

# Mining the cause of complaints by using a two-phase association rule technique

Elham Akhondzadeh<sup>a\*</sup> and Amir Albadvi<sup>b</sup>

Department of Industrial Engineering, Tarbiat Modares University, Tehran, Iran

[e.akhondzadeh.n@gmail.com](mailto:e.akhondzadeh.n@gmail.com) [albadvi@modares.ac.ir](mailto:albadvi@modares.ac.ir)

## Abstract

Complaint management is one of the most important dimensions of customer retention which is a crucial requirement for successful businesses. This element helps businesses to manage customers' needs and behavior (Ngai et al., 2009). One of the techniques that could be applied effectively in this domain is data mining. Ngai et al. (2009) introduced these tools, especially association rules, as techniques that could be applied to mine unseen patterns of complaints and discover the root of them. It helps organizations to obtain a deeper understanding of each customer's behaviors and needs and the factors that cause dissatisfaction. This knowledge is helpful in finding the root of complaints and managing customers' complaints. But, despite the advantages of such a study, there were relatively few articles covering complaint management and more research could be conducted on the application of data mining techniques in this domain. Based on the review study done by Ngai et al. (2009), only two papers have studied using data mining tools in complaint management. These two researches were conducted by Bae et al. (2005) and Larivière and Poel (2005) which used Self-organizing map and Survival analysis, respectively.

Bae et al. (2005) worked on the data of the customer complaints of a life insurance company by using statistical and data mining techniques. They could find the problematic areas which complaints happened, the relationship between the problems and the root of them. The results would be helpful for the decision makers to support the customers.

Larivière and Poel (2005) used survival analysis to analyze the customer complaints by a novel approach. The aim of this research was to predict the customer's post-complaint period. The case was a financial service provider. By using the survival analysis technique, they could monitor the effects of complaint management on the future behavior of customers and changing their behavior over time. The results show that handling the complaints effect the actual purchase behavior of customers and would be beneficial.

Regarding the usefulness of association rule in complaint management, this paper is focused on using this technique in this domain. The case study is one of the governmental systems in Tehran, the capital of Iran. This system is constructed in order to connect between the government and citizens. The main purpose is to manage the complaints against urban services and requirements. This system has been established to provide appropriate and timely responses to citizens' demands and complaints. The database of this system includes proper information about citizens' complaints and would be used to get a deep understanding of citizens' needs and to predict citizens' requirements. Based on the data, we can get exact information about the subject of complaint, the time of it and the demographic characteristics of people who had complained.

As mentioned, the analysis of the data could be very helpful in citizen relationship management. One of the important factors in CiRM is to establish a feedback system and analyze the feedbacks.

In this study, we apply the Apriori algorithm on the database of this system to extract the association rules between complaints, time of complaints and the geographical region that a complaint has occurred. In this way, the “Region” and “Time” are considered as the inputs and the “Subject” as the output of this algorithm. The last field shows the type of complaint. The minimum confidence is considered as 70%. About the support, it is better to consider different values for different subjects based on the type of the complaints. This is also could be notable in using association rule mining.

The results help us to get a deeper understanding of different kinds of the complaints in every region and section in our city. Furthermore, the prior citizens' needs and critical problems are identified in every geographical region in Tehran. These results could be helpful in managing urban complaints. Based on the results, we know that at which level of confidence, a special problem may accrues in a special geographical region at different period of time. This may be helpful for predictive purposes.

Furthermore, we try to find those kinds of subjects which occur together during a special period of time by using a two-step method of the Apriori algorithm. Firstly, we present a brief review of the association rule mining and Apriori algorithm. Then, the proposed two-step technique is discussed.

Association rule is one of the data mining techniques located in the domain of descriptive tools. This technique tries to find the items that exist together with a high frequency. The first application of this technique was in the area of customer relationship management, especially in the domain of market basket analysis. This is notable that the application of this technique is not limited to marketing problems and it is widely used for other decision making domains (Tan et al., 2006; Cock et al., 2005 and Ngai et al., 2009)

The general form of association rule is in the form of  $X \rightarrow Y$  where  $X \cap Y = \phi$ . There are two main metrics that are used for the validation and the strength of these rules. These metrics are support and confidence. The support metric shows that how often X and Y are exist together in the transaction database while the confidence metric determines how often items in Y appear in transactions that contain X (Tan et al., 2006).

Apriori algorithm is one of the common tools for mining association rules. This technique is a level-wise one and generates frequent item sets from size one to the longest frequent item sets. At each stage, the larger frequent item sets are generated based on the frequent previous ones. This characteristic helps to delete some of the options in scanning the whole data at each level. So, the scanning process is done more rapidly.

As mentioned before, the two-step technique is presented based on Apriori algorithm. The mining process is discussed step by step as follows:

At step 1, the Apriori algorithm is used to extract the association rules between complaints, time of complaints and the geographical region that a complaint occurred. The “Region” and “Time” are considered as the inputs and the “Subject” as the output.

At step 2, we try to find the subjects that occur together in a special period of time. Then, we convert the results of step 2 in to a new format of data that fields show the kinds of complaints and records show the preferred period of time.

At step 3, the Apriori algorithm is applied on the new format to mine the subjects that occur together in a special period of time. As a fact, we use the output of the Apriori algorithm as the input of this algorithm in the second stage.

The two-phase technique has a good performance in decreasing the number of fields and records and can be notable as a good solution in this problem. It also considers the important citizens' needs in the mining process and eliminates those kinds of subjects that are not important according to the citizens' opinion. The results of this method help us to handle the complaints by realizing the subjects which occur together in a period of time. As a fact, analysing the results make it possible to understand why such a complaint occurs. The

knowledge about the root of complaints helps the urban managers to handle the complaint before it occurs and this would increase citizens' satisfaction. This paper could be notable as the first studies on using association rule mining in complaint management. Another point that could be notable in this domain, is to develop the association rule mining algorithms to be able to find the cause and the root of subjects. The algorithms which have presented up to now are not capable of discovering the root cause.

**Keywords:** complaint management, Data mining, Association rule, Apriori algorithm.

### **References**

- Bae, S. M., Ha, S. H., & Park, S. C. (2005). A web-based system for analyzing the voices of call center customers in the service industry. *Expert Systems with Applications*, 28, 29–41.
- Cock, M. D., Cornelis, C. and Kerre, E. E. (2005), "Elicitation of fuzzy association rules from positive and negative examples. *Journal of Fuzzy Sets and Systems*", vol.149, pp.73–85.
- Ngai, .E.W.T., Xiu, L. and Chau, D.C.K. (2009). "Application of data mining techniques in customer relationship management: A literature review and classification", *journal of Expert Systems with Applications*, vol. 36, pp. 2592–2602.
- Larivière, B., & Poel, D. V. D. (2005). Investigating the post-complaint period by means of survival analysis. *Expert Systems with Applications*, 29, 667–677.
- Tan, P., Steinbach, M. And Kumar, V. (2006). *Introduction to Data Mining*, Addison Wesley, ISBN: 0-321-32136-7, pp. 171-180.