Proceedings of the World Congress on Engineering and Computer Science 2008
WCECS 2008, October 22 - 24, 2008, San Francisco, USA

Exploring Various Forms of Purchase Dependency in Retail Sale

Pradip Kumar Bala

***Abstract*—The present research paper explores into various forms of purchase dependency in retail sale with large number of items. In the existing works, dependency amongst the demands of a small number of items has been addressed under specific situations. However, for large number of items, it has remained unaddressed. Moreover, demand dependency is not the same as purchase dependency. Purchase dependency is a feature of consumer of behavior and it can be used as a valuable input in various decision making processes.**

Index Terms—**Data Mining, Negative Purchase Dependency, Purchase Dependency, Retail, Sequential Purchase Dependency.**

I. INTRODUCTION

Purchase dependency refers to dependency of purchase of an item or itemset by a retail customer on the purchase of another item or itemset by the same customer. Itemset is set of items with one item or more items. Purchase dependency can be within a transaction or it can be inter-transactional also. In inter-transactional purchase dependency, purchase of item or items in a transaction by a customer depends on the purchase of item or items in a previous transaction by him or her. Let us consider an example with two items, x and y, where x has an impact on the purchase of y within a transaction. This implies, some of the purchases (or, all purchases) of y depend on the purchase of x. With such purchase dependency, there is a chance that a customer intends to purchase y whenever he/she purchases x. In other words, purchase of x gives a lift to purchase of y. In fact, in many cases, y is purchased only when x is also purchased. Hence, stockout situation of x results in non-purchase of y in many cases, although the later is in stock. It will result in lost sale of not only x but also of y. This will also give rise to inventory holding cost of y as it remains in stock but does not sell in many cases due to purchase dependency on x.

According to [1], backorder in stockout situation is not always a realistic assumption. In retailing, backorder is not acceptable in a retail store when the retailers are in a competitive market and customers can easily turn to another retailer for purchasing goods. Hence, stockout results in a lost sale. In other language, unsatisfied demands result in lost sale in retail sale. Knowledge of purchase dependency can be used as an additional input in designing inventory replenishment policy, in designing effective promotional offers and in shelf designing in retail stores.

The author is working as Assistant Professor in Xavier Institute of Management, Bhubaneswar, PIN-751013, INDIA (phone: +91-674-3983857; fax: +91-674-2300995; e-mail: p_k_bala@rediffmail.com).

Purchase dependency in retail sale is important input in decision making and it has remained unaddressed in the existing works. A few works focus on identifying demand dependency in limited extent with a few items. **Demand dependency is different from purchase dependency as the former talks about the relationship between or amongst the periodic demands of the items, whereas the later talks about the purchase behavior of customer.** The present research paper focuses the problem of purchase dependencies amongst the large number of items in retail sale and explores into various forms of purchase dependencies.

II. LITERATURE REVIEW

A review of the existing literature on retail management has been made in this section. The retail industry faces stockout rates of 5-10% which results in sales losses of up to 4% corresponding to hundreds of millions of dollars for large retailers [2]. Various causes for stockout situations have been identified and they are mostly attributed to inefficiencies in in-store logistics due to the lack of inventory visibility. Retailers of products with limited shelf life face with the dilemma of stocking the right mix of standard product and customized stock keeping units in each product category [3].

Vendors have the obligation to maintain the inventory of retail stores in vendor managed inventory (VMI) in many places [4]. J.C. Penney has become a successful retailer due to several reasons including its supply chain infrastructure [5]. The paper in [6] develops a model of a periodic review multi-location inventory system that investigates the characteristics of optimal replenishment and transshipment decisions. The paper in [7] employs the technique of support vector machine (SVM) for demand forecasting. Various factors that affect the product demand such as seasonal and promotional factors have been taken into consideration in this model. A generic modeling framework has been proposed in [8] to address a number of issues, including stockouts, which continues and extends a recent stream of research aimed at integrating insights from modern inventory theory into the supply chain network design domain. Reinforcement learning (RL) techniques have been used in [9] to determine dynamic prices in an electronic monopolistic retail market. The market has been considered to consist of two natural segments of customers, captives and shoppers.

Demand dependency has been taken into account to a limited extent in a few articles. The article in [10] uses Markovian model for a two-item inventory system with correlated demands and coordinated replenishments, an expression for the long-run total cost per unit of time under the can-order replenishment policy has been obtained. The paper in [11] considers the level interdependency between two items and gives a new approach towards a two-item inventory model for deteriorating items with a linear stock-dependent demand rate. In 'level dependency' of two items as explained in the paper, the demand of an item depends in its own inventory level and the inventory level of the other item. The presence of one item reduces the demand for another item.

It is observed that purchase dependency has remained unrecognized and unaddressed in the existing works.

III. EXTRACTION OF PURCHASE DEPENDENCIES

Light has been thrown on different aspects of purchase dependencies in this section. Two relevant aspects of purchase dependency amongst the items in the context of retail management have been identified as given below.

(i) Purchase of one itemset having an impact on purchase of another itemset in the same sale transaction

(ii) Sequence of purchases where the same sequence gets repeated with a time gap (of a number of days/weeks/months etc.) between two consecutive transactions and purchase in a transaction depends on the purchase in the previous transaction or transactions.

Various types of purchase dependencies (within a transaction and inter-transactional) along with the suitable data mining techniques to extract these dependencies have been discussed for two cases as given below.

(i) Identifying Purchase Dependencies within a Transaction:

Association rules can be used to learn purchase dependencies within a transaction. In some cases, purchase dependency within a transaction may be related to the customer profile. This implies that there is dependency in purchase of these items for a specific customer profile. **Classification** and **clustering** can be used for learning the impact of **customer profile** on purchase dependency of one itemset on another itemset. Hence, the data mining techniques for classification and clustering of customers and products along with association rule mining may be useful for learning such purchase dependencies.

Various types of purchase dependencies may be observed in different time periods. This results in **time-dependent purchase dependency**. On different days of the week, there may be different purchase pattern of the customers. Purchase patterns on the weekdays and the weekends are generally different. Similarly, in different months or seasons, different patterns may be observed. Time-dependent purchase pattern may also be observed at different hours of the day. Purchase pattern in the evening hours happens to be different from the day hours. Segregated data with respect to time can be used to learn time-dependent purchase dependencies. For example, data of weekend and weekdays can be segregated for this purpose. Data mining tasks of association rule mining, classification, clustering etc can be employed on the segregated data for learning time-dependent purchase dependencies.

Similarly, purchase dependency may be different at different locations of a retail organization which results in **location-dependent purchase dependency**. There may be an impact of culture, climatic condition and other factors on the purchase pattern. To learn location-dependent purchase dependencies, data mining is to be performed on segregated data with respect to location or space.

(ii) Identifying Inter-Transactional Purchase Dependencies:

In inter-transactional purchase dependency, the purchase of item or items in a transaction depends on the item or items purchased in a past transaction. This may follow a sequence of purchases in two or more successive purchases. Sequence mining can be used to learn such inter-transactional purchase dependencies. Frequently repetitive sequence of purchases can be used to depict purchase dependency in a sequential form.

IV. EXPLORING VARIOUS FORMS OF PURCHASE DEPENDENCY

Various forms of possible purchase dependencies in retail sale with large number of items have been explored in this section which can be extracted through data mining. Prior to using these purchase dependencies in decision making, it must be examined whether they depict purchase dependency. Some of the different forms of purchase dependency are as given below.

(i) Purchase Dependency depicted by Association Rule:

In this section, a focus has been given on the use of association rules for depicting purchase dependency. Association rules are mined from the sales transaction data using data mining software with threshold values of support and confidence. The problem is to know which set of items are dependent on which other set of items while being purchased. In an inventory with very large number of items, association rule mining gives solution for finding the purchase dependencies amongst the items.

While considering an association rule, X=>Y, in any decision making process, there may be four situations at any point of time which are as given below.

(i) Both X and Y are in stock

(ii) X is in stock but Y is not in stock

(iii) X is not in stock but Y is in stock

(iv) Both X and Y are not in stock

The present research paper focuses on the situation (iii), i.e., when X is not in stock but Y is in stock. In 'X => Y', purchase of X gives a lift to purchase of Y. Under this situation, there is a chance that some customers with demand for both X and Y will not purchase Y in the absence of X as demand of Y depends on purchase of X in many cases. In fact, this purchase dependency is depicted by the association rule, 'X => Y'.

Association rules have been mined from 8,418 sale transactions of 45 grocery stock keeping units (SKUs) in a retail store in India at various threshold values of support and confidence. Some of the rules mined are as given below. Names of the items are as spoken by Indian shoppers.

(i) Kismis (100 g) => Basmati (1 kg) with support = 24.44%, confidence = 92.45% and lift = 2.37

(ii) Coffee – Nescafe (50 g) => Bournvita Cadbury (500 g) with support = 23.39%, confidence = 66.97% and lift = 1.75

(iii) Noodles – Maggi (200 g) and Tomato Sauce (Maggi) (200 g) => MDH Chicken Masala with support = 22.98%, confidence = 92.63% and lift = 2.76

It is expected that an association rule will depict purchase dependency. However, it is not necessary that each and every association rule will depict purchase dependency. In line with the fact that correlation does not always imply causality, it is found that association rule may not always imply a purchase dependency. That is why it is necessary to examine an association rule whether it depicts purchase dependency. To examine, whether an association rule, "X =>Y", depicts purchase dependency, observation should be made on the impact of presence and absence of X on the purchase of Y. Observations have been made through a designed experiment for the rule, Kismis (100 g) => Basmati (1 kg) in a retail store, and it is observed that the rule depicts purchase dependency where purchase of Basmati (1 kg) depends on the purchase of Kismis (100 g). As a result, many a times, the stockout of Kismis (100 g) leads to non-purchase of Basmati (1kg). This is how lost sale due to purchase dependency occurs in retail stores.

Hence, purchase dependency depicted by association rule, x => y has been described as given below.

Conditions:

(i) Customers come with a demand for x and y both

(ii) x is out of stock

(iii) y is in stock

Purchase Dependency:

Under the above conditions, the customer does not purchase y, although it is in stock. This results in lost sale of x and y both. This happens as a result of the fact that purchase of y depends on purchase of x.

(ii) Customer Profile-Induced Purchase Dependency depicted by Association Rule and Decision Tree for Classification

Customer profile-induced purchase dependencies can be learnt through mining association rules and through induction of decision tree rules, a classification technique in data mining.

Synthetic retail sale data with eleven (11) items has been used for the research in this part. Eleven items have been named as 'a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j' and 'k'. The data contains one thousand (1,000) past sale transaction data along with the profiles of the customers in each transaction. A row in the transaction data contains the items purchased by a customer along with a profile description of the customer.

Association rules obtained from data mining with threshold values of rule support and confidence to be 10% and 80% respectively are (i) (d, g) => (f), (ii) (f, g) => (d), and (iii) (d, f) => (g). Only three rules have qualified for the chosen threshold values of support and confidence. In each of these three rules, same three items (i.e., d, f and g) are involved.

With respect to the simultaneous purchase of all three items (i.e., d, f and g) by a customer in the same transaction, classification of the customers has been done based on the customer profile. For inducting decision tree using data mining, an 'additional field' has been created for simultaneous purchase of three items, d, f and g. If all three items have been purchased, the class is 'True' and represented by 'T'. If all three items are not purchased, class is 'False' and represented by 'F'. Hence, classification has been done in two categories, 'T' and 'F' with respect to simultaneous purchase of three items. A decision tree has been inducted from the database for the purpose of classification using the details of customer profile as inputs and the 'additional field' as output class. Considering threshold value of 80% confidence for a decision rule, only one rule qualifies. The rule (with 84.242% confidence) is – IF "income is less than or equal to $16,950 and sex is F (Female)", THEN "the customer purchases all three items, i.e., d, f, and g". No other rule can be induced from the decision tree as others do not satisfy the minimum confidence of 80%.

The following purchase dependency may be considered for decision making.

Conditions:

(i) Customer Profile: "income is less than or equal to $16,950 and sex is F (Female)"

(ii) Customers come with a demand for all three items (d, f and g).

(iii) Any one of the three items (d, f and g) is out of stock.

Purchase dependency:

Under the above conditions, the customer does not purchase any of the three items (out of d, f and g) even if the item/items, other than the out-of-stock item/items, is/are in stock. This results in lost sale of d, f and g all. This happens as a result of the fact that a customer with the profile, income is less than or equal to $16,950 and sex is female, purchases all three items together (as given by the rule inducted from the decision tree) and absence of one item results in non-purchase of other items whenever all three items are in demand (as given by the association rules).

(iii) Purchase Dependency in Sequential Form

In sequential purchase dependency, purchase of an item or itemset depends on the purchase of another item or itemset in the past transaction or transactions by the same customer. This has been explained below. Table 1 shows twenty (20) transactions of four items, I_1, I_2, I_3 and I_4, made by one customer. In this table, '1' denotes purchase and '0' denotes non-purchase of the corresponding item in the column. It is observed that in 83.33% cases, whenever the customer purchases I_2, he/she purchases I_3 and I_4 in the next transaction. This sequence is found in the following five cases in Table 1.

The customer purchases I_2 in transaction 1 and then he purchases I_3 and I_4 in transaction 2.

The customer purchases I_2 in transaction 4 and then he purchases I_3 and I_4 in transaction 5.

The customer purchases I_2 in transaction 11 and then he purchases I_3 and I_4 in transaction 12.

The customer purchases I_2 in transaction 14 and then he purchases I_3 and I_4 in transaction 15.

The customer purchases I_2 in transaction 18 and then he purchases I_3 and I_4 in transaction 19.

The sequence is not found only in the following case in

Table 1.

The customer has purchased I_2 in transaction 8 and he has not purchased I_3 and I_4 in transaction 9.

Hence, the sequence is observed in 5 out of six cases, i.e. in 83.33% cases.

TABLE 1: Transactions of a Customer

Transaction	I_1	I_2	I_3	I_4
1	0	1	0	0
2	0	0	1	1
3	1	0	0	0
4	0	1	0	0
5	0	0	1	1
6	0	0	0	1
7	0	0	1	0
8	0	1	0	0
9	1	0	0	0
10	0	0	1	1
11	0	1	0	0
12	0	0	1	1
13	0	0	0	1
14	0	1	0	0
15	0	0	1	1
16	1	0	1	1
17	0	0	1	0
18	0	1	0	1
19	0	0	1	1
20	1	0	1	1

Purchase dependency can be in the form of sequence of purchases as discussed in the above example. Here, purchase of I_3 and I_4 in a transaction depends on the purchase of I_2 in the previous transaction by the customer. The sequence can spread over more than two transactions also. Hence, the frequent sequences should be mined to learn the purchase dependencies in sequential form.

(iv) Negative Purchase Dependency

Negative purchase dependency can be explored where "purchase of one item or itemset leads to non-purchase of another item or itemset" or, "non-purchase of one item or itemset leads to purchase of another item or itemset". This covers the event of 'non-purchase' of item or items along with the 'purchase', and hence, it is more generalized. This difference must be noted as purchase dependency covers only the event of 'purchase' in the rest of the article.

TABLE 2: Negative Purchase Dependency

Transaction	I_5	I_6	I_7	I_8
1	1	0	0	0
2	1	0	0	1
3	1	1	1	1
4	0	0	1	0
5	0	1	1	1
6	0	1	1	1
7	0	0	1	0
8	0	0	1	1
9	0	0	1	1
10	0	0	0	1
11	1	0	1	1
12	0	0	1	1
13	1	1	0	0
14	1	0	1	1
15	0	0	1	1
16	0	0	1	1
17	1	0	0	0
18	0	1	0	0
19	0	1	0	0
20	1	0	1	1

As there is no scalable technique or methodology to discover negative purchase dependencies from large number of items, a scalable technique is needed for this purpose. However, for small number of items, negative purchase dependencies can be extracted using statistical analyses.

Table 2 shows twenty (20) transactions of four items, I_5, I_6, I_7 and I_8 in a retail sale store by different customers. Like in Table 1, in this table also, '1' denotes purchase and '0' denotes non-purchase of the corresponding item in the column. It is observed that in 12 transactions, either I_5 or I_6 or both have been purchased. Out of these 12 transactions, in 10 transactions, only one item, i.e., either I_5 or I_6 has been purchased and in 2 transactions both have been purchased. Hence, from this exploratory analysis, negative purchase dependency between I_5 and I_6 can be predicted in 10 out of 12 cases or, with 83.33 % confidence.

V. CONCLUSIONS

The current research paper delves into various types of purchase dependencies in retail sale. As purchase dependency in retail sale with large number of items has remained unaddressed in the existing literatures, the research paper explores into various types of purchase dependencies. Purchase dependencies can be used as valuable input in various decision making processes. The current paper opens the field of purchase dependency to the practitioners and researchers both. Several other forms of purchase dependency can be explored further.

REFERENCES

[1] Jokar, M.A., Zangeneh, S. (2006). Developing a model for a two-echelon two-item inventory system with lost sale and demand substitution. ICMIT 2006 Proceedings - 2006 IEEE International Conference on Management of Innovation and Technology 2, art. no. 4037155, pp. 926-930.

[2] Metzger, C., Meyer, J., Fleisch, E., Tröster, G. (2007). Weight-sensitive foam to monitor product availability on retail shelves. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 4480 LNCS, pp. 268-279.

[3] Shah, J., Avittathur, B. (2007). The retailer multi-item inventory problem with demand cannibalization and substitution. International Journal of Production Economics 106 (1), pp. 104-114.

[4] Xu, X., Cai, X., Liu, C., To, C. (2007). Optimal commodity distribution for a vehicle with fixed capacity under vendor managed inventory. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 4614 LNCS, pp. 329-339.

[5] Atkinson, W. (2006). J.C. Penney: Pioneer of supply chain efficiency. Apparel 47 (8), pp. 14-18.

[6] Archibald, T.W. (2007). Modelling replenishment and transshipment decisions in periodic review multilocation inventory systems. Journal of the Operational Research Society 58 (7), pp. 948-956.

[7] Yue, L., Yafeng, Y., Junjun, G., Chongli, T. (2007). Demand forecasting by using support vector machine. Proceedings - Third International Conference on Natural Computation, ICNC 2007 3, art. no. 4304723, pp. 272-276.

[8] Romeijn, H.E., Shu, J., Teo, C.-P. (2007). Designing two-echelon supply networks. European Journal of Operational Research 178 (2), pp. 449-462.

[9] Raju, C.V.L., Narahari, Y., Ravikumar, K. (2006). Learning dynamic prices in electronic retail markets with customer segmentation. Annals of Operations Research 143 (1), pp. 59-75.

[10] Liu, L., Yuan, X.-M. (2000). Coordinated replenishments in inventory systems with correlated demands. European Journal of Operational Research 123 (3), pp. 490-503.

[11] Bhattacharya, D.K. (2005). On multi-item inventory. European Journal of Operational Research 162 (3), pp. 786-791.