

# Statistical Analysis on the Reporting Cell Based Scheme for Mobile Computing

N. Shakeela<sup>1</sup>, and S. Thabasu Kannan<sup>2</sup>

## ABSTRACT

The tremendous development in mobile endorers starting late that it has realized the misusing of wireless network resources, especially, the information exchange bandwidth available. The aim of adroit abuse of the repressed available information transmission and to manufacture the cutoff of the network, reiteration of re-use thought is set up in cell frameworks which supported in growing the quantity of cells in the framework. This incited inconvenience in finding the position of a moveable customer in the network and development in the signaling cost. Likewise, Location administration oversees checking a dynamic mobile terminal in a specific zone though diminishing the expense realized in finding the mobile terminal. The present location management is finished by congregating the telephones in light of endorser density. A system called “Reporting Cell Based Schemes” is utilized to decrease the paging expense and area update cost. In this paper, the regression analysis on the input dataset is utilized for information depiction, parameter estimation, estimation and prediction of the variables. Though Hypothesis examination on the resultant dataset is utilized to locate the critical of the utilized technique as a part of the system of “Reporting Cell Based Schemes”.

**Keywords:** Regression Analysis, Hypothesis Test, Genetic Algorithm, Reporting Cell based Scheme, Location Management, Paging Cost , Total Cost, Network size.

## 1. INTRODUCTION

In our everyday life the organization of Wireless Communication [1] has turned crucial in the field of innovation. At present there is no world without wireless communication. Particularly, keeping up our center point to wireless communication, cell phones have bowed into an unavoidable bit of man. Because of the immense advancement of phones, the world has ended up being so little and it is connected with every corner and anteroom. The cutting edge development [2] has made unconstrained cell communication to the extent reasonableness, fabulousness of administration, transportability, different administrations and unwavering quality. Thusly, remote cell frameworks have been changed in accordance with transportability. There is a quick exponential improvement in cell communication since years and will continue continuing later on. This headway in client based communication structures is particularly related to abuse of radio reach.

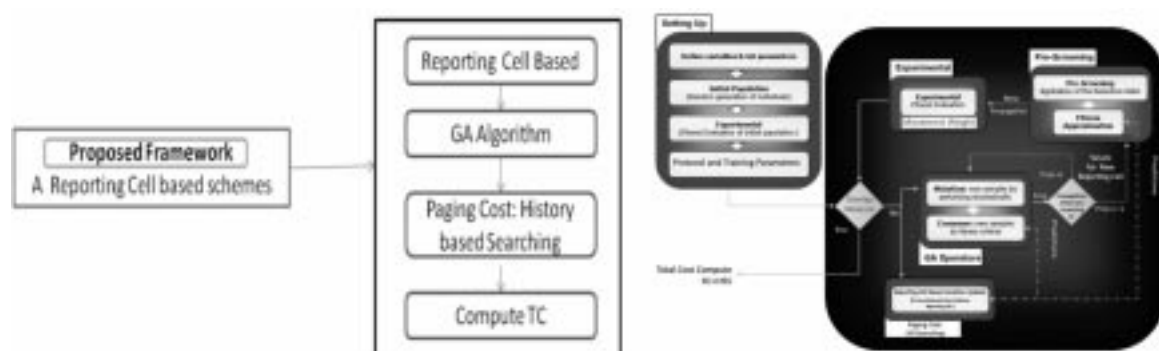


Figure 1: A proposed Framework for Location Management for Network and Framework setup using Genetic Algorithm and History Based Selection [3]

1 Research Scholar, PRIST University, Tanjavur, Tamilnadu, India

2 Principal and Research Advisor, Pannai College of Engineering and Technology, Sivagangai, Tamilnadu, India

This framework is intended to beat the issues in the dynamic location management in mobile environment i.e to minimize the total cost and paging expense is basically considered. The supplementary strides are included in this system.

**Reporting cell:** In Cellular system absolute system is isolated into cells. In this existing design a small number of cells are distributed in the engineering as a reporting cell and some other are non reporting cell. Finding of a best amount of reporting cells of the known design is a huge reporting the cell organizing issue. The problem is a hard optimization issue. For given cell design with N cells, the amount of feasible measures is  $2^N$ . For improvement we utilize genetic algorithm, it was produced to discover ideal solution for some optimizing issues. In this investigation, we deploy the genetic algorithm to position the best and ultimate answer for the reporting cell arranging issue. We show the outcomes got to reporting mobile arranging deploy the genetic algorithm [4]. In the previous researches, differential development, particle swarm optimization, replicated annealing and ant colony optimization were exploited for arrangement.

Genetic Algorithm (GA) [4] is a versatile heuristic search algorithm taking into account the organic developmental instrument of genetics and natural selection. GA is a evolutionary calculation which can discover great, conceivably ideal arrangements, to improvement issues with tremendous state spaces to be sought. It is globally probabilistic investigation method. GA is enlivened by Darwin's hypothesis about evolution - "survival of the fittest". The accompanying approaches are included in the genetic algorithm.

## 2. REGRESSION ANALYSIS FOR GIVEN INPUT DATASET

A statistical tool [5] that permits you to analyze how numerous independent variables are identified with a dependent variable [6]. When you have recognized how these various variables identify with your reliant variable, you can take data about the greater part of the independent variables and use it to make considerably more capable and exact expectations concerning why things are how they are.

- **Correlation Coefficient:** A solitary outline number that lets you know whether a relationship exists between two variables, how solid that relationship is and whether the relationship is sure or negative.
- **The coefficient of Determination:** A solitary outline number that lets you know the amount of variety in one variable is specifically identified with variety in another variable
- **Linear Regression:** A process that allows you to make predictions about variable "Y" based on knowledge you have about variable "X".
- **The Standard Error of Estimate:** A solitary outline number that allows you to tell how accurate your predictions are likely to be when you perform Linear Regression.

The equation of the plane of regression of  $Y_1$  on  $Y_2, Y_3, \dots Y_m$  is

$$Y_1 = a_{12.34\dots m} Y_2 + a_{13.24\dots m} Y_3 + \dots + a_{1n.23\dots(m-1)} Y_m \quad (1)$$

The formula for multiple Regressions as follows:

$$X' = b + a_1 Y_1 + a_2 Y_2$$

$X'$  = A predicted Value of X (Which is your dependent variable)

$b$  = The "X Intercept"

$a_1$  = The change in X for each 1 increment change in  $Y_1$

$a_2$  = The change in X for each 1 increment change in  $Y_2$

$Y$  = An Y score (Y is your dependent variable) for which you are trying to predict a value of X.

$$a_1 \left( \frac{r_{x,y1} - r_{x,y2} r_{y1,y2}}{1 - (r_{y1,y2})^2} \right) \left( \frac{SD_x}{SD_{y1}} \right)$$

$$a_2 \left( \frac{r_{x,y2} - r_{x,y1} r_{y1,y2}}{1 - (r_{y1,y2})^2} \right) \left( \frac{SD_x}{SD_{y2}} \right)$$

$r_{x,y1}$  = Correlation between highest value and average value,  $r_{x,y2}$  = Correlation between motivation and average value,  $r_{y1,y2}$  = Correlation between highest value and motivation value,  $(r_{y1,y2})^2$  = Coefficient of determination (r squared),  $SD_x$  = Standard deviation of X (dependent) variable,  $SD_{y1}$  = Standard deviation of Y variable,  $SD_{y2}$  = Standard deviation of second Y variable.

### 3. HYPOTHESIS ANALYSIS FOR OUTCOME DATASET

The chi-square (I) test is utilized to figure out if there is a critical contrast between the normal frequencies and the watched frequencies in one or more classifications [7].

#### A. Chi-Square Test for Independence

Keep in mind, subjective information is the place you gather information on people that are classes or names. At that point you would tally what number of the people had specific qualities. If you somehow managed to do a speculation test, this is your option theory and the invalid theory is that they are free. There is a theory test for this and it is known as the Chi-Square Test for Independence. In fact it ought to be known as the Chi-Square Test for Dependence, yet for recorded reasons it is known as the test for freedom [8].

total of columns B total happen

total number of individuals

1. State an alternative and null hypothesis and the level of significance  $H_0$ : the two variables are independent (this implies the one variable is not influenced by the other)  $H_A$ : the two variables are dependent (this implies the one variable is influenced by the other) Also, express your  $\alpha$  level here.
2. State and check the presumptions for the test of hypothesis.
  - a. The random sample is taken.
  - b. Expected frequencies for every cell are more prominent than or equivalent to 5 (The normal frequencies, E, will be figured later, and this supposition implies  $E \geq 5$ )
3. Discover the test measurement and p-value. Finding the test measurement includes a few stages. To start with the information is gathered and numbered, and afterward it is sorted out into a table (in a table every passage is known as a cell). These qualities are known as the observed frequencies, which the image for a observed frequencies is O. Every table is comprised of rows and columns. At that point every row is totaled to give a row total and every column is totaled to give a column total.

The null hypothesis is that the variables are independent. Utilizing the multiplication rule for independent occasions you can ascertain the probability of being one estimation of the main variable, A, and one estimation of the second variable, B (the probability of a specific cell  $P(A \text{ and } B)$ ). Keep in mind in a theory test, you accept that  $H_0$  is valid, the two variables are thought to be autonomous.

$$P(A \text{ and } B) = P(A) \cdot P(B) \text{ if } A \text{ and } B \text{ are independent} = \frac{\text{Number of ways A can happen}}{\text{total number of individuals}} \cdot$$

=

Now you want to find out how many individuals you expect to be in a certain cell. To find the expected frequencies, you just need to multiply the probability of that cell times the total number of individuals. Do not round the expected frequencies.

Expected Frequency (Cell A and Cell B) = E (A and B)

=

The symbol for chi-square [9] is  $\chi^2$

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

where O is the observed frequency and E is the expected frequency.

$\chi^2$  has diverse bends relying upon the degrees of freedom. It is skewed to the ideal for little degrees of freedom and gets more symmetric as the degrees of freedom expands. Since the test measurement includes squaring the distinctions, the test insights are all positive. A chi-squared test for independency is constantly right tailed p-value: Use  $\chi^2_{cdf}$  (lower limit, 1E99, df ) Where the degrees of freedom is df = (# of lines "1")\*(# of segments "1").

4. Conclusion : This is the place you compose reject Ho or neglect to reject Ho . The tenet is: if the p-value  $< \alpha$  , then reject Ho . On the off chance that the p-value  $\geq \alpha$  , then neglect to reject Ho
5. Interpretation: This is the place you interpreter in certifiable terms the conclusion to the test. The decision for a hypothesis test is that you either have enough confirmation to show HA is valid, or you don't have enough proof to show HA is valid.

#### 4. DATASET USED BY OUR PROPOSED FRAMEWORK

**Table 1: Input dataset of 4 X 4 Networks**

<i>Cell No</i>	<i>Movement Weight</i>	<i>Call arrival weight</i>
1	452	484
2	767	377
3	360	284
4	548	518
5	591	365
6	1451	1355
7	816	438
8	574	415
9	647	366
10	989	435
11	1105	510
12	736	501
13	529	470
14	423	376
15	1058	569
16	434	361

**Table 2: Input Dataset for 8 X 8 Network**

<i>Cell No</i>	<i>Movement Weight</i>	<i>Call arrival weight</i>	<i>Cell No</i>	<i>Movement Weight</i>	<i>Call arrival weight</i>
1	452	484	33	1357	1596
2	767	377	34	842	628
3	360	284	35	753	591
4	548	518	36	684	524
5	591	365	37	1068	486
6	1451	1355	38	951	573
7	816	438	39	519	357
8	574	415	40	624	842
9	647	366	41	245	652
10	989	435	42	584	685
11	1105	510	43	369	147
12	736	501	44	1234	987
13	529	470	45	379	816
14	423	376	46	618	937
15	1058	569	47	856	458
16	434	361	48	785	589
17	514	301	49	452	256
18	689	421	50	412	632
19	780	650	51	1047	1258
20	1205	854	52	1452	1236
21	1003	751	53	893	782
22	874	512	54	486	248
23	542	500	55	321	789
24	701	384	56	654	1032
25	321	789	57	879	1087
26	693	471	58	805	406
27	245	100	59	980	570
28	1478	369	60	740	684
29	552	348	61	654	450
30	987	1023	62	599	320
31	852	417	63	868	470
32	963	741	64	1010	585

**Table 3: A Novel Framework The GA Selects Reporting Cell In Dataset For 4 X 4 Networks**

<i>Cell No</i>	<i>Movement Weight</i>	<i>Call arrival Weight</i>
3	360	284
14	423	376
8	574	415
6	1451	1355

**Table 4: Selection of Reporting Cells for 8 X 8 Networks**

<i>Cell Number</i>	<i>Movement Weight</i>	<i>Call Arrival Weight</i>
9	647	366
18	689	421
37	1068	486
51	1047	1258
44	1234	987
26	693	471
61	654	450

## 5. EXPERIMENTAL RESULT AND ANALYSIS

### A. Result of Regression Analysis for given dataset 4X4 networks

**Table 5: Descriptive Statistics for the dataset 4 X 4 Networks**

	<i>Mean</i>	<i>Std. Deviation</i>	<i>N</i>
Cell No	8.500	4.7610	16
Movement Weight	717.500	301.2808	16
Call arrival weight	489.000	242.6778	16

**Table 6: Correlations of the 4 X 4 Networks**

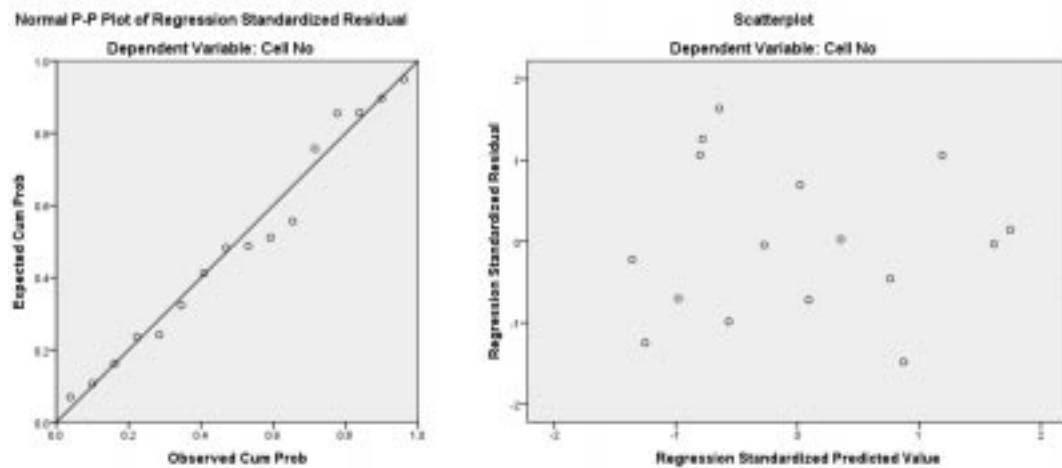
		<i>Cell No</i>	<i>Movement Weight</i>	<i>Call arrival weight</i>
Pearson Correlation	Cell No	1.000	.091	-.061
	Movement Weight	.091	1.000	.751
	Call arrival weight	-.061	.751	1.000
Sig. (1-tailed)	Cell No		.369	.412
	Movement Weight	.369		.000
	Call arrival weight	.412	.000	
N	Cell No	16	16	16
	Movement Weight	16	16	16
	Call arrival weight	16	16	16

**Table 7: ANOVA test result of 4 X 4 Networks**

	<i>Model</i>	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Sig.</i>
1	Regression	15.746	2	7.873	.316	.735 <sup>b</sup>
	Residual	324.254	13	24.943		
	Total	340.000	15			

**Table 8: Residuals Statistics for 4 X 4 networks**

	<i>Minimum</i>	<i>Maximum</i>	<i>Mean</i>	<i>Std. Deviation</i>	<i>N</i>
Predicted Value	7.107	10.293	8.500	1.0246	16
Std. Predicted Value	-1.360	1.750	.000	1.000	16
Standard Error of Predicted Value	1.251	4.792	2.004	.840	16
Adjusted Predicted Value	6.660	19.936	9.287	3.0575	16
Residual	-7.3934	8.1590	.0000	4.6494	16
Std. Residual	-1.480	1.634	.000	.931	16
Stud. Residual	-1.575	1.748	-.034	1.019	16
Deleted Residual	-13.9364	9.3398	-.7870	6.4136	16
Stud. Deleted Residual	-1.682	1.920	-.026	1.060	16
Mahal. Distance	.004	12.871	1.875	3.074	16
Cook's Distance	.000	2.390	.199	.587	16
Centered Leverage Value	.000	.858	.125	.205	16

**Figure 3a: Normal P-P Plot of Regression Standardized Residual chart of 4 X 4 networks and Scatterplot of Regression Standardized Residual for 4X4 Networks****B. Result of Regression Analysis for given dataset 8X8 networks****Table 9: Descriptive Statistics for 8 X 8 networks**

	<i>Mean</i>	<i>Std. Deviation</i>	<i>N</i>
Cell No	32.500	18.6190	64
Movement Weight	749.984	297.3300	64
Call arrival weight	595.750	293.2831	64

**Table 10: Correlations of the dataset 8 X 8 Networks**

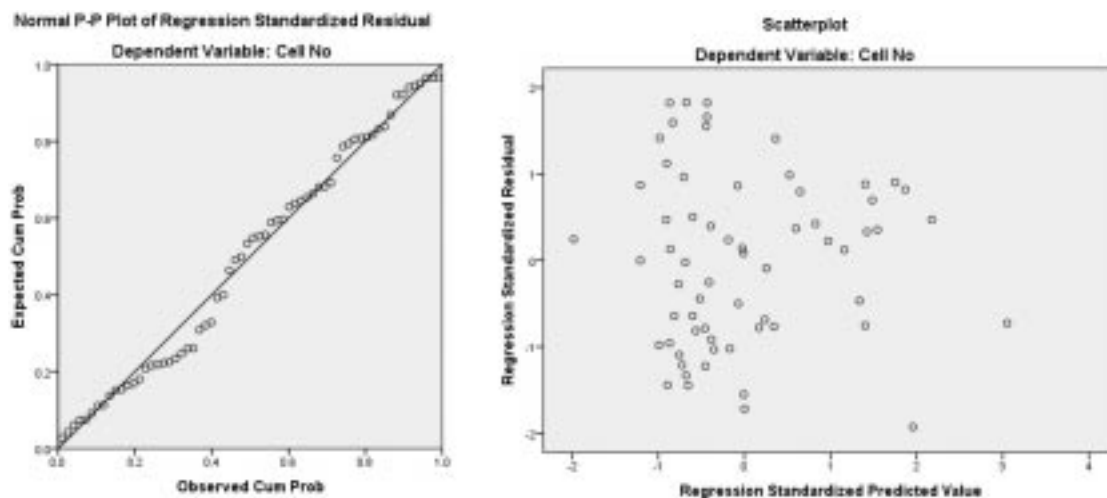
		<i>Cell No</i>	<i>Movement Weight</i>	<i>Call arrival weight</i>
Pearson Correlation	Cell No	1.000	.041	.225
	Movement Weight	.041	1.000	.527
	Call arrival weight	.225	.527	1.000
Sig. (1-tailed)	Cell No		.373	.037
	Movement Weight	.373		.000
	Call arrival weight	.037	.000	
N	Cell No	64	64	64
	Movement Weight	64	64	64
	Call arrival weight	64	64	64

**Table 11: ANOVA test result for 8 X 8 networks**

	<i>Model</i>	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Sig.</i>
1	Regression	1291.443	2	645.721	1.917	.156 <sup>b</sup>
	Residual	20548.557	61	336.862		
	Total	21840.000	63			

**Table 12: Residuals statistics for the dataset 8 X 8 networks**

	<i>Minimum</i>	<i>Maximum</i>	<i>Mean</i>	<i>Std. Deviation</i>	<i>N</i>
Predicted Value	23.530	46.321	32.500	4.5276	64
Std. Predicted Value	-1.981	3.053	.000	1.000	64
Standard Error of Predicted Value	2.295	8.301	3.743	1.344	64
Adjusted Predicted Value	22.380	49.683	32.538	4.8659	64
Residual	-35.3705	33.5480	.0000	18.0601	64
Std. Residual	-1.927	1.828	.000	.984	64
Stud. Residual	-2.083	1.853	-.001	1.006	64
Deleted Residual	-41.3039	34.5951	-.0382	18.9064	64
Stud. Deleted Residual	-2.143	1.892	.000	1.016	64
Mahal. Distance	.001	11.902	1.969	2.458	64
Cook's Distance	.000	.243	.016	.031	64
Centered Leverage Value	.000	.189	.031	.039	64

**Figure 5a: Normal P-P plot of Regression Standardized Residual chart of 8 X 8 networks, (b) Scatterplot of Regression Standardized Residual for the Dataset 8X8 Networks**

### C. Chi-Square Hypothesis Analysis for 4X4 Networks

The chi-square test is implemented on our proposed methods to know whether the methods will give effect when it is implemented for the location updating with reduced cost. The sample output for the chi-square test will be as follows:

**Table 13: Observed Frequency on the 4 X 4 Networks**

	<i>Location Management Metrics</i>			
	<i>Updation Cost</i>	<i>Paging Cost</i>	<i>Total Cost</i>	
Pre-Network for 16 cells	16	16	16	48
Proposed framework for 16 Cells	4	12	16	32
	20	28	32	80



**Table 14: Expected Frequency on the 4 X 4 networks**

	<i>Location Management Metrics</i>			
	<i>Updation Cost</i>	<i>Paging Cost</i>	<i>Total Cost</i>	
Pre-Network for 16 cells	12.000	16.800	19.200	48
Proposed framework for 16 Cells	8.000	11.200	12.800	32
	20	28	32	80

The number of degrees of freedom is calculated for an m-by-n table as  $(m-1)(n-1)$ , so in this case,  $(2-1)(3-1) = 1*2=2$ . To calculate the  $\chi^2$ , we then have a further table

**Table 15: The chi-square result on the 4 X 4 networks**

<i>Observed (O)</i>	<i>Expected (E)</i>	$ O-E $	$ O-E ^2$	$ O-E ^2/E$
16	12.000	2.667	7.111	0.533
16	16.800	-2.667	7.111	0.381
16	19.200	0.000	0.000	0.000
4	8.000	-2.667	7.111	1.067
12	11.200	2.667	7.111	0.762
16	12.800	0.000	0.000	0.000
Total				2.734

From the above table 15, at the significance level of 5%, since  $\chi^2 \geq$  critical value (i.e  $2.734 \geq 0.193$ ), it is concluded that the proposed framework for 4 X 4 networks rejects the null hypothesis ( $H_0$ ) and therefore it is significant to implement the proposed framework to reduce the total cost for 4 X 4 networks

#### D. Result of Hypothesis Analysis for Outcome dataset of 8 X 8 Networks

**Table 16: Observed frequency on the 8 X 8 networks**

	<i>Location Management Metrics</i>			
	<i>Updation Cost</i>	<i>Paging Cost</i>	<i>Total Cost</i>	
Pre-Network for 16 cells	64	64	128	256
Proposed framework for 16 Cells	7	24	31	62
	71	88	159	318

**Table 17: Expected Frequency on the 8 X 8 networks**

	<i>Location Management Metrics</i>			
	<i>Updation Cost</i>	<i>Paging Cost</i>	<i>Total Cost</i>	
Pre-Network for 16 cells	57.157	70.843	128.000	64
Proposed framework for 16 Cells	13.843	17.157	31.000	32
	20	28	48	96

The number of degrees of freedom is calculated for an m-by-n table as  $(m-1)(n-1)$ , so in this case,  $(2-1)(3-1) = 1*2=2$ . To calculate the  $\chi^2$ , we then have a further table .

**Table 18: The chi-square result on the 8 X 8 networks**

<i>Observed (O)</i>	<i>Expected (E)</i>	<i> O-E </i>	<i> O-E <sup>2</sup></i>	<i> O-E <sup>2</sup>/E</i>
64	57.157	6.843	46.823	0.819
64	70.843	-6.843	46.823	0.661
128	128.000	0.000	0.000	0.000
7	13.843	-6.843	46.823	3.383
24	17.157	6.843	46.823	2.729
31	31.000	0.000	0.000	0.000
Total				7.592

From the above table 18, at the significance level of 5%, since  $\chi^2 \geq$  critical value (i.e  $7.592 \geq 0.020$ ), it is concluded that the proposed framework for 8 X 8 networks rejects the null hypothesis ( $H_0$ ) and therefore it is significant to implement the proposed framework to reduce the total cost for 8 X 8 networks.

## 6. CONCLUSIONS

From the above results obtained in the table, it is concluded that the regression analysis on the input dataset of 8X8 network and 4X4 network generates the important cell point which gives more accurate prediction for location management. Thereby it reduces the paging as well as total cost. In the hypothesis analysis, the result obtained by using 4X4 network and 8X8 network rejects the null hypothesis, thereby it is concluded that the proposed technique using genetic algorithm gives better result than the existing techniques.

## REFERENCES

- [1] Yasir Mehmood, Carmelita Gorg, Maciej Muehleisen and Andreas Timm-Giel, "Mobile M2M Communication architectures, Upcoming challenges, application and future directions", *EURASIP Journal on Wireless Communication and Networking*, 2015.
- [2] Yun Chao Hu, Milan Patel, Dario Sabella, Nurit Sprecher, Valerie Young, "Mobile Edge Computing A Key Technology towards 5G", *White Paper-First Edition*, September 2015.
- [3] N. Shakeela and Dr. S. Thabasu Kannan, "A Novel Framework for Reducing the total cost through reporting call optimization in Mobile Computing", *IEEE Sponsored 3<sup>rd</sup> International Conference on Innovation in Information, Embedded and Communication Systems*, March 2016.
- [4] Jacek Malczewski, Claus Rinner, "Heuristic Methods", *Advances in Geographic Information Science-Springer*, 14-189, February 2015.
- [5] "Applied Regression Analysis: A Research Tool-Second Edition" by John O. Rawlings, Sastry G. Pantula, David A. Dickey.
- [6] "Multiple Regression: Inference for Multiple Regression and A Case Study" by *Freeman and Company*, 2009.
- [7] Vicki Sharp, "Chi-Square Test", *Statistics for the Social Sciences*.
- [8] Dr. Harry Benham, Gerard Carvalho, Michelle Cassens, "Student Perceptions on the Impact of Mobile Technology in the classroom", *Issues in Information Systems*, **15(11)**, 141-150, 2014.
- [9] Marie Diener, "Use of the Chi-Square Statistic", *Johns Hopkins University*.