An Analysis of the Socio-Economic Factors using Cellular Network Data

Thapelo Pholo, Karim Djouani and Yskander Hamam
Department of Electrical Engineering and Electronic Communications
Faculty of Electrical Engineering and Built Environment, Tshwane University of Technology,
P. O. Box 1020, Pretoria 0001
Tel: +27 12 382 4191, Fax: +27 12 382 4191
Email: pholot@gmail.com; {Djouani, Hamam}@tut.ac.za

Abstract-This paper analyses the social consumption behavior of subscribers in a major cellular network against the socio-economic data in South Africa. A central issue in this paper is to model the consumption versus socio-economic data and examine the correlation between the social data consumption and the socio-economic factors between two distinct geographical conditions; one an urban and the other a semi-urban.

Research has shown that using the mobile cellular network data set, we can be able to present a comparative analysis of the behavioral factors between users in urban and semi-urban societies. We consider the use of Cellular network data to model such a correlation that includes the daily load variations as well as user mobility. The study will also be able to shed light in the wide gap between connectivity and utilization between urban and semi-urban areas.

Index Terms—socio-economic factors, resource allocation, and social data.

I. INTRODUCTION

In recent years, the wireless cellular networks have witnessed what the industry has termed the "wireless revolution." In South Africa, the number of smartphones has reportedly more than doubled over the past twelve (12) months. Cellular network users have become 'data-hungry' as more and more applications are being developed and utilized daily. Almost each an every smartphone user accesses the internet and watches videos, accesses social networks frequently as they carry them everywhere. The services that are available to mobile users are growing exponentially and hence, mobile multimedia has become a reality.

The inability to accommodate this sudden increase in traffic is the problem facing the mobile network operators. Recently, this inability to serve subscribers is being strongly criticized and often comes with a penalty of some sort from the network regulator. The dilemma with the network operators is the trade-off between resource allocation in terms of the installation costs and the quality of service (QoS) as well as the network capacity enhancements.

Driven by these facts, we present the analysis and propose the use of cellular data to present and analyze how the Radio Resource Management (RRM) together with data analysis can be used to increase the efficiency of cellular networks, offer better incentives to customers and also maximize the revenue generated by the operator in terms of marketing strategies as a result of the user behavioral factors. We cannot rule out the fact that the data analysis can also be used to test the behavioral patterns that differentiate the urban, semi-urban and rural data consumption as well as the relative movement between them.

The techniques used to increase and optimize the efficiency of cellular networks are compared by focusing on the key Performance Indicators (KPIs) such as Handover success, call setup success rate and call drop rate etc. [1]. In addition to this, one of the most important factors which is the basis of this study is the correlation study. The study looks closely at the data and determines the social network data consumption. Correlation can most certainly be used to make informed decisions even though is not causation.

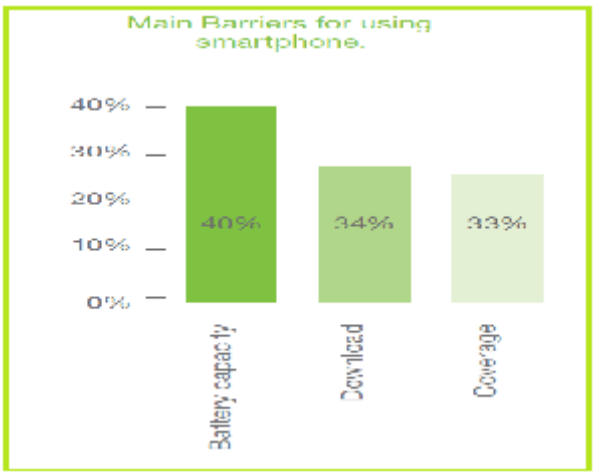

Figure 1. Main barriers for using smartphone [6].

II. DATA AND METHODOLOGY

Cellular network operators can use their Call Records Data (CRD) and analyze it to add business value to their operation. The CRD can be used to observe a precise state of the network and better understand the calling patterns of their subscribers in order to give users customized services in the pursuit of increasing the revenue.

There are a number of powerful tools that are used in cellular communications to analyze the caller information. The tools can be used to correlate and analyze the relationship between the urban, semi-urban and rural areas spread of subscribers with the use of KPIs.

The operator generates huge amounts of data and this is a serious challenge to analyze and hence our proposal to introduce a correlation model to analyze the raw data to

contribute to the informed decision making.

The network monitoring tools are statistical in nature and analyze the network based purely on the radio engineering parameters, Figure 2 and 3 refers. Our model seeks to analyze purely from the subscriber's point of view; subscriber location, used resources, social data, socio-economic factors, daily load variations as well as subscriber mobility.

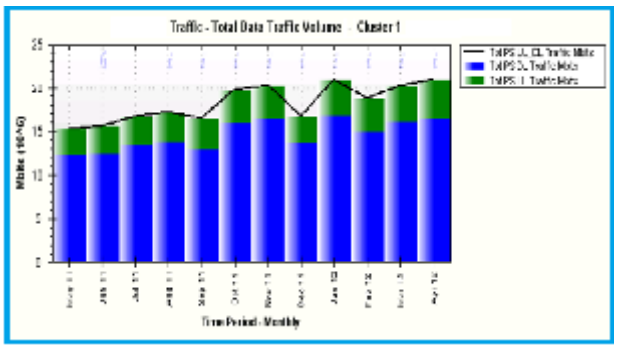

Figure 2 Traffic volume for a particular geographic location

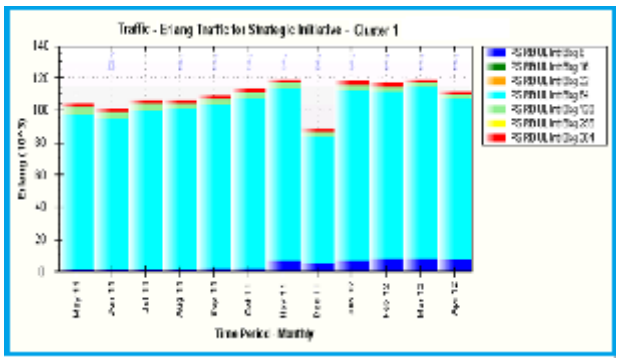

Figure 2 Traffic volume in Erlang

III. Proposed Model

The telecommunications demand model proposed by Taylor and Kridel in [1] will be used as the basis for the proposed approach in this study. However, the model will exclude the public-phone and fixed line as well as the gender variable and include the social data consumption, socio-economic factors and load variations, geographic location and mobility. We will use two sets of four weeks data records which will include all phone calls and short messaging services as well as other social data for the duration of the two months period, figure 3 refers.

The calls will be separated into three different categories; working hours, leisure hours and night-time hours. Calls made on weekends will also be distinguished. This will be able to show the characteristics of callers in these different times.

A. Geographic

The data will be divided into segments or regions based on the geographic status or location of the users; urban, semi-urban and rural regions. This will be based on the fact that a user is deemed to be lodging at a particular region by the majority of the time spent based on the cellular network base transceiver station in a week.

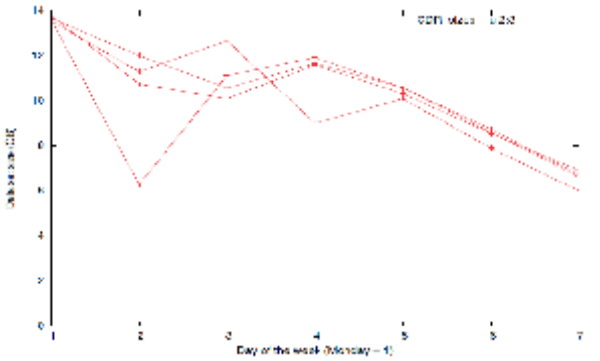

Figure3 Daily Call records data set for a month against day of week [5].

B. Socio-economic factors

Most of the customers will be based in urban areas and the least number will be in rural areas. Users in urban areas are economically advantaged than those in rural areas and hence we presume that a higher traffic and revenue will be generated in urban areas.

IV. Conclusions

This work presents an analysis as well as the model of how the CRD can be used to provide information into the socio-economic factors that influence the behavior of social data consumption between the urban, semi-urban and rural classes of cellular network customers. The work can also be used for marketing purposes as well as by other national departments like transportation.

V. References

1. Sofoklis A. Kyriazakos George T. Karetsos. "Practical Radio Resource Management in Wireless Systems". Artech House, Inc. Norwood, MA, '04.
2. Chabossou, A., Stork, C. and Zahonogo, Z. "Mobile Telephony Access and Usage in Africa", 2009.
3. F. Hollinger and M. Haller, "Kinship and social networks in modern societies: a cross-cultural comparison among seven nations", European Sociological Review 6:2, pp. 103-124, 1990.
4. M. Gonzalez, C. Hidalgo, and L. A. Barabasi, "Understanding individual human mobility patterns", Nature 453, pp. 779-782, 2008.
5. S. Hill, F. Provost, and C. Volinsky, "Network-Based Marketing: Identifying Likely Adopters via Consumer Networks", Statistical Science, 21:2, 256-276, 2006.
6. Scott Kirkpatricka, Alex Kulakovskya, Manuel Cebrianbc and Alex Pentland, "Social networks and spin glasses" UCSD, CA, 2011.
7. Smartphone usage in South Africa; An interesting trend. Accessed on 2012/06/10: http://www.formfunction.co.za/2011/12/27/smartphone-usage-in-sa-interesting-trends/

Thapelo Pholo received his M-Tech degree in 2008 and is presently studying towards his D-Tech degree at the Tshwane University of Technology (TUT). His research interests include optimization, Game theory and control systems.