

Personalizing mobile user subscription services using data mining

Aleksandar Karadimce¹, Dijana Capeska Bogatinoska¹

¹University of Information Science and Technology “St.Paul the Apostle”, Ohrid, R.Macedonia
aleksandar.karadimce@uist.edu.mk,
dijana.c.bogatinoska@uist.edu.mk

Abstract. Mobile companies today offer diversity of user subscription services to their subscribers in order to attract their attention. In order subscribers effectively to use that service they have to choose appropriate subscription service package. Choosing the appropriate subscription service is not always simple decision, using the knowledge discovery process can help customers to make the right choice. We suggest applying different data mining techniques to already aggregated user traffic from different types of services, stored in the data warehouse system. With the extraction of the useful information we can provide offers to subscribers that will be optimal with their used services (voice, SMS, MMS and Internet). Personalization of the user subscription services will contribute to more objective and transparent process of billing the subscribers.

Keywords: Personalized mobile services, data mining, mobile subscribers, Business Intelligence.

1 Introduction

Telecommunication companies, in battle to enlarge their coverage on the market, propose variety of user subscription services to their subscribers in order to attract their attention. When users sign a contract with telecommunication provider for mobile subscription services they have to choose from a predefined tariff models, either pre-paid or postpaid subscribers. These predefined pattern based subscription tariff models contain predefined amount of internet traffic, number of SMS or MMS events and specific amount of time for free voice calls in different mobile zones. Users should be able to use entirely their benefits defined with the subscription tariff model, or they have underutilized use of the benefits provided with the assigned subscription model. Proposed solution offers mobile subscribers appropriately to use the chosen subscription tariff model, providing them with personalized user subscription service.

Choosing the appropriate subscription service is not always simple decision, using the knowledge discovery process can help customers to make the right choice. Customizing the user subscription services according to personal requirements, it asks for knowledge to understand and analyze user data flow behavior. Everyday use of mobile devices generates different type of data traffic that can be used to analyze the user

requirements. The main contribution of this paper is these gathered user data traffic to be reused for personalizing mobile user subscription services that are offered, which will contribute to more objective and transparent process of billing the subscribers.

In section II of this paper we present personalized mobile user subscription services. The Section III gives overview of results from comparison of different data mining techniques. Section IV concludes the paper and presents some future work.

2 Personalized mobile user subscription services

Mobile customers in order to take advantage of the new benefits have to sign new long term loyalty contracts with mobile provider, as a subscriber to only one tariff model. On the other side the mobile provider itself cannot offer customized mobile subscription tariff model to every subscriber, as stated in its request for mobile services. Existing subscribers have to choose the predefined subscription tariff model and state the limit of monthly subscription with its appropriate charging rate plan. This way mobile provider knows subscribers predefined limit for monthly spending for mobile services. Existing research areas have already used the mobile events history records, call detail records (CDR), for marketing and fraud detection applications [1], [3]. Also data mining of large database systems has been a major challenge in the telecommunication companies, such as the survey of different available data mining techniques [2]. There have been different methodologies of data mining used for customer churn prediction based on either demographic features or billing or usage features, as shown in [4].

The importance of personalization is given in [5] that give an overview of issues that must be considered to leverage future technologies that can support more advanced personalization. It also highlights the distinction between customization and personalization, where customization is thought of as user controlled modifications of a service and personalization is machine-controlled. In [6] is given technical sense of profiling that using data mining we have certain degree of probability in order to customize individual decisions.

The process of creation subscriber personal profile enables businesses to provide highly individualized services for their subscribers and targeted advertising for their customers. This way mobile subscription service is more subscriber oriented, provides increased user experience and offers flexibility that the limit will be used more optimally, instead of the current pattern based mobile subscription services. Disadvantage in the current subscription service is that the existence of shorter events fragment the subscription limit, instead the calls that have long duration to be included to the subscription limit. The implementation of personalized mobile subscription services will provide new type of subscription service that will bring revenue increase on long term basis. Instead of applying different kind of promotion packages for the mobile tariff models, we suggest creating subscriber personal profile, where there will be modular discrete packages for different kind of service. This way only one discrete package for specific service can be assigned to a particular subscriber personal profile, which shall decrease the total amount of the user invoice.

3 Comparison of different data mining techniques

The advantage of using the data mining, as technique to extract knowledge, performed on already aggregated data relevant to the particular subscriber. We have done research using four different algorithms. The Microsoft Decision Trees algorithm is a classification algorithm that works well for predictive modeling [7]. On the other hand the Microsoft Clustering algorithm uses iterative techniques to group records from a dataset into clusters containing similar characteristics [7]. The Microsoft Logistic Regression algorithm is a regression algorithm that works well for regression modeling. The Microsoft Neural Network algorithm uses a gradient method to optimize parameters of multilayer networks to predict multiple attributes [7]. That dataset for the research are subscriber previous calls and data events that are being deposited by the mobile provider. Using the Microsoft SQL Server Management Studio we have conducted preparation for data mining research, on dataset sample of 10.000 CDR records, randomly generated and stored in MS SQL Server table called [CDR_traffic].

Especially important is the column Promotion that is populated indirectly using database update procedure, with discrete values 'LONG', 'MEDIUM' and 'SHORT', based on the values recorded in the column Duration. Estimation for the prediction value is saved in column Discount, receives values from 1 to 3 that represent affinity weight for discount. The highest value for discount means it should be applied highest priority distribution of the subscription limit. Microsoft Business intelligence development studio is other tool that was used to perform the task of conducting different data mining techniques. In order to create testing set that will be used in the data mining training, we have assigned 50% of 10.000 CDR records to be reserved for model testing. Using the Microsoft SQL Server Analysis Services Designer the process continued with comparison of the four different data mining models under the name "MINING AGG Traffic Data", see Fig 1.

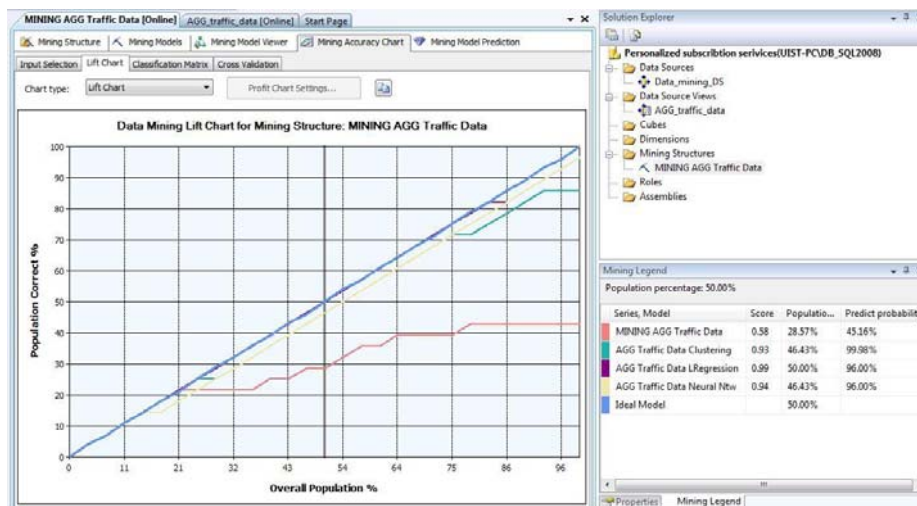


Fig. 1. Comparison of data mining models – lift chart.

Results from the compared four different data mining models in Mining Accuracy Lift Chart have shown that the best results, with score 0.99 closest to Ideal Model, are provided by using the Microsoft Logistic Regression. On the other hand the Microsoft Decision Trees algorithm, with score of 0.58, is not appropriate in the research. Observing the results from this research we can conclude that Microsoft Logistic Regression provides best results in determining the state of the predictable column for continuous and discrete input values.

4 Conclusion and future work

Personalization of the user subscription services means distribution of subscription according to user needs, depending on the type of service, zone or duration of the calls. It will contribute to more objective and transparent distribution of the subscription limit that gives customized charging to the subscribers. Used different data mining techniques to already aggregated user traffic from different types of services, provides new customized and flexible way subscribers to gain optimal charge for their used services that points to user oriented personalized subscription. The relatively small sample of data used in this research has provided results instantly; otherwise it should be considered that in real-time database systems, where there is thousand times more data, it would require more demanding resources. Upcoming research should involve implementation of realistic dataset, to overcome the real-time limitations of resource intensive environment, we suggest using the benefit of cloud computing.

References

1. Weiss G. M.: Data Mining in the Telecommunications Industry. In: O. Maimon and L. Rokach (.eds), *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*. Kluwer Academic Publishers, 1189-1201. (2005)
2. Chen M.S., Han J., Yu P. S.: Data mining: An Overview from a Database Perspective. In: *IEEE Transactions on knowledge and data engineering*, Vol 8, No 6, Dec 1996. 866- 883. (1996)
3. Qayyum S., Mansoor S., Khalid A., Halim K. Z., Baig A.R.: Fraudulent Call Detection For Mobile Networks. In: International Conference on Information and Emerging Technologies (ICIET). 14-16 June 2010. 1 – 5. (2010)
4. Khan A. A., Jamwal S., Sepehri M.M.: Applying Data Mining to Customer Churn Prediction in an Internet Service Provider. In: International Journal of Computer Applications (0975 – 8887) Vol 9, No.7, Nov 2010. 8-14. (2010)
5. Jorstad I., Thanh D. V., Dustdar S.: Personalisation of Mobile Services: In: IEEE International Conference on Wireless And Mobile Computing, Networking And Communications, 2005. (WiMob'2005), 22-24 Aug. 2005. Vol. 4. 59 – 65. (2005)
6. Nancy K. J., Wegner J. P.: Profiling the Mobile Customer – Privacy Concerns When Behavioural Advertisers Target Mobile Phones – Part I. In: Computer Law and Security Review. Computer Law and Security Review, 26(6), 595-612. Elsevier Publishing. (2010)
7. Microsoft Data Mining Algorithms (Analysis Services - Data Mining). [http://msdn.microsoft.com/en-us/library/ms175595\(v=sql.105\) <29.07.2012>](http://msdn.microsoft.com/en-us/library/ms175595(v=sql.105) <29.07.2012>)