Loyal to your city? A data mining analysis of a public service loyalty program

Sofie De Cnudde *, David Martens

Department of Engineering Management, Prinsstraat 13, B-2000 Antwerp, Belgium

ARTICLE INFO

ABSTRACT

Customer loyalty programs are largely present in the private sector and have been elaborately studied. Applications from the private sector have found resonance in a public setting, however, simply extrapolating research results is not acceptable, as their rationale inherently differs. This study focuses on data from a loyalty program issued by the city of Antwerp (Belgium). The aim of the loyalty card entails large citizen participation, however, an active user base of only 20 % is reached. Predictive techniques are employed to increase this number. Using spatial behavioral user information, a Naive Bayes classifier and a Support Vector Machine are used which result in models capable of predicting whether a user will actively use its card, whether a user will defect in the near future and which locations a user will visit. Also, a projection of spatial behavioral data onto even more fine-grained spatio-temporal data is performed. The results are promising: the best model achieves an AUC value of 92.5 %, 85.5 % and 88.12 % (averaged over five locations) for the predictions, respectively. Moreover, as behavior is modeled in more detail, better predictions are made. Two main contributions are made in this study. First, as a theoretical contribution, fine-grained behavioral data contributes to a more sound decision-making process. Second, as a practical contribution, the city of Antwerp can now make tailored strategic decisions to increase its active user base.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Since the evolution from a product-centric to a customer-centric approach, managing customer relationships has become a necessity for organizations [1]. Customers need to be understood on a per-customer basis in terms of their needs, preferences and potential [2,3]. Maintaining close ties to customers may lead to a competitive advantage and to customer loyalty [4]. In order for organizations to initiate and maintain connections to their customers, Customer Relationship Management (CRM) systems are used. Data mining forms an integral part of these systems and provides insight into raw customer data [5]. Consequently, organizations can closely monitor their customers and anticipate accordingly through proactive marketing such as offering personalized incentives [5]. An important source for CRM data constitutes the output of loyalty programs and more specifically from loyalty cards [6]. Loyalty card data is typically characterized by a large sample of transactional data connected to the customers, for which also socio-demographic data is available [7].

The analysis of loyalty programs has been largely researched in literature related to the profit sector, especially in retail [8]. In contrast, the behavior of customers in a non-traditional customer–supplier relationship such as the relationship between a customer and the government has not received that kind of attention [2]. The results from analyses performed in the profit sector cannot simply be extended to the non-profit sector due to their differing rationale. To our knowledge, the analysis of loyalty card data in the context of a public institution has not yet been researched. This study performs an analysis of such data and thereby is a first attempt to contribute to this gap. We state the contributions as follows: (1) gaining insight into the behavior of users of a government loyalty card and (2) attempting to learn non-trivial user information from this behavior with the help of data mining techniques. Both should contribute to the design of a decision support system able to lead to high-quality information for decision makers.

Concretely, in this design science paper [9], we analyze the behavior of users of the *A-Card*, a loyalty card issued by the city of Antwerp (Belgium) that can be used in public institutions and partnering institutions such as libraries and musea. The purpose of the card consists of promoting participation in cultural services offered by the city. The data consists of 4 million transactions concerning approximately 177,000 persons visiting one of 102 locations. Currently, there is no widespread use of the benefits that come along with the loyalty card. Moreover, users visit a relatively unilateral set of locations. This research presents an analysis of this government data that should empower the city to take appropriate strategic decisions.

The rest of this paper is organized as follows. Section 2 discusses relevant background and previous related work. Section 3 gives a detailed

* Corresponding author. Tel.: +32 3 265 50 14.
E-mail addresses: sofie.decnudde@uantwerp.be (S. De Cnudde), david.martens@uantwerp.be (D. Martens).

description of the data, the used data mining techniques and the evaluation criteria. Next, in Section 4, the experimental set-up of the research is presented. The results of the analysis are presented and discussed in Section 5. Lastly, Section 6 concludes our work and presents avenues for further research.

2. Background and related work

2.1. CRM

In the 1990s, organizations' customer focus shifted from a transactional view to a relational view [10]. CRM has since become a pivotal part of organizations. An understanding of customers' preferences may lead to the possibility of tailoring products and services to their wishes, which leads to higher customer satisfaction. This subsequently results in behavioral customer loyalty [11] and ultimately in profit. In order for these CRM systems to be of any value for the organization, data has to be collected and techniques have to be available to analyze this data [12]. Loyalty programs are one way of tying customers to an organization through incentives [4]. They are also a source for capturing behavioral patterns of the users which results in a large amount of transactional data.

2.2. High-dimensional, behavioral data

Since data collection and storage have become cheap, data is being gathered in vast amounts [13,14]. This data often consists of fine-grained behavioral data, such as click behavior of website visitors, payment transactions of a client of a bank or locations visited by a mobile phone user. These fine-grained features are testimonies of an individual's behavior. Using this data, predictions can be made concerning individual users as to whether they would be interested in a certain news article [15], a certain banking product [16] or a mobile ad [17]. An application in the public sector entails fraud detection in companies using their payment data [18,19]. These big high-dimensional feature sets lead to more complexity in the analysis [20], for which traditional dimensionality reduction techniques are not relevant given the low redundancy of the features. Traditional summarizing features such as socio-demographic variables contain more information on a per feature basis in comparison to behavioral features. The latter, however, each gradually add more information which leads to a more informative whole [21].

2.3. Analysis techniques in CRM

Once the data is collected, techniques have to be identified to extract information and knowledge. Data mining techniques are techniques that are able to extract patterns from the data which can subsequently lead to vital knowledge for the organization [22]. Ngai, Xiu and Chau performed a comprehensive literature review of data mining studies in the CRM field [8]. Here, the CRM field is subdivided into four stages: identification, attraction, retention and development of customers. The bulk of the research is positioned in the customer retention phase, and more specifically in personalizing marketing campaigns dependent on user's behavior (one-to-one marketing) and with the goal of tying customers to the organization (loyalty programs). The techniques mostly used in these studies consist of neural networks, decision trees and association rule techniques. Applications of data mining techniques in the context of CRM can be found in different sectors such as retail, the banking sector and telco companies [22]. In retailing, a typical application is market basket analysis in order to determine which products are frequently bought together. The banking sector is mainly interested in segmenting its customers in order to focus their marketing efforts towards the right customers. Another important application is fraud detection where suspicious transactions are identified. In the

telecommunications sector, churn prediction is of vital importance due to the highly competitive environment.

2.4. CRM in a public setting

The public sector deals with distinct challenges with respect to its private counterpart, among others providing basic services to its citizens. Lately, their modernization process has resulted in them adopting a more privatized character [23] along with private-like principles such as a focus on efficiency, competitiveness and profitability [24]. As time evolves, this transformation will result in applications originating from the private sector, migrating towards the public sector [25]. CRM applications in the public field, referred to as e-government, have until now mostly focused on providing equally accessible, transparent and omnipresent public services [26], on citizen engagement [27] and on policing [28]. To this end, diverse data can be used such as citizen feedback and complaints, social media data [29] and government data from legacy systems. The data traditionally used in big data government applications, however, is mostly textual [30]. Introducing behavioral data of some sort could open a new avenue of research and lead to surprising results, as it did in advertising research.

Despite some resemblances between the public and the private sector, two major differences have to be kept in mind. First, the 'customer' base of public services is diverse and dynamic [31,32]. Namely, the 'customer' base encompasses all citizens that belong to the governing institution. The customer base in the private sector only is a portion of this citizen base. Thus, the data collection opportunities are intrinsically larger and more diverse [33]. Secondly, the 'customers' cannot be valued based on profitability. In case of CRM in a government setting, customers must not only be served on an individual basis, an extra equality constraint is introduced [2]. Although the principles of CRM can be applied across sectors and industries [34,35], care must thus be taken as to whether results from research in a private setting can be applied to its public counterpart [35]. A government-issued loyalty card cannot be evaluated solely on its financial benefits for the institution. The element of so-called public value has to be taken into account [36].

2.5. Relevance

Underpinning the relevance of our study are the following arguments. First, the public sector is an omnipresent institution in everyday life and offers its services to all aspects of society with a diverse set of stakeholders with whom unique relationships are held, i.e. no profitability element is involved [24]. This distinct character results in interesting research opportunities. In the past, research concerning strategy [31], management practices [37], information technology [38] and information systems [39] has been conducted in the public sector and has been compared with practices in the private sector. Extending this research effort to loyalty programs may result in interesting discoveries. Secondly, the use of CRM in the public sector is quite unilateral at present. Using fine-grained behavioral data has proven to lead to good predictive results [21] and has been employed in advertising research [17,15,16,40]. Hence, investigating whether and to what extent these results are feasible with behavioral data originating from the public sector may give rise to interesting research conclusions. Finally, concerning the overall relevance of this work with respect to the government's cultural responsibility, cultural participation can help connect individuals and lead to civic engagement and social cohesion [41]. One of the goals of the *A-Card* is to encourage people to take part in the cultural life of the city of Antwerp and to this end analyze the behavior of users visiting public institutions in order to tailor the cultural offer. Getting insight into this behavior may thus lead to benefits for the city as a whole.