Behavioral Pattern–Based Customer Segmentation

Yinghui Yang
University of California, Davis, USA

INTRODUCTION

Customer segmentation is the process of dividing customers into distinct subsets (segments or clusters) that behave in the same way or have similar needs. Because each segment is fairly homogeneous in their behavior and needs, they are likely to respond similarly to a given marketing strategy. In the marketing literature, market segmentation approaches have often been used to divide customers into groups in order to implement different strategies. It has been long established that customers demonstrate heterogeneity in their product preferences and buying behaviors (Allenby & Rossi 1999) and that the model built on the market in aggregate is often less efficient than models built for individual segments. Much of this research focuses on examining how variables such as demographics, socioeconomic status, personality, and attitudes can be used to predict differences in consumption and brand loyalty. Distance-based clustering techniques, such as k-means, and parametric mixture models, such as Gaussian mixture models, are two main approaches used in segmentation. While both of these approaches have produced good results in various applications, they are not designed to segment customers based on their behavioral patterns.

There may exist natural behavioral patterns in different groups of customers or customer transactions (e.g. purchase transactions, Web browsing sessions, etc.). For example, a set of behavioral patterns that distinguish a group of wireless subscribers may be as follows: Their call duration during weekday mornings is short, and these calls are within the same geographical area. They call from outside the home area on weekdays and from the home area on weekends. They have several "data" calls on weekdays.

The above set of three behavioral patterns may be representative of a group of consultants who travel frequently and who exhibit a set of common behavioral patterns. This example suggests that there may be natural clusters in data, characterized by a set of typical behavioral patterns. In such cases, appropriate "behavioral pattern-based segmentation" approaches can constitute an intuitive method for grouping customer transactions.

BACKGROUND

The related work can be categorized into the following groups.

Market Segmentation

Since the concept emerged in the late 1950s, segmentation has been one of the most researched topics in the marketing literature. There have been two dimensions of segmentation research: segmentation bases and methods. A segmentation basis is defined as a set of variables or characteristics used to assign potential customers to homogenous groups. Research in segmentation bases focuses on identifying effective variables for segmentation, such as socioeconomic status, loyalty, and price elasticity (Frank et al 1972). Cluster analysis has historically been the most well-known method for market segmentation (Gordon 1980). Recently, much of market segmentation literature has focused on the technology of identifying segments from marketing data through the development and application of finite mixture models (see Böhning (1995) for a review). In general model-based clustering (Fraley & Raftery 1998; Fraley & Raftery 2002), the data is viewed as coming from a mixture of probability distributions, each representing a different cluster.

Pattern-Based Clustering

The definition of pattern-based clustering can vary. Some use this term to refer to clustering of patterns, e.g. pictures and signals. Others discover patterns from

the objects they are clustering and use the discovered patterns to help clustering the objects. In the second scenario, the definition of a pattern can vary as well. Wang et al (2002) considers two objects to be similar if they exhibit a coherent pattern on a subset of dimensions. The definition of a pattern is based on the correlation between attributes of objects to be clustered. Some other approaches use itemsets or association rules (Agrawal et al., 1995) as the representation of patterns. Han et al., (1997) addresses the problem of clustering-related customer transactions in a market basket database. Frequent itemsets used to generate association rules are used to construct a weighted hypergraph. Each frequent itemset is a hyperedge in the weighted hypergraph, and the weight of the hyperedge is computed as the average of the confidences for all possible association rules that can be generated from the itemset. Then, a hypergraph partitioning algorithm from Karypis et al., (1997) is used to partition the items such that the sum of the weights of hyperedges that are cut due to the partitioning is minimized. The result is a clustering of items (not transactions) that occur together in the transactions. Finally, the item clusters are used as the description of the cluster and a scoring metric is used to assign customer transactions to the best item cluster. Fung et al., (2003) used itemsets for document clustering. The intuition of their clustering criterion is that there are some frequent itemsets for each cluster (topic) in the document set, and different clusters share few frequent itemsets. A frequent itemset is a set of words that occur together in some minimum fraction of documents in a cluster. Therefore, a frequent itemset describes something common to many documents in a cluster. They use frequent itemsets to construct clusters and to organize clusters into a topic hierarchy. Yiu & Mamoulis (2003) uses projected clustering algorithms to find clusters in hidden subspaces. They realized the analogy between mining frequent itemsets and discovering the relevant subspace for a given cluster. They find projected clusters by mining frequent itemsets. Wimalasuriya et al., (2007) applies the technique of clustering based on frequent-itemsets in the domain of bio-informatics, especially to obtain clusters of genes based on Expressed Sequence Tags that make up the genes. Yuan et al., (2007) discovers frequent itemsets from image databases and feeds back discovered patterns to tune the similarity measure in clustering.

One common aspect among various pattern-based clustering methods is to define the similarity and/or the difference of objects/patterns. Then the similarity and difference are used in the clustering algorithms. The similarity and difference can be defined pairwise (between a pair of objects), or globally (e.g. within a cluster or between clusters). In the main focus section, we focus on the ones that are defined globally and discuss how these pattern-based clustering methods can be used for segmenting customers based on their behavioral patterns.

MAIN FOCUS OF THE CHAPTER

Segmenting Customers Based on Behavioral Patterns

The systematic approach to segment customers or customer transactions based on behavioral patterns is one that clusters customer transactions such that behavioral patterns generated from each cluster, while similar to each other within the cluster, are very different from the behavioral patterns generated from other clusters. Different domains may have different representations for what behavioral patterns are and for how to define similarity and difference between sets of behavioral patterns. In the wireless subscribers example described in the introduction, rules are an effective representation for behavioral patterns generated from the wireless call data; however, in a different domain, such as time series data on stock prices, representations for patterns may be based on "shapes" in the time series. It is easy to see that traditional distance-based clustering techniques and mixture models are not well suited to learning clusters for which the fundamental characterization is a set of patterns such as the ones above.

One reason that behavioral pattern-based clustering techniques can generate natural clusters from customer transactions is that such transactions often have natural categories that are not directly observable from the data. For example, Web transactions may be for work, for entertainment, shopping for self, shopping for gifts, transactions made while in a happy mood and so forth. But customers do not indicate the situation they are in before starting a transaction. However, the set of patterns corresponding to transactions in each category will be different. Transactions at work may be quicker and more focused, while transactions for entertainment may be long and across a broader set of sites. Hence, grouping transactions such that the patterns generated

Related Content

Reflecting Reporting Problems and Data Warehousing
Juha Kontio (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 1682-1688).*
www.irma-international.org/chapter/reflecting-reporting-problems-data-warehousing/11044

Enhancing Web Search through Query Expansion
Daniel Crabtree (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 752-757).*
www.irma-international.org/chapter/enhancing-web-search-through-query/10904

Applications of Kernel Methods
Gustavo Camps-Valls, Manel Martínez-Ramónand José Luis Rojo-Álvarez (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 51-57).*
www.irma-international.org/chapter/applications-kernel-methods/10797

On Interacting Features in Subset Selection
Zheng Zhao (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 1079-1084).*
www.irma-international.org/chapter/interacting-features-subset-selection/10955

Scalable Non-Parametric Methods for Large Data Sets
V. Suresh Babu, P. Viswanathand Narasimha M. Murty (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 1708-1713).*
www.irma-international.org/chapter/scalable-non-parametric-methods-large/11048