RADIO PROPAGATION MODELING OF JOS SUBURBS AT 900MHZ USING AN ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM

[1]Deme, A. C.*, [2]Usman, A. U. and [3]Choji, D. N.

[1]ICT Directorate, University of Jos, Jos-Nigeria
[2]Department of Electrical and Electronics Engineering, Federal University of Technology, P.M.B. 65, Minna , Nigeria
[3]Department of Computer Science, University of Jos, Nigeria.

***Corresponding Author: E-mail:** demeabraham@gmail.com ,demea@unijos.edu.ng

Date Manuscript Received: 07/01/15 *Date Accepted: 06/07/15* *Published: December 2015*

ABSTRACT
This study investigates radio propagation modeling of the suburbs of Jos, Nigeria, at an operating frequency of 900MHz, using an Adaptive Neuro-Fuzzy Inference System (ANFIS) Technique. Path loss values computed based on received power measurements obtained from Base Transceiver Stations situated across the suburbs of the city were used to train, validate and test the ANFIS model for capacity to predict path loss. Results indicated that the ANFIS model with a Root Mean Square Error (RMSE) value of 4.74dB offers an improvement in prediction accuracy over the COST 231 Hata model, which has an RMSE value of 5.90dB.

INTRODUCTION

Transmission between the base station and the mobile station is usually accompanied by some power loss, and this loss is known as path loss and it depends particularly on the carrier frequency, antenna height and distance (Ashis, 2012). There is a variety of models used by researchers and radio engineers to predict path loss across a particular terrain. A model that suits a given terrain (or environment) may not necessarily be suitable to another. Hence, radio engineers and researcher have been evolving new techniques for accurate prediction of path loss.

For the purpose of path loss prediction, a wide range of radio propagation models have been usedglobally. Some of these models include empirical and deterministic models. Empirical models are path loss prediction models that are based on observations and measurements alone. Unfortunately, these empirical models though easier to implement, are less sensitive to the environment's physical and geometrical structures and not so accurate while the deterministic models which though are more accurate are computationally inefficient and require more detailed site-specific information which is often difficult to come by (Abhayawardhana et al. 2005).

An Adaptive Neuro-Fuzzy Inference System (ANFIS) is an intelligent computation technique that is highlyefficient in handling complex non-linear problems. Because artificial intelligence is adaptive andrelies on observed data rather than on an analytical modelof the system, the resulting scheme is robust, efficient andcapable of reflecting changes in the wireless signal behavior and hence they mitigate the error in wireless signal prediction(Alotaibi et al., 2008).

Alotaibi *et al*. (2008) used ANFIS as wireless signal predictor on a private mobile network-terrestrial trunked radio (TETRA) network, where it was shown to outperform some empirical models and to be marginally better than RBF-NN predictor. Also, ANFIS was used by Turkan *et al.* (2010) to predict path loss based ondata obtained in the 900 MHz band in Harbiye region of Istanbul, Turkey. ANFISprediction error was shown to be less than that obtained using Bertoni–Walfish model. Faihan *et al.* (2007) developed a robust ANFIS based prediction model for the city of Riyadh city – Saudi Arabia. ANFISprediction error was shown to be less than that obtained using Bertoni–Walfish model (Walfisch and Bertoni, 1998).

This study investigated the applicability ofan ANFIS based predictor for path loss prediction across the suburbs of Jos, Nigeria. The ANFIS predictor is developed and trained on measured data obtained from Base Stations situated within the terrain. The prediction accuracy of the ANFIS predictor is compared with that of the empirical COST 231Hata based on root mean square error (RMSE) and the square of multiple correlations,R-square (R^2).

ADAPTIVE NEURO-FUZZY INFERENCE SYSTEMS

An ANFIS is a combination of an Artificial Neural Network (ANN) and a Fuzzy Inference System (FIS) to form an intelligent adaptive system capable of solving complex non-linear problems. ANNs are quite useful in modeling systems where there is no mathematical relationship between input and output patterns. This stems from the fact that, as systems that mimic the human brain, ANNs can be trained using input patterns and target output, and then used to predict a result given new set of inputs. Based on the concepts of fuzzy set theory, fuzzy if-then rules, and fuzzy reasoning, FIS is a computational network capable of modeling human knowledge and reasoning.

ANFIS was first proposed by Jang (1993) to combine the learning ability of ANNs with the ability of fuzzy systems to interpret imprecise information (Abraham *et al*., 2013). ANFIS is based on the first-order Takagi–Sugeno–Kang (TSK) model. A brief ANFIS architecture as described in (Abraham et al.,2013) is as follows: Figure 1 shows n example of such fuzzy inference system with two inputs, x and y and one output which is a function of the inputs. For the TSK inference system, the rule is constructed as follows:

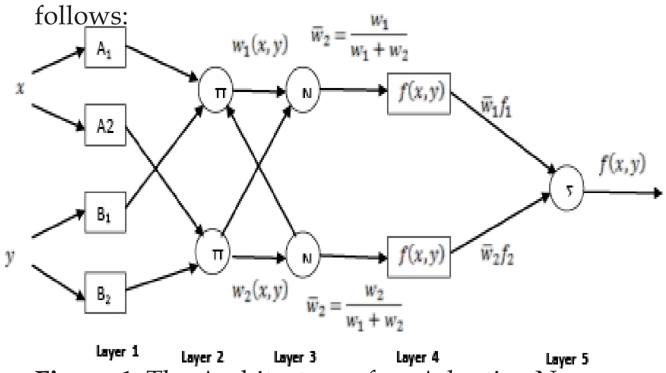


Figure 1: The Architecture of an Adaptive Neuro-Fuzzy Inference System

If x is A_i and y is B_i, THEN $f_i = p_i x + q_i y + r_i$

Where A_i and B_i are linguistic labels in the input spaces x and y respectively and f_i is a local function which depends on x and y.

Layer 1 is the fuzzification layer which generates membership grades for each linguistic label for any input value; these values are defined by membership functions. The common membership functions are the bell and triangular functions depicted in Figure 2. The bell function used in this work is described by the three parameters, a, b and c in equation (1).

$$\mu_{Ai}(x) = \frac{1}{1 + \left[\frac{x - c_i}{a_i}\right]^{2b_i}}$$

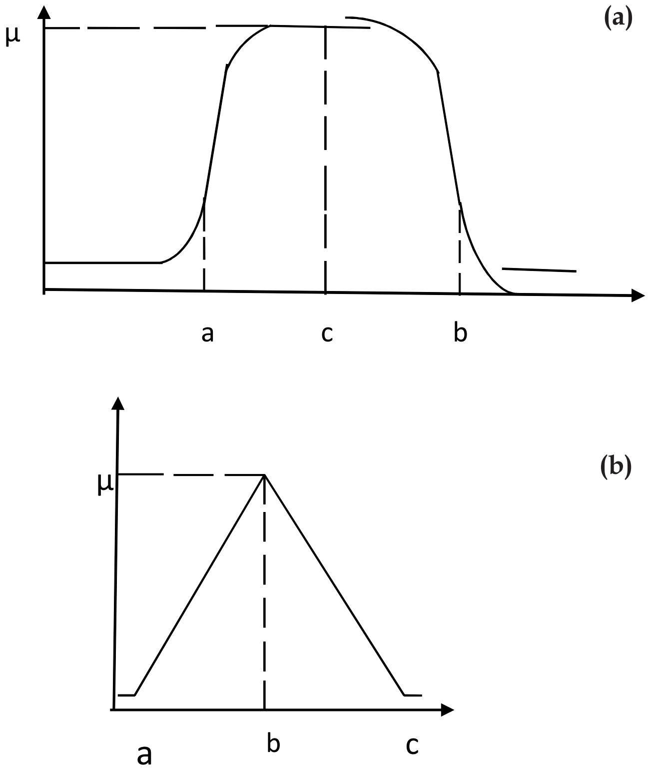


Figure 2: (a)bell and (b) triangular membership functions

Layer 2 is the application of the fuzzy operator and the output of every node in this layer is the product of all the incoming signals into the node as given by equation (2).

$$w_i = \mu_{Ai}(x_i) X \mu_{Bi}(y_i)$$

(2)

Layer 3 produces an output which is the so-called normalized firing strength of each rule according to equation (3):

$$\overline{w_i} = \frac{w_i}{\sum_{j=1}^{2} w_j}$$

(3)

Layer 4 is a layer of adaptive nodes each with a node function described by equation (4):

$$\overline{w_i} f_i = \overline{w_i}(p_i x + q_i y + r_i)$$

(4)

where, pi, qi and ri are called consequent parameters and the least-squares method is used to identify their optimal values.

Layer 5 is the defuzzification layer, that generates a crisp output given by equation (5).

$$f(x,y) = \sum_i \overline{w_i} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}$$

(5)

where, $\overline{w_i} f_i$ is the output of node i in layer 4 denoting the consequent part of rule i.

In the training algorithm suggested by Jang (1993), ANFIS uses a hybrid learning algorithm which is a combination of gradient descent and the least-squares approximation method in order to train the network. Gradient descent back-propagation algorithm is used for training the premise parameters and least-squares approximation is used for training the consequent parameters. The optimal consequent parameters are estimated in the forward pass, while the premise parameters remain fixed. In the backward pass, premise parameters are tuned, while the consequent parameters remain fixed.

THE COST 231 HATA MODEL

The COST 231 Hata (Purnima, 2010) Model is an extension of the Hata Model, which is also an extension of the Okumura Model. It was formulated to suit the European environments taking into consideration a wide range of frequencies (500MHz to 200MHz). The COST 231 Hata Model is one of the most widely used radio propagation models because of suitability for urban, semi-urban, suburban and rural areas. The COST 231 Hata Model is given by(Purnima, 2010).

Where,

- L = Median path loss in Decibels (dB)
- C=0 for medium cities and suburban areas
- C=3 for metropolitan areas
- f = Frequency of Transmission in Megahertz (MHz)(500MHz to 200MHz)
- h_B = Base Station Antenna effective height in Meters (30m to 100m)
- d = Link distance in Kilometers (km) (up to 20kilometers)
- h_R = Mobile Station Antenna effective height in Meters (m) (1 to 10metres)
- $a(h_R)$ = Mobile station Antenna height correction factor as described in the Hata Model for Urban Areas.
- For urban areas, a(hR) = 3.20 (log10 (11.75hr))2−4.97, for f >400 MHz
  For sub-urban and rural areas, a(hR) = (1.1log(f) - 0.7)hR - 1.56log(f) -0.8

PERFORMANCE EVALUATION INDICES

The statistical indices that form the bases for performance evaluation are based on Root Mean Squared Error (RMSE) and the square of multiple correlations, (R^2).

The RMSE is given by (Olasunkanmi *et al.*, 2014)

Where,

M – Measured Path Loss
P – Predicted Path Loss
N- Number of paired values

The square of multiple correlations is given by (Abraham *et a*l., 2013)

Where y_i is the measured path loss, is the mean of the measured path loss. R^2 can take on any value between 0 and 1, but can be negative for models without a constant, which indicates that the model is not appropriate for the data. A value closer to 1 indicates that a greater proportion of variance is accounted for by the model.

MATERIALS AND METHODS

Measurements were taken from 8 different Base Stations of the mobile network service provider (Mobile Telecommunications Network (MTN), Nigeria), situated within the within the Jos suburbs. The instrument used was a Cellular Mobile Network Analyser (SAGEM OT 290) capable of measuring signal strength in decibel milliwatts (dBm). Readings were taken within the 900MHz frequency band at intervals of 0.2 kilometer, after an initial separation of 0.05kilometer away from the Base Station.

The mobile network parameters obtained from the service provider include the following:

i) MeanTransmitter Height, H_T= 34 meters
ii) Mean Effective Isotropically Radiated Power, EIRP = 47dBm
iii) Transmitting Frequency, f_c = 900MHz

Creating the ANFIS Model Predictor

According to Alotaibi et al (2008), creating an ANFIS network involves specifying the number of network inputs, the number of fuzzy membership function (MF) per each input, the type of fuzzy MF, and the number of epochs. In this paper, the type of the MF chosen is the bell-shaped function, and the number of fuzzy MF per each input is 2 and the number of iterations is 30.

Path Loss Prediction using the ANFIS Model

The techniques adopted in this study include the following:

a) Splitting Base Station Data into 60% Training, 10% Validation and 30% Testing
This basically involves analyzing each base station separately by randomly splitting path loss data obtained from it into 60% training, 10% validation and 30% training. The essence of validation is to further refine the network construction.This technique simultaneously carries out a performance comparison of the ANFIS based models with the COST 231 Hata model on each base station.

b) Training with one Base Station data set and testing with a set from another
This is a test for generalization as described in (Abraham et al., 2014). The technique involves randomly training with data set from one Base Station and then testing with a data set from another Base Station. By implication, a given data set can both be used for training and testing.

RESULTS AND DISCUSSION

As stated in the previous section, the first comparative technique involves analyzing each base station data separately, by randomly splitting the data into 60% training, 10% validation and 30% training. Figures 3 to 6 show prediction comparisons ofthe ANFIS-based model with the COST 231 Hata model relative to the test output. It can be observed from Figures 3, 5 and 6 that the ANFIS based model exhibits a closer prediction, while Figure 2 shows that the COST 231 Hata outperforms the ANFIS model.

The performance statistics of the two models across all the Base Stations are presented in Table 1, which shows that the ANFIS model outperforms the COST 231 Hata counterpart on all but Base Stations 2 and 8. Geometric mean values show that the ANFIS model with an RMSE value of 4.18dB and R^2 of 0.6 is more accurate than the COST 231 Hata counterpart, which has an RMSE value of 5.87dB and a very poor R^2value of -0.10.

The sample graphical comparisons shown in Figures 7 to 10 are based on the second comparative technique. It can be observed that the ANFIS model exhibits a slightly closer prediction than the COST 231 Hata model as far as all the Figures are concerned. Results in Table 2 show that the ANFIS model gives a better prediction on all train/test pairings with the exception of BST6/BST2. Mean performance shows that there is a slight convergence in performance, with the ANFIS model just fractionally more accurate by about 0.56dB.

Finally, a combined performance evaluation based on the two comparative techniques shows that on the geometric mean, the ANFIS model with an RMSE value of 4.74dB offers an improvement over the COST 231 Hata model, which has an RMSE value of 5.9dB. Moreover, the ANFIS model's superior R^2 value of 0.72 indicates that it has higher correlation with the test data, compared with the COST 231 Hata model's R^2 value of 0.29.

MODEL	STAT.	BST1	BST2	BST3	BST4	BST5	BST6	BST7	BST8	GEOM. MEAN
ANFIS	RMSE(dB)	1.06	7.78	4.98	5.05	3.35	4.18	5.53	5.85	4.18
	R^2	0.99	0.11	0.78	0.73	0.92	0.78	0.69	0.56	0.60
COST 231 Hata	RMSE(dB)	4.89	6.32	6.03	6.62	6.52	5.32	6.08	5.42	5.87
	R^2	0.27	-0.75	0.04	-0.41	0.06	-0.06	-0.13	-0.14	-0.10

Table 1: Splitting data into 60% training, 10% validation and 30% testing

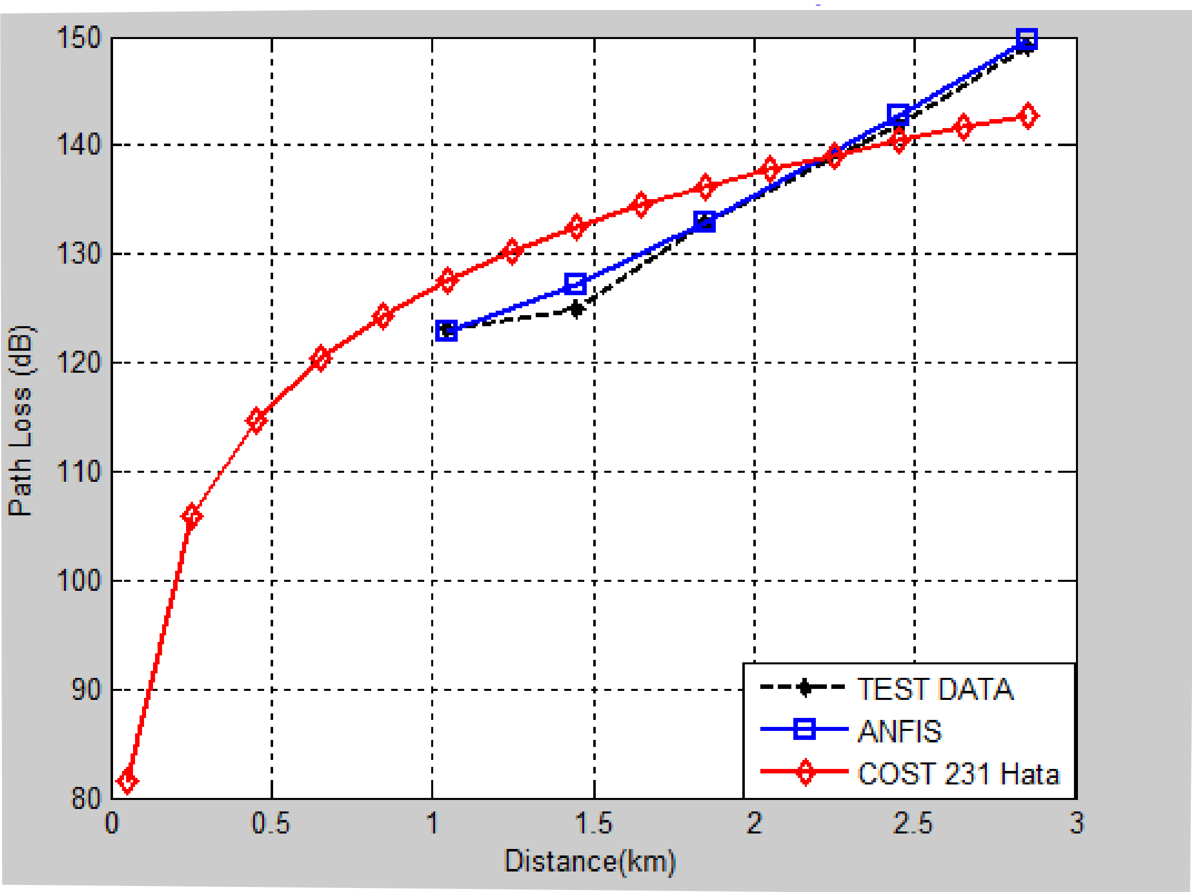


Figure 3 : Comparison on BST1

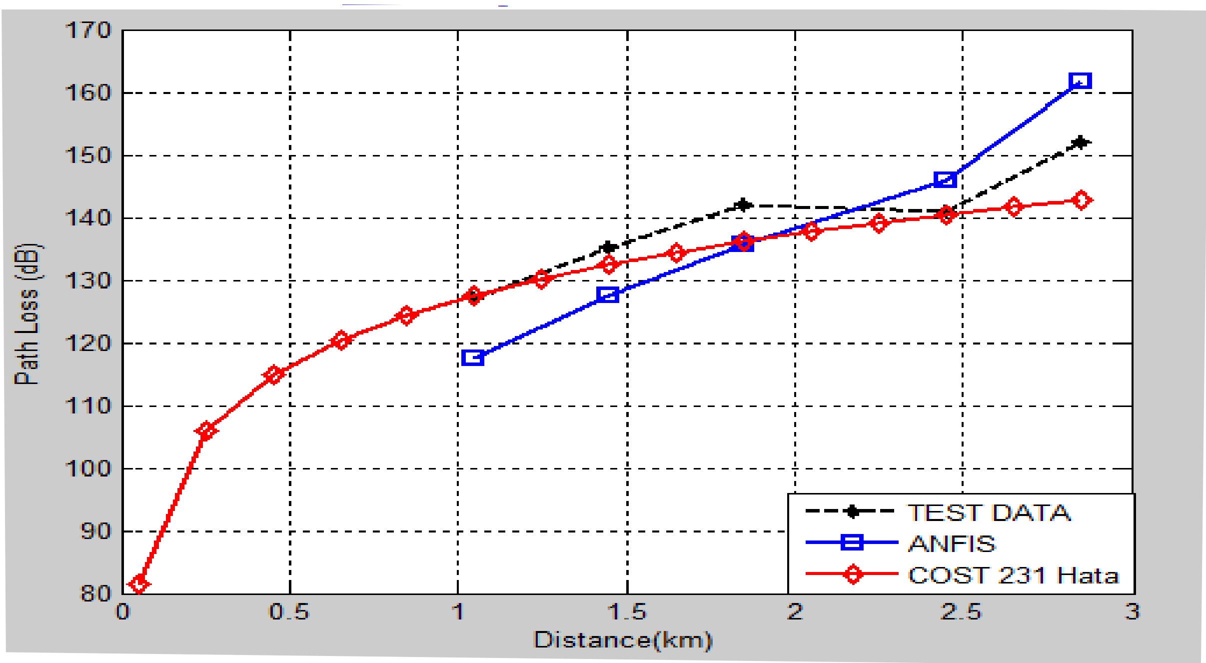


Figure 4 : Comparison on BST2

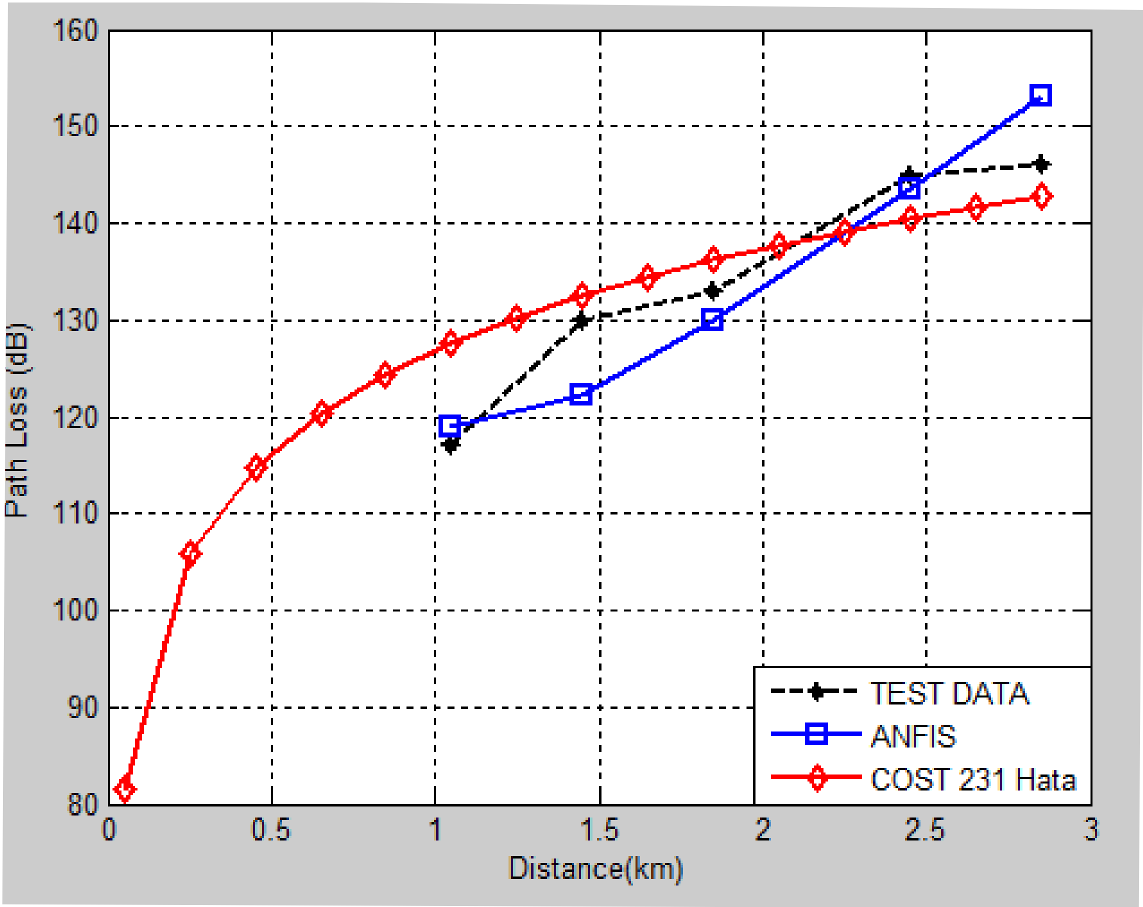


Figure 5 : Comparison on BST3

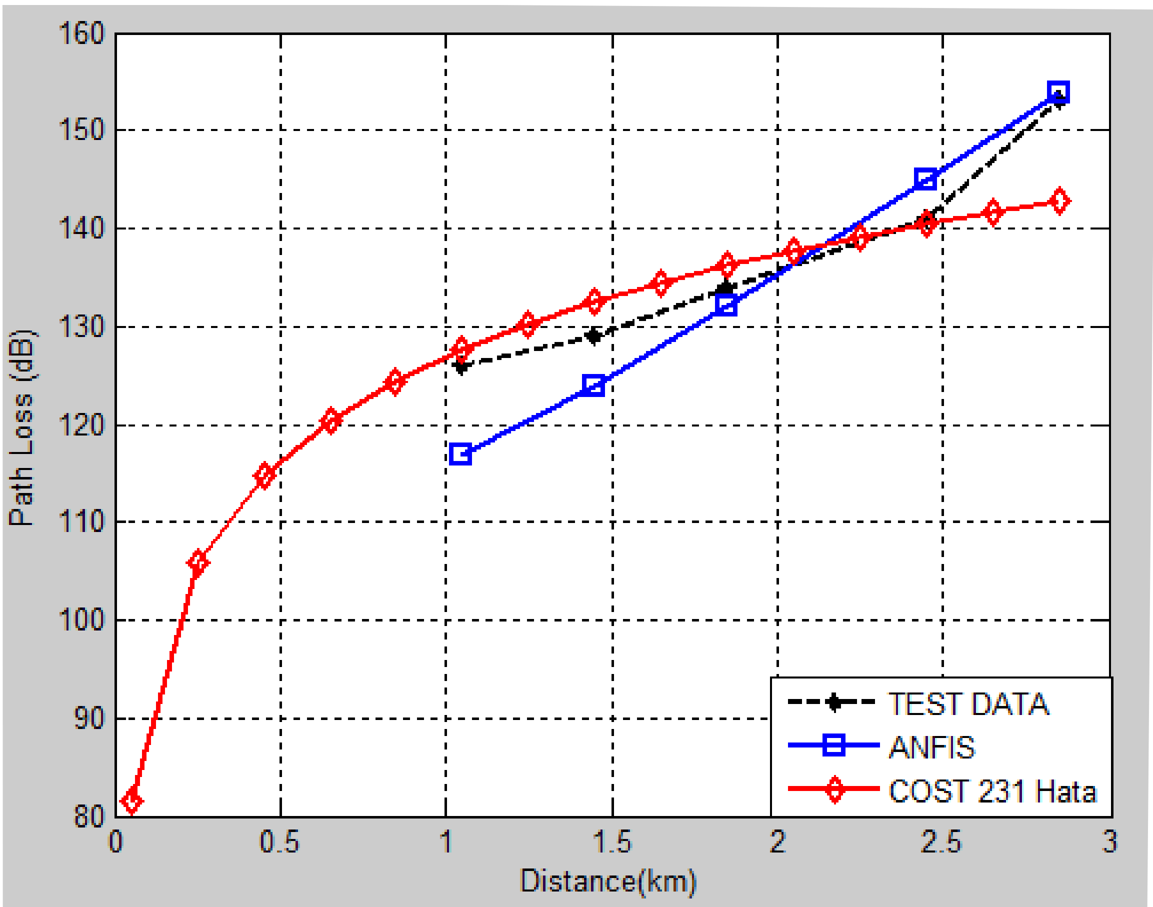


Figure 6 : Comparison on BST4

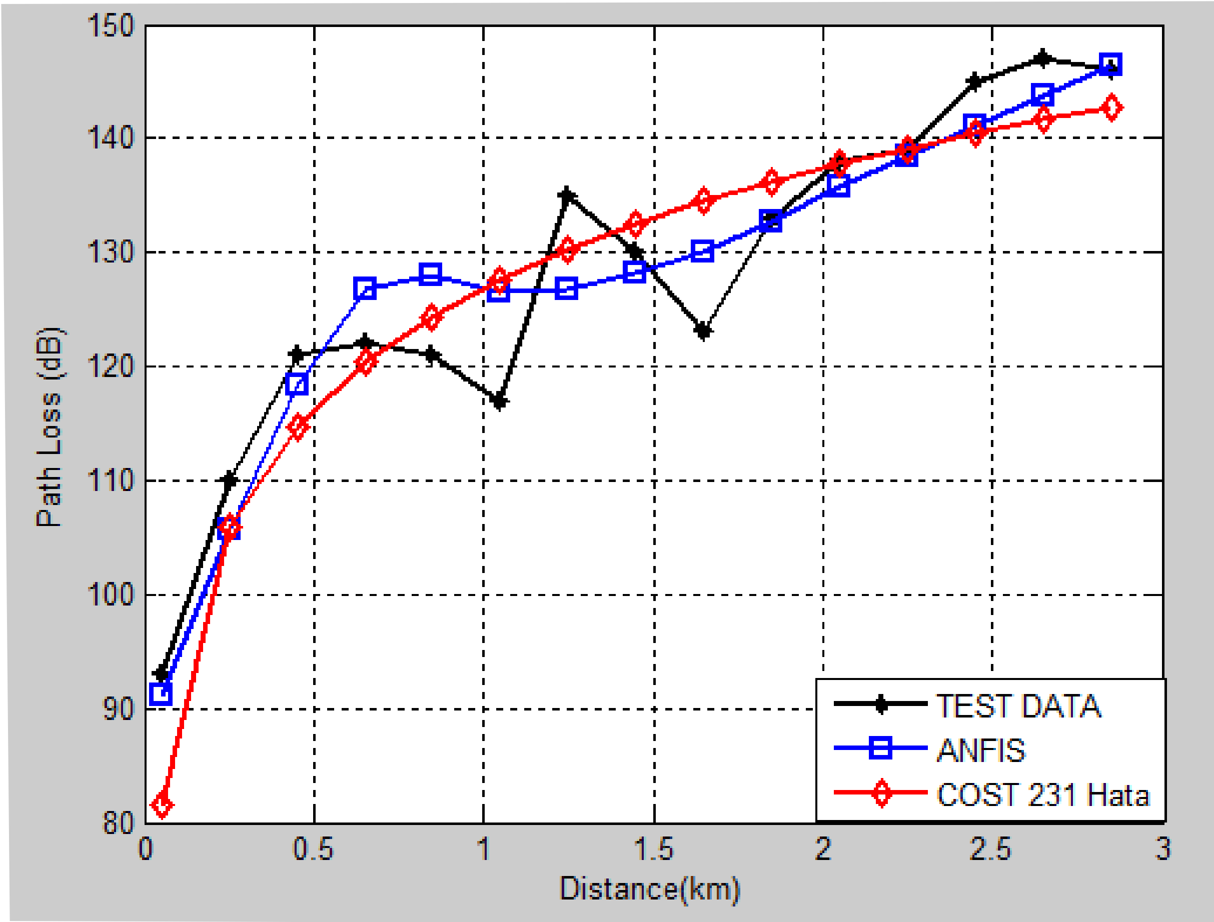


Figure 7: BST3/BST6 Pairing

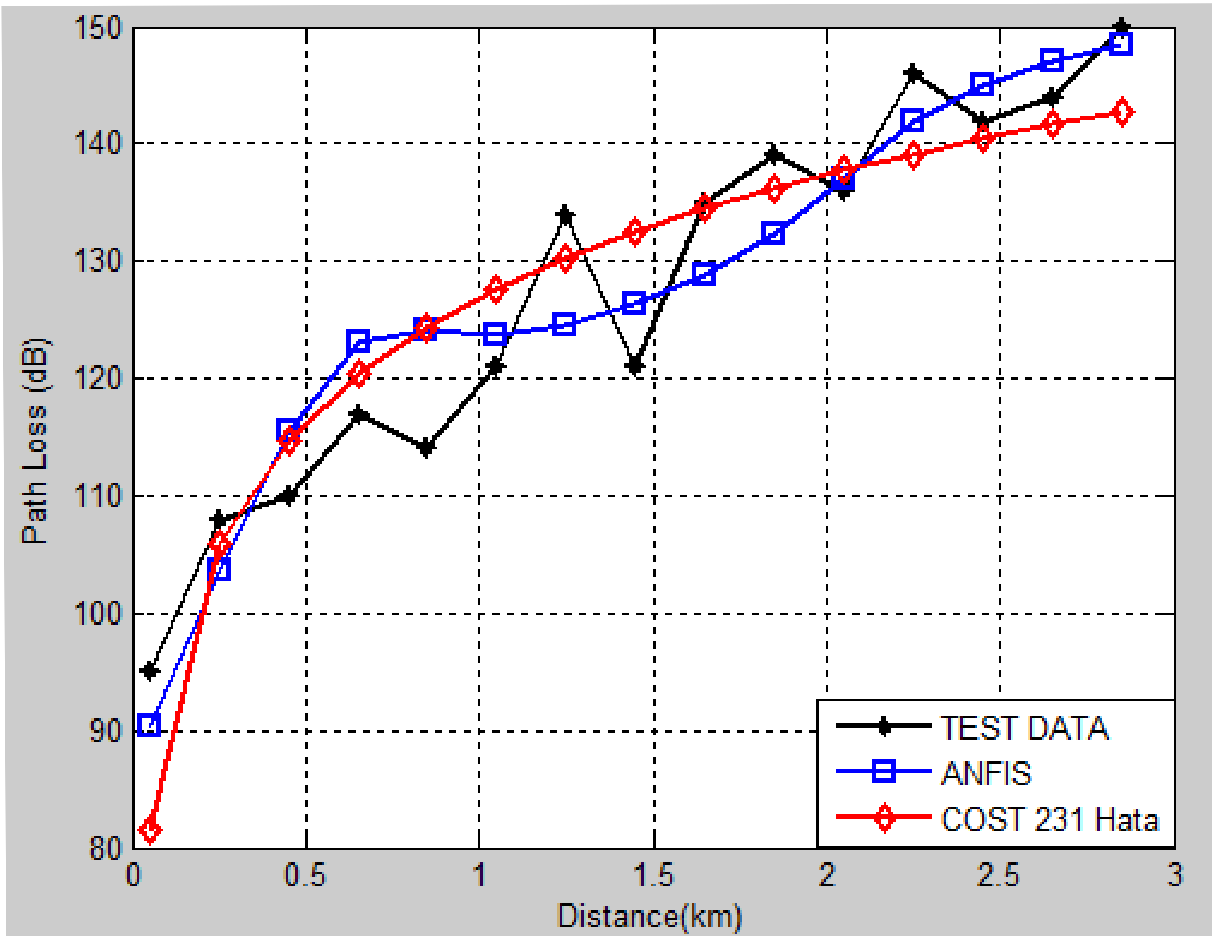


Figure 8: BST5/BST1 Pairing

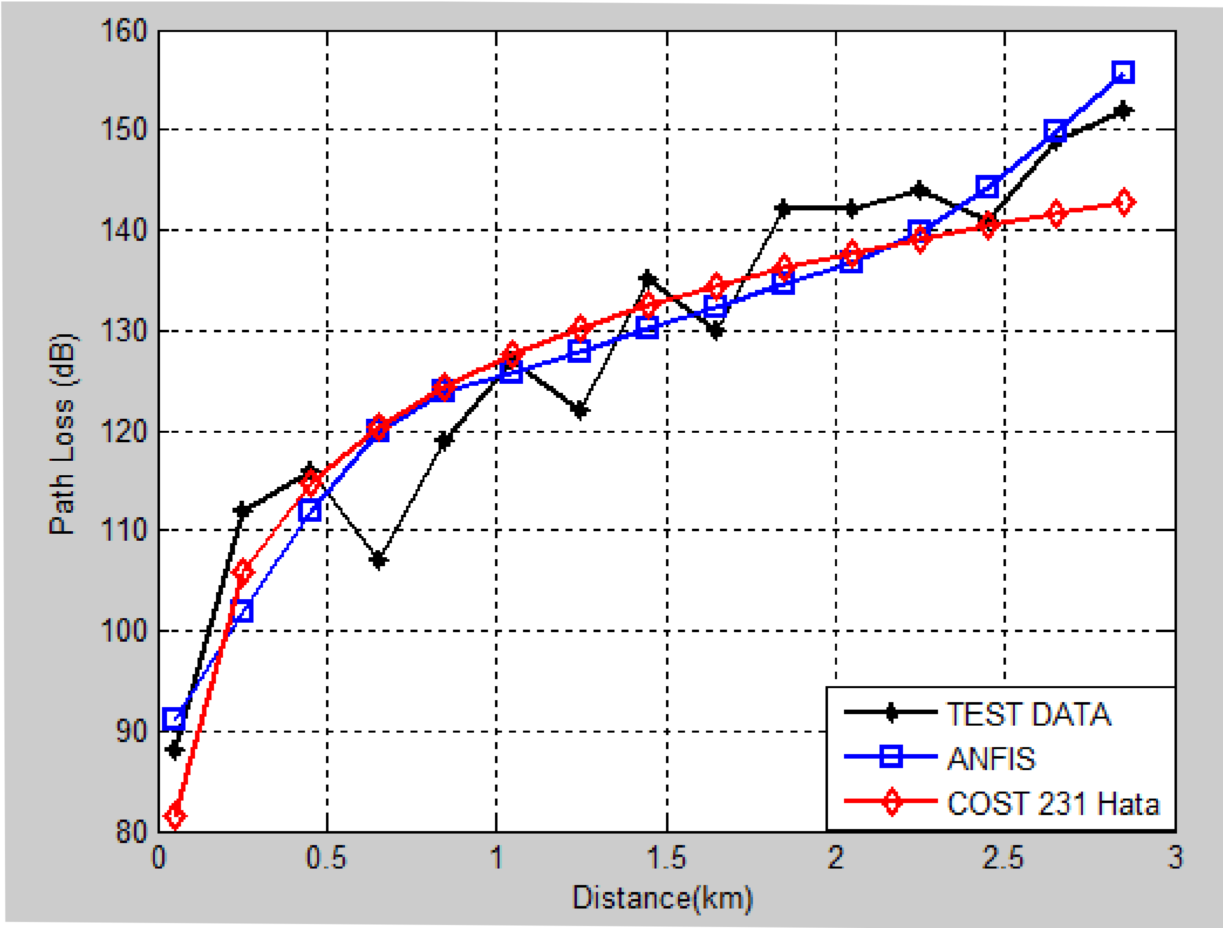


Figure 9: BST2/BST8 Pairing

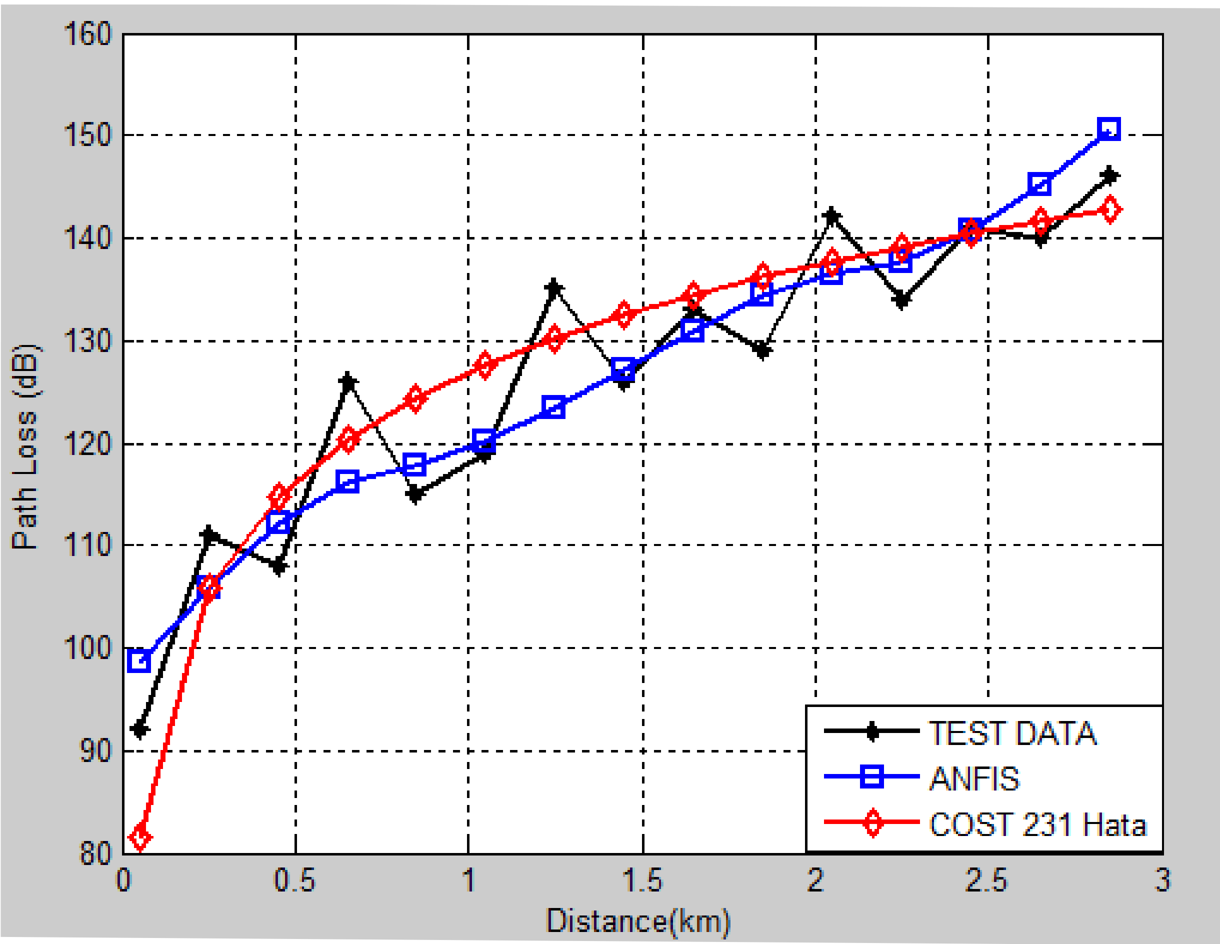


Figure 10: BST7/BST4 Pairing

MODEL	STAT.	BST3 /BST6	BST5 /BST1	BST2 /BST8	BST7 /BST4	BST6 /BST2	BST8 /BST5	GEOM. MEAN
ANFIS	RMSE(dB)	4.82	5.52	5.81	5.50	6.08	4.62	5.37
	R^2	0.89	0.88	0.89	0.86	0.83	0.9	0.87
COST 231 Hata	RMSE(dB)	6.03	6.52	6.32	6.08	5.33	5.42	5.93
	R^2	0.82	0.83	0.87	0.82	0.87	0.86	0.84

Table 2: Training with one Base Station data set and testing with a set from another

CONCLUSION

This study has successfully demonstrated the applicability of an Adaptive Neuro-Fuzzy Inference System technique for path loss prediction across the suburbs of the city of Jos, Nigeria. Prediction results based on the two comparative techniques implemented show that the ANFIS model with an acceptable RMSE value of 4.74dBis more accurate than the commonly used empirical COST 231 Hata model, which has an RMSE value of 5.9dB. Hence, the ANFIS model is recommended for path loss prediction across the area under investigation.

REFERENCES

Abhayawardhana, V.S., Wassel, I. J., Crosby, D., Sellers, M. P. &Brown, M. G. (2005). "Comparison of empirical propagation path loss models for fixed wireless access systems," 61th IEEE Technology Conference,Stockholm, pp. 73-77.

Abraham, U., Usman, O., Okereke, U. & Elijah E. Omizegba (2013).Macrocell path loss prediction using artificial intelligence techniques: *International Journal of Electronics*, 9, 6-9.

Alotaibi, F. D., Abdennour, A., & Ali, A. A. (2008). A robust prediction model using ANFIS based on recent TETRA outdoor RF measurements conducted in Riyadh city – Saudi Arabia. AEU – *International Journal of Electronics and Communications*, 62(9), 674–682.

Ashish, E. (2012). Pathloss Determination Using Okumura-Hata Model For Rourkela. A Thesis submitted to the Department of Electronics & Communication Engineering, National Institute of Technology Rourkela, Pp. 9.

Faihan, D., Alotaibi, A. A.&Adel, A. A.(2007). A robust prediction model using ANFIS based on recent TETRA outdoor RF measurements conducted in Riyadh city – Saudi Arabia. *International Journal of Electronics and Telecommunication*, 26, 545-575.

Olasunkanmi, F., Oseni, S. I., Popoola, R. O., Abolade, O. A. & Adegbola,O. (2014). Comparative Analysis of Received Signal Strength Prediction Models for Radio Network Planning of GSM900 MHz in Ilorin, Nigeria. *International Journal of Innovative Technology and Exploring Engineering*, 4(3), 45-50.

Purnima, K. S. (2010). Comparative Analysis of Propagation Path loss Models with Field Measured Data. *International Journal of Engineering Science and Technology*, 2(6), 2008-2013

Turkan, E. D., Berna, Y. H., & Apaydin, A. (2010). Fuzzy adaptive neural network approach to path loss prediction in urban areas at GSM-900 band. *Turkish Journal of Electrical Engineering &Computer Sciences*, 18(6), 1077–1094.

Walfisch, J. & Bertoni, H. L. (1998)."A theoretical model of UHF propagation in urban environments," IEEE Trans Antennas and Propagation, 36(12), 1788-1796.

DESIGNING AN OFFSET POISSON-GAMMA MIXTURE REGRESSION MODEL

Mbe, E. Nja[1*], E. C. Nduka[2] and P. U. Ogoke[3]

[1] Department of Mathematics, Federal University Lafia, Nasarawa State, Nigeria

[2] Department of Mathematics & Statistics, University of Port-Harcourt, Rivers State, Nigeria

[3] Department of Mathematics & Statistics, University of Port-Harcourt, Rivers State, Nigeria

[*] Corresponding Author/E-mail: mbe_nja@yahoo.com

Date Manuscript Received: 04/03/15 Accepted: 18/09/15 Published: December 2015

ABSTRACT

The problem of over-dispersion often encountered by Poisson distribution is what gave rise to the development of the Poisson-Gamma mixture distribution. This distribution is a mixture of a family of Poisson distributions with Gamma mixing weights. The log likelihood function of the Poisson-Gamma mixture distribution has been exponentiated in order to extract the canonical link of the mixture distribution. This link function is re-expressed as a sum of the loglinear link and an offset. By implication, modeling for the Poisson-Gamma mixture model amounts to modeling for the Poisson loglinear model and adding the offset term to the result. This model can be used to model HIV/AIDS infection rate for multiple sex partners. It is theoretically shown that the mixture distribution reduces to the Poisson distribution and that in large samples, the Poisson-Gamma mixture is approximated by the Poisson distribution.

KEYWORDS: ***Over-dispersion, Poisson distribution, Gamma mixing weights, Loglinear***

FULAFIA JOURNAL OF SCIENCE AND TECHNOLOGY | VOL. 1

INTRODUCTION

The need to use a Poisson-Gamma mixture distribution arises when the Poisson distribution exhibits over-dispersion. For the Poisson model, the mean and the variance are equal. If the variance of a Poisson model exceeds its mean, over-dispersion is said to occur. Over-dispersion is indicated if Pearson dispersion is greater than 1 Hibe, (2007). The problem with over-dispersion, common to most Poisson models is that it renders the parameter estimates biased. Where over-dispersion is a concern, the alternative models are:

(1) Quasi-Poisson
(2) Negative Binomial regression

MATERIALS AND METHODS

In this paper, we shall consider only the Negative Binomial regression as an alternative to the over-dispersed Poisson model. In designing an offset Poisson-Gamma mixture model, it is logical to use the order listed below:

*(i)*Poisson distribution
*(ii)*Gamma distribution
*(iii)*Poisson-Gamma distribution (Negative Binomial distribution)
(iv) Offset Poisson-Gamma mixture regression model

RESULTS AND DISCUSSION

THE POISSON REGRESSION

The Poisson regression is derived from the poisson probability mass function given Johnson, (2004), Freund, (1992) as

$P(X = x) = {\lambda^x e^{-x}}/{x!}$ and (Hilbe, 2007) as

$f(x_i\,;\,\lambda_i) = {e^{-t_i\lambda_i}(t_i\lambda_i)^{y_i}}/{x_i!}\ ,\ x = 0, 1, 2, \ldots, \lambda > 0$ ....(1)

where x_i is the count response, λ_i is the predicted count or rate parameter, t_i is the area or time in which counts enter the model

(1) reduces to $f(x_i, \lambda_i) = {e^{-\lambda_i}(\lambda_i)^{x_i}}/{x_i!}$

Where λ_i applies to individual counts devoid of

is the probability of observing any specific count y

The Likelihood function is

$$L(\mu_i\,, y_i) = \sum_{i=1}^{n}\{y_i \ln(\mu_i) - \mu_i - \ln(y_i!)\} \ldots\ldots\ldots\ldots\ldots\ldots$$

where μ_i is the predicted count.

(3) is used when the Poisson model is estimated by a

Generalized Linear Model. When estimation employ

full maximum likelihood update, we use

$$\mu_i = \exp(x_i, \beta)$$

and equation (4) becomes Hilbe, (2007)

$$L(\beta_i; y_i) = \sum\{y_i(x_i\beta) - \exp(x_i\beta) - \ln(y_i!)\} \ldots(5)$$

From (iii), $explogf = exp\{-\mu + xlog\mu - logx!\}$

Thus the canonical link is

$$\theta = \log(\mu) = X'\beta \ldots\ldots\ldots\ldots\ldots (6)$$

Equation (6) is called the log-linear model.

If time or space or size is a concern and we use a rate as $\frac{\mu}{t}$ instead of μ (6) becomes

$$\log({}^{\mu}/{}_{t}) = X'\beta - \log(t) \ldots\ldots\ldots(7)$$

where $\log(t)$ is called an offset.

Equation (7) is then termed an offset Poisson regression model.

GAMMA DISTRIBUTION

This distribution models the waiting time between Poisson distributed events. The probability of waiting time until the *nth* poisson (λ) event is a gamma probability density expressed as

$$P(x) = {\lambda(\lambda x)^{n-1}}/{(n-1)!}\, e^{-\lambda x}$$

where λ is the rate at which time changes.

If $X \sim$ Gamma $(K; \theta)$ the pdf

Becomes $P(x) = {x^{k-1}e^{-x/\theta}}/{\Gamma(x)\theta^k}$

for $x > 0,\ k,\ \theta > 0$

k = Shape parameter = no of occurrences of an event

θ = Scale parameter

By letting $\alpha = k$ (shape parameter) and $\beta = {}^{1}/{}_{\theta}$ (rate par

The pdf when $X \sim \Gamma(\alpha, \beta)$

becomes $(x, \alpha, \beta) = {\beta^\alpha X^{\alpha-1}e^{-X\beta}}/{\Gamma(x)}$

for $x \geq 0$ & $\alpha, \beta > 0$

$= {X^{\alpha-1}e^{-X/\beta}}/{\beta^\alpha\Gamma(\alpha)}$, for $x > 0,\ \alpha > 0,\ \beta > 0$

If $\beta = {}^{1}/{}_{\alpha}$

α =Poisson rate parameter

$g(x,\alpha,\beta) = 0$ elsewhere

$$g(x,\alpha,\beta) = logX^{\alpha-1}e^{-\alpha/\beta} - log\beta^{\alpha}\int_0^{\infty} y^{\alpha-1}e^{-y}dy$$

$$g(x,\alpha,\beta) = exp\{logf(x,\alpha,\beta)\}$$

$$= exp\left\{-\,{}^{1}/_{\beta}X - \alpha log\beta + log\Gamma(\alpha) + (\alpha-1)logx\right\}$$

The canonical link

$$\theta = {}^{-1}/_{\beta}$$

$$-\,{}^{1}/_{\beta} = E(X) = X'\beta$$

$$\text{But } \beta = {}^{1}/_{\alpha}$$

$$\Rightarrow -\frac{1}{\beta} = -\alpha$$

but $-\,{}^{1}/_{\beta} = \alpha'\beta$

$$-\alpha = \alpha'\beta$$

Thus Gamma regression for rates yields an offset Gamma regression as

$$^{\alpha}/_{t} = \alpha'\beta$$

$$-\alpha = t\alpha'\beta$$

t = time or size or shape.

POISSON-GAMMA MIXTURE DISTRIBUTION

The variance of the Poisson-Gamma mixture distribution differs from that of Poisson by

$var(X) = \mu + \mu^2 k$

The mean, $E(X) = \mu$ (same as poisson)

The distribution is also termed, Negative Binomial. This is because the negative binomial distribution is a mixture of a family of poisson distributions with gamma mixing weights. In this case, the poisson parameter is a random variable distributed as gamma.

The negetive binomial distribution with r number of failures, k number of success and having the probability of success p is given as

$$f(k;r;p) = P(X=x) = \binom{k-1}{k-r}(1-p)^r p^{k-r}, \quad k = r, r+1, \dots$$

By letting $P \to 0$ or $r \to \infty$, the Negative Binomial tends to the poisson distribution.

where the stopping parameter $r \to 0$ in a Negative Binomial distribution and the probability of success in each trial $p \to 0$, the mean parameter λ, is kept constant

The parameter P becomes $P = {}^{\lambda}/_{\lambda + r}$, $\Rightarrow \lambda = {}^{rp}/_{1-p}$

Under this parameterization,

The $pmf = f(k;r;p) = {}^{\Gamma(K+V)}/_{K!\,\Gamma(r)}\,P^K(1-P)^r$

$$= {}^{\lambda^K}/_{K!}\,{}^{\Gamma(r+K)}/_{\Gamma(r)(r+\lambda)^K} \cdot {}^{1}/_{(1+{}^{\lambda}/_{r})^r} \quad \dots (9)$$

where K is the number of successes in a sequence of independently and identically distributed Bernoulli trials. r is the number of failures.

$$\lim_{r\to\infty} f(K;r;p) = {\lambda^K}/{K!}\, e^{-\lambda}$$

λ = expected value or mean

Thus the poisson distribution is a limiting Negative Binomial distribution.

The Poisson-Gamma mixture $pmf,\ f(k;r;p)$ can be shown to be equal to that of the negative binomial as follows:

$$f(k;r;p) = \int_0^\infty f_{poisson(\lambda)}(K)\ .\ f_{Gamma\left(r,{1-p}/{p}\right)}(\lambda) d\lambda$$

$$= \int_0^\infty \frac{\lambda^K}{K!} e^{-\lambda}\ .\lambda^{r-1} e^{-\lambda^{(1-p)}/p} \Big/ \left(\frac{p}{1-p}\right)^r \Gamma(r)\ d\lambda$$

$$= {(1-p)^r p^{-r}}\Big/{K!\,\Gamma(r)} \int_0^\infty \lambda^{r+k-1} e^{-\lambda/p} d\lambda$$

$$= {(1-p)^r}\Big/{K!\,\Gamma(r)}\, p^{r+k}\Gamma(r+k)$$

$$= {\Gamma(r+k)}\Big/{K!\,\Gamma(r)}\ p^k(1-p)^r$$

Hence the Negative Binomial distribution is also the poisson-Gamma mixture distribution.

Under the parameterization$k\alpha$, $E(Y) = \mu = k\alpha$

$var\ (Y) = K\alpha + K\alpha^2 = \mu + {\mu^2}/{K}$

The probability mass function $P(Y = y; \alpha; k)$ becomes

$$P(Y = X; \alpha; K) = {(X+K-1)!}\Big/{X!\,(K-1)!}\ {\alpha^X}\Big/{(1+\alpha)^{X+K}}\ ,\quad X = 0,1,2,\ldots$$

by which we obtain the loglikelihood as

$$l = Xlog\left\{{\alpha}/{(1+\alpha)}\right\} - Klog(1+\alpha) + log\binom{X-1}{K-1}$$

The canonical link function is

$$\eta = \log\left({\alpha}/{1+\alpha}\right) = log\left({\mu}/{\mu+k}\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots(10)$$

Thus the generalized linear regression model arising from (10) is

$$log\left({\mu}/{\mu+k}\right) = \alpha + X'\beta$$

for a fixed K.

Using another parameterization with

$$f(x;k;p) = \binom{x-1}{k-1} p^k (1-p)^{x-k} \qquad x = k, k+1, k+2, \ldots$$

where p = probability of success

k = no of successes before rth failure

$$logf = klog\binom{x-1}{k-1} + klog\left({}^{p}/_{1-p}\right) + xlog(1-p)$$

$$expf = exp\left\{xlog(-p) + klog\left({}^{p}/_{1-p}\right) + log\binom{x-1}{k-1}\right\}$$

Canonical link, $g(p) = \log(1-p)$

$$= \log q$$

$$= X\beta$$

The offset model is achieved as

$$log\left({}^{q}/_{t}\right) = \log(q) - \log(t) = X\beta$$

$$\Rightarrow \qquad \log(q) = X\beta + \log(t)$$

q = probability that a person mates with an infected person

$t = {}^{1}/_{2}$ month

ESTIMATION IN POISSON-GAMMA MIXTURE MODEL

The poisson-Gamma mixture (Negative Binomial) X is a non-negative discrete random variable with probability mass function

$$P(X=x) = \begin{cases} \left({}^{r}/_{r+m}\right)^r \Gamma(r+k) \Big/ k!\,\Gamma(r) \left({}^{r}/_{r+m}\right)^k & m, r > 0 \\ 0 & elsewhere \end{cases}, X = 0, 1, 2, \ldots$$

m = mean = location parameter

r = dispersion parameter = heterogeneity

The variance $\sigma^2 = m + {}^{m^2}/_{r}$

when r is known, the Negative binomial distribution with parameter m becomes a member of the exponential family.

$\bar{x}$ = is a minimum variance unbiased estimator for m Al-Khasawneh, (2010).

For the two-parameter poisson-Gamma mixture distribution, both m and r are unknown. This situation is more practical.

ESTIMATION OF THE DISPERSION PARAMETER r

Both the method of the moments and maximum likelihood method individually impose constraints on the estimation of the dispersion parameter. Al-Khasawneh, (2010) combines both the method of moments and maximum quasi-likelihood estimation in a variety of ways using appropriate weights.

Method of moments

This is done by solving simultaneously, the following equations:

$$\widehat{m} = \bar{x} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (11)$$

$$s^2 = m + {}^{m^2}/_{r} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (12)$$

Thus $\hat{r} = {}^{\bar{x}^2}/_{s^2 - \bar{y}}$ ……………………………………….(13)

From (13) it is obvious that if the sample variance s^2 equals the sample mean, $\hat{r}$ cannot be determined.

$$\hat{\mu}_{MLE} = \bar{X}$$

$\hat{r}_{MLE\hat{\varphi}}$ is solution to

$$nln\left(1 + \bar{x}/_r\right) = n_1\left(1/_r\right) + n_2\left(1/_r + 1/_{r+1}\right) + n_3\left(1/_r + 1/_{r+1} + 1/_{r+2}\right) + \cdots \ldots\ldots\ldots\ldots\ldots\ldots (14)$$

Where n is the size, n_1 is the number of ones in the sample, n_2 is the number of twos in the sample and so on.

Equation (14) can be written as

$$\sum\sum 1/_{r+t} + nln\, r/_{r+\bar{X}} = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (15)$$

Equation (15) is solved by $IWLS$ or Newton-Raphson method to obtain $\hat{r}_{MLE}$.

$\hat{r}_{MLE}$ exist only for over-dispersed samples Levin and Reeds, (1977)

The re-parameterization of $\alpha = 1/_r$ was suggested by Piegorsch, (1990) and Anraku and Yanagimoto (1990), where α was estimated by Maximum Likelihood.

Piegorsch, (1990) obtained $\hat{\alpha}_{MLE}$ as follows:

$$P(X = x) = \begin{cases} \Gamma(x + \alpha^{-1})/_{X!\,\Gamma(\alpha^{-1})} \left(\alpha m/_{1+\alpha m}\right)^X (1 + \alpha m)^{-1/\alpha}, & m, \alpha > 0,\ x = 0,1,2, \ldots \\ 0 & elsewhere \end{cases}$$

$\ldots\ldots\ldots\ldots\ldots$(16)

$$\partial l/_{\partial m} = \sum_{i=1}^{n} \left(X_i/_m - 1 + \alpha X_i/_{1+\alpha m}\right) = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots (17)$$

$$\partial l/_{\partial \alpha} = \sum_{i=1}^{n} \left[\sum_{v=0}^{X_i - 1} 1/_{\alpha(1+\alpha v)} - n/_{\alpha^{-2}} \log(1 + \alpha m) + m\mu(\bar{X} + \alpha^{-1})/_{1+\alpha m} \right] = 0 \ldots\ldots\ldots (18)$$

$$m_{MLE} = \bar{Y}$$

Solving (18) at $m = \hat{m}$ yields $\hat{\alpha}_{MLE}$

THE OFFSET POISSON-GAMMA MIXTURE MODEL

The canonical link function of the Poisson-Gamma mixture regression model can be expressed as

$$\eta = log(1 - p) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (19)$$

Where $p = m/_{m+r}$ is the probability of success in each trial

$$\text{or} \quad \eta = log\left(m/_{m+k}\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (20)$$

where m is $E(X)$ = mean.

Equation (19) derives from a probability mass function defined as

$$f(k; r; p) = \Gamma(r + k)/_{k!\,\Gamma(r)}\, p^k (1 - p)^r$$

while (20) is derived from a pmf version of the Poisson-Gamma model expressed as

$$f(x; \alpha; k) = (X + K - 1)!/_{X!\,(K - 1!)}\, \alpha^X/_{(1+\alpha)^{X+K}}; \qquad X = 0,1,2, \ldots$$

where $E(X) = m = K\alpha =$ mean

$$Var(X) = K\alpha + K\alpha^2 = m + m^2/_K$$

It can be shown that expression (19) & (20) are equivalent.

Using equation (20), a generalized linear model having the Poisson-Gamma mixture model is expressed as

$$\eta = log\left({}^{m}/_{m+K}\right) = X\beta$$

Note that ${}^{m}/_{m+K} = p$

$$\Rightarrow \log(m) - \log(m+K) = X\beta$$

$$\log(m) = X\beta + \log(m+K) \dots\dots\dots\dots\dots\dots\dots\dots.(21)$$

$\log(m+K)$ is an off-set for the Poisson-Gamma mixture regression model.

Under the condition that $m + K = 1$, the Poisson-Gamma mixture model (21) reduces to a Poisson (m) loglinear model.

$$\hat{\beta}_{MQLE} = (X'WX)^{-'}X'WZ$$

Where W is a diagonal weight matrix with Poisson weights.

$w = {}^{m}/_{\phi}$, $\quad \boldsymbol{\phi} > 1 \quad$ (ϕ is over-dispersion parameter)

β estimates are Maximum Quasi Likelihood Estimates (MQLE).

DETECTING OVERDISPERSION IN LOGISTIC REGRESSION

Over-dispersion occurs when the observed variance is larger than the expected variance. Under the assumption that the logistic model is correct $D \sim \psi^2_{n-p}$. Collet (2003)

Thus Collet,(2003) $D > n - p = E(\psi^2_{n-p})$ can indicate over-dispersion, even though the condition $D > n - p$ can be caused by other reasons such as wrong link function, existence of large outliers, etc.

Illustration

For an HIV/AIDS study

We assume a community where multiple sex partners live. Let the meeting with a non-HIV/AIDS member of a community by a multiple sex partner be counted as a success and the meeting with a HIV/AIDS member be counted as a failure.

Let p represent the proportion of non-HIV/AIDS.

Let k = no of times a multiple sex member of the community meets non-HIV/AIDS members before meeting with the rth infected person.

Then r = no of times the multiple members meets with infected persons.

k is likely to be higher for higher populations.

$\therefore k$ is a function of the population (size). r is also going to be higher with higher populations.

Let $\alpha = {}^{1}/_{r}$, then α is a measure of Poisson over-dispersion. Recall that for the Poisson-Gamma mixture distribution, $var(X) = m + {}^{m^2}/_{r} = m + \alpha m^2$ thus as $r \to \infty$, over-dispersion$\to 0$ and the Poisson-Gamma mixture distribution tends to the Poisson distribution.

Thus in large populations the Poisson-Gamma mixture distribution will be approximated by the Poisson distribution.

CONCLUSION

The offset Poisson-Gamma mixture model has been formulated to show that the Poisson-Gamma mixture link is actually the sum of the Poisson loglinear link and an offset.

Thus modeling a Poisson-Gamma mixture regression model amounts to modeling a Poisson loglinear model and thereafter adding the offset term. It is also shown in this study that by using the fact that the Poisson overdispersion measure is a reciprocal of the number of failures, increasing the sample size reduces Poisson overdispersion.

REFERENCES

Al-Khasawneh, M. F. (2010). Estimating the Negative Binomial Dispersion Parameter. *Asian J. Maths Stat.*, 3, 1-5.

Anraku, K. & Yanagimoto, S. (1990). Estimation for the negative binomial distribution based on the conditional likelihood. Comm. Statist. *Simul. Comput*. 19: 771-786.

Collet, D. (2003). Modelling binary data (2nd edition). London: Chapman and Hall

Hilbe, J. M. (2007). Negative Binomial Regression. 2nd edition. Cambridge University Press. In press
Levin, B. & Reeds, J. (1977). Compound Multinomial likelihood functions are unimodal: Proof of a conjecture of I. *J. Good. Ann. Statist*; 5, 79-87.

Piegorsch, W. W., (1990). Maximum Likelihood estimation for negative binomial dispersion parameter . *Biometrics*, 46, 863-867.

Williams, D. (1982). Extra binomial variation in logistic linear models. *Applied statistics*, 31, 144-148.

EXAMINATION MANAGEMENT PORTAL FOR TERTIARY INSTITUTIONS: A CASE STUDY OF FEDERAL UNIVERSITY LAFIA

3

Jeffrey O. Agushaka., Absalom E. Ezugwu., Paulinus O. Ofem, Emmanuel O. Adetula, Christopher I. Eke & Daniel I. Yakmut

Department of Computer Science, Faculty of Science, Federal University Lafia
Corresponding author's email: agushaka@yahoo.com

Date manuscript Received: 12/07/15 Accepted: 18/12/15 Published: December 2015

ABSTRACT

A study was carried out on the analysis, design and implementation of a Web-based examination management system in Federal University Lafia. A step by step data flow diagram was created as a guide for application developers towards building a more flexible student information system for the university. Using the design, enhanced and intelligent computer software for result computation, integrated with a dedicated database for storing processed results and other relevant records was developed to simplify the University grading system and overcome the short-comings of existing examination packages. An empirical evaluation and comparison of the system with other similar systems shows that it will equally expedite the processing of students' records, semester results and transcripts at various levels and offer a faster, secured management and access to students' results On-line. This system is more efficient and provides substantial benefits for both the University management and students as observed in its implementation.

Keywords: *Examination Management System, Computer Software and Computation.*

FULAFIA JOURNAL OF SCIENCE AND TECHNOLOGY | VOL. 1

INTRODUCTION

Paper examinations still remain the primary means by which student learning in both undergraduate and postgraduate curricula are assessed, as they have unique advantages that may never be replaced in the nearest future. However, paper examinations are laborious to grade, tally, and record. They are also difficult to keep, especially after they are returned to the students. Often, electronic examination is considered a suitable alternative (grading, returning exams, etc). Despite this, more challenges still exist when administering electronic exams, including ensuring that everybody has a secure computer for the examination. Hence, the prevalence of paper examination is expected to continue for quite some time. However, examinations, results and the personal information about students are the key elements of an electronic student information management system. Given the various needs of the different categories of students in a university setting, each student record will need to be processed from time to time in order to meet those needs and the needs of other people outside the university who rely on such records. Some institutions still rely on manual systems for the archiving of students records thereby making its processing and retrieval cumbersome. According to Arekete and Osinowo (2009), student records in other cases are available in semi-electronic or computerized forms which still have the aforementioned inherent difficulties. These electronic formats are unfriendly because they require further cumbersome transformation processes in order to produce the much needed reports. In this current digital era, time is always of essence. There is therefore need to create systems that are seen to be efficient, flexible, reliable and able to produce accurate information as at when required.

Student Information Management Systems (SIMSs) have gone through tremendous changes with respect to technology. According to USDEOPEP (2009), since the early 1990s, the Web has become the technological infrastructure that has enabled the deployment of innovative SIMSs for tertiary institutions. Idogho *et al.*, (2011) reported the creation of an "interactive intranet portal for effective management in tertiary institution". Their work addressed problems that emanate from the processing of student results, payment of fees and management of library resources. They identified services that are relevant to the institution and sought to harmonize them into a single platform. However, their work discussed only two of such services which include result processing and library resource management. Also in Youh (2010), a "client server distributed database for student result processing" was also reported. Their work focused on allowing various academic units to maintain and control their data. They placed emphasis on the benefits of leveraging a distributed architecture as opposed to centralized database architecture. In another related study, Ayodele and Ezugwu (2010) reported the "design and implementation of student's information system for tertiary institutions using neural networks: an open source approach". Just like the previous researches, their work focused on the collection and storage of student records and a prompt processing of results and student transcripts and the end result being the availability of students result online.

This study compared the other similar systems and aims to expedite the processing of students' records, semester results and transcripts as well as offering a faster more secured access to results online.

MATERIALS AND METHODS

The best suitable alternative solution to resolving the stated problem is the creation of an Internet based application. The internet has become a melting pot for the business community, with web application as the key mechanism for success and growth. The choice of adopting web application lies solemnly on some of its important qualities and attributes among which include remote distributive nature (processing occurs at different physical locations), reliability, availability, interoperability and security.

SYSTEM DESIGN METHODOLOGY

Schedule Feasibility

Time evaluation is the most important consideration in the development of project. The time schedule that is required for the development of this project is very important since more development time affects machine time and cost that can cause delay in the development of other systems. A reliable Exam Management System can be developed in the considerable amount of time if scheduled feasibility is properly documented and followed in the cause of the design work. The figures covered in this section are the initial information gathering, which were later projected for a better design model of the proposed system implementation.

System Data Flow Models

The aim of designing a data flow diagram at the early stage of the system planning and implementation is to enable the developer have a wider perspective about the information and functional domain of the overall system functionalities. In the cause of this design process, the data flow diagram is usually refined to achieve a greater level of detailed information that encompasses the running application to be developed by the software designer. The Web-based Exam Management System data flow diagram is splinted into three levels; ranging from level 0 to level 3 as shown in Figures 1, 2, 3 and 4. The input entities involve Admin (exam officer and MIS staff), Instructors and Students, while the input data objects include: students' profile, course information and examination records. The labelled lines and arrows represent data type hierarchies

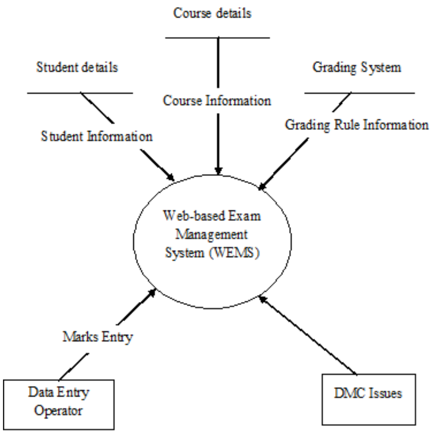


Figure 1: Proposed Exam Management System DFD Level 0

The figure above serves as the core or heart from which we build a fully functional system, with diverse functionalities. Our focus is to develop examination processing engine that will be deployed or migrated across universities, polytechnics and colleges of education. The running of such system should be platform independent and customizable to the taste of end users.

The system user comprises both the administrators, instructors and students interface/module, and each of these users will have different degree or level of right of assessments, depending on their respective roles. We will subsequently provide a detailed description of the running system.

In Figure 3, Examination Card denote a printed or hardcopy identification card containing the list of all courses registered by a student, and on it is also attached the bearer passport photograph. This might differ for some institutions, but WEMS offers a provision for generating examination identification card automatically for individual students.

Figure 6 shows the connection between students, courses and instructors. A connection is established between students, course and instructor because the three are related. This relationship can further be understood by having knowledge of the role of students, courses, and instructors within the context of the proposed system to be implemented. We choose to define a set of object/relationship pairs that defines the relevant relationships. For example, a student registers a course, receives lectures for the course, at the end, he or she takes an examination. The course is taught by an instructor who sets exams and

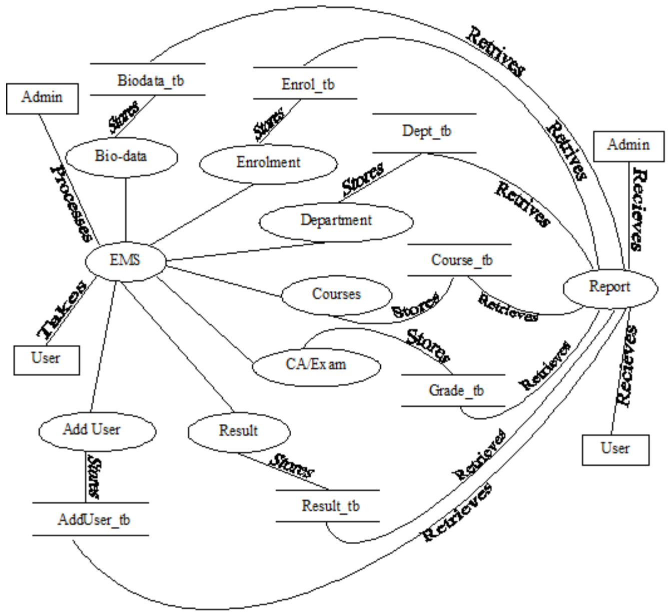


Figure 2: Exam Management System DFD Level 1

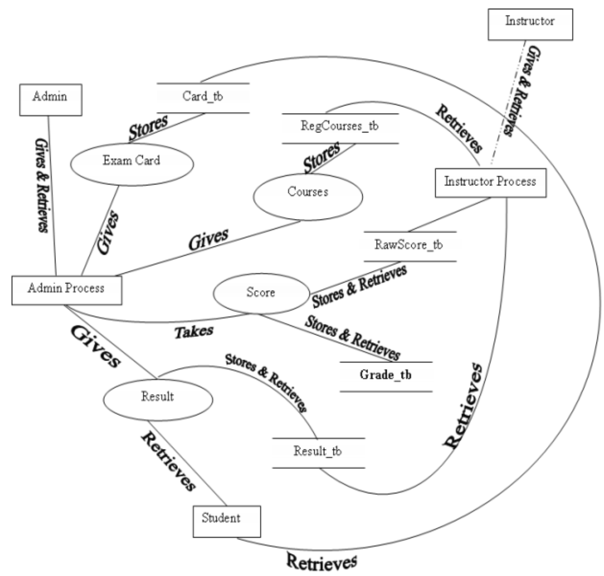
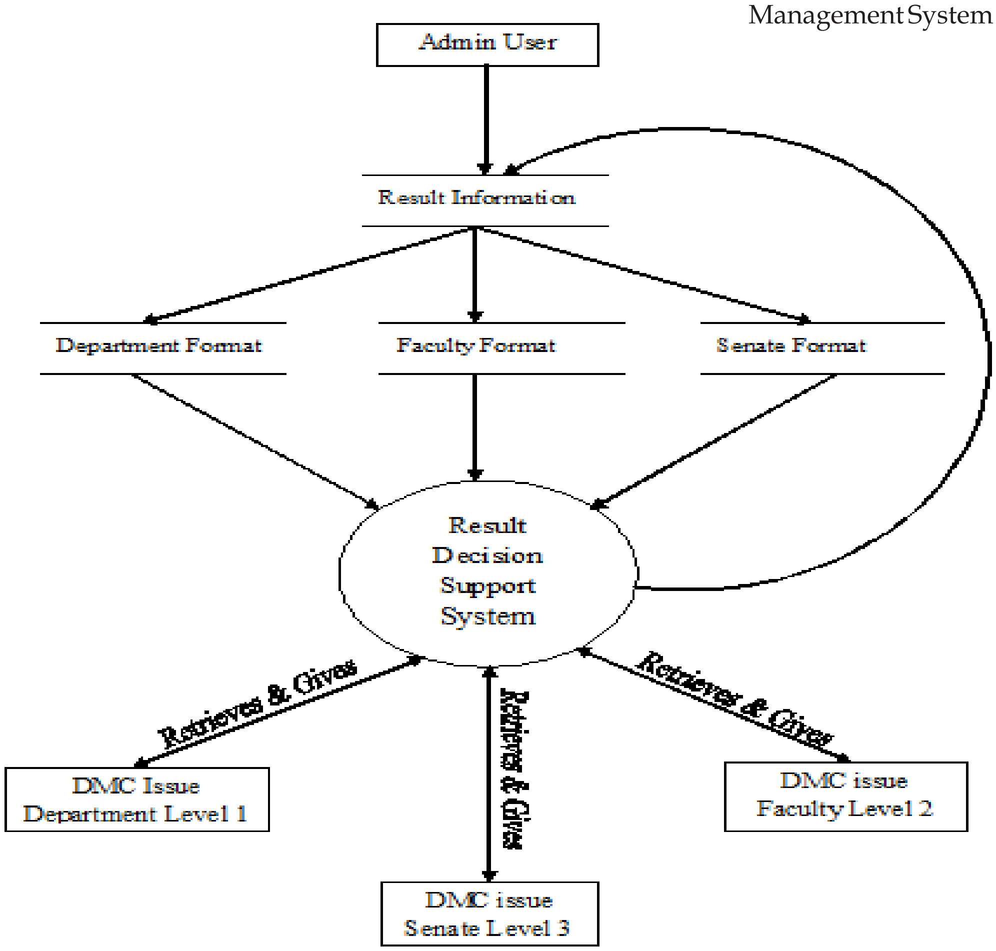


Figure 3: Proposed Exam Management System DFD Level 2

Figure 4: Proposed Examination Management System DFD Level 3



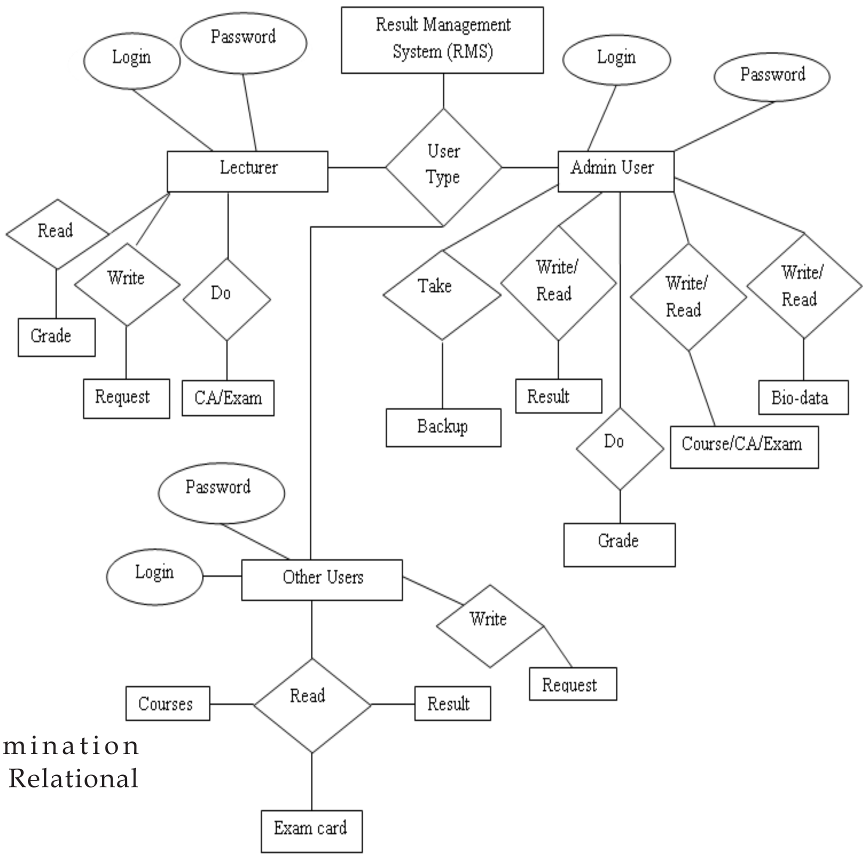


Figure 5: Proposed Examination Management System Entity Relational Diagram

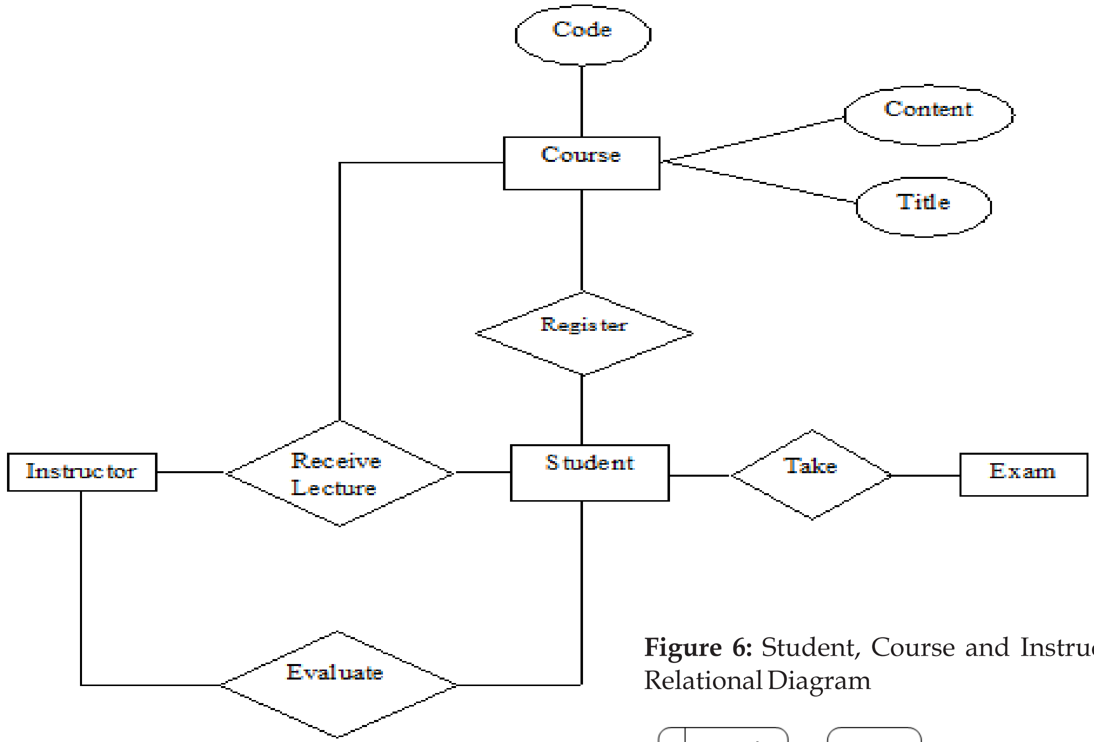


Figure 6: Student, Course and Instructor Entity Relational Diagram

System Design for Project Development

Fig. 7 is a task diagram depicting the work (= tasks) that should be performed to complete the design phase.

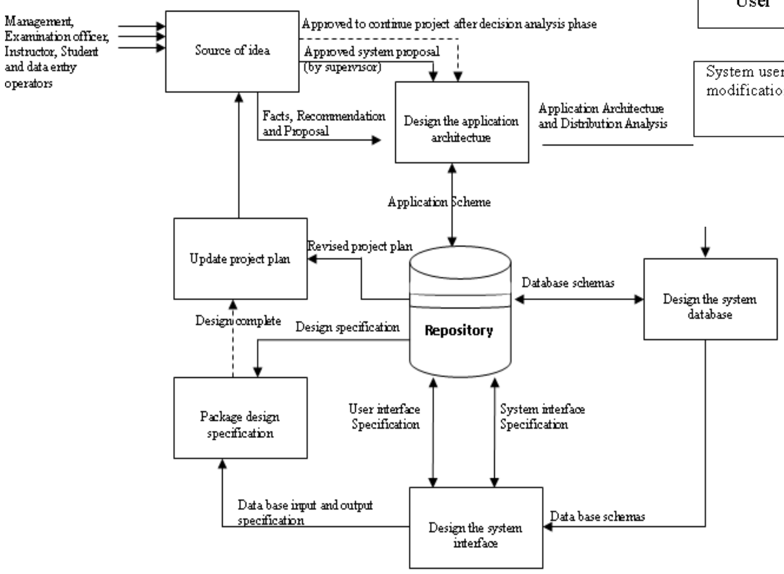


Figure 7: The System Design Task for the Software Project Development Life Cycle

Design of the Application Architecture

Given the data-models, process models and target solution, distribution decisions will need to be made. Figure 8 is a physical data flow diagram (PDFD) that is used to establish physical processes and data stores (databases) across a network of the proposed design.

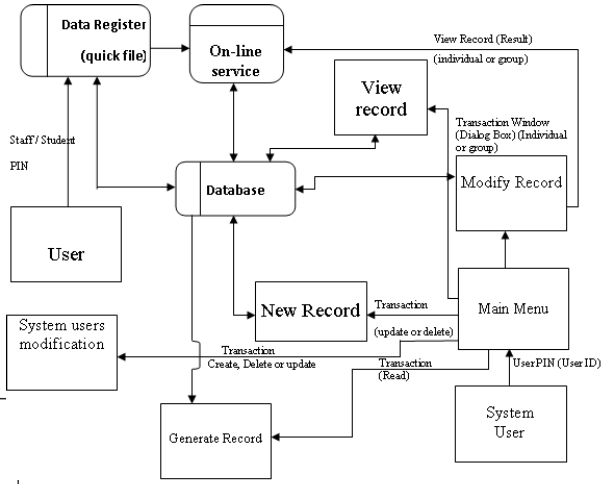


Figure 8: Application Architecture; Physical Data Flow Diagram

Design of the Database(s)

The resulting database schemas shown in Table 1, will serve as container were data are stored as records in tables on the server side database. It consists of a user table (login), Bio-data table, enrolment table, course table, grade table, department table etc. Each table contains a setup information data whose details are provided in the data tables that are shown on the next page:

Table 1: Sample course table from a database.

urse_id	course_code	course_title	credit_unit	course_type	semester	level
1	csc 111	Introduction to Computer System	2	core	1	1
2	csc 112	Computer Programming 1(BASIC)	2	core	1	1
3	csc 113	Computer Operation 1	2	core	1	1
4	csc 114	Data Structure I	2	core	1	1
5	csc 121	Electronic Data Processing	2	core	2	1
6	csc 122	Computer Programming II (FOTRAN)	2	core	2	1
7	csc 123	Computer Architecture	2	core	2	1
8	csc 124	The Teaching of Computer Science I	2	core	2	1
9	csc 211	Introduction to Symbolic Logic	2	elective	1	2
10	csc 212	Introduction to Operations Research	2	elective	1	2
11	csc 213	Database Management I (with Dbase III +)	2	core	1	2
12	csc 214	Data Sturcture II	2	core	1	2
13	csc 215	Word Processing (with WP 8 or higher or WS5)	2	core	1	2
14	csc2	Spreedsheet with Lotus 1-2-3/Excel	1	elective	2	2
15	csc 222	Computer Based System Application	1	elective	2	2
16	csc 223	The teaching of computer Science II	2	core	2	2
17	csc 225	Numerical Methods	1	core	2	2
18	csc 311	Operating System (os)	1	core	1	3
19	csc 312	The Microprocessor	1	core	1	3

Data Tables

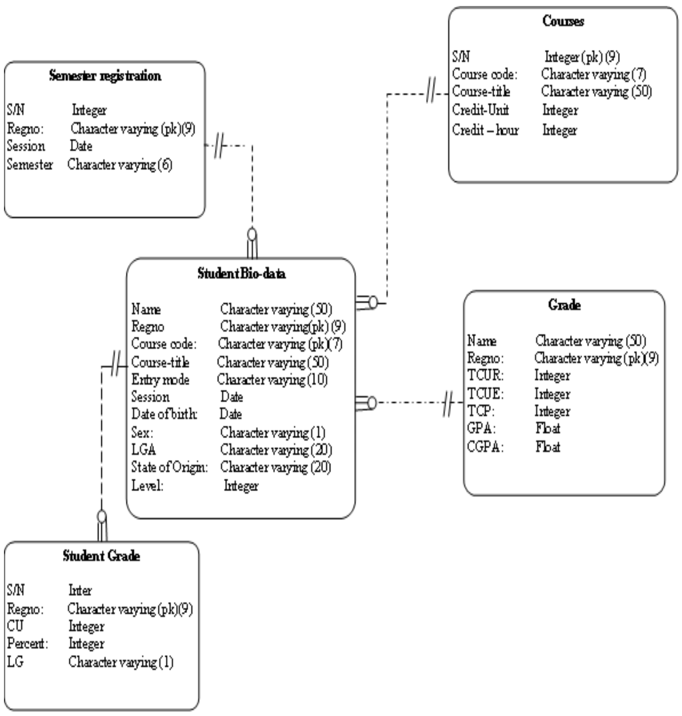


Figure 9: Physical Database Schema

System Design

From the Data Flow Diagram (DFD) depicted earlier on, the next step is the definition of the modules and their relationship to one another in a form called a structured chart, using a data dictionary and other structured tools. The design specifications that get generated at the end of this phase are technical in nature and contain:

i. User interfaces
ii. Databases and data structures
iii. Algorithms and program structures
iv. Equipment and other facilities required
v. Manual procedures that will be part of the implemented system.

Mathematical Model

The most important aspect of the new system is the grading module (mark information management system); the required parameters are basically the semester registration, course registration, assignment, and examination raw scores. The course credit unit registered is denoted by CC. The semester is broken into two; first and second semester and likewise those courses taken at the respective periods for first semester are usually denoted by odd number course codes, (MATH101, COSC201, STAT305, etc) while those for second semester are given even number course codes (MATH104, COSC206, STAT304, etc.) respectively. A student is expected to register for a minimum of 48 credit unit in 100 levels. Thus, the total number of credit unit expected at the end of the first year is given by:

$$\sum_{i=1}^{N} CC_i \leq 24 \qquad ....(1)$$

For first semester cumulative

$$\sum_{j=1}^{N} CC_j \leq 24 \qquad ....(2)$$

Second semester cumulative

$$\sum_{i=1}^{N} CC_i + \sum_{j=1}^{N} CC_j \leq 48 \qquad (3)$$

Where i, j = 1, 2, 3, …, N and N represents the total number of courses registered by a student per semester.

The credit unit point and grade point average are then computed based on the above perimeter as follows:

$$TCP = \sum_{i}^{N} CG_i \times CC_i \qquad (4)$$

$$GPA = \frac{\sum_{i=1}^{N} CG_i \times CC_i}{\sum_{i=1}^{N} CC_i} \qquad (5)$$

Where CC_i the course's credits, N is the total number of courses; CG_i is the course grade point of the course credit system evaluation policy.

The mathematical models which yield equations (4) and (5) will be used in the mark information management module, to compute grades. Subsequently, the processed marks information will be made available for transcript processing and semester result generation.

RESULTS AND DISCUSSION

The technological approach for the development of the new system is based on WAMP Server (Apache, MySQL, and PHP) open source solution Adewale (2006). The web server has what seems to be a simple straightforward job. It sits there over a network, running on top of the client's machine listening to requests that somebody on the web might make, and it responds to those requests and serves out the appropriate web pages. In reality, it is a bit more complicated than that, and because of the web that runs twenty four hours a day, seven days a week, the stability of the web server is a major issue. PHP and MySQL are popular pair for building dynamic web applications. PHP is the most widely supported and used web scripting language, and an excellent tool for building web database applications (Williams and Lane, 2004). MySQL is a client/server database that consists of a multithreaded SQL server that supports different back ends, several different client programs and libraries, administrative tools, and a wide range of programming interfaces Connolly & Begg (2002). It can hold up to 60,000 database tables with approximately 5 billion records).

The new system is built around a three-tier architecture model, fourth-tier could also be considered as a suitable alternative, depending on the availability of hardware/preference (see figure 10). At the base of the application is the database tier, consisting of the relational database management system that manages the data that users create, query and delete. This database tier is implemented using MySQL database server. Built on top of the database tier is the middle tier, which contains most of the application logics that have been developed using PHP as the scripting engine. This middleware works closely with the web server to interpret the request made from the World Wide Web, process this requests, interact with other programs on the server to fulfil the request, and then indicate to the web server exactly what to serve to the client browser. At the user end is the client tier, usually, a browser software (Internet Explorer, Mozilla Firefox, Opera or Safari) that interact with the application.

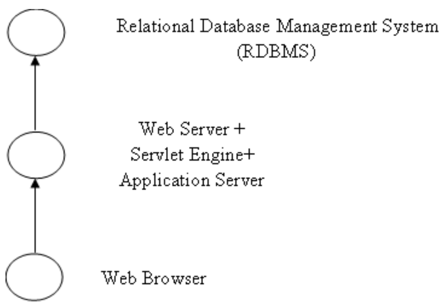


Figure 10: Three-tier architecture model

System Requirements

Table 2: System Requirements

Hardware Requirements		
Processor	*R*	*Disk Space*
Pentium III, Pentium IV or Higher	512 Mb or Higher	250
Software		
Operating System	*Da*	
Win-2007, Win-XP, Linux or any other higher version	My5	
Server		
Server Platform	*Bro*	
Apache WAMP Server	Internet Explorer, Firefox	
WAP Server	I	
Network		
Connectivity with University Campus		

System Description

We describe in brief some key roles being performed by the administrative officer, system users and students.

Admin: This module enables the administrator to carry out some administrative task such as addition of new students, enrolments, record modification, computation of result, system update and some basic system maintenance. There can be more than one administrator at a time, whose basic function is to manage the system.

Users: This component of the system carries out the interface or modules for all the users that access the system which includes instructors, student affairs and students. They can always access the system with a common web browser via the internet or local area network connectivity. Each user has an own interface which is always active on the client side of the system. Their respective role is further explained:

- **i.** Instructors interface: This is an interface where all the instructors can manage their own courses and exams. The instructors are assigned the tasks of grading students and submitting their continuous assessments and exams raw scores to the administrative officer in charge of result processing to handle. This module also provides the instructors with the opportunity of knowing the total number of students that registered for their courses.
- **ii.** Student interface: With this interface the students can perform three operations. First, they can use it to print their examination cards online, and second, preview or print out their semester results. Lastly, they can use the interface to report complaints regarding detected faults to the administrative officer in charge of processing.

Creating Student Profile: The bio-data of students are usually to be created during the first semester registration or at the beginning of the course and are not needed to be re-entered again for the rest of the semester a student has. The bio-data captures the basic students' information as required by the school authority which is needed for generating reports and other particulars that are required by other bodies who are involved within the academic cycles. As long as a student is already registered in the university portal, the system automatically imports some of the necessary profile needs such as a student's registration number, department etc, since some of these information are not given at the departmental level.

Student Enrolment: The students' enrolment is a repetitive process for every semester. The user is required to enrol students based on their registration status in the university on-line portal, which might include full time/part time, admission category, and exam details. Most importantly, it is at this level of functionality that the user selects the courses registered by the students.

The enrolment steps might include the following, depending on department choice:

- i. Session for which students are to be enrolled
- ii. Faculty/Department from which the course is taken
- iii. Course status (i.e. core, elective or cognate)
- iv. Exam date
- v. Courses under which students are to be enrolled
- vi. Semester
- vii. Admission category
- viii. Registration number

The examination management system flowchart shown in Figure 11 describes the activities diagram that takes place in the WEMS System. The various interactive scenarios that take place upon successful system login are also presented in chapter five of this in the form of screen shots.

- **a.** The system first checks the user name and password of the system user. If they are not valid, the login interface is re-presented again with a specific login error messages. If the username and password of the user are correct, then the interface presented to the user is based on the user type and access right. The admin user will have a different user interface presented to him as regards what other users will have. When an instructor log on with instructors credential, a different interface is presented and this looks different from that of a student.

- **b. Student Registration:** Figure 14 presents an interactive user friendly form to the admin officer in charge of students' registration in every semester to enter the necessary students' data into the system database.

- **c. Course Registration:** Students are expected to carry out their registration process online. They can login into the system using their registration numbers and a pin code given to them by the departmental examination officer as pass mark. After a successful login, they can select the appropriate course for that particular semester. There are rooms for course registration preview, but modifications of courses registered are not allowed not until after the course registration weeks are over.

- **d. Generate Exam Card:** The system will automatically generate examination cards for students, based on their registration status. The examination identification card will be systematically dispatched to the web portal and made available for student to download in the cause of their examination time table release.

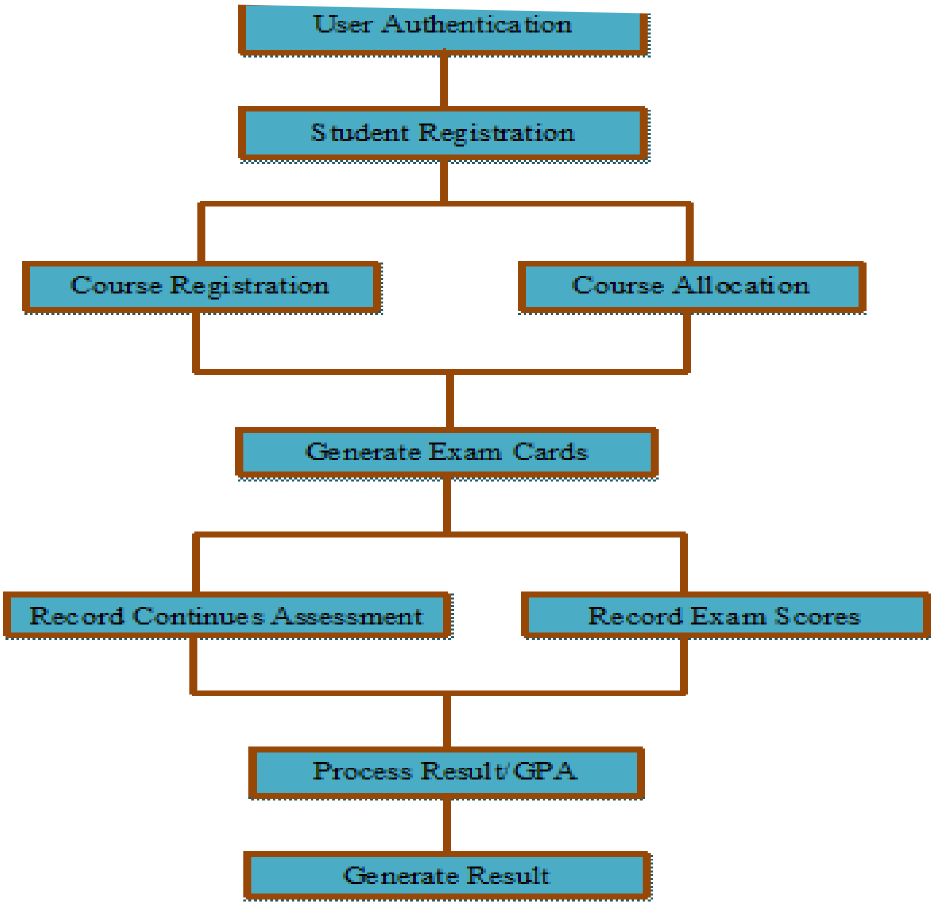


Figure 11: Examination Engine Activity Diagram

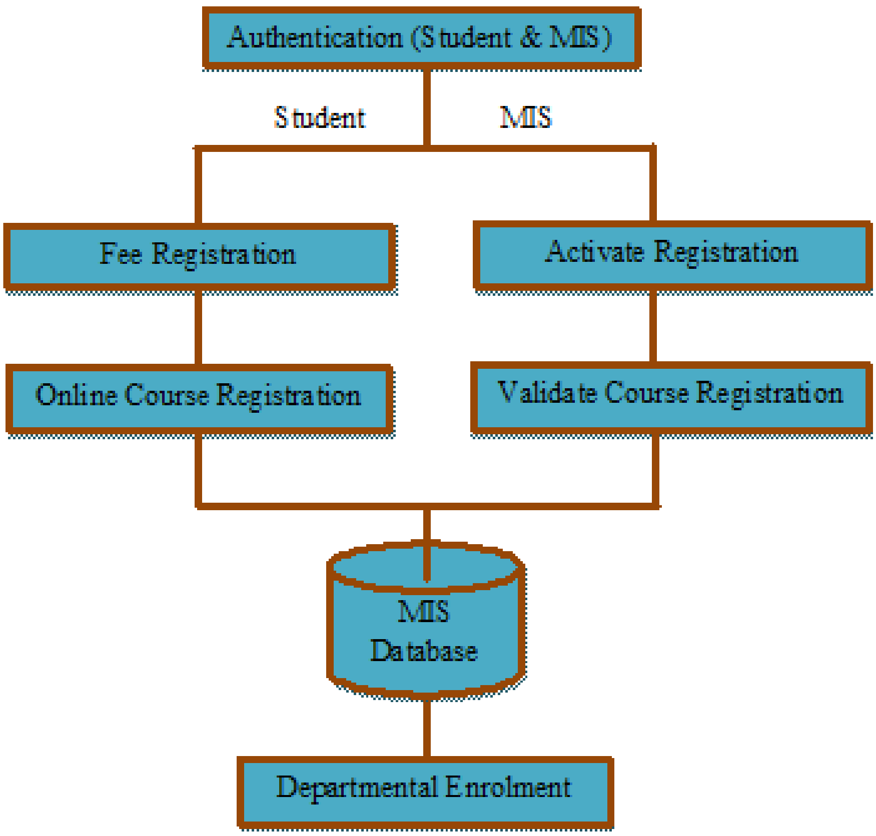


Figure 12: Online Registration Activity Diagram

e. Recording of Exam/CA: It is usually the work of the lecturers to make available their CA/Exam raw score by submitting it into the exam database via the mark-sheet interface meant for that purpose. The processed scores are later retrieved by the exam/admin officer who is in charge of result processing.

f. Process Result: Result processing is performed automatically by the system based on the set conditions inbuilt in the examination package. The system processes result by means of computing the overall grade obtained by students for every course he or she has registered and sat for during the semester. Also among the computed result options include GPA and CGPA. Once the processing is done, it is auto sent to the result database and stored for future report generation.

g. Generate Result: Results are generated in form of reports by querying, and this is often based on request. Students can only access their individual's result online via the university web portal in the form of semester result with their registration numbers, of which they can print out their results for documentation.

The system process assumes that every student should have performed the initial central registration and such registration being validated by MIS as required by the school authority. The departments enrol the students based on retrieved information from MIS central database. Figure 12 does not fall into the range of activities executed by WEMS, but rather it is shown to clarify the initial requirements among various stages that are involved in the process tasks.

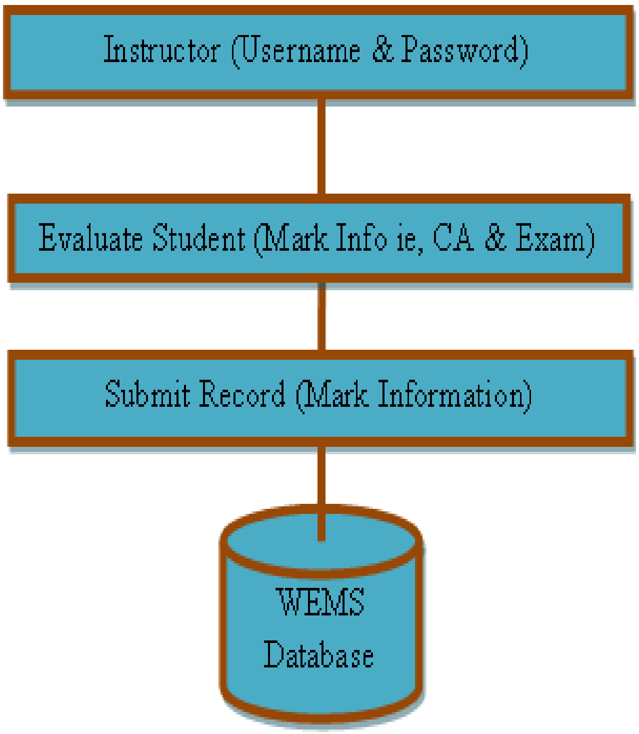


Figure 13: Result Entry Activity Diagram

This activity diagram presented in Figure 13 depicts a module that enables the Instructor to enter students' mark information into WEMS database. The record that is submitted by the instructor will serve as data input which is required by the administrative officer for further processing of the students' final results.

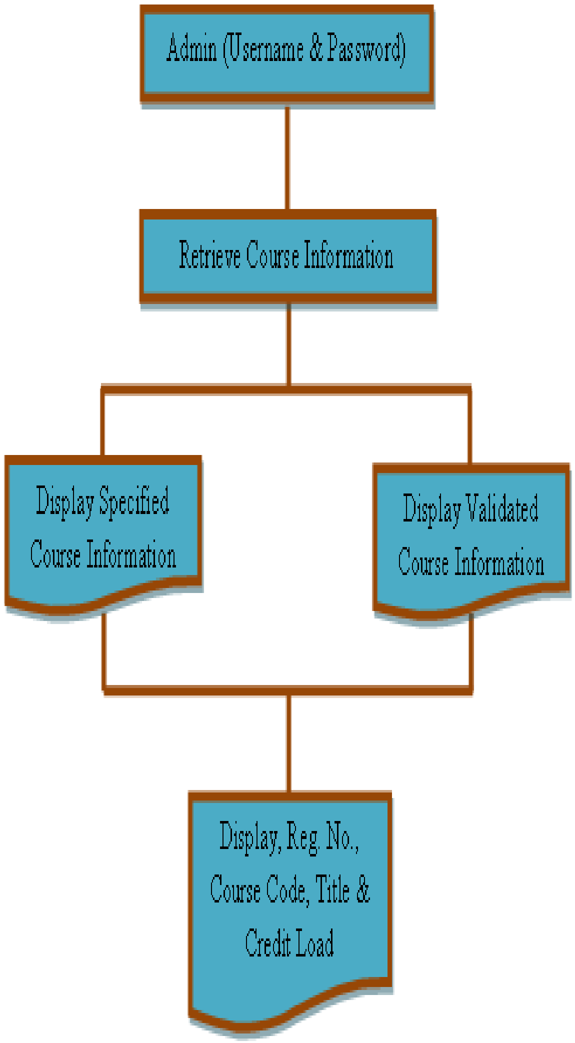


Figure 14: Departmental Course Report Activity Diagram

The departmental administrative officer can use the interface presented in Fig. 4.5 to produce lists of all students that registered for a particular course from within and outside the department.

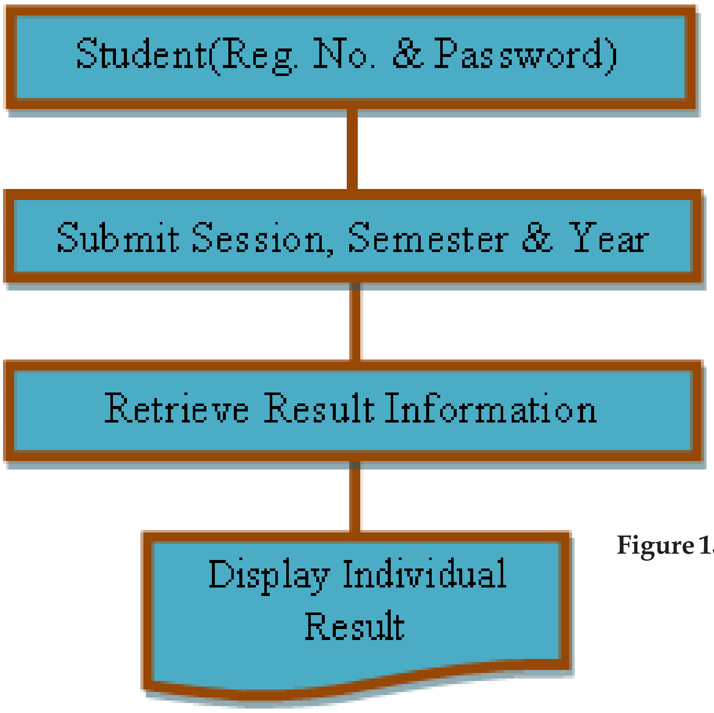


Figure 15: Online Result Checking activity diagram

This presents an activity diagram for students to check their results via online interface that is provided for such purposes. The system first requires the user to provide his/her registration number and password for authentication, after which access is granted or denied. For every successful login, the student sends quires (i.e., session, semester and year via a dropdown combo box). The system displays the students' page with the following information, students' personal data, and all courses and grades information that are obtained by the students for that session.

CONCLUSION

This model has presented a multitier academic management software models for managing students' academic records' in tertiary institutions. The implementation of the proposed system in Federal University Lafia was a success and this could be attributed to the initial models presented for the Web-based Examination System Package. The various detailed data flow and entity relationship diagrams assisted in the development of a mathematical model for WEMS result computation engine which has been implemented on a four-tier architectural model for a Web-based dynamic application, designed using the WAMP technology (Web server, Apache, MySQL and PHP), an open source solution. Because of the sensitive nature of the system, a role-based model access mechanism has been built into the new system to further boost security of the system. The Web-based Exam software serves a dual purposes; It serves as a stand alone student registration software and at the same time dedicated for remote computing/processing of students results from different departments of the university. It also has the features that will enable the user to enter data from any remote location, thereby facilitating the enrolment process for students academic records. WEMS, when integrated into the university portal, will present end users with an interface whereby students can automatically generate and print out their examination cards, check results and also print transcripts remotely.

REFERENCES

Adewale, O. S. (2006). University Digital Libreary. In *An Initiative to Improve Research, Teaching and Services* (pp. 34-63). Akure, Nigeria: Adeyemo Publishing House.

Arekete, S.A., Osinowo, M.O. (2009). "Modeling Rule-Based Information Systems with MS- Excel." *International Journal of Computer Science*, 1 (1), 71-78.

Ayodele O. A. and Ezugwu, A. E. (2010). "Design and Implementation of Students' Information System for Tertiary Institutions using Neural Networks: An Open Source Approach", *International Journal of Green Computing*, 1 (1), pp. 1-15.

Bal-Ilan, J. (2008), "Informetrics at the beginning of the 21st century — A review." *Journal of Informetrics*, pp 1–52.

Benchmark Minimum Academic Standards (Bmas) For. All Nigerian Universities, available at www.nuc.edu.ng/.../APPROVED%20PROGRAMMES%20IN%20SCIENCE

Cinfores, (2012) Web Portal for the University of Port Harcourt.

Connolly, T. M., and Begg, C. E. (2002). *Database System, A Practical Approach to Design, Implementation, and Management* (3rd ed.). Reading, MA: Addison Wesley.

Faculty Orientation Committee. (1996). *Faculty of Science Orientation Manual*. Zaria, Nigeria: Ahmadu Bello University.

Idogho O.P., Akpado K & Agajo J. (2011). "Interactive Intranet Portal for Effective Management in Tertiary Institution", *International Journal of Advanced Computer Science and Applications*, Vol. 2, No. 6, pp. 122-129.

Liu D.T., Xu, X.W., (2001), "A review of web-based data management systems", *Computer in Industries*, pp 251-262.

Qin, S.F., Harrison, R., West, A.A., Jordanovc, I.N., Wright, D.K. (2003) "A framework of web-based conceptual design", *Computers in Industry*, 153–164.

USDEOPEP, (2009). U.S. Department of Education Office of Planning, Evaluation and Policy Development. *Implementing data-informed decision making in schools: Teacher access, supports and use.* United States Department of Education (ERIC Document Reproduction Service No. Ed504191)

Williams, H. E., and Lane, D. (2004). *Web Database Applications with PHP and MySQL* (2nd ed.). Sebastopol, CA: O'Reilly Media.

Ye W. and Jeff O. (2002), *Modelling and Testing Web-based Applications* ISE-IR-02-08. http://www.cs.dal.ca/news/def-1376.htm (Accessed June, 2015).

Youh D.X. (2010). "Design and Implementation of a Client Server Distributed Database for Student Results Processing", *The Pacific Journal of Science and Technology*, Vol. 11, No. 2, pp. 288-295.

ON THE COMBINATORIAL PROPERTIES OF NILPOTENT AND IDEMPOTENT CONJUGACY CLASSES OF THE INJECTIVE ORDER-DECREASING TRANSFORMATION SEMIGROUP

[1]Ugbene, I. J.[*] and [2]Mbah, M. A.
[1]*Department of Mathematics, Federal University of Petroleum Resources, Effurum, Delta State*
[2]*Department of Mathematics, Federal University, Lafia, P.M.B. 146 Nasarawa State.*

**Corresponding author: Email-* ugbene.ifeanyi@fupre.edu.ng
Date Manuscript Received: 17/06/15 Accepted: 18/07/15 Published: December 2015

4

ABSTRACT

The elements of the injective order-decreasing transformation semigroup using path structure (Circuits and Proper paths) introduced by Stephen Lipscomb is presented. The elements were grouped according to their conjugacy class, the nilpotents and the idempotents were identified and the nilpotent and idempotent conjugacy class chain decompositions studied. Some observations were made and their order enumerated. Formulae were found for their order and combinatorial properties.

Keywords: ***Combinatorial properties, Nilpotent conjugacy classes, Idempotent conjugacy classes, Injective Order-decreasing Transformation semigroup***

FULAFIA JOURNAL OF SCIENCE AND TECHNOLOGY I VOL. 1

INTRODUCTION

Semigroup of injective order-decreasing transformations, a subsemigroup of the symmetric inverse semigroup(I_n)was introduced by (Umar, 1992) and defined as thus:

Let α be a transformation in I_n, it is said to be of order-decreasing if $(\forall x \in dom\ \alpha) x\alpha \leq x$. The orderof the injective order-decreasing transformations(ID_n)corresponds to sequence A000110: 1, 2, 5, 15, 52, 203, 877, 4140 … of the online Encyclopedia of integer sequences (OEIS). A formula for its order is given as $B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k$ where B_k is the k-th Bell's number. It is well known that an element $x \in I_n$ is nilpotent $(x^n = 0)$ for some $n > 0$. A property of nilpotent element among others is $x\alpha \neq x\ \forall\ x \in dom\alpha$.

In this paper, we are adopted the path notations invented by (Lipscomb,1996) for I_n which he defined as follows:

Let $\mathbb{N} = x_1, x_2, \ldots, x_m$ and $\alpha \in I_n$ have domain $d\alpha = x_1, x_2, \ldots, x_m$ and if $x_1\alpha = x_2, x_2\alpha = x_3, x_{m-1}\alpha = x_m, x_m = y$. Then α is a path. Having a circuit or a proper path depends on the value of y. If $y = x_1$ then $\alpha = (x_1, x_2, \ldots, x_m)$ is a circuit of length m. If $y \neq x_1$ then $\alpha = (x_1, x_2, \ldots, x_m, y]$ is a proper path of length (m + 1)

Various text or papers have slightly different path notations as can be seen in (Munn, 1957) where he used the notations "links" and "cycles" for proper paths and circuits respectively. He would write $(12)(345] \in I_5$ as (12)[345] where (12) is the cycle while [345] is the link. (Gomes and Howie, 1987) denoted a primitive nilpotent as "||1, 2, …, m||" while (Sullivan, 1987) in his study of semigroups generated by nilpotent transformations denoted a proper path of length (m + 1) as m-chains [1, 2, …, m + 1] and a circuit of length m as m-cycles (1, 2, …, m)

Let G be a group. An element $x \in G$ is said to be conjugate to an element $y \in G$ if there exists $g \in G$ such that $y = g^{-1}xg$. Since conjugation is important to group theory, it was only quite natural to have extended it to some certain classes of semigroups.

Theorem 1

Let $x, y \in I_n$. Then the following holds:

a. x is conjugate to y if and only if they have the same path structure
b. $x$is nilpotent if and only if its path structure are joins of only proper paths.
c. x is idempotent if and only if all the paths in its path structure is of length one

Proof:

a. Let $a = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 5 & 6 & - \end{pmatrix} \in I_6 = (123)(456]$ and
$b = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & - & 5 & 6 & 4 \end{pmatrix} \in I_6 = (123](456)$. $a$and$b$has a circuit of length 3 and a proper 3-path and hence falls in the same conjugacy class. Also we can find a permutation $g\ in\ permutation\ group$ such that $y = g^{-1}xg$ or $gy = xg$ by matching the paths in a vertical order
$$a = (123)(456]$$
$$b = (456)(123]$$
from where we can deduce $g = (14)(25)(36)$
b. Let φ be a proper path and σ a circuit. A circuit is an extension of a permutation and permutation group has no nilpotent, therefore $x = \varphi\sigma \Rightarrow x^n \neq 0$. Hence x is nilpotent iff its path structure are only proper paths
c. Assuming $\varphi\ or\ \sigma$ has length 2, $x\alpha \neq x\ \forall x \in dom\alpha$. This implies that$x\alpha^2 \neq x\alpha$.

The definition of Conjugacy in arbitrary semigroup seems not to be unique as can be seen in(Kudryavtseva and Mazorchuk,2009) where they compared three approaches of Conjugacy on semigroups. (Dauns, 1989) gave a definition for monoid and (Lallement, 1979) for free semigroups.

Theorem2

Let $x \in I_n$. Then the following holds

a. The index of x is the maximum of lengths of all the proper paths in it. The index is one if no proper path exist in x.
b. The period of x is the lowest common multiple of all the lengths of the circuits in x. The period is one if no circuit exist in x.

Proof: See (Lipscomb, 1996) pp 13.

MATERIALS AND METHODS

The conjugacy classes of $\alpha \in ID_n$were arranged according to the fix of αdenoted as $f(\alpha)$and defined by $f(\alpha) = |F(\alpha)| = |\{x \in X_n: x\alpha = x\}|$ for any number ofID_n.The nilpotent conjugacy classes are marked as $(*)$, while the idempotent conjugacy classes are marked as (@). The Index andperiodof each conjugacy class was also found. Some illustrations are given below

Table 1:Conjugacy classes of ID_1

$f(\alpha)$	Conjugacy classes	Index	Period
0	(1]*@	1	1
1	(1)@	1	1

Table 2: Conjugacyclasses ofID_2

$f(\alpha)$	Conjugacy classes	Index	Period
0	(1](2]*@	1	1
	(21]*	2	1
1	(1)(2]@	1	1
2	(1)(2)@	1	1

Table 3: Conjugacy classes of ID_3

$f(\alpha)$	Conjugacy classes	Index	Period
0	(1](2](3]*@	1	1
	(21](3]*	2	1
	(321]*	3	1
1	(1)(2](3]@	1	1
	(1)(32]	2	1
2	(1)(2)(3]@	1	1
3	(1)(2)(3)@	1	1

Table 4:Conjugacy classes of ID_4

$f(\alpha)$	Conjugacy classes	Index	Period
0	(1](2](3](4]*@	1	1
	(21](3](4]*	2	1
	(21](43]*	2	1
	(321](4]*	3	1
	(4321]*	4	1
1	(1)(2](3](4]@	1	1
	(1)(32](4]	2	1
	(1)(432]	3	1
2	(1)(2)(3](4]@	1	1
	(1)(2)(43]	2	1
3	(1)(2)(3)(4]@	1	1
4	(1)(2)(3)(4)@	1	1

RESULTS AND DISCUSSION

We observed some combinatorial relations between numbers associated with the nilpotent and idempotent conjugacy classes ofID_n. We define the following numbers:
N_n = the cardinality of nilpotent conjugacy classes of ID_n
E_n = the cardinality of idempotent conjugacy classes ofID_n
M_n = the total number of chains in the nilpotent conjugacy class chain decomposition of ID_n
L_n = the total number of chains in the idempotent conjugacy class chain decomposition of ID_n

Table 5:Table of combinatorial relations

n	N_n	E_n	M_n	L_n
1	1	2	1	2
2	2	3	3	4
3	3	4	5	6
4	5	5	7	8
5	7	6	9	10
6	11	7	11	12
7	15	8	13	14
8	22	9	15	16
9	30	10	17	18
10	42	11	19	20

The following results are products of Table 5 above.

Lemma 1: Let $\alpha \in ID_n$, the nilpotent conjugacy classes of ID_n is given as $|N_n| = \frac{1}{192}n^4 - \frac{1}{16}n^3 + \frac{29}{48}n^2 - n + 2$when n is even.

Proof: Using Mathematical Induction

We show that$|N_2| = 2$ is true $|N_2| = \frac{1}{192}2^4 - \frac{1}{16}2^3 + \frac{29}{48}2^2 = 2$

$|N_k|$is true for some positive even integer: $|N_k| = \frac{1}{192}k^4 - \frac{1}{16}k^3 + \frac{29}{48}k^2 - k + 2$

We assume it holds for$|N_{k+2}| = \frac{1}{192}(k+2)^4 - \frac{1}{16}(k+2)^3 + \frac{29}{48}(k+2)^2 - k$

Then $|N_{k+2}| = \frac{1}{192}k^4 - \frac{1}{16}k^3 + \frac{29}{48}k^2 - k + 2 + \left(\frac{1}{24}k^3 - \frac{1}{4}k^2 + \frac{11}{6}k\right)$
$= \frac{1}{192}(k+2)^4 - \frac{1}{16}(k+2)^3 + \frac{29}{48}(k+2)^2 - k.$

Lemma 2: Let$\alpha \in ID_n$, the idempotent conjugacy classes of ID_n is given as $|E_n| = n + 1$.

Proof: For each idempotent rank of $n = 0, 1, \ldots, n \; in \; ID_n$, there exist at least one idempotent element. Also idempotent elements of a particular rank fall under a conjugacy class. Thus we have $n + 1$ idempotent conjugacy classes

Lemma 3: The total number of chains in the nilpotent conjugacy class chain decomposition of ID_n is$M_n = 2n - 1$.

Proof:Let α be a nilpotent transformation inID_n with domain$X_n = 1, 2, \ldots, n$.

The number of images$(x\alpha) < X_n$ in any n ofID_n. Therefore number of total nilpotent conjugacy class chain decomposition of n will be $(n) empty\ maps + (n - 1)\ maps$ of other combinations of chain decomposition of n. Hence$M_n = 2n - 1$

Lemma 4: The total number of chains in the idempotentconjugacy class chain decomposition of ID_n is $L_n = 2n$

Proof: From Theorem 1.1 we have that x is idempotent if and only if all the paths in its path structure is of length one. It implies that it is either a circuit of length one or a proper 1-path of n. Generally, there are two kinds of path in n ways. Therefore $L_n = 2n$

CONCLUSION

The sequence of$N_n = 1, 2, 3, 5, 7, 11, 15, 22, 30, 45, \ldots \; where \; n = 1, 2, \ldots$ is the partitions of a positive integer A000041 of the Online Encyclopedia of Integer Sequences (OEIS). Also the idempotent and nilpotentconjugacy classes for ID_n as well as the total number of chains in the idempotent conjugacy class chain decomposition of ID_n has been shown and hence satisfies the part structure of length one.

AKNOWLEDGMENT.

Our appreciation goes to Dr. Makanjuola, S. O. of the Department of Mathematics University of Ilorin for his valuable insight into this area and his immense contribution towards this work.

REFERENCES

Clifford, A. H. & Preston, G. B. (1961).The Algebraic Theory of Semigroups, *Math. Surveys American. Mathematical Society, Providence*, 7(1), 67-79.

Dauns, J. (1989). Centres of Semigroup Rings and Conjugacy classes. *Semigroup Forum* 38(1), 355 – 364.

Gomes, G. M. S. & Howie, J. M. (1987).On the ranks of certain finite semigroups of transformations, *Math. Proc Cambridge Philos. Soc*, 101, 395-403.

Howie, J. M. (1995). Fundamentals of Semigroup Theory. Oxford Clarendon Press, New York, Pp. 645.

Kudryavtseva, G. & Mazorchuk, V. (2009). On three approaches to Conjugacy in Semigroups. *Semigroup Forum,* (78), 14 – 20.

Lipscomb, S. (1996). Symmetric Inverse Semi groups, Mathematical Surveys of The American Mathematical Society, 46, Providence, R. I.

Munn, W. D. (1957). The characters of the Symmetric Inverse Semi group. *Proc. Cambridge Philos. Soc.,* 53, 13 – 18.

Lallement, G. (1979). *Semi groups and Combinatorial Applications*, Wiley, New York, Pp. 235.

Sullivan, R. P. (1987). Semi group generated by Nilpotent transformations. *Journal of Algebra*, 6, 324 – 343.

Ugbene, I. J. & Makanjuola, S. O. (2012).On the number of Conjugacy classes in the Injective Order-preserving Transformation Semi group. *Icastor Journal of Mathematical Sciences, India*, 1(6), 45-55.

Umar, A. (1992).Semi groups of Order-Decreasing Transformations, Ph. D Thesis, University of St. Andrews.

Umar, A. (2010).Some Combinatorial problems in the theory of Symmetric Inverse Semi groups. *Algebra and Discrete Math*. (9), 115 – 126.

PROPORTIONAL EFFECT OF OUTLIERS ON OVER-DISPERSION

[1]Akeyede, I[*].,[2]Saleh, I. M. and [3]Babalola, O. A.
[1]*Department of Mathematics, Federal University Lafia, P.M. B 146, Lafia, Nigeria.*
[2]*Department of Mathematics, Federal University Lafia, P.M. B 146, Lafia, Nigeria.*
[3]*National Population Commission, Osogbo, Osun State*
Corresponding Author: E-mail: imamakeyede@gmail.com

Date Manuscript Received:10/06/15 Accepted:06/12/15 Published: December 2015

ABSTRACT

The impact of outlier on analysis of time series data in causing over-dispersion was examined. The problem of over-dispersion is central to all General Linear Models (GLM's) having discrete responses. If the estimated dispersion after fitting is not near the expected values, then the data may be over dispersed. One of the causes of over-dispersion is outlier. Outlier is a data which is unusual with respect to the group of data in which it is found. In this paper, data were simulated based on poison model using SPSS and first analysed to see whether the estimated parameters is unbiased of the fixed parameters. Thereafter, two different values of outliers, 10's and 20's were introduced to different percentages of the generated data and then analysed using the STATA package to observe the effect of the outliers being introduced on the data for small, moderate and large samples. The data simulated were replicated 300 times for all categories. The averages of the results were computed. The results showed that the higher the percentage of outliers the more the over-dispersion occurs in the models and the larger the sample size the less the over-dispersion.

KEYWORDS: Outliers, Over-Dispersion, Simulation,

INTRODUCTION

An outlier is an observation that lies outside the overall pattern of a distribution (Moore and Mccabe, 1999). A convenient definition of an outlier is a point which falls more than 1.5 times the interquartile range above the third quartile or below the first quartile. It can also occur when comparing the relationship between two set of data. According to Oxford Dictionary of Statistics (2008), outlier is an observation that is very different to other observations in a set of data. It is a data value which is unusual with respect to the group of data in which it is found. It may be a single isolated value far away from all others, or a value which does not follow the general pattern of the rest. Usually the presences of outliers indicate some sort of problem. This can be a case which does not fit the model under study or an error measurement. Outliers are often easy to spot in histograms. Since the most common cause of outlier is recording error, it is sensible to search for outliers by means of summary statistics and plots of the data before conducting any detailed statistical modeling or analysis. If there is only a single outlier present, then an effective test is the Dixon test. For data from a normal distribution, the test statistic of the Grubbs test, suggested by Grubbs (1969), could be used. The Rosner test for multiple outliers relies on ordering the n observation interms of their distance from the overall mean. Certain statistical estimators are able to deal with statistical outliers and are robust while others cannot deal with them. A typical example is the case of median, that can deal with outliers well, since it would not matter whether the extreme point is far away or near the other data points, as long as the central value is unchanged. The mean on the other hand, is affected by outliers as it increases or decreases in value depending on the position of the outlier. According to Hardin and Hilbe (2007), presence of outliers in data set may rise to apparent over-dispersion. Over-dispersion is a phenomenon that occurs with data fitted using the binomial, poison or negative binomial distribution. If the estimated dispersion after fitting is not near the assumed values, then the data may be overdispersed, the value is greater than the expected value. It is underdispersed, if the value is less than expected. It is generally caused by positive correlation between responses or by excess variation between response probabilities or counts. It also arises when there are violations in the distributional assumptions of the data (Breslow, 1990). The problem with over-dispersion is that it may cause underestimation of standard errors of the estimated coefficient vector. A variable may appear to be a significant predictor when in fact it is not. Usman and Oyejola (2013) emphasized that apparent over-dispersion may arise from any of the following:

(i) The model omits important explanatory predictors
(ii) The data contain outliers.
(iii) The data contain excess zero.
(iv) The model fails to include enough interaction terms.
(v) A predictor needs to be transformed (to the log or some other scale).

The assumed linear relationship between the response and the link function and predictor is misspecified. (Hardin &Hilbe 2007) A model may be overdispersed if the value of the Pearson (or χ^2) statistics divided by the degree of freedom is greater than 1.0. The quotient of either is called the dispersion. Small amounts of over-dispersion are of little concern; however, if the dispersion statistics is greater than 1.25 for moderate size models, then a correction may be warranted. Models with large numbers of observations may be overdispersed with a dispersion statistics of 1.05 (Hilbe 2007). This study therefore examined the effect of proportion of outliers and sample size in causing over-dispersion to set of data

MATERIALS AND METHODS

Proportional impact of outliers were studied by creating simulated data set for small, moderate and large samples which were taken to be 20, 50 and 100 respectively. For each sample size, we introduced 1, 2 and 3 different sets of outliers out of each of the values 20, 50 and 100 of the response y_i following the idea of Usman and Oyejola (2013). These were replicated 300 times. For instance, the numbers of values of outliers introduced in each sample represent 5, 10 and 15 percent of the observations for the sample size of 20. The values of y_i simulated range from 0 to 9.

Two sets of outliers were introduced into generated data. In the first set we added 10 to the first, first and second, and first, second and third respective values of y_i randomly in the different data generated. While in the second set, we added 20 the same way. Each constructed data set entails a specific cause of the over-dispersion observed in the display of the model output stated as follows;

Constant (β_o) = 0.9 and β_1 = 0.2, β_2 = -0.5, β_3 = 0.6 are coefficients of the predictors. $t = 0, 1, \ldots, \infty$, and $i = 1, 2, \ldots, 300$

Results Output of Sample Size of 20 without Outliers using Stata Codes

glmy$_i$ x_{1i} x_{2i} x_{3i}, family(Poisson) link(identity) nolognonrtolerance i=1, …, 300

A sample output of the above code is given as linear models. No. of obs = 20

Optimization : ML
Residual df = 16

Scale parameter = 1

Deviance = 21.936224
(1/df) Deviance = 1.371014

Pearson = 32.374656
(1/df) Pearson = 2.023416

AIC = 2.805261

Log likelihood = -38.630618
BIC = -5514.594

OIM

yi	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
x1	.3427364	.0334277	10.25	0.001	.2772193	.4082534
x2	-.5451369	.0453168	-12.03	0.003	-.6339563	-.4563175
x3	.3167947	.030732	10.31	0.000	.2565612	.3770283
cons	1.124692	.0344106	32.68	0.001	1.057248	1.192135

From the above result, the parameter estimates are significantly different from the parameters fixed for the model having a responsey_i, i.e. y with the first responses having 10 added to the value y. the Pearson dispersion statistics, however, has doubled to a value of 2.0234. The AIC and BIC Statistics are also inflated.

Given a small number of observations, a value of 2.0234 indicates a serious over-dispersion, of course, we understand that the source of the over-dispersion result from the 10-outlier. Adding another 10's counts to the observations we already made to the first observations produce multiple over-dispersion (see table 1-6 for the results). More so another value of outlier was introduced in the generated data.In this case, we added 20 to the first data. The codes used in the STATA to generate the responses with an outlier introduced on the same set of predictors yield the output given below.

Results Output of Sample Size of 20 with an Outlier '20' using Stata Codes

genyi = y
replaceyi = yi+20 in 1/1
glmyi x1i x2i x3i, family(Poisson) link(identity) nolognonrtolerance i=1, …, 300

A sample output of the above code is given as linear models. No. of obs = 20

Optimization : ML
Residual df = 16

Scale parameter = 1
Deviance = 37.79224
(1/df) Deviance = 2.362015
Pearson = 58.03245
(1/df) Pearson = 3.627025

AIC = 5.002345
Log likelihood = -38.4567137
BIC = -11.0712

OIM

yi	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
x1	.3754891	.0245773	11.21	0.007	.3.012568	.78956321
x2	-.7896587	.0675321	-10.53	0.010	-.7.231562	-.6853217
x3	.4517123	.0564325	12.45	0.000	.0457852	.5876543
cons	2.154321	.0312543	13.69	0.023	1.453278	1.786549

When another value of outlier was introduced in the generated data, i.e 10 added to the first data. The Pearson index increased to 3.627025 which is seriously overdispersed. Also there is a change in the parameter estimates and the AIC and BIC criteria increase to 5.002345 and -11.0712 respectively. The effect of 10% of the observation constituted outlier is remarkable.

RESULTS AND DISCUSSION

ANALYSIS WITH VALUE OF 10'S ADDED TO SOME PERCENTAGE OF RESPONSE

The analysis of data when 10 was added to the first, first and second, and first, second and third data could be seen clearly in the table 1-6. The data were simulated for each three sample size under consideration and analysed. Then a value of 10 was introduced to 5%, 10% and 15% of 20, 50 and 100 observations and they were analysed using the Stata code. Each set of simulations were replicated 300 times, the average of the results were taken and displayed in table 1-3 as follows.

Table 1: Proportional Effect of Outliers in Analysis for Data Set of Sample Size 20

Percentage B_3 Index	Pearson Constant Likelihood	Log-	AIC	BIC	B_1	B_2	
0%	0.9298 -30.6322.8052-57.9050.2053-0.5042				0.6025		1
5%	1.5673	-36.9933.2997-23.6720.3126-0.76540.4136				1.3218	
10%	2.3426	-37.8923.6712-21.7650.4312-0.71120.4678				1.5432	
15%	2.5630	-39.8974.2314-19.8650.45210.78560.5632				1.6754	

Table 2: Proportional Effect of Outliers in Analysis for Data Set of Sample Size 50

Percentage B_3 Index	Pearson Constant Likelihood	Log-	AIC	BIC	B_1	B_2
0%	0.9006	-21.7662.3312-77.125		0.2103-0.50020.61251.0012		
5%	1.5632	-36.8763.1007-53.6740.3451-0.67340.42371.1432				
10%	2.4321	-42.6743.6588-34.7780.4654-0.69870.48911.5651				
15%	2.6126	-49.9884.0654-23.7650.4897				
0.867540.54671.7733						

Table 3: Proportional Effect of Outliers in Analysis for Data Set of Sample Size 100

Percentage B_3 Index	Pearson Constant Likelihood	Log-	AIC	BIC	B_1	B_2
0%	0.8731	-20.98762.3001-82.1350.2189		-0.49820.6521		1.0066
5%	1.5632	-34.90253.0992-54.0020.3500		-0.6434	0.4743	1.2318
10%	2.6003	-41.77783.6172-35.7980.4672-0.6532			0.49981.5832	
15%	2.6715	-42.1723.7145	-25.087	0.4951	0.7864	0.56321.9865

Tables 1-3 show the effect of percentage of outliers introduced when compared with those with 0% outlier. It was observed that the Pearson's index increases with percentage increase in the number of outliers, hence, over-dispersion occurred as all values of the index are greater than 1.0 for all categories of percentage of outlier introduced and sample sizes. Also from the results all the parameters estimates are significantly different from the fixed parameters. The increase in AIC and BIC information criteria when percentage of outliers increase indicate worse model from smaller to higher number of outliers. It was also observed that the parameters' estimates, Pearson Index and Log-Likelihood increased while AIC and BIC decreased when sample size was increased for different percentage of outliers introduced.

ANALYSIS WITH VALUE OF 20'S ADDED TO SOME PERCENTAGE OF RESPONSE

The data were simulated for each three sample sizes under consideration and analysed. Then a value of 20 was introduced to 5%, 10% and 15% of 20, 50 and 100 observations and they wereanalysed using the Stata code. Eachset of data was replicated 300 times for different percentage of outliers. The average values of the estimated parameters Pearson index, Log-Likelihood and the Information criteria were taken and presented in table 3-6 as follows.

Table 4: Proportional Effect of Outliers in Analysis of Data Set with Sample Size 20

Percentage	Pearson Index	Log-Likelihood AIC BIC B_1 B_2 B_3 Constant
0%	0.92981	-30.6322.8052-57.9050.2053-0.50420.60251.0000
5%	3.0239	-37.9814.0078-12.8970.2328-0.96670.43211.3254
10%	3.9876	-50.8985.2245-7.87690.3897-0.90070.43691.3367
15%	5.8976	-51.7896.8965-2.89650.1567-0.89760.51562.8976

Table 5: Proportional Effect of Outliers in Analysis for Data Set of Sample Size 50

Percentage	Pearson Index	Log-Likelihood AIC BIC B_1 B_2 B_3 Constant
0%	0.9006	-21.766 2.3312 -77.1250.2103-0.50020.61251.0012
5%	3.4321	-91.0454.0156-79.9880.4116-0.7654-0.56781.339
10%	3.9876	-123.685.2245-20.8980.3814-0.78960.59871.9876
15%	5.1065	-151.676.1897-21.7770.4292-0.79980.60022.8764

Table 6: Proportional Effect of Outliers in Analysis for Data Set of Sample Size 100

Percentage B3 Index	Pearson Constant Likelihood	Log- AIC BIC B_1 B_2
0%	0.9236	-121.76 2.1267 -89.7680.2007-0.51120.61151.0102
5%	0.9256	-255.923.0651-67.9870.2145-0.54320.67521.0795
10%	0.9285	-187.903.1236-55.7690.2276-0.54780.67861.2435
15%	1.8931	-176.443.6751-43.7860.22580.22580.68952.2367

Tables 3-6 show the effect of percentage of outliers introduced when compared with those with 0% outlier. In this case, outlier 20 was added to 5%. 10 and 15% of data simulated. It was observed that, the Pearson's index increases with percentage increase in the number of outliers, hence, over-dispersion occurred as all values of the index are greater than 1.0 for all categories of percentage of outlier introduced and sample sizes. Also from the results all the parameters estimates are significantly different from the fixed parameters. The increase in AIC and BIC information criteria when percentage of outliers increase indicate worse model from smaller to higher number of outliers.

CONCLUSION

Tables 1-6 show the effect of percentage of outliers introduced when compare with those with 0% outlier. It was observed that the Pearson's index increases with increase in the number of outliers, hence, over-dispersion occurred as all values of the index are greater than 1.0 for all categories of percentage of outliers introduced. Also from the results all the parameters' estimates are significantly different from the model. The increase in AIC and BIC information criteria respectively, when percentage of outliers increase, indicate worse model from smaller to higher number of outliers. It was also observed that, the parameters' estimates, Pearson Index and Log-Likelihood increased while AIC and BIC decreased when sample size was increased for different percentage of outliers introduced. Therefore, the outlier has little effect on the model with increase in the sample size and indeed there is little over-dispersion.

The study concluded that the higher the number of outliers in a set of data the higher the over-dispersion especially at smaller sample sizes. However if the sample size increases with the same number of outliers there will be little over-dispersion. It is therefore recommended that if outliers are present in a data set the sample size should be increased.

REFERENCES

Breslow, N. E. (1990). Tests of hypotheses in Overdispersed Poisson regression and other quasilikelihood models. *Journal of American Statistical Association, 85*, 565-571.

Cox, D. R. (1983).Some Remarks on Over-dispersion. *Biometrica*, 70, 269-274

Grubbs, F. E. (1969).Procedure for detecting outlying observations in samples. *Technometrics*, 11, 1-21.

Hilbe, J. M. (2008).Negative Binomial Regression. Cambridge University Press, United Kingdom. Pp. 51-61

Hardin, J. W. & Hilbe, J. M. (2007).Generalized Linear Models and Extensions. SecoEdition, AStata Press Publication StataCorp LP, College Station, Texas. Pp. 49-50, 165-166, 221-222.

Lambert, D. &Roader, K. (1995).Over-dispersion diagnostics for generalized linear models. *Journal of American Statistical Association*, 90, 1225-1236.

Moore, D.S. & McCabe, G. P. (1999). Introduction to the practice of Statistics, 3rd edition, W. H. Freeman, New York: Pp. 571-578.

Oxford Dictionary of Statistics (2008).*A* Dictionary of Statistics, Second edition revised. Oxford University Press, USA.

Usman, M. & Oyejola, B. A. (2013).Models for Count Data in the Presence of Outliers and/or Excess Zero. *Mathematical Theory and Modeling*, 3(7), 94-103

STRUCTURAL AND OPTOELECTRONIC PROPERTIES OF ZNS NANOPARTICLES EXPOSED TO ORGANIC AND INORGANIC CAPPING AGENTS

6

[1]Kalu, O.[*], [2]Amah, A. N., [3]Mbah, M. A.

[1]Department of Physics, Material Science Division, Federal University, Lafia, P.M.B. 146 Nasarawa State.
[2]Department of Physics, Material Science Section, University of Agriculture, Makurdi, P.M.B. 2737, Benue State.
[3]Department of Mathematics, Federal University, Lafia, P.M.B. 146 Nasarawa State.
*Corresponding Author: Email- kalu.onyekachi@fulafia.edu.ng; konyekachi@gmail.com

Date Manuscript Received: Accepted: Published:

ABSTRACT

ZnS nanoparticles capped with Sodium Carboxyl-methyl Cellulose (CMC-organic) and Tetraethylorthosilicate (TEOS-inorganic) as capping agents were synthesized by chemical co-precipitation and sol-gel technique respectively. Structural, morphological and optical studies of the synthesizes nanoparticles were carried out using X-ray Diffraction (XRD), Scanning Electron Microscopy (SEM), UV-Vis spectroscopy and Fourier transforms infrared spectroscopy (FTIR) . Studies of the capped ZnS nanoparticles revealed zinc blende structure. Particle sizes calculated from Williamson-Hall (W-H) analysis of Uniform Deformation Model (UDM) were $2 \cdot 57$nm (ZnS:CMC) and $1 \cdot 58$nm (ZnS:SiO_2). The average strain values calculated from the model indicated that ZnS:SiO_2 has $10 \cdot 3 \times 10^{-3}$ while ZnS:CMC has $5 \cdot 5 \times 10^{-3}$. All samples had smooth surfaces due to passivation of the capping agents. Optical absorption studies shows a blue shift of the absorption edge with band gap values of $3 \cdot 73$eV and $5 \cdot 00$eV for ZnS:CMC and ZnS:SiO_2 which are higher than the bulk value of $3 \cdot 60$eV of ZnS nanoparticles due to quantum confinement effect. The vibrational modes associated with OH-stretching at $3484 \cdot 25\text{cm}^{-1}$ and the carboxylic group –COOH at $1703 \cdot 02\text{cm}^{-1}$ confirms the presence of capping agent from the Fourier transform infrared spectroscopy measurement.

Keywords: ZnS, Capping agents, Chemical route, Sol-gel, Nanoparticle sizes, Strain and Optoelectronic

FULAFIA JOURNAL OF SCIENCE AND TECHNOLOGY | VOL. 1

INTRODUCTION

Semiconductor nanoparticles as an advance research material for most dynamic area of scientific research have significant fundamental and commercial interest. ZnS as an important group II–VI semiconductor compound with average wide band gap energy of has shown excellent properties in luminescence, photochemistry and low toxicity compared to other semiconductors (Amah *et al.,* 2012). This is because of the wide range of its application in areas such as phosphors, optics, electronics, solar cells, catalyst and sensors, hence the need for intensive investigation in order to understand the workability of the device in the above aforementioned applications (Ravi, 2011; Ahemen *et al.,* 2013) .

Nanoparticles generally are zero dimensional structures which are size dependent. As they approach the nanoscale regime, quantization and domination of atoms at the surface becomes significant due to the large surface area of the nanoparticles. Therefore during synthesis or preparation of nanoparticles the choice of capping agent in engineering of size and surface structure is of both theoretical and practical importance due to the inherent mechanical, chemical, electrical, optical, magnetic, electro-optical and magneto-optical properties that are substantially different from those observe for their corresponding bulk (Gnanam and Rajendran, 2011).

These capping agents which is either organic or inorganic during synthesis determines the material to be synthesized, since they are responsible for the stability on the surface of the nanoparticles through covalent or ionic interaction as well as isolation of individual nanoparticles under ambient conditions (Kalu *et al.*, 2013).

Organic capping agent's which are polymeric in nature consist of molecules that bind to the surface of the particles with a network of hydroxyl head group and an alky tail chain. They restrict the growth of the nanometric size of the particles by stabilizing the nuclei and larger nanoparticles against aggregation, as well as accurate control of size and surface structure (Sperling and Parak, 2012).

Inorganic capping agents has been observed to be wide band gap semiconductors, whose primary aim is the elimination of surface related defect states at the particle surface during synthesis and the confinement of charge carriers to the core material. This usually results to increase quantum yield and luminescence efficiency. They also decrease poly-dispersity which enhances colloid stability at control growth rate. The lowering of density of trap states in surface passivated samples has been attributed to the effect capping agent, hence its application in luminescent semiconductors (Reiss *et al.,* 2002).

Also capped zinc sulfide nanoparticles with varying emission spectra and decay times have been reported due to passivation of the nanocrystal's surface which plays important role in the optical properties and the controlling of dispersed particle sizes (Kalu *et al.*, 2013).

Widening of a semiconductor band gap stemming from quantum size effect can be obtained when the semiconductor particle size reaches or become less than the Bohr radius. Therefore to prepare such nanoparticles at engineered size and surface structure requires limiting their growth at the beginning of their formation when capped (Kalu *et al.*, 2013; Pawar *et al.*, 2010,).

The variation of particle size is a crystal distortion due to the confinement caused by capping. The strain-induced broadening stemming from such distortions depends largely on the lattice parameters and the diffraction angle which can be evaluated. Lattice strain is therefore a measure of the distribution of the lattice constants arising from crystal imperfections (Mote *et al*, 2012). Crystallite size and lattice strain affect the Bragg peak in different ways. Such as increase in width of the peak, the intensity of the peak and shift peak 2θ to a position accordingly. (Kumar and Rao, 2013).

This research investigated the effect of organic and inorganic capping agents on the particle sizes of ZnS nanoparticles and analyzed the associated strain induced due to peak broadening using Williamson-Hall Model.

MATERIALS AND METHOD

The ZnS nanoparticles were prepared from analytical grade of zinc sulphate hepta-hydrate($ZnSO_4 \cdot 7H_2O$), sodium sulphide nona-hydrate($Na_2S \cdot 9H_2O$), toluene, ethanol, deionized water, Sodium carboxyl-methyl cellulose (CMC) and tetraethylorthosilicate (TEOS) .

The synthesis process was carried out by mixing 0.2M of CMC solution (1 % wt/wt in deionized water) with 50 ml of (4.39g) of $ZnSO_4.7H_2O$ under stirring for 25min. This was followed by adding (drop-wise) 50 ml of (4.80g) of $Na_2S.9H_2O$ using a pipette to the Zn^{2+}/ CMC precursor solution at room temperature under continuous stirring. Continuous stirring was uninterrupted until the last drop of the S^{2-} precursor solution was added for another 25min to allow for complete mixing of the reactants and formation of a white precipitate. The precipitate was then centrifuged at 3500 revolution per minute (rpm) for 20min to obtain a white precipitate which was washed several times with toluene/ethanol mixed solution. The white precipitate was filtered and then dried at 100^0C in a vacuum oven for 4 hours. The dried sample (ZnS:CMC) was grinded in an agate mortar and pestle to obtain a fine powder for characterization. Note that before drying, part of the washed precipitate was taken out and re-dispersed into ethanol for optical measurements. The above procedure was repeated using tetraethylorthosilicate (TEOS) as an inorganic capping agent with little modification during synthesis for the sample ($ZnS:SiO_2$).

was carried out using X-ray Diffraction (PANalytic X-ray Diffraction machine at Engineering Materials Development Institute Akure, Nigeria) with Cu Kα radiation (λ = 0.154056 nm), Scanning electron microscope (Carl-Zeiss MA 10 series machine at Sheda Science and Technology Complex Abuja, Nigeria) for the morphologies of the synthesized samples deposited on a carbon-coated grid for examination. Optical absorption spectra of the nanoparticles samples dispersed in ethanol were recorded at room temperature using a UV-Visible spectroscopy (JENWAY 6405 at Federal University Lafia, Nasarawa State Nigeria) and Fourier Transform Infrared spectroscopy between the range of $400-4000cm^{-1}$ were recorded (Shimadzu-IR 8400s at the National Research Institute of Chemical Technology, Zaria, Kaduna State, Nigeria) to verify the presence of capping agents.

RESULTS AND DISCUSSION

The result presented in Fig.1 shows that the XRD pattern of the samples ZnS, ZnS:CMC and ZnS:SiO_2 nanoparticle has a zinc blende (cubic) crystal structure respectively.

The spectrum as deployed in Fig.1 shows that the diffraction peaks at about 2θ values were $32.07^0, 51.96^0, 60.54^0$(ZnS) , $27.0^0, 51.72^0, 59.86^0$(Z whose corresponding diffraction planes are (111), (220) and (311). It is clear from Fig.1 that there are three major diffraction peaks. The highest intensity reflection peak is at about 27.0° for capped samples (ZnS:CMC , ZnS:SiO_2) with two other small intensity peaks at $51.72^0, 59.86^0$ and 49.08^0 58.54^0 respectively indicating that [111] is the preferred direction and at 32.07^0 for the uncapped sample (ZnS). The broadening of the XRD peak confirms the formation of the ZnS nanocrystals at particle sizes regime of 7.61nm, 2.57nm and 1.58nm for ZnS, ZnS-CMC and ZnS-SiO_2 respectively as calculated from the full-width half-maximum(FWHMs) for specific diffraction peaks using the Williamson-Hall (W-H) analysis of Uniform Deformation Model (UDM) (Mote *et al.*, 2012) . However the nature of XRD pattern for the ZnS-SiO_2 sample may be associated with Ostwald ripening and phase transformations occurring concurrently during the poly-condensation reactions followed by gel transforms into a solid mass.

The information concerning the strain and the particle size were obtained from the full width at half maximum (FWHMs) of the three diffraction peaks. The FWHMs (β) can be expressed as a linear combination of the contributions from the strain (ε) and particle size (D)nm of the Scherrer's equation which are components for the formation of Williams Hall analysis for particle size and lattice strain with the following relation shown below

$$\beta_{hkl} = \frac{k\lambda}{Dcos\theta} + 4\varepsilon \tan\theta \qquad (1)$$

$$\beta_{hkl}cos\theta = \frac{k\lambda}{D} + 4\varepsilon \sin\theta \qquad (2)$$

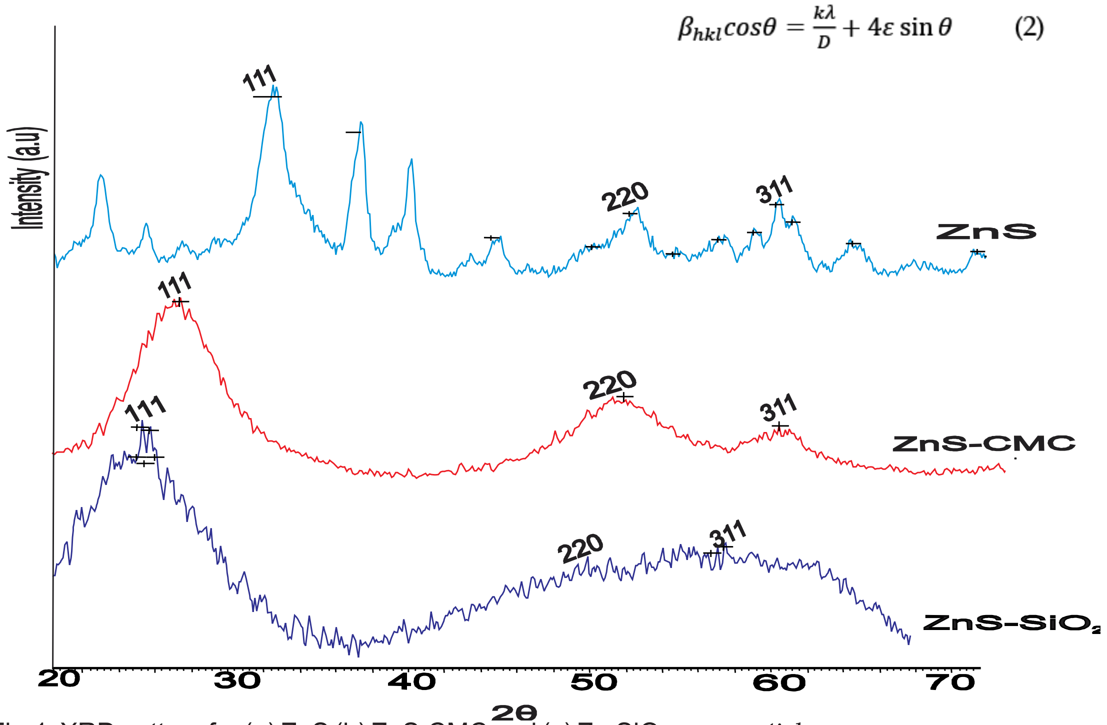


Fig.1. XRD pattern for (a) ZnS (b) ZnS:CMC and (c) Zn:SiO_2 nanoparticles

Equation (1) represents the Uniform Deformation Model (UDM) of Williamson-Hall equation for strain analysis whose plot separates the contribution due to strain (ε) and crystallite size (D) towards the line broadening.

The result in Fig.2 below represent the plot of $\beta_{hkl}cos\theta$ verse $4\sin\theta$ whose amount of strain (ε) is the slope of the linear fit and calculated crystallite size from the intercept on $\beta_{hkl}cos\theta$ axis.

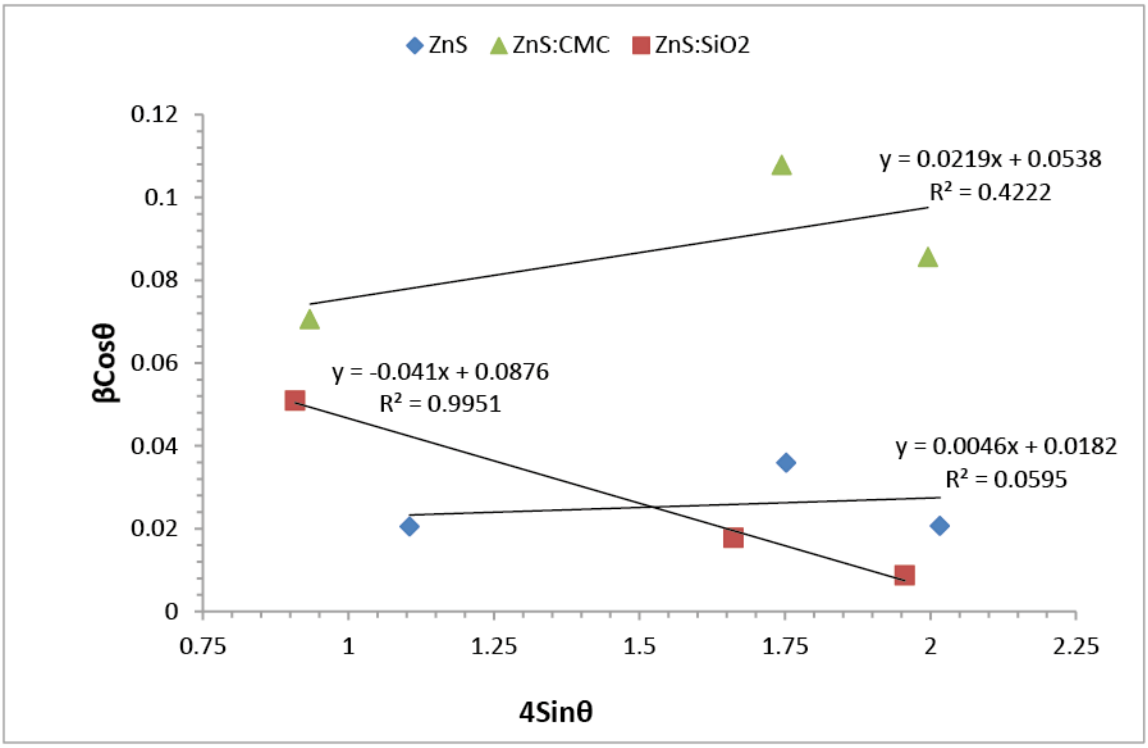


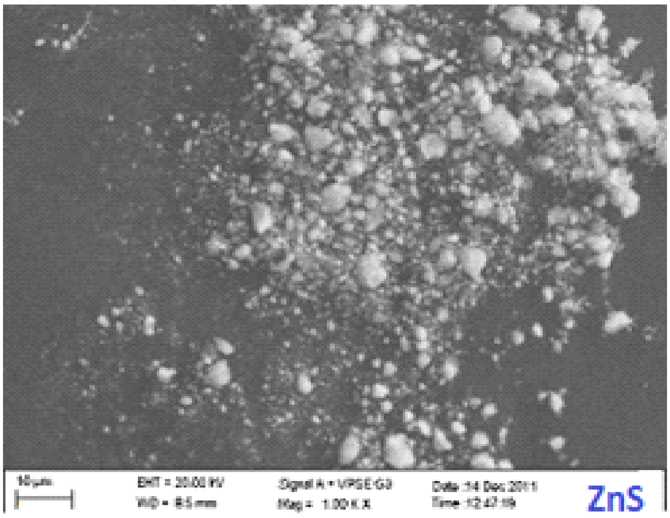
Fig.2. Williamson-Hall plots of powdered XRD data of ZnS, ZnS:CMC and ZnS:SiO_2

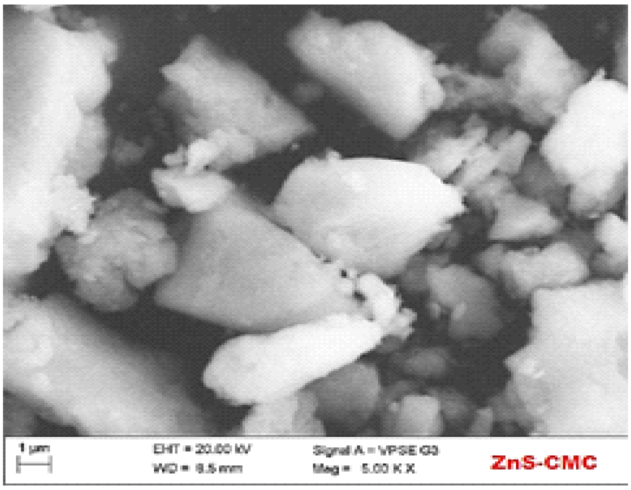
Table 1. Lattice parameter, structure, energy band gap, particle sizes and the strain values

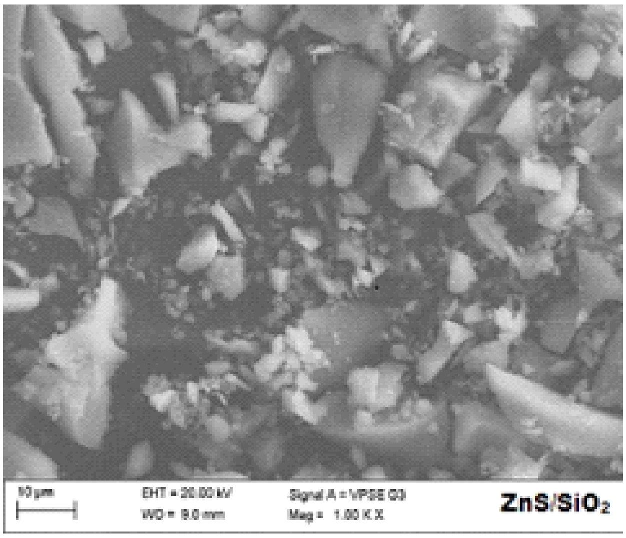
Sample	2θ	*hkl*	d_{hkl} Å	Lattice Constant Å	Structure	Energy band gap (eV)	Williamson-Hall method (UDM) Particle Size (nm)	Strain (ε)
ZnS	32.07	111	2.7886	4.8299	cubic			
	51.96	220	1.7584	4.9735		3.60	7.61	0.0011
	60.54	311	1.5281	5.0681				
ZnS:CMC	27.00	111	3.3002	5.7161	cubic			
	51.72	220	1.7662	4.9955		3.73	2.57	0.0055
	59.86	311	1.5439	5.1205				
ZnS:SiO_2	26.25	111	3.3925	5.8759	cubic			
	49.08	220	1.8547	5.2458		5.00	1.58	0.0103
	58.54	311	1.5755	5.2253				

The results presented in Table 1, show that the particles sizes obtained using Williamson-Hall Uniform Deformation Model (UDM) confirmed changes in particle sizes due to capping effect. Also the calculated values of the strain induced on the nanoparticles were directly proportional to the energy band gap and inversely proportional to the nanoparticle sizes. In comparison, it is observed that ZnS:SiO_2 has the smallest particle size, high induced strain and the largest energy band gap which indicates an excellent combination for tuning optoelectronic properties.

The results in Fig.3 shows the SEM micrographs of the organic and inorganic capped ZnS nanoparticles synthesized in this work. The morphology of the nanoparticles depends on the capping agent used such that the ZnS sample has a rough surface texture while that of ZnS:CMC and ZnS:SiO_2 revealed smooth surfaces confirming the presence of capping agent







The optical absorption spectra of the synthesized nanoparticles samples (ZnS, ZnS:CMC and ZnS:SiO_2) using the measurement from UV-vis spectroscopy are shown in Fig. 4a, with the absorption wavelengths ranging from 220-308nm. It was observed that at the same concentration of the capping agents the sample peak of (ZnS:SiO_2) shifted towards a low value 220nm wavelength(nm) compared to other sample.

The plot of the square of the absorption coefficient $(\alpha hv)^2$ verses energy (hv) for determining the energy band gap for all samples are presented in Fig.b4. The energy band gaps of the samples were estimated by extrapolating the straight-line portion of the spectrum to a zero absorption coefficient using the Tauc's relation of (equation 3) for direct band gap semiconductor (Ahemen *et al*, 2013).

$$(\alpha hv)^2 = (hv - E_g) \qquad (3)$$

The observed blue shift in the absorption edge is a reflection of the band gap increase owing to quantum confinement effect which can be associated to decrease in particle sizes at nanosize regime. The obtained optical band gap corresponding to the absorption edge for ZnS, ZnS:CMC and ZnS:SiO_2 samples were 3.60eV, 3.73eV and 5.00eV respectively as shown in Fig. 4(b). The band gap values of 0.13eV and 1.40eV were suggested to be as a result of blue shifted for ZnS:CMC and ZnS:SiO_2 samples as compared to the uncapped ZnS sample of energy band gap blue shifted at of 0.00eV (Murali-Krishna *et al.*, 2010). These results also indicate that at the same molar concentration, the capping ligands determine the actual value of electronic and optical properties of the nanoparticles.

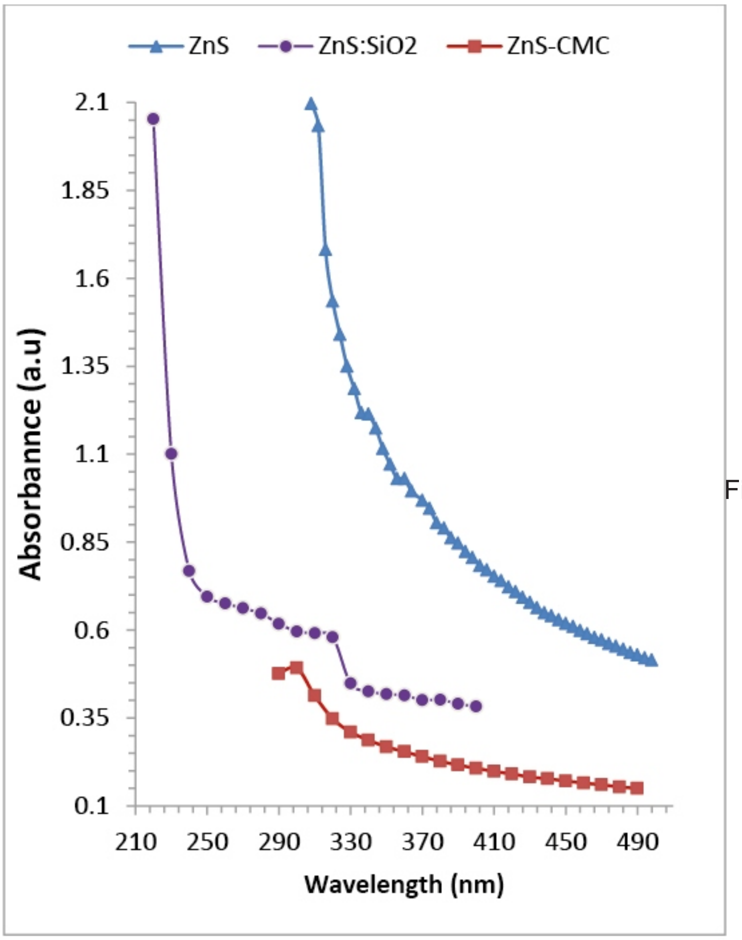


Fig. 4a: Optical Absorbance of ZnS, ZnS:CMC $ZnS:SiO_2$ nanoparticles

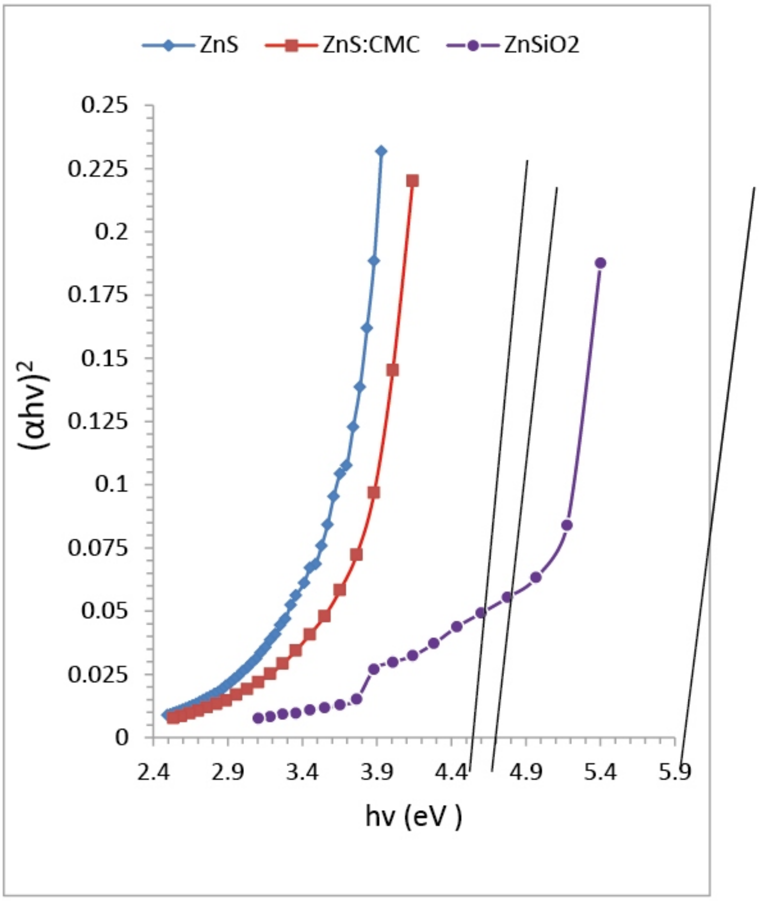


Fig.4b: Energy band-gap for ZnS, ZnS:CMC and $ZnS:SiO_2$

The characteristic absorption peak of the capping agent from FTIR confirmed the effective surface passivation of the nanoparticle in figures 5(a-b) for the capped samples only.

The result in Fig.5a reveals the spectra of capped ZnS:CMC with an associated band belonging to carboxylic groups shifted to a higher wavenumber of 1615.44cm^{-1} for C=O stretching and 1474.63$cm^{-1}$for C-O stretching. This may linked to their strong interaction with the Zn^{2+} ions forming a Zn–CMC complex during the in situ precipitation process. Fig.5b depicts the spectra of ZnS:SiO_2 nanoparticles with bending vibration mode at 1678.13cm^{-1} indicating C=O stretching and C-H out of plane bending at 942.26.cm^{-1} .The peak at 1383.01cm^{-1}may be due to -Si-C_6H_5 stretching.

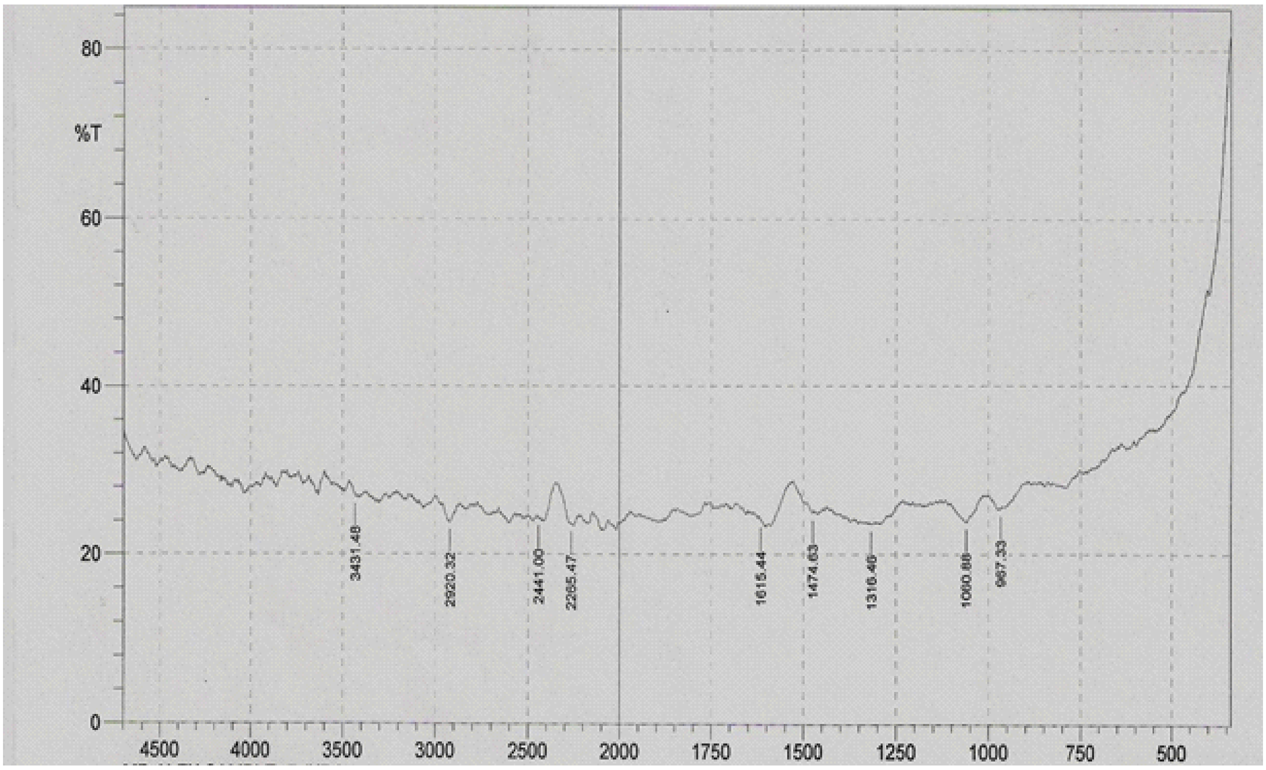


Fig. 5a: FTIR spectrum of ZnS/CMC **Wave number (cm^{-1})**

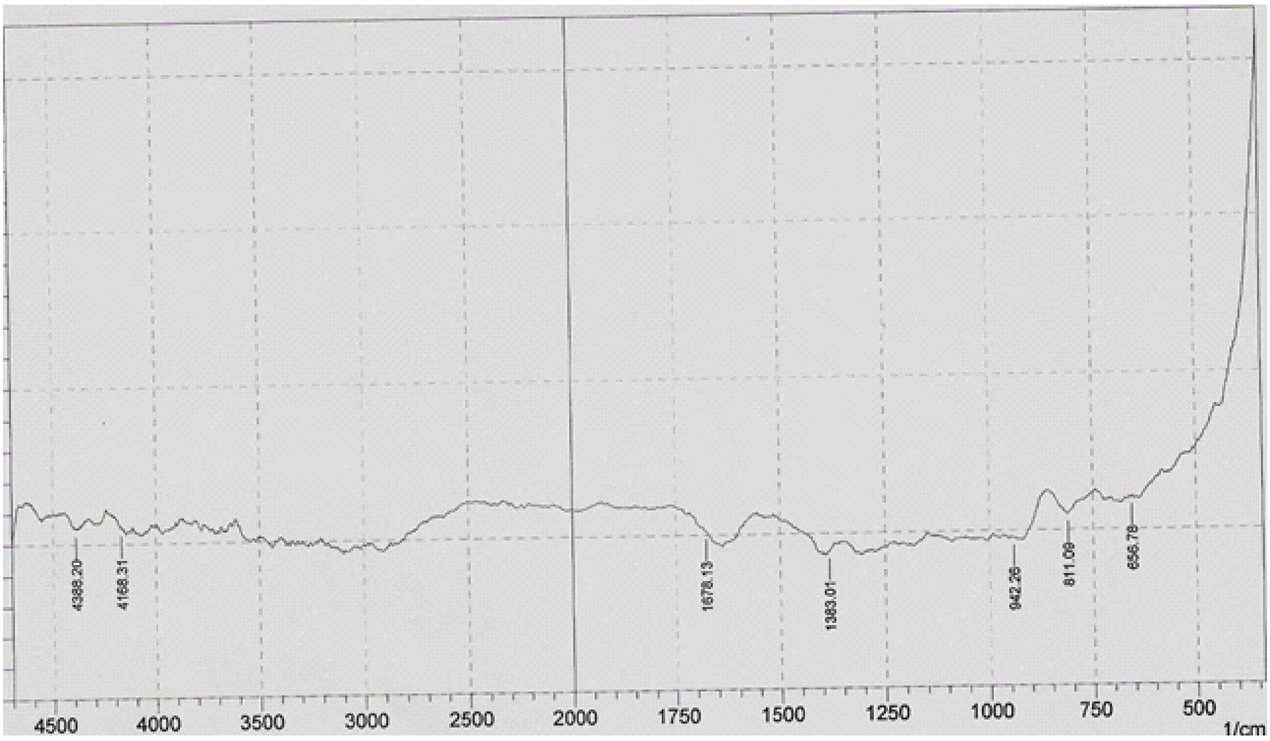


Fig.5b FTIR spectrum of ZnS:SiO_2 **Wave number (cm^{-1})**

CONCLUSION

The studies have revealed that organic and inorganic capping agents prevented particle agglomeration through its adsorption on the surface of nanoparticles. The particle sizes calculated from Uniform Deformation Model (UDM) are in range of 1.58nm to 2.57nm for the capped samples to that of 7.61nm for uncapped ZnS. The obtain energy band gaps shows a blue shift of 0.13eV to 1.40eV as compared to the absorption energy of the bulk ZnS.

This suggests the importance of capping agents in determining the actual value of electronic and optical properties of nanoparticles during nucleation and limiting the growth of the ZnS particles. In addition the values of crystallite sizes obtained are in good agreement with the values obtained for the energy band gap indicating the presence quantum confinement. In comparison it can be noted that the high strain value caused decrease in particle sizes evaluated which confirms the presence and good effect of capping agents.

AKNOWLEDGEMENT

The authors wish to thank the Vice Chancellor of Federal University Lafia, Prof. Ekanem Ikpi Braide, Dean Faculty of Science, Prof. Emmanuel Hala Kwon-Ndung and the Head of Physics Department, Prof. Sanusi Mohammed Liman for supporting this work.

REFERENCES

Ahemen, I., Meludu, O. & Odoh, E. (2013). Effects of pH Variation on Structural and Optoelectronic Properties of Zinc Sulfide nanoparticles Capped with Sodium Carboxymethyl Cellulose. *Journal of Nano and Electronic Physics,* 5(3), 03016.

Amah, A. N., Echi, M. I. & Kalu, O. (2012). Influence of Polyvinyl Alcohol and Alpha-methacrylic Acid as Capping Agents on Particle Size of ZnS Nanoparticles. *Applied Physics Research.* 4 (4), 1916- 9639

Gnanam, S. & Rajendran, V. (2011). Optical properties of capping agents mediate lead oxide nanoparticles via facile hydrothermal process. *International Journal of Nanomaterials & Biostructures,* 1(2), 12- 16.

Kalu, O., Amah, A. N. & Echi, M. I. (2013). Studies on the effect of capping agents on structural and optoelectronic properties of ZnS Nanoparticles. *Nigeria Journal of Physics,* 24, 96-101.

Kumar, V. S. & Rao, K.V. (2013). X-ray Peak broadening analysis and Optical studies of ZnO nanoparticles derived by surfactant assisted combustion synthesis. *Journal of Nano & Electronic Physics*, 5(2), 020-26.

Mote,V. D., Purushotham, Y. & Dole, B.N. (2012). Williamson-Hall analysis in Estimation of lattice strain in nanometer-sized ZnO particles. *Journal of Theoretical & Applied Physics,* 6(6), 2251-7235.

Murali-Krishna, D., Vijayalakshmi, R. P., Venugopal, R. & Reddy, B. K. (2010). Structural and Optical studies of $Zn_{0.98}Mn_{0.02}S$ Nanoparticles prepared by Chemical route with different capping agent. *Chalcogenide Letters*, 7 (5), 357-365.

Pawar, M. J., Nimkar, S. D., Nandurkar, P. P., Tale, A. S., Deshmukh, S. B. & Chaure, S. S. (2010). Effect of Sn^{2+} doping on Optical properties of thiourea capped ZnS nanoparticles. *Chalcogenide Letters,* 7 (2), 139 – 143.

Reiss, P., Bleuse, J. & Pron, A. (2002). Highly luminescent CdSe/ZnSe core/shell nanocrystals of low size dispersion. *Nano Letters*, 2, 781–784.

Sperling, R. A., & Parak, W. J. (2010).Surface modification, functionalization and bioconjugation of colloidal inorganic nanoparticles: *Phil.Trans. R. Soc.A*, 368, 1333 1383.

FULafia Journal of Science and Technology (FJST):

CALL FOR PAPERS

INTRODUCTION

FULAFIA JOURNAL OF SCIENCE AND TECHNOLOGY (FJST) is an International journal published by the Faculty of Science, Federal University Lafia. It is a peer-reviewed journal that publishes papers that contribute immensely to all areas of Agricultural, Biological, Environmental, Natural and Physical Sciences. The goal of this journal is to provide a global platform for scientists and academics from all over the world to promote, share, and discuss new issues and developments in different areas of Sciences.

The Journal welcomes the submission of manuscript(s) that meet the general scope and criteria of FJST. Our objective is to inform authors of the decision on their manuscript(s) within a few weeks of submission. Authors should submit their original manuscripts of research–based papers via email attachment to fulafjst@fulafia.edu.ng

Papers must be written in English Language, attached in a Microsoft word document, and should have thorough grammar and proper terminologies.

Following acceptance of a manuscript, the paper will be published in the next issue. Kindly support this initiative by publishing your papers with this upcoming and promising journal. Details of author's guide and Instructions are available on our website at www.fulafia.edu.ng.

COPYRIGHT

Authors are advised to ensure that articles submitted to the FULafia Journal of Science have never been submitted to any other journal. Papers will be subjected to editorial peer review process. By submitting a manuscript, the author agrees to transfer the copyright of the manuscript automatically to the FULafia Journal Science at the time of acceptance for publication. The decision of the editors on the acceptability of a paper or otherwise is final.

PREPARATION AND ARRANGEMENT OF MANUSCRIPTS

Manuscripts should be prepared in English language and should be clearly written. The subsections should include the Abstract, Introduction, Materials and Methods, Results and Discussion, Conclusion and Recommendations and References.

Manuscripts should be typed in Word form or Latex format in Time Roman 12 font size, double-spaced with margins of 4cm. A manuscript should be divided into cover page (title, names of author(s), author's affiliation, and address for correspondence and reprints, including e-mail addresses), abstract and keywords, introduction, materials and methods, results and discussion, conclusion, acknowledgement (if any) and references. A maximum of 10 printed pages is allowed.

TITLE

Title of manuscript should not be ambiguous and must be concise and reflecting the content of the paper. A maximum of 20 words is recommended for the title.

ABSTRACT

A concise and factual abstract is required. The abstract should state briefly the purpose of the research, the principal results and major conclusions. The abstract should be structured with a minimum of 250 words and a maximum of 300 words. Abbreviations should be avoided and if used should be explained at first mention.

INTRODUCTION
Should contain the objectives of the work and provide an adequate background. Avoid a detailed literature review.

MATERIALS AND METHODS
Provide sufficient detail to allow the work to be reproduced. Methods already published should be indicated by a reference: only relevant modifications should be described.

RESULTS AND DISCUSSION
Results should be clear and concise. Discussion should explore the significance of the results of the work and this should not be repeated. Avoid extensive citations and discussion of published literature.

CONCLUSION
The main conclusions of the study may be presented in a short Conclusions section, which may stand alone or form a subsection of a Discussion or Results and Discussion section.

ACKNOWLEDGEMENTS
Acknowledgements should be made in a separate section at the end of the article before the references. List here those individuals who funded or provided help during the research.

REFERENCES
References should be listed on a separate sheet and should be listed alphabetically without numbering at the end of the text. References should be arranged first alphabetically and then further sorted chronologically if necessary. More than one reference from the same author(s) in the same year must be identified by the letters "a", "b", "c", etc., placed after the year of publication. For example, Kwon-Ndung et al., 2009a, 2009b, 2009c).

All citations within the text should follow these formats:

1. Single author: the author's name (without initials) and the year of publication;
2. Two authors: both authors' names and the year of publication;
3. Three or more authors: first author's name followed by "et al." and the year of publication.

Groups of references should be listed first alphabetically, then chronologically. Examples: "as reported (Kwon-Ndung, 2009a, 2009b, 2012; Salihu and Chuku, 2009). Nja (2014) recently reported ……"

Only published papers can be mentioned in the manuscript. Reference to a journal publication, books, chapters in a book, thesis etc should be listed as follows:

Journal Article:
Edache, J.A., Ehiobu, N.G and Njike, M.C. (2003). Performance of laying Japanese quail (Cortunix cortunic japonica) fed different levels of protein under Nigerian environment. J. Agric Sci. & Tech. 13(2):110-116

Book
Helleiner, G.K. 1966. Peasant agriculture, government and economic growth in Nigeria. Homewood, Illinois, Irwin Press.
Chapter in an edited book: Jonathan, G. (1999). Agriculture and national development. In: Oduye, M.L. and Suswan, J.T. (Eds.), Politics and the agrarian revolution in Africa. Inter-Academy Publishers, Lagos, pp. 324-334.

ILLUSTRATIONS

Tables

Number tables consecutively in accordance with their appearance in the text. Place footnotes to tables below the table body and indicate them with superscript lowercase letters. Avoid vertical rules. Be sparing in the use of tables and ensure that the data presented in tables do not duplicate results described elsewhere in the article.

Figures

All illustrations, whether diagrams or photographs, are referred to as Figures. If any figures appear in colour, please note that they will only appear in colour in the online version; in the printed version they will be in black and white. If the quality of the colour figure supplied is not suitable to be produced in colour, it will also be shown in black and white in the online version. Figures should ideally be black and white, not colour, and numbered sequentially. However, if colour is essential to the figure please send a good quality colour image. Please place them at the end of the article, rather than interspersed in text.

Please prepare all figures, especially line diagrams, to the highest possible standards. Bear in mind that lettering may be reduced in size by a factor of 2 or 3, and that fine lines may disappear.

Units of Measurement

FULafia Journal of Science and Technology follow the Système International (SI) for units of measurement. Similarly IUPAC and Authors and editors are, by general agreement, obliged to accept the rules governing biological nomenclature, as laid down in the International Code of Botanical Nomenclature, the International Code of Nomenclature of Bacteria, and the International Code of Zoological Nomenclature. All crops, plants, insects, birds, mammals, etc. should be identified by their scientific names when the English term is first used, with the exception of common domestic animals.

Math formulae

Present simple formulae in the line of normal text where possible and use the solidus (/) instead of a horizontal line for small fractional terms, e.g., X/Y. In principle, variables are to be presented in italics. Powers of e are often more conveniently denoted by exp. Number consecutively any equations that have to be displayed separately from the text (if referred to explicitly in the text). Imperial units will be converted, except where conversion would affect the meaning of a statement, or imply a greater or lesser degree of accuracy.