

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

User Plan Recommendation using Mobile call Log Analysis

Salina Adinarayana, K. Yogeswara Rao

Associate Professor, Dept. of I.T, Shri Vishnu Engineering College for Women, Bhimavaram, Andhra Pradesh, India

Assistant Professor, Dept. of CSE, Gitam University, Visakhapatnam, Andhra Pradesh, India

ABSTRACT: In recent times, We can observe that there is a tremendous increase in mobile users with the usage of smart phones ,i-pads and tablet PCs for efficient communication with respect to both individual and official communication strata. Studies on mobile users statistics showing that there is also an emerging research area of analysing this new type of user data. In this paper, we perform call log analysis on mobile user data, peak time load of the operator, offers proving by the operator and recommended a suitable plan for the customer. We also investigate how customers are using proposed plan and rating the operators.

KEYWORDS: Call log statistics; Ratings; Peak load Analysis; Recommendation;

I. INTRODUCTION

As the demand for mobile phone grows, Mobile operators are providing new plans for the customers to generate revenue for expanding their network capacity. With the drastic explosion in the quantity of calls made/received and data usage, there is a need for analyses of patterns in these call logs to assist the user a best plan to select. Based on the feedback from customers on different issues on mobile usage statistics like call and data usage, mobile data explosion, operators are searching for the ways to improve their architectures and radio access technologies to handle this growth without any problem. Switching to Long Term Evaluation (LTE), adding more radio towers or small-scale base stations such as femtocells or picocells and using WiFi access points for offloading the burden on their network [2], [3], [4], [5] are some of the solutions operators are currently working on. According to Hindu Business Line[13] dated August 12 Customers are floating to Reliance stores has been increased by at least three times during the past 20-25 days, with customers mostly seeking LYF handsets and 4G SIM cards because of Reliance Jio 4G services and most of these customer visits are also translating into sales.

II. LITERATURE REVIEW

In paper [6] the author has proposed different techniques are proposed and experimented for analysing network traffic including neural network based techniques to data mining technique. In paper [8] the author has proposed - mobility, speed and network load to make a comprehensive performance analysis of the mobile ad-hoc routing protocols. Three major performance parameters - average throughput of the network, packet delivery fraction and number of Packets dropped are considered to determine the performance of the Ad-Hoc Routing Protocols. In Paper[9] The author provides a systematic overview of the main studies and projects addressing the use of data derived from mobile network operators to obtain location and traffic estimations of individuals, as a starting point for further research on incident and traffic management

III. RECOMMENDATION FRAME WORK

Many networks providers in the market have their own strategies and business plans to attract customers. Now a days retaining customers is a challenging task for the providers. As the competitors are coming with different plans retaining customers in the network is not easy. More over number portability for the customers is an easy way to migrate to new network. In India, now a days the network providers like Airtel, Idea, BSNL, Docomo and Uninor are facing big competition from Reliance with the launch of reliance Jio 4G as per our analysis.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

In the proposed recommendation framework It has the following stages/steps with respect to different network providers.

- Collect information regarding the number of calls made
- Collect information regarding time of calls made(out going)
- After collecting call log statistics, estimate peak load using Navie Bayes
- Recommend a plan on the basis of peak load estimation
- Customer after migrated to the selected plan, collect feedback on the new plan selected
- Give rating to network provider based on performance analysis

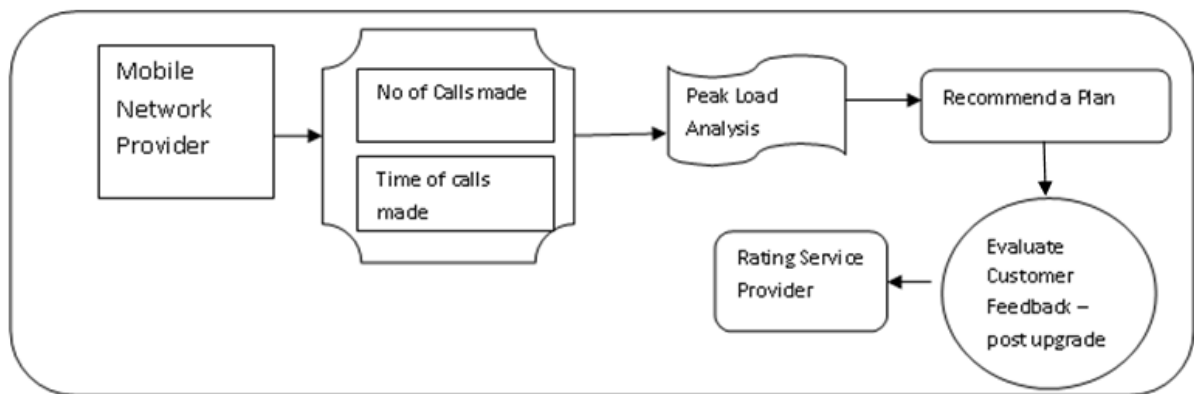


Figure 1: Plan Recommendation framework

IV. PLAN MIGRATION

Customers are too smart in selecting the network provider while they are migrating to a new plan in the existing network or during Number portability. To create awareness among customers and to retain customers the network providers regularly send details of offers and promotional plans through sms . An example offer-sms is shown in Figure-2.

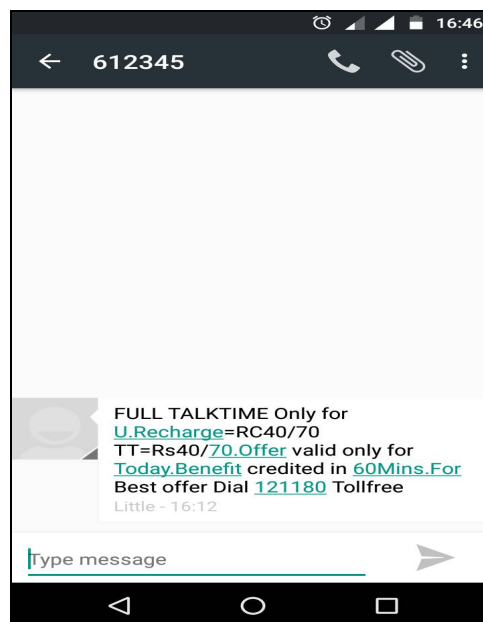


Figure 2: Offers-SMS

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

To study the call log we have used selected customers calls usage statistics as dataset. Based on the call log statistics of these customers shown below is a recommended offers of different network providers shown in table-1. Peak time scenario of each provider is illustrated in the figure-3.

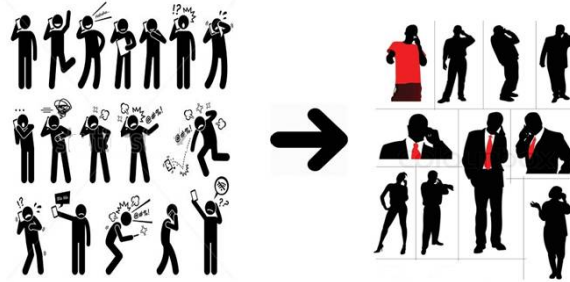


Figure-3. Identifying Peak time of each Provider-scenario

In the table-1 we have shown the recommended offers from the network operators for customers based on their mobile usage statistics. In table-1 we have represented the operator, mobile no, present plan, usage (outgoing duration) and recommended offers from the network operator. These offers are sent to customers periodically through SMS shown in figure-2, as and when the operators are proposed.

Table-1 Suggesting Plans Based On Usage

Sno	Operator	Plan	User	Outgoing Duration	Offers
1	Airtel	Bharti	7304637408	3229	Combopack(39/-) Nightpack(7/-) Tariffpvocher(170/-)
2	Airtel	Bharti	7641036117	41	Combopack(39/-) Nightpack(7/-) Tariffpvocher(170/-)
3	Airtel	Bharti	7640835719	52	Combopack(39/-) Nightpack(7/-) Tariffpvocher(170/-)
4	Airtel	Desi	7110730864	688	Combopack(47/-) Tariffpvocher(150/-) Combopack(85/-)
5	BSNL	Bhandan	7102745960	4083	0.015Rs/sec, 0.016Rs/sec, 0.013Rs/sec
6	BSNL	Bhandan	7102745961	22	0.015Rs/sec, 0.016Rs/sec, 0.013Rs/sec
7	BSNL	Nestam	7981267897	3629	0.013Rs/sec, 0.014Rs/sec, 0.012Rs/sec
8	BSNL	Nestam	7456622368	3834	0.013Rs/sec, 0.014Rs/sec, 0.012Rs/sec
9	DOCOMO	Gaurdian	7195393761	111	34/-,48days,1.6p/sec 61/-,56days,1p/sec 72/-,60days,1.2p/sec
10	DOCOMO	Gaurdian	7375968743	555	34/-,48days,1.6p/sec 61/-,56days,1p/sec 72/-,60days,1.2p/sec
11	DOCOMO	Raksha Plan	7839321359	160	18/-,56days,2p/sec 19/-, 90days,1.5p/sec 21/-,42days,1.6p/sec
12	IDEA	Stylish	7806391587	764	45/-,5days,35p/min 55/-,6days,36p/min 71/-,30days,45p/min
13	Uninor	Life Time	7651896055	503	26/-,14days,3000 Local calls 32/-,14days,unlimited 42/-,28days,5400 Local calls
14	Uninor	Short Time	7690700998	272	68/-,28days,9500 Local calls 109/-,14days,60,000 Local calls 229/-,28days,1,20,000Local calls

V. EXPERIMENTS AND RESULT ANALYSIS

This section explores the experimentation and results associated with the proposed framework. The results are processed using R language[7]. The dataset is prepared with live data of different users with 1000 entries which includes the parameters -caller, receiver, incoming/out-going, duration, date-time of call, operator-name, plan used and no. of messages of five operators namely Airtel, BSNL, Uninor, DOCOMO and Idea are taken in to consideration. A



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

sample of our dataset is shown here in table-2. The dataset is first pre-processed and then stored as .csv file and then analysed this csv file using R data processing [10].To perform data analysis we have used a set of standard statistical measures[11] and Navie bayes algorithm to estimate peak load. In search of statistical package or software for data analysis, I have reviewed pros and cons of different softwares[12] and finally we have used R software to evaluate the results of this work.

Table-2 Dataset for the framework

Sno	Caller	Receiver	incoming /out-going	Duration	Date-Time	Operator -name	Plan	No. of messages
1	7610039694	7434677419	Incoming	211	Wed Sep 15 19:17:44 +0100 2010	IDEA	Simple	666
2	7641036117	1666472054	Outgoing	31	Mon Feb 11 07:18:23 +0000 1980	Airtel	Bharti	1417
3	7641036117	7371326239	Incoming	45	Mon Feb 11 07:45:42 +0000 1980	Airtel	Bharti	1178
4	7641036117	7681546436	Outgoing	10	Mon Feb 11 08:04:42 +0000 1980	Airtel	Bharti	1105
5	7641036117	7681546436	Outgoing	0	Mon Feb 11 08:05:31 +0000 1980	Airtel	Bharti	380
6	7641036117	7681546436	Incoming	0	Mon Feb 11 08:06:18 +0000 1980	Airtel	Bharti	1555
7	7641036117	7981267897	Outgoing	0	Mon Feb 11 08:06:31 +0000 1980	Airtel	Bharti	1158
8	7641036117	7588304495	Incoming	124	Thu Sep 09 19:35:37 +0100 2010	Airtel	Bharti	1197
9	7981267897	7784425582	Outgoing	474	Thu Sep 09 18:43:44 +0100 2010	BSNL	Nestam	1942
10	7981267897	7743039441	Missed	0	Thu Sep 09 19:51:30 +0100 2010	BSNL	Nestam	161
...

Network load analysis is an important issue that has received much interest recently from computer network community. The network load analysis is one of the typical issues useful for monitoring network, network security, avoid congestion and increase speed of networks [6]. The peak load estimation of different network providers based on call log statistics and time of calling is shown in the graph (Figure-4). Here in the graph we have shown service provider wise analysis by taking the parameter no-of-call in Y-axis and the parameter time-of-calling in X-axis.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

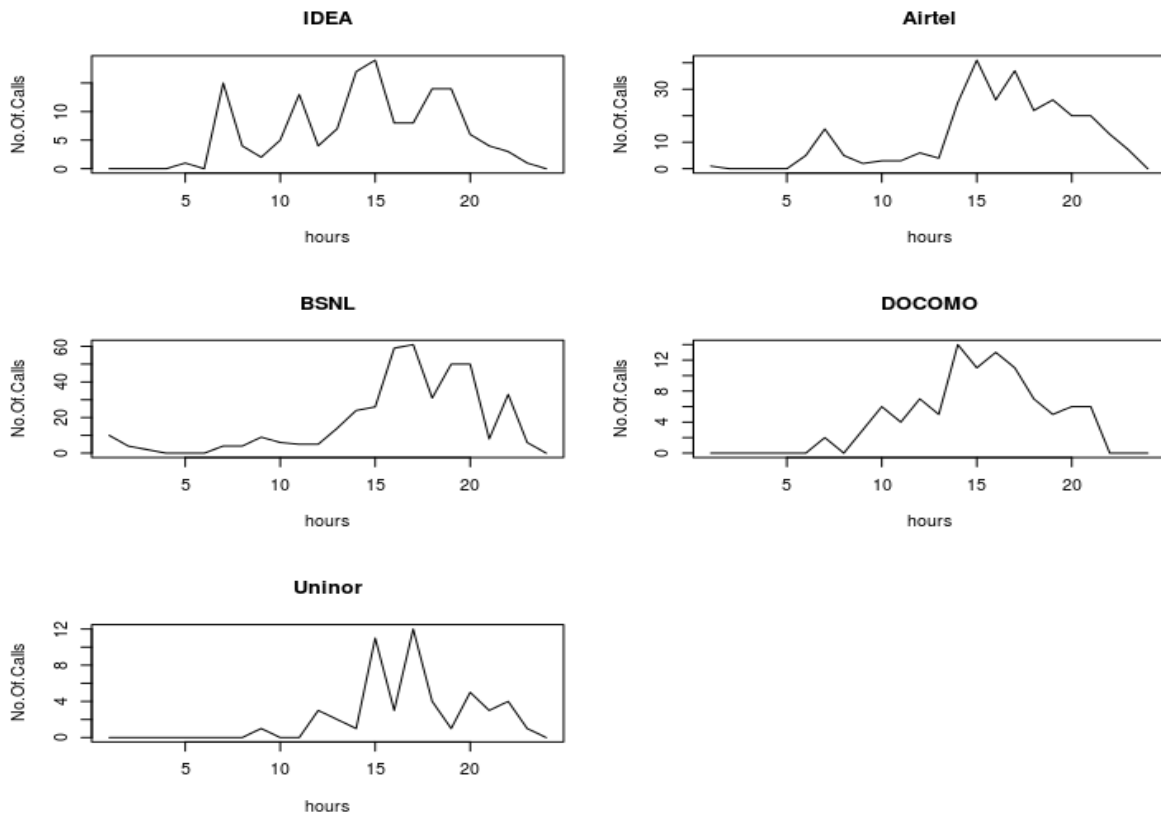


Figure-4: Peak load estimation graph

After the peak load estimation, the customer has ample number of recommendations out of which he can select a suitable plan for his mobile calls.

Based on the historical analysis of calls made by the users in different scenarios and after plan migration an analysis is conducted to suggest a new plan.

VI. EVALUATION METRICS

In the network peak load estimation, detection rate, false positive rate, accuracy and time cost metrics are used to measure the performance of the classifier on the proposed recommendation framework. During this process the standard classification algorithm naive bayes (NB) has used to estimate the peak load. Thereafter customers are classified network wise based on their usage patterns and recommend a better plan [1]. Performance analysis of the network providers is analysed after and before plan migration and given rating to the network provider based on call log statistics.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

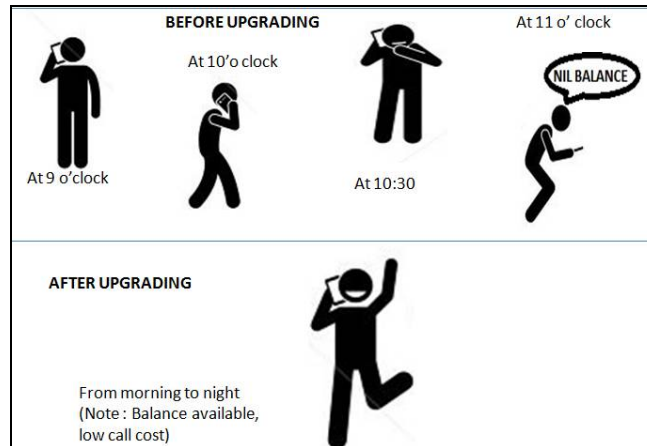


Figure-5.Before—After Plan Upgrading

In the figure-5 we have shown the clear difference between before and after plan up gradation and how the plan recommended is a better choice for the customers. Before plan up gradation, the customer has some problems w.r.t nil balance, call connection problems during peak network load. Whereas after upgrading to a plan suggested by the recommendation framework customer is enjoying the benefits of the plan in-terms of high connectivity and no worries on his pocket w.r.t charges. While recommending the new plan we have compared the present plan with new plan, if existing plan is better one, framework recommends the customer to continue with existing plan otherwise it suggests a new plan. This statistical analysis for plan recommendation is conducted using data of selected customers of BSNL, Uninor, DOCOMO, Airtel and Idea networks. These statistics are shown in Table-3. The table shows that some customers were upgraded to new plan where as other customers are already in higher plan so no need to change their plan.

Table-3 Upgraded Plan Statistics

	user	operator	Plan	direction	Update
1	7102745960	BSNL	Bhandan	143	Upgrade to Nestam
2	7102745961	BSNL	Bhandan	1	Upgrade to Bhandan if you have its ok
3	7304637408	Airtel	Bharti	118	Upgrade to Desi
4	7640835719	Airtel	Bharti	2	upgrade to Bharti if you have its ok
5	7641036117	Airtel	Bharti	4	upgrade to Bharti if you have its ok
6	7110730864	Airtel	Desi	30	You were already in the highest plan
7	7195393761	DOCOMO	Gaurdian	2	Upgrade to Gaurdian if you have its ok
8	7375968743	DOCOMO	Gaurdian	8	Upgrade to Gaurdian if you have its ok
9	7651896055	Uninor	Life Time	10	You were already in the highest plan
10	7456622368	BSNL	Nestam	116	You were already in the highest plan
11	7981267897	BSNL	Nestam	14	You were already in the highest plan
12	7946912731	Airtel	No Plan	29	upgrade to Bharti if you have its ok
13	7784425582	BSNL	No Plan	7	Upgrade to Bhandan if you have its ok
14	7817341678	DOCOMO	No Plan	35	Upgrade to Gaurdian if you have its ok
15	7607124303	IDEA	No Plan	42	Upgrade to Simple if you have its ok
16	7541486388	Uninor	No Plan	2	upgrade to Short Time IF you have its ok
17	7839321359	DOCOMO	Raksha Plan	11	You were already in the highest plan
18	7690700998	Uninor	Short Time	10	upgrade to Short Time IF you have its ok
19	7163185791	IDEA	Simple	18	Upgrade to Simple if you have its ok
20	7610039694	IDEA	Simple	15	Upgrade to Simple if you have its ok
21	7806391587	IDEA	Stylish	26	You were already in the highest plan

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

As per the recommendation frame work after migrating or upgrading to suitable plan, how effectively the new plans are using by the customers is evaluated. Feedback survey is conducted on “efficient usage of upgraded plans” with customers. The survey statistics of top three network providers in terms of efficient usage of the proposed plans is shown in figure-6.

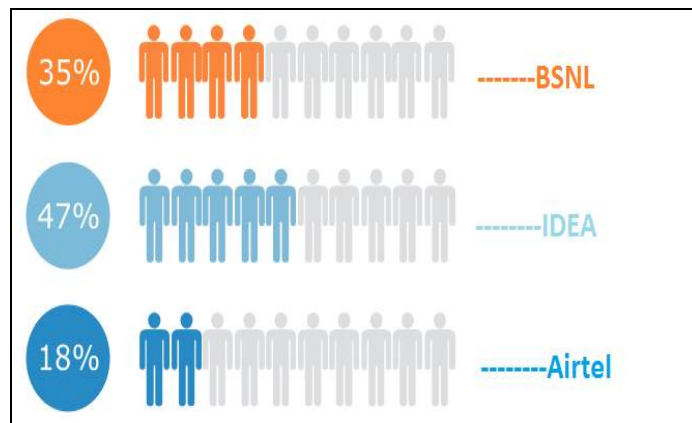


Figure-6: Efficient usage - proposed plan Statistics

After post usage of upgraded plan customers, ratings of the network operators is evaluated based on no of calls made and call-duration. A comparative analysis all five operators in terms of rating is shown in figure-7. As per this comparative analysis of rating BSNL scores top with 45% and DOCOMO scores least with 5%.

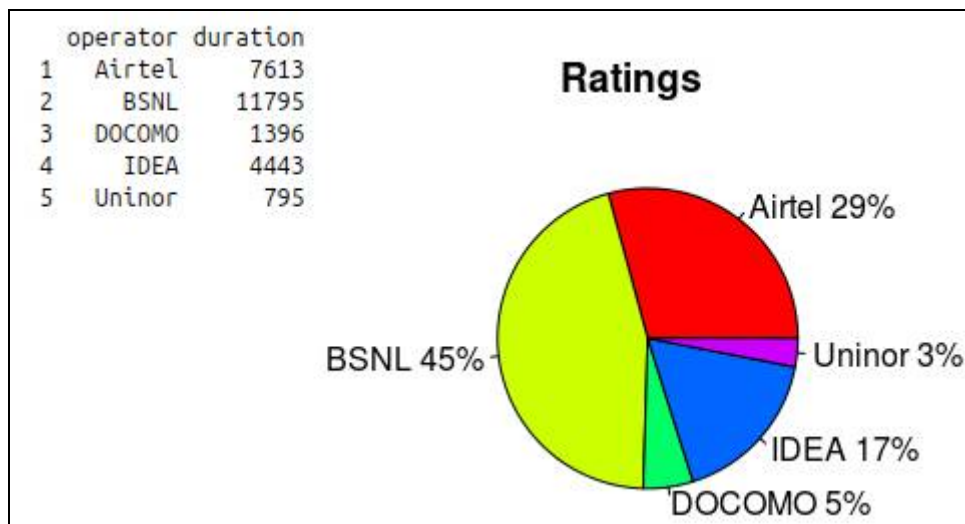


Figure-7: Network operators rating

VII. CONCLUSION

The simulation results showed that the proposed recommendation work well with five operators user data statistics taking different parameters into consideration. We have conducted this work with five operators with few parameters like call duration, peak time, no. of messages, present plan, caller, and receiver for analysis. In future work we will extend this paper with mobile apps usage statistics to classify the mobile users with a hybrid classification technique with more number of parameters and recommend suitable plan to enhance satisfaction level in mobile apps usage.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

REFERENCES

1. Lowd, Daniel, and Pedro Domingos. 'Naive Bayes models for probability estimation.' Proceedings of the 22nd international conference on Machine learning. ACM, 2005.
2. E. Bulut, and B. Szymanski, 'WiFi Access Point Deployment for Efficient Mobile Data Offloading' Proceedings of ACM International Workshop on Practical Issues and Applications in Next Generation Wireless Networks (PINGEN 2012) at Mobicom 2012., Istanbul, Turkey, August 26, 2012, pp. 45-50.
3. K. Lee, I. Rhee, J. Lee, S. Chong and Y. Yi, 'Mobile data offloading: how much can WiFi deliver?' Proceedings of the 6th ACM International Conference, ser. Co-NEXT 10, pp. 26:126:12, New York, NY, 2010.
4. S. Dimatteo, P. Hui, B. Han and V. O. K. Li, 'Cellular Traffic Offloading through WiFi Networks' Proceedings of 8th IEEE Intl Conf. Mobile Ad Hoc and Sensor Systems (IEEE MASS 11), 2011.
5. B. Han, P. Hui, and A. Srinivasan, 'Mobile data offloading in metropolitan area networks' SIGMOBILE Mob. Comput. Commun. Rev., vol. 14, pp. 2830, November 2010.
6. Joshi, Manish, and Theyazn Hassn Hadi. 'A Review of Network Traffic Analysis and Prediction Techniques.' arXiv preprint arXiv:1507.05722 (2015).
7. Rossiter, D. G. 'Introduction to the R Project for Statistical Computing for use at ITC.' International Institute for Geo-information Science & Earth Observation (ITC), Enschede (NL) 3 (2012): 3-6.
8. N.Sarmah, Y. Yang, H. Sharif, Y. Qian, 'Performance Analysis of Mobile Ad-Hoc Routing Protocols By Varying Mobility, Speed And Network Load' Masters Thesis, University of Nebraska-Lincoln, July 2014).
9. Steenbruggen J, Borzacchiello MT, Nijkamp P, Scholten H, 'Mobile phone data from GSM networks for traffic parameter and urban spatial pattern assessment: a review of applications and opportunities' GeoJournal(2011):78:1-21
10. Lumley, Thomas 'R Fundamentals and Programming Techniques', (2006).
11. Dennis D. Wackerly, William Mendenhall III, Richard L. Scheaffer, 'Mathematical Statistics with Applications' http://www.cs.cmu.edu/~bhiksha/courses/math.11691/fall2015/docs/mathematical_statistics_with_applications.pdf.
12. Abatan S. Matthew , Olayemi Micheal Sunday, 'The Role of Statistical Software in Data Analysis' International Journal of Applied Research and Studies (IJARS) ISSN: 2278-9480 Volume 3, Issue 8 (August - 2014)
13. Reliance Jio's launch hype -LYF phone sales spurt 'http://www.thehindubusinessline.com/info-tech/mobiles-tablets/reliance-jios-launch-hype-sees-lyf-phone-sales-spurt/article8981783.ece'.