

Identification and Modeling of Dominant Causes of Congestion In A Cellular Network

Akut E. K. B.Eng, M.Eng¹, S. M. Sani PhD, C. Eng, MIEE, MNSE²

¹(Department of Computer Engineering, Nuhu Bamalli Polytechnic, P. M. B. 1061. Zaria, Nigeria)

²(Department of Electrical and Electronics Engineering, Nigerian Defence Academy P. M. B. 2109. Kaduna, Nigeria.)

Abstract: In this paper, teletraffic data were collected from Airtel Nigeria Ltd. for a period of 18 months. Frequencies of occurrence of congestion in the network were derived from the teletraffic data using the Nigerian Communications Commission (NCC) benchmark. The dominant causes of congestion in the network were identified from the data and their value of occurrence was seen to be random in nature. A Poisson probability distribution model was therefore implemented for the prediction of congestion in the network. The identified four dominant causes of congestion and their percentage relative weights based on frequency of occurrence were: (i) X_1 : Faulty System (50.63%); (ii) X_2 : Faulty Trunk(s) (29.11%); (iii) X_3 : Power Failure (16.46%) and (iv) X_4 : Cable Cut (03.80%). MATLAB (M-file Data Acquisition Toolbox) programming environment was used to write a software program in C++ language that simulated the four dominant causes of congestion. Pearson Product Moment Correlation Coefficient formula was used to test the relationship between the measured and simulated congestion. The value $r = 0.7487$ was the correlation coefficient obtained between the measured and simulated results. This shows that the model is reasonably reliable.

Key words: Airtel, Congestion, Mobile network, Poisson, Probability.

I. Introduction

In telecommunications networks, congestion is the condition that arises when a system or a network experiences a level of offered calling activity or message traffic that exceed its capacity, [1]. In data networks, congestion occurs when the number of packets being transmitted through the network approaches the packet handling capacity of the network.

In cellular networks, congestion during peak periods and at hot spots is a major problem being faced today almost everywhere especially in urban areas and metropolitan cities. Increase in channel capacity may appear to be a natural solution to this problem. However, it is not economically viable due to the heavy cost of infrastructure involved [2]. Since the traffic demand is increasing on daily basis, the problem of congestion is going to persist. Consequently it needs to be effectively addressed.

1.1 Main Causes of Congestion

The main causes of congestion in the mobile telecommunications networks include:

(i) Irregularity in public power supply; (ii) Fiber cut, in the case of fiber optic link between the BSC and MSC or between MSC and MSC; (iii) Propagation impairments; (iv) Traffic build up (peak period); (v) Terminal equipment failure; (vi) Inadequate dimensioning of the switching network

1.2 Objectives of the Study

The objective of this work is in three fold: (i) to identify the dominant causes of congestion in a mobile network by using Messrs. Airtel (Nig) Ltd. telecommunications network in Kaduna as a case study; (ii) to develop a model for the prediction of congestion in the network; (iii) validated the model using Messrs. Airtel traffic data at the same location.

II. Literature Review

Some of the key studies undertaken by researchers on the causes and minimisation of congestion in networks are reviewed as follows:

2.1 Mr. Sriram [3] worked on the 'Dynamic Bandwidth Allocation and Congestion Control Scheme for Voice and Data Multiplexing In Wide Band Packet Technology' He presented results based on a simulation model which showed that the two schemes (novel bandwidth allocation method called T_1, T_2 -Scheme and voice block dropping scheme FIFO, together provide the desired performance in terms of: (i). very good voice quality; (ii). low delay and packet loss; (iii). efficient use of transmission bandwidth, and (iv). congestion control (protection in overload).

2.2 Mesrrs. David et al [4] reported on ‘Comparing Random Early Marking (REM) for a simple network with Random Early Drop (RED) and DropTail queue management system’. They optimized the approach to rate base flow control and made a performance comparison of the three internet Transport Control Protocol REM, RED and Drop Tail. The REM can achieve better throughput than RED and Drop Tail

2.3 Merrs. Yong et al [5] reported on ‘Accumulation-based Congestion Control.’ They generalised the Internet TCP Vegas and developed a model called the ACC FLUID MODEL, using (Monaco Linux scheme) accumulation. Their results showed that Monaco’s performance degrades with under provisioned buffer. Their reason was to detect and control network congestion.

2.4 Merrs. Chandrayana et al [6] researched into the ‘Uncooperative Congestion Control’ they proposed an analytical model for managing uncooperative flows in the internet by re-mapping their utility functions. They implemented the edge based re-maker in the Network Simulator (NS). They used Exponential Weighted Moving Average (EWMA) and Weighted Average Loss Indication (WALI). Their scheme works with any Active Queue System (AQM) scheme and also with Drop Tail queues. These works helps a great deal in controlling congestion, but, for more effective congestion management, the need for the prediction of congestion in telecommunications network is required.

III. Measurement Setup

A block diagram of the measurement setup Fig.1

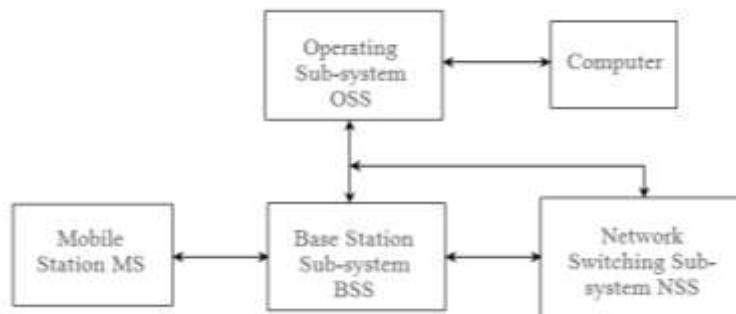


Figure 1. Block diagram of data collection setup. Source: Airtel (Nig) ltd.

Windows 2000 operating system is connected to the Operation Sub-System (OSS) which provides remote monitoring of the network performance and permits remote re-configuration. The computer is used to monitor and measure performance parameters of the traffic such as:

- (i) Call Set-up Success Rate (CSSR): Number of the unblocked call attempts divided by the total number of call attempts. Or $(1 - \text{Blocking Probability}) \times 100\%$. A call setup is an exchange of signaling information in the call process that leads to Traffic Channel (TCH) seizure. The Nigerian Communication Commission (NCC) standard for CSSR is $\geq 90\%$. [7][8]
- (ii) TCH Congestion Rate: This is the measure of congestion at the busy hour of the network when there is no congestion in the data network.
- (iii) SDCCH Congestion Rate: In data networks, Stand-alone Dedicated Control Channel congestion occurs when the number of packets being transmitted through the network approaches the packet handling [9] capacity of the network.
- (iv) Call Drop Rate: The Call Drop Rate is the number of dropped calls divided by the total number of call attempts. Or $(1 - \text{Call Completion Ratio}) \times 100\%$. A dropped call is a call that is prematurely terminated before being released normally by either the caller or called party (i.e., the call is dropped before the exchange of Released Message “RL_M” and Released Complete Message “RLC_M” in the signalling flow). The Nigerian Communication Commission (NCC) standard for CDR is $\leq 2.0\%$. [8][9]

Assuming that for any congestion to occur, the probability of any of these four causes surfacing is one, then the relative weights of the causes of congestion (with reference to the data collected), based on their relative occurrence and the class interval of congestion is given in Table 1[10]

Table 1 Weighted probabilities of dominant causes of congestion. Source: Airtel (Nig) ltd.

WEIGHTED PROBABILITY OF DOMINANT CAUSES OF CONGESTION				
Causes of Congestion	Denoted	Percentage Range		Percentage weight(s) (based on occurrence)
		CSSR (%)	CDR (%)	
Faulty System	X ₁	80 – 90	2.00 – 2.75	53.16
Faulty Trunk	X ₂	50 – 80	2.75 – 3.50	27.85
Power Failure	X ₃	0 – 30	4.00 – 100	16.46
Cable Cut	X ₄	30 – 50	3.5 – 4.00	03.80

The model is developed using the two major parameters namely (i) Call Set-up Success Rate (CSSR), and (ii) Call Drop Rate.

(CSSR) is the percentage of the number of the unblocked call attempts divided by the total number of call attempts:

$$CSSR = \frac{\text{Unblocked Call Attempts}}{\text{Total Number of Call Attempts}} \times 100\%. \tag{1}$$

The data collected reveals that there were 38 drops of CSSR below 90% within the sampling period. The threshold set by the Nigerian Communications Commission NCC for CSSR is $\geq 90\%$, [8]. Power failure or system failure mostly in the Base Station Sub-system (BSS), Network Switching Sub-system (NSS) or Operation Sub-System (OSS) causes congestion.

Measurement shows 40 records with Call Drop rate above 2.00% within the sampling period. Call Drop rate is given as:

$$\text{Call Drop Rate} = \frac{\text{Dropped Calls}}{\text{Total Number of Call Attempts}} \times 100\%. \tag{2}$$

The benchmark set by the Nigerian Communications Commission (NCC) for Call Drop Rate is $\leq 2\%$, [8]. High Call Drop Rate is caused by system failure, power failure in the Base Station Sub-system (BSS), Network Switching Sub-system (NSS) or Operation Sub-System (OSS) or excess subscribers accessing the network.

The readings of these parameters in a network give the measure of congestion.

IV. Development Of Model

4.1 Distributions of the Poisson Process

If you consider the arrival process of a call on a switching center as a point process i.e. considering arrivals within two non-overlapping time interval, Poisson process is used in a dynamical and physical way to derive the probability of this arrival. The probability that a given arrival pattern occurs within a time interval is independent of the location of the interval on the time axis.

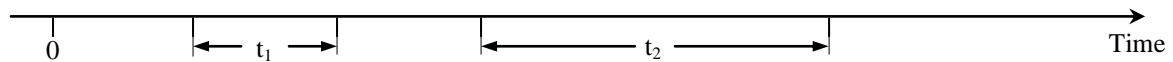


Figure 2 Deriving the Poisson process.

Let $p(v, t)$ denote the probability that v events occur within a time interval of duration t . The mathematical formulation of the above model is as follows:[11]

i. *Independence*: Let t_1 and t_2 be two non-overlapping intervals Fig. 2. Then because of the independence assumption we have:

$$p(0, t_1).p(0, t_2) = p(0, t_1 + t_2) \tag{3}$$

ii. (3) implies that the event “no arrivals within the interval of length 0” has the probability one:

$$p(0, 0) = 1 \tag{4}$$

iii. The mean value of the time interval between two successive arrivals is $1/\beta$:

$$\int_0^{\infty} p(0, t)dt = \frac{1}{\beta}, \quad 0 < 1/\beta < \infty \tag{5}$$

Here $p(0, t)$ is the probability that there are no arrivals within the time interval $(0, t)$, that is the probability that the time before the first event is larger than time t (the complementary distribution function).

iv. (5) implies that the probability of “no arrivals within a time interval of length ∞ ” is zero as it never takes place:

$$p(0, \infty) = 0 \tag{6}$$

Due to the stochastic nature of congestion occurrence as well as the large sampling size, the Poisson distribution function will be used to model the probability of congestion.

The modeling for this study will be based on the data obtained (May 2009 to October 2010).

Congestion is considered when the parameter readings {CSSR(%) and CDR(%)} crosses the benchmark given by the Nigerian Communications Commission (NCC). (i.e. 90% and 02% respectively).

Using the total element (sampling size):

$$n = 549 \tag{7}$$

Let congestion due to:

- | | | | |
|-----------------|----|-------|--|
| Faulty system | be | X_1 | |
| Faulty trunk(s) | be | X_2 | |
| Power failure | be | X_3 | |
| Cable cut | be | X_4 | where: X_1, X_2, X_3 and X_4 are random variables. |

The statistical mean of the Poisson distribution is given as: [12]

$$E\{X\} = \mu_x = \sum_{i=1}^n x_i P(X = x_i) = np \tag{8}$$

And the variance of Poisson distribution is given as:

$$E\{[X - \mu_x]^2\} = \sigma_x^2 = \sum_{i=1}^n (x_i - \mu_x)^2 P(X = x_i) = np \tag{9}$$

Let the probability that X within the sample space takes a value x_i which is independent of time and also independent of the next value x_{i+1} , then the probability of the four factors causing congestion, i.e. $X_j = x_i$ is given as $P(X_j = x_i)$.

Where $j = (1, 2, 3, 4)$ and $i = (1, 2, 3, \dots, 549)$

The rate of congestion occurrence of any of the factors of congestion is

$$P(X_1 = x_i) = 0.0747$$

$$P(X_2 = x_i) = 0.0401$$

$$P(X_3 = x_i) = 0.0237$$

$$P(X_4 = x_i) = 0.0055$$

Let these occurrences with their relative weights be given as

$$P(X_1 = x_i) = 0.0747 \times 0.5063 = 0.03782061$$

$$P(X_2 = x_i) = 0.0401 \times 0.2911 = 0.01167311$$

$$P(X_3 = x_i) = 0.0237 \times 0.1646 = 0.00390102$$

$$P(X_4 = x_i) = 0.0055 \times 0.0380 = 0.00020900$$

Summing the average outcome of the four factors that cause congestion, with respect to their relative weights, we have:

$$\sum_{j=1}^4 P(X_j = x_i) = \sum_{j=1}^4 p(X_j) = 0.05360374 \tag{10}$$

Using the following equation, the sum mean of the statistical means is:

$$\sum_{j=1}^4 np(X_j) = (np) = \mu = 549 \times 0.05360374 = 29.42845326 \tag{11}$$

Therefore, the Poisson model accepts

X_1, X_2, X_3, X_4 and n ,

Compute:

$$k = \sum_{i=1}^4 x_i * (\text{weight of } x_i) \tag{12}$$

$$P_{MN} = \frac{e^{-(np)} (np)^k}{k!} \tag{13}$$

Where $n = (1, 2, \dots, n)$ gives the number of days within which the probability is required

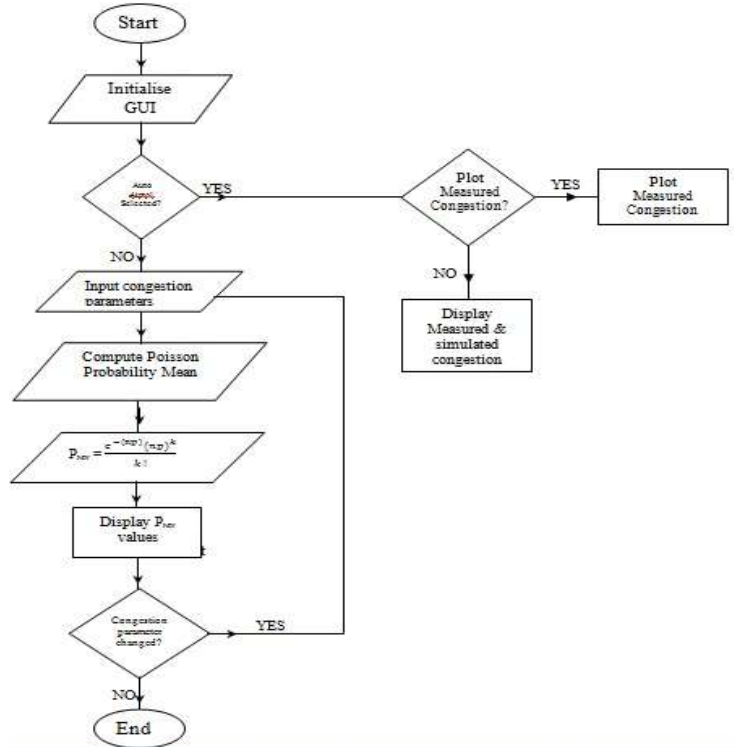
$$p = 0.05360374$$

The median of the distribution is given as: [12]

$$\begin{aligned} \text{Median} &\approx \left(\sum_{j=1}^4 np(X_j) + \frac{1}{3} - \frac{0.02}{\sum_{j=1}^4 np(X_j)} \right) \\ &\approx (29.42845326) + \frac{1}{3} - \frac{0.02}{29.42845326} \\ &\approx 29.762466 \end{aligned} \tag{14}$$

A program was written in C++ language using MATLAB (M-files Data Acquisition Toolbox) version 7 to implement (12) and (13).

The flowchart of the program is shown in Fig. 3. The flowchart begins by initializing the graphical user interface GUI, and then a selection of the task to be taken. A pop-up menu appears, requesting for the input parameters and the number of days (elements) to be entered. The program calculates the probability of congestion occurring in the network and also displays the probability of Transmission Channel (TCH) Congestion within these days. The program ends when there is no further entering of the congestion parameter.



V. Results

5.1 The measured congestion based on the data from Messrs. Airtel (Nig) Ltd. (see Appendix A) is shown in Fig. 4

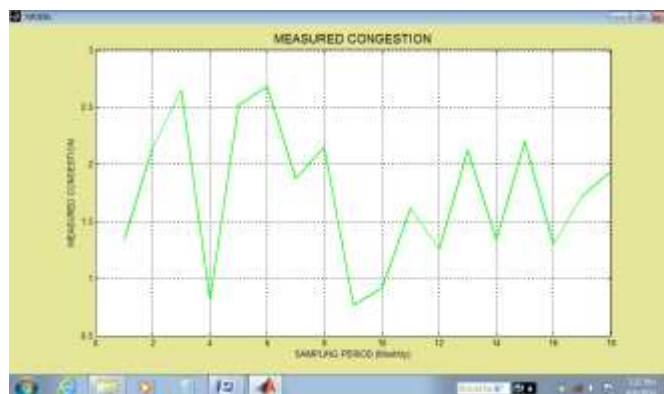


Figure 4. Measured Congestion

5.2 The simulated congestion based on (12) and (13) is shown in Fig. 5



Figure 5 Simulated Congestion

5.3 Comparison of Measured and Simulated Results.

The following equation is used to derive the monthly congestion:

$$\text{Measured congestion} = X_1 + X_2 + X_3 + X_4 \tag{15}$$

Where $X_1 = \sum_{i=1}^n (N_i * 0.5062)$

$$X_2 = \sum_{i=1}^n (N_i * 0.2911)$$

$$X_3 = \sum_{i=1}^n (N_i * 0.1646)$$

$$X_4 = \sum_{i=1}^n (N_i * 0.0380)$$

N_i = number(s) of occurrence in the month

In Fig. 6. the green graph depicted the measured congestion while the red graph depicted the simulated congestion using the model developed.

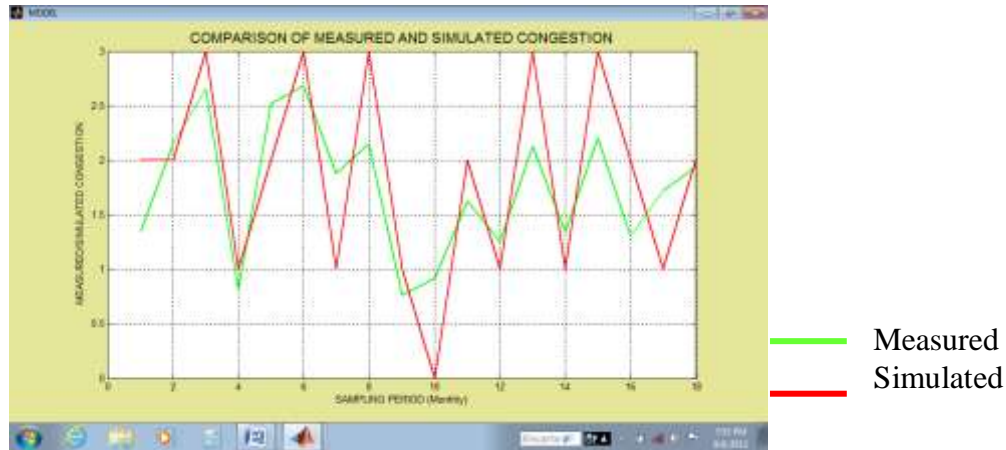


Figure 6. Comparison of Measured Congestion and Simulated Congestion

Reasons for low variation in the measured graph and high variation in the simulated graph is that the input values for X_3 and X_4 are reduced by the Poisson model due to their low weights and also the Poisson simulated output using MATLAB returns the smallest integer K such that the Poisson CDF evaluated at K equals or exceeds the probability value.

The Pearson Product Moment Correlation Coefficient formula was used to test the relationship between the measured congestion and simulated congestion. The following equation is used to obtain the correlation coefficient r of the two variations:[13]

$$r = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{\{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]\}}} \tag{16}$$

Table 5.1 Show values for the correlation coefficient r. X and Y are the measured and simulated congestions respectively. The table was used to derive the input parameters for (16).

Table 5.1 Values for the correlation coefficient formula r.

MONTH	X	Y	X ²	Y ²	XY
May'09	1.3417	2	1.800159	4	2.6834
Jun'09	2.1518	2	4.630243	4	4.3036
Jul'09	2.6580	3	7.064964	9	7.7040
Aug'09	0.8101	1	0.656262	1	0.8101
Sep'09	2.5178	2	6.339317	4	5.0356
Oct'09	2.6834	3	7.200636	9	8.0502
Nov'09	1.8733	1	3.509253	1	1.8733
Dec'09	2.1518	3	4.630243	9	6.4554
Jan'10	0.7640	1	0.583696	1	0.7640
Feb'10	0.9165	0	0.839972	0	0
Mar'10	1.6202	2	2.625048	4	3.2404
Apr'10	1.2526	1	1.569007	1	1.2526

Identification and Modeling of Dominant Causes of Congestion In A Cellular Network

May'10	2.1264	3	4.521577	9	6.3792
Jun'10	1.3417	1	1.800159	1	1.3417
Jul'10	2.2013	3	4.845722	9	6.6039
Aug'10	1.3021	2	1.695464	4	2.6042
Sep'10	1.7192	1	2.955649	1	1.7192
Oct'10	1.9367	2	3.750807	4	3.8734
	$\sum X =$ 31.3686	$\sum Y =$ 33	$\sum X^2 =$ 61.018178	$\sum Y^2 =$ 75	$\sum XY =$ 64.6942

Therefore

$$r = \frac{18(64.6942) - (31.3686)(33.0000)}{\sqrt{\{[18(61.018178) - (31.3686)^2][18(75.0000) - (33.0000)^2]\}}}$$

$$= \frac{129.3318}{\sqrt{(114.338138)(261.0000)}}$$

$$r = 0.7487$$

VI. Conclusion

From the data obtained from Messrs. Airtel (Nig) Ltd. the occurrences of the identified causes of congestion are random in nature and have a relative percentage weights of (i) X_1 : Faulty System (50.63%); (ii) X_2 : Faulty Trunk(s) (29.11%); (iii) X_3 : Power Failure (16.46%) and (iv) X_4 : Cable Cut (03.80%). Congestion occurrence appears to be high during festive period indicating that the switching processor is not adequately dimensioned to handle such traffic. A correlation coefficient of 0.7487 shows a close relationship between the measured and simulated congestions. This indicates that the prediction model is reasonable accurate.

References

- [1] Webster's New World Dictionary, copyright © 2008 by Wiley publishing inc. indianapolic, Indiana.
- [2] Dspace Vidyandhi Digital Library, University of Mysore <http://dspace.library.iitb.ac.in/jspui/handle/10054/426.html> (current March, 2009)
- [3] K. Sriram: "Dynamic Bandwidth Allocation and Congestion Control Scheme for Voice and Data Multiplexing In Wide Band Packet Technology". AT & T Bell Laboratories. Room 3H-607 Crawfords Comer Road Holmdel, New Jersey 07733. <http://www.gap.upv.es/index/php.html> (current April, 2009)
- [4] David L. & Steven L: "Random Early Marking for Internet Congestion Control" T melbourne information technologies australia, Department of EEE. The University of Melbourne Carlton, Victoria, Australia 3053. <http://www.kadhun.weebly.com/publications.html> (current April, 2009)
- [5] Yong X., David H., Kishore R., Arvind V. & Shivkumar K.: "Accumulation-based Congestion Control" Department of Electrical, Electronics & Computer Studies EECS. University of California, Berkeley, CA 94720. <http://www.arnetminer.org/viewpub.do.html> (current April, 2009)
- [6] Chandrayana K. & Kalyanaraman S. "Uncooperative Congestion Control" Rensselaer Polytechnic Institute, 110 Eight Street, Troy, NY 12180. <http://www.jiti.com/v09/jiti.v9n2.091.106.pdf> (current April, 2009)
- [7] Journal on "Telecommunication Transmission Engineering", 3rd ed., Vol. 2, Bellcore, Piscataway, New Jersey, 1991. pp 21 - 38
- [8] The Nigerian Communications Commission www.ncc.gov/index_e.htm (current, February, 2010)
- [9] Dropped Call (TCH Drop-SDCCH Drop) – TCH Drop Analysis <http://telefunda.blogspot.com/2010/10/dropped-calltch-drop-sdcch-drop.html> (current February, 2010)
- [10] Airtel Nigerian Limited. A Technical Report Handbook For Operations Department Kaduna, Nigeria. 1991
- [11] Handbook "Teletraffic Engineering" ITU-D Study Group 2. Geneva, January 2005 pp 104 – 105
- [12] "Poisson Distribution". From Wikipedia, the free encyclopaedia http://en.wikipedia.org/wiki/talk:law_%20poisson_distribution (Current April, 2011)
- [13] Razaq B. & Ajayi O. O. S. "Research Methods & Statistical Analyses" Haytee Press and Publishing Company Limited 154, Ibrahim Taiwo Road, Ilorin (2000) pp 151 - 155

Appendix

SUMMARY OF ALL THE CONGESTIONS IN THE DATA OBTAINED FROM MESSRS. AIRTEL (NIG) LTD.								
S/N	Call Set-up Success Rate (CSSR)				Call Drop Rate			
	Date	Value (%)	Date	Value (%)	Date	Value (%)	Date	Value (%)
1	10/05/09	85.56	23/09/10	87.57	4/05/09	3.03	18/08/10	6.65
2	16/05/09	86.95	2/10/10	40.00	4/06/09	2.95	19/08/10	0
3	4/06/09	84.74	19/10/10	74.87	12/06/09	2.05	9/09/10	0
4	20/06/09	73.94	30/10/10	88.83	20/06/09	2.15	23/09/10	0
5	28/07/09	87.57			14/07/09	2.04	1/10/10	2.25
6	29/07/09	88.25			16/07/09	2.70	2/10/10	3.33
7	18/08/09	68.92			28/07/09	2.43	21/10/10	3.25
8	1/09/09	85.81			1/08/09	2.57		
9	6/09/09	74.18			17/09/09	2.32		
10	9/09/09	0			26/09/09	3.94		
11	14/09/09	88.48			10/10/09	2.09		
12	23/09/09	0			28/10/09	2.34		
13	24/09/09	76.16			29/10/09	3.33		
14	10/10/09	70.09			1/11/09	2.57		
15	12/10/09	89.96			4/12/09	2.76		
16	27/10/09	69.45			15/01/10	5.07		
17	1/11/09	88.79			6/02/10	2.86		
18	16/11/09	87.02			20/02/10	3.67		
19	17/11/09	65.79			10/03/10	2.41		
20	3/12/09	87.31			11/03/10	2.79		
21	4/12/09	88.72			4/04/10	2.41		
22	15/12/09	76.15			16/04/10	3.02		
23	16/12/09	85.64			17/04/10	2.98		
24	1/03/10	89.92			28/04/10	5.53		
25	11/03/10	76.89			13/05/10	2.46		
26	8/05/10	86.22			23/05/10	2.21		
27	14/05/10	89.14			2/06/10	2.80		
28	17/06/10	89.81			8/06/10	2.57		
29	16/07/10	66.02			6/07/10	2.37		
30	19/07/10	89.87			17/07/10	4.19		
31	19/08/10	0			18/07/10	4.62		
32	3/09/10	89.85			30/07/10	2.11		
33	9/09/10	0			7/08/10	2.04		
34	23/09/10	0			12/08/10	2.94		