Journal of Information Technology Impact
Vol. 5, No. 3, pp. 129-138, 2005

Using a Priori Algorithm for Supporting an e-Commerce System

Mohammad Nazir Ahmad Sharif[1] Ng Moon Ching[2] Aryati Bakri[3] Nor Hidayati Zakaria[4]
Universiti Teknologi Malaysia
Malaysia

Abstract

The Internet technology has brought about a significant impact in doing business. It promotes the new way of doing business by enabling new system such as electronic commerce (e-commerce) to the worldwide users. Currently, the e-commerce system does not only provide electronic transactions like online payment, electronic cart shopping and ordering, and online tracking, but it must also be able to support a good relationship with their customers by providing a creative way in its business operations. It is because of many organizations having to maintain their customers by serving a good customer satisfaction. Lack understanding of the customers will cause an organization loss their customers and then would loss the company profit. This paper demonstrates the development of e-commerce system by focusing on the use of a Priori algorithm as supported feature in our e-commerce system. The feature is included to increase a good customer relationship management for the proposed system. It is hoped the proposed prototype would illustrate some practical ideas on how much advantages can be benefited from the e-commerce system and customer relationship management

Keywords: Business impact, customer relationship management, payments, transactions.

Introduction

The Internet is evolving rapidly. It has brought in a new style for doing business and also brings many unsolved issues on security, trust, and privacy that have been taking into account. However, if we compare to the United States, the Americans have been doing internet-based business for ages (Ghazali, Mohd Mohayidin & Shaari Abd Hamid, 1998). On the contrary, in the country like Malaysia, most of the people are still conservative due to seriously concern about security, trust and the reliable of online transactions. Nevertheless, this paper does not address the cultural issues on implementing e-commerce or to discuss why most of the developing countries are still lack in exploiting the internet-based business advantages. We believe that they are going to be there. Remember when the ATM cards were first introduced? People were worried about using them. They worry about the machine not giving them the money and all sorts of security concerns. But now, ATM cards have been accepted, because of the convenience. The same is going to happen with e-commerce too. It will just be a matter of time.

The aim of this paper is to demonstrate a practical application of e-commerce by paying special attention on how the development of e-commerce system must be able to include a Customer Relationship Management (CRM) feature. Although the proposed system is just a prototype application, we believe that it is capable to illustrate the application of current trend or technique in e-commerce development and then to spark a new idea in making other e-commerce system become better.

E- Commerce and Customer Relationship Management

E-commerce is used in its broadest sense: business activities conducted using electronic data transmission technologies such as those used on the Internet and World Wide Web (Schneider, 2003). The business activities may involve the customer transaction, selling, buying, product exchanging, services and information. The product and service are delivered to customer using physical or digital. For example, FedEx physically delivers books that customers shopping online at Amazon.com.

Customer relationship management (CRM) is an approach that recognizes that customers are the cores of the business and the company's success effectively depends on managing their relationship with them (Turban, 2005). It aims to identify the customer based on their needs, behaviors, preferences, buying patterns and so on. This information is used to identify the potential customer, predict the future sales, marketing, promotions, to set the suitable prices and so forth. Customer loyalty is important to increase the company profit and improve customer service.

E-commerce influence impacts the need for quality and accurate CRM (Turban, 2005). The term has been used to show the relationships between the CRM and e-commerce is electronic commerce customer relationships management (ECCRM) and some researcher call as technology-enabled relationship management. ECCRM refers to application of CRM in electronic commerce. According to Schneider (2003), ECCRM to obtain the information to specific customers, to identify and respond to specific customers behaviors and preferences, to give promotions and discount to individual customers based on preferences and buying patterns, to set the price of products and services to specific customers, to enhance the features of new product based on customer respond and to manage the relationship with the current customer. The CRM in web has been used to answering customer inquiries, providing search and comparison capabilities, providing technical information to customers, allowing customers to track order status and allowing customers to place an online order (Turban, 2005).

Today, CRM includes business intelligence or business analytics through data mining and OLAP technology (Turban, 2005). These tools can produce the analysis based on customer behavior, demographics, products and other factors. Data mining used to do the customer's fragmentation, prediction and so on. For instance, optimization algorithms used to process of calculating the price to get the maximize profit.

Data Mining – Association Rules

Data mining is a process of discovering meaningful patterns and trends by mining large amounts of data stored in database (Witten and Frank, 1999). The major attraction of data mining is its capability to build predictive. Data mining is not specified to any industry. It requires intelligent technologies and the willingness to explore the possibility of the hidden knowledge in the data to predict customer behavior.

Mining association rules from market basket data is discovering hidden knowledge from the database. Another phase for market basket data is transaction data. An association rule is the main data mining technique is used to trace the categorized variables pattern in the large data set. This technique mostly applies in the business area to analysis customer preferences and their purchase behavior (Rud, 2001). The transaction pattern hides in the large data set can be found out. With the detection of association rules, set product that purchased together frequently can be determined. Association rules mining search for the relationships between items in a data set. It finds association among sets of items in transaction database. For example, customer who buys product A tends to buy product B and product C as well. To find the association rules, a priori algorithm is applied.

There are many association rules that can be generated. Importance of an association rule can be measured by the value of support and confidence. The concept of frequent enough is a parameter of the algorithm, used to control the number of association rules discovered. Support specifies how frequently the items must appear in the whole data set before the items can be considered as a candidate association rule. Meanwhile, confidence specifies a minimal probability for the association rule. There is a support-confidence framework for the association rules. Let's say $I = \{i_1, i_2, \ldots, i_m\}$ is item set. D is transaction set where each transaction d is set for item; $d \epsilon I$. Transaction D that contains X is sub item set I. An association rule in the form $X \rightarrow Y$, where $X \epsilon I$, $Y \epsilon I$ and $X \cap Y = \phi$. The rule of $X \rightarrow Y$ will have the support, s in D if there are s% for $X \cup Y$ in transaction. Besides, this rule will maintain in the set transaction D with confidence c if there are c% of the transaction in D supports X and Y (Grossman et al., 2001).

A Priori Algorithm

A priori algorithm is an influential algorithm for mining frequent item sets from market basket data for Boolean association rules. It will extract the frequent item sets from the candidate item sets. Support for the frequent item sets must be greater than minimum support that defined by user. Generally, there are two main steps to implement a priori algorithm. Firstly, determine the frequent item sets. Secondly, generate the association rules from the frequent item sets that fulfil the requirement of minimum confidence. A priori algorithm contains a number of passes over the database. During pass k, the algorithm finds the set of frequent item sets L_k of length k that satisfy the minimum support requirement. The algorithm terminates when L_k is empty. A pruning step eliminates any candidate, which has a smaller subset. The pseudo code for a priori algorithm is shown in figure 1.0 as below.

```
Ck: candidate item set of size k,  Lk: frequent item set of size k

        L1 = {frequent items};

        For (k=1; Lk != null; k++) do begin

                Ck+1 = candidates generated from Lk;

                For each transaction t in database do

                        Increment the count of all candidates in

                        Ck+1 that are contained in t
```

Figure 1. A Priori Algorithm

The implementation steps for a priori algorithm is shown as following (Grossman et al., 2001):

i. Analyze all the transactions in a dataset for each items support count. Any item with support count less than minimum support count required is removed from candidate items.

ii. Initially each of the items is a member of a set of first candidate item sets. The support count of each candidate item in the item set is calculated and items with a support count less than the minimum required support count are removed as candidates. The remaining candidate items in the item set are joined to create second candidate item sets that each comprise of two items or members.

iii. The support count of each two member item set is calculated from the database of transactions and two member item sets that occur with a support count greater than or equal to the minimum support count are used to create third candidate item sets. The process in steps (i) and (ii) is repeated generating fourth and fifth candidate item sets until the support count of all the item sets are lower than the minimum required support count.

iv. All the candidate item sets generated with a support count greater than the minimum support count form a set of frequent item sets. These frequent item sets are then used to generate association rules with a confidence greater than or equal to the minimum confidence.

A priori algorithm recursively generates all the subsets of each frequent item set and creates association rules based on subsets with a confidence greater than the minimum confidence.

Implementation

The case study for this project is done at Four Star Chair Systems Enterprise. The major business activity at Four Star Chair Systems Enterprise is providing and selling various types of chairs especially office chairs for customers. The company runs their business conventionally and manually. Lack understanding of the customers causes the organization losses good and valuable business opportunity. Besides, the organization also faces with problem losing it own customers to other organizations. Therefore, this project is developed to provide an e-commerce and web based customer relationship management (CRM) system which organization can predict the customers' orders patterns. With the prediction, organization can provide suitable services to the targeted customers.

The use case diagram as depicted in figure 2.0 and figure 3.0 illustrate the functionality of our system. There are two main actors, online customer and administrator. The customers can register and then login into the system. After that, he can view catalog, searching for needed products, place orders, and also proceed to online payment. Interestingly, when customer views the catalog, he can browse the detail products and meanwhile the system will recommend what are others related products can be accessed which associate to current viewed product. To implement this CRM function, we apply a priori algorithm in the use case called "recommend product", which has an extend relationship with a "view catalog" use case (see figure 2.0). Meaning, system will only recommend associate products to customers if the product that is viewed by them exists in the frequent item sets in transactions.

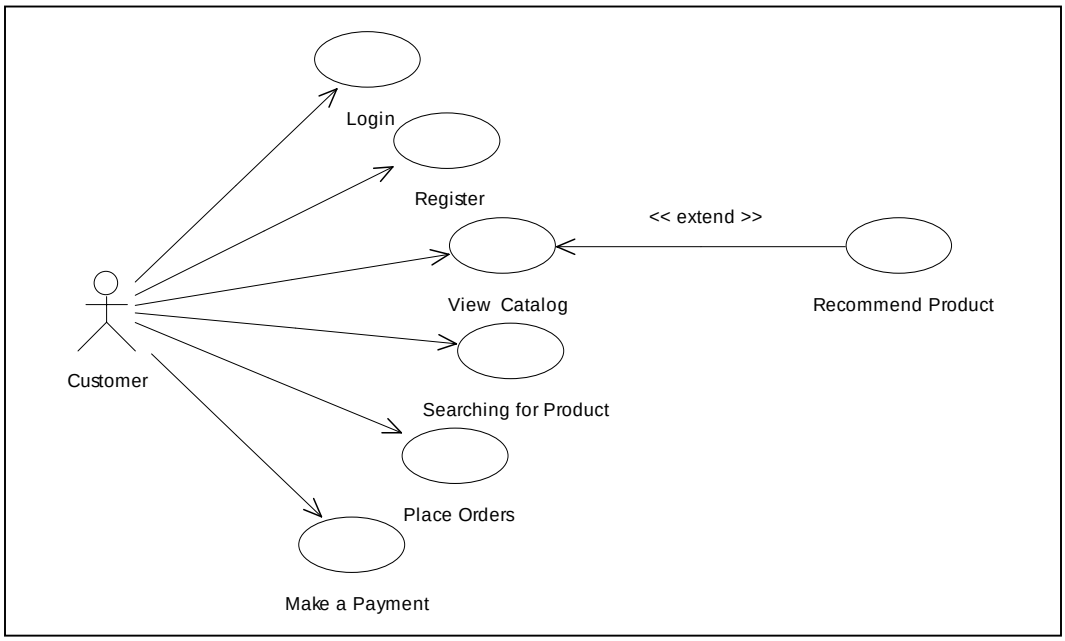

Figure 2. The use case diagram for customer's view

Therefore, not every time the customers view a product's detail will be provided with associate product. If the viewed product do not exist in frequent item sets in transactions, there are no associate product will be recommended to customers. For the administrator, he can also login into the system, send email, insert information, delete information and produce reports (see figure 3.0).

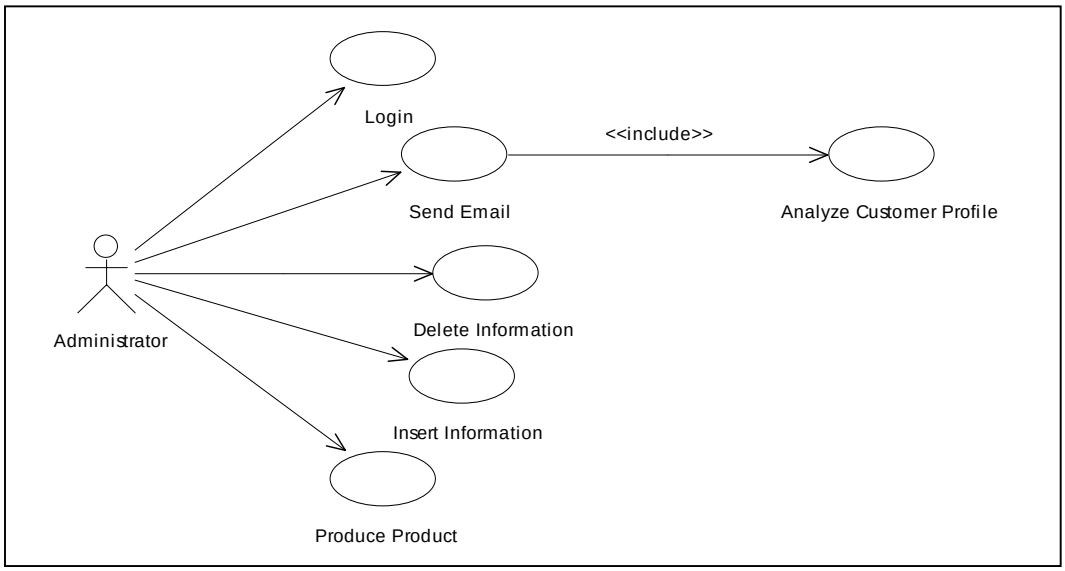

Figure 3. The use case diagram for administrator's view

In order to support another one CRM element, we propose an "analyze customer profile" use case, which is included in "send email" use case as shown by an include relationship (see figure 3.0). This is a kind of push technology that assists administrator to send online advertisement about new products or upcoming products to the right customers based on their profile. Customer profile is analyzed by using the a priori algorithm. The algorithm will find the frequent item sets and generate association rules. From the result, it will compare the customer profile with the product profile to extract the potential customers' email address. These two CRM features assist the customers doing online shopping, by helping them to buy the right product as demonstrated by "recommend product" and "analyze customer profile" use cases.

As discussed earlier, when customer views a product's detail, the system will recommend associate products to customers. This is also a guide for customers to make them get their desired products easily. The products recommended are the result of association rules and analyze customers profile with pass orders behavior. Besides, it also implements the concept of CRM where the system give the right offer to the right person, see figure 4.0 below.

Figure 4. A user interface for a product recommendation

In figure 5.0 below, it is a user interface for administrator to find frequent item sets and generate association rules.

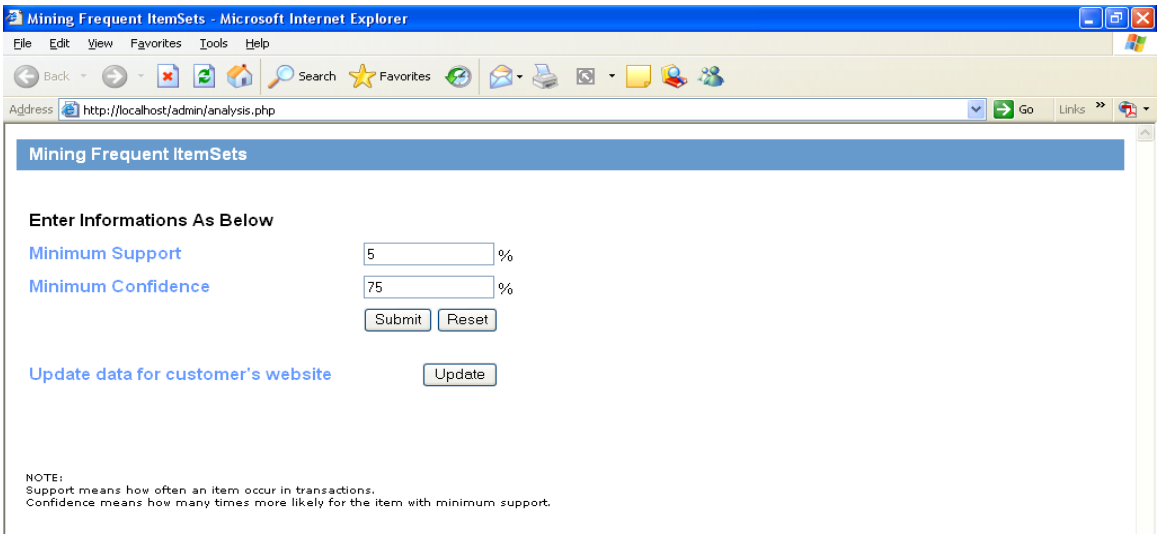

Figure 5. A user interface for generating association rules

Administrator needs to enter the value for minimum support and minimum confidence to find the frequent item sets and then generate association rules. With the result of frequent item sets and association rules, it will find the potential customers' email. When there are new releases, emails are sent to the potential customers. Here, it applies the CRM element by offering the products to the targeted customer at the right time through the right channel.

Conclusions

Nowadays, the Internet transactions are increasingly sophisticated and become more complex. Therefore, the development of e-commerce systems is always not an easy task. There is huge information that needs to be handled by an intelligent e-commerce system. Many ways can be done for supporting the capabilities of e-commerce. The best way is, to propose the e-commerce system that able to enhance customer satisfaction. However, being preferred by the customer is not easy, an e-commerce system must have some characteristics or system features that can support customer relationships. Thus, besides by having main components of e-commerce system such as shopping carts, online catalogs, orders and online payments, current development of e-commerce system should be able to include some elements of CRM. Today, a few technologies like data mining, ontology, OLAP and so on are recently recognized as an essential tool for supporting e-commerce system with CRM features. In this paper, we demonstrate on how e-commerce system can be developed with supported by CRM. It is hoped that, this simple prototype of e-commerce system will enable some ideas on e-commerce and CRM.

Bibliography

Berson, A., Smith, S., Therling, K. (2001) *Building Data Mining Application for CRM*, McGrawHill, USA.

Castagnetto, J., Rawat, H., Schumann, S., Scollo, C., Veliath, D. (2000) *Professional PHP Programming,* 3rd Ed, Wrox Press Ltd, USA

Forder, J & Quirk, P. (2001) *Electronic Commerce and the Law,* John Wiley & Sons, Australia.

Grossman, R.L., Kamath, C., Kegelmeyer, P., Kumar, V., Namburu, R. R. (2001) *Data Mining For Scientific and Engineering Applications,* Kluwer Academic Publishes, Boston, USA.

Ghazali, Mohd Mohayidin and Shaari Abd Hamid (1998) *Small and Medium Sale Enterprises in*

Malaysia: Technological and Managerial Capacity, Universiti Putra Malaysia: Kuala Lumpur.

Mohd Jan, Nawawi, Jamaluddin Che Sab and Johari Shaari (1990) "Strategies Toward Improving SMIs in Malaysia: With A Particular Reference to The Manufacturing Sector", *Proceedings of the International Conference on Small and Medium Scale Enterprises*, Universiti Utara Malaysia, Langkawi Island.

Mohd Khairuddin Hashim & Syed Azizi Wafa (2001) *Small & Medium Sized Enterprises In Malaysia,* Kuala Lumpur: Prentice Hall Inc.

Rud, O.P (2001), *Data Mining Cookbook: Modeling Data For Marketing, Risk, and Customer Relationship Management,* John Wiley & Sons, Canada.

Schneider, G. P. (2003) *Electronic Commerce,* 4th Ed, Thomson Course Technology, Canada.

Turban, E., Aronson, J.E., Liang, T.P. (2005) *Decision Support Systems and Intelligent Systems,* 7th ed, Prentice Hall, USA.

Turban, E., King, D., Lee, J., Warkentin, M., Chung, H.M. (2002) *Electronic Commerce 2002: A Managerial Perspective,* Prentice Hall, New Jersey, USA.

Witten, I. H., Frank, E. (1999) *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementation*, Academic Press, USA.

---

[1] Mohammad Nazir Ahmad is a lecturer at the Department of Information Systems, Universiti Teknologi Malaysia. He can be reached at: Department of Information Systems, Faculty of Computer Science & Information Systems, UTM 81310, Skudai, Johor, Malaysia. Email: nazir@fsksm.utm.my; Phone: +607 553 2433

[2] Ng Moon Ching is a graduate student of the Universiti Teknologi Malaysia with a degree in Computer Science. She can be reached at: moonching@gmail.com

[3] Aryati Bakri is a lecturer at the Department of Information Systems, Universiti Teknologi Malaysia. She can be reached at: Department of Information Systems, Faculty of Computer Science & Information Systems, UTM 81310, Skudai, Johor, Malaysia. Email: aryati@fsksm.utm.my; Phone: +607 553 2428

[4] Nor Hidayati Zakaria is a lecturer at the Department of Information Systems, Universiti Teknologi Malaysia. She can be reached at: Department of Information Systems, Faculty of Computer Science & Information Systems, UTM 81310, Skudai, Johor, Malaysia. Email: hidayati@fsksm.utm.my; Phone: +607 553 2426

Page left blank