Journal of International Scientific Publications:
Materials, Methods & Technologies, Volume 7, Part 2
ISSN 1313-2539, Published at: http://www.scientific-publications.net

THE STUDY AND CONSTRUCTION OF A SYSTEM FOR THE CREATION AND IMPLEMENTATION OF PROMOTIONAL OFFERS

Ana KTONA[1] Etleva Dashi[2] Jerina HOXHA[3] Afërdita SPAHIU[4]

[1]Informatics Department, Faculty of Natural Sciences, University of Tirana

[2]Department of Economy and Agrarian Policies, Faculty of Economy and Agribusiness, Agricultural University of Tirana

[3]Informatics Department, Faculty of Natural Sciences, University of Tirana

[4]Head of Core Support, AMC, Albania

Abstract

The development of information and communication technology creates possibilities for companies to maintain large amount of actual and historical data. The data keep information that can help companies answer to different questions about business performance and help them in decision making. Within these data there is information that will help companies to get new and keep existing customers.

Traditional methods of data analysis could not help in extracting the information that was unknown before from these data. Recently for analyzing the data with the aim to detect previously unknown information is being used data mining which is a combination of artificial intelligence, machine learning, statistics, and information processing systems. Data mining is known as a process that finds patterns and meaning of data in large datasets with the aim to discover crucial information for further use and to transform this information in an intelligible form.

The use of datamining techniques helps companies to determine the relationships between various factors such as price, the position of the product or staff skills. As a result, companies can create different offers and products for specific customers based on historical data of sales performed.

In our paper, we study the use of Association Rules technique in marketing and we apply it practically to the data gathered by a transaction processing system with the aim to increase existing customers' satisfaction and to attract new customers. The generated rules are implemented in a system that creates different offers and products for specific customers. We expect that the products and offers created by the system to have a positive impact in the increase of customers' satisfaction and attracting new clients.

Key words: *Data mining, association rules, market basket analysis*

ABOUT ASSOCIATION RULE MEANING

The interaction of a company with its costumers/clients generates large volumes of data. These data are first captured by the transaction processing system. Later the data can be collected, filtered and summarized to engage all in one Data Warehouse of customers. For customers a well-designed Data Warehouse keeps historical records of customers' interaction and serves as the memory of the company. Data Mining tools can be applied on these historical data to learn about the costumers. This acquired knowledge must be translated into action for bringing new customers and for retaining the

existing ones. This knowledge can be translated in offers and promotional products in a company. Offers and promotional products are important in fighting the competition, bringing new clients and retaining the existing ones. One of the ways to gain knowledge about the customers is market basket analysis, because in retailing most purchases are bought on impulse.[3]

Market basket analysis it's known also as association rule mining or affinity analysis. It is a data-mining technique that originated in the field of marketing to identify relationships between groups of products, items, or categories. [1], [2] Association rules represent models at the data without a specified goal.[4] Association rules are generated initially by Point Of Sales data and describe which products are bought together. POS data provides the purest demand stream from which to develop future plans. Collection and use of POS data is a critical part of the application of market basket analysis. Items have product codes that relate to the categories and so becomes possible their classification in hierarchical way, which helps in generalization of products. The complexity of a rule refers to the number of items it contains.[4] The more items included in transactions, the more is needed to generate the rule of a given complexity. The desired complexity of the rule determines also how specific or general the items must be. If high levels of the hierarchy are chosen, then this reduces the number of items. Hundreds of thousands of items can be reduced at one single general item. Even though their root is in the POS transactions, association rules can also be applied out of the retailed industry to find relationships between other types of baskets.

A market basket in our context shows what a customer has bought at a certain time. Order is an important element of market basket data. An order represents a single customer purchase. A full list of purchases made by all the customers ensures more information. It describes the most important of retail business: What goods the consumers are buying and when. Each customer buys a set of different products, in different amounts and at different times. Market basket analysis uses this information to ensure knowledge about who they are and why do they make certain purchases. Market basket analysis penetrates to the goods showing which products have the tendency to be bought together and which are more suitable for promotions. It can suggest new shop formats, determine which products are special and show when coupons can be used.[5] When these data connect to individual customers through a loyalty card, they become even more valuable

To understand the purchases there are some basic measurers:

Which is the average number of orders per customer?

Which is the average number of different items per order?

Which is the average number of items per order?

For a given product, which is the proportion of customers that have ever bought the product?

For a given product which is the average amount purchased in an order that the product is bought?

These measurers give a wide image of the business. The number of orders is usually a useful way to differentiate between customers. Good customers order more often than those not so good customers.

DATA AND METHODOLOGY

In the view of the main purpose of the paper was studied theoretical and empirical literature to analyze the role of Data Mining techniques and specifically that of association rules in bringing new clients and retaining existing ones in a company. The data were gathered by Transaction Processing System LogiPOS and preprocessed. LogiPOS system was implemented in Albania to Jumbo retails company.

From Jumbo database we will try to show LogiMark System effects. Features of data which did not affect the result were erased from the data.

Data were converted in the format csv. Association rules were applied on our data using weka environment.[6] Weka environment has 6 methods for finding the association rules. We applied FP Growth algorithm in our data. The FP-tree data structure and FP-growth algorithm for finding large item sets without candidate generation were introduced by Han et al. (2000) [7] following pioneering work by Zaki et al. (1997);[8] Han et al. (2004)[9] give a more omprehensive description.[6]

The obtained rules were implemented in LogiMARK system for creating offers and promotional products. Adding the association rules technique in LogiMARK system would make the system a decision making system enabling the creation of offers and promotional products in a dynamical way.

LogiPOS as well as LogiMARK are Windows applications (Widows Form Application) and are build using Microsoft Visual Studio. The language in which they are written is C#.NET.

Like the architecture of LogiPOS even the architecture of LogiMARK is a 2-layer architecture (2-tier), Client/Server. The logic of the application is set to Client and the data are set to the database layer (which can be on the same computer with the application or in a different computer).

The client layer consists of three layers:

Presentation/Interface - Presentation Layer

Business Logic - Business Layer

Data - Data Layer

Presentation Layer typically contains UI components and the logic of this layer. Business layer typically contains business logic, business workflow and components of business entities. Data layer contains data access components and service agent components.

ASSOCIATION RULES

Association rules appear in the form of rules for groups of products and as such are clear and useful. Association rules are easy to understand but are not always necessary. Three more used types of rules produced by association rules are: operational rules; irrelevant rules and unexplainable rules.

Operational rules are useful rules that contain operational information with high quality. Since the rule is easily understood, it suggests plausible causes and possible interventions.

Irrelevant rules are the cases when the result from market basket analysis can simply measure the success of other previous marketing campaigns.

Unexplainable rules seem to have an explanation that does not suggest a move. When applied in the market basket analysis, many of the results are often trivial or unexplainable. Trivial rules reproduce common knowledge for the business, wasting the opportunity to use more sophisticated techniques of analysis.

Association rules start with transactions that keep one or more products or services and some basic information for transactions. Each of these transactions provides information on which products are bought with other products, confidence. Another question is how better than chance is rule? To answer this is lift. Lift or improvement shows how much better is a rule in predicting the result than the assumption of the result. A similar measure is excess[3]. Since excess is measured with the same unit as the original sales, sometime is easier to work with it.

Three are the main issues on the establishment of the association rules.

Choosing the right set of items

Generating rules by deciphering numbers in the proliferation matrix C0.

Overcoming practical limits imposed by the hundreds of thousands of articles

Market basket analysis produces the best results when the items are located nearly with the same number as the number of transactions at the data. This helps the rules not to be dominated by the most frequent item. Products hierarchy can help at this point, items can be moved to the upper levels in the hierarchy, in order to become more frequent, and there is no need to move to the upper levels the items that appear more frequently.

Weka environment has 6 methods for finding the association rules.[6] Apriori method implements the apriori algorithm. The data for market basket analyses to b processed by this algorithm need to be coded in a specific way: each attribute has to represent an item in the basket. The attribute has to be declared as nominal with a single value (for example true when it is in the basket). Absence of an item in the basket is indicated by specifying a missing value (? when the input data are in the Attribute-Relation File Format). FPGrowth implements tree algorithm with frequent patterns. It is designed for the data of the market basket. All the attributes are expected to be nominal binary attributes, and the user can specify which of the two values will be treated as positive, which shows the presence or not in the basket (the default is to use the second value). The other four algorithm are Filtered Associator, Generalized Sequential Patterns, Predictive Apriori and Tertius.

THE ANALYSIS OF THE LOGIPOS AND LOGIMARK SYSTEMS.

LogiPOS is an application used to perform daily transactions at retail points. The application enables registration in real time of sales data, manages security aspects separating users in different profiles and giving the respective rights under these profiles. It also manages very well the reporting of data through a myriad of detailed reports on sales and staff performance. The data stored in the database can be transmitted to remote and integrated into other systems via different interfaces which enable the collection of information on a larger database for further processing. The system through advanced algorithms is able to manage price levels, discounts and rewards according to client credentials who is performing a purchase at that moment. It gets this kind of intelligence through advanced marketing module LogiMARK.

LogiMARK is an all inclusive system that performs a connection between several systems and stands between them as a connecting bridge. LogiMark is an application that helps marketing agents to create promotional offers. By using certain algorithms manages to create very interesting packages for clients through a cause-effects mechanism. It is used to create different marketing ‘recipes’ in order for them to be interpreted by LogiPOS and for those offers to be applied in real time. Recipes are different types of marketing promotional offers such as discount products, discount bills, creation of coupons with different values, that marketing agents create to bring the business a step forward. But to create these promotional offers and to make decisions for them, the marketing agents need to know. Knowledge which is related with the objects that they are going to consider as possible attributes of the promotional offer. As the market basket, the three main types of data that LogiMARK focuses on as input or cause channel are:

Customers/Clients

Items

Orders/Bills

For each of these input channels several key attributes are included such as: for customers are included attributes as birthday, gender, age, profession, if he is a new client or an existing one, status, city, country, type of card he obtains etc, as it regards to items some classification attributes are included such as category, group, barcode and two extra classifiers from the item category. Whereas for the bill is the attribute of the value of the bill. As for the output channel or effect, which is the object on which these promotions will be applied, except for the 3 main types of data that were included even at the input channel, including a message on the bill and printing of coupons with certain value.

The whole system takes wide information from LogiPOS system and stores it in database to analyze it in the opportunistic moment. The system collects information on sales, clients, offerts that are active in the system as well as the behaviors of the clients or their reaction to the offers. All this information, LogiMARK is capable to turn into knowledge, by highlighting customer behavior schemes or to predict how the client can react in certain marketing campaign based on historical data.

THE IMPLEMENTATION OF THE INFORMATION SYSTEMS LOGIPOS AND LOGIMARK

LogiPOS stores in a relational database the information for customers, suppliers, warehouses, articles, sales agents, system operators etc. All these indicators as we above mentioned were applied in Jumbo retail store in Albania.

Each of the above categories uses tables from database. With regard to the customer the system stores a set of data in database. Among them may be mentioned the designation, address, contact numbers, bank accounts etc. System in a relational way also stores for this customer the documents generated for him/her, such as orders, bills, warranty information on products, payment terms etc.

The supplier is also very similar in approach from the database. The system stores generalities as in the case of the Customer. In the same way are stored the orders made from the company to the supplier, the supplies done, cash balance etc.

For the warehouses the system stores the appropriate records where the warehouses are opened. These records contain less information than the customer and supplier card and serve to correctly identify the warehouses through transactions. Warehouse record contains ID, naming, address as well as information about the person in charge and some basic information about physical features such as surface, height, etc…

The articles have the largest card in the system. The system contains a lot of information about them such as ID, description, barcode, classificatory, sale and purchase prices, information on warranty and deadline, type of article, etc. The article like the warehouse, supplier or the customer is part of every document generated by the system.

System operators are those who perform transactions in the system. The system connects them automatically with every transaction performed. Besides the basic card, a relevant profile is completed for system level access. The system generates for them a number of reports which aim to monitor not only the performance, but also various aspects related to security.

Sales agents are the people who perform selling on the ground. The system stores for them the sales levels in order to reward them based on performance.

LogiMARK system takes information from ERP about the items, warehouses, clients etc. after converting them into the desired format is able to send this information to the sales system of

LogiPOS. In addition to the information on items LogiMARK sends in LogiPOS information on different offers created in the system based on a genuine logic of marketing.

To create promotional offers the marketing agents needed information to focus on and what they must not do. Question and uncertainties of agents found answers in the results of the application of the Association Rules technique on the data generated by LogiPOS. The results of Association Rules technique are obtained using Weka[6]. If these data are also associated with a client identification card, they increase their value for the information they will provide. These data passed in the preparation phase where was made the filtering from useless information, they were reorganized in order that the generated rules from Association Rule to have a high confidence. Through the association rules the company used the information obtained from them and identified the customers based on the products they buy, and understand why they buy certain products. At the same time the company got information on products such as, the most sold products, the products that move quicker, the products that are sold together as well as the products that can profit from promotions. Association Rules represent models at the data without any goal in mind, and then it is up to people, the marketing agent namely to determine if these models have meaning or not on data. All this information obtained through Association Rules application help the agent answer the questions he had. As well as making decisions on product placement on shelves so geographically positioning of items offering the customer/client products that are sold together in the same time saves time for the client and leads to sales growth, which products to put on coupons to promote sales of products, apply discounts, so in a few words to create more profitable promotional offers for the business. Through the outcomes obtained by Association Rules helps one determining the time to make cross-marketing. Following information received from the Association Rules results through LogiMARK marketing agents create different promotional offers discount invoice, to push products that have dropped the number of sales. Actually the system brought a lot of benefits for the company. They have sales improvement and reorganization of the department store. Also through LogiMARK they did the evaluation of the impact that these "suggestions" have. Through the reporting module that provides LogiMARK they compare promotional sales and non promotional ones and make further decisions about these promotions.

CONCLUSIONS

Promotional offers and products are important for fighting competition, bringing new clients and retaining existing ones. The acquired knowledge on customers can be translated in offers and promotional products in a company. One of the ways to gain knowledge is market basket analysis. The closest data mining technique with market basket analysis is automatic generation of association rules. Association rules using Weka environment were applied on data gathered by a Transaction Processing System LogiPOS which was implemented in Jumbo retail store in Albania. Weka environment has 6 algorithms for finding association rules. We applied FPGrowth algorithm which is designed for market basket analyses. The obtained rules were implemented in LogiMARK system in order to create offers and promotional products. Following information received from the Association Rules results through LogiMARK marketing agents create different promotional offers discount invoice, to push products that have dropped the number of sales. The company have sales improvement and reorganization of the department store. Through the reporting module that provides LogiMARK they compare promotional sales and non promotional ones and make further decisions about these promotions.

In the future we will implement association rules and other data mining techniques in the LogiMARK system in order to make it a decision making system. Adding association rules technique in the LogiMARK system would make the system a decision making system enabling the creation on offers and promotional products in dynamic way.

REFERENCES

[1] Herman Aguinis, Lura E. Forcum and Harry Joo, "Using Market Basket Analysis in Management Research", Journal of Management Vol. 39 No. 7, November 2013 1799 - 1824

[2] Larry Gordon, Partner of Fact point Group, "Leading Practices in Market Basket Analyses market Basket Analyses", 2008 , Retrieved from: http://factpoint.com/pdf2/1.pdf

[3] **Paco Underhill (1999)** Why We Buy: The Science of Shopping, **ISBN-10:** 0684849135

[4] Pang-Ning Tan, Michael Steinbach and Vipin Kumar, "Association Analyses Basic Concepts and Algorithms", Retrieved November 2013 from: http://www-users.cs.umn.edu/~kumar/dmbook/ch6.pdf

[5] Troy Raeder and Nitesh V. Chawla, "Market Basket Analysis with Networks" Retrieved November 2013 from: www3.nd.edu_~dial_papers_ASONAMJ10

[6] Ian H. Witten, Eibe Frank and Mark A. Hall, "Data Mining Practical Machine Learning Tools and Techniques Third Edition", Morgan Kaufmann, January 2011

[7] Han, J., Pei, J., & Yin, Y. (2000). Mining frequent patterns without candidate generation. In Proceedings of the ACM-SIGMOD International Conference on Management of Data (pp.1–12). Dallas. New York: ACM Press.

[8] Zaki, M. J., Parthasarathy, S., Ogihara, M., & Li, W. (1997). New algorithms for fast discovery of association rules. In Proceedings Knowledge Discovery in Databases (pp. 283–286). Newport Beach, CA. Menlo Park, CA: AAAI Press.

[9] Pei, J., Han, J., Mortazavi-Asi, B., Wang, J., Pinto, H., Chen, Q., et al. (2004). Mining sequential patterns by pattern-growth: The PrefixSpan approach. IEEE Transactions on Knowledge and Data Engineering, 16(11), 1424–1440.