International Journal of Computers and Informatics (IJCI)
Vol. (3), No. (1)

المجلة الدولية للحاسبات والمعلوماتية
الإصدار (3)، العدد (1)

January 2024

Association Rules Elicitation for Customer's Shopping on E-commerce Applications

Bassant Ashri Nouri
Ph. D. of Information Systems, Faculty of Computers and Information, Mansoura University, Egypt
sabryiaaeq@alexu.edu.eg

Mohamed Saad Gaafar
Lecturer of Marketing, Faculty of Economics, University of Benghazi, Libya
msaadsaad@uob.edu.ly

Abstract:

This research provides an approach for eliciting association rules for customers' shopping on e-commerce applications based on the Apriori algorithm. The proposed system shows and displays some offers and deals from various branches. The system provides the analytics for the seller as the demand for some varieties and the weakness in other varieties and the whole application is organized on the Cloud. The architecture includes three levels; the front-end, middle, and back-end. The front-end level of the site-based Mobile shopping application is made up of Android Mobile devices, to buy miscellaneous products from various nearby branches. The front-end level also displays the link between items purchased. The middle repository level provides a Web service to generate returns from a relational database.

The Exchanged information and data between application and servers is stored in the Cloud. The background level offers a Web server and a MySQL database. In this paper, we propose an architecture that reduces the communication overhead in existing Mobile Agent-based Distributed Association Rule Mining (MAD-ARM).

International Journal of Computers and Informatics (IJCI)
Vol. (3), No. (1)
IJCI
المجلة الدولية للحاسبات والمعلوماتية
الإصدار (3)، العدد (1)
January 2024

Keywords: E-commerce, Market Basket Analysis, Association Rules, Customer's Shopping.

1- Introduction

Software engineering aims to develop and adapt software systems to satisfy user needs schedules and budget constraints. In pursuit of this goal, a substantial amount of research has been conducted to improve the software development process. However, most software projects still continue to fail; this is reported in the Standish report, which states that only 28% of real-world projects are successful [16].

I know that mobile devices are constrained by processing power, battery life, and storage. However, cloud computing provides an illusion of infinite computing resources. Mobile cloud computing is a new platform combining mobile devices and cloud computing to create a new infrastructure, whereas cloud performs the heavy lifting of computing-intensive tasks and storing massive amounts of data. In this new architecture, data processing and data storage happen outside of mobile devices.

Cloud computing combined with data mining can provide powerful capacities of storage and computing and excellent resource management. Data mining in the cloud computing environment can be considered as the future of data mining because of the advantages of the cloud computing paradigm. Mobile cloud computing (MCC) is a new emerging research field.

Today's mobile devices have many advanced features such as mobility, communication, and sensing capabilities, and can serve as the personal information gateway for mobile users. However, when running complex data mining and storing operations, the computing, energy, and storage limitations of mobile devices demand an integrated solution relying on cloud-based computing and storage supported. The Mobile shopping application consists of Mobile devices with limited memory and

processing speed. The Cloud, on the other hand, provides a large storage and speed for data stores.

This paper is specially designed for product purchasing in the city for malls and a local market near the customer location. It uses Wi-Fi Global Positioning System (GPS) and Mobile network to get the current position of the customer and displays any registered supermarket's branch on the server from the customer's location. Association rule mining as a technique of data mining is used to find the offers associated with products. Technically, data mining is the process of extraction of interesting information or patterns from data in a large database. Association rule mining is widely used in market basket analysis. This method benefits retailers in numerous ways for marketing or planning shelf space.

2- Research Background

In this section, we review Mobile Cloud Computing (MCC), Market Basket Analysis (MBA), Association Rule Mining, and Location-based services.

2-1 Mobile Cloud Computing (MCC)

Mobile cloud computing consists of three modules: Mobile devices, communication network, and cloud as a server. The application that is large can be decomposed to smaller ones to process concurrently. This method is called application partition. Offloading is the process of transferring mobile applications to the cloud. This saves the device memory, processing power, and ultimately battery consumption. The classic facilities needed by a Mobile cloud client are, synchronization, push, i.e., updates the notifications sent by the cloud server, offline App automatically handles synchronization and notification, network, database, inter App Bus; and provides low-level coordination between applications.[15].

2-2 Market Basket Analysis (MBA)

Market Basket Analysis is a forming technique based on the theory that if you buy a certain group of items, you are more likely to buy another group of items. MBA uses this information to: (1) Understand why they make certain purchases, (2) Products that are purchased together, and (3) Products that might benefit from the promotion. This system has used the Market Basket Analysis method for analyzing the data. The following techniques are used in the analyzing process.[14]

2-3 Association Rule Mining

MAD-ARM is a mobile agent-based distributed data mining architecture. It contains a knowledge server, which works on the generation of association rules and data coming from different remote sites. The item sets are always upgrading on remote sites at the immobile agent. [3]

The association rules have been recently recognized as an important tool for knowledge discovery in databases. The problem of discovering association rules was first investigated in pioneering work in [4]. Here we examine a database of records which consist of both customer profile (such as salary and age) and behavior (such as buying decision) information).

The association rule problem was originally proposed for the case of binary itemset data.[9] The intuitive implication of the association rule X ~ Y is that the presence of the set of items X in a transaction also indicates a strong possibility of the presence of the itemset Y. The measures used to evaluate the strength of the rule are supported and confident. The supported of a rule X ~ Y is the fraction of transactions that contain both X and Y. The confidence of a rule X ~ Y is the fraction of transactions containing X which also contain Y. A considerable amount of research effort [5, 6], has been devoted to the problem of speeding up the itemset method for finding

association rules from very large sets of transaction data. Several variations of this problem have been discussed in [7, 8]. The quantitative association rule problem is a generalization of the problem of discovering association rules on sales transaction data, in which both categorical and quantitative attributes are allowed [10].

2-4 Location-based services

The Mobile location-based application for any shopping product was designed and developed to find nearby branches and stores in local markets, the association between the products purchased, display association to customer side screen, post order, and it is deployed on the Cloud (See Figure 1).

The grouping of web map service API and association rule mining using Mobile in the cloud has the potential to gather large-scale shopping habits of people, with lower data collection costs. The developed application model represents an environment for data analysis, and the proposed Algorithm is chosen dynamically for each handset. This is based on the environment where the data stream mining process runs on the user's Android handset. As the data streams continuously, possible concept drift is updated. There is a specific central Mobile decision agent that switches several other stream mining agents. Stream mining agents working on local Mobile phones decide the best possible algorithm to run on the local data.

3- Related Work

- In [11] this paper designed and established a location-based mobile shopping application for malls and local markets for Android platforms. This application shows nearby local markets and mall stores that are registered to the application. The main objective of marketing is achieved at a very low cost in comparison to advertisements, announcements, ground-level marketing, etc.
- In [12] this study has represented a recommended engine by using association rules. The system had been tested by existing data in terms of accuracy and

coverage. The best results have been determined for 25 days of data with 87.74% coverage and 16.43% accuracy. Processing time was 318 minutes for those data.

- In [13] this paper, we discussed the problem of online mining of profile association rules. Such rules may be very useful in developing relationships between consumer profiles and behavioral information. We discussed how to use multidimensional indexing to generate profile association rules in an online fashion.

4- The Proposed Architecture

Any product or application that uses the location data of mobile subscribers is called a location-based service. A location-based service like GPS uses latitude and longitude data. A location-based text mining approach categorizes texts into numerous categories based on the geospatial features, with the goal of discovering relationships between documents and areas.

There are three main modules in this architecture, including geographic data group and reprocessing, mapping forms into corresponding regions, and framing maximizing zones. Data mining and processing take place based on zones. The tourism industry has also taken the benefit from location-based services.

This application is designed and established using the cloud-based platform. It discovers the location of tourists, where they are directed or looking.

International Journal of Computers and Informatics (IJCI)
Vol. (3), No. (1)
IJCI
January 2024
المجلة الدولية للحاسبات والمعلوماتية
الإصدار (3)، العدد (1)

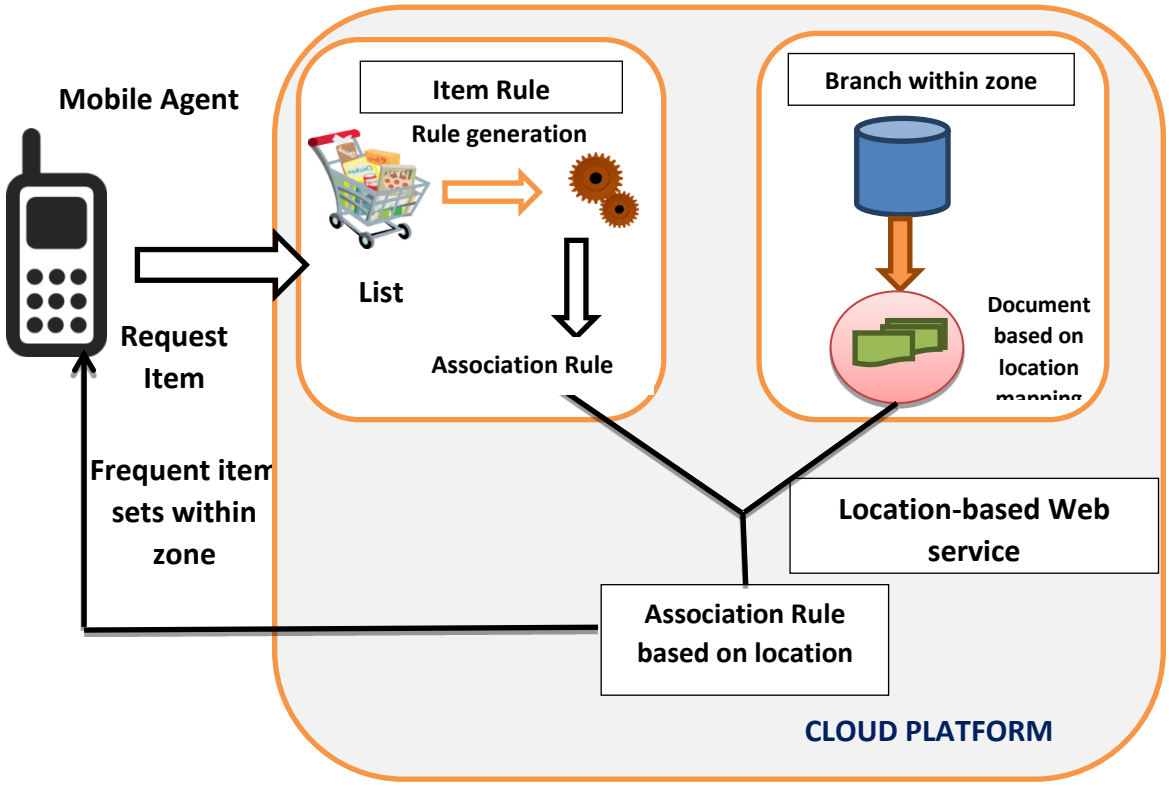

Figure 1 - Proposed Architecture Mobil Agent Association Mining System for Shopping

4-1 System Architecture

The architecture of the shopping application is divided into three parts Front end, Middleware level, and backend.

The Android open-source platform is used to design and develop shopping applications. For the end user, in front-end user can be able to select a particular supermarket branch in the city and see the offers available in the specific showroom at the mall. It also provides offers of local markets which are available to users in

nearby locations. The registered retailers can upload and remove their offers and advertisements from this application. When the user searches for specific offers of the showroom in a mall, then the request is sent to the middleware level that is to the Web Service. Web service acts like an interface for front-end and back-end. The data exchange between the front-end and back-end of the shopping application happens via the middleware level. The Android shopping application sends an HTTP Request, and Web Service will send a Query to fetch requested data from the MySQL database located on the cloud.

4-2 Software and Data Set

In this paper we used a data set from the supermarket that contains 4700 Records, every record has 23 Items for purchasing case registration. By using and preparing the rules for the Apriori diagram and testing this data by using the Weka Software tool. Weka is efficient and has a user-friendly user interface. It is fully implemented in Java language there and it runs on almost any computing platform. However, it can only run in the local environment.

4-3 Apriori Algorithm

- **Product set:** a set of items
- **k-product_set:** a product set which consists of k items.

  Frequent product_set (i.e., large product_set): a product_set with sufficient supported
- **Lk or Fk:** a set of large (frequent) k-product_sets
- **ck:** a set of candidate k-product_sets
- **Apriori property:** if item A is joined with item B, supported (A U B) = min (Supported(A), Supported(B))

In the process of Apriori, the following explanations are needed:

Definition 1: Suppose T= {T1, T2, … , Tm}, (m_1) is a set of transactions, Ti= {I1, I2, … , In},(n_1)

is the set of items, and k-product_set = {i1, i2, … , ik},(k_1) is also the set of k items, and k-product_set ⊆I.

Definition 2: Suppose_ (product_set), is the supported count of product_set or the frequency of occurrence of a product_set in transactions.

Definition 3: Suppose Ck is the candidate product_set of size k, and Lk is the frequent product_set of size k.

In the proposed approach, scan all transactions to create L1 which contains the products, the supported count, and the Transaction ID where the products are found. And then use L1 later as a helper to generate L2, L3 ... Lk. When creating C2, make a self-join L1 * L1 to construct two products set C (x, y), where x and y are the products of C2. Before scanning each transaction record to count the supported count of every candidate, use L1 to get the transaction IDs of the least supported count between x and y, and thus scan for C2 only in these specific transactions. The same thing for C3, construct three products set C (x, y, z), where x, y, and z are the products of C3, and use L1 to get the transaction IDs of the least supported count between x, y, and z, then scan for C3 only in these specific transactions and repeat these steps until no new frequent product_sets are identified. The whole process is shown in (Figure 2).

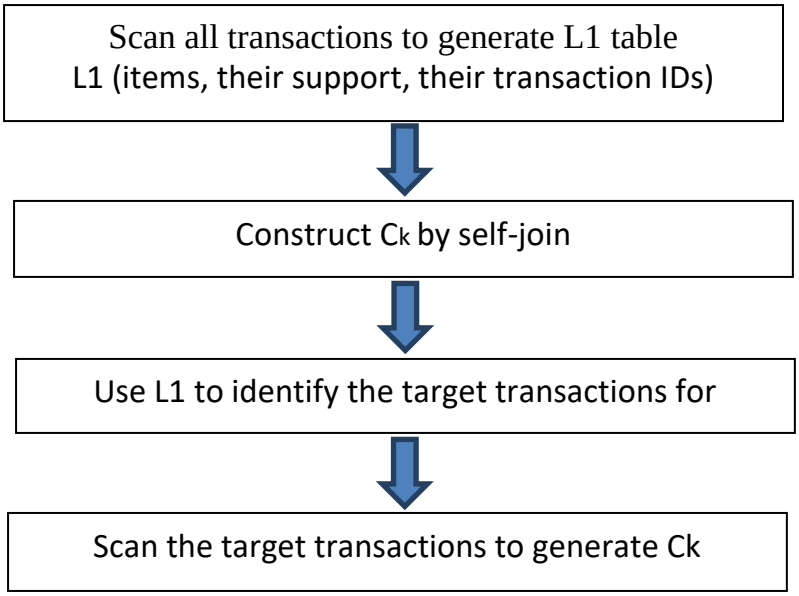

Figure 2: Steps for Ck generation

5- Conclusion

Based on our analysis of existing research on the application of recommended technology in Requirements Engineering, we now focus on a discussion of relevant issues for future research. Decision Supported & Preference Construction. Existing Requirements Engineering approaches often rely on the assumption of stable stakeholder preferences (e.g., in the context of requirements negotiation). The assumption of stable preferences is not applicable for Requirements Engineering scenario; in fact, related decision making follows an incremental preference construction process [15, 20]. In order to better integrate recommended technology into Requirements Engineering processes, we are in the need of deep knowledge about human decision strategies. Such a knowledge will help us to improve the decision supported quality. The integration of human decision strategies into

recommended systems research is a new and challenging field of research which requires a strongly interdisciplinary research approach [20].

We designed this architecture after survey of the research on the use of mobile in the purchase process using an application on the cloud computing systems and tried to improve this service, also to monitor the influence of the factor of space and distance between the consumer and supermarkets and calculated the distances between the place of contact of the consumer and the nearest branches of supermarkets and the lowest prices in the same Time and therefore there are more opportunities for the consumer. We designed a mobile shopping architecture based on the site for the very large supermarket has many Branches of the Android platform. The main objective of marketing is achieved at a very low cost compared to advertising, and marketing at the regional level and so on. Data is exchanged between different levels of structural design using a web service station that generates a JSON format for data transfer. With Mobile Computing Cloud Computing Mobile processing and storage move to the cloud as a server, helping to save battery consumption and extend performance or speed of implementation.

References

[1] Emil Almrot, Sebastian Andersson. A study of the advantages & disadvantages of mobile cloud computing versus native environment. Bachelor Thesis in Software Engineering May 2013.

[2] Raubal, M. and Rinner, C., 2004. Multi-Criteria Decision Analysis for Location Based Services. In S.A. Brandt (ed.) Proceedings of the 12th International Conference on Geoinformatics, pp. 47-53.

[3] A. Saleem Raja, E. George Dharma Prakash Raj, “MAD-ARM: Mobile Agent based Distributed Association Rule Mining”, International Conference on Computer

Communication and Informatics (ICCCI -2013), Jan. 09 – 11, 2013, Coimbatore, INDIA, DOI 10.1109/ICCCI.2013.6466112.

[4] Agrawal R., Imielinski T., and Swami A. 1993. Mining association rules between sets of items in very large databases. Proceedings of the A CM SIGMOD Conference on Management of Data, pages 207-216, Washington D. C.

[5] Agrawal R., and Srikant R. 1994. Fast Algorithms for Mining Association Rules in Large Databases. Proceedings of the 20th International Conference on Very Large Data Bases, pages 478-499.

[6] Savasere A., Omiecinski E., and Navathe S. 1995. An Efficient Algorithm for Mining Association Rules in Large Data Bases. Proceedings of the 21st International Conference on Very Large Data Bases. Zurich, Switzerland, pages 432-444.

[7] Hun J. and Fu Y. 1995. Discovery of Multiple-Level Association Rules from Large Databases. Proceedings of the 21st International Conference on Very Large Data Bases. Zurich, Switzerland, pages 420-431.

[8] Agrawal, R. and Srikant, R. Mining Generalized Association Rules. Proceedings of the 21st VLDB Conference. 1995

[9] Agrawal R., Imielinski T., and Swami A. 1993. Mining association rules between sets of items in very large databases. Proceedings of the A CM SIGMOD Conference on Management of Data, pages 207-216, Washington D. C.

[10] Srikant R., and Agrawal R. Mining quantitative association rules in large relational tables. Proceedings of the 1996 ACM SIGMOD Conference on Management of Data. Montreal, Canada.

[11] Shivani Rajput*, Rajeshree Pagar, Laxman Erande, Yogesh Alai. One Click Android Application for Shopping Based on Cloud. Proceedings of the 2015 IJSRSET | Volume 1 | Issue 5 | Print ISSN: 2395-1990 | Online ISSN: 2394-4099.

[12] Ozgur Cakir, Murat Efe Aras. A recommended engine by using association rules Procedia - Social and Behavioral Sciences 62 (2012) 452 – 456.

[13] Charu C. Aggarwal, Zheng Sun, Philip S. Yu. Online generation of profiles association rules - T.J. Watson Research Center Yorktown Heights, NY 10598 charu@watson.ibm.com, Duke University, Durham, NC-27706 sunz@cs.duke.edu, T.J. Watson Research Center Yorktown Heights, NY 10598 psyu@watson.ibm.com.

[14] Manpreet Kaur, Shivani Kang. Market Basket Analysis: Identify the changing trends of market data using association rule mining, Procedia Computer Science 85 (2016) 78 – 85.

[15] S M Shamim, Angona Sarker, Ali Newaz Bahar, Md. Atiqur Rahman. A Review on Mobile Cloud Computing. International Journal of Computer Applications (0975 – 8887), Volume 113 – No. 16, March 2015.

[16] Standish Group, Extreme CHAOS, Found at www.standishgroup.com, 2001.

[17] J. Golbeck. Computing with social trust. Springer, 2009.

[18] S. Lim and A. Finkelstein. Stakerare: Using social networks and collaborative filtering for large-scale requirements elicitation, in press. IEEE Trans. on Softw. Eng., pages 1–32, 2011.

[19] S. Lim, D. Quercia, and A. Finkelstein. Stakenet: Using social networks to analyze the stakeholders of large-scale software projects. pages 295–304, Cape Town, South Africa, 2010. ACM/IEEE.

[20] B. Mobasher and J. Cleland-Huang. Recommender systems in requirements engineering. AI Magazine, 32(3):81–89, 2011.

[21] M. Pazzani and D. Billsus. Learning and revising user profiles: The identification of interesting web sites. Machine Learning, 27:313–331, 1997.

[22] S. Marczak, I. Kwan, and D. Damian. Social networks in the study of collaboration in global software teams. Munich, Germany, 2007.

[23] J. Konstan, B. Miller, D. Maltz, J. Herlocker, L. Gordon, and J. Riedl. Grouplens: applying collaborative filtering to usenet news full text. Comm. of the ACM, 40(3):77–87, 1997.

[24] J. IyerandD. Richards. Evaluation framework for tools that manage requirements inconsistency. 2004.

[25] E. Tsang. Foundations of Constraint Satisfaction. Academic Press, London, 1993.

[26] R. Reiter. A theory of diagnosis from first principles. AI Journal, 23(1):57– 95, 1987.

[27] G. Ruhe, A. Eberlein, and D. Pfahl. Trade-off analysis for requirements selection. Journal of Software Engineering and Knowledge Engineering (IJSEKE), 13(4):354–366, 2003.

[28] A. Jameson, S. Baldes, and T. Kleinbauer. Two methods for enhancing mutual awareness in a group recommender system. In ACM Intl. Working Conference on Advanced Visual Interfaces, pages 48–54, Gallipoli, Italy, 2004.

[29] G. Ruhe and M. Saliu. The art and science of software release planning. IEEE Software, 22 (6):47–53, 2005.

[30] A. Schrijver. Theory of Linear and Integer Programming. John Wiley and Sons, 1998.

[31] D. McFadden. Rationality for economists. Jrnl. of Risk & Uncertainty, 19(1):73–105, 1999.

[32] S. Lichtenstein and P. Slovic. The Construction of Preference. Cambridge Univ. Press, 2006.

[33] I. Witten and E. Frank. Data Mining. Elsevier, 2005.

[34] Hans-J¨org Happel and Walid Maalej. Potentials and challenges of recommended systems for software development. In RSSE ’08: Proceedings of the 2008 international workshop on Recommended systems for software engineering. ACM, Nov 2008.

[35] A. Felfernig, G. Ninaus, H. Grabner, F. Reinfrank, L. Weninger, D. Pagano, and W. Maalej. An Overview of Recommender Systems in Requirements Engineering. Springer, pp. 315–332, 2013.