Customer Profiling based on Electronic Payment Transaction Data

Michiel Van Lancker[1], Annemie Vorstermans[1], and Mathias Verbeke[2]

[1] Faculty of Engineering Technology, KU Leuven, Belgium
[2] EluciDATA Innovation Lab, Sirris, Belgium

Abstract. Customer profiling allows companies to profile a customer or group of customers based on their transaction behavior. In this study, transactions of electronic payment cards are considered. Contrary to classic market basket analysis, this transactional data does not contain information about the products bought, but only the amount of the payment, as well as a number of details on the respective shop. In order to create meaningful customer profiles based on this limited amount of information, four different techniques were compared. While the construction of customer profiles for groups of customers was possible, it proved difficult to thoroughly validate the profiles and methods used without extra reference data.

Keywords: Transactional data mining, Customer Profiling, Clustering

1 Introduction

In recent years, data mining of transactional data has become widely implemented in a variety of businesses. The practice of building customer profiles from transactional data allows companies to profile a customer or group of customers based on their characteristics or behavior. These profiles can subsequently be used for different purposes, e.g., to improve customer service or tailor marketing campaigns. Essential to capture the true nature of the shopping behavior of a customer, is the extraction of relevant features [1]. Contrary to classic market basket analysis, in which information is available on the products bought, in this study the problem of customer profiling is considered in which only a limited amount of details are available.

2 Method

To construct profiles that represent the shopping behavior of groups of customers, relevant customer features need to be extracted from the data. Because no information on the products bought is present in the data set, transaction characteristics were transformed into customer features with a frequency-based approach, resulting in the extraction of 26 relevant features from four transaction characteristics, namely the time, location, transaction amount and shop type of a transaction [2]. To improve the quality of the data, two major changes were made. Firstly, shops were categorized per shop type based on similarity of transaction pattern and presence of keywords in the name of the shop. Secondly, because transactions can be performed in both a work-related or a leisure-related context, a context-based geographical representation for the customers was developed. Based on the resulting features, customers were clustered using two different algorithms: K-means clustering and agglomerative clustering. For each of these techniques, two approaches were tried: A first approach clustered once based on all customer features, while the second approach combined clusterings based on disjoint sets of customer features into one final clustering, similar to bottom-up subspace clustering [3]. The resulting profile per cluster summarizes the behavior of a customer as the difference (in number of standard deviations) of the cluster mean customer from the population mean customer for each feature.

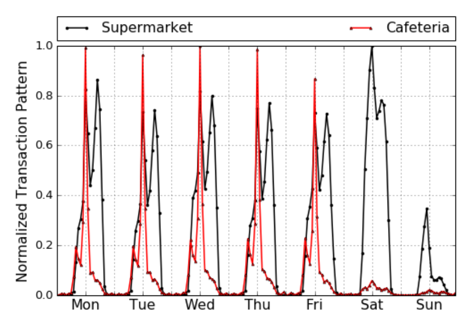

Fig. 1: Transaction patterns of shop categories show clear differences

3 Results

To analyze the results of different techniques, a comparison was made both between algorithms and between approaches. While K-means clustering slightly outperformed agglomerative clustering based on cluster quality, the difference was small and did not entail a better representation of the behavior of customer groups. Agglomerative clustering however offered multiple implementation advantages. Because of the deterministic hierarchical tree it constructs, the algorithm can be run once to build the tree and then generate multiple sets of clusters according to the needs of specific applications. Thus, agglomerative clustering can generate different levels of profiling without having to rerun the whole algorithm. K-means clustering does not offer this freedom. Furthermore, the effectiveness of K-estimation declines for data of higher dimensions and is thus less effective when clustering based on all customer features at once.

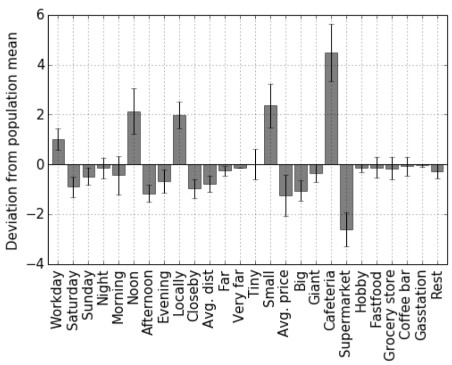

Fig. 2: The profile of a group of customers, showing the deviation from the population mean for all features, expressed in standard deviations

When comparing the results of clustering based on all customer features and clustering by combining clusterings based on disjoint subsets of features, multiple differences were observed. While the main advantage of the first approach is its simplicity, it is sub-optimal and generates a small number of clusters with rather generic profiles. Because the second approach is much more thorough, it is capable of finding more specific profiles. This is however paired with an explosion of the number of clusters: for K-means clustering 991 clusters were found, for agglomerative clustering 1322 clusters. It should be noted though that many clusters are very small and can be disregarded if needed. Overall, this approach can identify more specific customer behavior with only a slight extra effort compared to clustering based on all customer features. Additionally, this approach offers more freedom to the designer too: because clusterings from different subsets of features can be combined in multiple ways, different angles to customer profiling can be examined with only one run of the algorithms. The resulting customer profiles were evaluated in a qualitative way using domain knowledge. A quantitative analysis would be valuable to get deeper insights in the obtained results. This would however require an extensive customer study with manually labeled data which was out of the scope of this study.

4 Conclusion

Four different techniques for customer profiling based on transactional data from electronic payment cards were compared, in which only the transaction amount and a limited amount of information on the respective shop was available. Combined agglomerative clustering using subsets of customer features turned out to be the most interesting, because the qualitative analysis showed promising results and it offers a lot of freedom to the designer to tailor it to the needs of specific applications.

References

1. Jain, A. K.: Data clustering: 50 years beyond K-means. Pattern recognition letters, 31(8), 651-666 (2010).
2. Leonard, M., Wolfe, B.: Mining transactional and time series data. SUGI, 10-13 (2005).
3. Parsons, L., Haque, E., Liu, H.: Subspace clustering for high dimensional data: a review. ACM SIGKDD Explorations Newsletter, 6(1), 90-105 (2004).