Brand Switching Pattern Discovery by Data Mining Techniques for the Telecommunication Industry in Australia

Md Zahidul Islam

Charles Sturt University
School of Computing and Mathematics
zislam@csu.edu.au

Steven D'Alessandro

School of Management and Marketing
Charles Sturt University

Michael Furner

School of Computing and Mathematics
Charles Sturt University

Lester Johnson

Department of Management and Marketing
Swinburne University of Technology

David Gray

Department of Marketing and Management
Macquarie University

Leanne Carter

Department of Marketing and Management
Macquarie University

Abstract

There is more than one mobile-phone subscription per member of the Australian population. The number of complaints against the mobile-phone-service providers is also high. Therefore, the mobile service providers are facing a huge challenge in retaining their customers. There are a number of existing models to analyse customer behaviour and switching patterns. A number of switching models may also exist within a large market. These models are often not useful due to the heterogeneous nature of the market. Therefore, in this study we use data mining techniques to let the data talk to help us discover switching patterns without requiring us to use any models and domain knowledge. We use a variety of decision tree and decision forest techniques on a real mobile-phone-usage dataset in order to demonstrate the effectiveness of data mining techniques in knowledge discovery. We report many interesting patterns, and discuss them from a brand-switching and marketing perspective, through which they are found to be very sensible and interesting.

Keywords: decision tree, decision forest, ensemble of decision trees, data mining, brand switching, switching behaviour

1 Introduction

1.1 Mobile Phone Brand Switching

Australians love and hate their mobile phones. According to Euromonitor, in 2011 there were some 25.54 million mobile phone subscriptions, more than one for each member of the Australian population (*Mobile phone statistics* n.d.). The Telecommunications Industry Ombudsman (TIO) received a record 197,682 complaints in 2011, which was 18% higher than the number of complaints received in the previous year. (*Annual Report of the Telecommunications Industry Ombudsman* n.d.). The challenge for mobile service providers

then is to retain their customers by gaining their trust and satisfaction. In order to do so they need to assess the dynamics of brand switching behaviour.

A number of models have been developed to analyse the customer switching process (Bansal & Taylor 1999, Colgate & Lang 2001, Keaveney 1995, Lees, Garland & Wright 2007). The most general, the Keaveney model (1995, p79) identified eight major factors behind customer exit (or push factors) including: pricing, inconvenience, core service failure, service encounter failure, response to service failure, competition, ethical problems and involuntary switching.

1.2   The Problem with Existing Models

Bansal & Taylor (1999) investigated the impact of other factors (on switching behaviour) including switching costs, and service quality. Actual switching behaviour was also examined. The problem with many existing consumer behaviour models, is that they apply a standard model, to what is a heterogeneous market. There is also the possibility that a number of switching models may exist within a large market. To counter this issue researchers sometimes use latent class modelling (Vinzi, Trinchera, Squillacciotti & Tenenhaus 2008). However, even in this approach there is always an issue of model misspecification for different clusters.

1.3   A Data Mining Approach

An alternative to existing models is to let the data speak for itself. In other words, we can apply sophisticated data mining techniques to discover brand switching patterns, which can then be explained in a theory post-hoc manner. Such an approach may be more attractive to practitioners, who may be more interested in prediction and then explanation, than testing existing theories of explanation by evidence. In other words, data-mining techniques can discover previously hidden, non-obvious, interesting and valid patterns (i.e., logic rules) from a dataset without requiring any pre-assumed and domain specific knowledge.

Hung, Yen & Wang (2006) showed the usefulness of data-mining models in brand switching prediction, without exploring the knowledge these models discovered. Instead of discovering knowledge and understanding the patterns, they focussed on achieving a high quality prediction accuracy with the dataset they used. Wei & Chiu (2002) used data mining techniques to discover some knowledge, but they used only contractual data available with the service provider and did not carry out any data collection through survey. That is because knowledge discovery was not their primary objective. Instead, they focussed on building prediction models based on data mining techniques. Additionally, they suggested that future models make use of geographical and complaint information from customers. We use this information in the dataset used in our study. Ahn, Han & Lee (2006) also discovered some knowledge by creating models for brand switching based on statistical methods for the South Korean mobile industry. They found a variety of interesting statistics relating to demographics and phone usage. These however, did not use a data-driven approach, instead relying on hypotheses developed by the researchers.

This paper presents how a data mining approach can be successfully applied to a marketing context to reveal patterns of brand switching. We discover valid, interesting and previously unknown brand switching patterns and other customer behaviour patterns using a variety of decision tree and decision forest algorithms, showing the power of a data-driven approach.

1.4   Dataset, Context and Pre-processing

The dataset used in this study has been adapted from the results of a survey conducted for Amaysim in conjunction with Macquarie University in the report "State of the Mobile Nation" (Gray et al. 2012). The survey asked mobile phone users, who had either changed their mobile phone service provider in the last 12 months or had considered doing so, a variety of questions regarding their phone usage, contentment with a variety of factors relating to their provider, and their demographics. Twelve months after the initial survey, the participants were again asked a variety of questions through a *follow-up* survey, and it was determined if they did indeed switch.

The initial survey contained 9 questions regarding the respondent's demographics, 21 questions regarding the nature of their phone plan, around 50 questions regarding the relationship they have with their current and previous providers, and some other miscellaneous questions. Demographic questions included "*What is your age?*", "*What is the highest level of education you have completed?*" and questions regarding relationship and familial status - all of which are categorical. Questions regarding the phone plan include "*Who is your current mobile phone provider?*" and "*What kind of mobile service do you have?*"; they are a mix of numerical and categorical answers. Relationship questions include "*In general, to what extent do you trust mobile phone carriers to respect the rules and regulations protecting consumers, with 0 being little/no trust and 10 being very trustworthy?*" and are largely numerical ratings provided by the respondent.

The *follow-up* survey conducted 12 months after the original survey was considerably smaller, and included some demographics questions, and questions related to the respondent's experience over the last year. The period of time between the two surveys allows us to analyse *confirmed* switchers as well as confirmed non-switchers (those who had thought about switching in the last survey, but still didn't switch). The dataset with follow-up survey attributes appended was considered separate to the original dataset, but had the same pre-processing applied.

The first step of pre-processing the dataset involves merging some sets of questions into single attributes. For example, a set of multiple-choice questions are often merged into a single attribute having domain values covering all multiple choice answers. After this, records with missing values were removed from the dataset. Care was taken so that only records with missing values that affected the analysis were removed, as some missing values were naturally accounted for by the answers to previous questions. We did not use missing value imputation techniques (Rahman & Islam 2013, Rahman & Islam 2014) since the fundamental assumption of these techniques is that the missing values appear at random and do not depend on any other factor/s, which is not the case for our data set.

In order to explore a wide variety of patterns we prepare a number of sub-datasets mostly by selecting different subsets of attributes and then applying six (6) data mining algorithms on the sub-datasets. The process undertaken for this and descriptions of the datasets created are discussed in Section 3.

1.5 Major Contributions of the Study

The major contributions of this study are as follows:

- *The use of a real life dataset on mobile phone usage and brand-switching focussing in particular on the Australian market* – As far as we know this is the only study of its kind on an Australian market dataset.

- *Application of six (6) well known data mining techniques on the real life dataset in order to explore the effectiveness of the techniques in discovering valid patterns* (Section 3) - Different data mining techniques use various methods for extracting rules. In order to discover useful knowledge, it is important to know which algorithms are best for extracting interesting rules.

- *Comparison of two groups of data mining techniques, decision forests and decision trees, in discovering valid patterns from the dataset.* (Section 4) - This provides insights as to how and when to use these, and demonstrates the different amounts of knowledge in the form of rules accessible to these groups.

- *Identification and presentation of a number of interesting patterns/trends (Section 5).*

- *Evaluation of the discovered trends from a marketing perspective (Section 6).*

- *Discussion of the use of data mining techniques in collaboration with traditional statistical techniques in marketing (Section 6).*

Section 2 briefly introduces some data-mining techniques relevant to this study. Section 3 presents our data analysis approach. Section 4 discusses the effectiveness of various data-mining techniques in discovering valid rules (i.e., patterns/trends), and Section 5 presents some interesting trends discovered by the data-mining techniques. Finally, concluding remarks and discussion of the study's relevance to the marketing industry are presented in Section 6.

2   Data Mining Techniques - A Brief Review

Data-mining is used over a wide variety of disciplines (Fayyad, Piatetsky-Shapiro, Smyth & Uthurusamy 1996, Schultz, Eskin, Zadok & Stolfo 2001, Berry & Linoff 2004). The term encompasses everything from the collection of data, to preprocessing and analysis through the supervised and unsupervised learning. Supervised learning uses labelled data in order to make predictions, and classification. Decision tree and decision forest algorithms are examples of classification algorithms.

2.1   What is a Classifier?

A dataset D contains a set of records $R = \{R_1, R_2, ..., R_{|D|}\}$, where $|D|$ is the total number of records. Each record of D has $|A|$ attributes $A = \{A_1, A_2, ..., A_m\}$. $A_j \in A$ is an attribute that can be either numerical or categorical. The domain of a numerical attribute $A_j = [l, u]$ varies between a lower limit l and an upper limit u. The domain size of the attribute is $|A_j| = [u - l + 1]$ for integer values. Similarly, the domain of a categorical attribute can be presented as $A_j = \{a_j^1, a_j^2, ..., a_j^d\}$, where the domain size is $|A_j| = d$. R_i is the i-th record and $R_{i,j}$ is the j-th attribute value of the i-th record. $R_{i,j}$ takes a value from the domain of A_j; that is $R_{i,j} \in A_j$.

One of the categorical attributes A_j (e.g. the attribute Brand-Switching with two possible values Switched and Didn't Switch) can be considered as the class attribute, the values of which are used for labelling the records. A record $R_i; \forall i$ has a class value *Switched* or *Didn't Switch*.

The classification task aims to discover the logic rules that determine whether a record should have the class value *Switched* or *Didn't Switch*. An example of a logic rule can be "IF Gender = Male AND Age ≤ 40 THEN Brand-Switching = Switched" meaning that if any record R_i is a male aged 40 years or younger then he is likely to switch his service provider. A classifier discovers and uses the rules to predict new, unlabelled data in the future. A *black-box* technique such as artificial neural networks does not provide any logic rules that are human understandable, whereas decision trees (Quinlan 1993) avoid this issue of a black box algorithm.

2.2   Decision Trees, Forests and Logic Rules

A decision tree (Quinlan 1986) discovers a set of logic rules, where each record R_i of a training dataset (i.e., the dataset which the decision tree is built from) obeys/falls in one and only one logic rule. An example of a decision tree is shown in Figure 1. This decision tree is assumed to be trained on an imaginary mobile-phone-user dataset, which has a class attribute with two possible classes: "Switched" and "Didn't Switch".
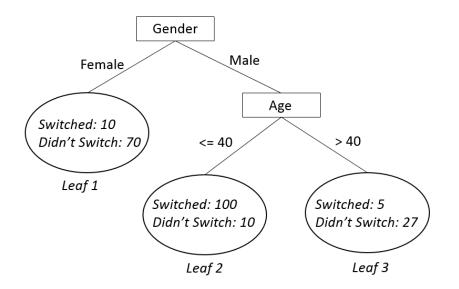
Figure 1: Example of a decision tree.

There are three leaves and each of them represents a logic rule. For example, Leaf 1 indicates a logic rule: *IF Gender = Female THEN Brand-Switching = Didn't Switch*. The validity of a rule is measured by two terms: support and confidence (Han, Kamber & Pei 2006). Support is defined as the ratio of the number of records falling into the leaf to the total number of records in the dataset. Confidence of a rule is defined as the ratio of the dominant class in a leaf to the other classes. The higher the support and confidence of a logic rule, the stronger (i.e., more valid) the rule is. A decision tree algorithm (Quinlan 1993) chooses the test attributes and corresponding splitting points (such as Age ≤ 40) *automatically* without requiring any input/assumption from a user/data miner.

A decision forest (Breiman 2001) being an ensemble of decision trees can typically discover higher number of valid and interesting patterns than a tree. It is important to discover those additional logic rules since based on them a service provider can get a better understanding on the reasons/trends of various important issues.

In this study we make use of the Random Forest (Breiman 2001), Bagging (Breiman 1996) and SysFor (Islam & Giggins 2011) forest algorithms as implemented by us. We also make use of the individual decision tree algorithms J48 (Quinlan 1993), SimpleCART (Breiman, Friedman, Olshen & Stone 1984) and REPTree (Witten & Frank 2005), all of which have implementations available in Weka (Hall, Frank, Holmes, Pfahringer, Reutemann & Witten 2009).

We classify logic rules/patterns into two main categories: valid and interesting. By *valid* rules we mean the rules that have high support and confidence. How high is high? We argue that this depends on the consideration of a domain expert or user. By *interesting* rules (Freitas 1999) we mean the rules that reveal interesting, previously unknown and non-obvious information. A number of interesting rules have been presented in Section 5.

3   Our Approach in Analysing the Dataset

Often an intelligent use of data mining techniques on a dataset can discover more (valid and interesting) rules/patterns than a straight forward application of data mining techniques. Therefore, we apply several data mining techniques on various sub-datasets prepared from the pre-processed dataset discussed in Section 1.4. Sub-datasets are prepared by selecting different class attributes and subsets of attributes.

We identify a set of attributes, each of which qualifies to be a class attribute in the sense that they have the ability to label the records. A class attribute is chosen to explore the causes of specific phenomena, such as switching. By selecting different class attributes for different sub-

datasets and applying a classifier on a sub-dataset, we force the classifier to choose logic rules in order to identify trends relating to the chosen class attribute. In this study we select the following class attributes: *Switching Status, Mobile Plan Type* and *Underspending*. Let us call the original pre-processed dataset (without follow-up questions) D, and the dataset with follow-up questions D'. The sub-datasets having Switching Status, Mobile Plan Type and Underspending as the class attributes are denoted by D_S, D_P and D_U, respectively.

We next use a variety of attribute subsets $A_x \subset A$, where $A_x; \forall x$ relates to a common question set, for example, demographics. The attribute subsets that we use in this study are based on *demographics, relationship with their telephone provider*, and *type of phone usage*. Additionally, we also select subsets of attributes (let us call them *class specific subsets*) that are related and specific to class attributes. Sub-datasets made of subsets of attributes relating demographics, relationship with their telephone provider, type of phone usage and class specific subsets are denoted by D^d, D^r, D^t and D^a, respectively.

Finally, we prepare sub-datasets using a subset of attributes and a class attribute for data analysis and knowledge discovery. For example, we prepare a sub-dataset with the *demographic* subset of attributes and *Switching Status* class attribute and denote the dataset by D^d_S. For each class attribute we choose a number of subsets of attributes and thus prepare a number of sub-datasets.

For the *Switching Status* class attribute we create 4 sub-datasets: D^d_S, D^r_S, D^t_S and D^a_S. Similarly, for the *Mobile Plan Type* class attribute we create 2 sub-datasets: D^d_P and D^a_P. For the *Underspending* class attribute we create 1 sub-dataset D^d_U. The design and selection of sub-datasets are our subjective decision.

For each sub-dataset we apply six (6) classification algorithms (i.e., data mining techniques) namely SysFor, Random Forest, Bagging, J48, SimpleCART, and REPTree. The code for SysFor is freely available online at http://csusap.csu.edu.au/~zislam/.

Based on the use of data mining algorithms on the sub-datasets and the whole dataset, we discover a huge set of logic rules from which we then select valid and/or interesting rules. Our definitions of valid and interesting rules are provided in Section 5. Additionally, we take advantage of the follow up survey (see Section 1.4) for a more useful data analysis. Let us denote the dataset containing *follow up* survey questions (in addition to the initial survey questions) by D'. We also prepare a dataset D'^{fu}, which only includes attributes from the follow up survey. We then prepare sub-datasets D'^a_S, D'^d_S, and D'^r_S, as well as D'^{fu}_S. All six data mining techniques are applied on these sub-datasets in order to explore valid and/or interesting rules.

We also take advantage of D' in identifying the customers who have definitely switched within the last 12 months (i.e., confirmed switchers) by carefully checking their mobile service providers in the initial survey and follow up survey. The class values of D'^a_S, D'^d_S, D'^r_S and D'^{fu}_S are then determined accordingly.

4   Data Mining Performance

In the process of conducting this study, a wide variety of logic rules regarding brand-switching and a variety of other key issues have been generated. From these rules, those which provide meaningful, non-trivial information are extracted. Rules that have high support and confidence are used to extract the meaningful information from the dataset.

In this study, the criteria of valid rules are user-defined, following the suggestions of previous literature (Freitas 1999). Rules with both high support and high confidence are considered to be valid. We use two different confidence thresholds: 70% and 90%, and two different support thresholds: 5% and 10%.

Table 1 shows the model performance over the generated sub-datasets from Section 3, grouped by D_S (Switching), D_P (Plan Type), D_U (Underspend) and D'_S (Confirmed Switchers). Each section of the table has the data-mining techniques in one of two groups, *Forest* or *Tree*. The average classification accuracy column shows the average classification accuracy over all sub-

datasets for a particular class attribute. For example, the Average Classification Accuracy column for SysFor in the *Switching* section represents the average classification accuracy on D^d_S, D^r_S, D^t_S and D^a_S. Similarly, the *No. Valid Rules Found* column represents the sum of valid rules (unique) found over these sub-datasets. Two identical valid rules are counted as one unique valid rule.

From Table 1, we can infer some interesting points. First, data mining techniques can discover a big number of valid and unique rules/patterns from the dataset. These rules can be used to classify the records against the class attributes, but are also the key in determining trends in the dataset. Second, forests can discover considerably more (valid and unique) rules/patterns than trees, according to our settings of valid rules. Third, classification accuracies of trees are similar to forests. So classification accuracies are not necessarily related to number of valid rules. Fourth, it appears (somewhat misleadingly) that RandomForest (and to a lesser extent Bagging) generate larger number of valid rules than SysFor. This does not necessarily indicate the superiority of these techniques for knowledge discovery over SysFor for the following reason. RandomForest and Bagging use bootstrap samples of the dataset on which they build trees. Since the bootstrap samples are often very similar to each other, they frequently discover very similar rules which have only slight differences in the splitting points of the attributes (despite the same attributes being tested). While technically "unique" rules, these are not necessarily useful more than once. This is in contrast to SysFor, for which the rules are diverse as the technique does not use bootstrap sampling.

Rules presented in Table 1 for each algorithm (such as SysFor and Bagging) are unique. That is, for Underspend SysFor discovers 62 unique and valid rules, but some of these rules can also be discovered by another algorithm. Therefore, Table 2 presents the total number of valid and unique rules found by all six algorithms for each class attribute. We can see that data mining techniques can discover a huge number of valid and unique rules from the dataset. The table also gives an idea on how quickly the number of valid and unique rules decreases with the increase of support and confidence. Table 2 provides an overview of the success of finding rules that fit our criteria of interestingness using all of the data-mining algorithms that are part of our methodology.

5 Observations and Findings Using Data Mining

In Section 4 we reported on the number of valid and unique rules selected from a huge set of rules that is obtained by the data mining techniques. In this section we present some *interesting* rules that are again selected from the huge set of rules. We group together similar rules into trends, which explain interesting components/observations of the data analysis. The findings reported in the following subsections are discovered automatically through data mining and not based on any hypothesis/assumption.

The survey was not originally designed by us. It only collected data on the customers who either switched their service providers within the last 12 months or considered switching seriously within the same time period. Hence, we only have information on the customers who considered switching and then switched, and the customers who considered switching but did not switch. We do not have any data on customers who did not consider switching their providers within last 12 months. The names of all telecommunication companies have been replaced with pseudonyms to maintain confidentiality.

In this study we use the data collected through the survey and extract knowledge from whatever data have already been collected. One of our goals is to demonstrate that data mining techniques can still discover valid and/or interesting rules from the dataset.

5.1 Demographics Based Analysis for Brand Switching

The goal of this section of the study is to test a variety of class attributes against only those attributes which represent demographic information, in order to explore how these class attributes (such as Switching Status) are dependent on the demographic information such as age, gender and place of residence. That is, we use the D^d_S dataset (see Section 3).

5.1.1 Trend: Age, Gender, and Brand-Switching

Out of 1392 records representing those surveyed who had either switched or considered switching, there were 149 participants (10.70% of the total participants) who were male and within the age bracket of 35 to 44. Interestingly, 73.83% of these respondents (i.e., 110 respondents) had switched in the last 12 months. That is, the male customers aged between 35 to 44 years have 73.83% probability of switching their provider once they consider switching.

A similar trend of switching is observed among the younger male participants. Male participants within the age bracket 18-24, 25-34 and 35-44 have 75.51%, 81.31% and 73.83% chance of switching respectively, once they consider it. Male participants aged between 18 and 44 have an overall 76.72% chance of switching once they consider it.

However, the trend of switching is significantly lower among the older male customers. Male customers within the age bracket 45-54, 55-64 and 65-above have only 58.27%, 52.87% and 52% chance of switching respectively, even when they consider doing it.

The overall chance of switching among the male customers aged between 18 and 44 is 76.72%, while the switching chance for male customers aged between 45 and above is only 54.23% (once they consider switching). In the dataset, there were 426 older (aged between 45 and above) male participants and 305 younger (aged between 18 and 44) participants.

From these results it is evident that males between the age of 18 and 44 are at considerable risk of switching. The providers may take different strategies for them as they are very vulnerable to switching. Older male customers are more stable than younger (less than 45 years old) male customers. Younger male customers being more switch-prone (vulnerable) may need more attention in order to increase retention.

The switching tendency among the women is significantly lower at only 37.97%, regardless of their age, compared to the young male customers (aged between 18 and 44) who have 76.72% chance of switching. Even within specific age brackets the switching tendency of women is lower than that of men. For example, once they consider switching the possibility of actually switching among men and women within the age bracket 25-34 are 81.31% and 40%, respectively. Similarly, within the age bracket 35-44 the actual switching possibility among men and women are 73.83% and 35.21%, respectively. Women that are married and with children have an even lower switching rate at only 25%.

It would seem that regardless of their age women are a lot more stable than young men in terms of the switching tendency. Perhaps the providers can have specific strategies for different groups to increase retention. The observations may help in building such strategies.

5.1.2 Trend: Professional Communities and Brand-Switching

The gender based difference of the switching pattern is also clearly evident within some professional communities. Once they have considered switching, the male respondents working in Finance and Insurance have 88.24% chance of actually switching, whereas female respondents working in the same sector have only 26.67% chance of actually switching. There are 17 male and 15 female respondents working in Finance and Insurance in the dataset.

A similar gender based difference in switching possibility exists among the students. While the possibility of actually switching within the male students is 62.5%, the actually switching possibility within the female students is 45%. Once they consider switching, the overall actual switching possibility among the students is only 52.78%. There are 32 male students and 40 female students.

Interestingly, the gender based difference is not observed among the Information Media and Telecommunication professionals. Within this community the actual switching possibilities for male and female (once they consider of switching) are 76% and 75%, respectively. While the switching tendency among the women is low overall, it is not the case for the women in Information Media and Telecommunication. It should be noted though that there were only

four women from the Information Media and Telecommunication profession surveyed, but the rules enjoy high confidence.

5.1.3   Trend: Marital and Familial Status, and Brand-Switching

Interestingly, it is observed that if male respondents aged between 55 and 64 are divorced then their actual switching possibility is 85%. In our dataset, there are 21 respondents within this category. Interestingly, 18 of them considered switching and actually switched in the 12 months prior to being surveyed.

On the other hand, men aged between 55 and 64 years who are married with their youngest child older than 5 years have an actual switching possibility of only 40%, even when they consider switching. There were 65 respondents within this group and 26 of them had switched in the last 12 months. This is even lower than the average for older men, which was only 54.23%.

Although generally older male customers are less likely to switch, their risk of switching increases dramatically if they are divorced (from 54.23% to 85%), and decreases dramatically if they are married with the youngest child aged 5 years or more (from 54.23% to 40%).

5.1.4   Trend: Income Ranges and Brand-Switching

Surprisingly, the analysis undertaken shows no patterns within different income ranges. While it has been previously observed that the patterns for switching or non-switching have been mainly driven by gender, this is not the case for income ranges - something which one would intuit differently. An interesting trend discovered however, is that women with an annual income of $40,000 and $49,999 have a lower switching possibility than most women at 21.93% compared to an overall average switching tendency of 37.97%.

Despite this being the case for almost all income ranges, a high non-switching tendency (87.5% possibility) is observed within the respondents having annual income between $100,000 and $109,000, and age between 55 and 64. There are 8 such participants, 7 of which did not switch in the last 12 months although they considered switching. When one compares this to the rest of men within this age bracket, who have a 47.13% non-switching tendency, this is a significant difference.

5.1.5   Trend: Education and Brand-Switching

Once they think of switching, the overall rate of switching for the university drop-outs is 31.51%. Those respondents with bachelor's degree as their highest education level who do not have children have an even higher actual switching rate at 37.25%. From this, that people who have been to university for a reasonable amount of time tend to switch a lot can be inferred. This finding suggests that advertising at universities and targeting university students to foster a sense of brand-loyalty may be a possible approach for the service providers.

5.2   Demographics Based Analysis for Properties Other Than Brand Switching

In this subsection we use the same demographic based attributes used in the previous study, but explore patterns other than just brand switching tendency. Therefore, we use some other attributes from the dataset as our class attribute in this section. This allows us to discover if there are any strong rules that predict other factors such as underspending and phone service type. These strong rules can be adapted into knowledge about these factors, just as has been done previously with brand-switching.

5.2.1   Trend: Relation between Underspending and Gender

A demographic analysis of the 1077 respondents who are not on a pre-paid plan shows that males are highly likely to underspend the "included value" on their plan. Males in this category have an underspending likelihood of 75.54%. This is opposed to females who have only 67% possibility of underspending. From this, it appears that men are slightly more likely to underspend the included value on their plans than women.

5.2.2 Trend: Mobile Service type and Familial status

A demographic analysis of 1540 survey-respondents pertaining to the type of mobile phone service shows that of all those customers with a child under 5 years old, a substantial amount of these (64.26%) are on a cap plan with a phone. The next largest service type for these customers is prepaid with 20.53% of respondents, a substantial difference. As a significant quantity (64.26%) of these customers are on a cap plan with a phone, it would make sense to directly market to people with young families for these services.

5.3 Relationship-Based Analysis for Brand-Switching

This section aims to explore interesting patterns of brand switching and the relationship a customer has with their phone company (and phone companies in general). The independent variables that are chosen in this section include those that represent various satisfaction metrics, number of times a customer has complained or had problems, and channels of complaint. The dependent variable (i.e., the class attribute) is Brand-Switching.

5.3.1 Trend: Perceived Industry Ethics and Brand-Switching

Interestingly those who believe that telecommunication service providers act ethically have a high switching tendency. Once they consider switching they have a switching probability of 65.73%. That is, 94 of these respondents switched after having considered switching while 49 of them did not ultimately switch. This finding makes it seem likely that despite customers being strongly convinced that their service providers are ethical, they may still be prone to switch. That is, just functioning ethically may not be good enough to retain customers.

Conversely it seems those who generally think that phone providers are ethical and trust that carriers follow the industry rules and regulations, and who haven't used their phone less frequently than usual in the past 12 months are very unlikely to switch (92.86% did not switch).

5.3.2 Trend: Brand-Switching and Decreased Plan Usage

A high quantity (48.48%) of confirmed switchers don't have a problem with their service, but claim that they use their service less. There are 31 customers who most strongly agreed that they use the service of the current provider less than the previous provider and 24 (77.42%) of them actually switched in the past 12 months.

5.3.3 Trend: Employment and Service Type

A very high quantity (70.11%) of customers surveyed who work in retail have a cap plan with phone. Also, a majority (83.39%) of respondents who claim to not be able to afford buying a phone outright are on cap plans with a phone. Very surprisingly, even if they do not consider phone companies to be extremely ethical, and cannot afford to buy a phone outright, but are happy with their monthly cost, they will most likely be on a cap plan with a phone (84% likely).

5.4 Other Relationship-Based Analysis

This study uses the relationship-based attributes to explore customer satisfaction and plan cost in relation to where plans are purchased.

5.4.1 Trend: Customer Satisfaction, Purchase Channel, and Plan-Cost

Of those respondents that purchased their service online and pay less than $6 a week for the service (75 customers), they are generally happy with the service and the length of the contract. This seems mostly due to the fact they are prepaid customers (52 of 75, 69.33%). Businesses probably need not focus resources on keeping these customers happy as they already largely satisfied, due to the nature of the service they have purchased.

5.5 Overall Preliminary Survey Analysis for Brand-Switching

Instead of analysing brand switching patterns based on groups of attributes such as demography and relationship related attributes, in this section we use a wide variety of independent variables.

5.5.1 Trend: Lifelong Switching Tendency

One of the strongest patterns is that although they have considered switching in the past 12 months, if a customer has never switched in the past then the possibility of his/her switching is only 1.15% - which is extremely low. There are 174 such respondents out of which only 2 switched in the last 12 months. This is reinforced by the fact that of the people surveyed in this category who were customers of one of the major established brands, none switched.

Another strong pattern is if customers have switched more than 12 times in the past then the probability of their switching is 100%. There are 9 such customers all of them have also switched in the last 12 months.

However, in comparison to this, those who have switched between 1 and 12 times in the past have a switching probability of 58.31% (705 switched and 504 did not switch).

Some customers have a strong predisposition to switching. If someone has switched more than 12 times in the past then he/she will switch again. Similarly those who have never switched remain unlikely to do so.

Previously it has been noted that male customers have a higher switching tendency. However, it is also possible to observe that even male customers who have never switched before have a low switching probability of only 1.75% (1 switched and 56 did not switch in the last 12 months). A similar trend is found among the female customers as well. Those who never switched before have a low switching tendency of 0.85% (1 switched and 116 did not switch).

The male customers who have switched before have a switching probability 68.84% (464 switched and 201 did not switch). In comparison, those female customers who switched before have a switching tendency of 45.96% (250 switched and 294 did not switch).

5.5.2 Trend: Company Size and Switching Probability

Our study shows that the switching probabilities for the major established telecommunication service providers are considerably lower than the smaller, less established providers. For the larger companies, listed here as *BigTelco#* there is an overall switching probability of only 27.29% if a respondent has considered switching in the last 12 months. This is especially low compared to the combined switching probability for the smaller providers, which was 85.2%. While in our survey dataset there are less customers for less established brands, the rate of customer attrition as a percentage of survey respondents is considerably higher in each case. For one group of the very smallest companies, this was close to 100% (these have been grouped together as *NicheCompetitors*). Other telecommunication companies have been listed as *OtherCompetitor#*, indicating that they are mid-sized. The individual results for each provider are as follows:

- For *BigTelco1*, the switching probability is 0% (0 switched and 4 did not switch).

- For *BigTelco2*, the switching probability is 23.14% (28 switched and 93 did not switch).

- For *BigTelco3*, the switching probability is 39.47% (150 switched and 230 did not switch).

- For *BigTelco4*, the switching probability is 44.13% (109 switched and 138 did not switch).

- For *BigTelco5*, the switching probability is 54.74% (150 switched and 124 did not switch).

- For *BigTelco6*, the switching probability is 56.98% (49 switched and 37 did not switch).

- For *OtherCompetitor1*, the switching probability is 58.62% (17 switched and 12 did not switch).

- For *OtherCompetitor2*, the switching probability is 71.43% (20 switched and 8 did not switch).

- For *OtherCompetitor3*, the switching probability is 75% (6 switched and 2 did not switch).

- For *OtherCompetitor4*, the switching probability is 80% (32 switched and 8 did not switch).

- For the *NicheCompetitors* group, the switching probability is 97.41% (113 switched and 3 did not switch).

One of the mid-sized competitors has discontinued their brand, and its services were taken over by one of the major established providers since the survey was conducted.

5.6 Overall Preliminary Survey Analysis for Other Factors

In this section we use the same set of attributes as in the previous section, but we change the class attributes to find patterns on other factors such as the method of plan purchase and contract-length.

5.6.1 Trend: Method of Purchase

Customers who purchase services from supermarkets are overwhelmingly on "prepaid services" (93.24%). Of these prepaid customers, most have not upgraded or downgraded their service in the last 12 months (89.86%).

Customers who report that they would not buy a handset outright as they could not afford it are overwhelmingly on "cap plans with phones" (81.25%). These customers account for 36.62% of all of those on "cap plans with phones".

There are less purchases (23) of "broadband bundles" "over the phone" than there are "cap plans with phones" (116). However, "over the phone" sales account for 26.7% of all bundle sales. Another 36.05% of broadband bundles are sold in telecommunication carrier stores. "Over the phone" sales seem to be one of the most useful methods to obtain broadband bundle customers. Inside telecommunication carrier stores is also a great place to obtain these customers.

The majority (63.04%) of non-prepaid services purchased online who pay more than $6 a week are for "cap plans with phones". Most "over the phone" sales of services are for "cap plans with phones". Of 163 customers that purchased their plans this way, 116 purchased "cap plans with phones" (71.17%). This accounts for 16.34% of all "cap plan with phone" purchases in the dataset used which contains 1540 records.

5.6.2 Trend: Contract Length and Customer Happiness

45% of customers that pay between $3 and $10 a week for their cap plan with phone are unhappy with the length of their contract. 42.5% of all customers are unhappy with the length of their contract. Customer retention may be increased by allowing more flexible contract lengths in order to increase customer satisfaction. Almost half of these low cost cap plans are unhappy due to their contract length.

25.71% of prepaid customers that pay between $3 and $10 a week are unhappy or ambivalent with the quality of the phone, although are willing to pay for a phone outright to avoid entering into a contract.

Finally, the overwhelming majority (70%) of people happy with their contract length who make more than 10 calls a week are on cap plans with phones. Of this group most (95.4%) are happy or ambivalent to their SMS rates.

Class Attribute	Algorithm Type	Technique	Avg. Classification Accuracy	No. Valid and Unique Rules Found
Switching	Forest	SysFor	60.19	12
		RandomForest	61.34	6
		Bagging	61.14	4
	Tree	J48	61.17	0
		REPTree	72.97	3
		SimpleCart	70.69	0
Plan Type	Forest	SysFor	71.54	53
		RandomForest	68.89	70
		Bagging	69.04	38
	Tree	J48	74.21	0
		REPTree	69.57	3
		SimpleCart	71.19	0
Underspend	Forest	SysFor	69.01	62
		RandomForest	68.83	59
		Bagging	68.88	33
	Tree	J48	69.47	0
		REPTree	70.26	2
		SimpleCart	69.47	0
Confirmed Switchers	Forest	SysFor	80.24	124
		RandomForest	76.21	197
		Bagging	76.39	82
	Tree	J48	80.19	0
		REPTree	77.02	1
		SimpleCart	76.2	0

Table 1: Analysis of Data Mining Algorithm Performance

Class Attribute	C=0.7 S=0.05	C=0.7 S=0.1	C=0.9 S=0.05	C=0.9 S=0.1
Switching	17	3	7	2
Plan Type	155	76	73	30
Underspend	71	24	0	0
Confirmed Switchers	244	132	14	3

Table 2: Sums of unique rules found for each class attribute

5.6.3 Trend: Low Plan Usage and Customer Service

The majority (76.84%) of customers that only make 1 or 0 calls on average per week are prepaid customers. 97.89% of these customers indicate that customer service is important to them when changing mobile phone providers.

5.7 Follow-Up Study Based Analysis

The original customers surveyed were surveyed once again a year later, and data was collected based on this. In this study, only the data available from the questions asked in this follow-up survey was used.

5.7.1 Trend: Issues Surrounding Confirmed Switchers

People who claim to be happy with their current service provider, but who say that mobile service providers did not live up to their expectations at all over the last 12 months, are all confirmed switchers.

40% of confirmed switchers claimed to have had network issues that lead to the switch. It would appear that network issues are a major component of why people switch.

5.7.2 Trend: High Plan Usage and Switching

Those pre-paid customers that make 13 or more calls a week were shown to be 70.87% likely to switch.

5.7.3 Trend: Fixed Line Bundles

Most people (90.41%) with a fixed line bundle did not switch.

6 Conclusion

In this study we analyse the brand-switching and customer-behaviour patterns through data mining techniques without using any other switching models and domain knowledge. We find that the discovered patterns are interesting, valid and useful when we evaluate them from a marketing point of view. We also compare this to highly used statistical methods from the marketing community, and find that data-mining techniques, especially decision forests, are useful for finding "deep" patterns within the data and provide a compliment to existing analyses.

The lower rates of switching by women compared to men, can be explained by organisational commitment theory, which suggests women are more likely to be loyal to an organisation than men (Wahn 1998). Mobile phone switching research also suggests that men are more likely to consider alternative offers than women, and therefore would be more likely to switch (Wang & Acar 2006).

The pattern that suggests that older (45 years and over) male customers have lower switching tendency can be explained by relationship age with the provider (Ranaweera & Menon 2008). Older males are more likely to have longer relationships with a provider and are less likely to switch.

We discover in this paper that better educated consumers are more likely to switch. A possible reason behind the pattern is the fact that educated people have better capacity to compare alternative offers.

The research clearly shows the importance of service failures in switching. Around 40% of consumers switch because of network issues. Interestingly customer service becomes more of an issue for consumers who have a low usage levels. For consumers who make no more than one call per week, poor service is almost the only reason to switch providers. The relationship here may also be the other way round, as research suggests that consumers who encounter poor service, are less satisfied and will therefore use the service less prior to switching (Bolton & Lemon 1999). Therefore, providers should focus on consumers who make fewer calls, as they are quite likely to be dissatisfied and are also likely to switch.

Our study also discovers that switching history of a customer can be important in analysing future switching behaviour. It was found in the previous research that frequent switchers in the past are likely to change providers again (Roos & Gustafsson 2007). This is because such consumers feel empowered to do so and act in more active fashion when selecting their

providers. It is important for telecommunications marketers to know previous switching history, particularly if the consumer had switched in the 12 months prior to making the change. This indicates that such consumers are problematic, in terms of repeat purchases.

The last and perhaps most interesting finding of our study is the lower switching rates (ranging between 0% and 39.47%) among the big three telco providers. We also find that the switching rates among small telco providers are high ranging between 56.98% and 100%. This pattern reflects the notion of double jeopardyï¿½, whereby brands with low share, such as these small telcos, are associated with lower brand loyalty, than companies with larger market shares (Bongers & Hofmeyr 2010, Labeaga-Azcona, Lado-Couste & Martos-Partal 2010).

Overall this research shows that patterns of provider switching are not uniform in the market and that reasons for switching are heterogeneous depending on customer background and company type. We believe that unlike models of switching which aim to fit an overall explanation to a market or dataset (Bansal & Taylor 1999, Colgate & Hedge 2001, Colgate & Lang 2001, Keaveney 1995, Lees, Garland & Wright 2007, Levesque & McDougall 1993), this study provides a fruitful avenue for the development of a number of explanations of consumer behaviour in this context. Importantly, our results show consistency of empirical generalisations of "double jeopardy", found in other brand switching studies, which suggests that data mining approach can provide consistent results with that which use a more complex theoretical explanation.

References

Ahn, J. H., Han, S. P., & Lee, Y. S. (2006). Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry. *Telecommunications Policy*, **30**(10), 552-568.

Annual Report of the Telecommunications Industry Ombudsman (n.d.), http://www.tio.com.au/. Accessed: 2015-04-24.

Bansal, H. S. & Taylor, S. F. (1999), 'The service provider switching model (spsm)', *Journal of Service Research* **2**(2), 200 – 218.

Bernard, S., Adam, S. & Heutte, L. (2012), 'Dynamic random forest', *Pattern Recognition Letter* **33**(12), 1580 – 1586.

Berry, M. J. & Linoff, G. S. (2004), *Data mining techniques: for marketing, sales, and customer relationship management*, John Wiley & Sons.

Bolton, R. & Lemon, K. (1999), 'A dynamic model of customers' usage of services: Usage as an antecedent and consequence of satisfaction', *Journal of Marketing Research* **36**, 171 – 186.

Bongers, M. & Hofmeyr, J. A. N. (2010), 'Why modeling averages is not good enough', *Journal of Advertising Research* **50**(3), 323 – 333.

Breiman, L. (1996), 'Bagging predictors', *Machine Learning* **24**(2), 123–140.

Breiman, L. (2001), 'Random forests', *Machine Learning* **45**(1), 5–32.

Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. (1984), *Classification and Regression Trees*, Wadsworth International Group, Belmont, California.

Colgate, M. & Hedge, R. (2001), 'An investigation into the switching process in retail banking services', *International Journal of Bank Marketing* **19**(5), 201–212.

Colgate, M. & Lang, B. (2001), 'Switching barriers in consumer markets: An investigation of the financial services industry', *Journal of Consumer Marketing* **18**(4), 332 – 347.

Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P. & Uthurusamy, R. (1996), *Advances in knowledge discovery and data mining*, The MIT Press.

Freitas, A. A. (1999), 'On rule interestingness measures', *Knowledge-Based Systems* **12**(5), 309–315.

Geurts, P., Ernst, D. & Wehenkel, L. (2006), 'Extremely randomized trees', *Machine Learning* **63**, 3 – 42.

Gray, D., D'Alessandro, S. & Carter, L. (2012), *The state of the mobile nation*, Technical report, Faculty of Business and Economics, Macquarie University.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. & Witten, I. H. (2009), 'The weka data mining software: an update', *ACM SIGKDD Explorations Newsletter* **11**(1), 10–18.

Han, J., Kamber, M. & Pei, J. (2006), *Data mining: concepts and techniques*, Morgan Kaufmann.

Ho, T. K. (1998), 'The random subspace method for constructing decision forests', *IEEE Transaction on Pattern Analysis and Machine Intelligence* **20**(1), 832 – 844.

Hung, S. Y., Yen, D. C., & Wang, H. Y. (2006). Applying data mining to telecom churn management. *Expert Systems with Applications*, **31**(3), 515-524.

Islam, M. Z. & Giggins, H. (2011), Knowledge discovery through sysfor: a systematically developed forest of multiple decision trees, *in 'Proceedings of the Ninth Australasian Data Mining Conference-Volume 121'*, Australian Computer Society, Inc., pp. 195–204.

Keaveney, S. M. (1995), 'Customer switching behavior in service industries: An exploratory study', *Journal of Marketing* **59**(2), 71 – 82.

Labeaga-Azcona, J. M., Lado-Couste, N. & Martos-Partal, M. (2010), 'The double jeopardy loyalty effect using discrete choice models', *International Journal of Market Research* **52**(5), 633 – 652.

Lees, G., Garland, R. & Wright, M. (2007), 'Switching banks: Old bank gone but not forgotten', *Journal of Financial Services Marketing* **12**(2), 146 – 156.

Levesque, T. J. & McDougall, G. H. (1993), 'Managing customer satisfaction: The nature of service problems and customer exit, voice and loyalty', *Asia Pacific Journal of Quality Management* **2**(2), 40–58.

Mobile phone statistics (n.d.), http://www.portal.euromonitor.com/Portal/Pages/Statistics/Statistics.aspx. Accessed: 2015-04-24.

Quinlan, J. R. (1993), *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Mateo, CA.

Rahman, M. G. & Islam, M. Z. (2013), 'kDMI: A Novel Method for Missing Values Imputation Using Two Levels of Horizontal Partitioning in a Data set', in *Proc. of the 9th International Conference on Advanced Data Mining and Applications* (ADMA 13), Hangzhou, China, 14-16 December 2013, pg. 250-263. (Eds.).

Rahman, M. G. & Islam, M. Z. (2014), 'Fimus: A framework for imputing missing values using co-appearance, correlation and similarity analysis', *Knowledge-Based Systems* **56**, 311–327.

Ranaweera, C. & Menon, K. (2008), 'For better or for worse: Moderating effects of relationship age and continuance commitment on the service satisfaction-word-of-mouth relationship', in *Advances in Consumer Research - Latin American Conference*, Vol. 2.

Roos, I. & Gustafsson, A. (2007), 'Understanding frequent switching patterns', *Journal of Service Research* **10**(1), 93 – 108.

Schultz, M. G., Eskin, E., Zadok, E. & Stolfo, S. J. (2001), 'Data mining methods for detection of new malicious executables', in *Proceedings of the IEEE Symposium on 'Security and Privacy, 2001.* IEEE, pp. 38–49.

Vinzi, V. E., Trinchera, L., Squillacciotti, S. & Tenenhaus, M. (2008), 'Rebus-pls: A response-based procedure for detecting unit segments in pls path modelling', *Applied Stochastic Models in Business and Industry* **24**(5), 439 − 458.

Wahn, J. C. (1998), 'Sex differences in the continuance component of organizational commitment', *Group and Organization Management* **23**(3), 256 − 266.

Wang, A. & Acar, S. A. (2006), 'Information search and mobile promotions', *International Journal of Mobile Marketing* **1**(2), 80 − 87.

Wei, C. P., & Chiu, I. T. (2002). Turning telecommunications call details to churn prediction: a data mining approach. *Expert Systems with Applications*, **23**(2), 103-112.

Witten, I. H. & Frank, E. (2005), *Data Mining: Practical machine learning tools and techniques*, Morgan Kaufmann.