Hybrid Behavioural Features for Churn Prediction in Mobile Telecomm Networks with Data Constraint

Naimisha Kolli, Varshini Reddy, and N Balakrishnan

I. Introduction

As the competition in the mobile telecommunications market is burgeoning constantly, customers look for lucrative pricing and high quality service and, thus will not hesitate to switch providers if they are satisfied with the current services they receive. In today's mobile telecommunications market, customers want competitive pricing and high quality service, and hence do not hesitate to switch providers, a phenomenon known as churning. As the cost of customer acquisition is far greater than that of customer retention, telecomm operators focus their marketing strategies on targeted customer retention campaigns. This requires building an accurate churn prediction model for identifying customers most prone to churn. The results of these analyses can be used by telecomm operators to improve targeted marketing, product design, and detect possible causes that lead to churn and potential fraud.

The primary idea is to create a customer profile from various data sources available for modeling including call patterns, contractual information, billing, payment, customer service, demographic profile and credit card data and then predict the probability of churn based on these attributes of the customer [1,2,3,4]. An obvious drawback of these approaches is that they focus exclusively on the individual customer profile and require access to numerous other sources of information apart from Call Data Records (CDRs). More importantly, these models do not take into account any social influence. Another line of research [5,6,7] deals with the underlying social network in a mobile call graph where the decision to churn is based solely on social influence. They show that reasonable prediction performance can be achieved by using only link information.

Behavioural models help to obtain a diagnostic view of the underlying structure to predict a specific action. Applications of behavioural models have assumed grave significance in the area of information security, spanning from malware detection to network and IoT security. In our present work, we recognize the importance of the role-played by social ties to better understand the underlying behavior of customers, and incorporate a novel feature of the social aspects of customers' social group along with the traditional individual customer profiles to extract hybrid behavioural novel features with potential practical implications.

As precision plays an important role in churn prediction, one of the main ways of improving prediction performance is with the use of appropriate feature sets that have high predictive power. The prediction model built is based on features that can be extracted from Call Detail Records (CDRs) exclusively. Throughout this work no demographic or contractual information was used, regardless of the data available being restricted to only a period of one month. In presence of these data constraints, we will show through our experiments that the proposed model achieves good prediction performance using the hybrid behavioural feature set extracted from the CDRs as well as mobile social graphs. To evaluate the feasibility of new features, we performed experiments using gradient boosted trees algorithm on the customers for identifying potential churners and report the Receiver Operating Characteristics (ROC) curves and Precision-Recall (PR) curves with and without these novel features. The results demonstrate the improved prediction performance with the inclusion of these combined features.

Naimisha Kolli (e-mail: naimishak@iisc.ac.in)
Varshini Reddy (e-mail: varshinib@iisc.ac.in)
N Balakrishnan (e-mail: balki@iisc.ac.in)

II. Data Description

In this paper we analyze the Call Data Records (CDRs) provided by one of the largest mobile operator in India. The data contains detailed records of calls made by a segment of customers active during a one-month period in a large metropolitan city in India. The data set consists of 400+ million CDR entries. Each CDR contains all the details pertaining to a call such as the time, duration, call type, cell tower, originating number, destination number, etc. of the call as maintained by the mobile operator for billing purposes. For our experiments, we have identified 106,002 non-churners. Moving forward, we have considered two cases for churner subsets:

Case I — Churners are those customers who have been active for the first two weeks, but inactive for the next two weeks. The number of customers are 89,745.

Case II — Churners are those customers who have been active for the first three weeks, but inactive for the last (fourth) week. The number of customers are 97,859.

III. Churn Prediction Modelling

The following block diagram, i.e. Fig. 1, gives an overview of our churn prediction model. Based on the information contained in the CDRs regarding who called whom, time at which the call was made, we have constructed a mobile call graph for each day in our dataset with customers as nodes and the calls between the customers as edges. We use the Louvain community detection algorithm [8], which is one of the best methods to efficiently partition graphs with millions of nodes to get the communities in our customer mobile graphs for each day in the dataset.

The feature extraction phase is divided into two sets. The first feature set relates to the features used to describe call patterns of a customer by aggregating his call usage while the second set of features correspond to the social group or community related features of the customers extracted from the previously constructed mobile call graphs. Based on the premise that changes in call usage and social groups are likely indicators for churn, we have calculated the rate of change in each of the above mentioned features using a mathematical formula defined as 'Change(f)' in Fig. 1 for a feature f and customer c, at time interval i, and a small constant ε. Features calculated in this manner will be able to capture subtle changes in behavioral patterns and thus possibly improving the performance of churn prediction. We have randomly split the dataset into training and testing datasets in 80:20 ratios for both Case I and Case II. We build a gradient boosted tree classifier to learn the model to classify potential customers who might churn.

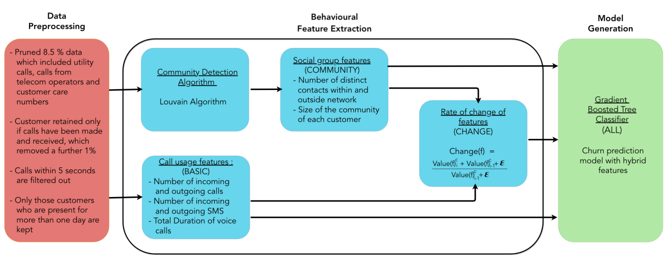

Fig. 1. Churn Prediction Modelling - Block diagram

IV. Experiment and Results

We build various models to evaluate the effectiveness of our hybrid behavioural feature subsets including call usage patterns, social group features and the rate of change in these features in predicting churn from customer behavioral patterns. For each case considered, we build models based on call usage features only (BASIC), combined call usage and changes in call usage features (BASIC+CHANGE), hybrid basic model with call usage combined with community or social aspect features (BASIC+COMMUNITY) and combined hybrid model with call usage, community and changes in both call usage and community features (ALL) and report the performance of these models for both Case I and Case II.

In both the cases considered it can be seen that the combined hybrid model with call usage, community and changes in call usage features (BASIC+COMMUNITY), the distribution of churners from non-churners is better observed than BASIC models. With the addition of rate of change features, we see the model predicting the churners with increased confidence probabilities for both Case I and Case II depicted by Fig. 2 and Fig. 3 respectively. Also, the distribution of churn probabilities are better separated in the BASIC+CHANGE model, which seem to point at improved class separation in the learned model when the novel hybrid features of both call patterns and social group features are considered for classification.

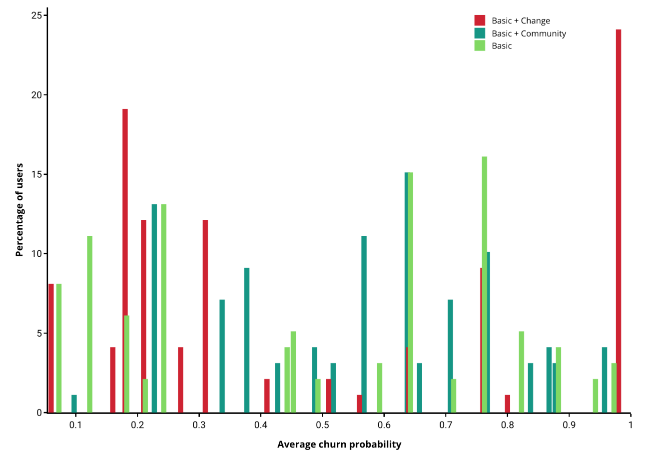

Fig. 2. Average churn probability plots for BASIC, BASIC + COMMUNITY and BASIC + CHANGE

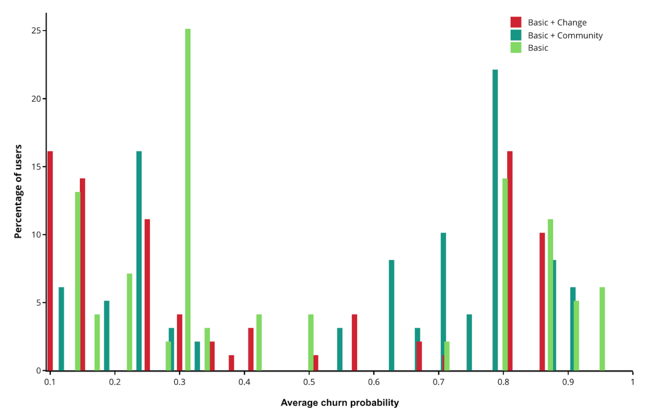

Fig. 3. Average churn probability plots for BASIC, BASIC + COMMUNITY and BASIC + CHANGE

From Fig. 4 and Fig. 5 for Case I and Fig. 6 and Fig. 7 for Case II, we can discern significant improvement in the BASIC+COMMUNITY model ROC curve and PR curve over the BASIC model ROC curve and PR curve. This shows that adding the social group based features to the call usage patterns gives a significant improvement in prediction performance and that establishes the social influence as a key indicator for churn in customers. It can also be observed that there is a significant improvement in the ALL model ROC curve and PR curve over the BASIC+COMMUNITY model ROC curve and PR curve. This shows the improved predicted power of our hybrid feature sets and significance of using the changes in community or social group sizes for predicting churn.

In Fig 4. BASIC model ROC (AUC=0.58) curve, BASIC+COMMUNITY model ROC (AUC=0.86) curve and ALL model ROC (AUC=0.98) curve

In Fig 5. BASIC model PR curve (Precision=0.53, Recall=0.57 at threshold=0.5), BASIC+COMMUNITY model PR curve (Precision=0.71, Recall=0.88 at threshold=0.5) and

ALL model PR curve (Precision=0.94, Recall=0.89 at threshold=0.5)

In Fig 6. BASIC model ROC (AUC=0.81) curve, BASIC+COMMUNITY model has ROC (AUC=0.89) curve and ALL model ROC (AUC=0.98) curve

In Fig 7. BASIC model PR curve (Precision=0.76, Recall=0.69 at threshold=0.5), BASIC+COMMUNITY model PR curve (Precision=0.76, Recall=0.96 at threshold=0.5) and ALL model PR curve (Precision=0.89, Recall=0.96 at threshold=0.5)

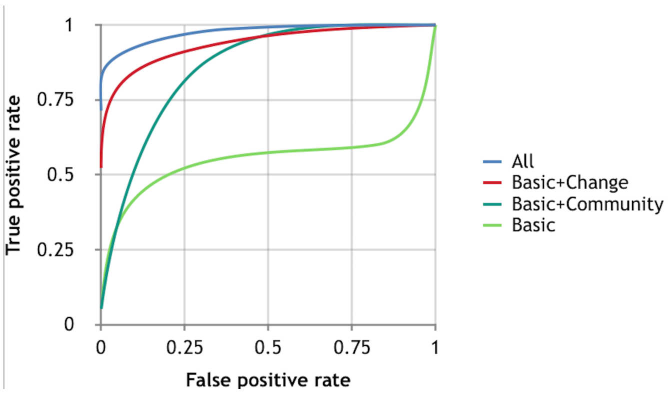

Fig. 4. Comparison of ROC curves for 4 models — BASIC, BASIC+COMMUNITY, BASIC+CHANGE, ALL

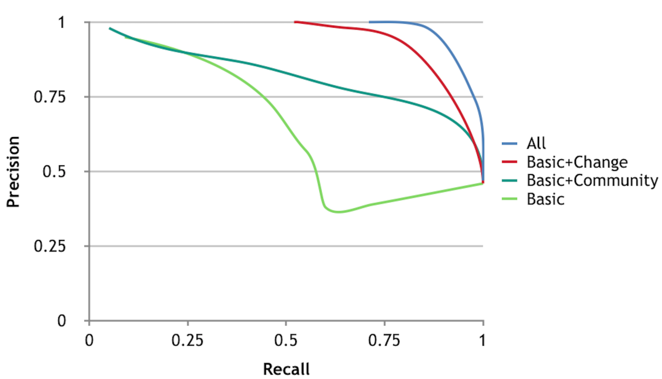

Fig. 5. Comparison of Precision-Recall curves for 4 models —BASIC, BASIC+COMMUNITY, BASIC+CHANGE, ALL

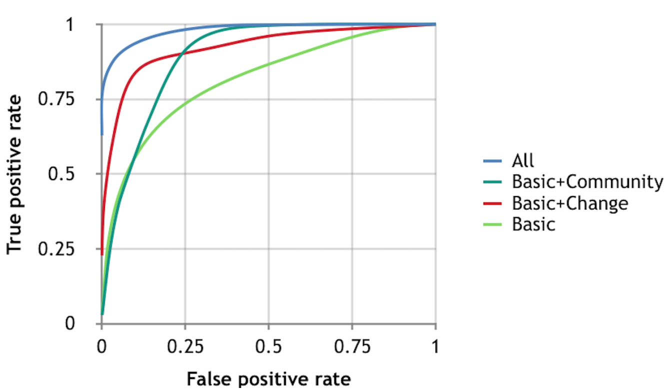

Fig. 6. Comparison of ROC curves for 4 models — BASIC, BASIC+COMMUNITY, BASIC+CHANGE, ALL

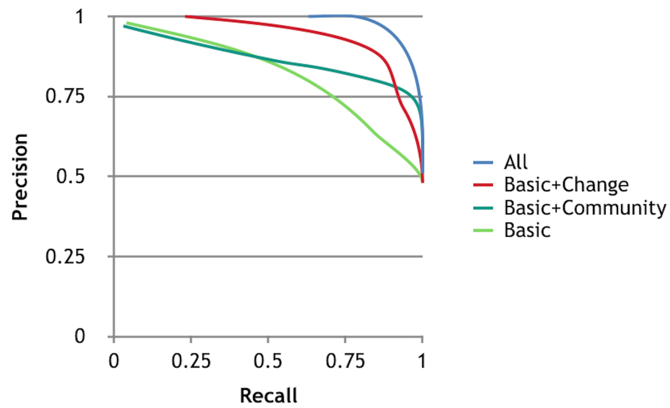

Fig. 7. Comparison of Precision-Recall curves for 4 models —BASIC, BASIC+COMMUNITY, BASIC+CHANGE, ALL

V. Conclusion and Future Work

Our experiments show that our proposed features based on social influence and changes in both call usage patterns and social groups are good indicators of churn and proven to perform better than models without these features on the same dataset. The predictive power of our hybrid behavioural feature set augment the churn prediction model in the face of data constraints.

It is of special interest to extend our model to be able to handle detection of extreme churn events associated with competitive schemes and technological advantages mooted by competitors. This work can be extended to handle detection of extreme churn events associated with competitive schemes and also in detecting compromised IoT nodes or nodes in a WSN by identifying similar behavioural features.

References

1) P. Datta, B. Masand, D. R. Mani, and B. Li, “Automated cellular modeling and prediction on a large scale,” Artif. Intell. Rev., vol.14,no.6, pp. 485–502, December 2000.
2) Coussement and D. V. Poel, “Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques,” Expert Syst. Appl., vol. 34, no. 1, pp.313–327, January 2008.
3) D. A. Kumar and V. Ravi, “Predicting credit card customer churn in banks using data mining,” Int. J. Data Anal. Tech. Strategies , vol. 1,no. 1, pp. 4–28, August 2008.
4) P. Spanoudes and T. Nguyen “Deep Learning in Customer Churn Prediction: Unsupervised Feature Learning on Abstract Company Independent Feature Vectors,” CoRR, abs/1703.03869, March 2017
5) A. A. Nanavati, S. Gurumurthy, G. Das, D. Chakraborty, K. Dasgupta, S. Mukherjea, and A. Joshi. On the structural properties of massive telecom call graphs: findings and implications. In CIKM ’06: Proceedings of the 15th ACM international conference on Information and knowledge management, pages 435–444, New York, NY, USA, ACM, 2006.

6) K. Dasgupta, R. Singh, B. Viswanathan, D. Chakraborty, S. Mukherjea, A. A. Nanavati, and A. Joshi. Social ties and their relevance to churn in mobile telecom networks. In EDBT '08: Proceedings of the 11th international conference on Extending database technology, pages 668– 677, New York, NY, USA, ACM, 2008
7) Y. Richter, E. Yom-Tov, and N. Slonim, "Predicting customer churn in mobile networks through analysis of social groups," in Proc. SDM 10, Soc. Ind. Appl. Math., Philadelphia, PA, pp. 732–741, 2010.
8) V. D. Blondel, J.L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. In Journal of Statistical Mechanics: Theory and Experiment, page 10008, 2008.