

Impact of Data Mining on Telecommunication Company Revenues

I. A. Kamani C Samarasinghe, S. R. Kodituwakku, Y. P. R. D. Yapa

Abstract: Rapid advancement of the technology has made the telecommunication sector very competitive. In order to keep up with the competition, telecommunication operators have to identify the exact needs of the customers and offer services in-line with customer needs. The aim of this research is to investigate the applicability of data mining in identifying customer needs and how it can be adapted to increase the revenue of telecommunication companies. The objectives of this study include an investigation into the relationship between data mining practices and customer behavior patterns, relationship between customer needs and products or services, relationship between new product design initiatives and revenue increases of the companies, the impact of data mining on the revenue of telecommunications companies, and the development of a data mining framework to improve the overall Average Revenue Per User (ARPU) levels in the industry in addition to the designing of a Business Intelligence (BI) tool to enhance decision making processes for improving the overall ARPU levels of the industry. Firstly, the conceptual model is developed based on the feedback of a sample of employees who hold positions in the telecommunications sector. This model has four main variables; data mining, customer behavior, product and increased revenue. Secondly, a preliminary study was carried out to test the variables and to find out how data mining can be applied to identify customer needs and how companies can benefit from using data mining techniques in their businesses. Next, a Data Mining framework was developed to make sure that the expected results could be received from the data mining exercise in place. Finally, a Business Intelligence tool was developed to validate the data mining framework. The preliminary study revealed a clear relationship between the variables of the conceptual framework. Furthermore, it was evident that data mining could lead to better business decisions, apart from the other key benefits of using it, such as timely delivery of services and an increase in customer satisfaction which may affect the revenue of the company. The post survey validation from the target users (managers of telecom companies) indicated that the proposed BI tool is capable of retrieving much needed information for business decisions, which would lead to increased revenues of the companies. The long term results are likely to be positive in this context and it is also evident that the role of data mining can be expanded by the companies and that this practice could eventually lead to companies providing markets with the exact requirements.

Index Terms: Business Intelligence, data mining, revenues, telecommunication

I. INTRODUCTION

The Sri Lankan telecommunications industry is growing at a modest rate and based on the Fitch rating estimates the growth is in the mid-single digit range as a percentage.

Revised Version Manuscript Received on June 20, 2016.

I A Kamani C Samarasinghe, M.Phil, Post Graduate Institute of Science (PGIS), University of Peradeniya, Sri Lanka..

Saluka R. Kodituwakku, Professor, Department of Statistics and Computer Science, Faculty of Science, University of Peradeniya, (Peradeniya). Sri Lanka.

Roshan D. Yapa, Lecturer, Department of Statistics and Computer Science, Faculty of Science, University of Peradeniya, (Peradeniya). Sri Lanka.

This growth will continue as a result of the increase in demand for voice services, primarily at the enterprise level. The mobile telephony growth will be flat to modest due to the high level (84%) of telephony penetration levels in the country. This indicates that companies are operating in a very competitive market and that the market is highly penetrated. Thus, companies will face intense competition from each other and will have to ensure they gain more customers in order to increase market share. Rapid advancement of the technology has made the telecommunication sector a very competitive area of business. In order to keep up with the competition, telecommunication operators have to identify the exact needs of the customers and offer services in-line them.

Data mining is a very popular concept in the current context and many companies mine data to understand customer behavior patterns. These patterns can be used to predict the how they are likely to behave in the future. As a result, data mining is a very useful tool in the context of data manipulation practices. Data mining practices would allow companies to understand their customers' needs and serve them accordingly. With increased data mining practices, the companies are in a position to understand how the customers behave and what drives the ARPU of the companies. Once this is understood, the companies are in a position to design products in line with the requirements of the customers. Currently, many of the products seem to be developed to manage network capacity rather than to fulfill customer expectations. However, data mining practices will facilitate the understanding of the real market needs which can then be used in order to device means of fulfilling them. With the products falling in line with the exact requirements of the customers, the ARPU would increase as the customers would increase the use of services that are compatible with their expectations. This would, subsequently, increase the revenues of the company. However, the nature of the relationship may change from industry to industry and the impact created on the relationship has to be understood in detail. Thus, understanding data mining and the possibility of its contribution towards increased revenues has been the main purpose of the study. The main research question is "Will data mining practices contribute towards increased revenue in the telecommunications industry in Sri Lanka?"

The aim of this research is to find out how data mining can be applied to identify customer needs and how telecommunication companies can benefit from using data mining techniques in their business.

II. RELATED WORK

Companies are known to have a considerable amount of information for processing. Segura, Castro, Domínguez, Campos, and Prieto show that if companies know how to

Impact of Data Mining on Telecommunication Company Revenues

utilize this information efficiently, they can get better insights into the customer profiles and their needs [1]. It is likely that there are heterogeneous customers who have different needs. The customer needs are driven by a number of facts and data mining can be used to provide insights into these areas. With the companies gaining such awareness, it is likely that they can create better products and services packages to the customers which are compatible with their needs. Pendharker and Rodger explains that this ensures that the demand for products or services offered by the company would increase as they are designed according to the needs of the customers [2]. The eventual result is that the companies are able to establish a good reputation in the market due to the fact that they are meeting the exact needs of the customer. Larson states that business intelligence is about the delivery of accurate, useful information to the appropriate decision makers with necessary timeframe to support effective decision making [8]. Inmon suggests that data warehousing is the process of collecting data to be stored in a managed database in which the data are subject-oriented and integrated, time variant, and non-volatile so that it can support decision making processes [9].

The telecommunication industry is a global industry and many of the companies have the global trends in place. For instance, the trends developed in the markets in developed countries have been replicated to a greater extent in those found in the developing nations as well. This indicates that telecommunication companies in countries such as Sri Lanka have been able to gradually build their businesses according to the requirements of the market by ensuring that the adaptation of global market trends are taken as key guiding modes by such companies. McAdams, Camp and Divakaruni, demonstrate how this could indicate the fact that data mining has become highly irrelevant in the context of the telecom industry [11]. It is clear that this particular industry requires a large amount of data. These data have to be collected and processed with the intention of assuring that the market developments are being fully understood. All these developments would eventually indicate that data mining is a crucial factor when identifying the market needs for a particular product or service. [12]. Data mining as a technique allows the industry to understand the market in a better manner. It is likely that the industry has a number of complications and these complications can reveal the eventual details of the market needs as well as the customer spending patterns etc. Simpson shows that importance of identifying these patterns and offering suitable packages in line with the expectations of the customers. Such an approach would ensure that the companies are able to gain an understanding of the volatile nature of customer needs.

By seeking to comprehend the market requirements, the general perception is that such companies would be able to build their position in the existing market [13]. This indicates that the companies who have the systems in place for the identification of the customer needs are in a position to prosper as opposed to those which do not. In such a scenario, data mining can be used as a process that the companies can use for the purpose of digging into the historical storage of data that they already have in store. The information obtained can then be processed, analyzed and used for the benefit of the company. Thus, in the long term, companies can benefit from

these findings.

Obviously, data mining is a practice which cannot be used by all the companies in any industry. This is because to engage in such an activity a company needs to have the suitable levels of relationships with the customers and it should also be in a position to store and use the data obtained through these interactions. The companies who can resort to data mining should invest in the development of the tools needed for the task since there are special tools used for this particular purpose. Having them would facilitate the activity of being able to tabulate the available data and identify the market trends in addition to being able to forecast potential shifts in market needs.

It is important to note that in most instances the customers themselves do not have very specific awareness of what their actual needs are. This is the result of them not being able to specify and identify the main features of what they actually require. This indicates the fact that the companies will have to identify exactly what is needed by the customers and ensure that they are fulfilled. Therefore, data mining should be done using the most appropriate tools in a given context.

With regard to the telecommunication industry in the context of Sri Lanka, we can say that it has become a very competitive industry at present. It is of great importance that the players in the industry should be aware of data mining and its benefits and adapt them in order to address customer needs appropriately. Since data mining would indicate the main features associated with the market space that each of the company is occupying it has the potential of providing useful insights into the industry and its requirements. As a result such companies adapting this technique would be able to benefit from it. This would help them to encroach the market and improve their market share as a result of being able to eat into the market shares of the telecommunication sector.

It is very much likely that in the long run, such companies will become value players who provide products and services in line with the needs of the markets. Thus, the customers will obtain them as they will realize that their needs are fulfilled satisfactorily. All these aspects indicate that telecommunication companies in the context of Sri Lanka will be able to benefit from engaging in data mining related exercises.

III. METHODOLOGY

Firstly, a conceptual model was developed to indicate the main information areas that the study seeks to provide such as market insights tested through feedback provided by sample of employees who hold high rank positions in the telecommunications sector. This conceptual model has four main variables; an independent variable (data mining), two intermediate variables (customer behavior and product) and a dependent variable (increased revenue). Secondly, a preliminary study was carried out to test the variables and to find out how data mining can be applied to identify customer needs and how companies can benefit from using data mining techniques in their businesses. Thirdly, a theoretical Data Mining Framework was developed based on the finding of pre-survey data to make sure that the expected results could be received from the data mining exercise in place.

Since the conceptual model could not be successfully used by the end users, it was decided to develop a BI tool based on

the proposed data mining framework. The developed tool was validated by passing it to the end users. In developing the tool, a few widely used BI solutions were observed before selecting the Pentaho BI solution which was chosen since it is totally free and can be developed in a small environment for academic purposes. Finally, a post survey validation test was conducted to check the acceptance of the proposed framework and the tool. The pre and post survey discussions were held based on a structured questionnaire. The pre-survey questionnaire was developed in line with the operationalized variables. The concepts dissected to different components which contributed to the construction of the complete concept in an appropriate context. The Lickert scale based measuring technique was used with scores ranging from 1-5 and 1 being “strongly disagree” to 5 being “strongly agree”, to convert the qualitative information to quantitative form. This facilitated the comparison of the findings associated with these areas. The hypotheses were developed based on the needs of the study. The size of the sample was 100. The data collected were analyzed in SPSS 16.0 version. Different analytical and descriptive techniques were used to analyze the data.

IV. RESULTS

Out of 100 employees 92 responded the questionnaire. Therefore, response rate was 92%. To understand the integrity of data within a variable, Cronbach’s Alpha was used and Table 1 indicates the data integrity scores in every instance. The information in Table 1 indicates that all the scores are above 0.90 which confirms that the scores are high in general. This data integrity levels associated with all the variables of the discussion remain high and therefore confirms that the data can be used for the purpose of further analysis. The core findings indicate the nature of the relationships between the variables that are in discussion. It is clear that the conceptual model indicates that with a high level of data mining activities, the companies are able to identify customer behavior and market requirements. This enables the companies to develop products and services that are consistent with these expectations. Once the developed products are consistent with the customers’ needs, they are likely to purchase additional quantities of these products and services due to the fact that they meet the exact requirements expected. As a means of identifying the role that data mining plays in the context of the companies, data mining variables related to the concept in the conceptual model was tested and an average score of 3.12 was observed.

TABLE 1: CHRONBACH’S ALPHA SCORES (RESEARCH DATA)

Variables	Cronbach's Alpha score
Data mining	0.960
Customer behavior	0.919
Product design	0.939
Increased revenue	0.933

The customer behavior identification approaches were important, as they can indicate how the customers behave. This can allow the companies to develop suitable long-term solutions consistent with the requirements of the customers. Thus, an appropriate understanding of customer behavior remains one of the key areas of discussion. To understand

customer behavior, the variable related to the conceptual model was tested and an average score of 3.05 was observed.

Once the customer needs are identified, it is important to compile the information identified for the purpose of developing products and services and to use them for their further development to ensure the satiation of customer needs. To understand the product design consistent with customer requirements, variables related to the concept in the conceptual model was tested and an average score of 3.15 was observed.

A company could be able to generate revenue from the activities that are discussed above. If there is a clear relationship between the developments of products consistent with customer requirements, then it would ensure that the market needs are met. . As a result, the company should be able to enjoy increased revenue over a period of time. To understand the revenue increase impact, the variable related to this concept in the conceptual model was tested and average score of 3.15 was observed. Then, a step-by-step analysis was undertaken for the purpose of establishing the nature of the relationship between the variables of the pre-study conceptual model.

To understand the relationships between the variables of the conceptual model, Pearson’s correlation scores were tested. The correlation score between data mining and customer behavior variables was 0.979. This indicates an almost perfectly positive relationship. Next, Customer behavior and product design variables were tested and correlation score was 0.973. Pearson’s correlation score associated with the variables of the conceptual model Product design and revenue increase was 0.954. Finally, the relationship score of 0.998 was observed between data mining and revenue increase variables of the conceptual model and it indicated a highly positive relationship between the variables.

The theoretical data-mining framework developed based on the findings of the pre-study conceptual model for testing data was used to make sure that the expected results could be received from the data mining exercise in place. The usage habits, categorization of the customers and historical analysis were identified as main steps in the proposed data mining framework. They are considered as the main areas that could contribute towards the data mining activities in the appropriate context.

The project was focused on developing an open-source Pentaho Business Intelligence platform to test the validity of the conceptual model and data mining framework. After completing the tool design, pilot demonstrations were carried out to introduce the system to decision makers. Finally system validation was done using post survey questionnaire, to check the acceptance of the tool. The questionnaire was distributed among decision makers at a telecommunication company. The responses acquired from the post survey were further analyzed to formulate business decisions. Table 2 summarizes the responses and results which show that a vast majority of respondents were extremely satisfied with the system. Furthermore, the results of the system validation proved the acceptance of the tool.

Impact of Data Mining on Telecommunication Company Revenues

**TABLE 2 : ANALYSIS OF POST SURVEY QUESTIONNAIRE
(RESEARCH DATA)**

Question Number	I strongly disagree with the statement	I do not agree with it	I neither agree nor disagree with the statement	I agree with the statement	I strongly agree with the statement
The Telco BI system is easy to use	0.00%	0.00%	0.00%	22.20%	77.80%
The BI system provides sufficient information to make better decisions	0.00%	0.00%	5.60%	44.40%	50.00%
It is easy to retrieve data from the system.	0.00%	0.00%	0.00%	38.90%	61.10%
The output is in line with the business requirements	0.00%	0.00%	17.60%	23.50%	58.80%
The BI system satisfies my requirements	0.00%	0.00%	11.10%	38.90%	50.00%

V. DISCUSSION

The study was aimed to identify the nature of the data mining practices in the telecommunications sector in Sri Lanka. This study was conducted to identify the data mining practices currently in place and to propose an improved data mining framework.

The above results clearly indicated the fact that there is a relationship between the variables of the conceptual framework. As a result, for companies focusing on data mining activities, using it would enable them to identify the customer behavior patterns accurately. Once the customer behavior patterns are identified, companies would be able to develop products and services which are consistent with these requirements. After their successful development, companies are in a position to guarantee that the products and services provided by them are truly compatible with the requirements of the customers. Hence, it is obvious that telecommunication industries that opt to undertake data mining initiatives would be able to benefit from these activities.

The preliminary study revealed a clear relationship between the variables of the conceptual framework. Moreover, it was evident that data mining could even lead to better business decisions. Other advantages of this activity are the timely delivery of services and high customer satisfaction levels which all contribute towards increased revenue generation. The post survey validation from the target users (managers of telecom companies) indicates that the proposed BI system is capable of retrieving information needed for making business decisions, which leads to increased revenue of the company. The results of the system validation provide positive acceptance of the proposed framework and the tool. The long term results are likely to be positive in this context and it is also evident that the role of data mining can be adapted for the betterment of the

companies. Such practices would eventually lead to the fulfillment of customer requirements. The results indicate that data mining practices are helpful to the telecommunications sector of the country

The study encountered a number of limitations, which need to be considered. In academia, a limited number of research are being carried out with regard to this particular area of study. One reason for this is that although a large part of the data needed lies within corporations, they are unwilling to share them due to the competition among them. More knowledge should be made available by these corporations to stimulate research in the business intelligence area by academics. The topic is very wide in the real world, in the context of academia, we have been able to show our understanding in this area but several directions for future research is possible from where we stopped and the work performed in this project provides basis for future research. In addition, further studies on innovative designs and testing of new concepts, sustainable and cost-effective technologies, and structures for the telecommunication sector to achieve highest average revenue per user, are possible and recommended as future work. Also the work performed in this project can be retested using scientific Data mining functions, and this technique is recommended for prospective research work. In the future, possible integration of the business intelligence window into tool is possible.

ACKNOWLEDGMENT

I would also like to thank my MPhil supervisors at Post Graduate Institute of Science (PGIS), University of Peradeniya, Prof S. R. Kodithuwakku and Dr. Y. P. R. D. Yapa, for all their valuable input, support and encouragement throughout this research.

REFERENCES

- Segura, A., Castro, C., Domínguez, V., Campos, P.G., and Prieto, M., (2011) "Using data mining techniques for exploring learning object repositories", *Electronic Library*, the, Vol. 29 Iss: 2, pp.162 – 180
- Pendharkar, P.C., and Rodger, J.A., (2000) "Data mining using client/server systems", *Journal of Systems and Information Technology*, Vol. 4 Iss: 2, pp.72 – 82
- Data mining using SAS/EM: A Case Study Approach, SAS, Institute Inc. 2003.4
- Jiawei Han and Micheline Kamber *Data Mining Concepts and Techniques* Beijing Higher Education Press, 2001:279-299.
- Elder J F & Abbott D W, *A Comparison of Leading Data Mining Tools*, KDD-98.
- Jiawei Han, Micheline Kamber, Fan Ming, Meng Xiaofeng (interpret), *Data Mining: Concepts and Technologies* [M], Beijing Machinery Industry Press, 2001, 8
- Ren Mingshu, *Study on Web Mining and Electronic Commerce*, Shandong University of Science and Technology, 2002.
- Larson, B (2006) *Delivering Business Intelligence with Microsoft SQL Server 2005*, McGraw-Hill/Osborne
- Inmon W.H., (1993) *Building the Data Warehouse*, A Wiley QED publication, John Wiley and Sons, Inc. New York 123-133
- Liu, S.S., and Chen, J., (2009) "Using data mining to segment healthcare markets from patients' preference perspectives", *International Journal of Health Care Quality Assurance*, Vol. 22 Iss: 2, pp.117 – 134
- McAdams, A., Camp, J., and Divakaruni, S., (2000) "The evolution of US telecommunications infrastructure", *info*, Vol. 2 Iss: 2, pp.107 – 110
- Lee, Y., Yen, S., and Hsieh, M., (2005) "A lattice-based framework for interactively and incrementally mining web traversal patterns", *International Journal of Web Information Systems*, Vol. 1 Iss: 4, pp.197 – 208
- Simpson, S., (2010) "Governing information infrastructures and services in telecommunications", *Aslib Proceedings*, Vol. 62 Iss: 1, pp.46 – 56

14. Sacripanti, A.M., (1999) "Liberalizing telecommunications in Italy: the role of the regulator", info, Vol. 1 Iss: 5, pp.449 – 453
15. Premalatha, S., and Baskar, N., (2012) "Implementation of supervised statistical data mining algorithm for single machine scheduling", Journal of Advances in Management Research, Vol. 9 Iss: 2, pp.170 – 177



I A Kamani C Samarasinghe is pursuing her M.Phil. degree at the Post Graduate Institute of Science (PGIS), University of Peradeniya, Sri Lanka. Her main research area is Data Mining.



Saluka R. Kodituwakku is a Professor at Department of Statistics and Computer Science, Faculty of Science, University of Peradeniya Peradeniya, Sri Lanka. He has obtained his Ph.D (Computer Science) from RMIT University, Melbourne, Australia. His research interests include Database Systems, Software Engineering and Distributed Computing.



Roshan D. Yapa is a senior lecturer at Department of Statistics and Computer Science, Faculty of Science, University of Peradeniya Peradeniya, Sri Lanka. He has obtained his PhD (Information Engineering), from Hiroshima University, Japan. His research interests include Database Systems, Image Processing and Statistical Data Mining.