, Iraq minaabdalkareem7@gmail.com

Abstract

This work aims to study experimentally the effect of different acceptance angles for the compound parabolic concentrator (CPC) on the performance of solar collector. The effect of three values of acceptance angles of (30º, 45º, 60º) for the solar collector has been studied. The diameter of the absorbing tube for the solar collectors of (1.5 inches) has been used with the different values of the acceptance angles. A halogen light of (1000) Watt has been used as an energy source. The relationship between the time with the readings of water temperature inside the absorber tube of the solar collector and shows three regions of variation for all the three collectors. Where in the first region, it varied sharply, in the second region, it varied gradually and in the third region it remained constant. Comparing the performance results between the different solar collectors with different acceptance angles show that the collector with an acceptance angle of 45º gives the best performance. Where the percentage increase in the water temperature difference for the collector with an acceptance angle of 45º comparing with the other acceptance angles (30º, 60º) is (10.67 and 45.33) respectively. While in the case of heat energy are (10.66 and 45.33) and for the solar collector efficiency are (36.74 and 36.77) at the time of 120 min.

Key words: CPC, solar collector, compound parabolic concentrator, acceptance angles and diameter of absorbing tube.

دراسة عملية حول تأثير زاوية القبول على اداء السخان الشمسي ذو القطع المكافئ المركب

الملخص

يهدف هذا البحث لدراسة تأثير زوايا قبول مختلفة لمركز شمسي ذات قطع مكافئ مركب على اداء مجمعات شمسية عمليا .حيث تم دراسة تأثير ثلاث زوايا قبول مختلفة هي (30^o,45^o,60^o) على اداء المجمعات الشمسية . تم استخدام قطر انبوب امتصاص للمجمع الشمسي مقداره (1.5 انج) مع زوايا قبول مختلفة . وقد استخدام ضوء مصباح الهالوجين بطاقة 1000 واط كمصدر للطاقة . لقد اظهرت العلاقة بين الزمن والقراءات لدرجات الحرارة الماء داخل انبوب للمجمع الشمسي بوجود ثلاث مناطق للتغير لكل المجمعات الشمسية الثلاثة . حيث يكون التغيير في المنطقة الاولى حاد . ويكون التغيير في المنطقة الثانية تدريجي وفي المنطقة الثالثة يكون التغيير ثابت تقريبا. بمقارنة نتائج الاداء بين المجمعات الشمسية المختلفة عند زوايا قبول مختلفة وجد بأن المجمع الشمسي عند زواية قبول ل45^o يعطي افضل نتائج اداء مقارنة بألاخرين . حيث ان اعظم نسبة زيادة عند المنطقة الثالثة عند زمن 120 دقيقة في الفرق في درجة حرارة الماء للمجمع الشمسي ذات زاوية قبول 45^o مقارنة للمجمعان الشمسيان الاخران عند زوايا قبول (60^o,30^o) هما (45.33,10.67) على التوالي و بينما في حالة الطاقة الحرارية هما (10.66,45.33) وكذلك لكفاءة المجمع الشمسي هما (36.74,36.77).

الكلمات الدلالية : مجمع شمسي , مجمع شمسي ذات قطع مكافئ مركب , زاوية قبول , قطر انبوب الامتصاص

1. Introduction

Solar energy collector is a special kind of heat exchanger which transforms solar radiation energy to useful heat energy. The solar collector is the main component of any solar system [1,2]. A. Rabl, et. al. [3], analyzed the design parameters of solar collectors with (CPC). They studied the effect of the type of the receiver, the optimum gap between the reflector and the receiver to reduce thermal and optical losses, and also the effect of receiver with envelope of glass. Also, they studied the effect of misalignment between the mirror and the receiver, and the effect of the difference between fluid temperature and absorber plate temperature. Los Alamos [4] analyzed and compared the arc-Lengths of reflector and the ability to reflect of non-direct radiation and absorb by the absorber tube of the solar collectors with simple parabolic concentrator and with compound parabolic concentrator. In the case of simple parabolic concentrator the absorber tube diameter can make larger to intercept some defocused radiation... B.Norton, et. al. [5] described the thermal performance and design of modified solar collectors with (CPC). The modified collector with (CPC) give a better performance than the convectional collector with (CPC) by eliminating and or reducing the gab optical losses and heat losses, which enhancement of the thermal efficiency and optical efficiency. D. Suresh, et. al. [6] evaluated the characteristics of performance (optical and thermal) of a truncated CPC by using two-stage solar collector systems. The 5:1 cones with CPC with an acceptance angle of 25^o and ratio of concentrator of 5.6x and untruncated. They tested cones have the same

dimensions but with different reflector surface, shell materials and employing different methods to heat removal. The experimental results show that the CPC with higher reflectivity surface have a higher optical efficiencies. S senthilkumar et.al [7] studied theoretically and experimentally the performance of the collectors of two dimensional and three dimensional (CPC). They found the performance of the thermal and optical efficiencies of the three dimensional (CPC) better than that of the two dimensional (CPC) due to greater concentration ratio. Rene Tchinda [8] presented a mathematical model to calculate the thermal performance of solar air heater with a truncated CPC having a flat absorber. The effected parameters as the flow rate of air mass, the length of the collector and wind speed on the solar collector performance of the air heater was investigated. The predication results show a good agreement with experimental results. GAO Hong-yu, et. al. [9] Predicted and analyzed the performance of solar collector with CPC and evacuated tube as absorb tube and also without and with top cover of glass. The comparative results without and with glass top cover show that for low operating temperature the collector without cover had a higher thermal efficiency, while in the case of the high operating temperature the collector with top glass cover had a higher thermal efficiency. Mansi G. Sheth and Dr. P. K. Shah [10] designed and developed a solar collector with CPC and absorber of flat plate. A. Borah, et. al. [11] Studied experimentally the performance of solar collector with CPC and the solar collector with flat plate having the same size for drying product. They observed that the mean temperature rise in the solar collector with CPC of 9.5 ^{o}C greater than the mean temperature in solar collector with flat without tracking, while in the case of manual tracking the mean temperature rise about 11.2 ^{o}C.

This paper presents a system of test rig of solar collectors with the CPC have been designed and tested experimentally to study the performance, namely (heat energy and solar collector efficiency) for the solar collector with different acceptance angle of (30^{o}, 45^{o}, 60^{o}).

2. Experimental work

The experimental work was carried to study the performance of different solar collectors with different compound parabolic concentrators (CPCs). Figure (1) shows the experimental rig arrangement with the major parts of it. They are the halogen light (1000 Watt) as an energy source, solar collector and thermometers. Figure (2) shows the solar collector which consists of an external structure, reflector plate of a compound parabolic concentrator and absorber tube. Figure (3) shows the collectors CPC with different acceptance angles. Table (1) shows

the area of different solar collector with CPC with different acceptance angles. The reflector plate made from the stainless steel . The absorber tube made from copper with thermal conductivity of (k=385 W/m.K) and painted by black paint. The CPC's external structure is made from rigid polyurethane foam with (1 cm) thickness. A halogen light has been used as an energy source which consumes (1000) Watt. A thermometer with (-10~+110C) range has been used to measure the temperature inside the absorber tube with a high accuracy. To ensure the accuracy of manufactured of a compound parabolic concentrator so, a laser beam projected to the surface of CPC plate and reflected to the absorber tube as shown in figure (4).

Table (1) the area of solar collector with CPC with different acceptance angles

Acceptance Angle (deg)	**Length (cm)**	**Wide (cm)**	**Area (cm^2)**
30◦	24.4	22	536.8
45◦	17.3	22	380.6
60◦	14.7	22	323.4

Figure (1): The test rig

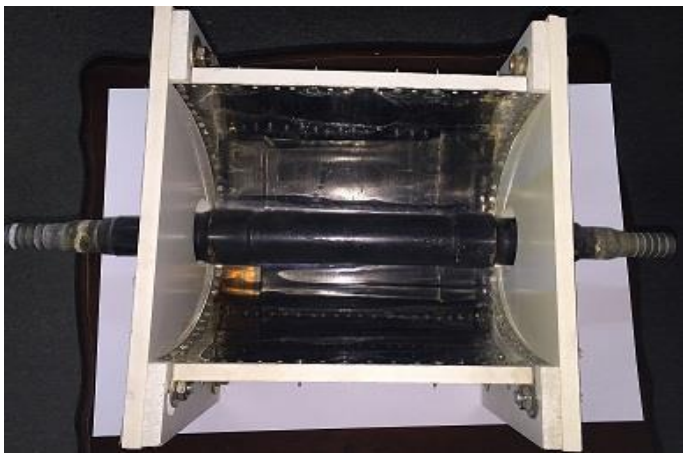
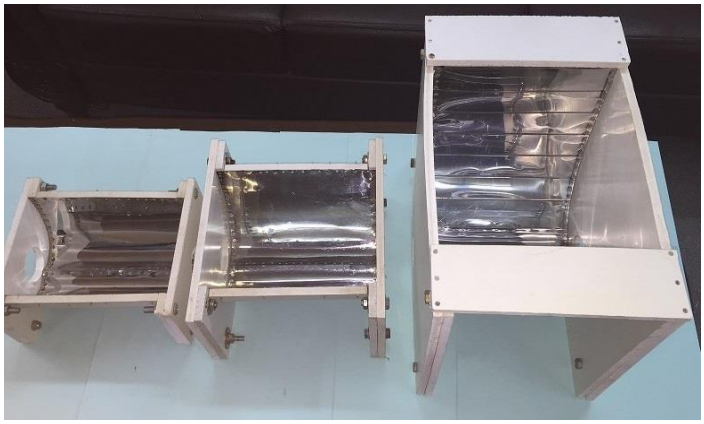


Figure (2): CPC solar collector with (45◦) acceptance angle and pipe diameter (1.5 inch)



(a): 60^{o} (b): 45^{o} (c): 30^{o}

Figure (3): CPC with difference, acceptance angle (30◦, 45◦, 60◦)

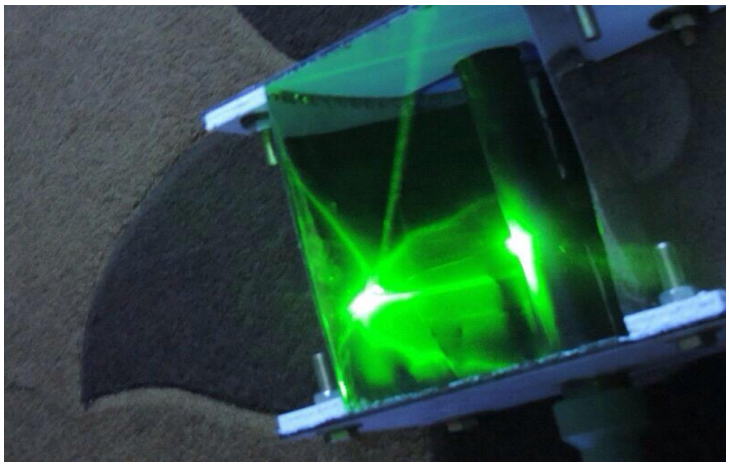


Figure (4): laser test

3. Compound parabolic concentrator design:

The surface curvature of the compound parabolic concentrator constructs from two curves [12], the first one from A to B and the second one from B to C as shown in figure (5). Three independent variables (r ,θ_{ac} and θ),

Where

R = f (r ,θ_{ac} ,θ)

For a certain surface of the compound parabolic shape the values of radius of the absorber tube (r) and the acceptance angle (θ_{ac}) kept constant and the value of the angle (θ), which it's values change from zero to the ($\frac{3\pi}{2} - \theta_{ac}$) on the y-axis.

The curves of the CPC are symmetric about the y-axis and can obtained from the following equations [13].

a) The equation of first part of the curve from A to B is :

$$R = r.\theta \quad for\ |\theta| \leq \theta_{ac} + \pi/2 \qquad \text{……(1)}$$

b) The equation of the second part of the curve from B to C is :

$$R = r.\frac{\theta+\theta_{ac}+\pi/2-\cos(\theta-\theta_{ac})}{1+\sin(\theta-\theta_{ac})} \quad for\ (\theta_{ac} + \frac{\pi}{2} \leq |\theta| \leq \frac{3\pi}{2} - \theta_{ac}) \qquad \text{……(2)}$$

Where R is the distance from point D on the absorber tube to point S on the surface of CPC.

The concentration ratio of the CPC is a function of the acceptance angle given by following equation [13]

$$CR = \frac{1}{\sin(\frac{1}{2}\theta_{ac})} \qquad \text{……(3)}$$

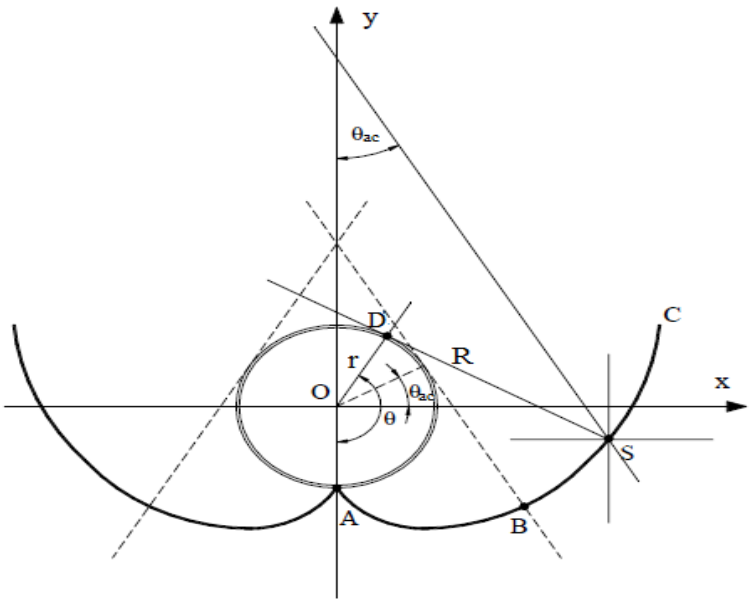


Figure (5): The CPC shape [13]

The calculations the performance of the solar collector are :

a) Mass of water (m_w)

$$m_w = \rho_w \,.\, \pi . r^2 . L \qquad \ldots(4)$$

Where:

m_w: mass of water (kg/sec)

ρ_w: Water density (kg/m^3)

r: The radius of the absorber tube (m)

b) Useful heat energy (Q_u)

$$Q_u)_{water} = (m_w \,.\, (C_p)_{water} \,.\, (\Delta T))/t \qquad \ldots\ldots(5)$$

Where:

$(C_p)_{\text{water}}$: Specific heat capacity of water (J/kg.K)

ΔT: Temperature difference between temperatures inside the absorber tube with respect to time (^{o}C).

t: time (sec)

$Q_u)_{water}$: heat energy (W)

c) Efficiency solar collector (η)

$$\eta = \frac{Q_u)_{water}}{A.I_H} \qquad \ldots\ldots(6)$$

Where:

A: Area of the collector (m^2)

I_H= Halogen lamp intensity(W/m^2)

d) Percentage in prove in the performance

The percentage in the performance between (45˚ and 30˚) the accuracy is calculated, as following

$$\Delta T\% = \frac{\Delta T_{45} - \Delta T_{30}}{\Delta T_{45}}$$

$$Q\% = \frac{Q_{45} - Q_{30}}{Q_{45}}$$

$$\eta\% = \frac{\eta_{45} - \eta_{60}}{\eta_{45}}$$

4. Results and Discussion

Figure (6) shows the reading values of the water temperature difference inside the absorbing tube with respect to time of heating by the halogen light of 1000 Watt as an energy source, for different values of acceptance angles (30^o, 45^o, 60^o) of CPC solar collectors. The figure shows that the water temperature difference of the solar collector with an acceptance angle of 45^o of CPC is increasing more than the others due to the higher concentration ratio. Also the figure shows there are three different regions of variation of increasing of water temperature difference with respect to the time for all different solar collectors. The first one is a sharp increasing, the second region is a gradually increasing and the third one is an approximately constant. These are due to the cold bodies gained heat quickly and the hot bodies gained heat slowly.

Figure (7) shows the relationship between the heat energy and the time for different solar collectors. The heat energy for different collectors have higher values at the lower time and decrease when the time increase, because the heat energy function of water temperature difference and the time and in this case the water temperature difference it's has a higher value in the first than other regions as shown in figure (6). Also figure (7) shows that the collector with an acceptance angle of 45^o has a higher value of heat energy due to high heat temperature difference than the other collectors.

Figure (8) shows the comparison between the solar collector efficiency and the time for different collectors. In this study the solar collector efficiency is a function of heat energy and the area of the collector when the heat intensity is kept constant. The efficiency of the collector

with a CPC of the acceptance angle of 45^{o} has a higher value than the others due to the high value of heat energy. The efficiency of the collector with a CPC of the acceptance angle of 60^{o} is a higher than the collector with an acceptance angle of 30^{o} due to a smaller area of the collector with an acceptance angle of 60^{o} as shown in table (1).

The figures (6,7 and 8) show that the solar collector with a CPC of the acceptance angle of 45^{o} give better performance than the other collectors. The figures (9,10 and 11) show the percentage increase in the water temperature difference for the collector with an acceptance angle of 45^{o} comparing with the other acceptance angles (30^{o}, 60^{o}) is (10.67 and 45.33) respectively. While in the case of heat energy are (10.66 and 45.33) and for the solar collector efficiency are (36.74 and 36.77) at the time of 120 min.

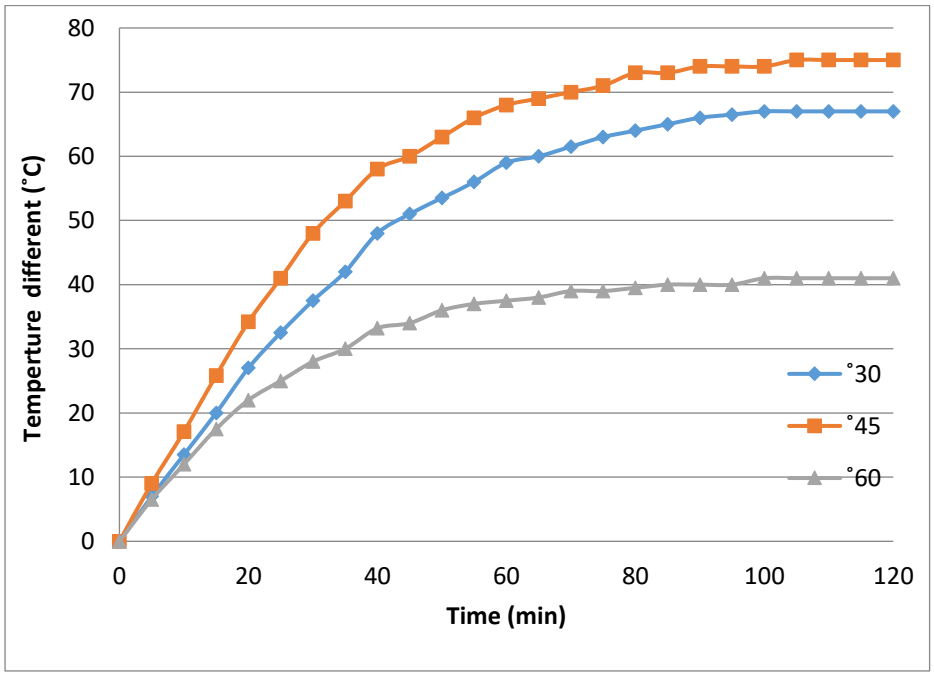


Figure (6): The relationship between the difference of water temperatures and time for different acceptance angles

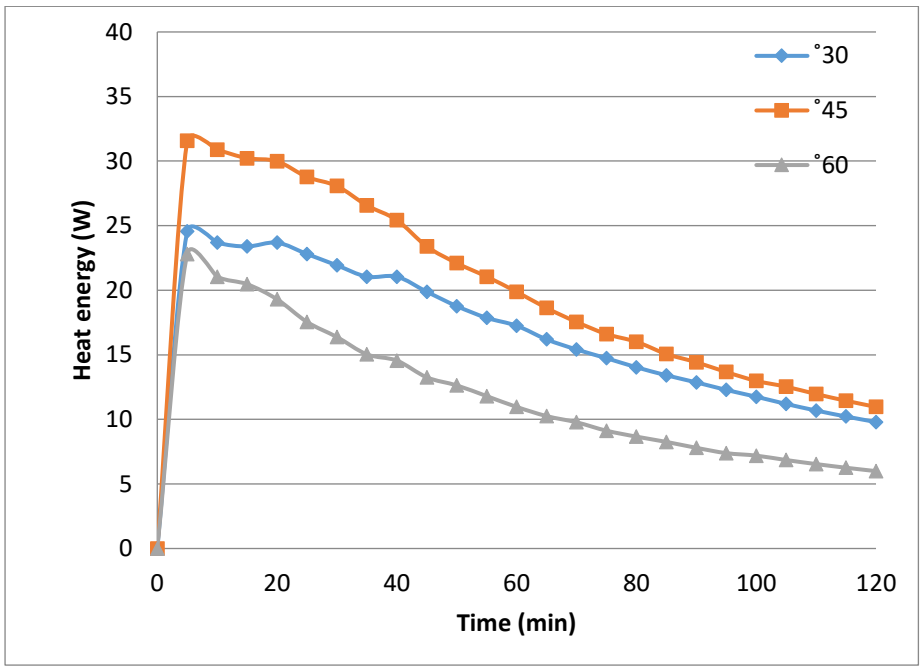


Figure (7): The relationship between the heat energy and time for different acceptance angles

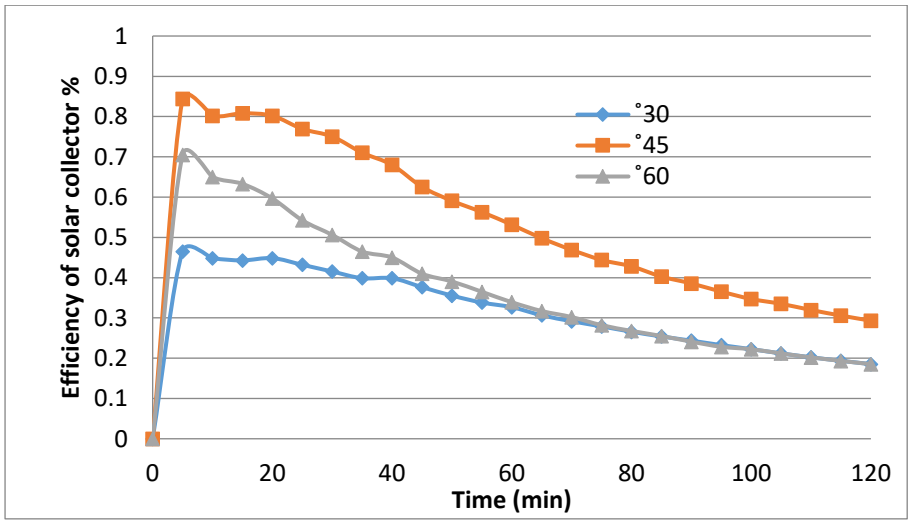


Figure (8): The relationship between the solar collector efficiency and time for different acceptance angles

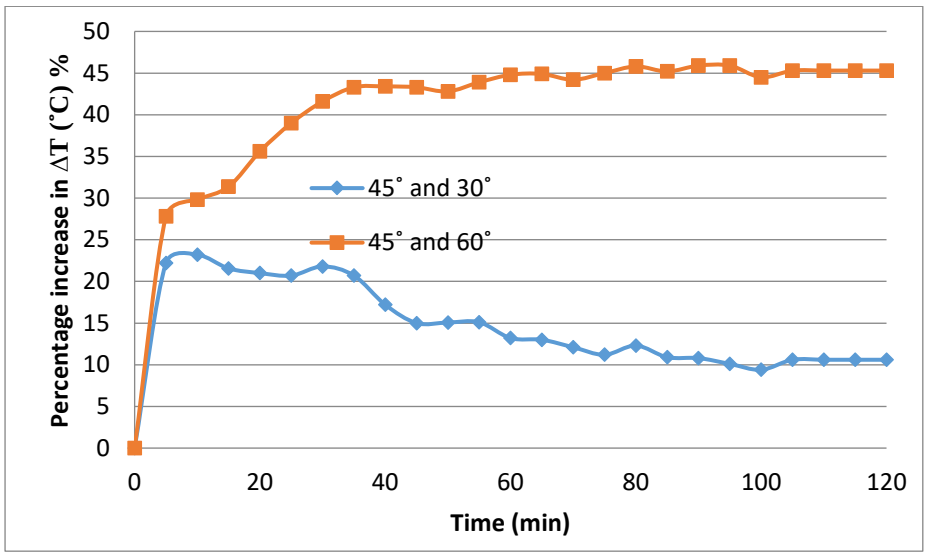


Figure (9): The relationship between the percentage increases in temperature difference and time.

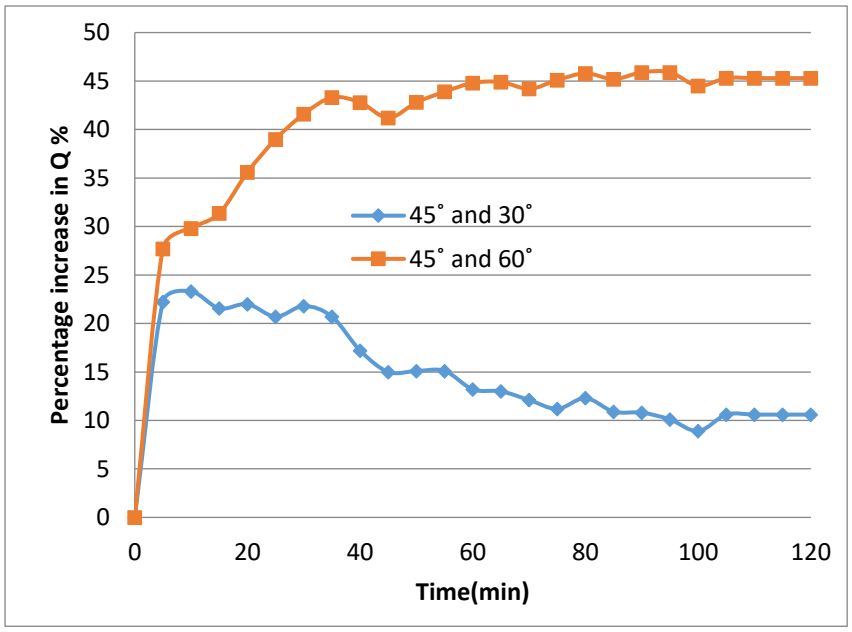


Figure (10): The relationship between the percentage increase in heat energy and time

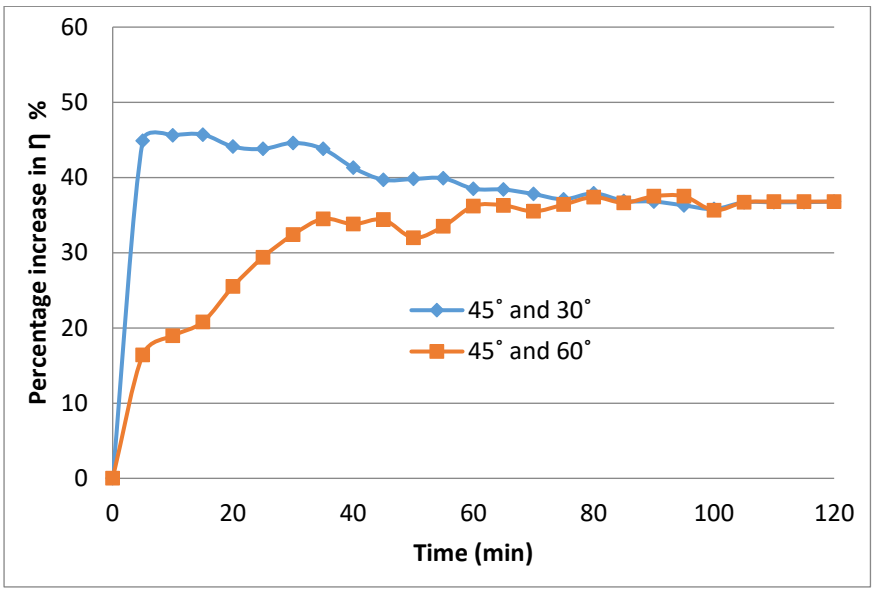


Figure (11): The relationship between the percentage increase in solar collector efficiency and time

5. Conclusion

This section shows some of the important conclusions of this work:

1- The best value of acceptance angle is (45^o) and it is considered the ideal value for absorbing the maximum amount of solar radiation.

2-The heat energy increases with increasing of water temperature difference.

3- The efficiency of the CPC solar collector depends on heat energy and cross sectional area of the CPC, so the collector with an acceptance angle of 30^o gives a lower value of thermal solar collector efficiency because of the large-sectional area of acceptance angle of 30^o and the smaller sectional area at acceptance angle 60^o.

4- the percentage increase in the water temperature difference for the collector with an acceptance angle of 45^o comparing with the other acceptance angles (30^o, 60^o) are (10.67 and 45.33) respectively. While in the case of heat energy are (10.66 and 45.33) and for the solar collector efficiency are (36.74 and 36.77) at the time of 120 min.

References

[1] Soteris Kalogirou. (2009)." Solar energy engineering: processes and systems. "Academic press.

[2] G.N. Tiwari "Advanced Solar Distillation systems ", (2017), Springer Nature Singapore Pte Ltd.

[3] A. Rabl , N.B. Goodman and R. Winston,(1979)," Practical design considerations for CPC solar collectors " Journal of Solar Energy, 22,4,(373–381).
[4] Los Alamos (1978) "A comparison of compound parabolic and simple parabolic concentrating solar collectors", Journal of Solar Energy, 22, 1, (21-25).
[5] B.Norton, D.E. Prapas, P.C. Eames and S.D. Probert (1989) "Measured performances of curved inverted-Vee, absorber compound parabolic concentrating solar-energy collectors ", Journal of Solar Energy,43,5,(267–279).
[6] D. Suresh J. O'Gallagher and R. Winston, (1990)," Thermal and optical performance test results for compound parabolic concentrators (CPCs)", Journal of Solar Energy, 44, 5, (257–270)
[7] S senthilkumar, K perumal and P S S srinivasan,(2007)," Optical and thermal performance of a three dimensional compound parabolic concentrator for spherical absorber",34,3,(369-380)
[8] Rene Tchinda, (2007) "Thermal behaviour of solar air heater with compound parabolic concentrator", Energy Conversion and Management
[9] GAO Hong-yu, Wang Hui-tao and Wang Hua, (2011)," Thermal Performance Analysis of A Novel Compound Parabolic Concentrator Solar Collector" International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE) December 16-18, Changchun, China
[10] Mansi G. Sheth and Dr.P.K.Shah,(2013)," Design and development of compound parabolic concentrating solar collector with flat plate absorber", International Journal of Innovative Research in Science, Engineering and Technology,2,8,
[11] A. Borah, S.M. Khayer and L.N. Sethi, (2013)," Development of a Compound Parabolic Solar Concentrator to Increase Solar Intensity and Duration of Effective Temperature." International Journal of Agriculture and Food Science Technology, 4, (161-168)
[12] D K Dixit, (2016),“Heat and Mass Transfer” McGraw Hill education (India) Private Limited.
[13] J.Blanco Galvez and S.Malato Rodriguez, "Solar photochemistry technology" solar energy conversion and photo energy system

A REVIEW ON TRIAGING FOR PATIENTS IN THE EMERGENCY DEPARTMENT: OVERVIEW, AIMS, GUIDELINE, AND CHALLENGES

Abdulniser Khald Hamzah[1]*, Louloua Mustafa Qasim[1,2], Golodov Valentin A.[3]

*Corresponding E-mail: nasernaser310@gmail.com

[1]Department of Computer Engineering Techniques Al-Nisour University Baghdad ,Iraq
[2]South Ural State University, School of Electrical Engineering and Computer Science

Abstract

Congestion in emergency departments is an overall issue. Triaging includes an underlying arranging of patients who arrive at the Emergency Department (ED), keeping in mind the end goal is to organize the most emergency patients and to guarantee giving them the proper and quick healthcare needed. This study plans to survey the models, guidelines, advancement of the electronic triage procedures of the triaging frameworks. Likewise, the goals of this investigation are to experimentally investigate and show the huge of utilizing triaging in various healthcare environment, to highlight the fundamental issues in the triage framework and to open research issues that guide the researches to improve the efficiency of remote healthcare monitoring frameworks by enhancing triaging processes.

Keywords: *Triage, Guidelines and Standard of Triage, E-Triage*

Introduction

Triage is an important task in an emergency department (ED). Triage is viewed as the rating of patient's clinical urgency (Farrohknia *et al.*, 2011a). Rating is important to recognize the order in which patients ought to be given care in an ED when demand is high. Triage scales mean to streamline the time of patients as per the seriousness of their health condition, keeping in mind to treat the most extraordinary symptoms as fast as necessary and to decrease the negative effect on the visualization of a drawn out postponement before treatment. ED triage is a relatively modern phenomenon, introduced in the 1950s in the United States (Gilboy *et al.*, 1999). Triage is a complex decision-making process, and several triage scales have been designed as decision Sport systems (Bullard *et al.*, 2008) to guide the triage nurse to a correct decision. To ensure patient safety and provide quality services, hospitals should be certain that each patient entering the ED relives appropriate care at the right time. Triaging is important for

patients' appraisal to appropriately anticipate the information's required for every patient and perceive abnormal vital signs.

Triage is originally from the French word trier signifying 'to sort,' is a procedure of prioritization (Bottrill *et al.*, 2008). Indeed, it is utilized to distinguish the level of urgency of care and to treat patients in light of their triage level (Farrohknia *et al.*, 2011b). Triage is the arrangement of arranging patients by order of treatment need in extensive scale emergencies (Jentsch *et al.*, 2013).It is likewise the way toward surveying and organizing look after all patients present in the ED (Innes *et al.*, 2011). Triage protocols are helpful to decide the need of patient's treatments in light of the seriousness of their conditions (Niswar *et al.*, 2013). An important goal of triage in the ED is to recognize patients who can safely wait and the individuals who can't (Brown and Clarke, 2014). Nurses who are experienced can examine patient's health condition as well as takes note of any progressions, and decides the patient's priority for admission to the ED and the essential treatment (Seising and Tabacchi, 2013).

Evaluation of the Medical Guidelines

Guidelines are verified on the basis of environment and type of disease, as shown in Tab 1.

Table 1: Review of Evaluation Studies for Medical Guidelines

Ref/year	Aim	Conclusion
(La Vonne *et al.*, 2015)	Compare the Canadian Triage and Acuity Scale (CTAS) protocol with the Australian Emergency Mental Health Triage System protocol to evaluate psychiatric patients in the ED.	The Australian Emergency Mental Health Scale provides less ambiguous mental health-specific triage guidelines compared with the CTAS.
(Hamamoto *et al.*, 2014)	Measure and compare the effectiveness of nursing triage before and after the introduction of the Japanese Triage and Acuity Scale (JTAS).	The JTAS positively affects nursing triage and medical care.

Ref/year	Aim	Conclusion
(Parenti *et al.*, 2014)	Conduct a systematic review to check the level of validity and reliability of the Manchester Triage System (MTS).	The quality of reporting in reliability and validity studies of the MTS is good.
(Gräff *et al.*, 2014)	Evaluate the German version of the MTS with the English version with respect to presentation diagrams and change indicators, which significantly affect the assigned category.	The German version of the MTS is a reliable and valid instrument for a first assessment of emergency patients in the ED.
(Ng *et al.*, 2010)	Compare the prioritization and resource usage of the four-level Taiwan Triage System and the standardized five-level CTAS among ED patients.	CTAS exhibits better discrimination in ED patient triage and shows greater validity when predicting hospitalization compared with the Taiwan Triage System.

The main finding of this section is that researchers continue to improve existing guidelines and that numerous studies were already conducted to demonstrate the evaluation of certain guidelines in different hospitals and cities. These studies aim to improve the performance of the guidelines. No restricted guidelines should be used, and the guidelines can be improved.

Review of Triage System Standards

In this section, different triage systems are reviewed. Those systems are being used in the ED in different countries and different hospitals. The reviewed systems are demonstrated in Tab 2.

Table 0: State of the art triaging systems

Triage System	Type	Description
Early Warning Scorecard (EWS) (Christensen *et al.*, 2011b) / 2011	Paper-Based	A simple scoring system for bedside monitoring to serve as a clinical add-on using routinely collected vital sign data.

Triage System	Type	Description
Modified Early Warning Scorecard (MEWS) (Zarabzadeh *et al.*, 2013)	Paper-Based	MEWS is a modified version of EWS. Vital sign data is manually sampled and recorded.
Electronic Modified Early Warning Scorecard (eMEWS) (Zarabzadeh *et al.*, 2013)	Digital	(eMEWS) represents a shift from paper-based warning Scorecard (eMEWS) systems to electronic systems. eMEWS is often designed and developed around the paper-based MEWS guidelines and associated protocols.
Emergency Severity Index (ESI) (Seising and Tabacchi, 2013)	Paper-Based	Designed for use in ED triage by the US Department of Health & Human Services.
Manchester triage system (MTS) (Pinto Júnior *et al.*, 2012)	Paper-Based	Developed in 1994 by a group of professionals specializing in triage. Classifies risks into five categories: immediate (red); very urgent (orange), urgent (yellow), standard (green) and non-urgent (blue).
Bispebjerg Early Warning Score (BEWS) (Christensen *et al.*, 2011a)	Paper-Based	BEWS used by the triage nurses inside ED to identify critically ill and severely injured patients who should be received by an Emergency or Trauma Call (TC).
Simple triage and rapid treatment (START) (Sakanushi *et al.*, 2013)	Paper-Based	It is widely used since it classifies casualties into four categories without medical equipment.

According to Tab 2 s six out of seven triage system that has been reviewed used paper triage, while only one used digital triage. Furthermore, the majority of these systems has been used inside ED department. The main conclusion in the section is there is a remarkable trend toward digitalization the triage system. The significance of the Emergency Department triage assessment has lead researchers to study and developed digital triage procedures based on medical guidelines. According to Krupinski and Bernard (2014), the advancement of rules and standards for telemedicine is a vital and important process that guarantees viable and safe conveyance of value healthcare. Associations, for example, the American Telemedicine Affiliation (ATA) organize the advancement of standards and rules and demonstrate that telemedicine will keep on growing as a vital and suitable strategy for enhancing access to medicinal services all through the world. In an ideal world, rules for giving healthcare services

in the traditional manner and telemedicine should be standardized. However, the current guidelines serve an assortment of valuable capacities and will keep on being perceived in telemedicine. A critical part of triage is the consistency in triage basic leadership and the utilization of triage classifications.

Several factors contribute to consistency in triage decision-making (Innes *et al.*, 2011). One of the main factors is "*the number of vital signs (features) of the patient*" used in triaging. Certain guidelines are also used in the ED to triage patients, as demonstrated by (Godfrey *et al.*, 2000). The guidelines differ from one disease to another, thus indicating that no unique guidelines exist in triaging patients for all diseases. Another factor called "*type of disease*" also affects the accuracy of triage decision.

Electronic Triage Technique

Remote triage is a process of triaging patients using electronic devices, medical sensors, and computing algorithms based on medical guidelines (Sakanushi *et al.*, 2013; Salman *et al.*, 2017). Many studies were conducted on this topic. The study by Gao and White (2006) developed a triage system to develop the efficiency of emergency response. The proposed system comprises E-triage tags, base stations laptops, wearable vital sign sensors, pervasive tracking software, and personal digital assistants (PDAs) to help documentation and correspondence. Their investigation condensed designing contemplations for advances that ought to work under limitations of medicinal emergencies. Despite the advantages in the study by Gao and White (2006), several drawbacks are notable; firstly, the triage process depends on only one feature, which affecting triaging accuracy. In most cases, a patient may have two levels of simultaneous priorities. Secondly, the system does not address the scalability challenge. The ability of the server to respond to all patients' requests was not made clear. Thirdly, the algorithm, model design, and implementation of their method were not analyzed mathematically or logically. Finally, no evaluation and validation method was presented.

In the study by Jorma *et al.* (2013), one mobile triage system called "mTriage" was proposed to determine the applicability of the radio frequency identification technology and a "mobile triage" system in a simulated multi-casualty situation. The mobile triage exhibited the potential to contribute for the management of mass casualty situations. However, one limitation of this study is its focus on the medical personnel's opinion. The main limitation in the study by

Cameron *et al.* (2014) is that this research was conducted in a hospital, not in a telemedicine environment. Therefore, this study is beyond the scope of our study. The study by Xiong *et al.* (2012), presents a general model of the application of telemedicine. The proposal of the new tele medical triage system overcomes many limitations of the current paper triage tags. The authors acknowledged that the use of telemedicine for triage in future form casualty incidents is expected to grow considerably. Nevertheless, the study has some limitations, such as; firstly, the study investigated the scope of potential use of telemedicine in disaster response, however the authors did not propose a telemedicine model. Secondly, the study lacked an efficient algorithm to extract patients' vital data, as well as a triaging algorithm. Thirdly, the implementation of the model requires various suspicions about arrival patterns in local facilities, injury types, and portrayal of different treatment procedures and time-subordinate mortality.

Finally, although the simulation was able to describe the overall medical response process (Sequence of events and resource requirements), increasingly complex models are necessary to represent complicated issues arising in reality, such as the priority issue.

the model requires various suspicions about patients' entry designs in nearby offices, injury types, and portrayal of different treatment procedures and time-subordinate mortality. In the study by Mercadal *et al.* (2012), a double multi-agent architecture was suggested that empowers the triage of casualties in emergency situations and the programmed updating of their health condition. The proposed scheme combines Wireless Sensor Network (WSN), an electronic triage tag, and a double multi-agent system to accomplish low cost and efficiency of the system that is not infrastructure based. Initial results of the WSN roaming by agents are presented. However, the configuration of the triage tag based on the vital signs was not addressed. Moreover, the study focuses on routing among wireless nodes, the implementation includes many assumptions.

As shown above, several techniques have been used to triage patients out of the hospital. However, the latest and the most efficient one is the electronic triage technique (E-triage). According to the technical vision of the remote triage algorithm, the required triage algorithm should have the following specifications; firstly, triage patients remotely, wherever they are. Secondly, use new symptoms and features for triaging users. Thirdly, improve the guideline scores for chronic diseases through the use of sensory data and text. Finally, improve the

accuracy of triaging through an efficient mathematical theory so that the triaging level is selected by an efficient algorithm and not by human opinion.

Triage System Problems and Challenges

According to the description on the triaging process explained in the previous sections, triaging systems face many challenges. Rising number of patients in Emergency departments can lead to overcrowding, which frequently results an organizational problems. Triage aims to predict the severity of disease and organize patient flow (Azeredo *et al.*, 2015). Moreover, triaging is a complex decision-making process. Thus, several triage scales were designed to correspond to decision support systems (Seising and Tabacchi, 2013) and direct the triage nurse to a accurate decision. EDs around the world face continuous increases in patient visits (Seising and Tabacchi, 2013). Reducing the number of patients in the ED with NSCs leads to the effective management of medical resources. As the aging population will increase further, providing quality cost-efficiency care for these patients with several and complex medical conditions (Alemdar and Ersoy, 2010; Momani and Abualkishik, 2014; Seising and Tabacchi, 2013; Zughoul *et al.*). In the ED, the limited number of medical professionals requires the efficient exploitation of human resources to make the right decisions and leads to the need for develop automatic decision-making systems, such as those in the ED triaging process (Meri and Hasan, 2017; Seising and Tabacchi, 2013). Moreover, the number of triage nurses affects the number of patients that can be concurrently serviced at the triage desk (Yuan and Herbert, 2014).

Furthermore, Experts had argued that waiting patients should be reassessed occasionally (Claudio *et al.*, 2014; Derlet, 2002). Since, it has been recorded that patient's situation can be changed over time, which may lead changes in patient's triage status (Claudio *et al.*, 2011; Wolf *et al.*, 2006). Nevertheless, this may be impossible in an understaffed or an overcrowded ED (Derlet, 2002). In fact, many hospitals and HealthCare centers only have a verbal policy on patient reassessment however nurses do not act on it because, in many cases, overcrowding keeps them busy with new patients (Claudio *et al.*, 2011).

Conclusion

This paper provide a comprehensive insights on patient triaging systems, processes, standards, and guidelines. Besides, this paper shows the issues and challenges surround the

triaging process inside and outside the hospital. Furthermore, this review highlighted the use of electronic triage, which is a process of triaging patients using electronic devices, medical sensors, and computing algorithms based on medical guidelines. A state of the art of electronic triage systems were provided in order to show the current trends and the related problems. Some researchers made suggestions towards improving the present triage systems, however, the major drawback continues to revolve around its dependency on the vital indicators originally measured by ED staff that are not followed up periodically over the critical waiting period.

Reference

Alemdar, H., and Ersoy, C. (2010). Wireless sensor networks for healthcare: A survey. *Computer Networks, 54*(15), 2688-2710.

Azeredo, T. R. M., Guedes, H. M., de Almeida, R. A. R., Chianca, T. C. M., and Martins, J. C. A. (2015). Efficacy of the Manchester Triage System: a systematic review. *International emergency nursing, 23*(2), 47-52.

Bottrill, M. C., Joseph, L. N., Carwardine, J., Bode, M., Cook, C., Game, E. T., et al. (2008). Is conservation triage just smart decision making? *Trends in Ecology & Evolution, 23*(12), 649-654.

Brown, A.-M., and Clarke, D. E. (2014). Reducing uncertainty in triaging mental health presentations: Examining triage decision-making. *International emergency nursing, 22*(1), 47-51.

Bullard, M. J., Unger, B., Spence, J., and Grafstein, E. (2008). Revisions to the Canadian emergency department triage and acuity scale (CTAS) adult guidelines. *Cjem, 10*(2), 136-151.

Cameron, A., Rodgers, K., Ireland, A., Jamdar, R., and McKay, G. A. (2014). A simple tool to predict admission at the time of triage. *Emergency Medicine Journal*, emermed-2013-203200.

Christensen, D., Jensen, N. M., Maaloe, R., Rudolph, S. S., Belhage, B., and Perrild, H. (2011a). Low compliance with a validated system for emergency department triage. *Dan Med Bull, 58*(6), A4294.

Christensen, D., Jensen, N. M., Maaløe, R., Rudolph, S. S., Belhage, B., and Perrild, H. (2011b). Nurse-administered early warning score system can be used for emergency department triage. *Dan. Med. Bull, 58*, A4221.

Claudio, D., Bravo-Llerena, W., Okudan, G. E., and Freivalds, A. (2011). *Waiting in the Emergency Department: Dynamics of Patients' Vital Signs.* Paper presented at the IIE Annual Conference. Proceedings, 1.

Claudio, D., Kremer, G. E. O., Bravo-Llerena, W., and Freivalds, A. (2014). A dynamic multi-attribute utility theory–based decision support system for patient prioritization in the emergency department. *IIE Transactions on Healthcare Systems Engineering, 4*(1), 1-15.

Derlet, R. W. (2002). Triage and ED overcrowding: two cases of unexpected outcome. *The California Journal of Emergency Medicine, 3*(1), 8.

Farrohknia, N., Castrén, M., Ehrenberg, A., Lind, L., Oredsson, S., Jonsson, H., et al. (2011a). Emergency department triage scales and their components: a systematic review of the scientific evidence. *Scandinavian journal of trauma, resuscitation and emergency medicine, 19*(1), 42.

Farrohknia, N., Castrén, M., Ehrenberg, A., Lind, L., Oredsson, S., Jonsson, H., et al. (2011b). Emergency department triage scales and their components: a systematic review of the scientific evidence. *Scand J Trauma Resusc Emerg Med, 19*(42), 1-13.

Gao, T., and White, D. (2006). *A next generation electronic triage to aid mass casualty emergency medical response.* Paper presented at the AMIA Symposium Proceedings, 6501-6504.

Gilboy, N., Travers, D., and Wuerz, R. (1999). Re-evaluating triage in the new millennium: a comprehensive look at the need for standardization and quality. *Journal of Emergency Nursing, 25*(6), 468-473.

Godfrey, B., AVNER, L. I., LIU, C. C., TANABE, D., SIMMONS, J. L., ZELINKA, M., et al. (2000). Emergency Medical Guidelines, Third Edit. Sunshine Act of Florida. p.245.

Gräff, I., Goldschmidt, B., Glien, P., Bogdanow, M., Fimmers, R., Hoeft, A., et al. (2014). The German Version of the Manchester Triage System and its quality criteria–first assessment of validity and reliability. *PLoS One, 9*(2), e88995.

Hamamoto, J., Yamase, H., and Yamase, Y. (2014). Impacts of the introduction of a triage system in Japan: a time series study. *International emergency nursing, 22*(3), 153-158.

Innes, K., Plummer, V., and Considine, J. (2011). Nurses' perceptions of their preparation for triage. *Australasian Emergency Nursing Journal, 14*(2), 81-86.

Jentsch, M., Ramirez, L., Wood, L., and Elmasllari, E. (2013). *The reconfiguration of triage by introduction of technology.* Paper presented at the Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services, 55-64.

Jorma, J., Heli, L., Janne, E., and Ville, H. (2013). Experiences of Using a Mobile RFID-Based Triage System. *Journal of Aeronautics & Aerospace Engineering, 2013*.

Krupinski, E. A., and Bernard, J. (2014). *Standards and guidelines in telemedicine and telehealth.* Paper presented at the Healthcare, 74-93.

La Vonne, A. D., Zun, L. S., and Burke, T. (2015). Comparison of Canadian triage acuity scale to Australian Emergency Mental Health Scale triage system for psychiatric patients. *International emergency nursing, 23*(2), 138-143.

Mercadal, E., Robles, S., Martí, R., Sreenan, C. J., and Borrell, J. (2012). Double multiagent architecture for dynamic triage of victims in emergency scenarios. *Progress in Artificial Intelligence, 1*(2), 183-191.

Meri, A., and Hasan, M. K. (2017). Assessing the Determinants of Cloud Computing Services for Utilizing Health Information Systems: A Case Study. *International Journal on Advanced Science, Engineering and Information Technology, 7*(2), 503-510.

Momani, F., and Abualkishik, A. M. (2014). Factors Influencing Students' Intention to Adopt Mobile Blackboard. *International Journal of Science and Research, 3*(5), 29-32.

Ng, C.-J., Hsu, K.-H., Kuan, J.-T., Chiu, T.-F., Chen, W.-K., Lin, H.-J., et al. (2010). Comparison between Canadian triage and acuity scale and Taiwan triage system in emergency departments. *Journal of the Formosan Medical Association, 109*(11), 828-837.

Niswar, M., Ilham, A. A., Palantei, E., Sadjad, R. S., Ahmad, A., Suyuti, A., et al. (2013). Performance evaluation of ZigBee-based wireless sensor network for monitoring patients' pulse status.

Parenti, N., Reggiani, M. L. B., Iannone, P., Percudani, D., and Dowding, D. (2014). A systematic review on the validity and reliability of an emergency department triage scale, the Manchester Triage System. *International journal of nursing studies*, *51*(7), 1062-1069.

Pinto Júnior, D., Salgado, P. d. O., and Chianca, T. C. M. (2012). Predictive validity of the Manchester Triage System: evaluation of outcomes of patients admitted to an emergency department. *Revista latino-americana de enfermagem*, *20*(6), 1041-1047.

Sakanushi, K., Hieda, T., Shiraishi, T., Ode, Y., Takeuchi, Y., Imai, M., et al. (2013). Electronic triage system for continuously monitoring casualties at disaster scenes. *Journal of Ambient Intelligence and Humanized Computing*, *4*(5), 547-558.

Salman, O., Zaidan, A., Zaidan, B., Naserkalid, and Hashim, M. (2017). Novel Methodology for Triage and Prioritizing Using "Big Data" Patients with Chronic Heart Diseases Through Telemedicine Environmental. *International Journal of Information Technology & Decision Making*, 1-35.

Seising, R., and Tabacchi, M. E. (2013). *Fuzziness and Medicine: Philosophical Reflections and Application Systems in Health Care: A Companion Volume to Sadegh-Zadeh's Handbook of Analytical Philosophy of Medicine* (Vol. 302): Springer.

Wolf, L. D., Potter, P., Sledge, J. A., Boxerman, S. B., Grayson, D., and Evanoff, B. (2006). Describing nurses' work: combining quantitative and qualitative analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *48*(1), 5-14.

Xiong, W., Bair, A., Sandrock, C., Wang, S., Siddiqui, J., and Hupert, N. (2012). Implementing telemedicine in medical emergency response: concept of operation for a regional telemedicine hub. *Journal of medical systems*, *36*(3), 1651-1660.

Yuan, B., and Herbert, J. (2014). Context-aware hybrid reasoning framework for pervasive healthcare. *Personal and ubiquitous computing*, *18*(4), 865-881.

Zarabzadeh, A., O'Donoghue, J., O'Connor, Y., O'Kane, T., Woodworth, S., Gallagher, J., et al. (2013). Variation in health care providers' perceptions: decision making based on patient vital signs. *Journal of Decision Systems*, *22*(3), 168-189.

Zughoul, O., Jani, H. M., Shuib, A., and Almasri, O. Privacy and Security in Online Examination Systems. *Journal of Computer Engineering (IOSR-JCE) e-ISSN*, 2278-0661.

NEW DESIGN OF LOWER LIMB REHABILITATION ROBOTICS SYSTEM BASED ON BP NEURAL NETWORK

Dhurgham Aiham Kadhim Alshakarchi[1], Firas T. Al-Maliky[2]

Alkafeel University College , Computer Engineering Technique dept., Najaf /Iraq
E-mail(1): dhurgham.alshakarchi@alkafeeluc.edu.iq
E-mail(2): firas.almaliky@alkafeeluc.edu.iq , firasthair@yahoo.com

Abstract

In this research a new design for planar robot used for lower limb rehabilitation submitted to people who suffered from stroke. First the mechanical design is presented and discussed also the mathematical model of the planar surface is placed (torque and angle of planar). Then the control system build on ATM mega 250 (arduino) all data collected are processed on neural network which it was built based on number of patients' data . The results shows great flexibility for both treat stroked people and therapist

Keywords: artificial neural network, lower limb, brain stork, rehabilitation, BP neural *network*

Introduction

The concept of Neural Networks was inspired from biological nervous systems of human being Artificial Neural Networks (ANN) are calculative models for pattern recognition and machine learning. It is simple way for operate parallel elements that introduced to the Networks. The functions of ANN Artificial Neural Networks are greatly depended on the connections between parallel elements so anyone can adjust the connections of elements to accomplish a particular function. The process of modifying of connections is called learning of Artificial Neural Networks. so that a specific input/targets the Artificial Neural Networks can be trained to get outputs of the network that match or equal the targets.. The NN will be already trained for comparing the output and target and supplying information during that until output and target will be matched.

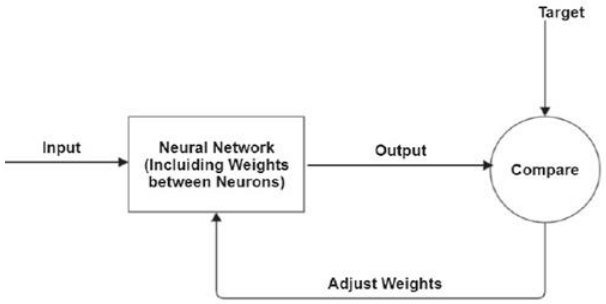


Fig.1: Neural Network block diagram

Literature review

Abdul-Basset A. AL-Hussein (2017) A composite PD and sliding mode neural network (NN)-based adaptive controller, for robotic manipulator trajectory tracking, is presented in this paper. The designed neural networks are utilized to sacrificial the robotics nonlinearities , also make more it effective this will improve the functionality of filtering errors depending on PD and sliding controller of neural system . the theorem used to establish system stability was Lyapunov theorem the pursuing error limitedness. The enhance Lyapunov tool is used to derive the neural network weights updating law. Erhan AKDOGAN (2016). a lower limb rehabilitation robot was modeled and controlled for upper limbs. The ret rotted robot system is able to do, active, active-assistive and passive curative exercises. On the other side, it works as active-assistive exercises by muscular activation. A network –established on human machine interface, which can support self-care, was created to control the robotic system. The robotics system interface contains valuation unit, are not able to use only torque and ROM, but also muscular activation of patients for the estimation results. The, estimation and therapy can be perceived by a system for both lower and upper limbs using parameters such as ROM, generated muscular activation, and torque of the limbs]. Erhan Akdog˘an & etal (2011). They use three degrees of freedom therapeutic exercise robot and demonstrate design and control of robot used in therapy for patient that required rehabilitation after a muscle disorder , spinal cord injury stroke , a surgical operation, or SCI. for controlling robot, a ‘‘Human–Machine Interface’ with a rule-based control structure was developed. The robot can perform all passive and active exercises as well as learning ability of specific exercise without the physiotherapist by using Human–Machine Interface. Further, if a patient interact against robot's manipulator during exercises, the flexibility of manipulator change the position depend on feedback data.

This benefit of manipulator flexibility will serve therapeutic exercise and physiotherapist in terms of motion capability. Chen JIA (2015) in This study feasibility of new a model to characterize nonlinearity of mechanical impedance of the ankle within a particular range of frequency and the root mean square (RMS) value of the Electromyography of muscles of ankle by using ANN. A lower limb rehabilitation robot Ankle bot was used to apply dummy -random mechanical disturbance to the ankle and explore the angular displacement of ankle to assessment data of mechanical impedance for ankle. Meanwhile, the surface Electromyography signals from the chosen muscles were observed and recorded. in this research final ANN models are built or constructed for two degrees of freedom. The results of analysis of the ANN model showed the feasibility of developing models with adequate accuracy and to define the mechanical impedance of the human ankle in terms of lower extremity muscles' EMG statistical properties.

Strategy physical therapy

The training are the backbone of treatment for people had foot drop caused by stoke. The benefit of training contain decrease spasticity, and increasing ability of muscle. The training are classified according to foot motion Kazem & etal :

class 1 & 2 : the patient will sit and contacting heel's edge to the table then trying rotate the ankle CW or CCW (roll motion) angle and then back to neutral position as shown in Figure (2). Do it ten times and then revers the operation Figure (2).

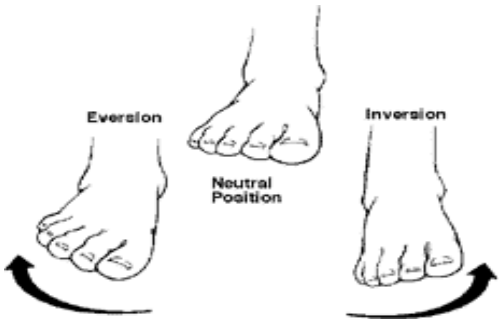


Fig.2: Exercises 1 and 2

class 3 & 4 : the patient will sit and contacting heel's edge to the table ”then trying translate the ankle up or down angle and then back to neutral position as shown in Figure (3) Do it ten times and then revers the operation

.

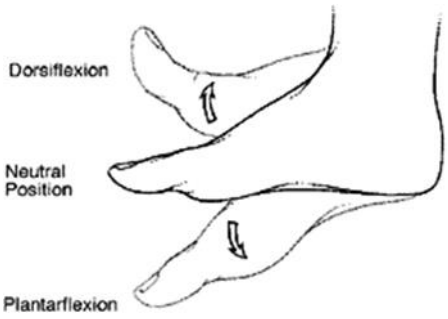


Fig.3: Exercises 3 and 4

Class 5 & 6: the patient will sit and contacting heel's edge to the table then trying rotate the ankle CW or CCW (yaw motion) angle and then back to neutral position as shown in Figure (4). Do it ten times and then revers the operation. Do this exercise ten times a day.

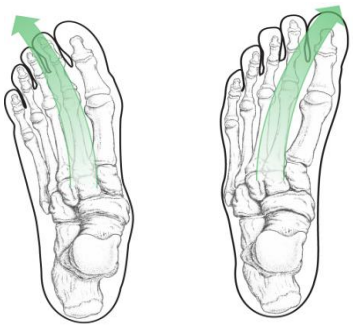


Fig.4: Exercise 5, 6

A base Cartesian coordinate frame XYZ is fixed at the center of the base platform with the Z-axis pointing vertically upward and the X- axis pointing towards the ball joint $\boldsymbol{B_i}$**.** Similarly, a coordinate

frame xyz is assigned to the center of the upper platform, with the z- axis normal to the platform and the x-axis pointing towards the ball joint P_i. Hence, the coordinates of the pin joints in XYZ frame are

$$d_1 = \begin{bmatrix} R \\ 0 \\ 0 \end{bmatrix} \qquad d_2 = \begin{bmatrix} \frac{-1}{2}R \\ \frac{\sqrt{3}}{2}R \\ 0 \end{bmatrix} \qquad d_3 = \begin{bmatrix} \frac{-1}{2}R \\ \frac{-\sqrt{3}}{2}R \\ 0 \end{bmatrix} \quad \ldots\ldots\ldots\ldots (1)$$

The position of the center of moving platform is

$$P_0 = \begin{bmatrix} 0 \\ h \\ 0 \end{bmatrix} \qquad \ldots\ldots (2)$$

when h is the hight of moving platform respect to the base

Surface Motion of proposed Robot .

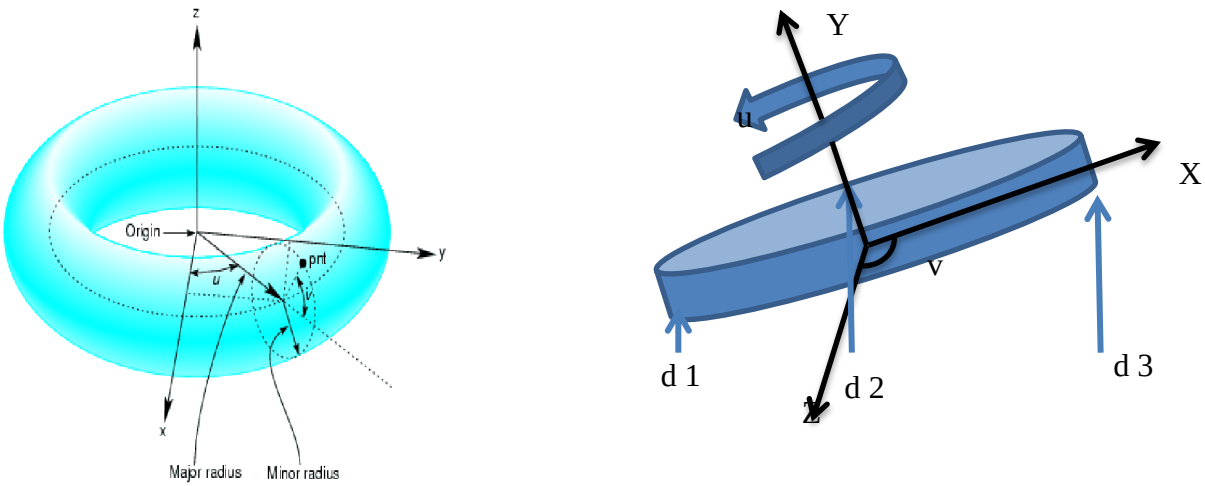


Fig. 5 : shows Cartesian coordination of

Suppose any point on surface of robot the parametric equation which represents the movement of the lower limb for patient Cartesian equation for a torus takes the form

(if any of x, y and z are interchanged, a similar, rotated torus plate will result). It can be obtain the parametric equation by setting $z=c\sin v$ as in the diagram above The Cartesian equation becomes as Joel R. Hass-Thomas (2017) .

$$(b-\sqrt{x^2+y^2})^2=c^2-c^2\sin^2 v=c^2(1-\cos^2 v)=c^2\cos^2 v \quad \ldots\ldots (3)$$

$$c^2\sin^2 v+(b-\sqrt{x^2+y^2})^2=c^2 \quad \ldots\ldots(4)$$

Square rooting both sides gives $b-\sqrt{x^2+y^2}=c\cos 2v$ …..(5)

Hence $\sqrt{x^2+y^2}=b-c\cos 2v$ Now recognize that for constant v, this equation defines a circle around the z – axis of radius $b-c\cos v$, so set $x=(b-c\cos 2v)\cos u$ and $y=(b-c\cos 2v)\sin u.$

Table1: The practical actuated length with exercises

Exercise	d 1		d 2		d 3	
	Length	Change	Length	Chan	Length	Change (mm)
0	0.3027	-	0.3027	-	0.3027	-
(1)	0.3027	0	0.3357	33	0.2697	-33
(2)	0.3027	0	0.2787	-24	0.3267	24
(3)	0.2297	-73	0.3387	36	0.3387	36

(4)	0.3487	46	0.2797	-23	0.2797	-23

Differentiating the expression gives the velocity relations as Schilling, R. J. (1996)

$d_i\, w_i \times s_i + \dot{d}_\iota\, s_i = w_p \times p_i$ We can get linear velocity by multiplying dot product with s_i as:

$$[\dot{D}] = (p_i \text{ x } S_i)\begin{bmatrix}\dot{\alpha}\\ \dot{\beta}\\ 0\end{bmatrix} \quad \ldots\ldots \quad (6) \qquad [\dot{D}] = [J]^{-1}\begin{bmatrix}\dot{\alpha}\\ \dot{\beta}\\ 0\end{bmatrix} \quad \ldots\ldots (7)$$

Where $\dot{\alpha}, \dot{\beta}$ is the angular velocity about x and y axis respectively as Figure (5) The linear velocities of link can be represented as :$[\dot{D}] = \begin{bmatrix}\dot{d}_1\\ \dot{d}_2\\ \dot{d}_3\end{bmatrix}$ ……..(8)

Training robotics systems BP neural network algorithm analysis

The aim behind controlling BP algorithm is to make plate of upper robotics system to do all exercise to be by the therapist in the same time with all constraint. The training of neural network, lots of keep scenes constructed in middle size match platform, data, totally 400 groups successful pass samples data recorded to BP neural network training; before training, normalization handling of samples data should be done, from which initial value of weight adopt random other learning parameters $\alpha = 0.5 = \eta$, through nearly 1 hour learning, network tends to convergence state**,** the errors of network error is roughly equal to 0.001, its convergence process

Results

Figure (6) shows the training error of a network with one hidden layer trained on a small data segment. The network has properly learned how to produce steering angles that match the targets (Figure 6-a). However, notice the paths in Figure (6-b) and (6-c). The blue paths show the simulated result of a robotics during using our neural network to produce desired velocity and angles of platform commands. After 1000 iterations, the network has learned a reasonable

concept of obstacle avoidance for this situation. After 200 iterations, though, the network is over fitted to the target line

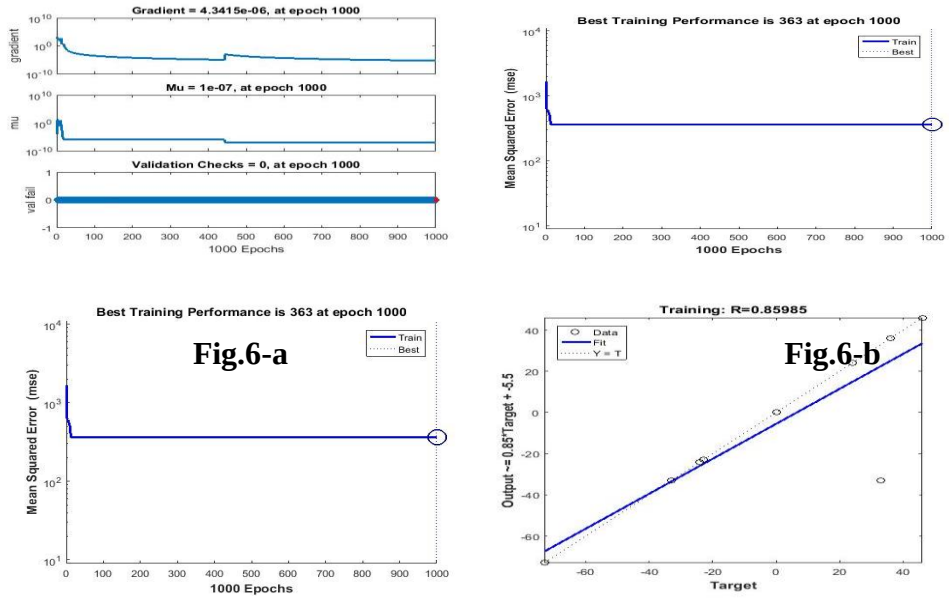


Fig.6-a **Fig.6-b**

Fig. 6-c **Figure 6-d**

Fig.6: training results of Neural Network

Conclusion

In this study, a therapy training robot was presented and controlled for lower-limb rehabilitation. The robot manipulator was controlled through a (Arduino card). Manual and self-assistance training performed by the patient and also all other active and passive trainer can be supported by this devices with this interface. It was explained through experiments done hospital subjects that the RM can perform the necessary training as well as mimic manually exercises performed by the patience .

References

[1] AL-Hussein, A. B. A. (2017). Neural Network-Based Adaptive Control of Robotic Manipulator Application to a Three Links Cylindrical Robot. Iraqi Journal for Electrical & Electronic Engineering.

[2] AKDOĞAN, E. (2016). Upper limb rehabilitation robot for physical therapy: design, control, and testing. Turkish Journal of Electrical Engineering & Computer Sciences, 24(3).

[3] Akdoğan, E., & Adli, M. A. (2011). The design and control of a therapeutic exercise robot for lower limb rehabilitation: Physiotherabot. Mechatronics, 21(3), 509-522.

[4] Jia, C. (2015). Mechanical impedance of ankle as a function of electromyography signals of lower leg muscles using artificial neural network (Doctoral dissertation, Michigan Technological University).

[5] Dzahir, M. A. M., & Yamamoto, S. I. (2014). Recent trends in lower-limb robotic rehabilitation orthosis: Control scheme and strategy for pneumatic muscle actuated gait trainers. Robotics, 3(2), 120-148.

[6] Chowdhury, R. H., Reaz, M. B., Ali, M. A. B. M., Bakar, A. A., Chellappan, K., & Chang, T. G. (2013). Surface electromyography signal processing and classification techniques. Sensors, 13(9), 12431-12466.

[7] Kazem, B. I., Morad, A. P. D. A., & Hasan, K. M. 3-DOF Parallel robotics System for Foot Drop therapy using Arduino.

[8] Tsai, L. W. (1999). Robot analysis: the mechanics of serial and parallel manipulators. John Wiley & Sons.
[9] Schilling, R. J. (1996). Fundamentals of robotics: analysis and control. Simon & Schuster Trade.

[10] J. Dul and G. E. Johnson, "A kinematic model of the human ankle," Journal of Biomedical Engineering, vol. 7, pp. 137-143, 1985.

[11] S. Siegler, J. Chen, and C. D. Schneck, "The three-dimensional kinematics and flexibility characteristics of the human ankle and subtalar joints- Part I: Kinematics," Journal of Biomechanical Engineering, vol. 110, pp. 364-373, 1988.

[12] C. S. Parenteau, D. C. Viano, and P. Y. Petit, "Biomechanical properties of human cadaveric ankle- subtalar joints in quasi-static loading," Journal of Biomechanical Engineering, vol. 120, pp. 105-111, 1998.

[13] George B. Thomas Jr., Maurice D. Weir, Joel R. Hass-Thomas' "Calculus Early Transcendental" 14th edition , 2017 .

[14] R. P. Kleipool and L. Blankevoort, "The relation between geometry and function of the ankle joint complex: A biomechanical review," Knee Surgery, Sports Traumatology, Arthroscopy, vol. 18, pp. 618-627.

[15] Bullen, D. (n.d.). Tools To Treat Foot Drop on ADVANCE for Physical Therapy & Rehab Medicine. ADVANCE for Physical Therapy & Rehab Medicine. Retrieved February 28, 2013, from http://physical- therapy.advanceweb.com/Features/Articles/Tools-To-Treat-Foot-Drop.aspx

[16] Richard Sandor, M.D., Camino Medical Group, Orthopedics Scott Brone, P.T., C.S.C.S., Physical Therapy, 2007
[17] KOK-MENG LEE AND DHARMAN K. SHAH " Dynamic Analysis of a Three-Degrees-of-Freedom In-Parallel Actuated Manipulator",IEEE Journal of robotics and automation,vol 4,No.3,june1988.

[18] Lung-Tsia"robot analysis:mechanics of the serial and parallel manipulator", 3rd edition,1999.

[19] ROBERT J.SCHILLINH " FUNDAMENTALS OF ROBOTICS ANALYSIS AND CONTROL"

[20] Vala´sˇek, M., Bauma, V., Sˇika, Z., Belda, K., and P´ısˇa, P. Design-by-optimization and control of redundantly actuated parallel kinematics sliding star. Multibody Syst. Dyn., 2005, 14(3/4), 251–267.

[21] G. Liu, J. Gao, H. Yue, X. Zhang and G. Lu, "Design and Kinematics Analysis of Parallel Robots for Ankle Rehabilitation," Proceedings of the IEEE/RSJ Int. Conf on Intelligent Robots and Systems, Beijing, pp. 253-258, October 9 – 15, 2006.

[22] Michael McRoberts,"Beginning Arduino", ISBN-13 (electronic): 978-1-4302-5017-3, second edition, 2013.

[23] Patrick Di Justo and Emily Gertz," Atmospheric Monitoring with Arduino", ISBN: 978-1-449-33814-5, First Edition,November 2012.

[24] Simon Monk, "30 Arduino Projects for the Evil Genius", ISBN: 978-0-07-174134-7, 2010.

[25] Online, [Cited: Sep 10, 2014.], available at: http://arduino.cc/en/Main/arduinoBoardMega2560

[26] Online, [Cited: Sep 10, 2014.], available at: www.hacktronics.co.in/blog/1-using-the-l298-motor-driver-to-control-the-robotics-vehicle.

[27]I. BoldeaSyed A. Nasar,"Linear Electric Actuators and Generators," Cambridge University press, pp. 33–60, 1997.

[28] Hamid D. Taghirad , "Parallel Robots: Mechanics and Control," First Edition, 2013.

[29] John J. Craig, "Introduction to Robotics, Mechanical and control," Third Edition, 2005.

[30] Bl.Theraja, Kl. Theraja, "text book of Electrical Technology," 26th Edition, 2003.

Specification parameters of WLAN performance with MatLab Simulink model of IEEE 802.11.

Dr. faeza abas abid
Middle technical university, institute of technology, electronic technical department
Corresponding author : **Dr_faeza@yahoo.com**

Abstract

Wireless networks have gained popularity, providing users flexibility and mobility in accessing information. The IEEE 802.11 Wireless Local Area Network (WLAN) standard has become the dominant architecture in practice. In this paper IEEE 802.11 standard implemented in Mat lab Simulink, where Allowing to modify beacon frame parameters such as beacon interval, SSID, supported rates, etc. also include or exclude the CF parameter set element from the generated beacon frame, DBPSK and DQPSK modulation techniques BER performance included in this paper to illustrate the performance of the model in AWGN and Rayleigh channels.

Keywords: IEEE stander, IEEE 802.11, MAC layer, DBPSK, DQPSK.

محكاة بواسطة الماتلاب لأداء محددات عناصر الشبكة الاسلكية محلية المساحة لموديل IEEE 802.11

أ.م.د. فائزة عباس عبد
الجامعة التقنية الوسطى / معهد تكنولوجيا بغداد / قسم التقنيات الالكترونية

الخلاصة

الشبكات الاسلكية اكتسبت شعبية واسعة وذلك لتوفير المرونة للمستخدمين والحركة في مداخل المعلومات.
المعياري IEEE 802.11 المعياري اصبح المعمارية المسيطرة في الجانب العملي، في هذا البحث IEEE 802.11 صمم ونفذ بواسطة برنامج المحكاة الماتلاب حيث يسمح هذا الموديل بتحديث عناصر اطار beacon ، مثل الفترة ،تم beacon من العنصر المتولد في اطار CF ، معدلات المجهزة، الخ. وكذلك ظمنت او اقصيت عناصر SSID فحص اداء تقنيات التظمين DBPSK و DQPSK في الموديل تحت تاثير القنوات AWGN و Raylieg.

Introduction

Wireless networks are emerging as a significant aspect of networking; wireless local area networks WLANs, Bluetooth, and cellular systems have become increasingly popular in the business and computer industry, with consequent security issues. WLANs, especially the Institute of Electrical and Electronics Engineers (IEEE) 802.11 networks, are becoming common access networks in private and public environments. The freedom of movement and

simplicity in its implementation have made WLANs popular in the home and businesses sectors, as well as hotspots 1 such as airports and cafes. The increasing availability of, and therefore increasing reliance on, wireless networks makes it extremely important to maintain reliable and secure communications in the wake of network component failures or security breaches. However, recent news reports on a number of attacks against wireless networks, especially WLANs, have alarmed wireless adopters, developers, and intended users. The broadcast nature of wireless communication links makes them unique in their vulnerability to security attacks and their susceptibility to intentional threats. Organizations that want to deploy a secured WLAN infrastructure are challenged by the flaws in the existing wireless mechanism design, such as the wired equivalent privacy (WEP) protocol. WLANs provide greater flexibility and scalability than traditional LANs. Unlike a wired LAN, which requires a wire to access the network, a WLAN facilitates network transmissions of data from computers and other components through an access point (AP). An AP typically provides a range (cell or area coverage) of 100 metres. IEEE 802.11 is an international standard providing transmission speeds ranging from 1 Mbps to 54 Mbps in either the 2.4 GHz or 5 GHz frequency bands. The 802.11b is the dominant WLAN technology at present [WECA, 2001b], and provides an expected data throughput of 5.5 Mbps . High performance radio LAN is a European Telecommunications Standards Institute (ETSI) standard operating in the 5 GHz frequency band; Hiper LAN/1 has a transmission speed of 19 Mbps, while Hiper LAN/2 operates at 54 Mbps. Hiper LAN/2 supports quality of service (QoS) and is based on an infrastructure topology, whereas Hiper LAN/1 is more suitable for forming ad-hoc networks.

IEEE 802.11 Standard

The standard provides three *physical* (PHY) layers and one *medium access control* (MAC) layer for deploying wireless communication in local networks (Fig. 1). As for the *logical link control* (LLC) layer, there is no difference between wireless (802.11) and wired (802) LANs, such as the IEEE 802.3 Ethernet network. The MAC protocol provides two service types: asynchronous using the *distributed coordination function* (DCF), and synchronous using the *point coordination function* (PCF) that is contention-free.[1]

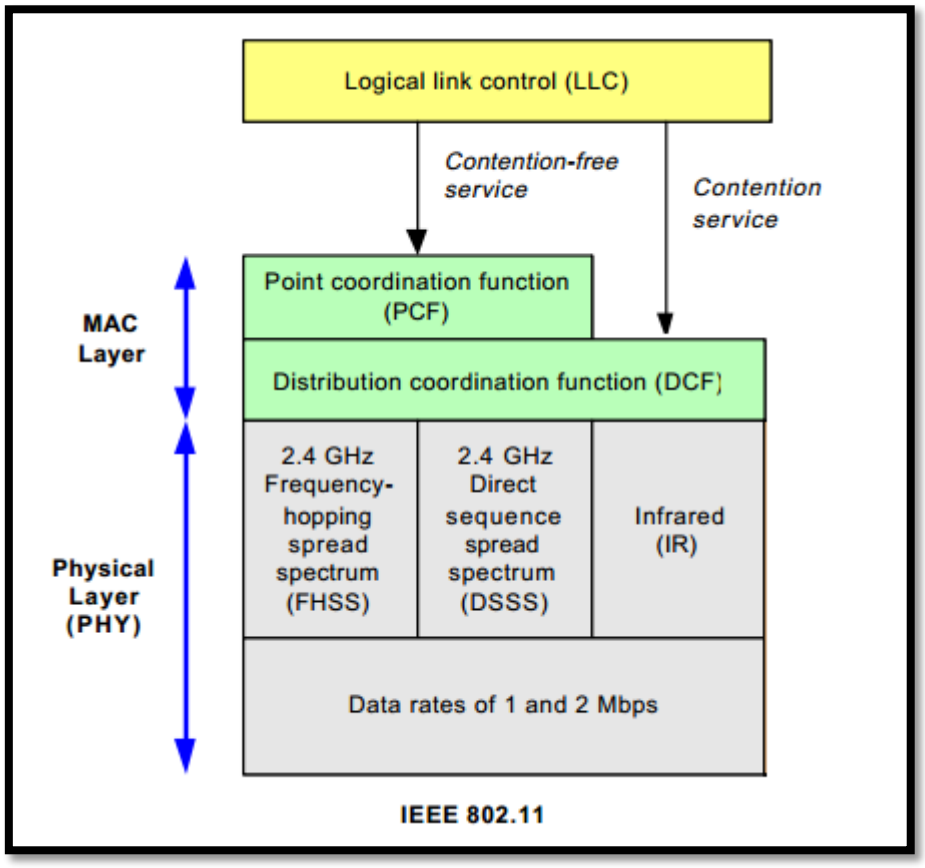


Fig.1: The 802.11 Protocol Stack

The 802.11 standard is a family of specifications originally providing a 1 to 2 Mbps data transmission rate using either the frequency-hopping spread spectrum (FHSS) or direct sequence spread spectrum (DSSS), After revisions, the standard includes 802.11a, operating in a 5 GHz frequency band at 54 Mbps, and 802.11b and IEEE 802.11g, operating in a 2.4 GHz frequency band at speeds of 11 Mbps and 54 Mbps, respectively. [2,4]

The 802.11 standard takes advantage of radio spectrum technologies, allowing multiple users to share the radio frequencies without end- user licenses. Specifically, it makes uses of the 2.4 GHz Industrial, Scientific, and Medical band (ISM) band for 802.11 and 802.11b networks, and the 5 GHz Unlicensed National Information Infrastructure (UNII) band for 802.11a-based networks. The International Telecommunication Union (ITU) defines both bands. However, interference issues remain, especially in the 2.4 GHz band; if the technology interferes with an authorised operation such as an airline radio frequency, it will cease to operate. In addition, there

is no protection from other technologies, such as Bluetooth, accessing 802.11 frequencies.A portal is a logical point, which is required to integrate the 802.11 architecture with existing wired LANs facilitating the transmission. An AP with software implementation can offer the portal service. [5,6]

MAC Layer

The 802.11 specifications provide asynchronous (DCF) and contention-free (PCF) services. The asynchronous service is always available whereas the contention-free service is optional. DCF implements the basic access method of the 802.11 MAC protocol, or path sharing. The PCF provides contention- free service, which implements a polling access method. It uses a point coordinator (PC), usually the AP,
which cyclically polls stations, giving them the opportunity to transmit. Thus the access priority provided by a PCF may be utilized to create a contention-free access method. The PC controls the frame transmissions of the stations in order to eliminate contention for a limited period of time. Unlike the DCF, the implementation of the PCF is not mandatory. Furthermore, the PCF itself relies on the asynchronous service
provided by the DCF (Fig.1). All physical layers support one common MAC layer. Task Group 802.11e focuses on enhancing the MAC layer for QoS.[7,8]

Physical Layers

The physical layer processes data to and from radio signals over the airwave. In other words, it handles the transmission of the frame via the air interface. The standards define three alternative physical layers (Figure 2-2):

1. Frequency-hopping spread spectrum (FHSS)
2. Direct sequence spread spectrum (DSSS)
3. Infrared (IR)

The first three physical layers belong to radio spread spectrum technology that operates in the 2.4 GHz band, while OFDM operates in the 5 GHz band. IR operates in the 300-428 GHz band and operates at a slow speed with line-of-sight connection. IR has become a legacy protocol thus will not be addressed in this discussion [8].

1. FHSS

FHSS modulates data signals with a carrier signal that hops from one frequency to another, using time as its measurement, over a wide range of frequencies. The carrier frequency (between 2.4 and 2.483 GHz) is changed periodically to avoid collisions. A collision occurs only when both a narrowband system and the spread spectrum signalsare transmitting at the same frequency simultaneously. A hopping code is used to decide the order of data transmission and which frequency to hop to. FHSS provides a maximum data transmission speed of 2 Mbps [9].

2. DSSS

DSSS combines a data signal at the sending station with a higher data rate bit sequence, known as the chipping code or processing gain. This chipping code reduces interferences by dividing the user data according to a spreading ratio, enabling a faster data transmission rate of 11 Mbps. It sets a specific string of bits to be sent for each data bit. A redundant bit pattern is included in the chipping code to increase resistance to interference.

WLAN IEEE 802.11 Simulink Model

IEEE 802.11 have been implemented with Matlab Simulink as shown in fig.2, where the model has three main parts: Model Parameters, where the information can be selected and sent within the beacon frame and the channel parameters, 802.11 System, which is formed of the transmitter, channel, and the receiver, Results, where there is a viewing to the most important received information, such as CRC flags, packet type, and received SSID.

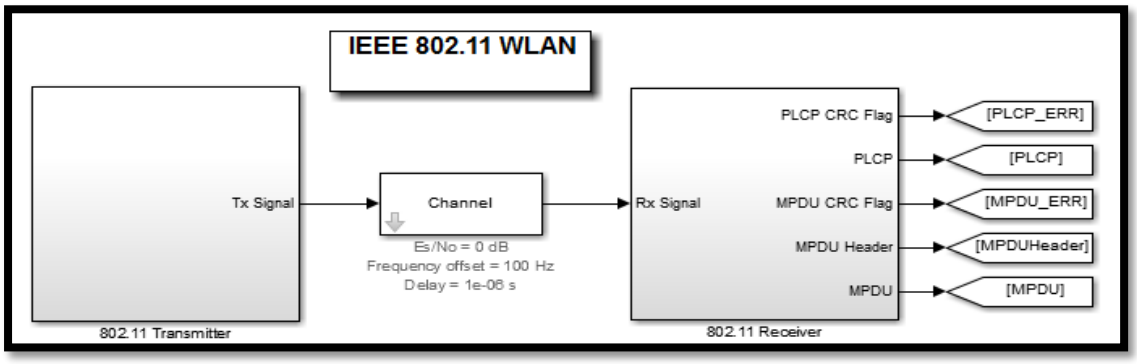


Fig. 2: IEEE 802.11 Simulink model

The transmitter implements a lightweight MAC sublayer and physical layer that constructs the PHY layer frame as shown in fig. 3. It transmits each PHY frame using several consecutive channel frames.

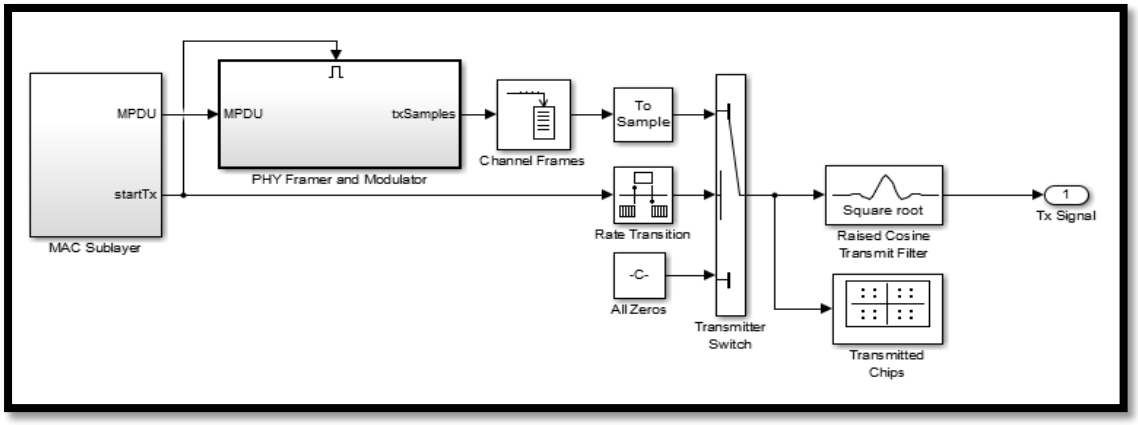


Fig. 3: the transmitter block in the IEEE 802.11 Simulink Model.

The MAC sublayer constructs the beacon frame, which is a type of a management frame. Fig.4 shows the management frame format. This frame is also called a MAC protocol data unit (MPDU).

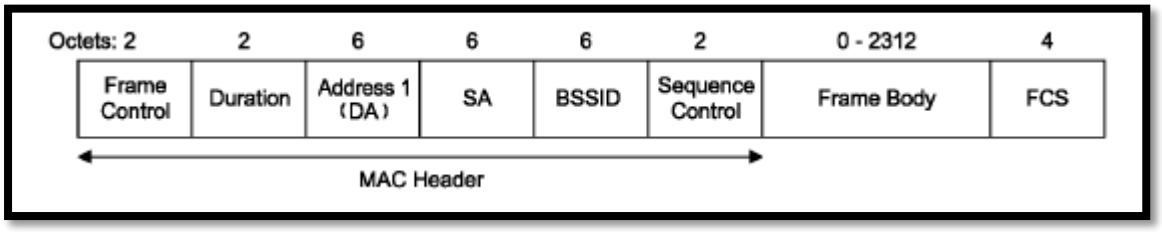


Fig. 4: MAC sublayer beacon frame

The implements a beacon frame with a frame body that contains the following fields: Timestamp, Beacon interval, Capability, SSID elements, Supported rates element, Direct sequence (DS) parameter set element, Contention Free (CF) parameter set element (optional) The transmitter turns on periodically to transmit a beacon frame. The beacon interval parameter determines the transmission period.

The PHY Framer and Modulator creates as shown in figure 5 and 6, the PLCP protocol data unit (PPDU) by adding the physical layer convergence procedure (PLCP) preamble and header to the MPDU. The PLCP preamble contains 128 bits of ones (SYNC), which are later

scrambled. The receiver determines the presence of a PPDU frame using this SYNC signal. The LENGTH field of the PLCP header determines the MPDU frame length. PLCP header also contains a 16-bit CRC.

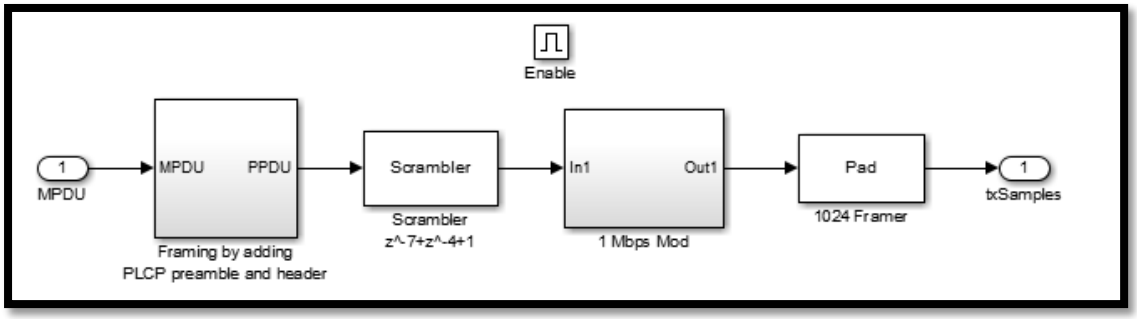


Fig. 5: PHY Framer and Modulator block

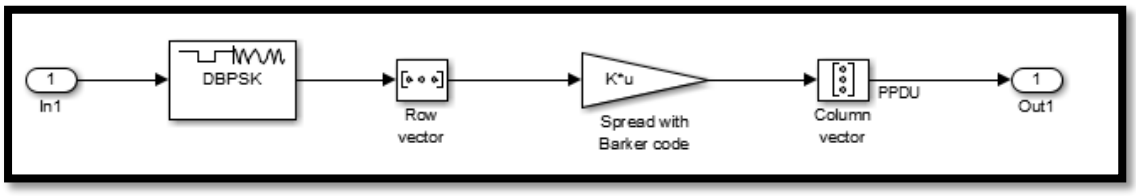


Fig.6: 1Mbps Mod block

The transmitter scrambles the PPDU frame, modulates using differential binary phase shift keying (DBPSK or DQPSK) at a rate of 1 Mbps, and applies spreading using a length 11 Barker code. This subsystem also pads extra chips to the spread symbols to reach a 1024-bit maximum PPDU length. This way, the system forces the MAC layer to work at a period of 1024 microseconds, which is a time unit (TU).

The transmitter emits 128 modulated symbols (1408 chips), which is a channel frame, at a time. Modulated symbols pass through a square root raised cosine pulse shaping filter. Pulse shaped symbols are sent through an AWGN channel that also applies a frequency offset and delay.

A **Barker code** or **Barker sequence** is a finite sequence of *N* values of +1 and −1, with the ideal autocorrelation property, such that the off-peak (non-cyclic) autocorrelation coefficients

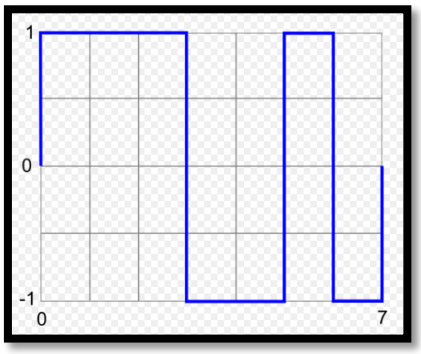


Fig. 7: Graphical representation of a Barker-7 code

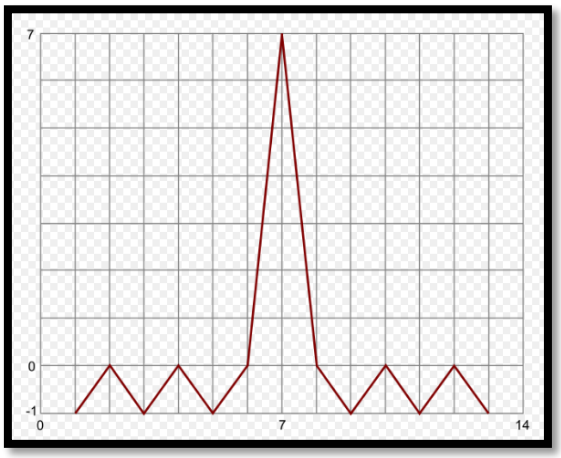


Fig. 8: Autocorrelation function of a Barker-7 code

A Barker code has a maximum autocorrelation sequence which has side lobes no larger than 1. It is generally accepted that no other perfect binary phase codes exist [3]. (It has been proven that there are no further odd-length codes [3], nor even-length codes of $N < 10^{22}$.) [3,4].

Barker codes of length N equal to 11 and 13 are used in direct-sequence spread spectrum and pulse compression radar systems because of their low autocorrelation properties (The side lobe level of amplitude of the Barker codes is 1/N that of the peak signal).[3,4] A Barker code resembles a discrete version of a continuous chirp, another low-autocorrelation signal used in other pulse compression radars.

The positive and negative amplitudes of the pulses forming the Barker codes imply the use of bi-phase modulation or binary phase-shift keying; that is, the change of phase in the carrier wave is 180 degrees.

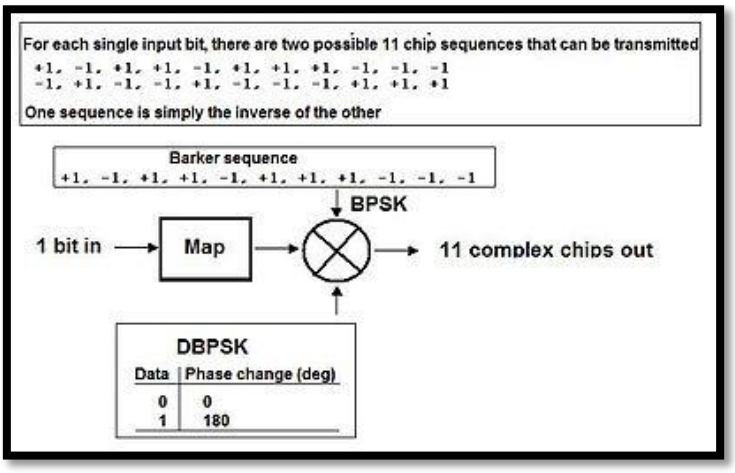


Fig. 9: Barker code used in BPSK modulation

Similar to the Barker codes are the complementary sequences, which cancel side lobes exactly when summed; the even-length Barker code pairs are also complementary pairs. There is a simple constructive method to create arbitrarily long complementary sequences.

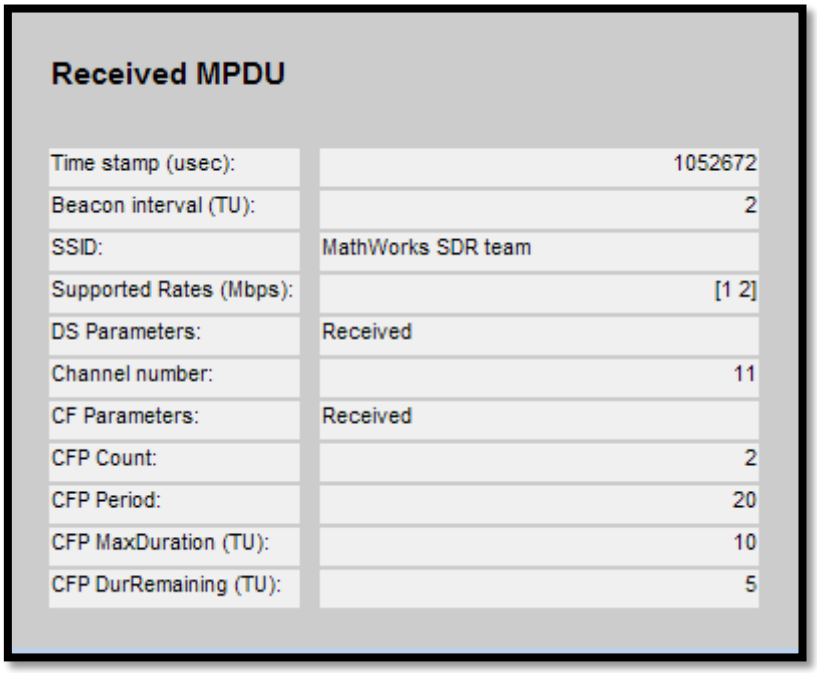
For the case of cyclic autocorrelation, other sequences have the same property of having perfect (and uniform) side lobes, such as prime length Legendre sequences and maximum length sequences (MLS). Arbitrarily long cyclic sequences can be constructed.

In wireless communications, sequences are usually chosen for their spectral properties and for low cross correlation with other sequences likely to interfere. In the 802.11b standard, an 11-chip Barker sequence is used for the 1 and 2 Mbit/sec rates. The value of the autocorrelation function for the Barker sequence is 0 or −1 at all offsets except zero, where it is +11. This makes for a more uniform spectrum, and better performance in the receivers [3,4].

Results

Simulink model that implemented in this paper allowing to modify beacon frame parameters such as beacon interval, SSID, supported rates, etc. also include or exclude the CF parameter set element from the generated beacon frame as shown in table 1. seting the noise level (Es/No), delay, and frequency offset introduced by the channel.

Table 1: Received MPDU



Received MPDU

Time stamp (usec):	1052672
Beacon interval (TU):	2
SSID:	MathWorks SDR team
Supported Rates (Mbps):	[1 2]
DS Parameters:	Received
Channel number:	11
CF Parameters:	Received
CFP Count:	2
CFP Period:	20
CFP MaxDuration (TU):	10
CFP DurRemaining (TU):	5

The simulation also displays the scatter plots of transmitted and received chips, and the despread symbols as shown in figure 10-12 respectively. You can see the rotating scatter plot due to the frequency offset. The differential demodulator is able to track the phase change and correctly demodulate the symbols.

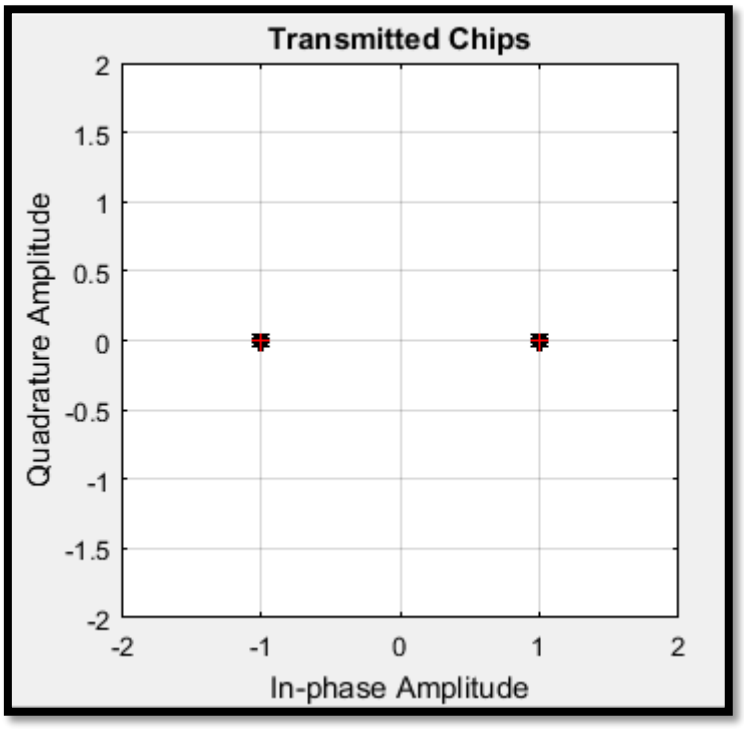


Fig. 10: Scatter plot of transmitted chip

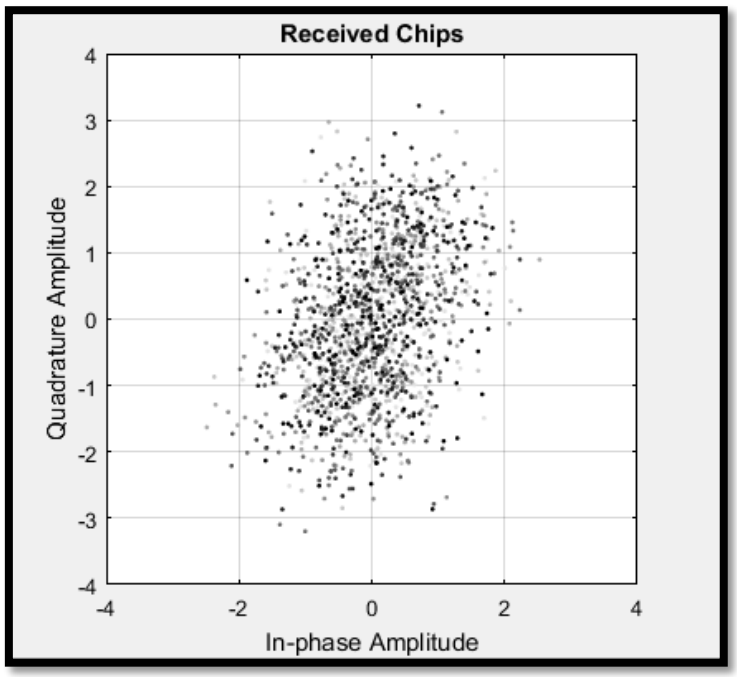


Fig. 11: Scatter plot of received chip

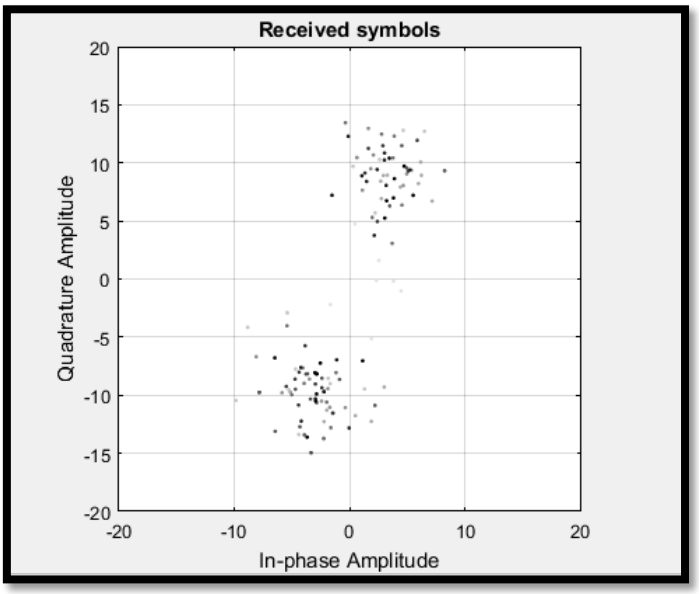


Fig. 12: the despread symbols.

Modulation techniques that used in the Simulink model which are DBPSK and DQPSK are tested under AWGN channel, figure 13 shown BER parameter of both modulation techniques where DBPSK shows butter BER with 2dB gain in E_b/N_o. while in Rayliegh channel both technique shows same BER as shown in figure 13.

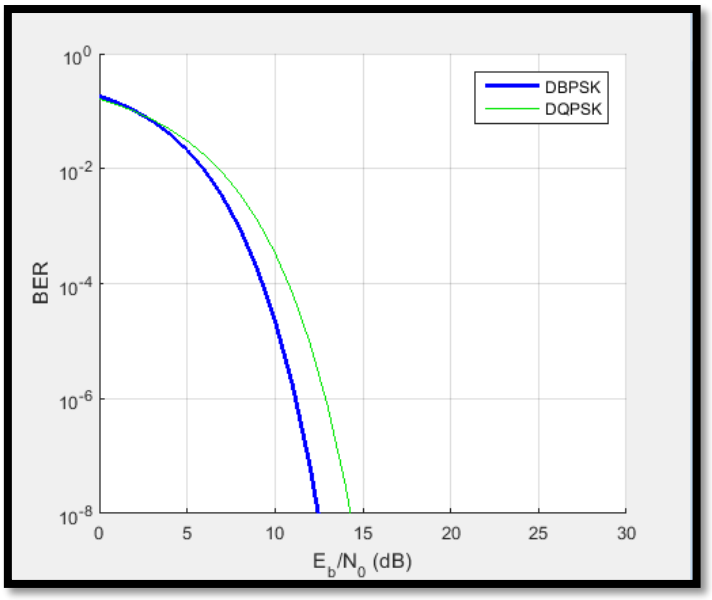


Fig. 12: DBPSK and DQPSK in AWGN channel

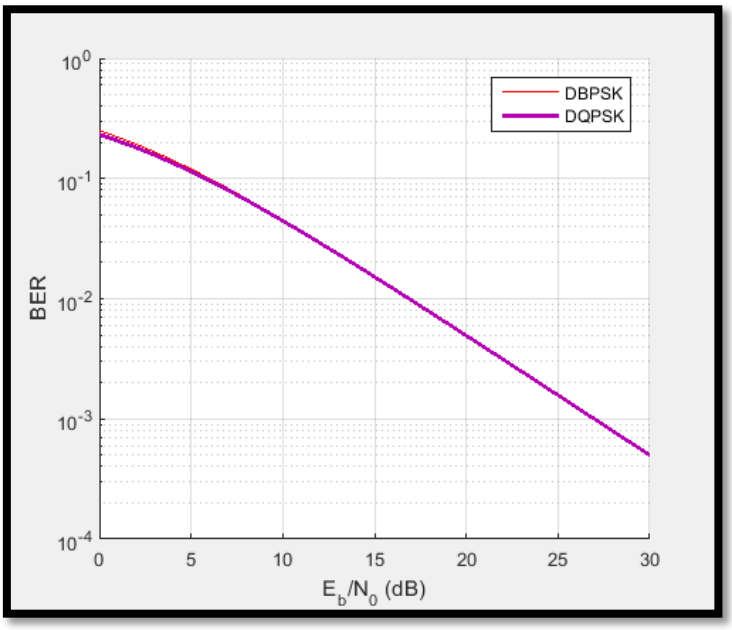


Fig.13: DBPSK and DQPSK in Rayliegh channel

Conclusion

In this paper IEEE 802.11 standard implemented in Mat lab Simulink where transmitter, channel and receiver designed. Allowing to modify beacon frame parameters such as beacon interval, SSID, supported rates, etc. also include or exclude the CF parameter set element from the generated beacon frame and displaying the scatter plots of transmitted and received chips, and the despread symbols, Modulation techniques that used in the Simulink model which are DBPSK and DQPSK are tested under AWGN channel and Rayliegh channel, DBPSK shows butter BER with 2dB gain in E_b/N_o. while in Rayliegh channel both technique, Wireless networks have gained popularity, providing users flexibility and mobility in accessing information. IEEE) 802.11 networks, are becoming common access networks in private and public environments.

Reference

[1] Fitchard, Kevin. "Wi-Fi Alliance gobbles up WiGig; plans to certify devices this year". Archived *from the original on 5 March 2016*. Retrieved 8 January 2016.

[2] Fleishman, Glenn (December 7, 2009). "The future of WiFi: gigabit speeds and beyond". *Ars Technica.* Archived *from the original on December 13, 2009*.

[3] Retrieved 2009.Borwein, Peter; Mossinghoff, Michael J. "Barker sequences and flat polynomials". In James McKee; Chris Smyth. Number Theory and Polynomials. LMS Lecture Notes. **352**. Cambridge University Press.2008.

[4] Merrill I. Skolnik, McGraw–Hill "Introduction to Radar Systems", 3rd Edition, 2001.

[5] Sun, Weiping; Choi, Munhwan; Choi, Sunghyun. "IEEE 802.11ah: A Long Range 802.11 WLAN at Sub 1 GHz" *(PDF). Journal of ICT Standardization. July 2013.*

[6] Chen, J. C. Measured Performance of 5-GHz 802.11a Wireless LAN Systems. Atheros Communications, Inc. 27 August.2001.

[7] Enterasys Network. Wireless Demilitarised Zone (WDMZ)-Enterasys Network's Best Practices Approach to an Interoperable WLAN Security Solution. Enterasys Network.2002.

[8] Goth, G. (2002). Wireless Security Still Ad Hoc and Add-On. DS Online Exclusive. IEEE Distributed Systems Online, 3(7). 7 July.2002.

[9] Tittelbach-Helmrich, K., Miletić, E., Wcislek, P. and Stamenković, Z. MAC Hardware Platform for RF-MIMO WLAN. Proceedings of 53rd IEEE International Midwest Symposium on Circuits and Systems, Seattle, 339-342. 2010.

Modulation and RS-CC rate specifications in WiMAX IEEE 802.16 Standard with MATLAB Simulink model

Fatima Faydhe al-azzawi
Middle technical university, institute of technology, electronic technical department
Corresponding author- Email : Fatima_faydy1981@yahoo.com

Abstract

The telecommunication industries development have focused on an intenive use of broadband systems, whicheareecharacterizedebyehighequalityefeatures. So a new technologies with high transmission abilities have been designed. Base on WiMAXeconcept. a wireless transmission infrastructure that allows a fast deployment as well as low maintenance costs. Based on the IEEE 802.16-2004 standard, in this paper an end -to-end baseband model of the physical layer of a wireless metropolitan area network (WMAN) have been implemented using MATALB Simulink, based on IEEE® 802.16-2004 standard. Supporting all of the mandatory coding and modulation options. As will as illustrates Space -Time Block Coding (STBC), an optional transmit diversity scheme specified for use on the downlink . it illustrates the use of digital pre-distortion , a technique for extending the linear range of a nonlinear amplifier. The modulation sets have been changed and tested in the model to reduce running time and power consumption. Finally, measurements of Error vector magnitude EVM and modulation error ratio MER were performed.

Keywords: IEEE 802.16-2004 standard, WiMAX, WMAN, wireless communication, broadband systems.

محددات التظمين ومعدل تشفير ريدسلمون وشفرة الالتواء في (WiMAX IEEE 802.16) مع موديل ببرنامج المحاكاة الماتلاب

م. فاطمة فايض محمد
الجامعة التقنية الوسطى / معهد تكنولوجيا بغداد / قسم التقنيات الالكترونية

الخلاصة

في تطوير صناعة الاتصلات عن بعد تم التركيز على الاستخدام المركز لانظمة النطاق العريض، تقنيات حديثة مع قدرات والذي هو تركيب نقل لاسلكي يسمح بسرعة عالية في النقل وكلفة WiMAXنقل عالية تم تصميمها، من اهمها مفهوم

، في هذا البحث نموذج من النهاية للنهاية ل للطبقة الفيزيائة ل IEEE 802.16 2004قليلة للتركيب. يعتمد على داعما كل خيارات IEEE 802.16 2004 نفذت باستخدام برنامج المحاكاة الماتلاب تبعا الى النظام العالمي WMAN و مخطط للتنوع النقل الاختياري المحدد المستخدم في الهابط، كما يوضح (STBC)التشفير والتضمين كما توضح استخدام تقنية التشويه المسبق الرقمي و تقنية تمديد المعدل الخطي للمكبرات اللاخطية كما تم تغيير نوع التظمين المستخدم و نسبة خطأ التظمين تم حسابها.EVMلتقليل زمن التنفيذ واستهلاك القدرة. واخيرا حسابات قيمة الخطأ المتجهة

1.Introduction

The broadband wireless access has become the best way to meet escalating business demand for rapid internet connection and integrated "triple play"services . In addition to not only topographic but also technological limitations, wireless solution alternatives have been found. That is the very base of the WiMAX concept :a wireless transmission infrastructure that allows a fast deployment as well as low maintenance costs. Based on the IEEE 802.16 - 2004 standard, WiMAX allows for an efficient use of bandwidth in a wide frequency range, and can be used as a last mile solution for broadband internet access [1, 2].

WiMAX vs. Wi-Fi and UMTS

Wi-Fi is a technology for wireless local area networking with devices based on the IEEE 802.11 standards provides which is provides a high data rate, only on a short range distances and with a slow user movement.while , UMTS which is (Universal Mobile Telecommunications System (UMTS) a third generation mobile cellular system for networks based on the GSM standard) that is offers larger ranges and vehicular mobility with lower data rate, it provides lower data rates, and requires high investments for its deployment. WiMAX tries to balance Wi-Fi and UMTS as shown in figure (1), thus providing vehicular mobility (included in IEEE 802.16e),and high service areas and data rates. While WiMAX will complement Wi-Fi and UMTS in some of the possible scenarios where these systems are not sufficiently developed, i.e. they face several problems in the deployment and they do not offer enough capacity to servee for all possible users, WiMAX will compete with Wi-Fi and UMTS , in other possible scenarios, where, in general, the costs in the deployment, maintenance , or just the supply of the service would not be profitable.

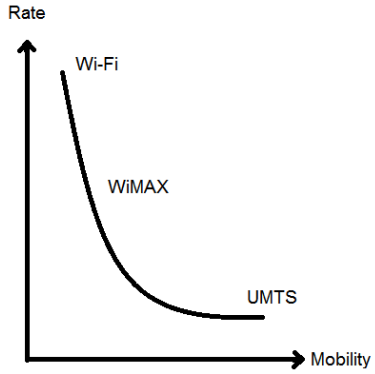


Fig. (1): WiMAX position between Wi-Fi and UMTS

I. WiMAX Model

This model as shown in fig. (2) shows an end -to- end baseband model of the physical layer of a wireless metropolitan area network (WMAN), according to the IEEE ® 802.16 - 2004 standard [1]. More specifically , it models the OFDM -based physical layer, called Wireless MAN - OFDM, supporting all of the mandatory coding and modulation options. It also illustrates Space-Time Block Coding (STBC), an optional transmit diversity scheme specified for use on the downlink. Finally, it illustrates the use of digital pre-distortion , a technique for extending the linear range of a non-linear amplifier.

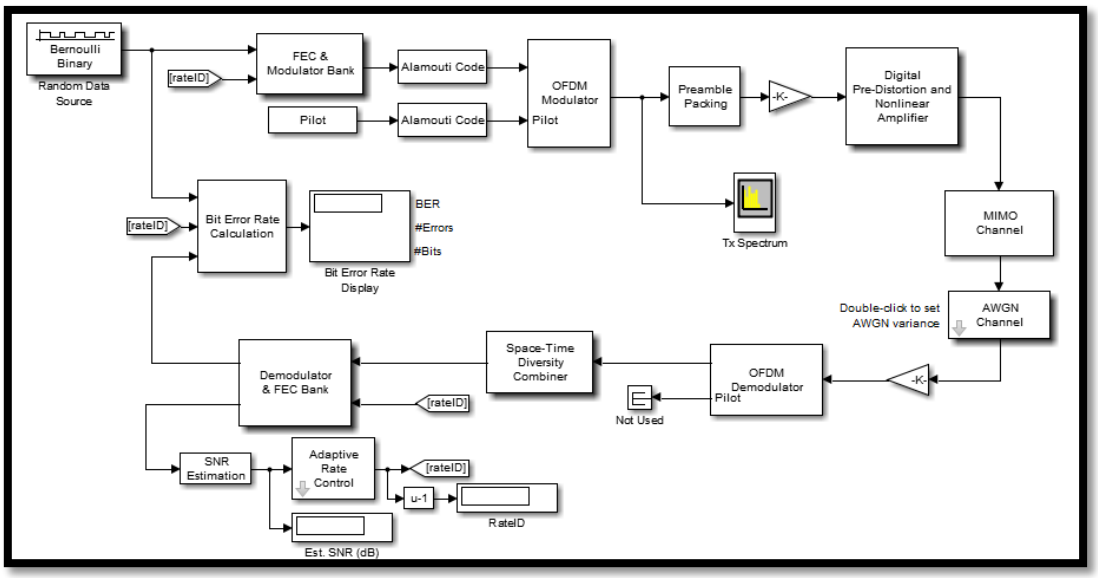


Fig. (2): WiMAX Simulink model

In fig. 2 a random bit data will be generated and mapped to a Forward Error Correction (FEC), consisting of a Reed - Solomon (RS) outer code concatenated with a rate - compatible inner convolutional code (CC) and data interleaving (Interleaver block permutes the symbols in the input signal, interleaver accepts a set of symbols and rearranges them, without repeating or omitting any of the symbols in the set), Modulation, using one of the BPSK, QPSK, 16-QAM or 64-QAM constellations specified. The model also use an adaptive - rate control scheme based on SNR estimates at the receiver to vary the data rate dynamically based on the channel conditions. The models use the standard - specified set of seven rates for OFDM - PHY, each corresponding to a specific modulation and RS - CC code rate as denoted by rate _ ID (see the following table as per [1]).

Rate ID	Modulation RS-CC rate
0	BPSK 1/2
1	QPSK 1/2
2	QPSK 1/2
3	16-QAM 1/2
4	64 QAM 3/4
5	64 QAM 2/3
6	64 QAM 3/4

While [9] gives the block sizes and code rates used for the different modulations.

Table 2: WiMAX modulation and coding schemes [9]

Rate ID	Modulation	RS code	CC code rate	Overall code rate
0	2-PAM	(12, 12, 0)	1/2	1/2
1	4 QAM	(32, 24, 4)	2/3	1/2
2	4 QAM	(40, 36, 2)	5/6	3/4
3	16 QAM	(64, 48, 4)	2/3	1/2
4	16 QAM	(80,48, 4)	5/6	3/4
5	64 QAM	(108,96,6)	3/4	2/3
6	64 QAM	(120,108,6)	5/6	3/4

FEC and modulator bank block in Simulink model detail illustrated in figure 3, where seven rates for OFDM - PHY, each corresponding to a specific modulation, each modulation block in Simulink model details shown in figure 4 and 5 where rate ID0 and rate ID1-6

respectively illustrated in those figures. If any changing in the modulation sets of table 1 and 2 only the block of specified type of modulator will changed [2,3].

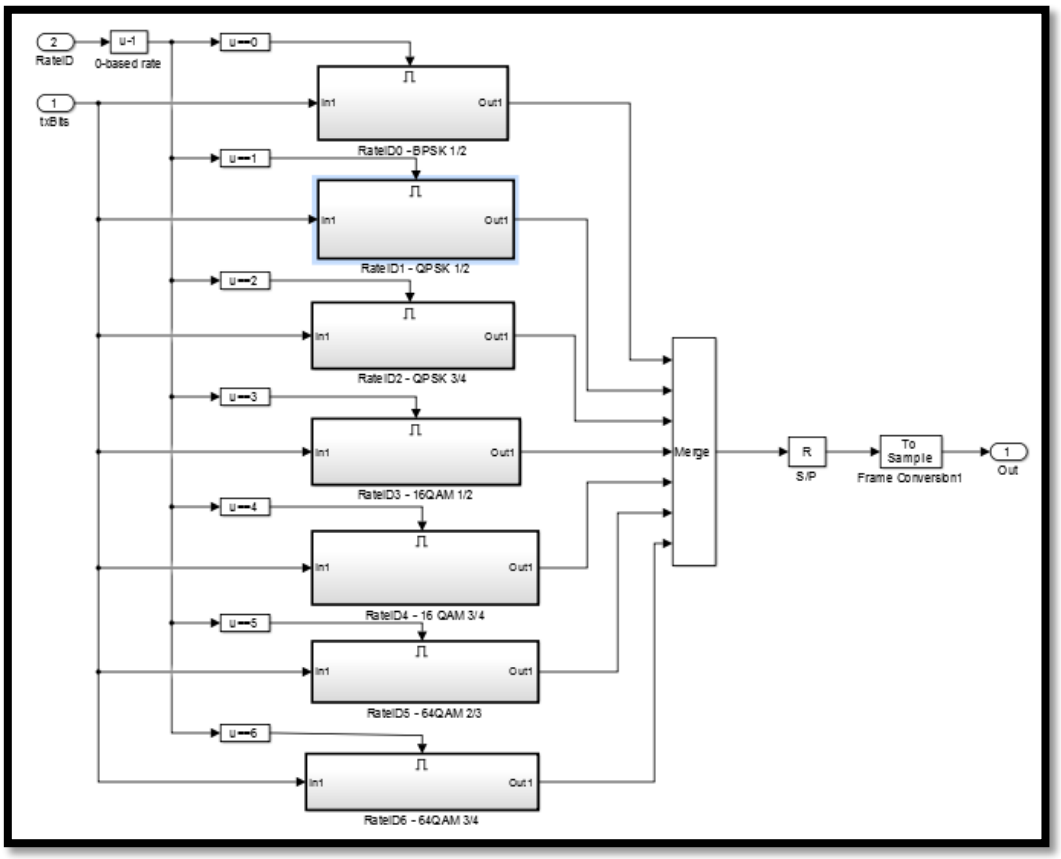


Fig. (3): FEC and modulator bank Simulink block.

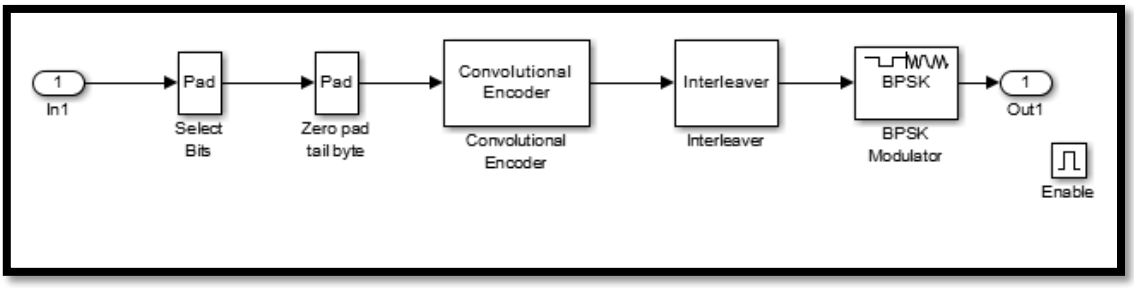


Fig. (4): Rate ID0- BPSK ½

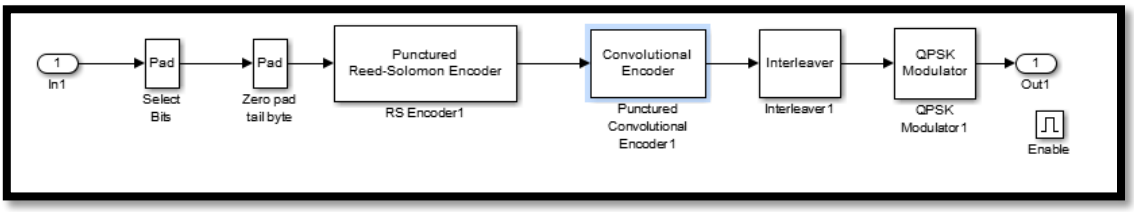


Fig. (5): Rate ID1 - BPSK 1/2

Space - Time Block Coding using an Alamouti code [3]. This implementation uses the OSTBC Encoder and Combiner blocks in the Communications System Toolbox™. Orthogonal Frequency Division Multiplexed (OFDM) transmission using 256-point FFTs and a variable cyclic prefix length. Among the 256 subcarriers, 192 are for data, 8 are for pilots, 28 are left guard band, 27 are right guard band and 1 is DC null. A single OFDM symbol length preamble that is used as the burst preamble. For the STBC model, both antennas transmit the single symbol preamble [4,5]. An optional memoryless nonlinearity that can be driven at several back off levels. An optional digital pre - distortion capability that corrects for the nonlinearity. A Multiple - Input - Single - Output (MISO) fading channel with AWGN for the STBC model. You can choose a non - fading, flat - fading, or dispersive multipath fading channel for the non - STBC model. [6,8]

OFDM receiver followed by channel estimation using the inserted preambles. For the STBC model, this implies diversity combining as per [3,6, 9, 10]. Hard - decision demodulation followed by DE interleaving (recovers a signal that was interleaved using the Interleaver block), Viterbi decoding (The Viterbi Decoder block decodes convolutional encoded input data input symbols to produce binary output symbols), and Reed-Solomon decoding as shown in figures 6, 7 and 8.

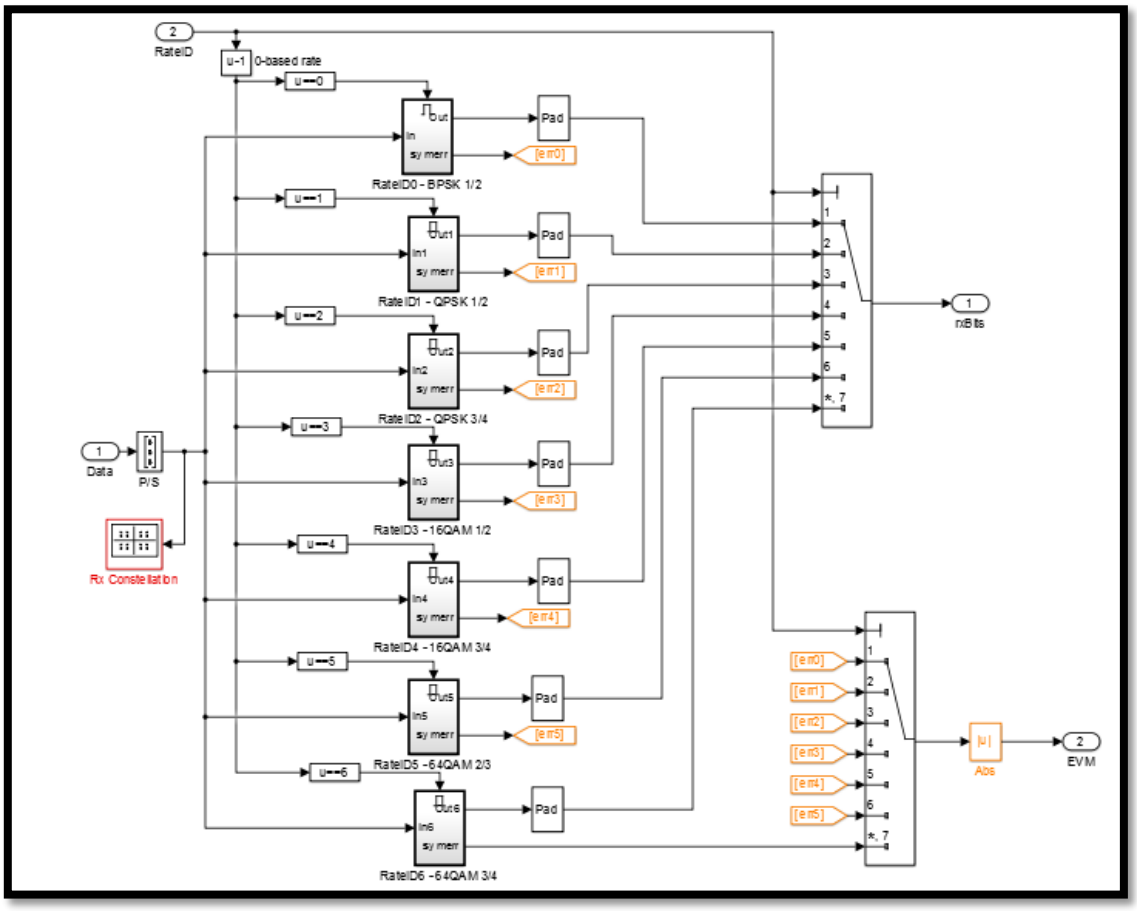


Fig. 6: demodulator and FEC bank Simulink block.

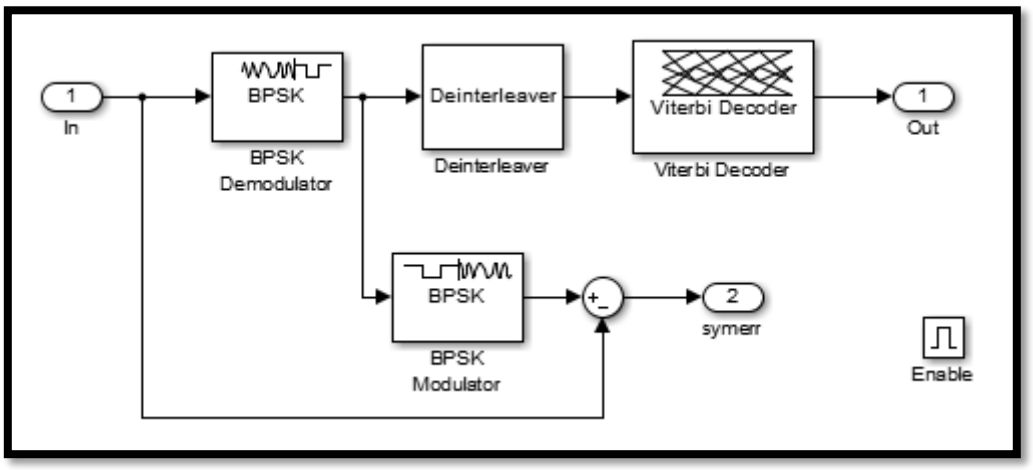


Fig. 7: Rate ID0- BPSK 1/2

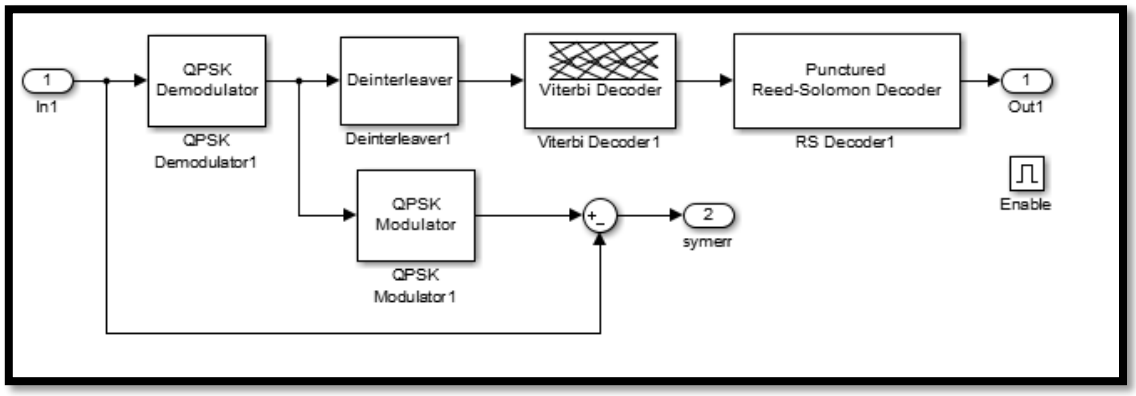


Fig. 8: Rate ID1 - BPSK 1/2

BER equations of modulation type used in the model

1. BER equation of M-PAM in AWGN channel [4-6].

$$P_s = 2\left(\frac{M-1}{M}\right) Q\left(\sqrt{\frac{6}{M^2-1}\frac{kE_b}{N_o}}\right)$$

$$P_b = \frac{2}{M\, log_2 M} \times \sum_{k=1}^{log_2 M} \sum_{i=0}^{(1-2^{-k})M-1} \left\{ (-1)^{\left\lfloor \frac{i\, 2^{k-1}}{M} \right\rfloor} \left(2^{k-1} - \left\lfloor \frac{i 2^{k-1}}{M} \right\rfloor \right) Q\left((2i+1) \sqrt{\frac{6 log_2 M}{M^2 - 1}\frac{E_b}{N_o}} \right) \right\}$$

Where:

P_s : Symbol error rate (SER)

P_b : Bit error rate (BER)

M : Size of modulation constellation

K : Number of bits per symbol → k= $\log_2$ M

$\frac{E_b}{E_0}$: Energy per bit - to - noise power - spectral - density ratio

Q function : $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp(\frac{-t^2}{2})\, dt$

2. BER equation of M-QAM in AWGN channel [4, 5, 6].

$$P_s = 4\frac{\sqrt{M}-1}{\sqrt{M}} Q\left(\sqrt{\frac{3}{M-1}\frac{kE_b}{N_o}}\right) - 4\left(\frac{\sqrt{M}-1}{\sqrt{M}}\right)^2 Q^2\left(\sqrt{\frac{3}{M-1}\frac{kE_b}{N_o}}\right)$$

$$P_b = \frac{2}{\sqrt{M}\, log_2\sqrt{M}} \times \sum_{k=1}^{log_2\sqrt{M}} \sum_{i=0}^{(1-2^{-k})\sqrt{M}-1} \left\{ (-1)^{\left\lfloor \frac{i2^{k-1}}{\sqrt{M}} \right\rfloor} \left(2^{k-1} - \left\lfloor \frac{i2^{k-1}}{\sqrt{M}} + \frac{1}{2} \right\rfloor \right) Q\left((2i+1) \sqrt{\frac{6log_2 M}{2(M-1)} \frac{E_b}{N_o}} \right) \right\}$$

3. BER equation of BPSK in AWGN channel [4].

$$P_S = P_b = Q\left(\sqrt{\frac{2E_b}{N_o}}\right)$$

4. BER equation of QPSK in AWGN channel [4].

$$P_S = 2Q\left(\sqrt{\frac{2E_b}{N_o}}\right)\left[1 - \frac{1}{2}Q\left(\sqrt{\frac{2E_b}{N_o}}\right)\right]$$

$$P_b = Q\left(\sqrt{\frac{2E_b}{N_o}}\right)$$

Simulation results

Modulation sets that chosen in model either table 1 or tow examined under AWGN channel with E_b/N_o from 0 to 30 dB as shown in figure 9 for table 1 and figure 10 for table 3

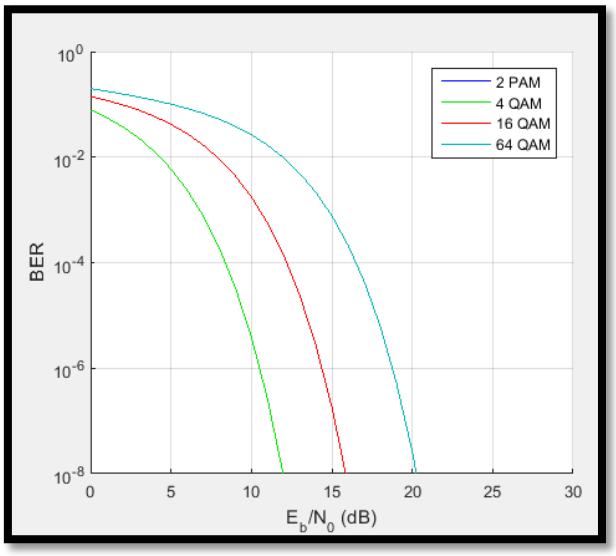


Fig. 9: BER of modulation sets in table 1

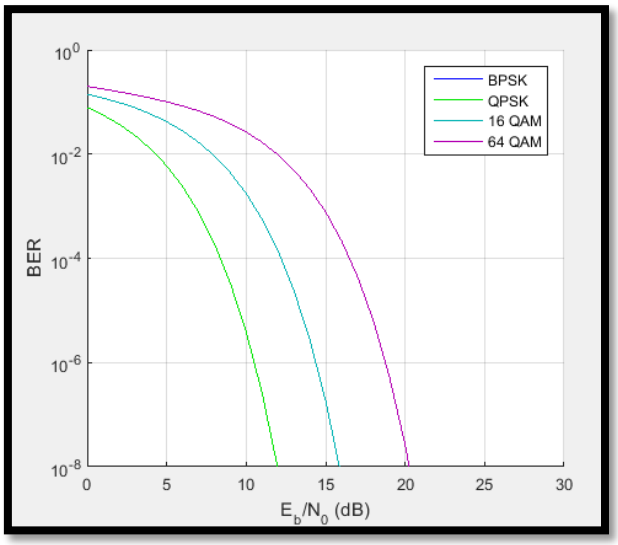


Fig. 10: BER of modulation sets in table 2

From figure 9 BER of both 2PAM and 4QAM are identical while figure 10 shows BPSK and QPSK same BER results in AWGN channel.

Power Spectral Density (PSD) which is simply the average noise power per unit of bandwidth, spectrum of transmitted signal in term of PSD is illustrated in figure 11 and 12, where figure 11 shows PSD of transmitted signal of 2 antennas in OFDM block without using raised cosine windowing between OFDM symbols, while figure 12 shows the spectrum with raised cosine

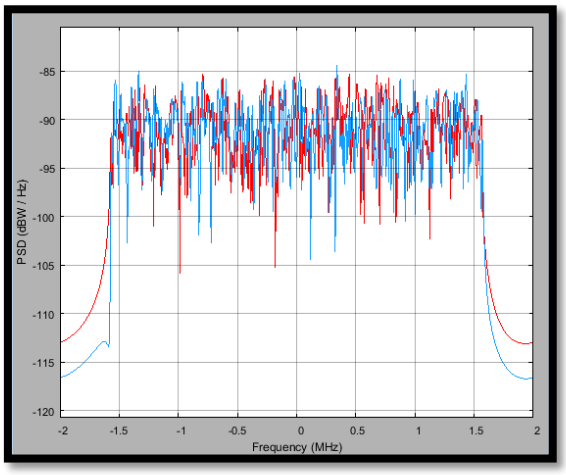


Fig. 11: transmitted spectrum

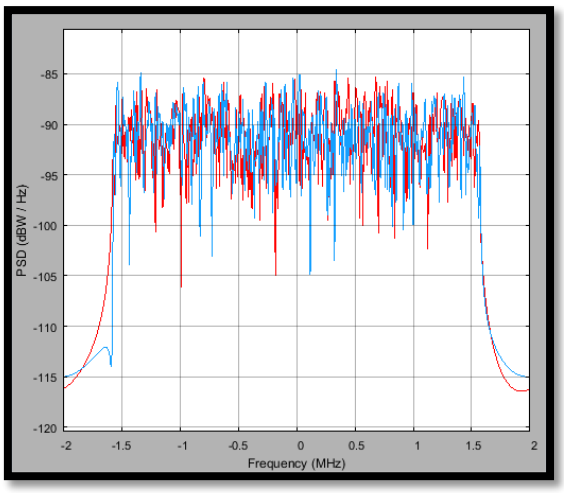


Fig. 12: transmitted spectrum with raised cosine windowing between OFDM symbols

Multi input single output power spectral density spectrum is shown in figure 13, while figure 14, 15 shows the received constellation and Received signal trajectory with EVM and MER measurement.

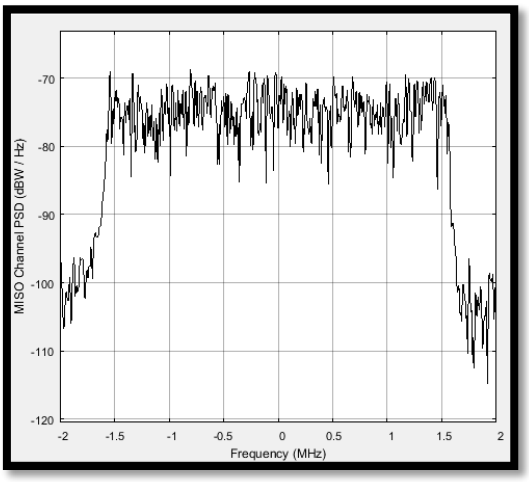


Fig. 13: MISO channel spectrum

The Constellation Diagram block plots constellation diagrams, signal trajectory, and provides the ability to perform Error vector magnitude EVM and modulation error ratio MER measurements.

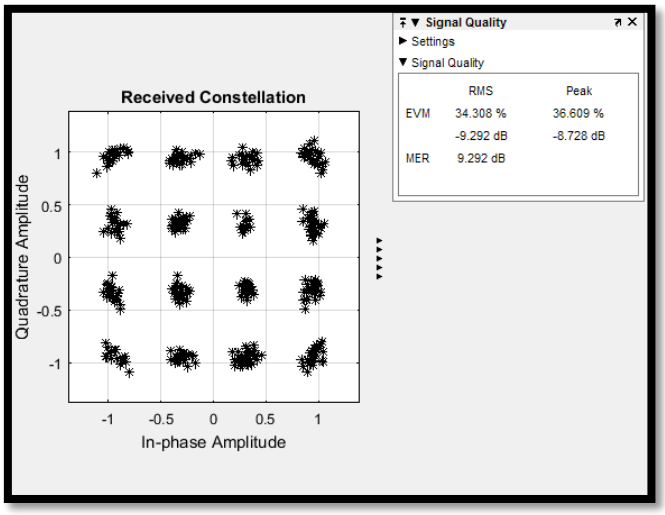


Fig. 14: Received constellation with EVM and MER measurement

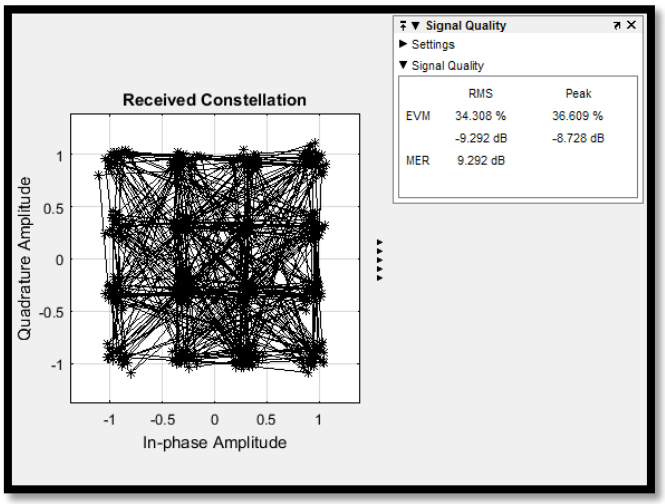


Fig. 15: Received signal trajectory with EVM and MER measurement

Conclusion

IEEE® 802.16-2004 standard have been implemented using MATLAB Simulink, where the model supporting all of the mandatory coding and modulation options. It also illustrates Space - Time Block Coding (STBC), an optional transmit diversity scheme specified for use on the downlink. it illustrates the use of digital pre-distortion, a technique for extending the linear range of a nonlinear amplifier. Also modulation sets have been changed and tested in the model to reduce running time and power consumption. Finally, measurements of Error vector magnitude EVM and modulation error ratio MER were performed, the results shows good performance in term of BER and power spectral density of both transmitted signal and received signal.

Reference

[1] IEEE Standard 802.16-2004, "Part 16: Air interface for fixed broadband wireless access systems," October 2004.http://ieee802.org/16/published.html.

[2] IEEE 802.16 Broadband Wireless Access Working Group, "Channel models for fixed wireless applications," IEEE 802.16a-03/01, 2003-06-27.

[3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," IEEE Journal on Selected Areas in Communications, Vol. 16, No. 8, Oct. 1998, pp. 1451-1458.

[4] Proakis, J. G.,*Digital Communications*, 4th Ed., McGraw-Hill, 2001.

[5] Simon, M. K., and Alouini, M. S., *Digital Communication over Fading Channels – A Unified Approach to Performance Analysis*, 1st Ed., Wiley, 2000.

[6] Cho, K., and Yoon, D., "On the general BER expression of one- and two-dimensional amplitude modulations", *IEEE Trans. Commun.*, vol. 50, no. 7, July 2002, pp. 1074-1080.

[7] Carl Eklund, et.al., "WirelessMAN: Inside the IEEE 802.16 Standard for Wireless Metropolitan Area Networks," IEEE Press, 2006.

[8] J. G. Andrews, A. Ghosh and R. Muhamed, "Fundamentals of WiMAX: Understanding Broadband Wireless Networking," Prentice Hall, 2007.

[9] J. Pino Lacosta: "*WiMAX: una alternativa d'accés a les xarxes*," Master Thesis, Universitat Oberta de Catalunya, Enginyeria Informàtica, June 2004.

[10] T. S. Rappaport: "*Wireless Communications: principles and practice*," Prentice Hall, 1996.

[11] Hussain A. Alhassan, 'A Study on the Performance of IEEE 802.16-2004 Includes STBC', ASEE 2014 Zone I Conference, April 3-5, 2014.

[12] Dahlman, E., S. Parkvall, and J. Skold. *4G LTE/LTE-Advanced for Mobile Broadband*.London: Elsevier Ltd., 2011.

وقائع المؤتمر العلمي الدولي الثاني لكلية النسور الجامعة للفترة 5-6/ 5/ 2018
Vol: 4 Year: 2018

Improvement of Mathematical Model of Scheduling Blocks for Resource Allocation Type 0 in Downlink of lit Technology

Mohammed K. Hassan[1]& Aymen M. Khodayer [2] & Haider D. Kamil [*3]

mohammednofa@yahoo.com[1], aymenaldulaimi@yahoo.com[2] ,
dr.haideraljanabi@gmail.com [3]

(1) Al-Rafidain University College, Department of Computer Engineering
(2) Al-farahidi University College, Department of Communications Techniques Engineering
(3)Al-Nisour University College - Department of Computer Techniques Engineering

Abstract

The results present the mathematical model for scheduling blocks allocation in the downlink of the LTE technology with support of different types of quality of service, organized according to RAT 0. The basis of the proposed solution is a model described. The novelty of the improved model consists in ensuring the balanced allocation of scheduling blocks among user equipment's taking into account requests priority and channel quality indicator (CQI). This allowed facilitating the differentiated quality of service for users both with consideration of their priorities and features of the signal-to-noise conditions.

Keywords: *LTE, channel quality indicator, RAT 0, Scheduling blocks, time-frequency resource, radio resource management.*

تحسين النموذج الرياضي في جدولة الكتل لتخصيص الموارد من النوع صفر (RAT 0) في الوصله الهبوط (Downlink) في شبكةLTE تكنولوجيا

و م.د. أيمن محمد خضير[2] و م.د.حيدر ضياء كامل [3] أ.م.د. محمد خضير حسن[1]
كلية الرافدين الجامعة، قسم هندسة الحاسبوب[1] ، كلية الفراهيدي الجامعة، قسم هندسة تقنيات الاتصالات[2]، كلية النسور الجامعة، قسم هندسة تقنيات الحاسبات[3]

الخلاصة

تقدم هذا النتائج النموذج الرياضي لجدولة تخصيص الكتل **(Scheduling blocks)** في الوصلة الهابطة **(Downlink)** لتقنية **LTE** مع دعم أنواع مختلفة من جودة الخدمة**(QOS)**، منظمة وفقًا لـ **RAT (0)**. أساس الحل المقترح هو نموذج تم وصفه في [5، 8، 10]. تتمثل التجديد في النموذج المُحسَّن من ناحية

ضمان التوزُّع المتوازن لكتل الجدولة**(Scheduling blocks)** بين المستخدمين (**Users**)مع الأخذ بعين الاعتبار طلبات الأولوية ومؤشر جودة القناة (**CQI**). وقد سمح ذلك بتوفير جودة الخدمة المتميزة للمستخدمين (**Users**) مع مراعاة أولوياتهم وخصائص ظروف الإشارة إلى الضوضاء**(SNR)**.

1. Introduction

In this regard, improving theoretical principles of technological mechanisms of networks resources management is one of the effective ways to improve indicators of quality of service of users' requests. In the analytical technology, such resources traditionally include symbols (time resource); frequency subcarriers (frequency resource); resource elements, resource blocks and scheduling blocks (time-frequency resource). To organize the access of user equipment's to the frequency and time resources of the downlink, the technology of Orthogonal Frequency-Division Multiplexing (OFDM) is used in the LTE technology [1,2]. However, the mechanisms for resource management in the downlink from the eNodeB side in the LTE technology, as well as in WiMAX networks, are not standardized, which allows developers of network equipment and appropriate protocols to offer their solutions in this area. At the eNodeB of the LTE network, the Radio Resource Management (RRM) system is responsible for allocation of network resources. It is practically realized as a scheduler of requests for service coming from UEs [2, 3]. The goal of the RRM is to provide prompt response to a change in the signal-to-noise conditions, as well as structural and functional parameters of the LTE network itself: the number and allocation of user equipment's, requirements of users' requests to the level of QoS, types of provided services, characteristics of provided traffic, etc.

Adaptation to the changes of parameters in the signal-to-noise conditions means that the eNodeB collects information on radio channel parameters from a UE as values of the Channel Quality Indicator (CQI) [3]. However, each UE transmits a CQI to eNodeB formed on the basis of the analysis of SNR values (Fig. 1) near any eNodeB [1, 4]. In its turn, the choice of scheme of the used Adaptive Modulation and Coding (AMC) depends on values of CQI (Table 1) [1, 11].

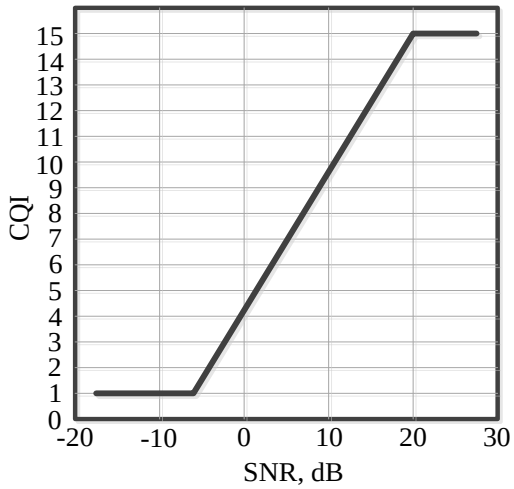


Fig. 1. The model of CQI-SNR ratio [1, 4]

Table 1: Correspondence of schemes of adaptive modulation and coding to CQI index values

CQI index	**Modulation**	**Symbol bit load**	**Code rate**	**Effectiveness, bit/s/Hz**
1	QPSK	2	78/1024	0.1523
2	QPSK	2	120/1024	0.2344
3	QPSK	2	193/1024	0.3770
4	QPSK	2	308/1024	0.6016
5	QPSK	2	449/1024	0.8770
6	QPSK	2	602/1024	1.1758
7	16 QAM	4	378/1024	1.4766
8	16 QAM	4	490/1024	1.9141
9	16 QAM	4	616/1024	2.4063
10	64 QAM	6	466/1024	2.7305
11	64 QAM	6	567/1024	3.3223
12	64 QAM	6	666/1024	3.9023
13	64 QAM	6	772/1024	4.5234
14	64 QAM	6	873/1024	5.1152
15	64 QAM	6	948/1024	5.5547

2. Analysis of specifics of quality of service ensuring in LTE network

The 3GPP LTE technology has been initially designed as a multiservice solution aimed at providing users with a set of info-communication services, which have a certain level of quality of service. The traditional solution of QoS ensuring problems is based on the classification and labeling (prioritization) of flows. Then, on the basis on the identified type of

traffic and assigned priority values, in the LTE network the solution for the problems of management of the available in the downlink time-frequency resource should be performed.

In order to provide ensuring of QoS requirements to a set of application, different Evolved Packet Core (EPS) flows of the transport network are formed. All possible EPS flows can be divided into two groups: flows with Minimum Guaranteed Bit Rate (GBR) and flows with Non-Guaranteed Bit Rate (Non-GBR). Let us consider the basic features of these types of flows:

1. **GBR flows.** The flows of this type have a certain value of the Minimum Bit Rate, which is set during the procedures of flow formation or its change. Therefore, transmission with the rate, which is higher than the set one, under a free radius is possible. In addition, there could be established a limitation on Maximum Bit Rate (MBR). The flows of this type are used, for example, in VoIP traffic transmission.
2. **Non-GBR flows.** The flows of this type do not guarantee the Minimum Bit Rate. These flows are used, for example, to transmit data traffic (FTP, e-mail).

At the section between the eNodeB and UE, QoS parameters ensuring is performed by the eNodeB. To do this, each flow is assigned with a QoS Class Identifier (QCI). Each QCI determines values for the following QoS parameters: the priority, allowed delay and number of lost packets. All the possible QCI values, as well as QoS parameters, related to a particular QCI, are determined in [97, 98], which should provide similar processing of flows of the same type at the equipment of different vendors. (Table 2) shows the standardized CQI values and their characteristics.

It should be noted that the value of the allowed delay at the section between P-GW and UE is determined [5]. In realization of data transmission with the guaranteed bit rate, the eNodeB must manage the resources in the dynamic mode. Services of classes QCI 1, 2, 3 and 7 are the services provided to the user in real time over the UDP/IP protocols. The main limiting factor for their realization is the allowed packet delay.

Table 2: Standardized characteristics of QoS Class Identifier (QCI)

QCI	**Type of service**	**Priority value**	**Allowed value of average packet delay**	**Allowed possibility of error**	**Examples of services provided**

				s and packet losses	
1	Guaranteed Bit Rate (GBR)	2	100 ms	10^{-2}	Conversational Voice
2		4	150 ms	10^{-3}	Conversational Video, Live Streaming
3		3	50 ms	10^{-3}	Real Time Gaming
QCI	Type of service	Priority value	Allowed value of average packet delay	Allowed possibility of errors and packet losses	Examples of services provided
4	Guaranteed Bit Rate (GBR)	5	300 ms	10^{-6}	Non-Conversational Video
65		0.7	75 ms	10^{-2}	Mission Critical user plane, voice Push To Talk
66		2	100 ms	10^{-2}	Non-Mission-Critical user plane, voice

					Push to Talk
5	Non-Guaranteed Bit Rate	1	100 ms	10^{-6}	IMS Signaling
6		6	300 ms	10^{-6}	Buffered Streaming Vide based on TCP
7		7	100 ms	10^{-3}	Live Streaming, Interactive Gaming
8		8	300 ms	10^{-6}	Buffered Streaming Video based on TCP
9		9			
69		0,5	60 ms	10^{-6}	Mission Critical delay sensitive signalling
70		5.5	200 ms	10^{-6}	Mission Critical Data

The PERL (Packet Error Loss Rate) value shows reliability of packet transmission, which is evaluated according to the relative value of rejected packets. The value of PERL $\leq 10^{-6}$ is achieved in packet delivery using the TCP/IP protocol. Signal traffic has the highest priority. Class 9 is applied by default in delivery of TCP/IP traffic (reading files from the Internet, E-mail, video) to nonprivileged users.

As a rule, values of priority and allowed delay determine in which way the scheduler at the eNodeB will process data packets. If the value of allowed losses will be 10^{-6}, then the acknowledged transmission is used.

The transmission of a service flow of a particular service is performed due to a bearer channel of the appropriate QCI. However, the basic parameters, which characterize the bearer, are:

- for GBR classes the guaranteed and maximum bit rate, which cannot be exceed;
- set and saved priority.

The total bit rate of flows over all the channels is set up for bearers with non-guaranteed bit rate. At the same time, the bearers of the GBR classes are allocated links.

The dynamic allocation of the channel resource is performed by the scheduler of the eNodeB, which should operate considering the following parameters:

- state of radio channels with the corresponding user equipment's;
- bearer attributes;
- characteristics of transmission process over bearers;
- signal-to-noise ratio.

When transferring data over the network, user flows must undergo several interfaces (LTE-Uu, S1, S5/S8) before they will be transmitted to the external network or UE. At each EPS interface, flows are displayed on the flows of lower layers, which have their own IDs. Each node provides matching of identifiers of the flow at the various interfaces (to which the node is relevant). Note that on the interfaces S1 and S5/S8, the flow is determined by the identity of the GTP tunnel. Packets belonging to the same flow of EPS are always processed in the same way.

The classification of incoming packets in order to determine the flow, to which they belong, is performed using the TFT (Traffic Flow Template). These templates use such IP header information as IP-addresses of a sender and receiver, and the TCP-port number.

3. Initial data, input terms and definitions

The developed model is aimed at the use of RAT 0, which implies combining RBs into the so-called resource block groups (RBGs) allocated to UE. The number of RBs included into one RBG (p) depends on the used frequency

and bandwidth. If the division of the number of RBs by the parameter p does not give integer values, the extreme RBG will have a size smaller than p.

In the proposed model, in accordance with the features of LTE technology, the following initial data assumed to be known [1, 5, 11]:

- N is the number of UEs;
- K is the number of RBs formed during the transmission of a single time slot. In the LTE technology the number of RBs depends on the bandwidth of the frequency channel and it can have values: 6, 15, 25, 50, 75, 100;
- K_S is the number of subcarriers for data transmission in a single RB in a Scheduling Block (SB). The given parameter of the frequency diversity between subcarriers Δf and it must satisfy the condition $K_S \Delta f = 180$ KHz. K_S can take the values of 12 and 24, which correspond to the frequency diversity between subcarriers Δf of 15 KHz and 7,5 KHz;
- N_{symb}^{RB} is the number of symbols, which form a single resource block. The parameter N_{symb}^{RB}=7 in case of using a normal cyclic prefix (CP). The length of the normal CP of the first OFDM symbol is T_{CP}^{1}=5,2 μs, and from the second to sixth OFDM symbol – T_{CP}^{2-6}=4,7 μs (Fig. 2). When the extended CP is used (T_{CP}=16,7 μs), the RB consists of six OFDM-symbols (N_{symb}^{RB}=6);
- T_{RB}=0,5 ms is the transmission time of a single RB;
- T_{SB} =1 ms is the transmission time of a single SB;
- T_{SF}= 1 ms is the time of one subframe transmission;
- N_{SF}^{RB}=2 is the number of RBs, which are formed at the same subcarriers and allocated to a UE during the transmission of one subframe;
- $R_c^{n,m}$ is the code rate used in the coding of a signal of the n-th UE at the subcarriers of the m-th SB;

- $k_b^{n,m}$ is the bit load of the n-th UE symbol at the subcarriers of the m-th SB;
- channel division type – FDD or TDD and the used frame configuration;
- R_{req}^n is the bit rate, which must be allocated to the n-th UE;
- M is the number of subframes used to transmit data in the LTE downlink;
- p is the number of RBs included into a single RBG.

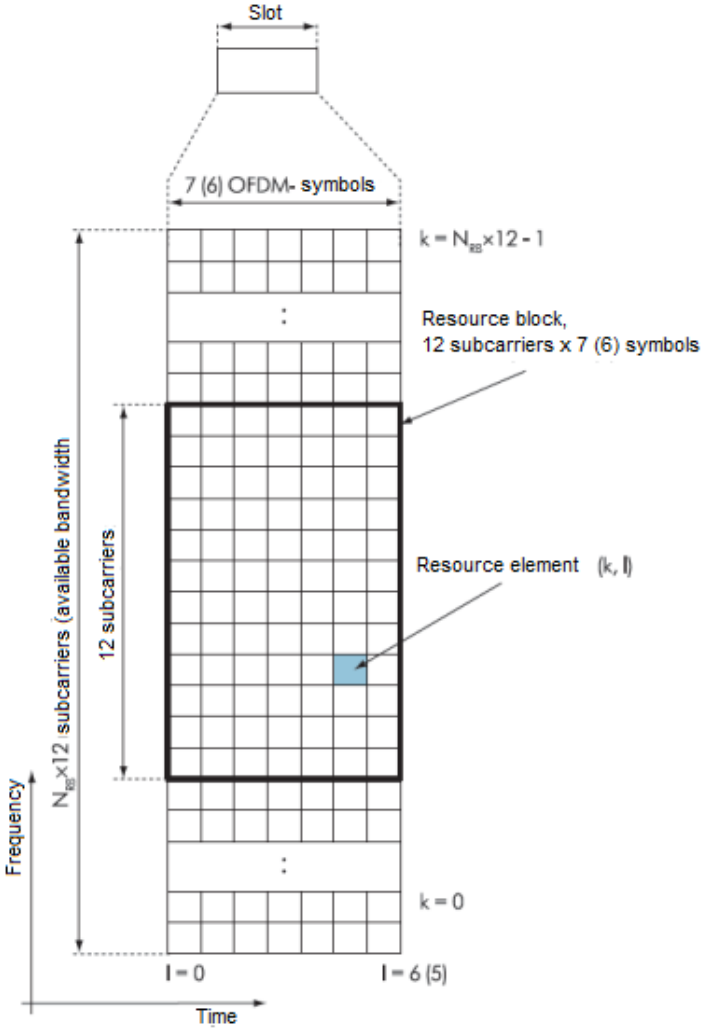


Fig. 2: The LTE resource grid under the standard step of subcarriers of 15 KHz

4. Mathematical model for allocation of scheduling blocks in LTE downlink organized according to RAT 0.

It is important to consider that according to the RAT 0 of the LTE technology, allocation of resource blocks to user equipment can be performed only during all the

length of a subframe. Then the resource blocks, which correspond to identical subcarriers, form the so-called scheduling block, and the problem of time-frequency resource allocation in the downlink of the LTE technology can be formulated as a problem of scheduling blocks (SBs) allocation [1,5, 11].

To consider the structure of a frame or subframes, as well as the number of subframe, allocated to transmit data by UEs in the downlink, the notion of downlink configuration matrix is introduced into the mathematical model [1]. In the given rectangular matrix H, the number of lines corresponds to the number of possible configurations of the frame (L), and the number of columns is equal to the appropriate number of subframes (M) in the frame as a whole. Then

$$H = \left\| h_{l,m} \right\|, (l = \overline{0, L-1};\ m = \overline{0, M-1}), \qquad (1)$$

where

$$h_{l,m} = \begin{cases} 1, \text{ if the } m\text{th subframe is used to transmit the information in the downlink} \\ \text{under the } l\text{th configuration;} \\ 0, \text{ otherwise.} \end{cases}$$

The configuration matrix of the downlink is as follows

$$H = \begin{Vmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{Vmatrix}. \qquad (2)$$

In the proposed model, a set of Boolean variables will be used as control variables, which regulate the process of scheduling blocks allocation among user equipment's:

$$x_{k,m}^{n}=\begin{cases}1, \text{if the } kth \text{ scheduling block at the } m\text{th subframe} \\ \quad \text{is allocated to the } n\text{th UE;} \\ 0, \text{ otherwise.}\end{cases} \quad (3)$$

According to the physical sense of the introduced variables and on the basis of specifics of the LTE network functioning during calculation of the control variables $x_{k,m}^{n}$, it is important to fulfill several key conditions, which will further act as existing restrictions imposed to these variables in solving the optimization problem. To provide the allocation of the *k*-th SB of the downlink during the transmission of the *m-th* subframe to not more than one UE, the control variables are imposed with the restrictions of the type:

$$\sum_{n=1}^{N} x_{k,m}^{n} \le 1, \quad (k=\overline{0,K-1};\ m=\overline{0,M-1}). \quad (4)$$

The fulfillment of the given condition guarantees that at the physical OSI layer the same frequencies will not be used for data transmission to different UEs at the same period of time. The condition of allocating only downlink scheduling blocks to the user equipment in accordance with the used frame configuration, configuration matrix (1.1) is:

$$\frac{\sum_{k=0}^{K-1}\sum_{n=1}^{N} x_{k,m}^{n}}{M} \le h_{l,m}, (m=\overline{0,M-1}). \quad (5)$$

The most important condition of the proposed improved model is the condition of balanced allocation of scheduling blocks to the *n*th UE in such a quantity, which will ensure the required transmission rate in the downlink under the used modulation and coding scheme:

$$\frac{(Pr_n+1)}{(\mathrm{CQI}_n+1)}\left[\sum_{k=\frac{K}{2}-3}^{\frac{K}{2}+2} R_{SB}^{0} x_{k,0}^{n} + \sum_{k=\frac{M}{2}-3}^{\frac{K}{2}+2} R_{SB}^{5} x_{k,5}^{n} + \sum_{k=0}^{\frac{K}{2}-4} R_{SB}^{0-9} x_{k,0}^{n} + \right.$$

$$+ \sum_{k=\frac{K}{2}+3}^{K-1} R_{SB}^{0-9} x_{k,0}^{n} + \sum_{k=0}^{\frac{K}{2}-4} R_{SB}^{0-9} x_{k,5}^{n} + \sum_{k=\frac{K}{2}+3}^{K-1} R_{SB}^{0-9} x_{k,5}^{n} +$$

$$\left. + \sum_{k=0}^{K-1}\sum_{m=0}^{4} R_{SB}^{0-9} h_{l,m} x_{k,m}^{n} + \sum_{k=0}^{K-1}\sum_{m=6}^{M-1} R_{SB}^{0-9} h_{l,m} x_{k,m}^{n} \right] \geq \beta \cdot R_{req}^{n}$$

(6)

when $n = \overline{1, N}$.

In expression (6), the following parameters have been used [1]:

$$R_{SB}^{0} = \frac{\left(\left(N_{symb}^{RB} N_{SF}^{RB} - N_{PDCCH} - N_{PSS,SSS} - N_{PBCH}\right) K_s - N_{pilot} N_{SF}^{RB}\right) R_c^{n,k} k_b^{n,k}}{KT_{SF}}$$

Is the bandwidth of SBs formed on the zero subframe and transmitting Primary Synchronization Signals (PSSs), Secondary Synchronization Signal (SSSs) and Physical Broadcast Channel (PBCH) using $N_{PSS,SSS}$=2 and N_{PBCH} =4 symbols accordingly;

$$R_{SB}^{5} = \frac{\left(\left(N_{symb}^{RB} N_{SF}^{RB} - N_{PDCCH} - N_{PSS,SSS}\right) K_s - N_{pilot} N_{SF}^{RB}\right) R_c^{n,k} k_b^{n,k}}{KT_{SF}}$$

is the bandwidth of SBs formed at the fifth subframe and transmitting PSSs and SSSs;

$$R_{SB}^{0-9} = \frac{\left(\left(N_{symb}^{RB} N_{SF}^{RB} - N_{PDCCH}\right) K_s - N_{pilot} N_{SF}^{RB}\right) R_c^{n,k} k_b^{n,k}}{KT_{SF}}$$

is the bandwidth of SBs formed at all the subframes and containing only the Physical Downlink Control Channel (PDCCH) out of all the service information; N_{PDCCH}=3 is the number of symbols in each subframe allocated to transmit the PDCCH;

N_{pilot}=4 is resource elements allocated to transmit pilot signals in a single RB. Besides, to form a single RE, one subcarrier and one OFDM-symbol are used.

β is a control variable, which characterizes the lower bound of the required QoS level for requests of users regarding the bit rate;

Pr_n is the priority of service requests coming from the *n*-th user equipment (Table 2); CQI_n is the numerical value of the channel quality indicator toward the *n*-th user equipment's.

The calculation of the desired variables (3) according to the above introduced limiting conditions (4) -(6) is reasonably to perform during solving the optimization problem, which is oriented at ensuring the maximum of the lower bound of meeting UEs' requirements to the bandwidth allocated to them in the LTE downlink:

$$\max_{x,\beta} \beta. \qquad (7)$$

The novelty of the proposed improvement of the model is that during balancing of the number of scheduling blocks allocated to UEs, both request priority (Pr_n) and parameters of signal-to-noise ratio (CQI_n) are explicitly taken into account (6). According to the performed improvement, a large amount of the time-frequency resource will be allocated to UEs, which has higher priority (lesser value of Pr_n), as well as better value of CQI_n. On the one hand, this allows providing multiservice nature of solutions for differentiated quality of service, and on the other hand, excluding the situation, when UEs with high QoS requirements but low SINR/CQI practically block the performance of the LTE network by monopolizing practically all the available time-frequency resource.

The number of control variables determines the scalability of the optimization problem solved and fully depends on the number of UEs in the LTE network, the number of subframes in a frame, as well as the number of scheduling blocks, which are formed during the transmission of one subframe. Considering this, the number of control variables ($x_{k,m}^{n}$) corresponds to $N \times M \times K$. As a result of determining the

numerical values of variables (3), allocation of scheduling blocks and assignment of subframes to user equipment's, during which data will be transmitted from eNodeB to UE in the downlink, is performed.

Given that the introduced control variables $x_{k,m}^{n}$ (3) are Boolean and the variable β is real-valued, then the formulated optimization problem, which is connected with the maximization of the objective function (7) under linear restrictions (4)-(6), is related to the class of Mixed Integer Linear Programming (MILP). To solve this problem Rounding-off, Branch-and-bound, simulated annealing methods [6, 12] can be used. In the thesis the modified method, which is the basis of the intlinprog function included into the Optimization Toolbox package of the Matlab system, has been used [7].

5. Analysis of methods for solution of problem of scheduling blocks allocation represented by optimization problem of mixed integer linear programming

As it is said in the previous Subchapter, the solution of the problem of resource block allocation is narrowed down to the solution of mixed integer linear programming (MILP), for which the MatLab «Optimization Toolbox» has been used [7] (version 14 and higher), which is represented by the «intlinprog» subprogram.

In solving the problems of the mixed integer linear programming, it is necessary to minimize the objective function in the linear form

$$\min_{x} \overset{\rho}{f}^{t} \overset{\rho}{x}, \qquad (8)$$

under fulfillment of a number of conditions of linear restrictions, which are represented as equalities and inequalities

$$A \cdot \overset{\rho}{x} \le \overset{\nu}{b};\ Aeq \cdot \overset{\rho}{x} = \overset{\nu}{b}eq;\ \text{and}\ \overrightarrow{lb} \le x \le \overrightarrow{ub}, \qquad (9)$$

where $\overset{\nu}{x}$ is the vector of control variables, and the number of its coordinates is determined by the number of scheduling blocks, subframes and user equipment's, including variable β, and it is respectively equal $K \cdot M \cdot N + 1$; Matrix *A* and vector $\overset{\nu}{b}$ are formed from the inequalities (4)-(6);

Matrix *Aeq* and vector $\vec{beq}$ are not determined (*Aeq*=[] or *beq*=[]) because there are no conditions in the form of equalities in the proposed model;

$\vec{lb}$ and $\vec{ub}$ are column-vectors of bounds, namely $\vec{lb}$ consists of zeros corresponding to the lower bound of the control variables values, and $\vec{ub}$ has all unit values, except for the explicitly undetermined higher bound of the variable β («inf»).

Moreover, $\vec{f}$ is a weight vector, in which there is only one non-zero element, which corresponds to the variable β, equal to -1 as a result of the need for maximization of the objective function (7).

Let us consider the features of the used algorithm. The subprogram «intlinprog» uses the following base strategy to solve the problem of the MILP. In addition, the «intlinprog» can obtain the solution of the problem at any of the represented stages. In case when the problem is solved at one of the stages, the «intlinprog» does not perform the following ones.

Thus, among the stages we can point out the following:

1. Reducing the scalability using pre-processing method of the linear programming.
2. Solution of the initial problem with weakened restrictions (non-integer) using methods of linear programming.
3. Performance of pre-processing by the method of integer linear programming, limiting solution under weakened restrictions, obtained by the method of linear programming for solving mixed integer problem.
4. Further tightening of the method of linear programming by Cut Generation restrictions for solving the mixed integer problem.
5. Determination of integer solutions using heuristics.
6. Use of the Branch-and-Bound algorithm for the systematic search of optimal solutions. This algorithm solves the simplified linear programming problem with a limited range of possible values of integer variables. At the same time, it is trying to create a sequence of updated bounds under the optimal value of the objective function.

In turn, the Branch-an-Bound method consists of a sequence of subproblems that are trying to converge to the solution of MILP. The solution of these subproblems make it possible to obtain the sequence of the upper and lower bounds of solution $f^{t}(x)$. The first upper bound is any feasible solution, and the first lower bound of is the solution to the problem with the weakened restrictions.

Conclusions

1. The framework of the LTE technology provides several resource allocation types that use certain procedures, while one of the types is the RAT 0. However, the known models describing the allocation according to RAT 0 do not take into account the mutual influence of requests' priorities and the channel quality, which is evaluated, for example, by the value of the CQI toward the user equipment.

2.Thus, the mathematical model for the allocation of scheduling blocks in the LTE downlink with support of various types of quality of service, organized within RAT 0, the novelty of the model is in ensuring balanced allocation of scheduling blocks among user equipment's based on the priority of requests and channel quality indicator (CQI). This allowed ensuring differentiated quality of service of users taking into account both the priority of their requests and characteristics of signal-to-noise ratio.

3. Implementation of the model (1) - (7) optimizes the allocation of scheduling blocks in the LTE downlink with support of various types of quality of service under RAT 0 on the basis of the calculation of the control variables that determine the order of scheduling blocks allocation. Given that the input control variables $x_{k,m}^{n}$ (3) are Boolean, and the variable β is real-valued, then the formulated optimization problem related to the maximization of the objective function (7) in the presence of linear constraints (4)-(6) belongs to the class of mixed integer linear programming (MILP). The model (1) -(7) is linear, which makes its use quite effective in terms of computational realization.

4. To solve the optimization problem of scheduling blocks allocation in the LTE downlink on the basis of RAT 0, the Optimization Toolbox has been used, which is widely tested and well-proven in practice tool of the MatLab system. In order to solve

the problem of mixed integer linear programming, the «intlinprog» subprogram has been used; it is based on the combined use of a number of computational methods (Rounding-off, heuristic method, method of Branch-and-bound) at various stages of the program performance.

5. The proposed model can be used directly in solving the problem of scheduling blocks allocation in the LTE downlink with support of various types of QoS organized within RAT 0 and under the mathematical description of more complex management processes of time-frequency resource in the LTE network, for example, associated with designing and planning of a wireless communication system.

References

[1] Dahlman E. 4G: LTE/LTE-Advanced for Mobile Broadband, Second Edition / E. Dahlman, S. Parkvall, J. Skold. – [2nd Edition] – Academic Press, 2013. – 544 p.

[2] 3GPP TS 36.211 V13.0.0 (2016-01). 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation (Release 13). Valbonne, France, 2015. – 143 p.

[3] 3GPP TS 36.213 V12.4.0 (2015-02). 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures (Release 12). Valbonne, France, 2015. – 227 p.

[4] Mehlfuhrer C. Simulating the long-term evolution physical layer / C. Mehlfuhrer, M. Wrulich, J.C. Ikuno, D. Bosanska, M. Rupp // 17th European Signal processing conference (EUSIPCO 2009). – Glasgow: EURASIP, 2009. – P. 1471-1478.

[5] Garkusha S. V. The Model of Time and Frequency Resource Scheduling in Downlink LTE / S.V. Garkusha // 2013 International Siberian Conference on Control and Communications (SIBCON) : proceedings, Krasnoyarsk, Russia. September 12–13, 2013. – Krasnoyarsk : Siberian Federal University, 2013. – P. 1–4.

[6] Lopez C. MATLAB Optimization Techniques / C. Lopez. – [1st Edition] – A press, 2014. – 292 p.

[7] Al-Dulaimi A.M. Balancing Model of Resource Blocks Allocation in LTE Downlink / Aymen M. Al-Dulaimi, Essa M. Al-Azzawi, Ali I. Al-Anssari // International IEEE Conference on Electronics and Information Technology (EIT`2016). Proceedings. –

Odesa: Odesa National Polytechnic University. Ukraine, Odesa, May 23−27, 2016. – PP. 9-12.

[8] Lemeshko O.V. Priority Based Balancing Model of Resource Allocation in LTE Downlink / O.V.Lemeshko, A.M.K.Al-Dulaimi // Scholars Journal of Engineering and Technology. – 2016. – 4(4). – P. 169-174.

[9] Garkusha S. Model of Resource Allocation Type 1 for LTE Downlink / S. Garkusha, A.M.K. Al-Dulaimi, H.D. Al-Janabi // Proceedings of the IX International Conference on Antenna Theory and Techniques (ICATT), 21-24 April 2015, Kharkiv, Ukraine. – Kharkiv: Kharkiv National University of Radioelectronics. – P. 279-281.

[10] Garkusha S. Result Research Model of Scheduling Block Allocation in Downlink LTE / S. Garkusha, A. Al-Dulaimi, H. Al-Janabi // Modern Problems of Radio Engineering, Telecommunications and Computer Science. Proceedings of the international Conference TCSET'2014. – Lviv-Slavske, Ukraine, February 25 - March 1, 2014: Publishing House of Lviv Polytechnic, 2014. – P. 498-500.

[11] Gordejuela-Sanchez F. LTE Access Network Planning and Optimization: A Service-Oriented and Technology-Specific Perspective / F. Gordejuela-Sanchez, J. Zhang // IEEE Global Telecommunications Conference, 2009. GLOBECOM 2009, Nov. 30 2009-Dec. 4 2009. – P. 1-5.

[12] Rao S.S. Engineering Optimization: Theory and Practice / S.S. Rao. – [4th Edition] – Wiley, 2009. – 840 p.

Spiking Neural Network Controller Design for Lumber Spine and Leg Movements of Human Rehabilitation Robot

Janan J. Abdulkareem[1], Mohammed Yousif Hassan[2]

M.ScStudent, Department of Control& systems Engineering, University of technology, Baghdad, Iraq[1]
Associate Professor, Department of Control &systems Engineering, University of technology, Baghdad, Iraq[2]

Abstract

This paper presents the design and simulation of a Spiking Neural Network (SNN) to control the joints position for seven degrees of freedom rehabilitation robot. The design of SNN is performed using Spike Response Model (SRM). The Back- Prop algorithm is applied for learning the SNN. The coding and decoding is applied for converting real numbers into spikes. The controller is designed so that the rehabilitation robot follows position medical trajectory paths in the rehabilitation process with minimum error in position. The designed structure of SNNs is single input-single output. The PI controllers are also simulated on the same structure to approve that SNNs controllers works correctly.

Keywords: *Rehabilitation Robot, SNNs Controller, PI Controller.*

الخلاصة

المسارات الطبية وكل شبكه يتم في هذا البحث تصميم و نمذجة الشبكات العصبية المتصاعدة كمسيطرات تتبع مسؤولة على ألسيطرة على مفصل واحد. لتعليم الشبكة ألعصبية تم استخدام طريقه (back –prop algorithm). صممت تقنية الشبكات العصبية الاصطناعية لنموذج الحركية العكسية للروبوت من أجل تقليل الوقت الحسابي لإيجاد زاويا يتم .الروبوت بصوره دقيقة وكإدخال من الشبكة العصبية هو موقع محور وإخراج الشبكة العصبية هو زوايا الروبوت يعمل بشكل صحيح.SNNs على نفس البنية لاعتماد وحدات التحكم PIأيضًا محاكاة وحدات التحكم

1. Introduction

Biological neurons process input information and transform them into output through sequences of spikes. The transmission of information through the neurons is carried out by the use of their electrical and chemical properties. To model the input-output relation of biological neurons, a number of mathematical models known as artificial neuron models are proposed [1].

Neuron models of the first two generations do not employ individual spikes, but their output can be seen as normalized firing rates of the neuron within a certain period of time. This type of coding is called *rate coding*.

The Spiking Neuron (SN) models are considered to be the third generation of the Artificial Neural Networks, which have a better resemblance to the biological neuron than the earlier models. Similar to their biological counterparts, in SN models, the magnitude of the spikes contains no information; rather, all information is encoded in the timing of the individual spikes [2]. This coding scheme is called *pulse encoding*.

One of the major works in the region of SNNs [3] has shown that they can approximate any continuous function using individual spike times. The same work has explained that SN are computationally more powerful than neurons with sigmoid activation function. Furthermore, SNNs have the ability of fast adaptation [4]. The abovementioned advantages and an increasing interest in temporal computation have encouraged the use of SNNs for many different applications.

The training of SNNs has been considered in many research projects [5]. A Back Propagation (BP)-based learning algorithm, called Spike Prop has been presented in [4] Where it is assumed that the value of the internal state of the neuron increases linearly in the infinitesimal time around the instant of neuronal firing. To improve the performance of BP, a number of modifications are suggested to the basic algorithm, such as Back Propagation with momentum, QuickProp [6].

2. Spiking Neural Networks

The dynamic behavior of a spiking neuron is close to its biological counterpart. Figure 1 demonstrates the structure of the SNN, in which each connection between a presynaptic and a postsynaptic neuron is modeled by multiple synapses. Each synaptic connection is characterized by weights that represent the strength of the connection and each synapse has a different delay. The inputs and outputs of a spiking neuron are described by a series of firing-times, called a spike-train. Each firing-time describes the time a neuron has sent out a pulse [6].

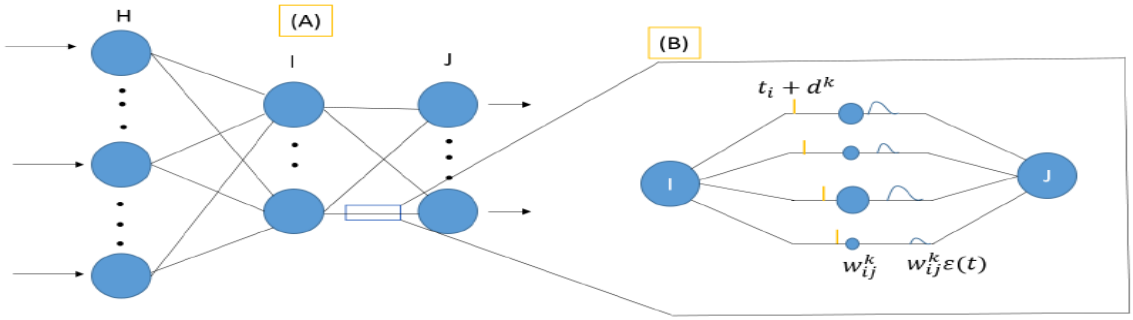


Fig. 1. Structure of Spiking Neural Network: (A) Feed Forward Spiking Neural Network (B) Connection Consisting of Multiple Delayed Synaptic Terminals [6].

The potential of a spiking neuron is modeled by a dynamic variable and works as a leaky integrator of the incoming spikes: newer spikes contribute more to the potential than older spikes. If this sum is higher than a predefined threshold, the neuron fires a spike.

Different models were developed for mathematical formulation of information transmission in spiking neurons. These are the Integrate-and-Fire Model, Izhikevich Model, Hodgkin- Huxley Model, and Spike Response Model [7]. The Spike Response Model (SRM) is the most widely exploited model because of its mathematical simplicity.

In this paper, the SRM model is used to describe how the sequence of incoming spikes is processed to produce a new spike train leaving the neuron [8]. It is assumed that any neuron generates at most one spike during the simulation interval and fires when the internal state variable reaches a threshold.

Assume that feedforward operation, where each neuron in hidden layer is checked if it is spiked or not. Each neuron can spike at maximum once only during the time interval, when its membrane potential (m_i) exceeds threshold value (ϑ). If the neuron (i)is spiked at (t_i^f) , the algorithm goes to the next neuron i+1 in the same layer, else SNN algorithm calculates the membrane potential $m_i(t)$ using the following equation:

$$m_i = \sum_{h=1}^{NH} \sum_{k=1}^{D} w_{hi}^k(t)\varepsilon(t - t_h^f - d^k) \qquad (1)$$

The response function (ε) shows the spike effect of the presynaptic on the membrane potential of the postsynaptic neuron. Various mathematical formulations are used as a response function, but the form of the function is always a short rising part followed by a long decaying part and ε (t) =0 for $t \leq 0$ to insure causality, as shown in Figure (2), and is calculated as [9]:

$$\varepsilon(t) = \frac{t}{\tau} e^{(1-\frac{t}{\tau})} \text{for } t>0 \qquad else\ \varepsilon(t) = 0 \qquad (2)$$

Where, τ is the membranes potential decay time constant in second.

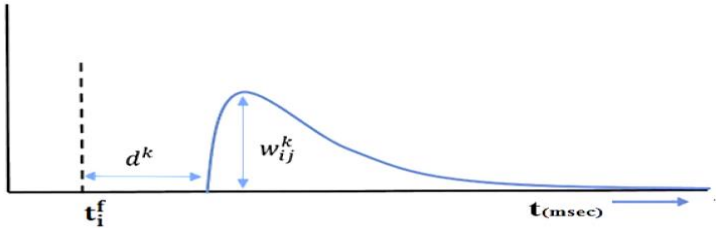


Fig. 2: The Potential of Postsynaptic Neuron j [9]

Compute the error using the Least Mean Squared Error as in equation

$$E = \frac{1}{2}\sum_{j\in J}(t_j^f - t_j^d)^2 \quad (3)$$

Where, t_j^d is the desired spike time for output neuron j

The backward part initiates, where the synapses weights of connection will be adjusted. Counter to feedforward, the Back-Propagation initiates from output layer and returns to the hidden layer. The synapses weights of output layer are adjusted according to equations (4) to (7):

$$\frac{\partial}{\partial t}\left(y_i^k\right) = y_i^k\left(\frac{1}{t - t_i^f - d_k} - \frac{1}{\tau}\right) \quad (4)$$

Where, $$\frac{\partial}{\partial t(y_i^k)} = 0 \; for \; t \le t_i^f + d_k$$

And, y_i^k represents the spike response function for a single sub-connection terminal.

$$\delta_j = \frac{T_j - t_j^f}{\sum_{i=1}^{NI}\sum_{k=1}^{D} w_{ij}^k(R)\frac{\partial}{\partial t}(y_i^k)} \quad (5)$$

Where, δj is the delta function

The updated weight of the output layer is:

$$w_{ij}^k(t+1) = w_{ij}^k(t) - \Delta w_{ij}^k \quad (6)$$

Where, $\Delta w_{ij}^k(R) = \eta.\delta_j.y_i^k$ (7)

And, η is the learning rate.

The synapses weight between hidden and input layer will be updated using equations (8) to (10):

$$\delta_i = \frac{\sum_{j=1}^{NI}\delta_j\sum_{k=1}^{D} w_{ij}^k(R)\frac{\partial}{\partial t}(y_i^k)}{\sum_{j=1}^{NH}\sum_{k=1}^{D} w_{hi}^k(R)\frac{\partial}{\partial t}(y_h^k)} \quad (8)$$

$$w_{hi}^k(t+1) = w_{hi}^k(t) - \Delta w_{hi}^k \quad (9)$$

Where, $\Delta w_{hi}^k(R) = \eta.\delta_i.y_h^k$ (10)

3. Coding and Decoding:

When the state of spiking neuron *j* described by its potential *mj* (*t*) crosses a certain constant threshold value **ϑ**, the neuron fires a spike, which is described by its spike time *tj*. However, in most engineering problems the computations must be performed on analog data, which leads to the development of methods for encoding analog signals into spike trains. The following formula is employed

encoding the data of inputs to spikes time and passes these spikes to hidden layer using equation (11) [6]:

$$t_h^f = t_{max} - round\left(\mathrm{t_{min}} + \frac{(\mathrm{Rin}-\mathrm{R_{min}})(\mathrm{t_{max}}-\mathrm{t_{min}})}{(\mathrm{R_{max}}-\mathrm{R_{min}})}\right) \qquad (11)$$

Where, *Tmax* and *Tmin* are the maximum and minimum time intervals, where they are different for the input and the desired output, *Rmax* is larger than the largest value in the input patterns and the output pattern, *Rmin* is the smaller than the smallest value in the input patterns and output patterns, and *Rin* is the current value of data.

The output from the network must be decoded from time to real information. This can be done using the following equation [6]:

$$t_j^f = \left(\frac{(\mathrm{Tmax}-t_j^f-\mathrm{Tmin})(\mathrm{R_{max}}-\mathrm{R_{min}})}{(\mathrm{T_{max}}-\mathrm{T_{min}})}\right) \qquad (12)$$

4. Controllers Performance

In this paper the development of the position controls system for rehabilitation robot is discussed. The robot model has high non-linearity due to Pneumatic Muscle Actuator that used for the robot movement so that the designed controllers are used to make the robot follows the medical trajectory with minimum error. PI controllers are connected with the robot to control it and explained that the SNNs controllers are given result with minimum tracking error best than conventional error. The inputs for the controllers are the error of angles and the outputs are control signals connected with pneumatic muscle actuator of robot to control the high non-linearity. The general specifications of SNN using Modified Spike-Prop training algorithm (MSP) are presented in the Table (1).

Table 1: SNN Specification

Joint No.	Controller Network Topology	τ (sec)	ϑ	D	Δt (sec)	tmax-tmin (I/P)	tmax-tmin (O/P)
1	1,4,1	1	0.5	3	0.1	10-1	10-1
2	1,6,1	1	0.5	2	0.1	20-1	20-1
3	1,4,1	1	0.5	3	0.1	10-2	20-1
4	1,5,1	2	0.7	3	0.1	9-1	20-1
5	1,7,1	2	0.5	3	0.1	9-1	9-1
6	1,5,1	6	1	3	0.1	15-1	15-1
7	1,7,1	2	0.5	3	0.1	10-1	10-1

where the network topology represents the number of units in Input/Hidden/Output,τ is time constant in(*msec*), ϑ is the threshold value for spiking, D is the time delay of synapses per connection,Δt is the Step size in *msec* and tmax-tmin (i/p) is time range for input spike in (*msec*). The structure of the SNN is shown in Figure (3).

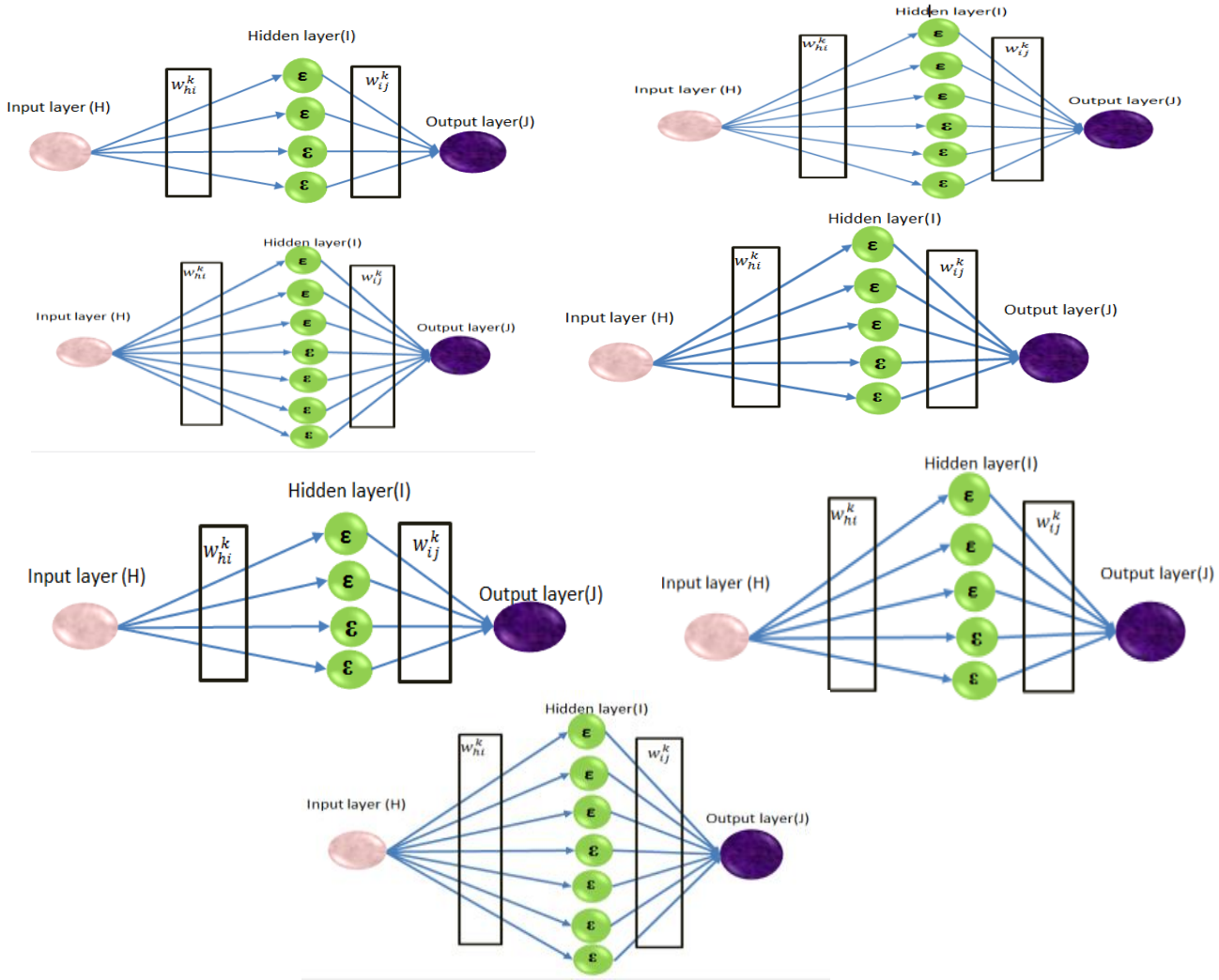


Fig. 3: Structure of SNN position controllers.

4. Simulation of Controlled System Rehabilitation Robot

In this work, two medical trajectories were applied. The two trajectories axes are generated by using a cubic interpolation where its trajectory axes are x, y and z.t_o is the initial time and t_f is the final time of motion. The trajectory is the time of planning in such a way that the robot arrives at the final configuration in total time tf = 30 sec. The planned trajectory is [11]:

$$\theta(t) = \theta_i(t^\circ)[\frac{\theta id(tf)-\theta id(t^\circ)}{tf^2}]t^2 - 2[\frac{\theta id(tf)-\theta id(t^\circ)}{tf^3}]t^3 \quad (13)$$

Where, (i=1,2,3,4,5,6 and 7) is the index of the generated angle for the first desired trajectory. This trajectory is transformed into Cartesian coordinates using the direct kinematics in order to obtain the desired position trajectory, Xd= $[xd\ \ yd\ \ zd]^T$.This trajectory represents the motion of knee flexion/extension and ankle flexion/extension joint in a maximum range movement and its axes equations are:

x=-a5.cos(q6d) (14)

y=a1 (15)

z=- (a6.sin(q6d) +d8.sin(q6d+q7d)) (16)

The planned second trajectory is explained below. This trajectory is transformed into Cartesian coordinates using the direct kinematics in order to obtain the desired position trajectory. This trajectory represents the motion of hip joint flexion /extension and ankle flexion /extension and its axes equations are shown below:

x=0 (17)

y=a1 (18)

z=a3+ (a5+a6). cos (q6d) +d8.cos (q6d+q7d) (19)

where q5d, q6d, and q7d are the desired position for the fifth joint (hip), sixth joint (knee) and seventh joint (ankle)joint respectively for two trajectories. These types of trajectories are required and important in rehabilitation training because they give motion for a large number in muscles of human leg.

The Simulink of the complete Position controlled system for rehabilitation robot using SNN is shown in Figure (4). It consists of sub blocks of the dynamic model, PMA, SNN position controller, inverse kinematic of system, and the reference trajectory and its inverse.

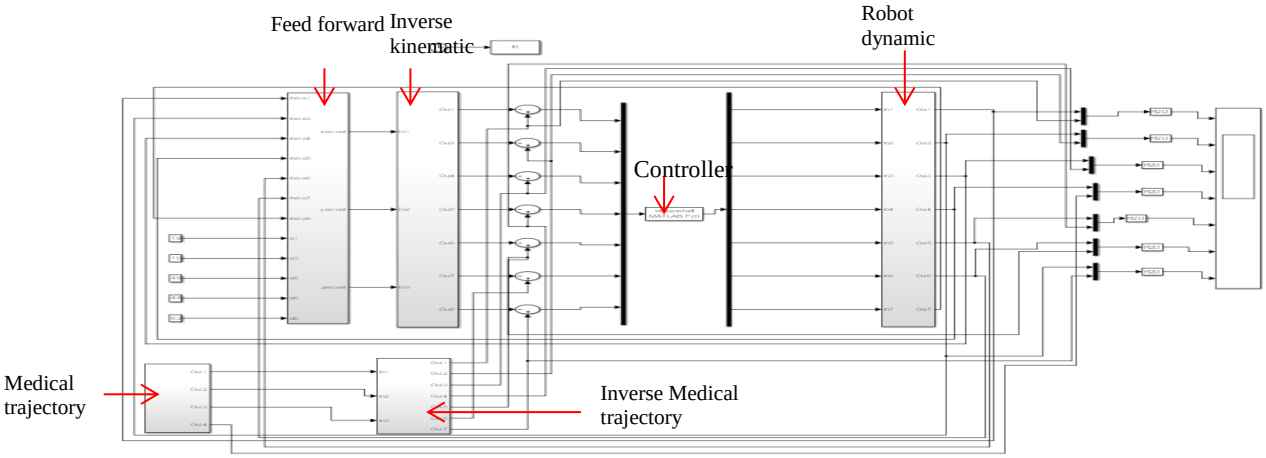


Fig. 4: Simulink of position-controlled system for the rehabilitation robot.

The disturbance represented in this work is stated as [12]:

$\tau dis = A * \sin(wt)$ (20)

where the frequency of disturbance is assumed to be w=10×π rad/s and amplitude is A=4 $Volt$

4.1 First Trajectory without Applying Disturbance

By applying the first reference trajectory with angles initially are: joint 6 is 0°and joint 7is 0°. The destination angle of joint 6 is -65° and joint 7 is -40°.

In this work, the criterion to measure the error in trajectory is by using Root Mean Squared Error (RMSE) [11]:

$$\text{RMSE} = \frac{1}{N}\sum_{1}^{N}\sqrt{e_{th1}^2 + e_{th2}^2 + e_{th3}^2 + e_{th4}^2 + e_{th5}^2 + e_{th6}^2 + e_{th7}^2} \quad (21)$$

Where,$e_{th1}^2, e_{th2}^2, e_{th3}^2, e_{th4}^2, e_{th5}^2, e_{th6}^2, e_{th7}^2$is the error in trajectory in the first though seventh joint angles respectively and N is the number of samples. The acceptable gains of PI controller are shown in Table (2), and the joint angle response are shown in Figure (5), the joint angle response using SNN controller are shown in Figure (6).

Table 2: Best PI gains for the first trajectory without applying disturbance.

No. of joint	1	2	3	4	5	6	7
KP	0.125	0.9	0.9	0.5	13.2	4.78	0.2
KI	7.1	2.5	0.09	2.3	0.07	0.36	0.15

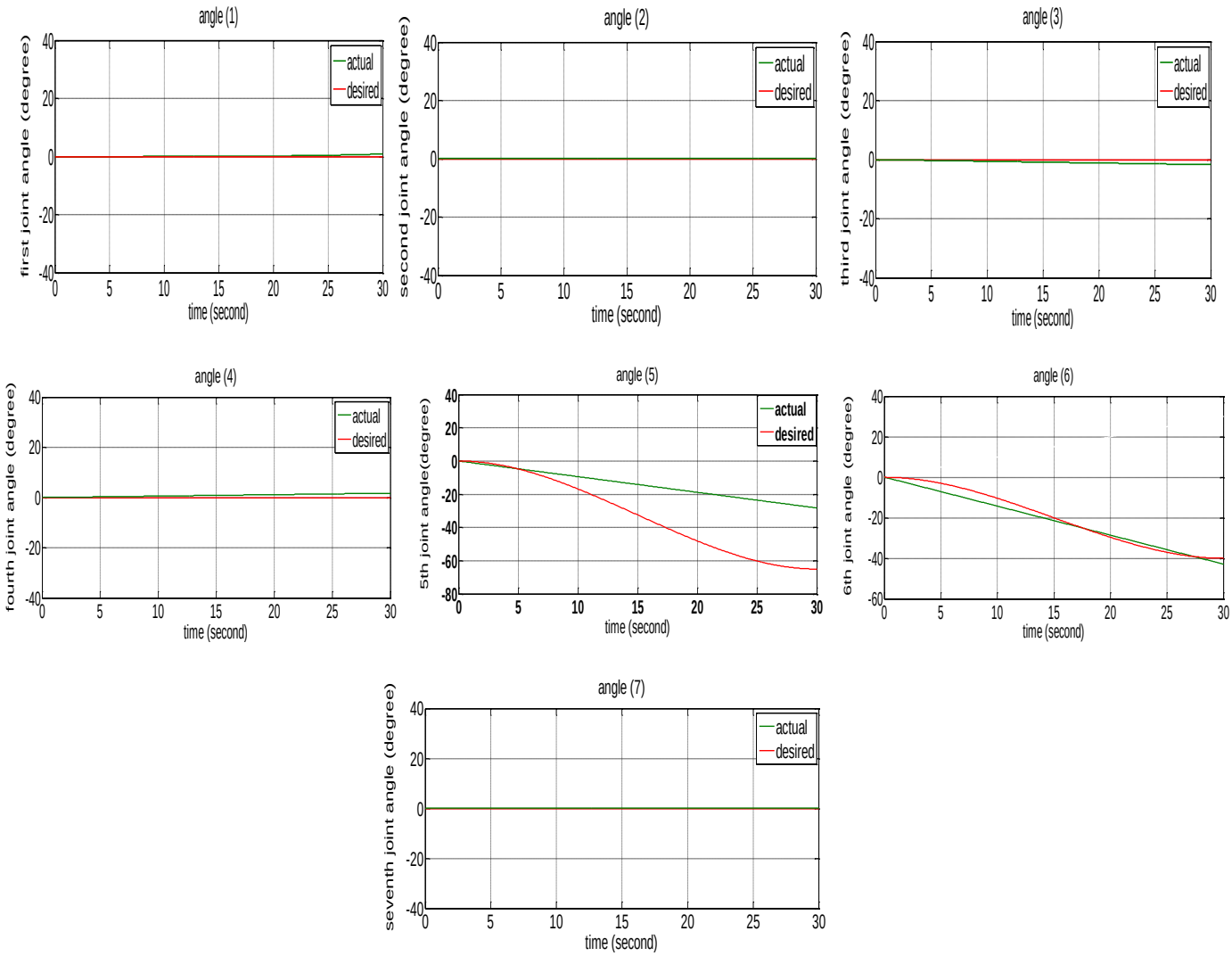


Fig. 5: Joint angles responses for the first reference trajectory without applying disturbance (PI controller).

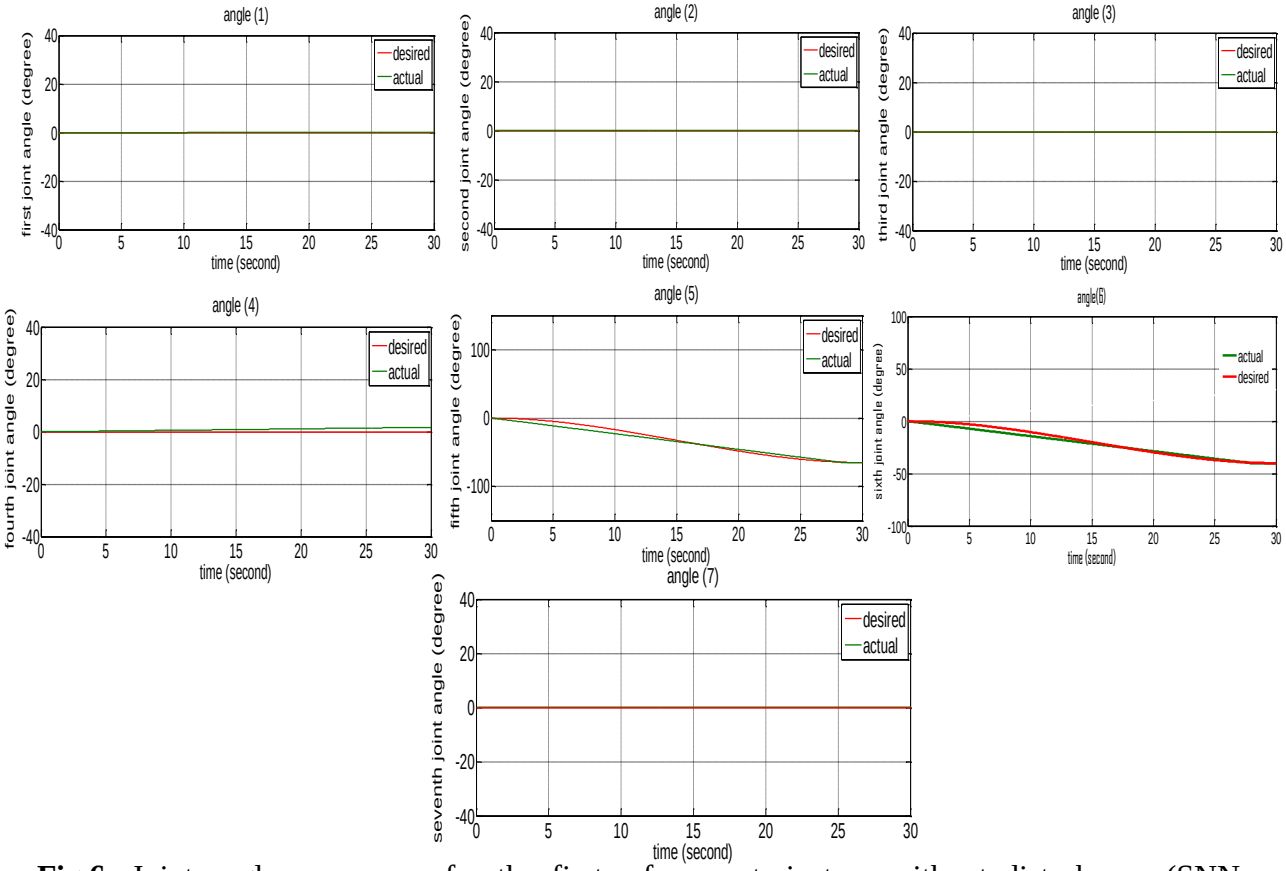


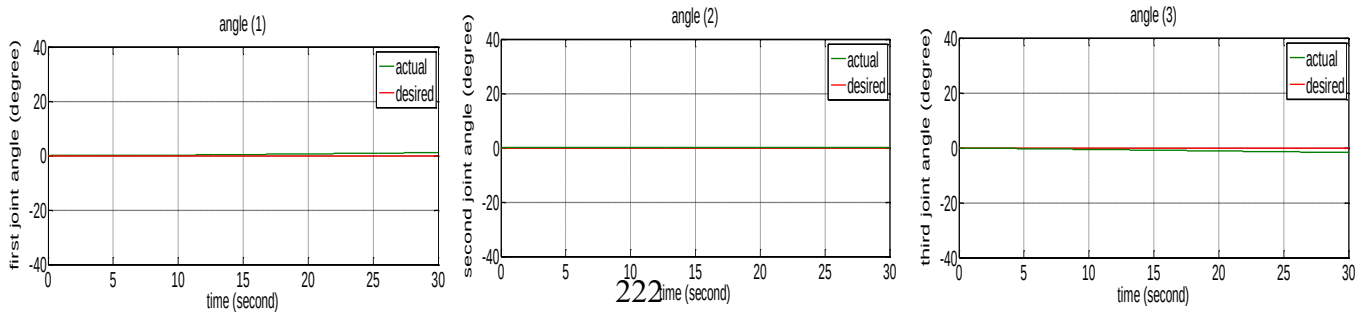
Fig.6: Joint angles responses for the first reference trajectory without disturbance (SNN controller).

4.2 First Trajectory with applying Disturbance

Assuming the disturbance mention equation (21) is applied, the best gains of PI controller are shown in Table (3). The response for each joint angle using PI and SNNs controllers are shown in Figure (7) and (8) respectively.

Table 3: Best PI Gains for the First Trajectory without Applying Disturbance.

No. of joint	1	2	3	4	5	6	7
kp	1.25	0.1	0.4	0.125	3.2	7.4	0.3
ki	0.9	0.5	0.8	0.3	0.2	0.01	0.024



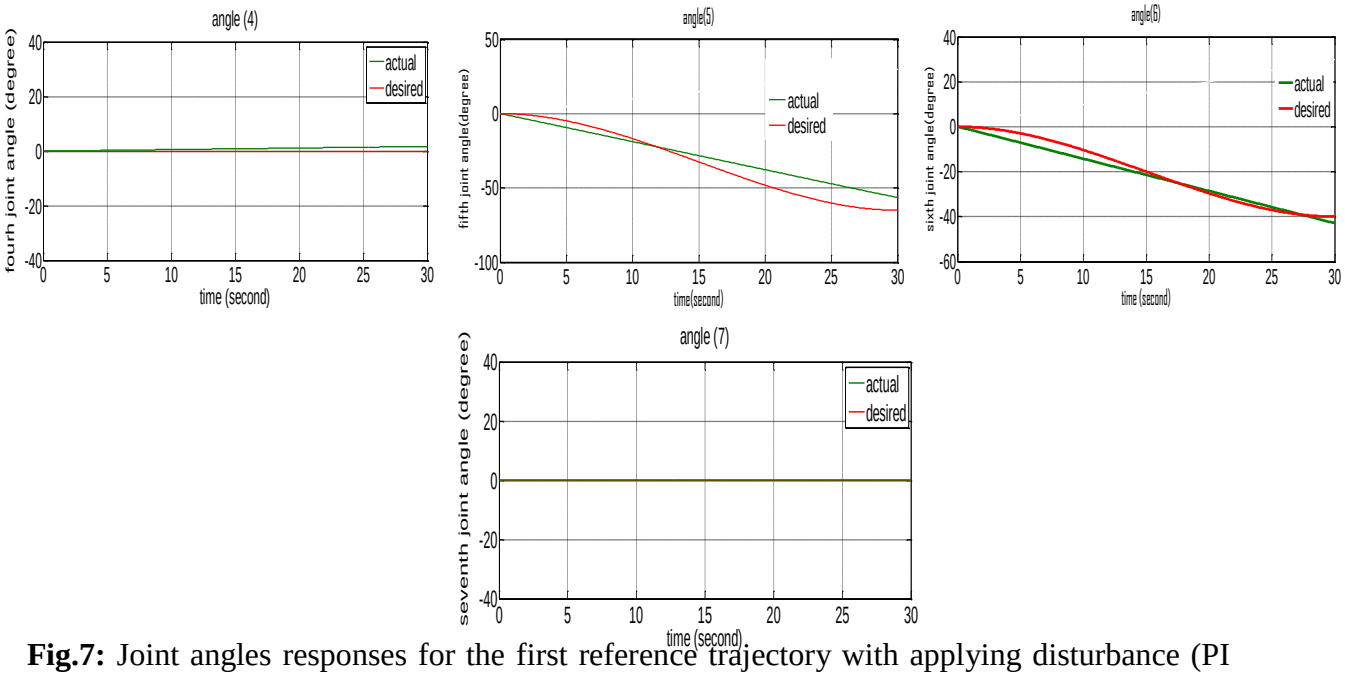


Fig.7: Joint angles responses for the first reference trajectory with applying disturbance (PI controller).

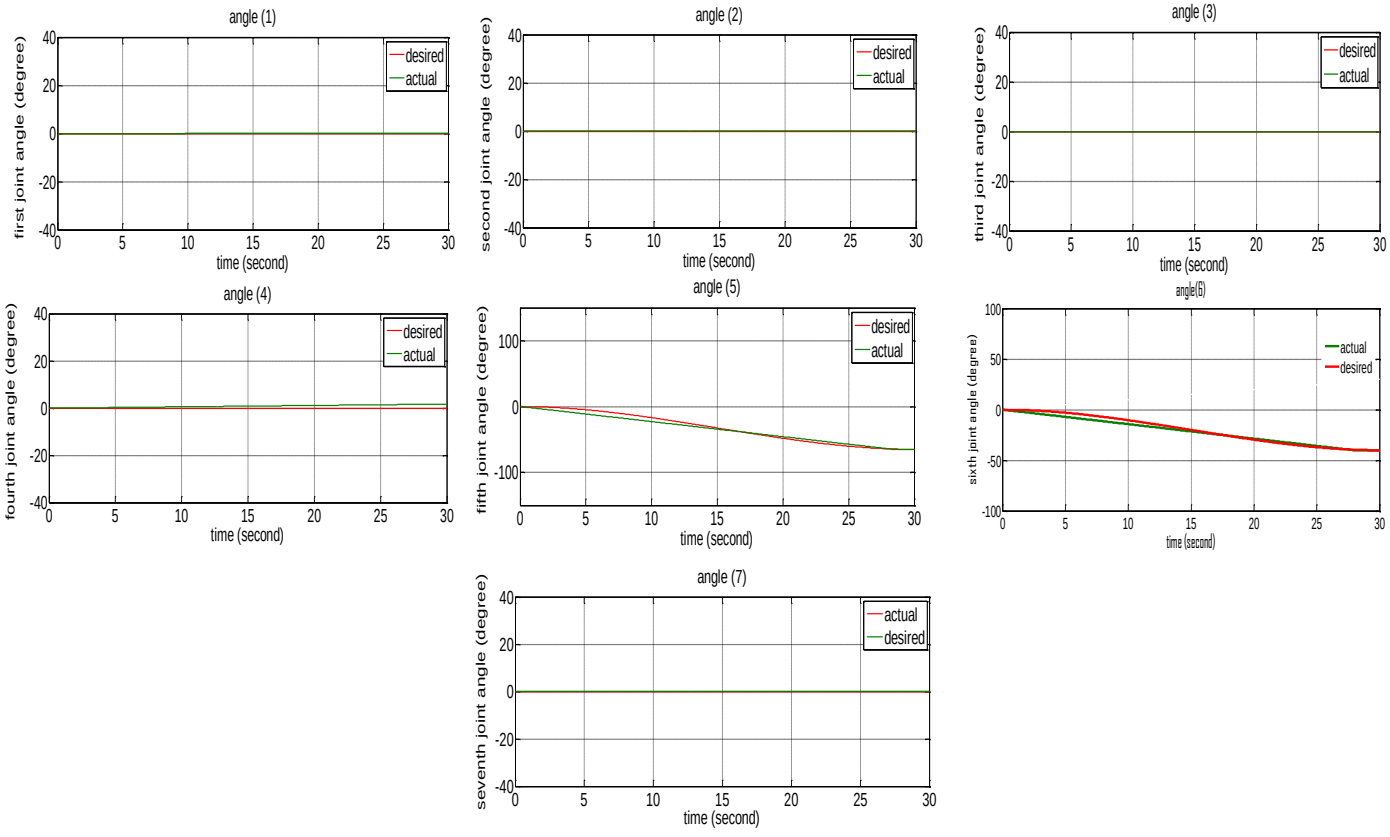


Fig.8: Joint angles responses for the first reference trajectory without disturbance (SNN controller).

4.3 Second Trajectory Without Applying Disturbance

The parameters of SNNs controllers are tuned as in Table (1)**.** The best gains of PI controller are shown in Table (4). The response for each joint angle using PI and SNNs controllers are shown in Figures (9) and (10) respectively.

Table 4: Best PI gains for the second trajectory without applying disturbance.

No. of joint	1	2	3	4	5	6	7
KP	2.2	13	1.5	12	0.25	38	0.04
KI	0.9	2.5	0.1	0.09	2.3	10.4	0.5

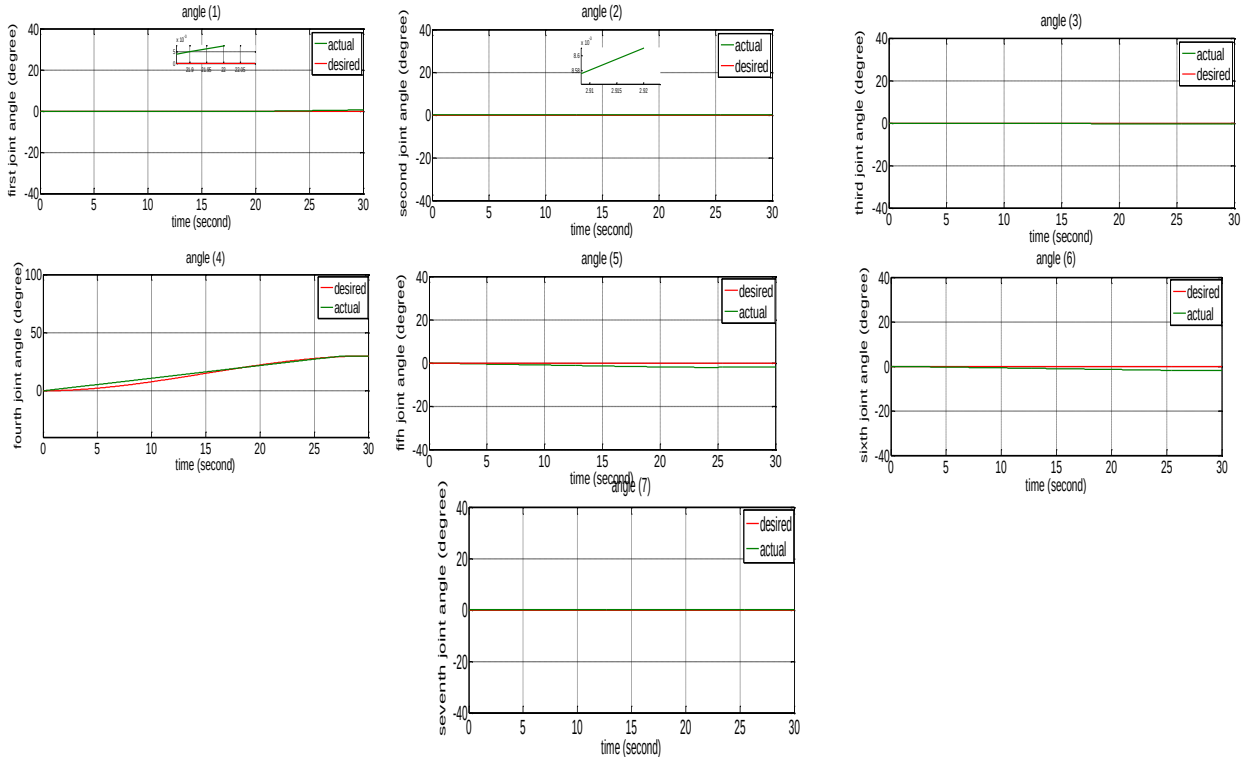


Fig.9: Joint angles responses for the second reference trajectory without applying disturbance (PI controller).

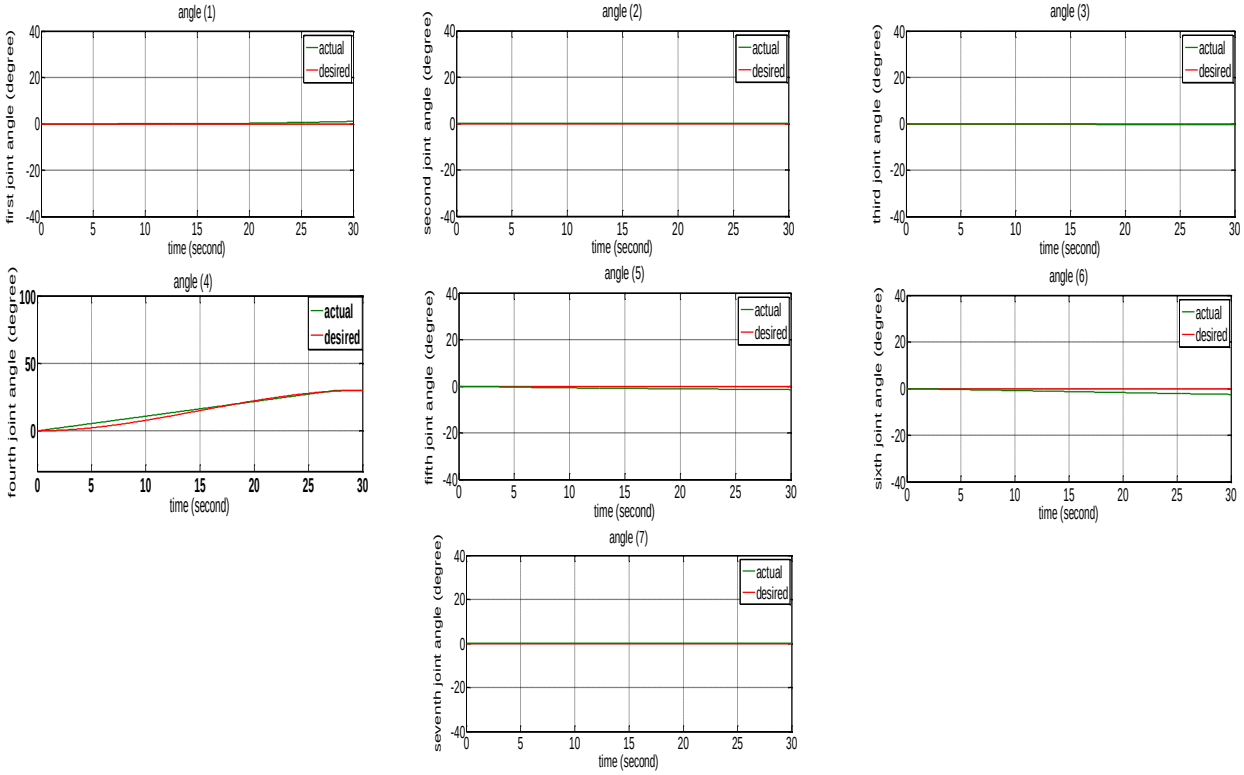


Fig.10: Joint angles responses for the second reference trajectory without disturbance (SNN controller).

4.4 Second Trajectory with Disturbance

The reasonable disturbance mentioned in equation (21) is applied. The best gains of PI controller are shown in Table (5). The responses for each joint angle using PI and SNNs controller are shown in Figures (11) and (12) respectively.

Table 5: Best PI gains for the Second Trajectory with Disturbance.

No. of joint	1	2	3	4	5	6	7
KP	9	2.3	2.8	5.3	4	21	24
KI	7.5	4.1	0.34	0.2	0.23	2.4	0.9

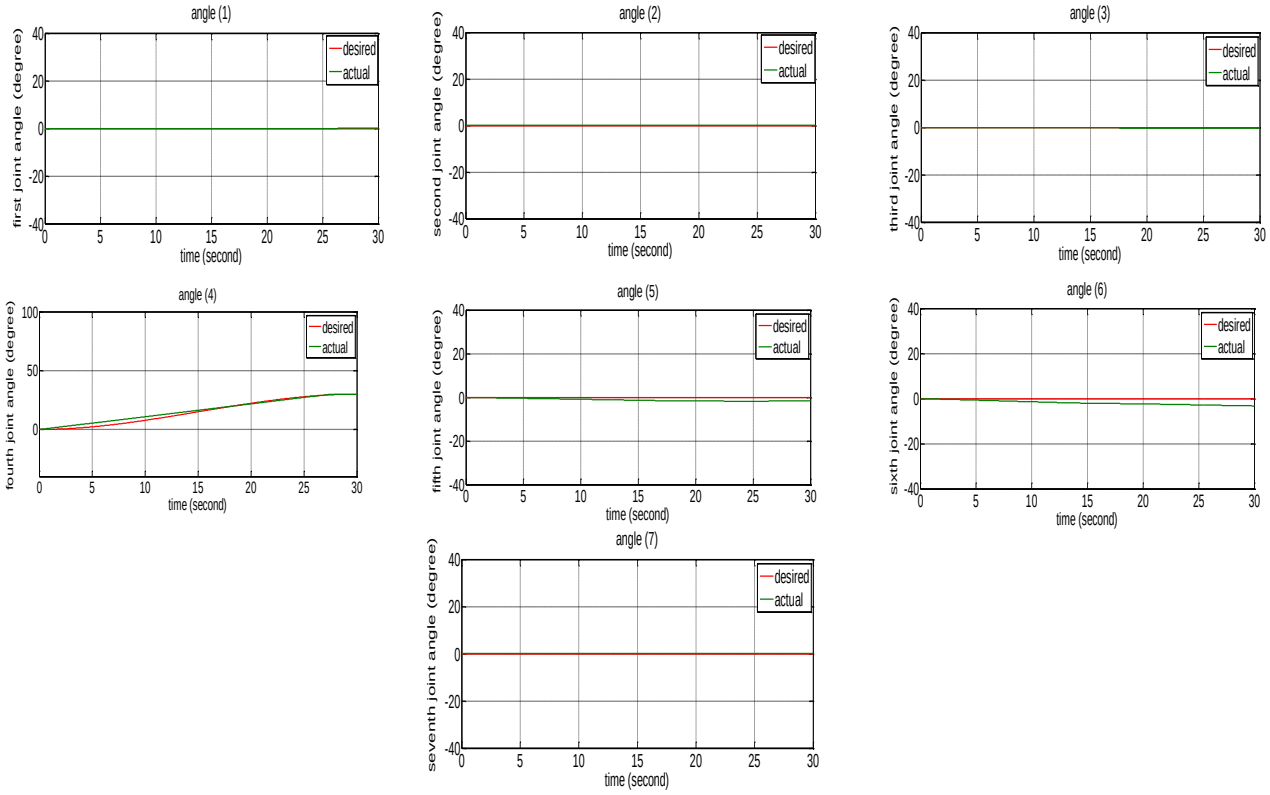


Fig.11: Joint angles responses for the second reference trajectory with applying disturbance (PI controller).

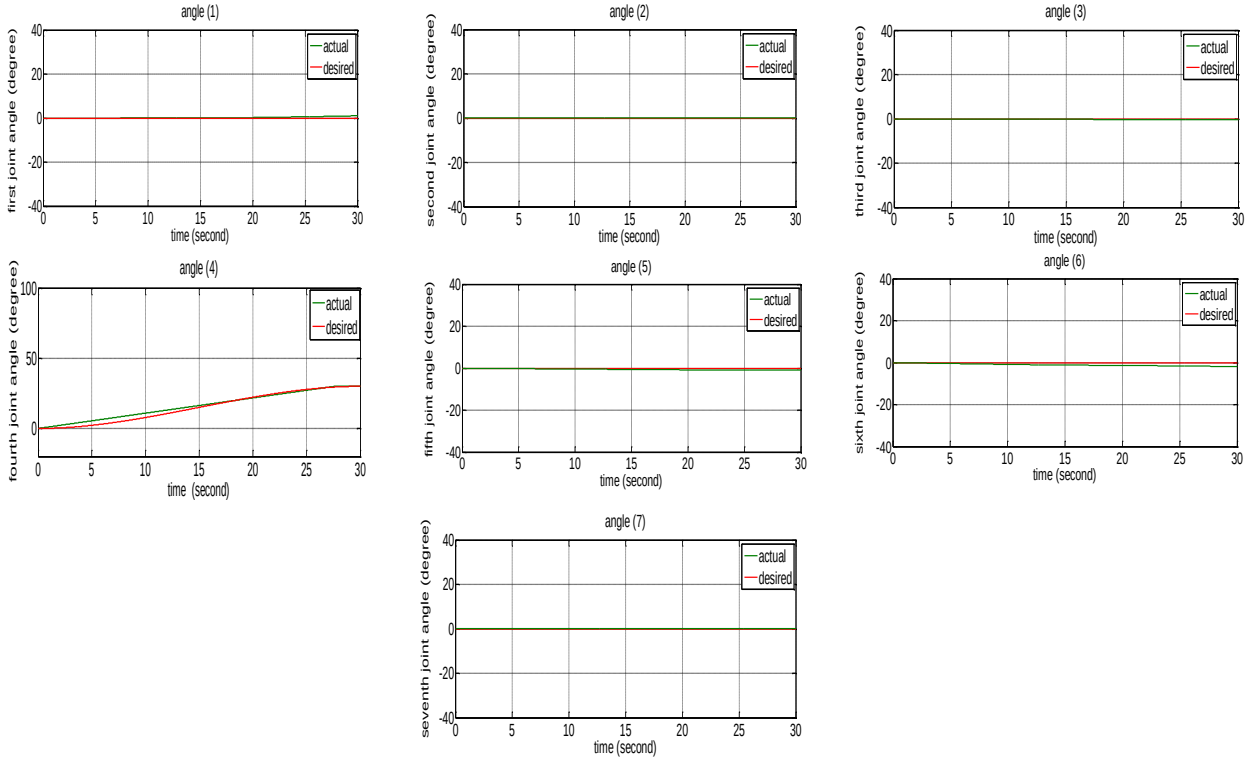


Fig.12: Joint angles responses for the second reference trajectory with disturbance (SNN controller).

5. Conclusions

In this paper, it can be shown from the previous results that the PI controller cannot keep the tracking trajectory, and the controller parameters (KP and KI) must be tuned with every change in trajectory and changing disturbance. Thus, Spiking Neural Network can solve this problem by training it to reach the best trajectory tracking performance using the same weights and strength. The SNN can adopt any trajectory and give the best performance with minimum error. This Table shows the enhancement in responses for the two trajectories applied.

The SNN designed was single input –single output which is equal to the conventional PI controller and this gives an advantage that the structure of controller is smaller and speedy in working.

References:

[1] S. Haykin, *Neural Networks: A Comprehensive Foundations*, Maxwell Macmillan, 1994.
[2] W. Maass, "Networks of spiking neurons the third generation of neural network models", Neural Networks, vol. 10, no. 9, pp. 1659-1671, 1997.

[3] W. Maass, "Fast Sigmoidal Networks via Spiking Neurons", Neural Computation, vol. 9, no. 2, pp. 279-304, 1997.
[4] S.M. Bohte, "Spiking Neural Networks", PhD. Thesis, 2003.
[5] S. Ghosh-Dastidara and H. Adeli, "Improved spiking neural networks for EEG classification and epilepsy and seizure detection", integrated Computer-Aided Engineering, vol. 14, pp. 187-212, 2007.
[6] S.M. Bohte, J.N. Kok and J.A. La Poutre, "Error backpropagation in Temporally Encoded Networks of Spiking Neurons", Neurocomputing, vol. 48, no. 1-4, pp. 17-37, 2002.
[7] S. M. Kennoch, D. Liu and L.G. Bushnell, "Fast Modifications of the Spike Prop Algorithm", International Joint Conference on Neural Networks, vol. 3970-3977,2006.
[8] W. Gerstner and W. M. Kistler, *Spiking Neuron Models: Single Neurons, Populations, Plasticity*, Cambridge University Press, 2002.
[9] O. Booij, "Temporal Pattern Classification using Spiking Neural Networks", M.Sc. Thesis, University Van Amsterdam, August 2004.
[10] M. Y. Hassan, J. J. Abdul Kareem," Design and Modeling of Six DOF Lumbar Spine and Lower Limb Rehabilitation Robot Structure", The First International Conference for Engineering Researches,2017.
[11] M. Y. Hassan, Z. A. Karam, "Modeling and Force-Position Controller Design of Rehabilitation Robot for Human Arm Movements", Eng. & Tech. Journal, Vol. 32, Part (A), No.8, 2014.
[12] Z. A. Karam, Design and Implementation of Rehabilitation Robot Arm, M.Sc. Thesis, University of Technology, Dec. 2013.

Variable Rounds Block Cipher Algorithm Design Based on Feistel Block Cipher

Louay F. Hasan

Al-Nisour University College (Private Sector)-Baghdad/ Iraq

E-mail: **lfhalbaghdady@gmail.com** Mob: (+964)7700484448

Abstract

One of the most important application of the secrecy and authenticity of secrete information is the cryptography. A cryptographic system based on Feistel cipher structure uses the same algorithm for both encryption and decryption. In each round, half of the input data goes through unchanged. In the proposed algorithm the left half L and the right half R both are goes through an operation depends on the round encryption key and L, R thus both L and R goes through changed in each round, this leads to possibility use one round only to encrypt/decrypt block of input data.

The proposed algorithm designed according to the second Kerckhoff's principle which is stated that a cryptographic system should be secure even if everything about the system, except the key, is public knowledge. virtually all the contemporary encryption algorithms such as DES, AES, etc. are applied this principle. This principle leads means that message depends solely on the security of the secret encryption key. In modern era, cryptosystem needs to cater to users who are connected to the Internet. In such cases, using a secret algorithm is not feasible, hence Kerckhoff's principles became essential guidelines for designing algorithms in modern cryptography.

The Feistel algorithm, such as DES, in real implementation instead of using the whole encryption key during each round, a round-dependent key a subkey is derived from the encryption key. This means that each round uses a different key, although all these subkeys are related to the original key. But in the proposed algorithm each round subkey is independent and unrelated to another different subkeys, this making the proposed system stronger, more secure and makes it more difficult to attack.

Keywords: Cryptosystems, Block Cipher, Proposed Algorithm, Encryption, Decryption, Round, Key, Plaintext, Ciphertext

الخلاصة

واحد من أهم التطبيقات في مجال سرية وموثوقية المعلومات هو التشفير .خوارزميات التشفير التي تستند في تصميمها الى نظام التشفير Feistel تستخدم نفس الخوارزمية في التشفير وحل الشفرة في كل جولة ، حيث تمر نصف بيانات الإدخال من دون اي تغيير . وفي الخوارزمية المقترحة ، فإن النصف الأيسر L والنصف الأيمن R يمران بعملية تعتمد على مفتاح التشفير للجولة و L و R وبالتالي يتغير كل من L و R خلال كل جولة ، وهذا يؤدي إلى إمكانية استخدام جولة واحدة فقط تشفير / وحل تشفير بيانات الإدخال. الخوارزمية المقترحة مصممة وفقًا لمبدأ كيرشوف الثاني الذي ينص على أن نظام التشفير يجب أن يكون آمنًا حتى إذا كان كل شيء عن النظام ، باستثناء المفتاح ، معروف ومعلن. يتم تطبيق جميع خوارزميات التشفير المعاصرة مثل DES و AES وما إلى ذلك .وفي هذا المبدأ يؤدي إلى أن الرسالة تعتمد فقط على أمان مفتاح التشفير السري .في العصر الحديث ، يحتاج نظام التشفير إلى تلبية احتياجات المستخدمين المتصلين بالإنترنت . في مثل هذه الحالة ، فأن استخدام خوارزمية سرية غير ممكن ، لذلك أصبحت مبادئ كيرشوف أساسية لتصميم الخوارزميات في التشفير الحديث.

خوارزمية Feistel ، مثل DES ، في التنفيذ الحقيقي بدلاً من استخدام مفتاح التشفير بالكامل أثناء كل جولة ، يتم اشتقاق مفتاح فرعي للجولة اعتماداً على مفتاح التشفير .وهذا يعني أن كل جولة تستخدم مفتاحًا مختلفًا ، على الرغم من أن جميع المفاتيح الفرعية هذه مرتبطة بالمفتاح الأصلي .ولكن في الخوارزمية المقترحة ، يكون كل مفتاح فرعي مستقلاً وغير مرتبط بمفاتيح فرعية أخرى، مما يجعل النظام المقترح أكثر قوة وأكثر أمانًا ويجعل الهجوم أكثر صعوبة.

Introduction

Fundamentally, there are two types of cryptosystems based on the manner in which encryption-decryption is carried out in the system, Symmetric Key Encryption and Asymmetric Key Encryption. A block cipher is one of symmetric cryptographic system, in a symmetric-key cryptography Figure-1, which involves both sender and receiver using a common key. Block cipher takes a block of plaintext bits and produce a block of ciphertext bits, in general both pair has the same size. The strength of block cipher algorithm depends on the key length and the algorithm complexity [1], while the strength of encryption algorithm doesn't directly affect by that block size [2].

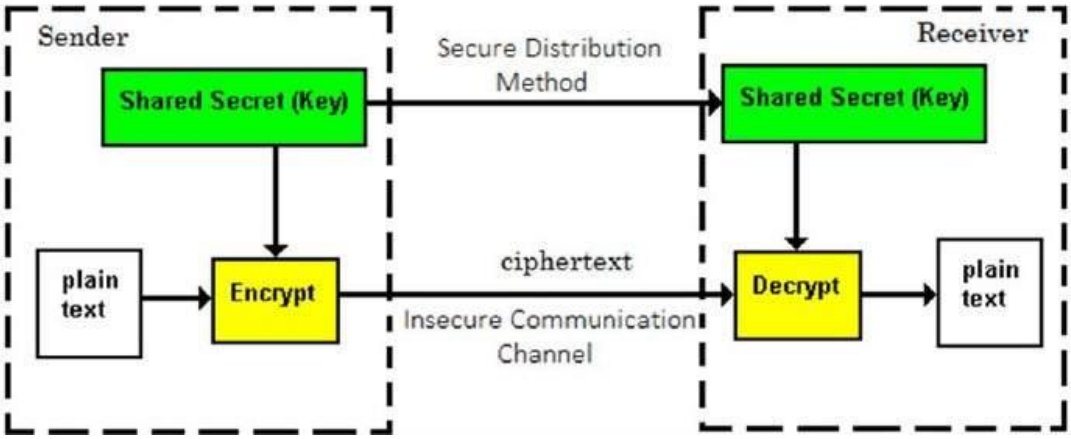


Fig.1: The typical block diagram of a block cipher

Feistel Algorithm Structure

Feistel algorithm is a design model from which many different block ciphers are derived [3]. DES is just one example of a Feistel Cipher.

Encryption Process

The encryption process uses the Feistel structure consisting multiple rounds of processing of the plaintext. Feistel Structure is shown in Figure-2. The input block to each round is divided into two halves that can be denoted as L and R for the left half and the right half. In each round, the right half of the block, R, goes through unchanged. But the left half, L, goes through an operation that depends on R and the encryption key.

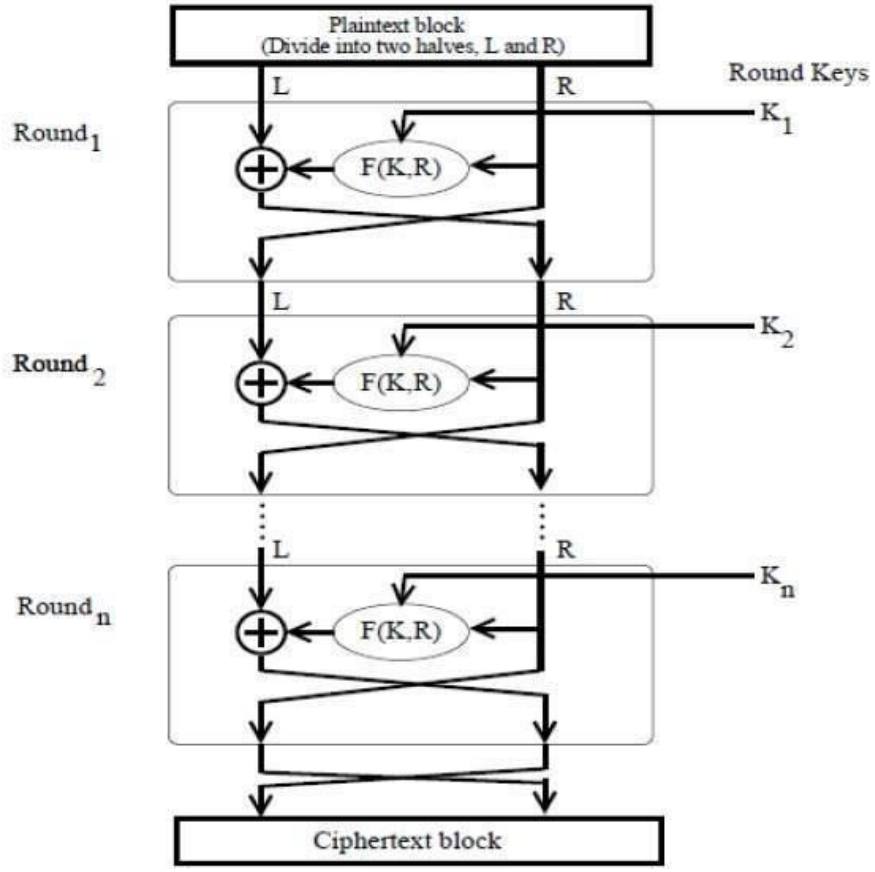


Fig. 2: The Feistel Algorithm Structure

First, we apply an encrypting function f that takes two inputs, the key K and R. The function produces the output fR,K. Then, we XOR (Exclusive OR) the output of the logical function with L. In implementation of the Feistel Cipher, DES round dependent key a subkey is derived from the encryption key. This leads that each round uses a different key, although all these subkeys are related to the original key.

The permutation step at the end of each round swaps the modified L and unmodified R. Therefore, the L for the next round would be R of the current round, and vice versa. Above substitution and permutation steps form a round. The number of rounds is specified by the algorithm design [3]. Once the last round is completed then the two sub blocks, right R and left L are concatenated in this order to form the ciphertext block. The difficult part of designing a Feistel Cipher is selection of round function f [4].

Decryption Process

The process of decryption in Feistel cipher is almost similar. Instead of starting with a block of plaintext, the ciphertext block is fed into the start of the Feistel structure and then the process there after is exactly the same as described in the given illustration.

The process is said to be almost similar and not exactly same. In the case of decryption, the only difference is that the subkeys used in encryption are used in the reverse order. The final swapping of L and R in last step of the Feistel Cipher is essential. If these are not swapped, then the resulting ciphertext could not be decrypted using the same algorithm.

Number of Rounds

The number of rounds used in a Feistel Cipher depends on desired security from the system. More number of rounds provides more secure system. But at the same time, more rounds mean the inefficient slow encryption and decryption processes. Number of rounds in the systems thus depends upon efficiency and security tradeoff [3].

Confusion and Diffusion

Shannon suggested the use of confusion and diffusion as two techniques that ciphers should implement to deter cryptanalysts [5].

Diffusion Dissipate the statistical properties of the plaintext over the whole ciphertext [6]. Concretely, for a block cipher, if each bit of the output block changes with 50% probability for any flipped bit in the input block, the cipher has good diffusion properties. For good diffusion properties to exist, the cipher has to employ shuffling or reordering of bits (permutation operations).

Confusion refers to the property that the ciphertext (and its statistical properties) is not easily relatable to the key [6]. This Means the Key must involve non-linear substitution algorithms. the Feistel Cipher structure design combines permutations and substitutions effectively.

The popular Block Cipher Algorithms

There is a various number of block ciphers algorithms that are publically known. Most popular and prominent block cipher:

Data Encryption Standard (DES) Algorithm

Now considered as a broken, due primarily to its small key size [7]. The plaintext and the ciphertext are 64-bit block size, while the secret key is 56 -bit long (eight bytes, in each byte seven bits are used; the eighth bit can be used as a parity bit). Iterates a round-function16 times in 16 rounds. The round-function mixes the data with the key [8].

Triple DES (3DES) Algorithm

this algorithm based on repeated DES [8,9]. It is still a respected encryption algorithm but inefficient compared to the new faster block ciphers available [5]. The plaintext and the ciphertext are 64-bit block size, while the secret key is defined three keying options, firstly all three keys are independent and key size=3x56=168
bits, secondly K1 and K2 are independent, and K3 = K1 and key size 112 bits. Thirdly all three keys are identical, i.e. K1 = K2 = K3.

Advanced Encryption Standard (AES) Algorithm

It is a relatively new block cipher based on the encryption algorithm Rijndael [10] allows many block sizes and key sizes. AES restricts it to Block Size 128 bits and allows 128-bit, 192-bit, 256-bit keys. AES number of rounds is 10, 12 or 14, for a 128-bit, 192-bit or 256-bit key, respectively.

IDEA Algorithm, International Data Encryption Algorithm a strong Block cipher algorithm[11,12] with plaintext and the ciphertext are 64-bit blocks, while the secret key is 128-bit long. IDEA encrypt block in 8 rounds.

Twofish Algorithm, it is based on the Blowfish block cipher algorithm with a block size of 64 bits. This block cipher algorithm uses block size of 128 bits and accepts keys of any length, up to 256 bits and round number is 16 [2,13].

Serpent Algorithm, Strong block cipher algorithm. It is slower but has more secure design than other block cipher [14] uses block size of 128 bits and key size of either 128,192, or 256 bits, it consisting of 32 rounds.

The Proposed Algorithm

In the Proposed encryption algorithm encrypts the block of plaintext and produce a block of ciphertext.

We have $c = E_k(m)$ …… (1)

where:

_ m is the plaintext block

_ k is the secret key

_ E is the encryption function

_ C is the ciphertext

In addition, we have a decryption function D such that:

$m = D_k(C)$ …… (2)

Encryption Process

Plaintext is divided into blocks of 192-bit fixed length and every block is encrypted one at a time. The encryption process of a plaintext block depicted in Figure-3

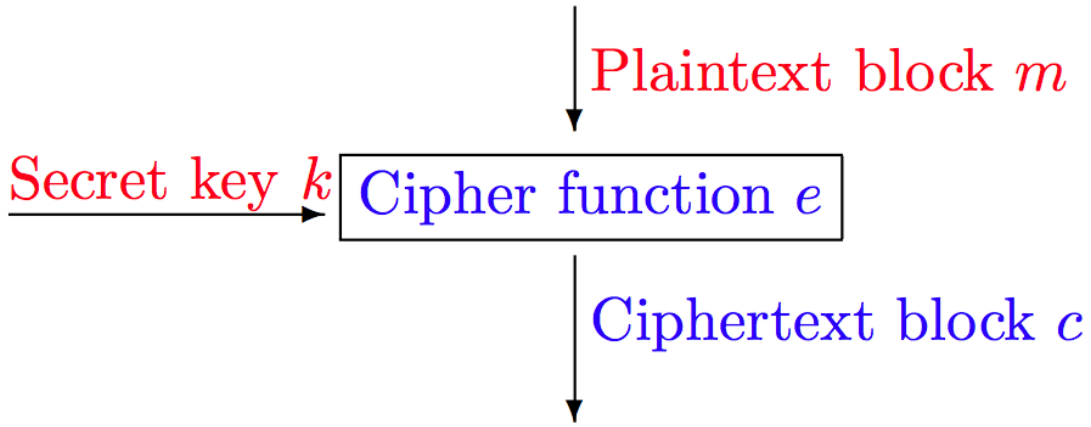


Fig. 3: The Encryption Process

Round Plaintext Block

The encryption algorithm encrypts block 192 bits of plaintext at a time, plaintext block is splitted into left and right equal half of 96 bit for each, both halves are divided into three 32 bits (4 bytes), the left half groups L_{00}, L_{01}, L_{02} and the right half groups R_{00}, R_{01}, R_{02}.The plaintext if necessary may need to pad, the length of plaintext is mostly not a multiple of the block size. For example, a 150-bit plaintext provides two blocks of 64 bits each with third block of balance 22 bits. The last block of bits needs to be padded up with redundant information so

that the length of the final block equals to block size of the scheme. In our example, the remaining 22 bits need to have additional 42 redundant bits added to provide a complete block. The process of adding bits to the last block is referred to as padding. Too much padding makes the system inefficient [15]. Also, padding may render the system insecure at times if the padding is done with same bits always.

Round Key Block

round key or subkey block size 176 bits (22 byte), each byte in the block swapped the least significant nibble with the most significant nibble as in the *Figure-4.*

27 26 25 24 31 30 29	19 18 17 16 23 22 21	11 10 9 8 15 14 13 12	3 2 1 0 7 6 5

(a

31 30 29 28 27 26 25	23 22 21 20 19 18 17	15 14 13 12 11 10 9 8	7 6 5 4 3 2 1

(b

Fig.4: Subkey swapping (a) the original subkey (b) the subkey block after swap

First 96 bits (12 byte) grouped into three 32 bits k00, k01, k02, 72 bits (9 byte) as input to F-Function, the remaining 8 bits are input to G-Function.

Mixing exclusive-or operation between the left half round block input and the round key.

$Y_{i,j}=L_{i,j} \oplus k_{i,j}$, ……(3)

where i=0, 1, 2, j=0, 1, 2, 3, bitwise XOR denoted $\oplus$

The encryption key depicted

K=n*k …… (4)

where n is number of rounds and k is a round key

F-Function

It consists of 3 x-box each x-box has 4 s-boxes of 256 random bytes in contents of each, In an s-box each output byte can be represented as a function of its input bits from eq. (1):

$R_{ij}=F(Y_{ij})$ …… (5)

Y_{ij} is the address of byte in $S\text{-box}_{i,j}$

K_{con} 72 bits key is splitted into 3 x 24 bits, each block of 24 bits to control in X-box, which is divided into 4x6 bits K_{con00}, K_{con01}, K_{con02}, K_{con03}, $K_{coni,j}$ represented as (w y x c b a) to control the produced byte from S-box as in table-1

Table1: Round SubKey K_{con} 6-bit Group

w	y	x	$Z=c*2^2+b*2+a$
0	0	0	R byte Shift left Z bits
0	0	1	R byte Swap Ln, Mn
0	1	0	R byte Swap Ln, Mn
0	1	1	R byte Shift left Z bits
1	0	0	R byte Swap Ln, Mn
1	0	1	R byte Shift right Z bits
1	1	0	R byte Shift right Z bits
1	1	1	R byte Swap Ln, Mn

F-Function output bits can be represented as a function of its input bits K_{con} and from:

$X_i=F(Y_i,K_{coni})$ ……(6), where i=0, 1, 2

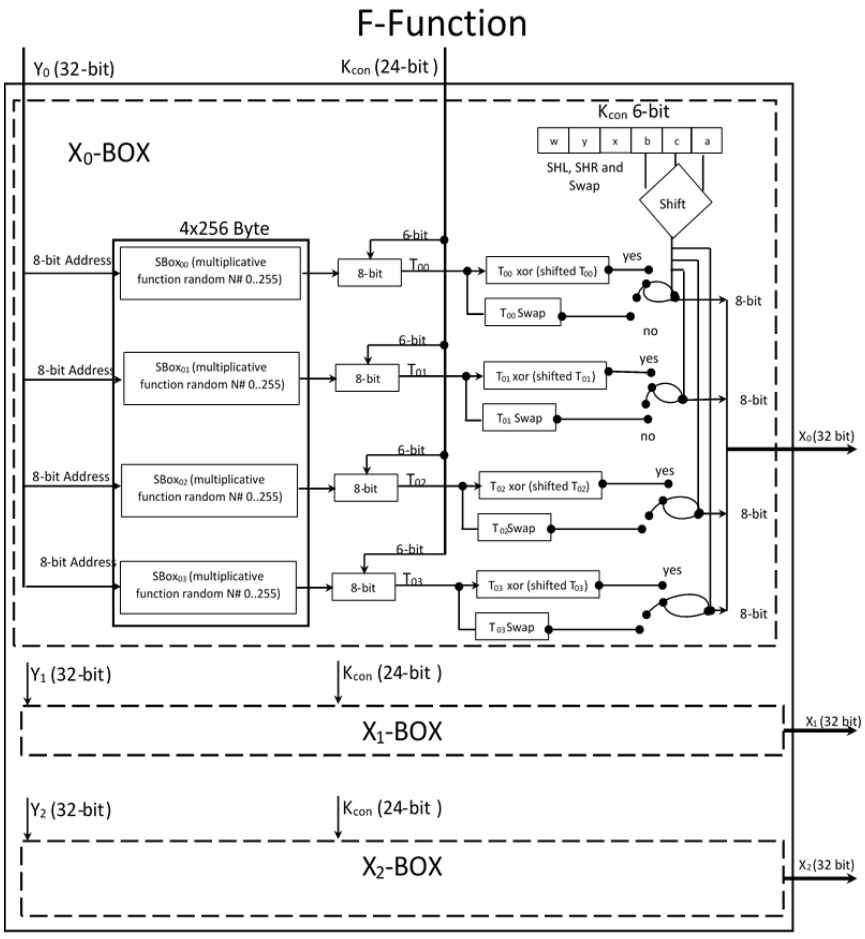
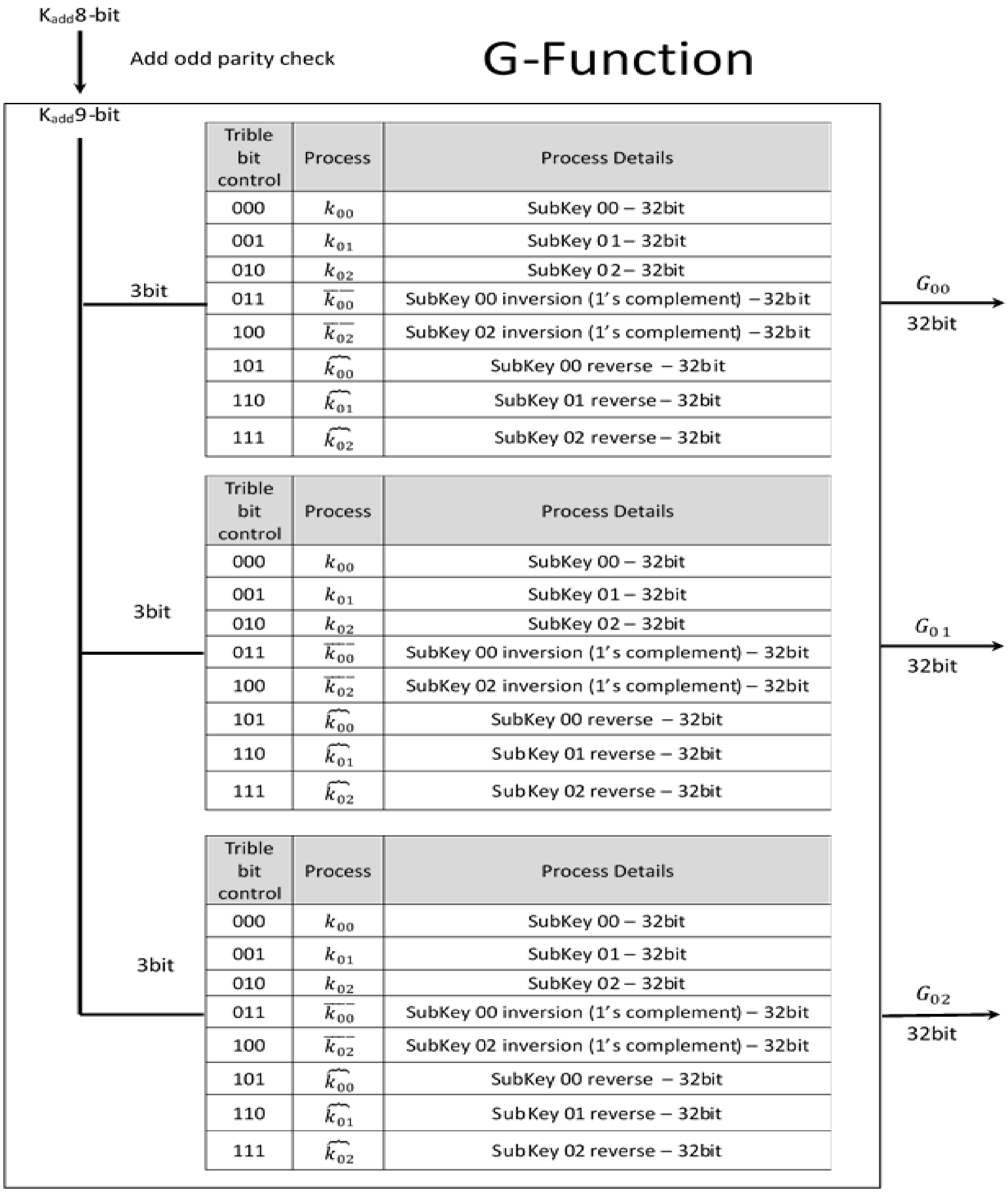


Fig.5: F-Function

G-Function

8 bits round key K_{add} is expanded into 9 bits by adding odd parity check bit to the most K_{add} byte, then splitting the produced 9 bits into three 3 bits, G-Function output byte can be represented as a function of its input bits:

$G=F(EK_{add},K_{00},k_{01},k_{02})$ ……………(7)



Trible bit control	Process	Process Details
000	k_{00}	SubKey 00 – 32bit
001	k_{01}	SubKey 01 – 32bit
010	k_{02}	SubKey 02 – 32bit
011	$\overline{k_{00}}$	SubKey 00 inversion (1's complement) – 32bit
100	$\overline{k_{02}}$	SubKey 02 inversion (1's complement) – 32bit
101	$\overset{\frown}{k_{00}}$	SubKey 00 reverse – 32bit
110	$\overset{\frown}{k_{01}}$	SubKey 01 reverse – 32bit
111	$\overset{\frown}{k_{02}}$	SubKey 02 reverse – 32bit

Trible bit control	Process	Process Details
000	k_{00}	SubKey 00 – 32bit
001	k_{01}	SubKey 01 – 32bit
010	k_{02}	SubKey 02 – 32bit
011	$\overline{k_{00}}$	SubKey 00 inversion (1's complement) – 32bit
100	$\overline{k_{02}}$	SubKey 02 inversion (1's complement) – 32bit
101	$\overset{\frown}{k_{00}}$	SubKey 00 reverse – 32bit
110	$\overset{\frown}{k_{01}}$	SubKey 01 reverse – 32bit
111	$\overset{\frown}{k_{02}}$	SubKey 02 reverse – 32bit

Trible bit control	Process	Process Details
000	k_{00}	SubKey 00 – 32bit
001	k_{01}	SubKey 01 – 32bit
010	k_{02}	SubKey 02 – 32bit
011	$\overline{k_{00}}$	SubKey 00 inversion (1's complement) – 32bit
100	$\overline{k_{02}}$	SubKey 02 inversion (1's complement) – 32bit
101	$\overset{\frown}{k_{00}}$	SubKey 00 reverse – 32bit
110	$\overset{\frown}{k_{01}}$	SubKey 01 reverse – 32bit
111	$\overset{\frown}{k_{02}}$	SubKey 02 reverse – 32bit

Round Encryption Algorithm

Mixing XOR operation between the left half round input and the round key are input to F-Function, the output Xi from (3) mixing right half round input as in equation (5)

$T_{0i}=R_{0i} \oplus X_{0i}$ …… (8)

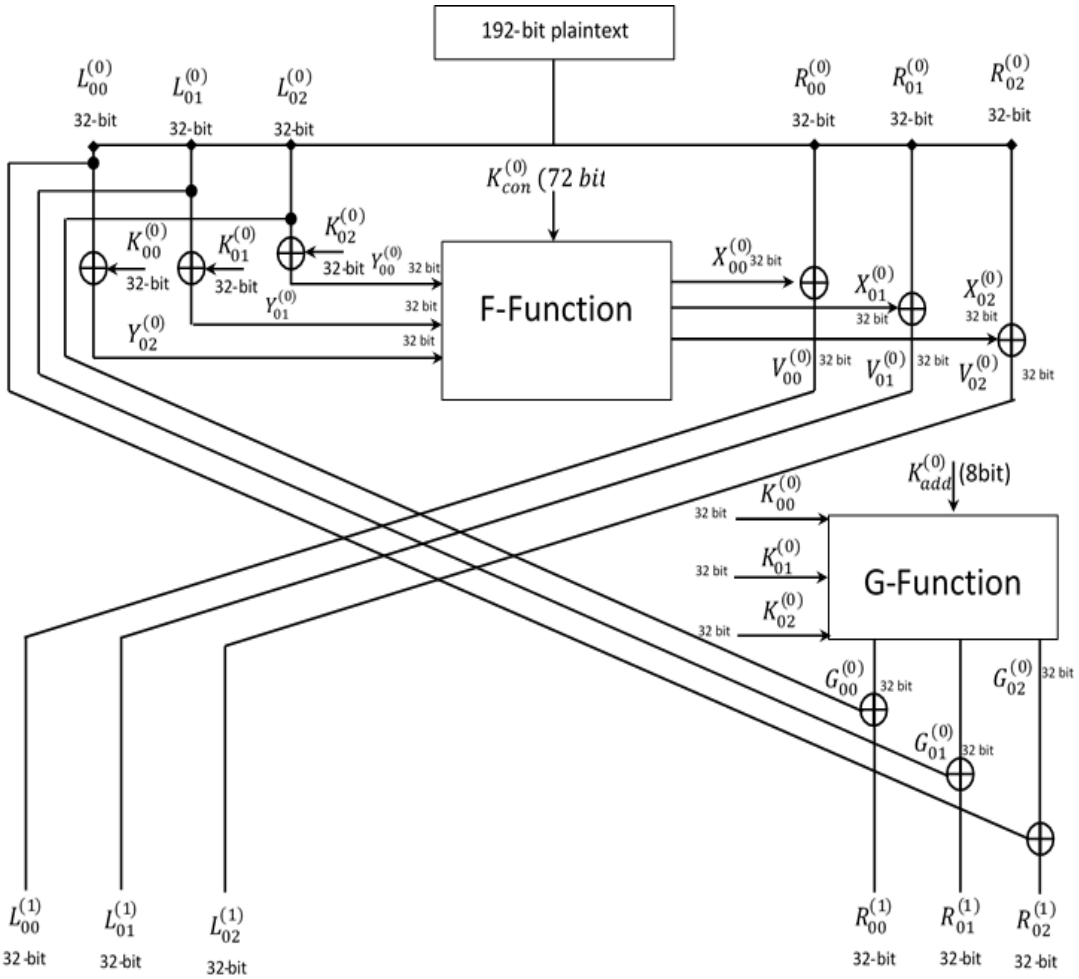
Mixing XOR operation between the left half round input and the resulted output from G-Function G_i from (4) as in equation (6).

$H0i=L0i \oplus G0i$ …… (9)

Switch the left half result with right half result

$L_{1i}=T_{0i}$ and $R_{1i}=H_{0i}$

An iterated block cipher involves repeated use of a round function and dependent on the number of rounds, for every subkey the round function is designed as invertible; if not decryption is impossible [16].



Round Ciphertext Block

Cipher text is splitted into left and right half of 96 bit for each, both halves bits are divided into 3 blocks of length 32 bits or 4 bytes (for each block), the left half blocks L_{00}, L_{01}, L_{02} and the right half blocks R_{00}, R_{01}, R_{02}.

Each plaintext byte in the block the least significant nibble swapped with the most significant nibble as in the Figure 2 Round F-Function Input

$Y_{0i}=K_{0i} \oplus L_{0i} \oplus G_{0i}$ …… (10)

Round Decryption Algorithm

Decryption is given by eq. (2), both sides need to know the key k, but k needs to be kept secret [16, 17]. Mixing XOR operation between the left half round input and the resulted output from G-Function G_i as in eq. (11).

$H_{0i}=L_{0i} \oplus G_{0i}$ …… (11)

Mixing XOR operation between the left half round input and the round key are input to F-Function, the output Xi mixing right half round input as in eq. (12)

$T_{0i}=R_{0i} \oplus X_{0i}$ …… (12)

Switch the left half result with right half result

$L_{1i}=T_{0i}$ and $R_{1i}=H_{0i}$ …… (13)

An iterated block cipher involves repeated use of a round function and depends on the number of rounds

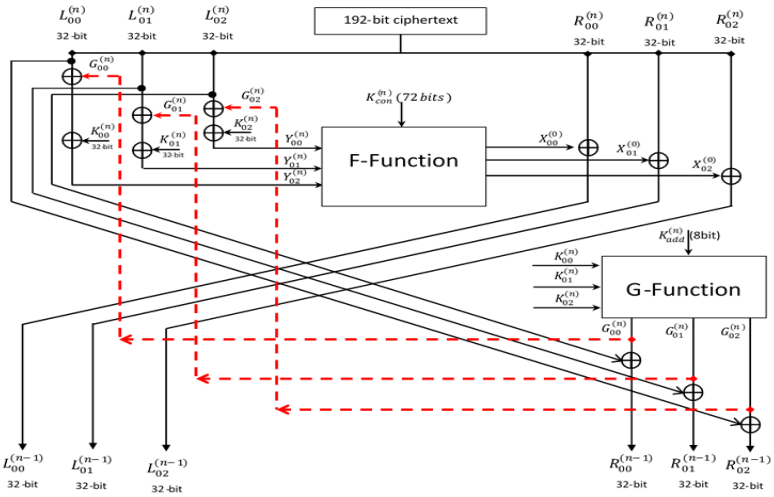


Fig.8: Decryption Algorithm Structure

Iterated Block Ciphers

An iterated block cipher involves repeated use of a round functions. The design makes a for strong encryption invertible functions easy to implement by repeatedly using it. Each use of the round function employs a subkey derived from the encryption key k, but unrelated with other round keys.

Simulation Results

The software implementation of the proposed encryption algorithm are tested and executed for variable rounds from 1 to 18 in both modes encryption/decryption and the processes success in all cases, the following sections takes 2 states in case of one round and two rounds operation modes.

One Round Encryption Mode

Round Number (1)

==========

The Original Round Key: 0BCBFD053C448EF365D0F60C67B39A9EDA5477BDFDB4

The Swapped Round Key: B0BCDF50C344E83F560D6FC0763BA9E9AD4577DBDF4B

the Left half : 00011122233344455566777

the Right half : 888999AAABBBCCCDDDEEEFFF

G-Function input Parameters: K0= B0BCDF50 K1= C344E83F K2= 560D6FC0 Kadd= 4B

G-Function Output: T0= 560D6FC0 T1= 560D6FC0 T2= B0BCDF50

F-Function input Parameters : Y00= B0 Y01= BD Y02= CE Y03= 72 Y10= E0 Y11= 77 Y12= AC Y13= 7A Y20= 03 Y21= 6B Y22= 08 Y23= B7

Kcon00= 05 Kcon01= 00 Kcon02= 0E Kcon03= 09 Kcon10= 00 Kcon11= 02 Kcon12= 0D Kcon13= 05 Kcon20= 05 Kcon21= 00 Kcon22= 0F Kcon23= 07

F-Function output : X00= 0F X01= 42 X02= 80 X03= 75 X10= 42 X11= C6 X12= 04 X13= 5A X20= 18 X21= 09 X22= 39 X23= E7

Block Ciphertext:

8 7 C B 1 9 D F E 9 7 D C 8 9 7 C 5 E 7 D 6 1 8 5 6 0 C 7 E E 2 7 5 3 E 2 B 8 5 E 5 D A B 8 2 7

One Round Decryption Mode

Round Number (1)

==========

The Original Round Key: 0BCBFD053C448EF365D0F60C67B39A9EDA5477BDFDB4

The Swapped Round Key: B0BCDF50C344E83F560D6FC0763BA9E9AD4577DBDF4B

the Left half : 560C7EE2753E2B85E5DAB827

the Right half : 87CB19DFE97DC897C5E7D618

G-Function input Parameters: K0= B0BCDF50 K1= C344E83F K2= 560D6FC0 Kadd= 4B

G-Function Output: T0= 560D6FC0 T1= 560D6FC0 T2= B0BCDF50

F-Function input Parameters : Y00= B0 Y01= BD Y02= CE Y03= 72 Y10= E0 Y11= 77 Y12= AC Y13= 7A Y20= 03 Y21= 6B Y22= 08 Y23= B7

Kcon00= 05 Kcon01= 00 Kcon02= 0E Kcon03= 09 Kcon10= 00 Kcon11= 02 Kcon12= 0D Kcon13= 05 Kcon20= 05 Kcon21= 00 Kcon22= 0F Kcon23= 07

F-Function output : X00= 0F X01= 42 X02= 80 X03= 75 X10= 42 X11= C6 X12= 04 X13= 5A X20= 18 X21= 09 X22= 39 X23= E7

Block Plaintext:

0 0 0 1 1 1 2 2 2 3 3 3 4 4 4 5 5 5 6 6 6 7 7 7 8 8 8 9 9 9 A A A B B B C C C D D D E E E F F F

Two Rounds Encryption Mode

Round Number (1)

==========

The Original Round Key: 4A4120DA31402CBCB89314212914A2A0E19D8FB116A7

The Swapped Round Key: A41402AD1304C2CB8B39411292412A0A1ED9F81B617A

the Left half: 000111222333444555666777

the Right half: 888999AAABBBCCCDDDEEEFFF

G-Function Input Parameters: K0= A41402AD K1= 1304C2CB K2= 8B394112 Kadd= 7A

G-Function Output: T0= 5BEBFD52 T1= 8F3F2986 T2= A41402AD

F-Function Input Parameters: Y00= A4 Y01= 15 Y02= 13 Y03= 8F Y10= 30 Y11= 37 Y12= 86 Y13= 8E Y20= DE Y21= 5F Y22= 26 Y23= 65

Kcon00= 00 Kcon01= 00 Kcon02= 04 Kcon03= 0A Kcon10= 02 Kcon11= 00 Kcon12= 00 Kcon13= 01 Kcon20= 00 Kcon21= 01 Kcon22= 0D Kcon23= 00

F-Function output: X00= 41 X01= CC X02= 1F X03= C4 X10= 40 X11= CD X12= 41 X13= 3A X20= 24 X21= CB X22= 01 X23= 4A

Round Number (2)

==========

The Original Round Key: 260E01A7726FA606DB5CB5E53B8DD6E4792CF85BA07E

The Swapped Round Key: 62E0107A27F66A60BDC55B5EB3D86D4E97C28FB50AE7

the Left half : C945866EEB768DF7F925EEB5

the Right half : 5BEAEC70AC0C6DC3F17265DA

G-Function input Parameters: K0= 62E0107A K1= 27F66A60 K2= BDC55B5E Kadd= E7

G-Function Output: T0= 0CDD414B T1= 27F66A60 T2= 62E0107A

F-Function input Parameters : Y00= AB Y01= A5 Y02= 96 Y03= 14 Y10= CC Y11= 80 Y12= E7 Y13= 97 Y20= 44 Y21= E0 Y22= B5 Y23= EB

Kcon00= 0C Kcon01= 00 Kcon02= 00 Kcon03= 0D Kcon10= 00 Kcon11= 09 Kcon12= 07 Kcon13= 02 Kcon20= 00 Kcon21= 00 Kcon22= 00 Kcon23= 00

F-Function output : X00= C3 X01= C3 X02= 0F X03= 47 X10= 89 X11= 7C X12= 16 X13= 43 X20= 49 X21= 96 X22= F4 X23= 20

Block Ciphertext:

9 8 2 9 E 3 3 7 2 5 7 0 7 B 8 0 B 8 E 4 9 1 F A C 5 9 8 C 7 2 5 C C 8 0 E 7 9 7 9 B C 5 F E C F

Two Rounds Decryption Mode

Round Number (1)

==========

The Original Round Key: 260E01A7726FA606DB5CB5E53B8DD6E4792CF85BA07E

The Swapped Round Key: 62E0107A27F66A60BDC55B5EB3D86D4E97C28FB50AE7

the Left half : C598C725CC80E7979BC5FECF

the Right half : 9829E33725707B80B8E491FA

G-Function input Parameters: K0= 62E0107A K1= 27F66A60 K2= BDC55B5E Kadd= E7

G-Function Output: T0= 0CDD414B T1= 27F66A60 T2= 62E0107A

F-Function input Parameters : Y00= AB Y01= A5 Y02= 96 Y03= 14 Y10= CC Y11= 80 Y12= E7 Y13= 97 Y20= 44 Y21= E0 Y22= B5 Y23= EB

Kcon00= 0C Kcon01= 00 Kcon02= 00 Kcon03= 0D Kcon10= 00 Kcon11= 09 Kcon12= 07 Kcon13= 02 Kcon20= 00 Kcon21= 00 Kcon22= 00 Kcon23= 00

F-Function output : X00= C3 X01= C3 X02= 0F X03= 47 X10= 89 X11= 7C X12= 16 X13= 43 X20= 49 X21= 96 X22= F4 X23= 20

Round Number (2)

==========

The Original Round Key: 4A4120DA31402CBCB89314212914A2A0E19D8FB116A7

The Swapped Round Key: A41402AD1304C2CB8B39411292412A0A1ED9F81B617A

the Left half : 5BEAEC70AC0C6DC3F17265DA

the Right half : C945866EEB768DF7F925EEB5

G-Function input Parameters: K0= A41402AD K1= 1304C2CB K2= 8B394112 Kadd= 7A

G-Function Output: T0= 5BEBFD52 T1= 8F3F2986 T2= A41402AD

F-Function input Parameters : Y00= A4 Y01= 15 Y02= 13 Y03= 8F Y10= 30 Y11= 37 Y12= 86 Y13= 8E Y20= DE Y21= 5F Y22= 26 Y23= 65

Kcon00= 00 Kcon01= 00 Kcon02= 04 Kcon03= 0A Kcon10= 02 Kcon11= 00 Kcon12= 00 Kcon13= 01 Kcon20= 00 Kcon21= 01 Kcon22= 0D Kcon23= 00

F-Function output : X00= 41 X01= CC X02= 1F X03= C4 X10= 40 X11= CD X12= 41 X13= 3A X20= 24 X21= CB X22= 01 X23= 4A

Block Plaintext:

0 0 0 1 1 1 2 2 2 3 3 3 4 4 4 5 5 5 6 6 6 7 7 7 8 8 8 9 9 9 A A A B B B C C C D D D E E E F F F

Conclusions

The proposed block cipher algorithm was designed to give the desired security according to number of rounds. The number of round varies from 1 to the extent that it achieves the degree of security required. More number of rounds provide more secure system. The use of two invertible round functions with the left and right half iterates a round-functions n-times in n-rounds. The round-functions mixed the data with the key n times provided the possibility of using a variable round. Normally all block size is a multiple of 8 as it is easy for implementation as most computer processor handle data in multiple of 8 bits.

References

[1] Mansoor Ebrahim, Shujaat Khan, Umer Bin Khalid," Symmetric Algorithm Survey: A Comparative Analysis", International Journal of Computer Applications (0975 – 8887) Volume 61– No.20, January 2013.

[2] Henk C.A. van Tilborg،Sushil Jajodia," Encyclopedia of Cryptography and Security", Springer Science & Business Media, second edition, 2014.

[3] Valerie Nachef, Jacques Patarin, Emmanuel Volte," Feistel Ciphers: Security Proofs and Cryptanalysis", Springer, February 2017.

[4] Lars R. Knudsen،Matthew Robshaw, "The Block Cipher Companion", Springer Science & Business Media., 2011.

[5] Keith M. Martin," Everyday Cryptography: Fundamental Principles and Applications", Oxford University Press, 2012.

[6] Atul Kahate," Cryptography and Network Security", Tata McGraw-Hill Education., 2003.

[7] Eli Biham, Adi Shamir, " Differential Cryptanalysis of the Data Encryption Standard ", Springer Science & Business Media, December 2012.

[8] 6+ Source Wikipedia ,"Block Ciphers: Data Encryption Standard, Blowfish, Triple Des, Advanced Encryption Standard, International Data Encryption Algorithm, Block Cipher, Rc", General Books, 2013, ISBN 1230550313, 9781230550312.

[9] NIST Special Publication 800-67 Revision 1, "Recommendation for Triple Data Encryption Algorithm (TDEA) Block Cipher", January 2012.

[10] Joan Daemen, Vincent Rijmen, "The Design of Rijndael: AES - The Advanced Encryption Standard", Springer Science & Business Media, March 2013.

[11] Simon Singh," The Code Book: The Secrets Behind Codebreaking", Random House Children's Books, 2001.

[12] Thomas Baigneres،Pascal Junod،Serge," A Classical Introduction to Cryptography Exercise Book", Springer Science & Business Media, 2006.

[13] Bruce Schneier," The twofish encryption algorithm: a 128-bit block cipher", J. Wiley, April 1999.

[14] Chuck Easttom, " Modern Cryptography: Applied Mathematics for Encryption and Information Security ", McGraw Hill Professional, October 2015.

[15] Bill Buchanan, "Cryptography", River Publishers, 2017.

[16] 24th International Conference, Ottawa, ON, Canada, August 16-18, 2017, "Selected Areas in Cryptography – SAC 2017", Revised Selected Papers (Lecture Notes in Computer Science), Springer, January 10, 2018.

[17] Kazuo Sakiyama, Yu Sasaki, Yang Li, "Security of Block Ciphers: From Algorithm Design to Hardware Implementation" 1st Edition, Wiley, April 25, 2016.

SMART EMOTION DETECTION BY USING NEURAL NETWORK

Louloua M. Al-saedi*[1]& Golodov Valantin A.[2]& Abdulniser K. Al-maini[3]

(1,3)**Al-Nisour University College - Department of Computer Techniques Engineering**
(2) South Ural State University, School of Electric Engineering and Computer Science
loulouamustafa@yahoo.com[1], Golodovva@susu.ru[2], Nasernaser310@gmail.com[3]

Abstract

Human emotion detection is important to improve the interaction humans and computers, by allowing computers to shape their behavior according to the mood of the human. This project, present a study of various feature extraction methods called Gabor features, Histogram of Gradients, Haar-like features, moments coupled with different machine learning algorithms were Support Vector Machines (SVM), Random Trees (aka Random Forests TM). The Artificial Neural Networks (ANN) to recognize and identify human emotion was using facial expressions. The classifier on a database of around 5 images of 10 subjects emoting Anger, Disgust, Fear, Happiness, Neutral, Sadness, and Surprise. Correspondingly, conduct experiments to test how accurately humans identify emotion. The results show that the best performance obtained using Gabor features coupled with a linear SVM with an accuracy. However, the outcome find that humans show an accuracy and outperform of every extractor-learner combination that implemented**.**

Keywords: Face detection, Facial expressions**,** Artificial Neural Networks**.**

الخلاصة

يعد اكتشاف العواطف البشرية أمرًا مهمًا لتحسين التفاعل بين البشر وأجهزة الكمبيوتر، من خلال السماح لأجهزة الكمبيوتر بتشكيل سلوكها وفقًا لمزاج الإنسان. يقدم هذا البحث، دراسة حول طرق استخلاص ميزة مختلفة تسمى ميزات Gabor، رسم بياني للدروس، ميزات تشبه الحوافر، لحظات مقترنة بخوارزميات مختلفة لتعلم الآلة، آلات دعم المتجهات (SVM)، والأشجار العشوائية) المعروفة أيضًا باسم .(Random Forests TM كانت الشبكات العصبية الاصطناعية (ANN)للاعتراف والتعرف على المشاعر البشرية تستخدم تعبيرات الوجه. المصنف على قاعدة بيانات من حوالي 5 صور من 10 مواضيع تحاكي الغضب والاشمئزاز والخوف والسعادة والحياد والحزن والمفاجأة. بالمقابل، قم بإجراء تجارب لاختبار مدى دقة تعريف البشر للعاطفة. تظهر النتائج أن أفضل أداء تم الحصول عليه باستخدام ميزات Gabor إلى جانب SVMخطي بدقة. ومع ذلك، فإن النتيجة توصلت إلى أن البشر يبدون دقة وتفوقًا في الأداء لكل تركيبة من مستخلصات المتعلم.

1. Introduction

In the earlier years have seen computers come into every aspect of our lives. An active area of research is in improving the interactions between humans and computers. Our project aims to improve the human computer interaction by providing techniques for a computer to identify human emotion, and to tailor its behavior accordingly. Detection of human emotion can improve interactions with machines in everyday life. For instance, a personal robot can detect the emotions of its user, and respond accordingly. Smart houses can detect the mood of the residents, and adjust parameters like lighting, air conditioning, power usage of personal equipment etc. accordingly. Similarly, a smart car can detect when the driver is incensed, automatically pull over and stop; thereby preventing accidents caused due to road rage.

Some existing applications of emotion detection are nViso [1] which captures and analyzes the emotional response and visual attention of consumers for applications such as market research and brand management. Samsung researchers have developed a smart phone that can infer the user's emotional state based on how the user operates the phone [1]. In 2013, Cheng, J., Deng, Y., Meng, H. and Wang, Z propose the GP-GPU acceleration service for continuous face and emotion detection system [1]. For real-world scenario of continuously monitoring of movie scene promising results were achieved. The system was initially tested in MATLAB. It was proven that GPU acceleration can speed up the processing by 80 times comparing to CPU. This system can provide the detected emotional state every 1.5 second [1]. In 2015, the Microsoft Oxford API cloud service provides the recognition of emotions based on facial expressions [1]. This API provides the confidence across a set of emotions for each face in the image, as well as bounding box for the face. The emotions detected are anger, contempt, disgust, fear, happiness, neutral, sadness, and surprise. These emotions are understood to be cross-culturally and universally communicated with facial expressions. Recognition is experimental and not always accurate [1].

2. Basic of Image processing

Image processing involves changing the nature of an image in order to either:

1. Improve its pictorial information for human interpretation,
2. Render it more suitable for autonomous machine perception.

We shall be concerned with digital image processing, which involves using a computer to change the nature of a digital image (see below). It is necessary to realize that these two

aspects represent two separate but equally important aspects of image processing. A procedure which satisfies condition (1)—a procedure which makes an image “look better” may be the very worst procedure for satisfying condition (2). Humans like their images to be sharp, clear and detailed; machines prefer their images to be simple and uncluttered.

Examples (1) may include Enhancing the edges of an image to make it appear sharper; an example is shown in (Figure 1) Note how the second image appears “cleaner”; it is a more pleasant image. Sharpening edges is a vital component of printing: in order for an image to appear “at its best” on the printed page; some sharpening is usually performed [2]

(a) The original image (b) result after "sharpening"

Fig.1: Example 1 Image sharpening

Removing “noise” from an image; noise being random errors in the image. An example is given in (Figure 2) Noise is a very common problem in data transmission: all sorts of electronic components may affect data passing through them, and the results may be undesirable. As we shall see in chapter 5 noise may take many different forms; each type of noise requiring a different method of removal. Removing motion blur from an image. An example is given in (Figure 3). Note that in the deblurred image (b) it is easier to read the number plate, and to see the spikes on the fence behind the car, as well as other details not at all clear in the original image (a). Motion blur may occur when the shutter speed of the camera is too long for the speed of the object. In photographs of fast moving objects: athletes, vehicles for example, the problem of blur may be considerable [2].


(a) The original image (b) after removing noise

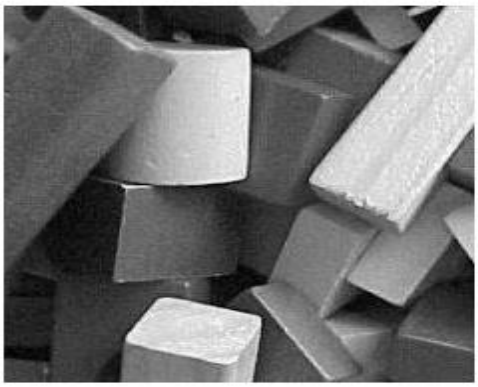
Fig.2: removing noise from an image


(a) The original image (b)after removing the blur

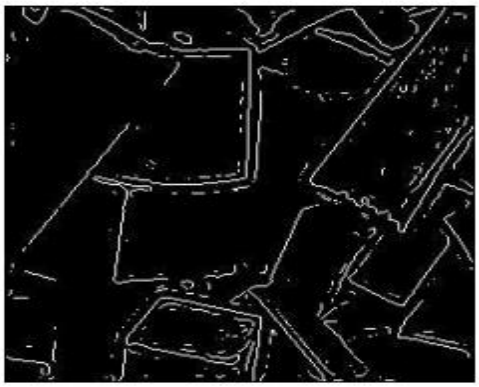
Fig.3: image deburring

Examples (2) may include: Obtaining the edges of an image. This may be necessary for the measurement of objects in an image; an example is shown in (Figures 4). Once we have the edges we can measure their spread, and the area contained within them. We can also use edge detection algorithms as a first step in edge enhancement, as we saw above

From the edge result, we see that it may be necessary to enhance the original image slightly, to make the edges clearer.



(a) The original image



(b) its edge image

Fig. 4: Example 2 finding edge in an image.

3. An Artificial Neural Network

It is a computational system inspired by the Structure Processing Method Learning Ability of a biological brain Characteristics of Artificial Neural Networks:

1-A large number of very simple processing neuron-like processing elements

2-A large number of weighted connections between the elements

3-Distributed representation of knowledge over the connections Knowledge is acquired by network through a learning process.as shown in (Figure 5) [3].

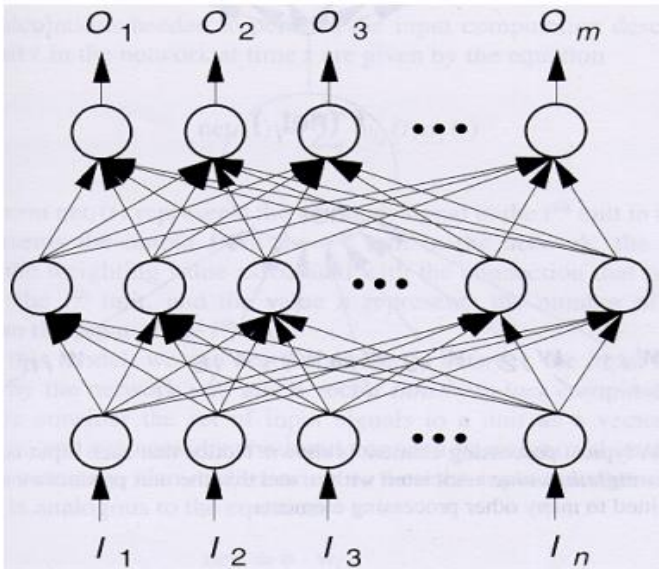


Fig. 5: neural network represents

- **Types of artificial Neural Networks:**

- Massive Parallelism.
- Distributed representation.
- Learning ability.
- Generalization ability.
- Fault tolerance [3].

4. Literature Review

Human-like robots and machines that are expected to enjoy truly intelligent and transparent communications with human can be created using automatic facial expression recognition with a set of specific desired accuracy and performance requirements. Facial expression recognition deals with the problem of classifying facial images into expression classes. It has been of interest to a growing number of researchers and much progress has been made during the last two decades.

Expression recognition involves a variety of subjects such as perceptual recognition, machine learning, and affective computing etc. One case study uses skin color range of human face to localize face area. Then high frequency noise is removed by masking with a low pass filter from the preprocessed image and skin color blocks are detected. After face detection, various facial features are identified by calculating the ratio of width of multiple regions in human face [5].

Finally, the test image is partitioned into a set of sub-images and each of these sub-images is matched against a set of sub-pattern training set [5,6]:

- Partitioning is done using Aw-SpPCA algorithm. Given as input any emotion of face, this pattern training set will classify the particular emotion (Md. Zahangir Alom et al., 2012). Face component extraction by dividing the face region into eye pair and mouth region and measurement of Euclidean distance among various facial features is also adopted by a case study.
- Similar study is done by Neha Gupta to detect emotions. This research includes four steps: pre-processing, edge detection, feature extraction and distance measurement among the features to classify different emotions This type of approach is classified as Geometric Approach (DewiAgushinta, 2011).

- Another research includes Face detection method using segmentation technique. First, the face area of the test image is detected using skin color detection. RGB color space is transformed into YCbCr color space in the image and then skin blocks quantization is done to detect skin color blocks. As next step, a face cropping algorithm is used to localize the face region. Then, different facial features are extracted using segmentation of each component region eyes, nose, and mouth). Finally, vertical & angular distances between various facial features are measured and based on this any unique facial expression is identified. This approach can be used in any biometric recognition system (Neha Guptaand, 2013).
- A template matching based facial feature detection technique is used in a different case study (MRS). Different methods of face detection and their comparative study are done in another review work.
- Face detection methods are divided into two primary techniques: Feature based & View based methods (Mrs. Sunita Roy et al., 2012).
- Gabor filters are used to extract facial features in another study. This approach is called Appearance based approach. This classification based facial expression recognition method uses a bank of multilayer perceptron neural networks.
- Feature size reduction is done by Principal Component Analysis PCA) (Lajevardi and Lech, 2008). Another study represents a robust face detection and gender classification strategy in color images under non-uniform background. This is done by localizing human face region in the given image and detecting facial features after converting the given RGB image to YCbCr color space for skin detection. Here, first mouth region is detected from the localized face region. From this, eye regions and nose regions are located and they are used as feature points.
- Gender classification is done using these features on images with different sizes. Linear support vector machine is used as the classifier which gives best classification rate (Liao and Medioni, 2005; Kanade et al., 2000

5. Mathematical Model of Ideal System for Facial Expression Recognition and Face Detection

There are three main factors to construct a Facial Expression Recognition system, namely face detection, facial feature extraction, and emotion classification. An ideal emotion

analyzer should recognize the subjects regardless of gender, age, and any ethnicity. The system should be invariant to different lightening conditions and distraction as glasses, changes in hairstyle, facial hair, moustache, beard, etc. and also should be able to “fill in” missing parts of the face and construct a whole face. It should also perform robust facial expression analysis despite large changes in viewing condition, rigid movement, etc. A good reference system is the human visual system, the current systems are far from ideal and they have a long way to achieve these goals [6].

Most systems detect face under controlled conditions, such as without facial hair/glasses, any rigid head movement, the first frame should be a neutral emotion etc., and thus nowadays, arbitrary face detection has drawn great intention. Normally the face detection is done in 2 ways. In the holistic approach, the face is determined as a whole unit, while in an analytic approach only some important facial features are detected.

After the face is detected, there are 2 ways to extract the features. In the holistic face model, a template-based method is used. In the analytic face model, featured-based methods will be use to track the facial features while people are showing the facial expression. In our system, we mainly focus on the emotion classification part, not on face detection or on facial feature extraction. For the extraction of the facial features we

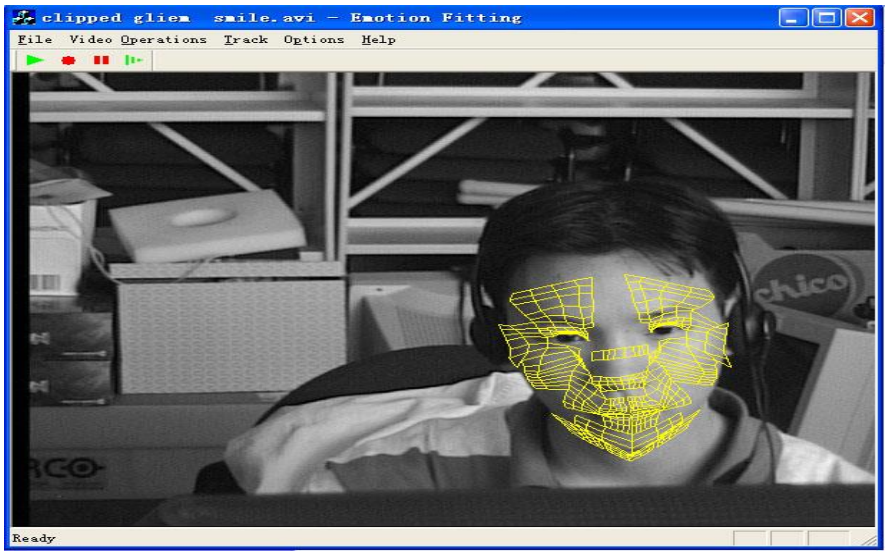


Fig.7: A snap shot of the real time face tracker.

On the right side is a wireframe model overlaid on a face being tracked, the example is from the authentic database we created use the real time facial expression recognition system developed by Sebe and Cohen (Figure 7). This system is composed of a face tracking part, which outputs a vector of motion features of certain regions of the face. The features are used as inputs to a classifier.

This face tracker uses a model-based approach where an explicit 3D-wireframe model of the face is constructed. In the first frame of the image sequence, landmark facial features such as the eye corners and mouth corners are selected interactively. The generic face model is then warped to fit the selected facial features. The face model consists of 16 surface patches embedded in Bezier volumes. The surface patches defined this way are guaranteed to be continuous and smooth. Changing the locations of the control points in the Bezier volume can change the shape of the mesh. Before describing the Bezier volume, we begin with the Bezier curve. Given a set of control points$b_0, b_1, \ldots\ldots, b_n$, the corresponding Bezier (or Bernstein-Bezier curve) is given by

$$x(u) = \sum_{i=0}^{n} b_i \; B_i^n(u) \quad = \sum_{i=0}^{n} b_i \; \binom{n}{i} \; u^i \, (1-u)^{n-1} \qquad (1)$$

Where the shape of the curve is controlled by the control points *bi* and *u* ranging between [0,1]. As the control points are moved, a new shape is obtained according to the Bernstein polynomials $B_n(u)$in the above equation. The displacement of a point on the curve can be described in terms of linear combinations of displacements of the control points.

The Bezier volume is a straightforward extension of the Bezier curve and is defined by *V=BD* written in matrix form. In this equation, *V* is the displacement of the mesh nodes, *D* is a matrix whose columns are the control point displacement vectors of the Bezier volume, and *B* is the mapping in terms of Bernstein polynomials. In other words, the change in the shape of the face model can be described in terms of the deformations in *D*.

Once the model is constructed and fitted, head motion and local deformations of the facial features such as the eyebrows, eyelids, and mouth can be tracked. First the 2D image motions are measured using template matching between frames at different resolutions. Image templates from the previous frame and from the very first frame is both used for more robust

tracking. The measured 2D image motions are modeled as projections of the true 3D motions onto the image plane.

From the 2D motions of many points on the mesh, the 3D motion can be estimated by solving an overdetermined system of equations of the projective motions in the least squared sense. The recovered motions are represented in terms of magnitudes of some predefined motion of various facial features. Each feature motion corresponds to a simple deformation on the face, defined in terms of the Bezier volume control parameters. We refer to these motion vectors as Motion-Units (MU's). Note that they are similar but not equivalent to Ekman's AU's and are numeric in nature, Representing not only the activation of a facial region, but also the direction and intensity of the motion [7].

The MU's used in the face tracker are shown in (Figure 8) and are described in (Table 1).

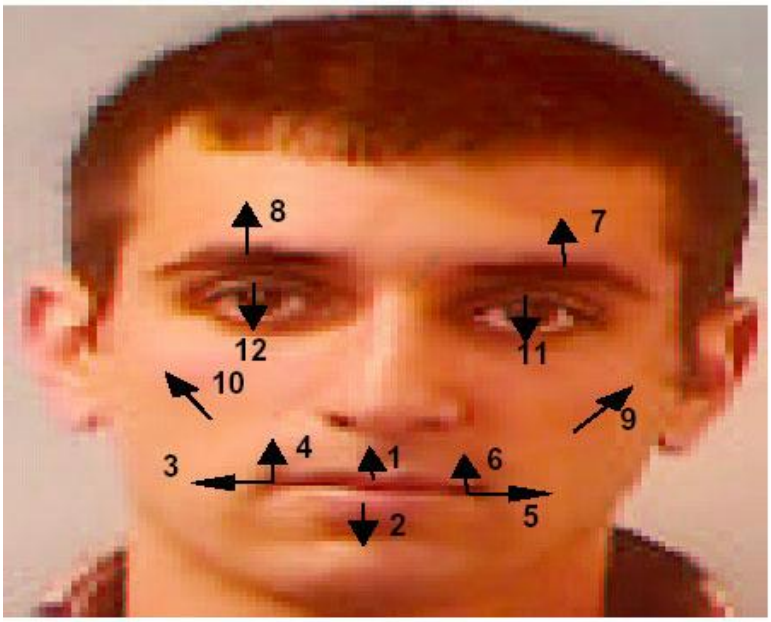


Fig.8: The 12-facial motion measurement

Table 1: Emotion units used in the face tracker

1	Vertical movement of the center of upper lip
2	Vertical movement of the center of lower lip
3	Horizontal movement of left mouth corner

4	Vertical movement of left mouth corner
5	Horizontal movement of right mouth corner
6	Vertical movement of right mouth corner
7	Vertical movement of right brow
8	Vertical movement of left brow
9	Lifting of right cheek
10	Lifting of left cheek
11	Blinking of right eye
12	Blinking of left eye

Each facial expression is modeled as a linear combination of the MU's

$$V = B[D_0\ D_1\ \ldots\ldots D_M]\begin{bmatrix} P_0 \\ P_1 \\ P_M \end{bmatrix} = BDP \qquad (2)$$

Where each of the *D* corresponds to a MU, and the *Pi* are the corresponding magnitudes (or coefficients) of each deformation. The overall motion of the head and face is:

$$R(V_0 + BDP) + T \qquad (3)$$

Where R is the 3D-rotation matrix, T is the 3D-translation matrix, and V0 is the initial face model. The MU's are used as the basic features for the classification scheme [7].

6. Testing phase recognizing for different cases of emotions:

This phase can be performed to measure the classification rate. The inputs to this phase are the models that were built during training phase and the test images for which the emotions are to be recognized. Here again only the face region is used as rest of the image do not contribute information about the emotion. In a typical real time scenario, the input image would be detected face image from an earlier face detection phase, and its shown in (Figures 9,10,11).

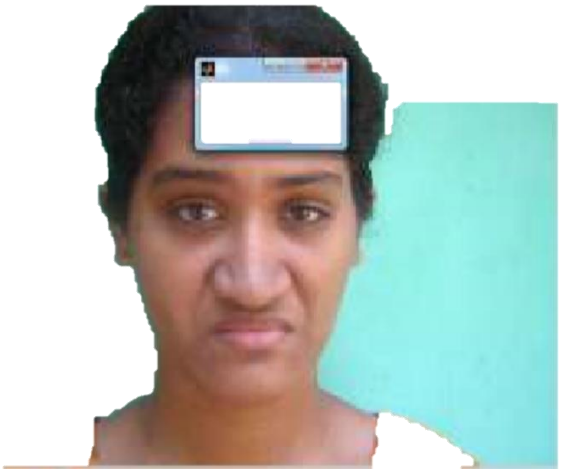


Fig.9: Shows the result for disgust image.

Fig.10: Shows the result for happy image

Fig. 11: Shows result for angry image.

Conclusion

In this research discussed tutorial about image processing and neural Network and how can use this combination to detect the emotion of human, when talking about emotion research, its mean the area of psychology, which is researching and provides solutions in the psychiatric status of the individual. Until recently, this science was based on philosophy, but in recent years with the increase of technology, a new parameter induced. By the new digital evolution,

neuroscience enhanced studies related with emotions, by introducing new methods in investigating brain functions. From another point of view, music is commonly known as an emotional stimulant that prompts memories, enhances brain activity and stimulates the mind Future studies can be focused on extending the Data set used in the study. The sample images of different group of subjects with different ages and races may also be considered. The larger sample of images would in effect yields a more representative classification emotion model. Nevertheless, the emotion classification model developed in this study can support the development of Intelligent Tutoring System in particular, and E-learning System in general.

Reference

[1] Shamla Mantri, Kalpana Bapat, "Neural Network Based Faced Recognition Using Matlab", *International Journal of Computer Science Technology*, Vol. 1, No. 1, pp. 6-9 February 2011.

[2] An Introduction to Digital Image Processing with MATLAB Notes Alasdair McAndrew.

[3] https://www.cse.unr.edu/~bebis/MathMethods/.../lecture.pdf.

[4] REVIEW OF FACE DETECTION SYSTEMS BASED ARTIFICIAL NEURAL NETWORKS ALGORITHMS Omaima N. A. AL-Allaf Assistant Professor, Faculty of Sciences & IT, Al-Zaytoonah University of Jordan, P.O. Box130, Amman, (11733), Jordan, Vol.6, No.1, February 2014.

[5] Global Journal of Computer Science and Technology: F Graphics & Vision Volume 15 Issue 3 Version 1.0 Year 2015 Type: Double Blind Peer Reviewed International Research Journal Publisher: Global Journals Inc. (USA) Online ISSN: 0975-4172 & Print ISSN: 0975-4350 ((Face Detection and Expression Recognition using Neural Network Approaches)) By Nisha & Sandeep Dahiya.

[6] S. Zafeiriou, C. Zhang, Z. Zhang, A Survey on Face Detection in the wild: past, present and future, Computer Vision and Image Understanding (2015).

[7] N. Sebe and M.S. Lew, Toward Improved Ranking Metrics, IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1132-1143, 2016.

Effect of Light Pulses Generator in Free Space Optics Application

Muthana. Y. Al-Dour

Al-Nisour University College/Department of Computer Technical Engineering/Baghdad-Iraq.

E-mail: **muthana.eng@nuc.edu.iq** Mob: +9647715227636

Abstract

There is a wide range of techniques for generating pulses with durations of nanoseconds, picoseconds, or even femto seconds with lasers. Such short durations make light pulses very interesting for many applications, such as telecommunications or ultra precise measurements of various kinds soliton waves system basically depend on the generation of the Gaussian optical pulse generator, which can transfer data to more than 350 Km as in this paper, with very stable system which is designed by optisys software simulation. The system intended in this paper by means of free space optics (FSO) as the media channel. Two users in the input channel at frequency 1556nm, and CW laser source transmitted at frequency 1550 nm. Maximum distance transmit data in this project is 350Km, at 155MHz data rate, also 200Km maximum distance at 1GHz data rate.

Keywords: *Gaussian pulse generator, FSO, Mux, Demux, coupler, Ground, optical null, CW Laser*

1-Introduction

Soliton waves: When both the dispersive and the nonlinear term are present in the equation the two effects can neutralize each other. If the water wave has a special shape the effects are exactly counterbalanced and the wave rolls along undistorted. The soliton shape can be found by direct integration of the KdV equation

$$U(x,t) = a\,\mathrm{sech}^2[b(x-vt)] \qquad (1)$$

With $b= (a/12)^{1/2}$ and $v=3a$. The constant a is the only free parameter in the solution. It defines the amplitude and the width in such a way that a large (tall) soliton will be narrow, while a low

soliton will be broad. The constant *v* defines the velocity of the soliton. Since *v=3a* a tall soliton will move faster than a low one. [1]

A soliton is a self-reinforcing solitary wave packet that maintains its shape while it propagates at a constant velocity. Solitons are caused by a cancellation of nonlinear and dispersive effects in the medium. (The term "dispersive effects" refers to a property of certain systems where the speed of the waves varies according to frequency.) Solitons are the solutions of a widespread class of weakly nonlinear dispersive partial differential equations describing physical systems.

A single, consensus definition of a soliton is difficult to find. Ascribe three properties to solitons:

1. They are of permanent form;
2. They are localized within a region;
3. They can interact with other solitons, and emerge from the collision unchanged, except for a phase shift.

More formal definitions exist, but they require substantial mathematics. Moreover, some scientists use the term *soliton* for phenomena that do not quite have these three properties (for instance, the 'light bullets' of nonlinear optics are often called solitons despite losing energy during interaction).

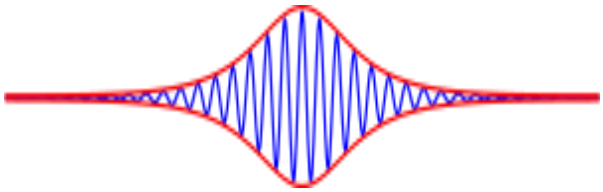


Fig. 1: *Hyperbolic secant (sech) envelope soliton for water waves*

A hyperbolic secant (sech) envelope soliton for water waves. The blue line is the carrier signal, while the red line is the envelope soliton

1-1 Solitons in fiber optics

Much experimentation has been done using solitons in fiber optics applications. Solitons in a fiber optic system are described by the Manakov equations. Solitons' inherent stability makes long-distance transmission possible without the use of repeaters, and could potentially double transmission capacity as well.

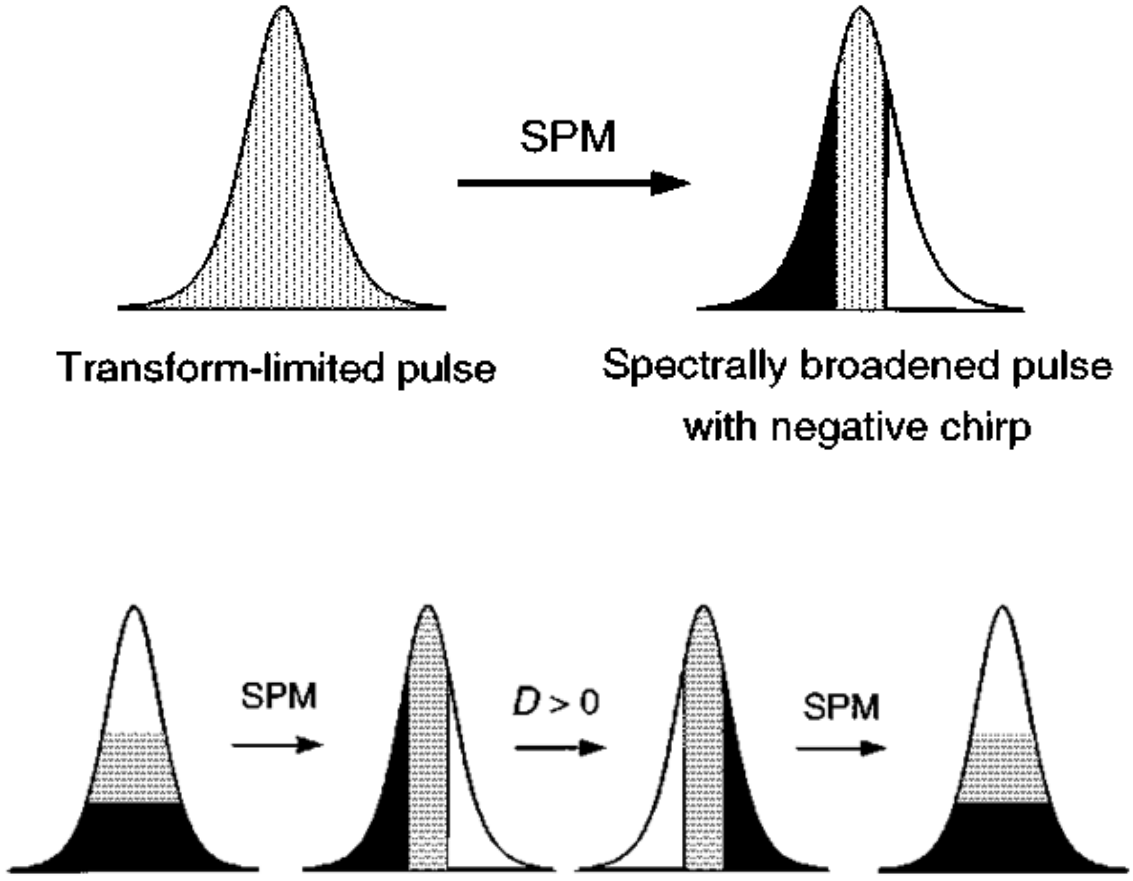


Fig. 2: *Soliton shape*

1-2 Wavelength Division Multiplexing (WDM) Technique

Wavelength Division Multiplexing (WDM) is a technique in which simultaneous transmission of signals occurs at different optical wavelength. In some applications of this technique, several optical signals combine, transmit together, and will be separated again based

on different arrival times. To handle this technique, optical signals from separating lasers can be combined together or a broadband optical signal from for example a light emitting diode spectrally be sliced into smaller pulses. The use of multiple channels allows increased overall data transmission capacities without increasing the data rates of the single channels, where the time slot per bit must be reduced. Even if the bandwidth of the data modulator is limited, this can be done by using a train of ultra-short pulses (rather than a continuous optical wave) as the input of the modulator [2].

1-3 Nonlinear Effect as Self-Phase Modulation (SPM)

In a Kerr effect medium such as fiber optics, high intensity of light causes a phase delay having the similar temporal shape with the intensity. This nonlinear phenomenon occurs for a beam is called self phase modulation (SPM) which is generated by its intensity. This effect refers to nonlinear changes of the refractive index given by

$$\Delta n = n^2 I \qquad (2)$$

Where, $n2$ is the nonlinear index and the optical intensity is shown by I. Therefore, This phase shift is a temporal dependence effect, whereas the transverse dependence leads to the effect of self-focusing.

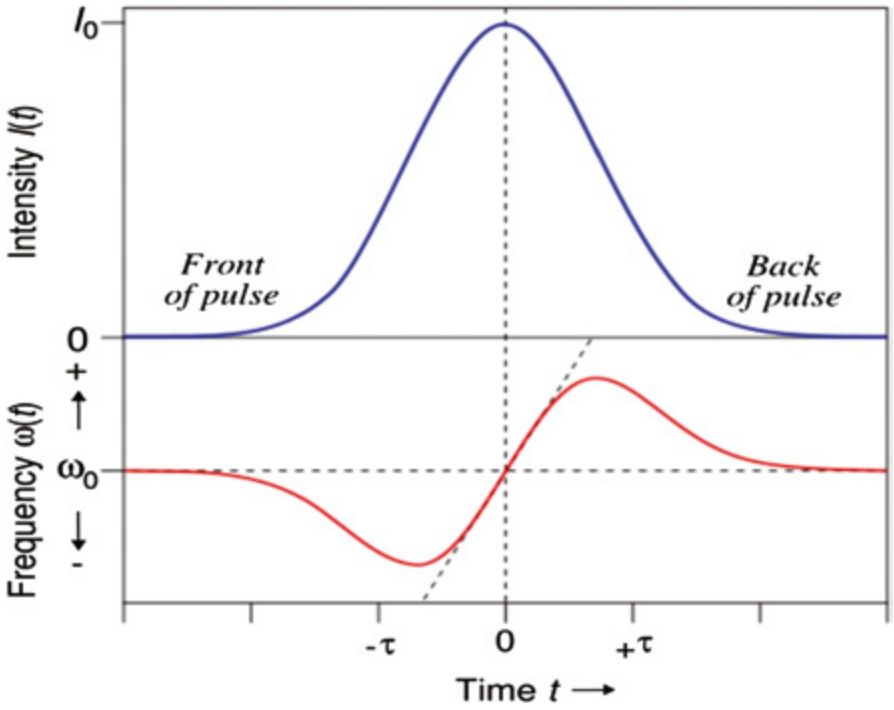


Fig.3: *Spectral broadening of a pulse due to SPM [3]*

Unlike the way that SPM affects the phase of the propagating pulse such phase changes in semiconductor lasers do not follow the temporal intensity profile. Therefore, this effect is declared for the pulse of picoseconds to a few nanoseconds. SPM is very efficient in mode-locked femto second lasers with the Kerr nonlinearity effect medium. In materials with negligible or zero dispersion effects, the nonlinear phase shift is unstable, thus soliton pulse mode is employed which is a result of balancing SPM and dispersion [4, 5]. The intensity of a Gaussian ultra short pulse at a time (t) can be expressed by

$$I(t) = I_O \, EXP\,(-t^2/\tau^2) \qquad (3)$$

Where *I*0 and τ are the peak intensity and pulse duration. In a Kerr type medium,
The refractive index is given by:

$$n(I) = n_o + n_2 I \qquad (4)$$

Where, *n*0 and *n*2 are the linear and nonlinear refractive indices.

An optical pulse is a flash of light. Lasers and related devices have been found to have an amazing potential for generating light pulses with very special properties:

There is a wide range of techniques for generating pulses with durations of nanoseconds, picoseconds, or even femto seconds with lasers. Such short durations make light pulses very interesting for many applications, such as telecommunications or ultra precise measurements of various kinds [7].

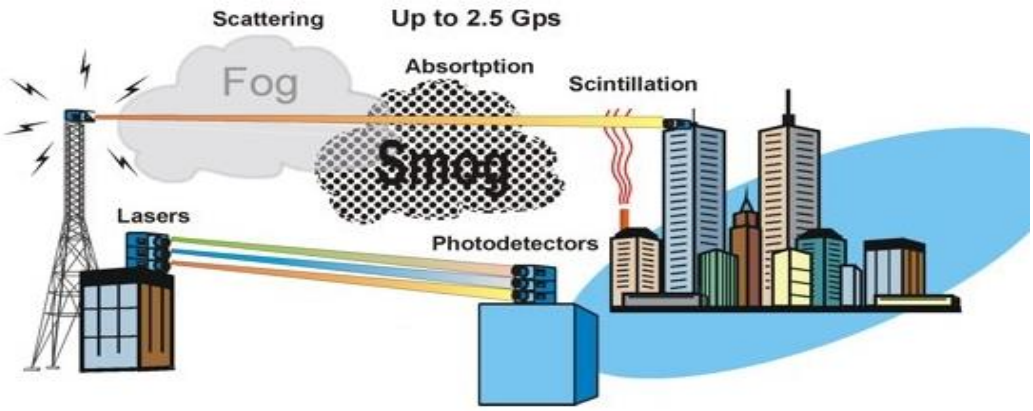


Fig .4*:* Free Space Optics - FSO Diagram

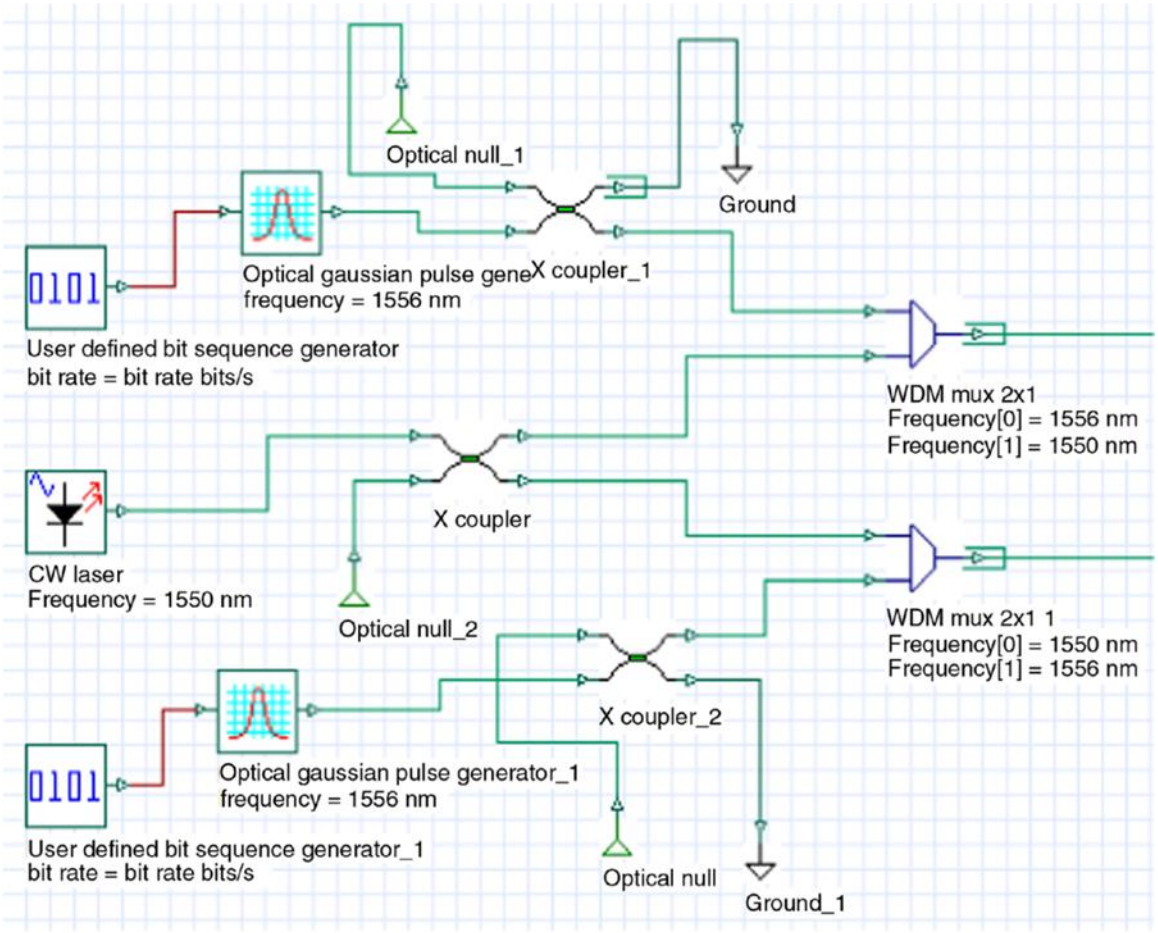
Free space optical communication systems use an optical carrier signal to transfer information through the air (free space) between two or more optical receivers or transceivers.

Figure .4. Show free space optical (FSO) transmission. This diagram shows that free space optical transmission systems lose some of their energy from signal scattering, absorption and scintillation. Optical signal scattering occurs when light signals are redirected as they pass through water particles. Optical signal absorption occurs as some optical energy is converted to heat as it strikes particles (such as smog). Scintillation occurs when heated (such as from smokestacks) air cause a bending of the optical beam. This example shows that it is possible to transmit multiple light wave signals on different straightforward method was used in optical communication in order to convert from electrical to optical binary information. Electrical bit "1" was associated with a higher optical while bit "O" was associated to a lower optical intensity [6].

2- Simulation Setup

The simulation has been carried out by using opt system software version 7.

In this occupation we used optical sech pulse generator as a source to transmit data over 300 km length through one loop control and a traveling wave SOA (0.15 an injection current). The systems build as figure 7. This figure shows the system design of soliton wave effect in optical fiber.



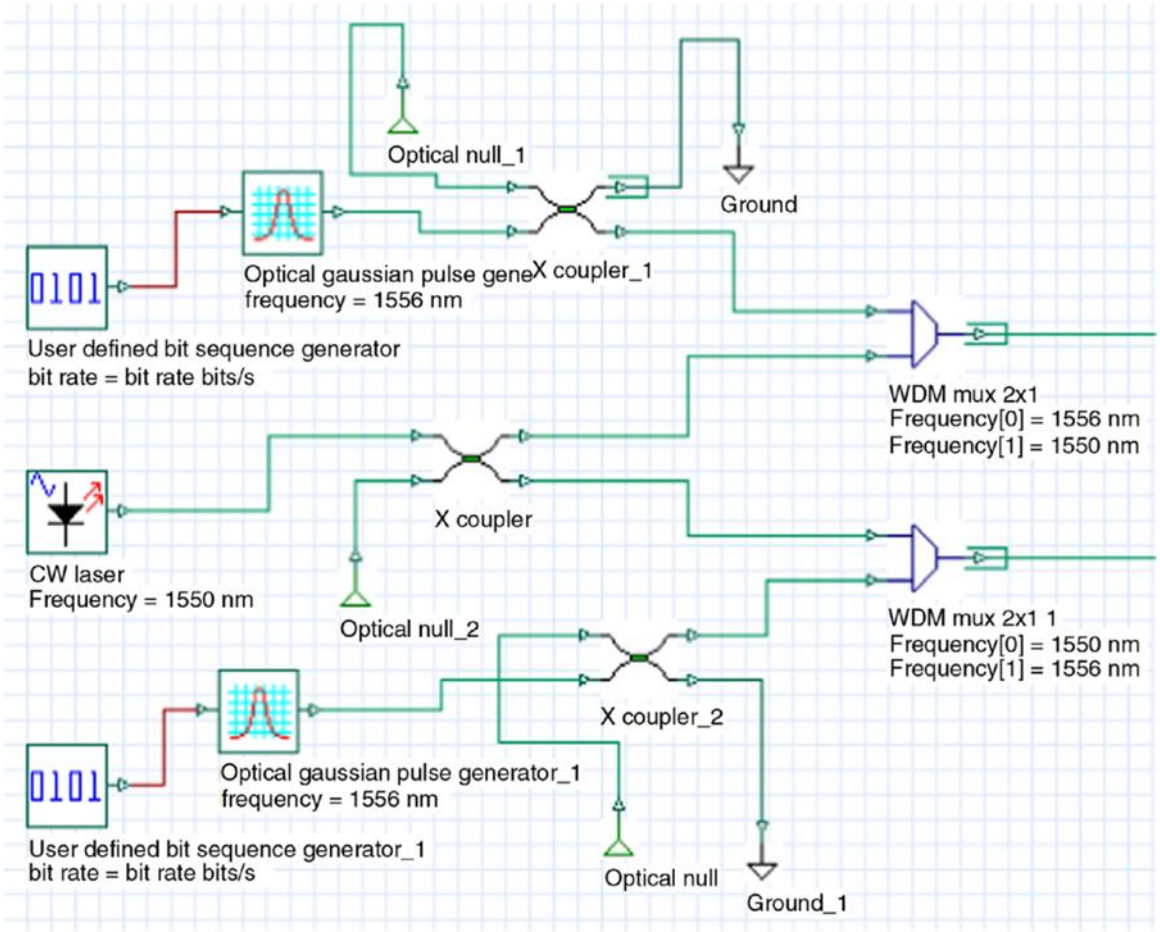


Fig. 5: Gaussian optical pulse generator project layout

For this system the parameters was setting as;

- Multiplexing channel from the transmitter is 1556nm, 1550nm, for the two users.
- FSO channel is setting at 1556nm.
- Demultiplexing channel at the receiver is 1556, 1550, for the two users.
- Coupling coefficient = o.5.[8].

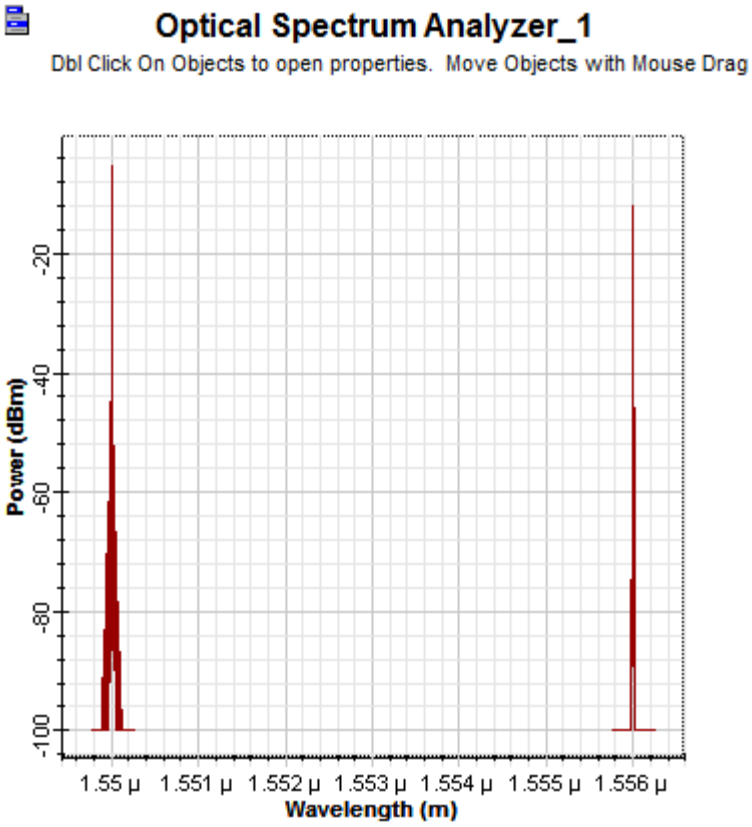


Fig.6: Two signal spectrum sending from the transmitter channel

The two input signals is prepared to be transmitted in the range of 1550 nm to 1556 nm, time interval frequency with the same power equal to -18 dBm. Figure referring to the shape of the signals to be transmitted from the transmitter channel.

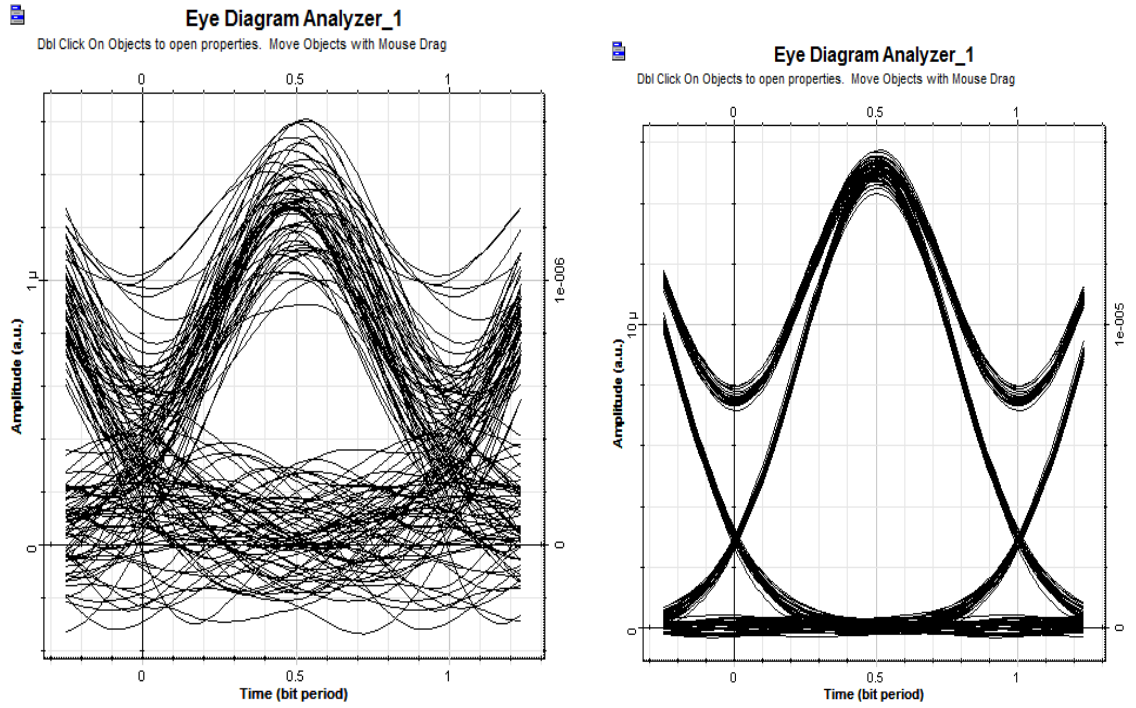


Fig .7: Eye diagram at 350 Km fiber length and 100Km

The eye diagram analyzer is a very good observation device to illustrate the output results of the designed system. From this diagram it seems clearly that at minimum distance 100Km, the BER recording the value of $5.1x10^{-194}$ and at maximum distance 350 Km, the value of BER is $7.1x10^{-5}$.this results is at 1GHz Data Rate. This is a very good results can be achieved by this system. Parameter relative with bit error ratio or quality of the communication system.

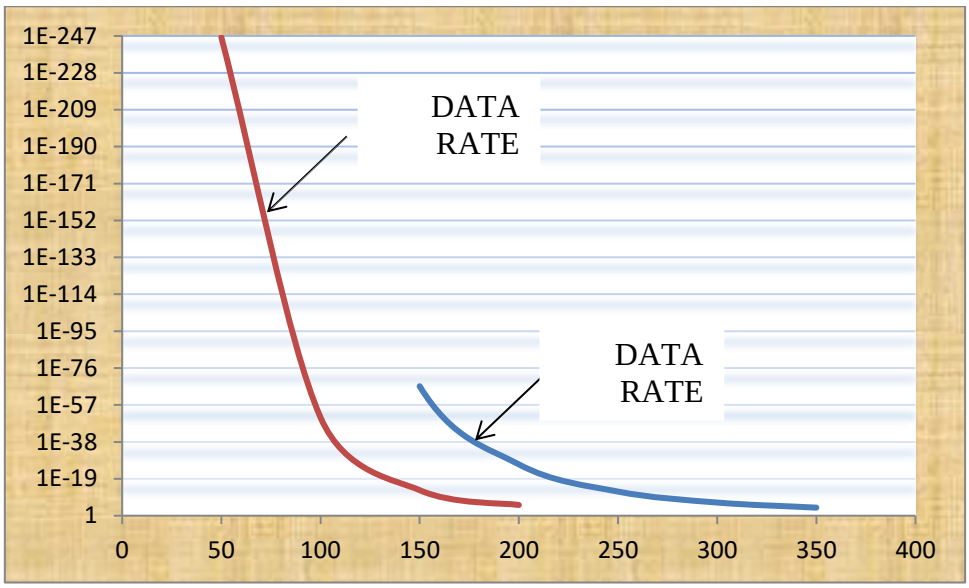


Fig. 8: Relation between fiber length and BER

This figure referring to the fiber length distance which starting from 50Km and reaching to the maximum distance at 350 Km at 1GHz, while the distance reach to the 200Km,at data rate equal to 155MHz.

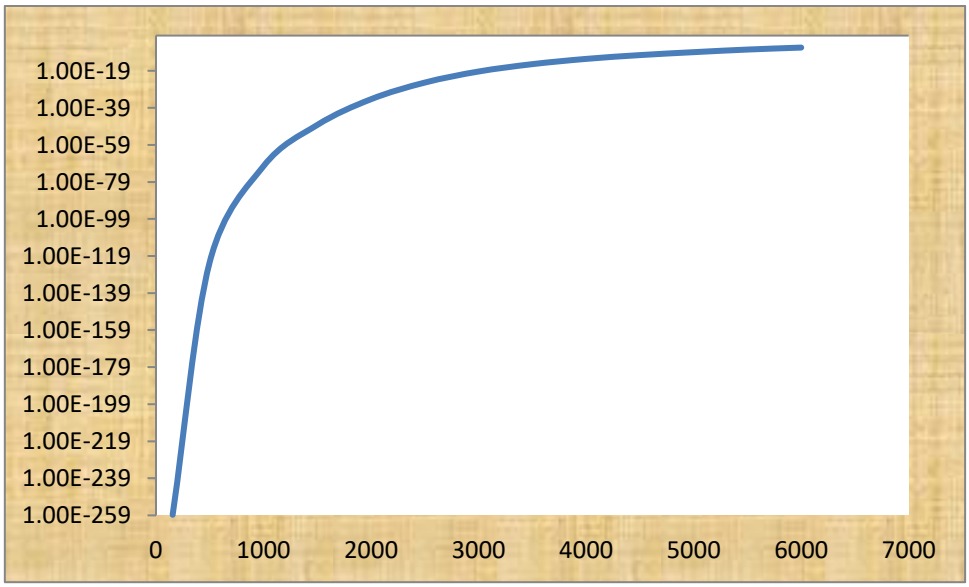


Fig 10: Relation between data rates and BER for 100Km distances

In figure 10, we test multi band of data rate can this system operating with high stability one. The distance of transferring data will be contain constant at value of 100Km,then record the

values of data rate Vs Bit Error Rate ,to achieve the figure of 10. This high data rate is illustrated can use more than 5GHz this is the first time using this value of data rate and giving a stable system, only occurs in fiber optics because of high velocity of laser beam.

Conclusion

As we have discussed that to transfer data (voice, sound and data) through two buildings, we can use a modified leaser with high data rate up to 2 Gbps using soliton system in free space with the use of Gaussian optical pulse generators. This will be reducing the width of the signal to more than 3ps increasing the data rate to be transferred. From the results optioned at 350 Km distance using free space optics (FSO) to transfer data to the receiver's.

Meanwhile, the (sech) optical pulse generator which is equal to the square root of Gaussian. is the perfect solution in Application based on FSO has a great potential to become one of the optical communication suitable system for OCDMA implemented in FTTH application? And by using the FSO the OCDMA network will be expanding directly.

Reference

[1] Marchese SV, Südmeyer T, Golling M, Grange R, Keller U (2006) Pulse energy scaling to 5 μJ from a femto second thin disk laser. Opt Lett 31(18):2728–2730.

[2] Meijerink a (2005) Coherence multiplexing for optical communication systems. University of Twenty, the Netherlands.

[3] Singh SP, Singh N (2007) Nonlinear effect in optical fibers; origin, management and applications. Prog Electromagnet Res PIER 73:249–275.

[4] Krtner F, Keller U (1995) Stabilization of soliton like pulses with a slow saturable absorber. Opt Lett 20(1):16–18.

[5] R. Paschotta, *Field Guide to Laser Pulse Generation*, SPIE Press, Bellingham, WA (2008).

[6] Bloom, S. *The physics of Free Space Optics*.AirFiber Inc.1-22. (2002).

[7] D. Y. Tang *et al,"* higher-order vector soliton"2008.

[8] Muthana.Y. Aldouri *et al*, Int. Journal of Computer Science & Mobile Computing, Vol.5 Issue.1, Jan- 2016, pg. 149-160.

Using Clustering Data Mining for Web Transmission in Telecommunication Industry

Saad K. Majeed
Al – Nisour College University / Baghdad- Iraq
Corresponding author: saad.eng@nuc.edu.iq

Abstract

The data mining applications for any industry depend on two factors: the data that are available and the business problems facing the industry. In this paper we discuss two topics that contact with these factors: the first is the telecommunication services, including computer and Web data transmission, mobile service and other data traffic. The second is using data mining for mobile communication clustering, statistic, optimization, machine learning, artificial intelligence, and database. Data mining is the process of extract information from large database and discover knowledge by using statistical analysis and modeling techniques for discovering patterns and hidden relations in these databases.

The following project are demonstration for which data mining may improve telecommunication service:

. Multidimensional analysis of telecommunication data;
. Fraudulent pattern analysis and the identification of unusual pattern;
. Multidimensional association and sequential pattern analysis;
. Mobile telecommunication services;
. Use of visualization tools in telecommunication data analysis;

Keywords: Data mining, Clustering algorithm, Industries, Web and internet services.

الخلاصه

تطبيقات تقنية التنقيب عن البيانات في المجالات الصناعيه تعتمد على عاملين الاول البيانات المتوفره والثاني المشكله التي تواجه العمليه الصناعيه . في هذه المقاله سوف نناقش نقطتين اساسيتين ذات ارتباط في العاملين اعلاه: الاول الخدمات الاتصاليه اللاسلكيه تتضمن الحوسبه ونقل البيانات عبرشبكة الوب, خدمات الموبايل وحركة سير البيانات. الفقره الثانيه استخدام تقنيات التنقيب عن البيانات في عملية عنقدة اتصالات الموبايل , الاحصائيات, الأمثليه, تعليم الماكنه , الذكاء الأصطناعي و قواعد البيانات .

التنقيب عن البيانات هو استخلاص المعلومات من قواعد بيانات كبيره الحجم مع اكتشاف المعرفه ضمن تلك القواعد بأستخدام التحليل الاحصائي و تقنيات النمذجه لكشف الأنماط و العلاقات المخفيه في تلك القواعد البيانيه. المقاله الأتيه استعراض لكيفية الربط بين تقنيات عنقده التنقيب عن البيانات في مجال النقل عبر الشبكات الاسلكيه وتشمل:

. التحليل متعدد الابعاد لبيانات الاتصالات اللاسلكيه.

. الانماط التحليليه الشاذه مع تحديد الانماط غير المستخدمه.

. الارتباطات متعددة الابعاد وانماط التحليل المتسلسل.

. خدمات الاتصالات بأستخدام الموبايل.

. استخدام الوسائل المرئيه في تحليل البيانات اللاسلكيه.

1. Introduction

The internet has become the largest data repository, facing the problem of information systems. In the same time, more people use the World Wide Web as their main source of information. The existence information, in combination with the dynamic and heterogeneous nature of Web, makes information retrieval a complex process for the average users. Data mining, Search engines, and Web Directories have been developed in order to help the users quickly and easily satisfy their information need. One of the techniques that can play an important role towards the achievement of these objectives is the Clustering Data Mining (CDM). The increasing importance of CDM and the variety of its applications has led to a wide range of algorithms with different quality and complexity in field of telecommunication industry.The contribution of this project is a review and a comparison of the existing CDM approaches in telecommunication industry.

2. Motivation for Clustering Data Mining

Clustering or cluster analysis is one of the main data analysis techniques and deals with the organization of partitioning a set of objects in a multidimensional space into groups, called clusters. Each cluster contains objects that are very similar to each other and very dissimilar to objects in other clusters [1].

Fig. 1 shows the conceptual structure that we can use to identify the performance of various prediction models by using data mining techniques. Fig. 2 shows the proposal structure for creating a predictive model.

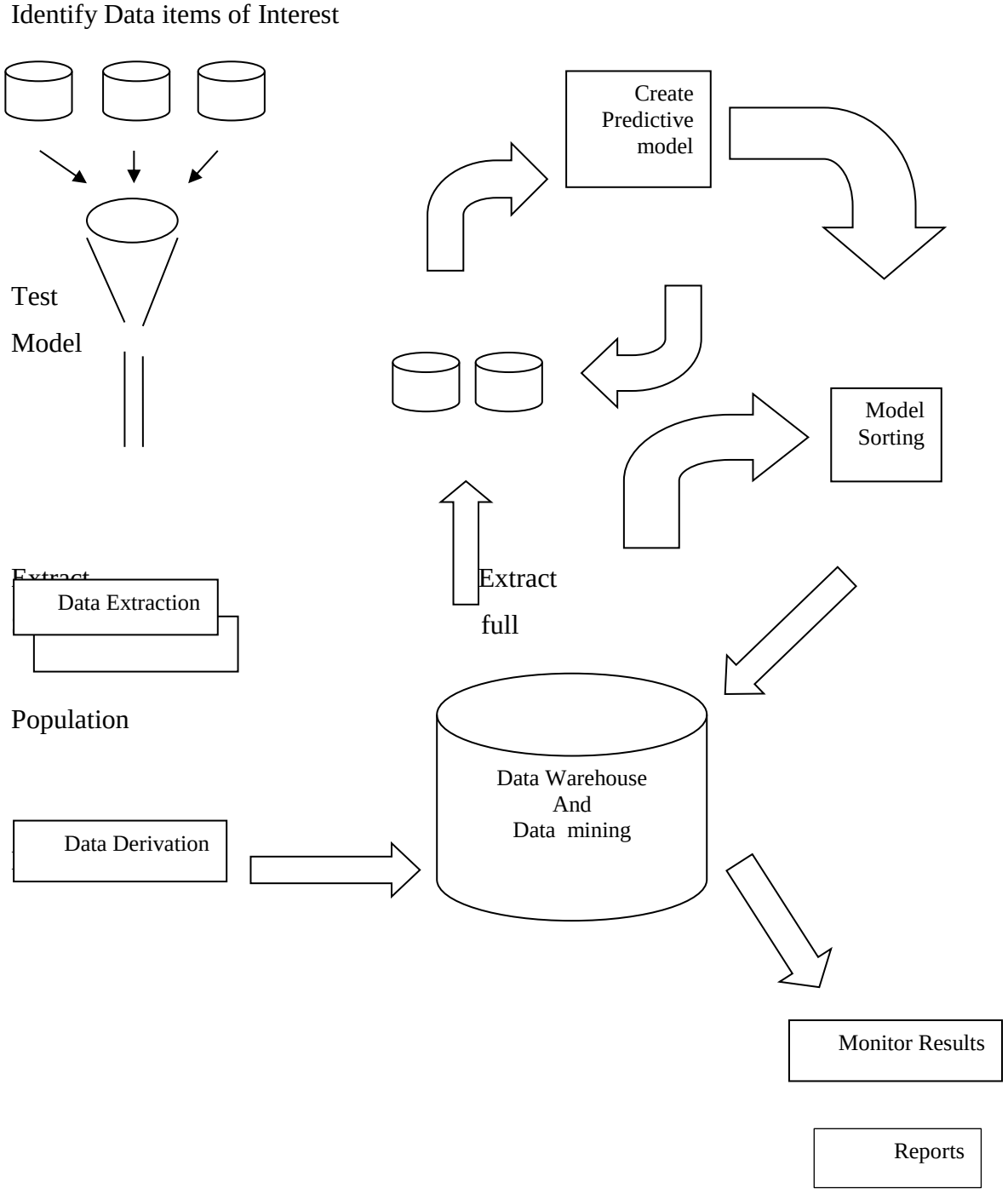


Fig. 1: The conceptual structure of predication models by using data mining technique

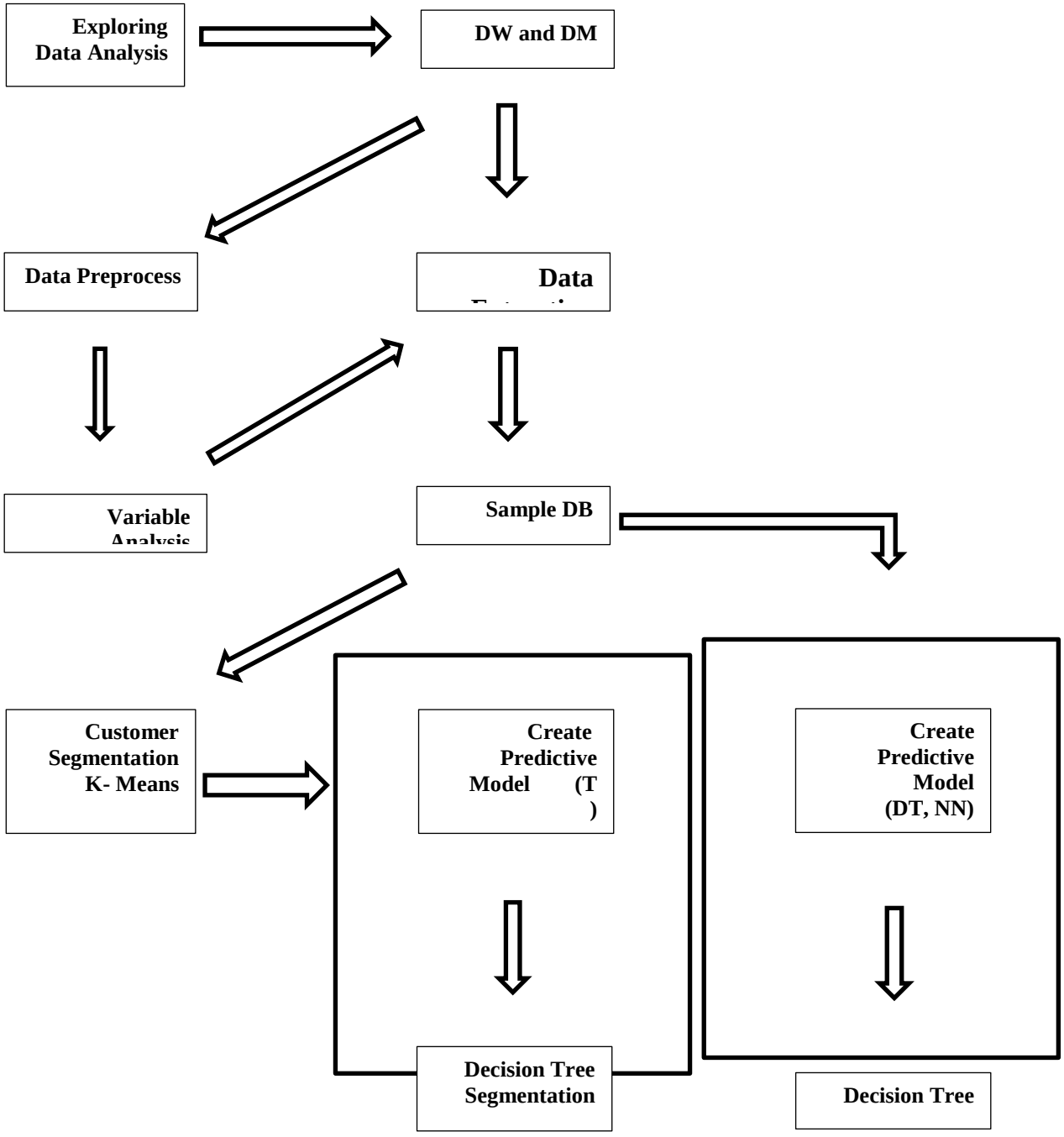


Fig. 2: The proposal structure for creating a predictive model

The most distinct characteristic of data mining is that it deals with very large data set. This requires the algorithms used in data mining to be scalable because they were initially

developed for other applications than data mining become a datamining research. In this paper there is a fast clustering algorithm used to cluster web data transmission in telecom industry. This algorithm is an extension to the K-means algorithm [2]. Compared to other clustering methods the K-means algorithm and its variants [3] are efficient in clustering large data sets, which is very suitable for data mining. In [4] we have an algorithm called K-prototypes,to cluster large data sets with mixed numeric and categorical values.

3. Mining database on Web

There are various ways to carry out Web information, one way is to integrate mining tools with database on the web. This approach work well when the data is in relational databases then SQL- based mining tools could be applied to the virtual relational database. In Web there are unstructured data like imagery data and video data and using relational interface to all such database may be difficult. For mining such data this research develops new tools to mine multimedia data and then focus on developing tools to mine such d ata on the web. Fig. 3 illustrates some of the Web mining concepts [5]

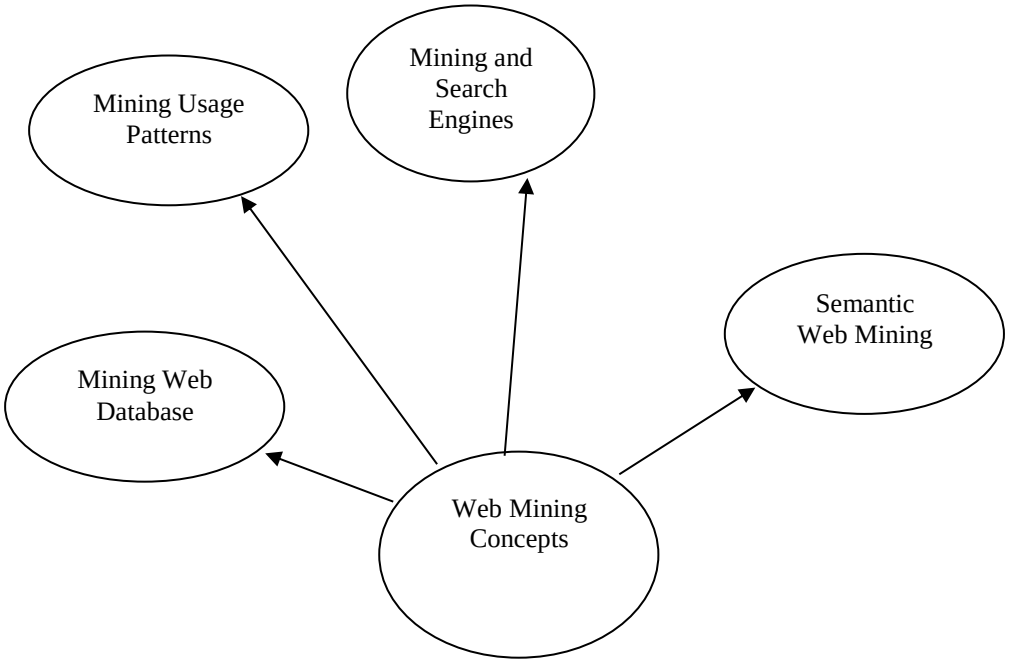


Fig.3: Web mining concepts

4. K-means algorithm

It is a learning algorithm with different data analysis applications widely used for data mining and machine learning in telecommunication industry. The main goal is to classify data into groups of information each group is consists of the separation of information of specific database into k different clusters which combine every data entries. The k-means algorithm [6] is built upon four basic operations: (1) selection of initial k means for clusters, (2) calculation of the dissimilarity between an object and the mean of a cluster, (3) allocation of an object to the cluster whose mean is the nearest to the object, (4) Re-calculation of the mean of a cluster from the objects allocated to it so that the intra cluster dissimilarity is minimized. The aim of algorithm is to minimize the cost function:

$$E = \sum_{I=1}^{k} \sum_{j=1}^{n} Y_{i,j}\, d(Xi, QI) \qquad \ldots\ldots\ldots\ldots\ldots (1)$$

Where n is the number of objects in data set X. xi E X, QI is the mean of cluster I, and $Y_{i,j}$ is an element of a partition matrix Ynxi . d is a dissimilarity measure.

The k-means algorithm has the following important properties:

1. It is efficient in processing large data sets.
2. It often terminates at local optimum techniques such as genetic algorithm.
3. It works only on numeric values.
4. The clusters have convex shapes so that it is difficult to use with non-convex shapes.

One difficulty in using the K-means algorithm is to specify the number of clusters at the cost of performance.

5. K-modes algorithm

This algorithm is a similar of the k-means algorithm. In this algorithm there was three major modifications to the k-means: using different dissimilarity measures, replacing k means with k modes, and using a frequency based method to update modes. These modifications organized with following formula [7]:

$$d(X, Y) = \sum_{j=1}^{m} \partial(Xj, Yj) \qquad \ldots\ldots\ldots\ldots\ldots (2) \quad \text{for Dissimilarity}$$

n

D(Q , X) = ∑ d(Xi , Q) …….. (3) for replacing

i=1

nck,j

Fr (Aj = Ck,j I X) = —— …… (4) for frequency update modesN

6. Proposed Algorithm

In this section as we have seen, in applying the k-means method, two main problems are limitation: the formation of cluster center and the calculation of dissimilarity between objects and cluster centers. These problems completely solved in k-modes algorithm by using the simple matching dissimilarity measure, and replacing the means of clusters by modes. These modifications meet another problem: The multidimensional analysis

and fraudulent pattern analysis for mobile and visualization tools. In the first problem following representation of C is defined by Q= (q1,q2,….,qm) with

qi = {(Cj , Fcj) I Cj E Dj) …… (5)

where Fcj is the relative frequency of class Cj within C . In the second problem, the dissimilarity d(X , Q) is mainly dependent on the relative frequencies of class values within cluster. We can seen that:

m

d(X , Q) =∑ (1 – Fxj) …….. (6)

j=1

The proposed algorithm for clustering large set of data include following steps:

1. Initialize a k-partition of D randomly.
2. Calculate k representation for each cluster.
3. For each Xi , calculate the dissimilarities.

4. Repeat step 3 until no object has changed clusters after a full cycle test of data set.

7. Experimental Results

The purpose of this experiment was to test scalability of k-means and k-modes algorithms for clustering large sets of data in telecommunication industry. We selected a large data set from Telecommunication Company about problems with web transmission. The data set consists of 100000 records, each being described by 34 categorical attributes in which 4 have more than 1000 categories each. We tested two scalabilities of the algorithm, the first one is scalability of the algorithm against the number of clusters for a given numberof objects and the second is the scalability against the number of objects for a given number of clusters. Figure 4 and 5 show the result produced using a single processor of 4000 computers. The results are very advantages because they show clearly a liner increase in time as both the number of clusters and the number of records increase. The algorithm is much faster than its previous cause it needs many less iterations to converge.

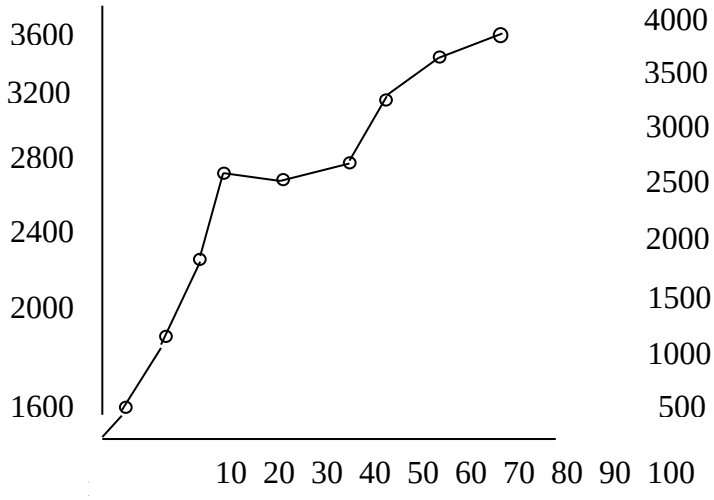


Fig.4: Scalability to the number of Cluster

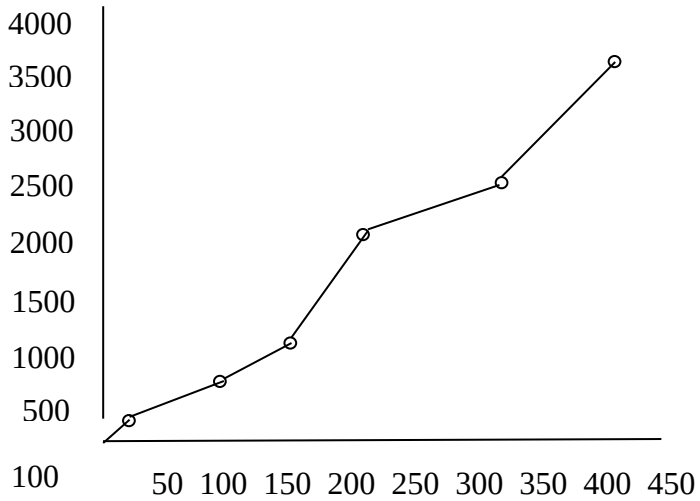


Fig.5: Scalability to the number of records

8. Conclusions and future Work

The first advantage of k-means algorithm in data mining applications is its efficiency in clustering large data sets and Its use is limited to numeric values.The k-modes algorithm has removed this limitation with preserving its efficiency.In this paper suggest a new algorithm is a combine of two previous algorithms for remove the limitation with them. The new algorithm can make the following extensions:

1. Replacing means of clusters with modes.
2. Using new dissimilarity measure to deal with categorical object.
3. Using a frequency based method to update modes of cluster.

The future work plan is to develop and implement a parallel algorithm to clustering data sets with large number of objects for a number of data mining applications such as heterogeneous sets of objects.

References

[1] Jiawei Han, Micheline Kamber, Data mining Concepts and Techniques , 2001 Acadimc Press USA.

[2] Jianping Z. and Xiyu A. " Research and application of k-means algorithm based on clustering analysis" Application Rec. Computer 2007.

[3] John Burkardt " k-means clustering " ICAM Interdisciplinary Center for Applied Mathematics " 2009.

[4] Jinchao Ji, Tian Bai " An improved k-prototype clustering " 2013 Journal homepage: www.e/sevier.com

[5] Zdravko M., Daniel T. L. " Data mining in the Web " Wiley Iner science , Puplication 2007.

[6] Brian T. " k-means Clustering " http//fconyx. Ncifcrf gov/ 2006
Neha Sharma , Nirmal Gaud " k-modes Clustering Algorithm for Categorical Data " International Journal of Computer Application 2015.

تحديد اعراض مرض التوحد بأستخدام الشبكة العصبية المطورة (BPV)

م. سالي علي عبد اللطيف[1] و عمران طه علي[2]

الجامعة المستنصرية / كلية التربية - قسم علوم الحاسبات

Samh_forever2000@yahoo.com[1], omran-taha2123@yahoo.com[2]

الخلاصة

لقد شهدت الحقبة المنصرمه لاسيما السنوات الاخيرة من القرن العشرين ومطلع القرن الحالي سلسلة واسعة من تدهورات الاوضاع الاقتصادية والاجتماعية والسياسية الامر الذي انعكس سلبا على واقع المجتمع العراقي بما فية العامل النفسي والصحي وكان من بين هذه الانعكاسات ظهور مرض التوحد وتفشية بصوره ملموسة بين فئة الاطفال.لذا كان لزما على الباحثين والمختصين ان يأخذوا المسألة على محمل الجد لكي يوجدوا حلولا جذرية من شأنها ان تعمل للحيلولة من انتشاره ،او التخفيف منه على الاقل. في هذا البحث تم معالجة جزئيه معينه من هذا المرض حيث تبنى فكرته على تشخيص المرض من خلال ما يطرأ على المريض من اعراض و يتم ادخالها الى شبكة عصبية مطورة تدعى (الشبكة العصبية المتغيرة) تعتمد فيها على

(Back propagation & simulated annealing & single link list).

يتم تدريب الشبكة بأدخال اعراض المريض اليها وبعد التدريب والتعلم ومقارنتها بالاعراض القياسية يتم تحديد و تشخيص المرض مبكرا بصورة ادق من تشخيص الطبيب.

وقد اثبتت هذه الطريقه نجاحا ملموسا من خلال نتائج التجارب التي اجريت على بعض المرضى.

المقدمة

يعتبر مرض التوحد احد الامراض التي تصيب الاطفال والذي يتسبب بتدمير خلايا المخ .وقد اجريت علية العديد من الاختبارات الا انها لم تصل الى النجاح المطلوب سواء كان على مستوى تحديد اعراض المرض او معالجته. [1] وقد تم استخدام الشبكات العصبية في كثير من التطبيقات ومنها التطبيقات الطبيه وغير الطبية وذلك بسبب معمارية الشبكه في امكانيتها للتدريب والتعلم ومن خلال النتائج المتحصل عليها اثبتت نجاحها في التوصل للنتائج بصوره ادق واسرع من الطرق التقليدية. [4]

التوحد Autism: هو احد الاضطرابات التابعة لمجموعة من اضطرابات التطور المسماة باللغة الطبية اضطرابات طيف التوحد(Autism Spectrum Disorders ASD). [2]

اسباب مرض التوحد

1- **اسباب وراثية:-** بعض الجينات تتسبب بالذاتوتية مما يؤدي الى جعل الطفل مصابا بمرض التوحد وايضا تؤثر على دماغه
2- **اسباب بيئية :-** وذلك ينتج عن اسباب كالتلوث البيئي او فيروسات معينة او غيرها
3- **اسباب اخرى:-** مثل المشكلات التي تتخلل عملية الولادة او عن المشكلات التي قد تحدث من خلال مخاطر الولادة.

أعراض التوحد

1- المهارات الاجتماعية: لا يستجيب لمناداة اسمه ,لا يكثر من الاتصال البشري المباشر ,غالبا يبدو انه لا يسمع محدثه .
2- المهارات اللغوية : يبدأالكلام (نطق الكلمات) في سن متأخرة مقارنة بالاطفال الاخرين ,يفقد القدرة على قول كلمات او جمل معينة كان يعرفها في السابق , يقيم اتصالا بصريا حينما يريد شيئا ما .
3- السلوك: ينفذ حركات متكررة مثل اكمزاز,الدوران او التلويح باليدين,ينمي عادات وطقوس يكررها دائما ,دائم الحركة ,يصاب بالذهول او الانبهار من اجزاء معينة من الاغراض مثل دوران عجل في سيارة لعبة.
4- الشعور بالحزن الشديد بسبب تغيرات بسيطة.
5- ترديد اصوات الاخرين لفظ حدوي تكرار الكلمات او الجمل .
6- تظهر احيانا اعراض جسدية تشمل مشاكل الجهاز الهضمي ومشاكل النوم.
7- تظهر تأثيرات على اجزاء من المخ تتحكم في العواطف وحركات الجسم.

8- علاج التوحد

لا يوجد الى يومنا هذا علاج يلائم كافة مصابي مرض التوحد بالمقدار ذاتة كما يمكن القول ان علاج هذا المرض متنوع جدا ومتعدد ايضا وينقسم علاجة الى عدة اقسام:

1- العلاج الدوائي: العلاج بالادوية التي تركز على علاج المشكلة الموجودة في (الخلايا العصبية) ومحاولة السيطرة عليها وتخفيضها .
2- علاج مشكلة اللغة والنطق:- التعامل معهم بطريقة خاصة تساعدهم على التواصل مع غيرهم والاندماج مع الاخرين .
3- العلاج البديل.

اعمار الذين يصابون بالتوحد

1- (0-1) سنة :لا يستجيب لصوت والده ,لا يستجيب عندما يدعى بأسمه ,لن ينظر للناس بأعينهم, لا يبتسم ولا يستجيب(لا يتفاعل) للاشارات الاجتماعية من البيئة.

2- (2-6): في هذا العمر تكون اكثر وضوحا لا يلفظ الكلمات ,لا يلعب العاب التخمين ,لا يكون جملا ,يفقد المهارات اللغوية ,لا يبدي اهتمام عندما يشير شخص الى شيء مثل طائرة تحلق فوق رأسه[2] .

4- الشبكات العصبية

هي تقنيات حسابية مصممة لمحاكاة الطريقة التي يؤدي بها الدماغ البشري مهمة معينة ، مكونة من وحدة معالجة بسيطة هذه الوحدات ماهي الا عناصر حسابية تسمى عصبونات

(Neurons Nodes) [5]

انواع الشبكات العصبية

1-الشبكات العصبية ذات التغذية الامامية(Networks Neural Forward Feed):-تكون هذه الشبكات خالية من الحلقة المعلقة بين الوحدات المكونة لها وتكون هذه الشبكات الاكثر استخداما.

2-الشبكات العصبية ذات التغذية الراجعة(Networks Neural Back Feed):-هي الشبكات التي لمخرجاتها طريقا خلفيا مرة اخرى لتصبح مدخلات لبعض افضل النتائج .

3-الشبكات العصبية ذات الترابط الذات (Networks Neural Associative auto). **[6]**

5- القوائم الموصولة

هي نوع من هياكل البيانات وتتألف من مجموعة من الخلايا المترابطة بينها وكل عنصر فيها يسمى عقدة وهذه العقدة فيها حقلين , الحقل الأول يستخدم لتسجيل القيم , اما الحقل الثاني فهو مؤشر يؤشر لعنوان العقدة التالية او السابقة او NULL في حال كانت هذه الحلقة هي الحلقة الاخيرة.

انواع القوائم المتصلة :-يوجد عدة انواع اهمهما :

- القائمة المتصلة المفردة single Linked List
- لقائمة المتصلة المزدوجة double Linked List
- القائمة المتصلة الدائرية circular Linked List

القائمة المتصلة المفردة (single Linked List) :وتعرف بأنها مجموعة من العقد ، كل عقدة مرتبطة مع العقدة التي تليها ،وهكذا حتى نصل الى عقدة تشير الى NULL والتي ستكون العقدة الاخيرة. [3]

المحاكاة التأقلمية

تعد تقنية المحاكاة التاقلمية هي تقنية احتمالية لتقريب الحل الامثل لوظيفة معينة. على وجه التحديد، هو ميتاهوريستيك لتقريب الحل الامثل في مساحة بحث كبيرة [7]. وغالبا ما يتم استخدامها عندما تكون مساحة البحث منفصلة (على سبيل المثال، كل الجولات التي تزور مجموعة معينة من المدن). أما بالنسبة للمشاكل التي يكون فيها إيجاد

حل عالمي تقريبي أكثر أهمية من إيجاد حل محلي دقيق في فترة زمنية محددة، فقد يفضل أن تكون المحاكاة التاقلمية بديلا [8].

خوارزمية المحاكاة التأقلمية

الطريقة المقترحة

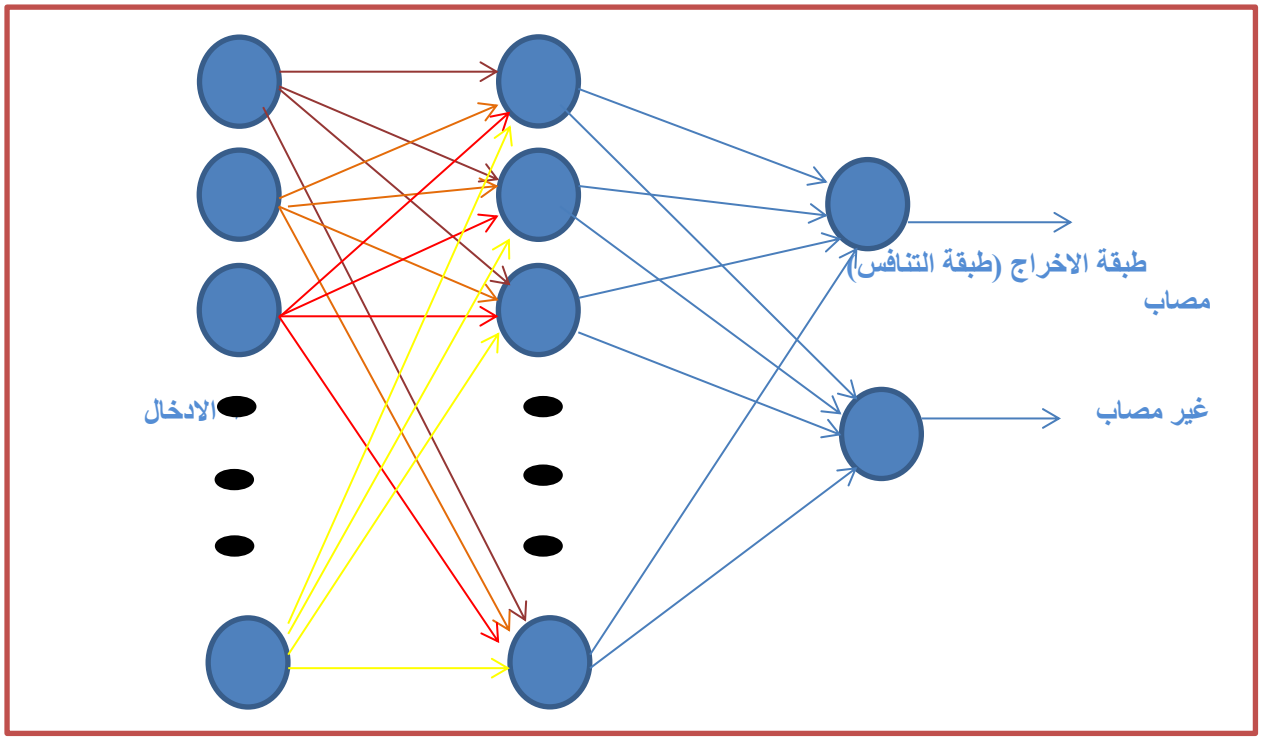
- Let $s = s_0$
- For $k = 0$ through k_{max} (exclusive):
  - $T \leftarrow \text{temperature}(k/k_{max})$
  - Pick a random neighbour, $s_{new} \leftarrow \text{neighbour}(s)$
  - If $P(E(s), E(s_{new}), T) \geq \text{random}(0, 1)$:
    - $s \leftarrow s_{new}$
- Output: the final state s

كما ذكر سابقا فأن من المشاكل التي يعاني منها الاطباء تحديد المرض بصوره مبكرة من خلال الاعراض التي تظهر على الاطفال .

ولحل هذه المشكلة تم اقتراح طريقة تتلخص في استخدام الشبكات العصبية المطورة لتكون مناسبة لتحديد المرض من خلال نفس هذه الاعراض.

تتلخص الطريقة المقترحة ببناء شبكة عصبية من نوع التغذية الرجعية (Back propagation) حيث يتم تنفيذ الجزء الاول وهو الانسياب خلفا لتدريب الشبكة اما الجزء التالي يتم فيه استخدام التحاسبات التأقلمية والتي يتم من خلالها تعليم الشبكة ولتحديد عقد الطبقة المخبأة تم استخدام هيكلية القوائم الموصولة لتحديد العقد ذات الاوزان الكبيرة والاوزان الصغيرة والتي تحذف فيما بعد لتحديد المرض.

تتألف الشبكة من طبقة ادخال(input) تمثل اعراض المرض والطبقة المخبأة (hidden) والطبقة الاخيرة هي طبقة الاخراج (التنافس) (output) يتم من خلالها تحديد هل ان هذا الطفل مصاب ام غير مصاب .



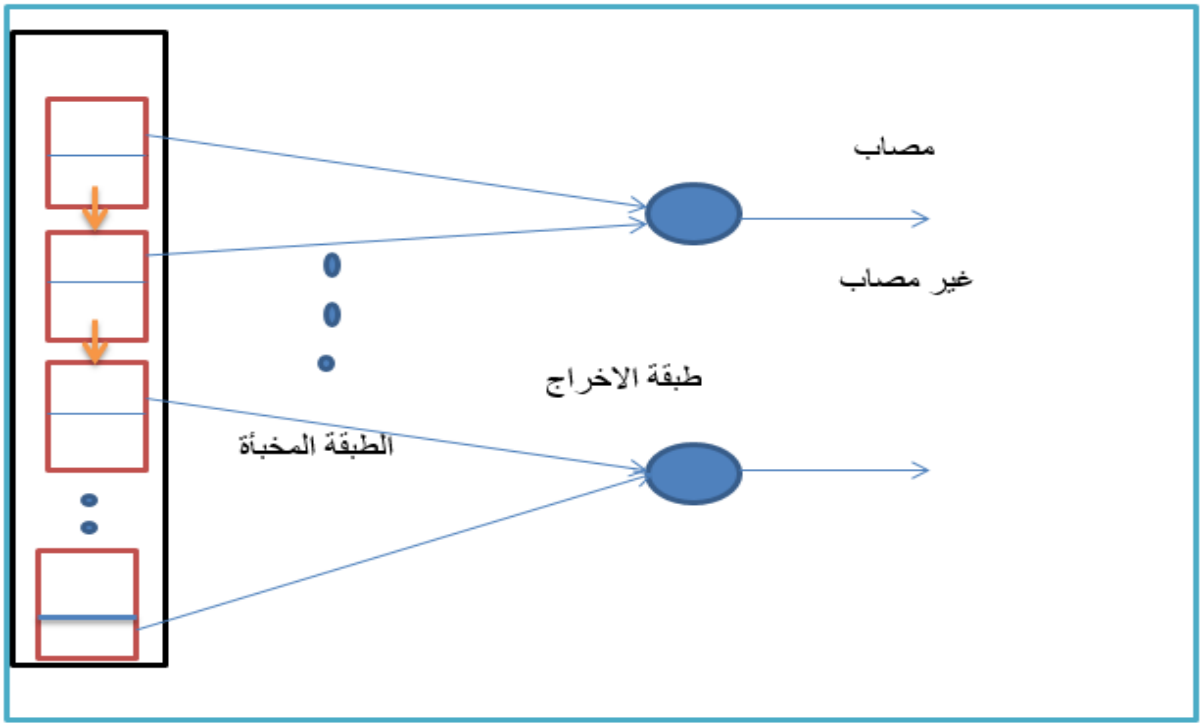
الشكل 1 : مخطط الشبكة العصبية المقترحة

يتم تدريب الشبكة العصبية من خلال توليد اوزان عشوائية وبعد عملية التدريب تأتي المرحلة الثانية (مرحلة التعلم) وتستخدم بها التحاسبات التأقلمية لاعطاء تغيير كبير في البداية بالاوزان وتستمر عملية التغير الى ان تثبت تدريجيا لتحصل على افضل نتائج وبوقت اقل من مرحلة التغذية الرجعية التي تقع في مشكلة القاع المحلي وتأخذ وقت اطول .

لتحديد النتيجة النهائية للمرضى في اصابتهم او عدمها من خلال التحكم بعدد العقد المخبأة وذلك بأستخدام القوائم الموصولة الاحاية حيث تكون هناك مجموعتين من العقد مجموعه(A) ومجموعه(B) وهناك قيمة عتبة اذا كانت القيمة الناتجة للعقد اصغر من قيمة العتبة يتم حذف العقد وتستمر الى اخر الطبقة وبالنتيجة نقارن المجموعتين فأذا كانت مجموعة (A) اكبر من مجموعه(B) فأوزان الاخراج للمجموعة (A) تعطى لها قيمة اضافية والعكس صحيح.

يتم حساب عدد العقد المخبأة الكلي من خلال المعادلة التالية:

عدد المخبأة=عدد المدخلات 1+2X.....(1)



الشكل(2) يوضح عمل الطبقة المخبأة

- خوارزمية التغذية الراجعة Start algorithm of BP

الخطوة الاولى : تهيئة المعلومات.

الخطوة الثانية : حساب الخطأ error & learning rate

الخطوة الثالثة : حساب الطبقة المخفية من المدخلات مع وزنها.

الخطوة الرابعة : حساب المخرجات من المخبأة مع وزنها.

الخطوة الخامسة : حساب الخطأ من (النتائج الفعلية).

الخطوة السادسة : تعديل مدخلات الطبقة المخبأة واوزان الاخراج المخبأة.

الخطوة السابعة : كرر الخطوة 2اذا (Loop<=max)

End algorithm

- خوارزمية شبكة الانسياب الخلفي المتغيرة(BPV)

Start algorithm of BPV

الخطوة الاولى : تهيئة باراميرات للشبكة

الخطوة الثانية :حساب العقد المخبأة من طبقة المدخلات واوزانها

الخطوة الثالثة : حساب عقد المخرجات من الطبقة المخبأة واوزانها

الخطوة الرابعة :اذا الوزن < من خط العتبة فأن اخذ الناتج من (Wg/t) Δ

الخطوة الخامسة :كرر الخطوة الرابعة اذا (loop1>=max1)

الخطوة السادسة :حساب الاوزان الجديدة للشبكة

الخطوة السابعة :حساب قيم العقدة المخبأة الجديدة

الخطوة الثامنة :اذا Value<threshold2 فأن امسح العقدة

الخطوة التاسعة :كرر هذه الخطوة عندما (loop2>max 2)

الخطوة العاشرة :اذا (output1<output2) فأن الابناء تصاب بالعدوى

End algorithm

-النتائج

من خلال تنفيذ هذه الخوارزمية وبأخذ عدد من العينات (25) عينة عشوائية من الاطفال اعمارهم تتراوح بين (2-6) سنوات ، هذه العينات قسم منها مصابين واخرون غير مصابين, وبعد ادخال الاعراض الى الشبكة والتي يبلغ عددها (7) اعراض رئيسية وتدريبها تم الوصول الى نتائج صحيحة بنسبة 95 بالمئة . والجدول رقم (1) يوضح الباراميترات وقيمها المستخدمة في الشبكة

جدول 1: يوضح الباميترات المستخدمة في الشبكة

	الباميترات المستخدمة لخوارزمية BPV	قيمتها
1	• عدد الدورات الخارجية	15
2	• اقل عدد من الدورات لتقليل درجة الحرارة	10
3	• معامل درجة الحرارة (الفا)	0.9
4	• اقصى عدد دورات لتبديل الاوزان الجديدة بالقديمة	5
5	• اقصى عدد للعقد المخبأة	15
6	• عدد ارتباطات الاوزان الابتدائية	98

7	• مدى الاوزان الممولة عشوائيا	(0.5 -0.5-)

جدول 2 : يوضح النتائج حسب قيم الباميترات وعدد الدورات الكلية

عدد الدورات / حجم العينات	عدد الارتباطات	الجهد الكلي للشبكة	نسبة النجاح
5	11	275	%90
10	11	550	%95
15	11	825	%97
20	11	1100	%95
25	11	1375	%96
5	13	390	%90
10	13	780	%92
15	13	1170	%94
20	13	1560	%96
25	13	1950	%97
5	15	525	%90
10	15	1050	%93
15	15	1575	%95
20	15	2100	%97
25	15	2100	%97

من خلال جدول (2) يتبين لنا ان النتائج تتحسن بزيادة عدد الدورات و عدد العينات تزداد نسبة النجاح وهذا يدل على ان تعقيد الشبكة يزيد من نسبة تعلم الشبكة بشكل ممتاز بالمقابل جهد الشبكة يزداد وهذا يزيد من الوقت المخصص للتنفيذ ويزداد تعقيد الخوارزمية.

الاستنتاجات والتوصيات

من خلال تطبيق النظم تم التوصل الى انه يمكن تحديد اعراض المرض بصورة مبكرة بأستخدام هذه الخوارزمية التغذية الرجعية المتغيرة وهي نوع من انواع الشبكات العصبية المطورة .

وقد اثبتت هذه الخوارزمية كفائتها وذلك من خلال نسبة النجاح المتحققة بعد تطبيق النظام والتي بلغت 97% وذلك بعد تطبيق النظام على (25) حالة مابين مصاب وغير مصاب.

وكفكرة مستقبلية يمكن تطوير النظام ليتم تطبيقه في تحديد المرض ومعالجته بنفس الوقت وذلك من خلال مجموعة من الاختبارات توجه للطفل ومن ثم قياس مدى استجابته من خلال هذه الشبكة ومن ثم اعادة توجيه الاختبار بأسلوب مختلف وهكذا الى ان يتم تحسن المريض.

المصادر

[1] Caronna EB, Milunsky JM, Tager-Flusberg H. Autism spectrum disorders, clinical and research frontiers. Arch Dis Child. 2008; 93(6):518–23.

[2] DSM-5 News and Updates". Autism Speaks. Retrieved 27 August 2014.

[3] Okasaki Chris, "purely Functional Random- Access Lists (PS)", In Functional Programming Languages and computer Architecture,

(May 7/2015).

[4] Stuart Russell and Peter Norving , "Artificial Intelligence:Amodren Approach",3rdEd.,Wiley India,(2015).

[5] John Wiley & Sons, "The Roots of Backpropagation. From Ordered Derivatives to Neural Networks and Political Forecasting. New York, (1994)

[6] DREYFUS, STUART E, "Artificial neural networks, back propagation, and the Kelley-Bryson gradient procedure". Journal of Guidance, Control, and Dynamics. 13 (5): 926–928. doi:10.2514/3.25422. ISSN 0731-5090, (1990).

[7] "Sequential Monte Carlo samplers - P. Del Moral - A. Doucet - A. Jasra - 2006 - Journal of the Royal Statistical Society: Series B (Statistical Methodology) - Wiley Online Library". Journal of the Royal Statistical Society, Series B. 68: 411–436. doi:10.1111/j.1467-9868.2006.00553.x. Retrieved 2015-06-11.

[8] M. P. Vecchi and S. Kirkpatrick, "Global wiring by simulated annealing", IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 1983.