Understanding the Internet Usage Habits of the Students of University of the Visual & Performing Arts through Data Mining

I A. Kamani C Samarasinghe, Saluka Kodituwakku, Roshan D. Yapa

Abstract— This Data mining has been a commonly used tool in the telecommunication sector. This is due to the useful insights that it can provide to assess the user preferences and optimize the service offerings consistent with user expectations. Data mining provides the required insights to the Internet and data usage habits of the students of the University of the Visual & Performing Arts (UVPA). It is clear that the data usage demand has been accelerating over the past few years and mobile data usage has been in the forefront of growth. Reduced prices as well as increased Internet usage options along with benefits achieved from the attributes of increased speed had augmented the usage of the data services. The study was a quantitative study and the information was collected from a random set of students who had registered with the University for their Degree program. The information collected had been processed consistent with the requirements to gain insight on the Internet usage habits of the students. The study represents data of 133 students who are from 17 districts of the country. The findings indicate that 83% of the households are Internet ready and Close to 50% of the households possess desktops followed by over 40% of households employing smartphones, which facilitate Internet access. Meanwhile, several students use dongles for Internet connectivity. The Internet is utilized for various purposes; where the purposes of online education and social networking are the two prominent areas.

Index Terms— Data mining, Telecommunication, University of the Visual & Performing Arts.

I. INTRODUCTION

Data mining has been a very common topic of discussion during recent times due to the increasing need for companies to identify their customer requirements and habits [3]. If a company undertakes data mining practices, this indicates that they will be able to identify how the customers use their products and services as well as how the changes in customer preferences can affect their business. This understanding will provide the necessary groundwork for companies to develop products and services consistent with the expectations that they have. This indicates that data mining is an important area for complex industries.

One of the industries that find data mining to be useful is the telecommunications sector. This is due to the fact that telecommunication companies provide various services to the public and the public will have their own usage patterns. The companies will have to identify these trends and ensure that they specifically cater to these needs [5].

This will result in increased quality of services to the customers. The purpose of the paper is to identify the Internet usage habits amongst the students of the University of the Visual & Performing Arts (UVPA). This enables the companies who provide such services to ensure that they target the requirements of the students at a specific level.

The role of data mining becomes significant in the context of identifying the factors related to the industry and how one variable would impact the other [9]. Thus, they will be able to identify future market trends by studying the demand for internet usage amongst the students and take appropriate steps to improve their services.

II. RELATED WORK

The complexity of the telecommunication industry is growing considerably. It is important to note that the demands for various services are changing considerably along with the changes taking place in technology. It is this change in technology, which drives the changes in the services sector which require to be provided [4]. When it comes to the industry, they provide voice based services as well as text messaging and data services. The demand for the data services is growing at a considerable pace and the rest of the areas such as the mobile and fixed voice communication is gradually becoming less profitable in nature [6]. The following chart indicates the growth experienced in telecommunication data services demand in Sri Lanka over the past few years.
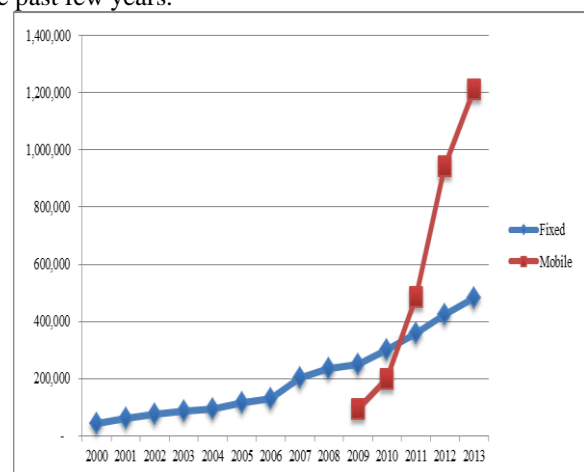

Fig 1. Demand for fixed and mobile data services (Telecommunication Regulatory Commission of Sri Lanka, 2013)

The above chart depicts that the demand for mobile data services is growing at a considerably faster pace in comparison with fixed data services. This indicates accessing of data through smartphones as well as other wireless broadband devices. It is clear that this is primarily driven by the fact that the IT equipment as well as the broadband services have experienced increase in speed, while the prices have considerably reduced due to growth in competition in this area.

There are many competitive points that the companies are seeking to develop; one of the areas is the network speed [2]. It is clear that faster the network responses, faster the user would be able to access the required data. On the other hand, the network should be able to manage all the resources and optimize towards the platforms that the customers use in order to access data. Thus the challenge is to ensure to increase the user experience and guarantee that the required service quality is provided to the users [7].

It is also important to identify the hardware and the software platforms associated with the users. It is clear that the identification of Internet usage habits of the user will allow the companies to develop suitable marketing plans consistent with the needs of the users as well. This will ensure that the users will be able to receive the benefits as well as the companies will be able to gradually consolidate their position in a given market [8]. For instance, if the users possess the required facilities at home, it is best that a home-based solution is developed specifically targeting the education needs of the user etc. Thus the expected value proposition of the user should be targeted.

The mobile phone usage information in the context of Sri Lanka is not an area where research is available on the public domain. While there is possibility that company based literature is available in this area of consideration, they are not publicly available. On the other hand, the market specifics can change depending on the different customer communities. This is a common denominator for most of the telecommunication companies. For instance, the students will have specific telecommunication service requirements. The discussion will provide the required insights to these specific requirements that are in place.

It is interesting to note that data mining requires comprehensive data collection systems. In the instance where the data collection systems are not comprehensive or accurate, it is likely that the data collected by the systems may not provide the expected results [1]. The eventual impact in this context would be negative in nature. The companies will require ensuring that the necessary system with the required accuracy levels is delivered ensuring that the eventual outcome is likely to be positive in nature in this context. This is likely to benefit the companies on the long run.

III. METHODOLOGY

The methodology involved in this case is a quantitative analysis and the data has been collected from 133 randomly selected students based on a questionnaire to identify the data usage habits. Random sampling is likely to provide a fair representation of the customers in focus and the questionnaire remains a focused method of collecting well-structured information from the customers in a quantitative context. Data mining provides data, which are descriptive in nature as the purpose of the exercise is to identify the usage and the related

habits. It is important that field material remains short and the information that is collected is not sensitive in nature. Various data analysis methods have been used to identify the Internet usage habits of the students. The telecommunication companies can use these findings to structure the services they provide.

IV. RESULTS

The students who were subject to discussions are representing 17 districts of the country and this indicates that they provide a fair representation towards the majority of the areas of the country. The highest representation is from Kalutara and Colombo districts with a count of 15% each while Badulla represented 10.5%. The sample is notably female biased with only 10.5% being male. This could be due to the nature of the population biasness of the University in focus. The primary age group in consideration is the 20 – 25 group and the representation of the other age groups remains minimal.

When the students were inquired as to the place they frequent to obtain Internet facilities, the majority comprising of 54.1% stated that they use Internet at home. This was followed by a further 25.6% who did not have any such specific place while another 11.3% indicated that they prefer to use an Internet café. This indicates that majority of the students would like to access their data services at home. However, in order to access these data services at home, they require having the necessary facilities installed at home. The following chart indicates the home based facilities that the students possessed to access Internet.
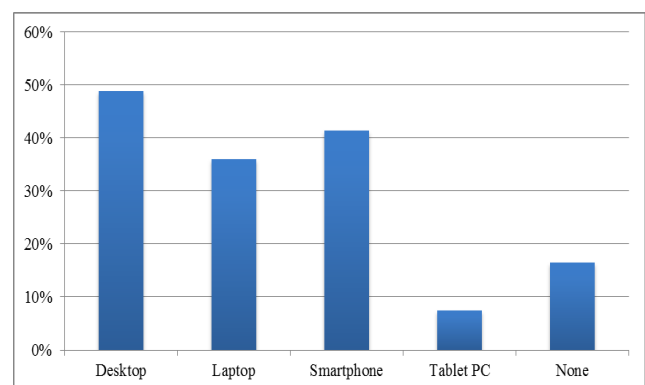

Fig 2: Facilities available at home to access Internet (research data)

It is interesting to note that only 17% of the households did not have the facilities to connect to the Internet. This indicates that most of the homes are either connected or Internet ready homes. This facilitates the service providers with a very unique opportunity to ensure that they penetrate these households. However, the It literacy of the students also have to be assessed. This can be identified through the studies they have undertaken on IT as well as the email connectivity they have.

The majority of the students, comprising of 58.5% have attended some form of formal education related to IT whilst 53.3% have email connections.

This indicates that the majority of the students are in possession of a certain formal qualification or a connection to indicate that they are IT literate.

When it comes to the nature of the courses they have attended, approximately 31.1% have attended basic computer courses. Thus the knowledge that they have is at the basic user level. This indicates that the formal education in the IT sector as well as the exposure to IT devices at home is the combination which drives and contributes towards shaping of IT based services.

It is also important to note as to how the students connect to the Internet. This is important due to the fact that identification of the method of connection will define the methods that the service providers may use in order to improve the level of service that they provide. The following chart indicates the main methods used in connection.
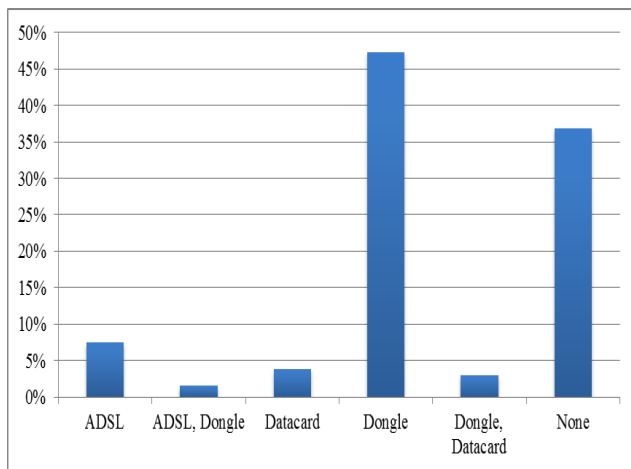

Fig 3: Methods of connection (research data)

It is interesting to note that the most preferred method of connection is the wireless connection and the Dongles are the frequently used method in accessing the Internet. There are approximately 52% of the users who would use dongles for their Internet access. On the other hand, there is another formidable market share yet to be penetrated reaching 37% who do not have their own means of connecting. The companies have the opportunities to introduce suitable data packages to these groups of students as well.

The role of ADSL is not very prominent despite this facility being the only fixed line broadband method available to the customers. The limited availability areas as well as issues when it comes to mobility might be the factors that require to be taken into consideration in this context. It is clear that the students use Internet for multiple purposes; whilst one of the main objectives that have been identified is for the purpose of on-line education. This is one of the prominent aspects and for the students; this is one of the most important areas in consideration. On the other hand, the usages for social networking as well as online chat services also play a prominent role. When it comes to the average time of usage of the Internet per day, the following chart can indicate the details.
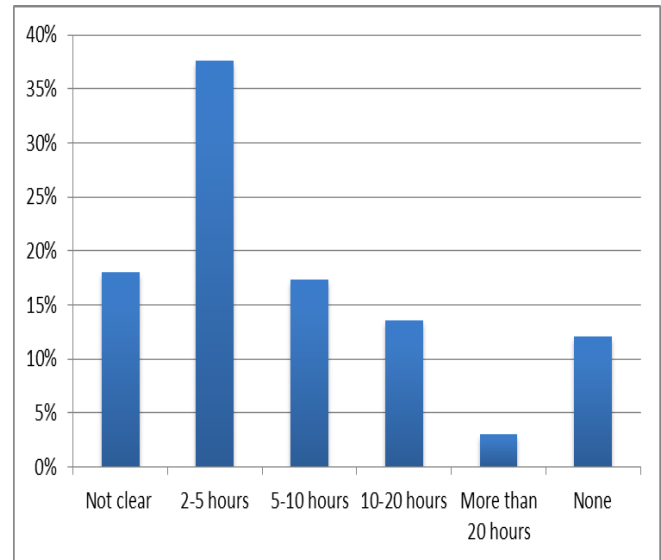

Fig 4: The average length of Internet usage (research data)

The above chart indicates the length of usage of the Internet. It is clear that many use a period of 2 to 5 hours of Internet on a daily basis. This indicates that in general the usage time is limited and they do not wish to be continuously on line. However, it is also important to note the fact that there are certain others who still ensure that they continue to be online. This is evident by the fact that there are students who seek to be online for over 10 hours a day as well. Thus, there are time sensitive users of the data connections as well as those who are not time sensitive in this area of consideration.

V. DISCUSSION

The above findings indicate that the students have drivers' requirements when it comes to their Internet usage habits. It is interesting to note that many households possessing Internet facilities are not yet ready to receive such facilities. Therefore in homes where students are found, Internet usage is likely to be higher than the national average.

Thus the companies who offer data services could provide such households. Furthermore, the students are at the stage of selecting the connections. It is clear that there is a considerable group of students who does not have a model of connection on their own selection. This indicates that they will eventually purchase their own data connections and if the companies will be able to build a relationship with this group of customers, it is likely that the relationship will remain long term in nature. Thus the companies seeking to consolidate their position in this market should provide them with long-term benefits.

The most used methods for data connection is a dongle. It is clear that this will ensure portability of the connection as the students may need to access internet in the university as well as at home. This freedom is provided through a wireless and dongle based data connection that they possess. It is likely that this segment would continue to dominate the data connections in the future. This is the reason that the mobile data connection growth is accelerating in comparison with the fixed data connections.

ACKNOWLEDGMENT

REFERENCES

1. Baxter, G., (2003). Challenge and Change in the Information Society. Journal of Documentation. Vol. 59, Iss: 6, pp.731 – 734.
2. Elragal, A. and El-Gendy, N., (2013). Trajectory data mining: integrating semantics. Journal of Enterprise Information Management. Vol. 26, Iss: 5, pp.516 – 535.
3. Gargano, M.L. and Raggad, B. G., (1999). Data mining - a powerful information creating tool. OCLC Systems & Services. Vol. 15, Iss: 2, pp.81 – 90.
4. Jareevongpiboon, W. and Janecek, P., (2013). Ontological approach to enhance results of business process mining and analysis. Business Process Management Journal. Vol. 19, Iss: 3, pp.459 – 476.
5. Lee, S.J. and Siau, K., (2001). A review of data mining techniques. Industrial Management & Data Systems. Vol. 101, Iss: 1, pp.41 – 46.
6. Mutula, S. M., (2002). Current developments in the Internet industry in Botswana. Electronic Library. The, Vol. 20, Iss: 6, pp.504 – 511.
7. Nemati, H. R. and Barko, C.D., (2003). Key factors for achieving organizational data-mining success. Industrial Management & Data Systems. Vol. 103, Iss: 4, pp.282 – 292.
8. Ozgulbas, N. and Koyuncugil, A.S. (2006). Application of Data Mining Method for Financial Profiling. Social Responsibility Journal. Vol. 2, Iss: 3/4, pp.328 – 334.
9. Sharma, S. Goyal, D. P. and Mittal, R. K., (2007). Evaluation model for data mining software: an empirical investigation of ICICI bank. Journal of Advances in Management Research. Vol. 4, Iss: 2, pp.63 – 68.
10. Telecommunication Regulations Commission (TRC), (2013). Statistics. [Online] Available at:<http://www.trc.gov.lk/index.php/information/statistics.html> (Accessed March, 2014).
11. Viktor, H. L. and Arndt, H., (2006). Combining data mining and human expertise for making decisions, sense and policies. Journal of Systems and Information Technology. Vol. 4, Iss: 2, pp.33 - 56

AUTHORS PROFILE

I A Kamani C Samarasinghe is pursuing her MPhil degree at Post Graduate Institute of Science (PGIS), University of Peradeniya, Sri Lanka. Her 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 Ph.D (Computer Science) from RMIT University, Melbourne, Australia. He is having about more than 20 years of teaching experience. Database Systems, Software Engineering and Distributed Computing are his research areas.

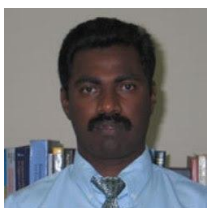**Roshan D. Yapa** is a senior lecturer at Department of Statistics and Computer Science, Faculty of Science, University of Peradeniya Peradeniya. Sri Lanka. He has obtained PhD (Information Engineering), from Hiroshima University, Japan. He is having about more than 12 years of teaching experience. Database Systems, Image Processing and Statistical Data Mining are his research areas.