

Computing the Potential Profit of a Sales Item from Cross-Selling Relationships

Tatsuya Mori, Katsutoshi Kanamori, and Hayato Ohwada

Abstract—This paper describes a method for computing the potential profit of a sales item from cross-selling relationships produced by association rule mining. This method generates a true ranking by which most valuable items are top-ranked as contributing to the increase of total profits even if each item is unprofitable. Such ranking is effective in real situations where some items are loss leaders in daily cross-selling. Unprofitable items in the head of a rule are likely valuable for selling more profitable items. Such potential profit is simply defined and computed in terms of the confidence factors of association rules, thus, efficient and easy implementation is possible. Moreover, presentation by ranking is simple enough to suggest a marketing strategy for sales promotion and advertising.

Index Terms—Association rule, sales analysis, cross-selling relationships.

I. INTRODUCTION

The sales section is selling combining a certain items and the items relevant to it, and it has obtained many profits rather than selling alone. From this sales method, it is thought that items have not only the value of the sales profits in a simple substance but the potential value of yielding the sales profits of other items.

The view over value of items is various, we devised new commodity value paying attention to the combination of item by which simultaneous purchase is carried out. For example, although those who generally buy television have a high possibility of buying a DVD player simultaneously, it is thought that a possibility that those who buy the DVD player will buy television is low (since a possibility of already having television is high). We think that television has the work which induces purchase of the DVD player from this example. It is thought that television has not only the items profits of itself but the added values of inducing purchase of the DVD player.

The association rule[1] utilizes the technology of data mining to determine the relevance of an item. The index of degree of support and confidence of the association rule can be used to clarify the kinds of items that are highly likely to be purchased simultaneously, or whether any relationships exists between items.

We propose a method of using the association rule for computing the value of an item to induce the purchase of other items, as a result of cross-selling relationships.

For example, rice balls brought in a monthly profit of

¥1,000 profits in one month, and tea brought in a monthly profit of ¥1,500. If the purchase of the tea is induced by the purchase of a rice ball, some profits from tea purchases are the result of rice ball purchases.

Supposing ¥500 in the item profits of tea is generated by purchase of the rice ball, in the potential items profits based on cross-selling relationships, and a rice ball, tea will change to ¥1,000 ¥1,500.

This method generates a true ranking by which most valuable items are top-ranked as contributing to the increase of total profits even if each item is unprofitable. The sales section can select loss leaders in daily cross-selling by using such ranking.

II. PREVIOUS RESEARCH

In recent years, studies have focused on the purchase relevance of items.

Tom Brijs[2] and others addressed the problem that value in business cannot be evaluated merely by an association rule. They developed (PROFSET), a model for choosing items that fitted sale paying attention to cross-selling which sells both another item related to the purchase candidate of a certain item.

Feng Tao[3] and C.H.Cai[4] proposed techniques for determining the relationships between items by weighting each item. These studies emphasize knowing the importance of items, does not have a clear idea of how important it is. In addition, they show that cross-selling relationships is related to profit but do not know how much.

Yagi[5] and others stressed the importance of the relationships between items that are purchased. This research, statistically extracts a predominance association rule and analyzes the importance of an item is analyzed using algorithms, (e.g. HITS and the page rank method). However, these studies assign a score of 1 to all items as rather than considering the importance of the item after calculation of gross profits.

III. METHOD

This chapter explains the proposed method of determining the value that induces other purchases of items, based on cross-selling relationships.

Profit is defined as having deducted cost from the gross margin (what deducted the cost price from sales) of items. Cost includes various aspects, such as transportation expense and management expenses. We define sales gross profit as what deducted sales charges from the profit on sales.

Manuscript received September 15, 2012; revised November 11, 2012.
The authors are with Tokyo University of Science, Japan (e-mail: tatuyamori3@gmail.com).

A. Calculation of Value That Induces Other Purchases

Here, we discuss computation of the value that induces the potential purchase of other items, using the association rule. The features of this method are the point of delivering each item profits based on the association rule, and calculating per transaction of sales data. For example, when the association rule “riceball \Rightarrow tea, conf = 50%” exists between rice ball and tea, we perform calculation of delivering 50% of the item profits of the rice ball to the rice ball from the tea. By delivering item profits of ¥500 to the rice ball from the tea, the value which induces other purchases is ¥1,000 of rice ball, ¥1,500 of tea. These profits delivered shall be called profits relevant to other purchase; the value which induces other products is calculated by adding profits relevant to other purchase to the original item profits. The feature of the proposal technique is explained below by dividing into three.

1) Calculation in transaction units

The delivery of item profits calculated in a transaction unit. If the delivery of item profits calculated using gross profit, the delivery of item profits of transaction which unrelated to association rule calculated. Accordingly, this method calculates the delivery of item profits in a transaction so that the items which sold alone are not included in the item profits.

2) One-on-one association rule

With one-on-one association rule, the profit relevant to other purchases based on cross-selling relationships is defined as follows.

- The association rule makes item X a head or a body.
- The kinds of item that make item X a head are m ; their items are Y_1, Y_2, \dots, Y_m and their confidence levels are $\epsilon_1, \epsilon_2, \dots, \epsilon_m$.
- The kinds of item that make the item X a body are n , and their goods is each Z_1, Z_2, \dots, Z_n , their confidence is each Y_1, Y_2, \dots, Y_n .

Profits relevant to other purchase based on purchase relevance X_β are calculated as follows.

$$X_\beta = \sum_{i=1}^m Y_i \epsilon_i - \sum_{j=1}^n Z_j Y_j \tag{1}$$

The following example uses the transaction set in Table1. One-on-one association rule and confidence from this table are “A \Rightarrow B:conf =60%”, “B \Rightarrow A:conf =60%”, “A \Rightarrow C:conf=40%”, “C \Rightarrow A:conf =30%”, and “B \Rightarrow C:conf =40% “C \Rightarrow B:conf =30%”.

TABLE I: THE EXAMPLE OF A TRANSACTION SET

Item ID	Profits unit price
A	¥1000
B	¥1100
C	¥500

Transaction ID	Itemset (unit sales)	Transaction ID	Itemset (unit sales)
1	A(9),B(10),C(20)	6	C(2)
2	B(5),C(10)	7	C(2)
3	B(2)	8	A(2),B(1)
4	A(4),B(2)	9	A(4)
5	A(5),C(10)	10	C(1)

If item profits are delivered in transaction ID 1, the item A will receive profits from the item B and the item C and the item B, C process in the same way. As for profits relevant to other purchase of each item, the profit of item A change to ¥-600 and the item B change to ¥-2,600 and the item C changes to ¥-2,000.

3) Multi-association rule

With a multi-association rule, two or more items exist in the head or body of the association rule. After delivering profits relevant to other purchases, it is necessary to assign them to a node from the node (a head or a body) of an association rule to each item. The item profits of the node that consists of two or more items are simply the sum of the item profits of the items in the node, and related profits of purchases are allocated based on the number of sales for each item.

With the multi-association rule, the definition of the profits relevant to other purchases based on cross-selling relationships is as follows.

- The multi-association rule is “ $X \Rightarrow Y : conf = \epsilon\%$ ”.
- The number of items contained in the head X of the association rule is m , and the number of items contained in the body Y of the association is n .
- The items contained in the head X are X_1, X_2, \dots, X_m , items profits are $X_{p1}, X_{p2}, \dots, X_{pm}$, and numbers of items are $X_{c1}, X_{c2}, \dots, X_{cm}$.
- The items contained in the body Y are Y_1, Y_2, \dots, Y_n , Yn , items profits are $X_{p1}, X_{p2}, \dots, X_{pn}$ and numbers of items are $Y_{c1}, Y_{c2}, \dots, Y_{cn}$.

In this case, the profits relevant to other purchases β received and passed to head X from body Y , and the profits relevant to other purchases $X_{\beta m}$ and $Y_{\beta m}$ assigned by each item are calculated as follows.

$$\beta = \sum_{i=1}^m X_{pm} \times \epsilon \tag{2}$$

$$X_{\beta m} = \beta \times \frac{X_{cm}}{\sum_{i=1}^m X_{cm}} \tag{3}$$

$$Y_{\beta m} = - \left(\beta \times \frac{Y_{cn}}{\sum_{i=1}^n Y_{cn}} \right) \tag{4}$$

The following example uses Table 1. The multi-association rule and confidence of transaction ID 1 are “AB \Rightarrow C:conf = 30%”, “AC \Rightarrow B:conf =50%”, “BC \Rightarrow A:conf = 50%”, “A \Rightarrow BC:conf = 20%”, “B \Rightarrow AC:conf =20%”, “B \Rightarrow AC:conf = 20%”, “C \Rightarrow AB:conf = 10%”.

If the profits relevant to other purchases are calculated according to the definition, the profit of item A becomes ¥-4,067, that of item B become ¥-1,768, and that of item C become ¥-5,835.

B. Final Value that Induces the Purchase of Other Items Based on Cross-Selling Relationships

The final value that induces the purchase of other product

based on cross-selling relationships is calculated below.

STEP1 An association rule is extracted from transaction data.

STEP2 The kind of product is turned up for every transaction and they are turned up whether to exist in the extracted association rule.

STEP3 Based on a definition, profits relevant to other purchase of each items is calculated the case where the association rule which fulfills the conditions of Step2 is discovered.

STEP4 The purchase-related profits of each item are determined using all the association rules relevant to the transaction and are added to the original product Profits of each item.

STEP5 Steps 2, 3, and 4 are repeated for all transactions, and the product profits of each product are totaled.

The following example uses Table I. The value that induces purchases of other items in transaction ID 1 in Table 1 can be calculated by adding the profits relevant to other purchases for which it calculated by 3.3.2 and 3.3.3 to the original product profits. As a result of this calculation, the profit of item A becomes ¥4,333, that of item B becomes ¥11,832, and that of item C becomes ¥13,835.

From these calculations for all transactions, the value that induces the purchase of other items becomes ¥20,912 for item A, ¥25,453 for item B, and ¥25,135 for item C.

Essentially, the sales gross profit of item was gaining high profits gross profits decreased in the order of A, B, and C, but the value that induces the purchase of other products based on cross-selling relationships decreased in the order of B, C, and A. It can thus be said that commodity value changed with the information of the profits relevant to other purchases.

IV. EXPERIMENT

In this research, the experiment for checking whether the value that induces other purchases is actually calculated based on sales gross profit, sales volume, and confidence of the association rule. The Apriori algorithm [6], a popular and efficient serial algorithm, was used for extracting the association rule.

We used data on Food Mart, which is a sample cube of Microsoft SQL Server Analysis Services. A total of 34,077 transactions and 1,559 items exists in the sales 1998 dectable of Food Mart. Moreover, 10,281 customer data records are stored in the customer table. The degree of minimum support to which the minimum count number of an item is set to 4, 5, 6, and 7 from this data; 0.5%, 1%, 5%, 10%, 20%, 40%, 60%, and 80% of the minimum confidence were given and the association rule was extracted. The value that induces other purchases based on cross-selling relationships was calculated using the association rule of these 32 patterns.

The implement was performed using MySQL and java.

The data of Food Mart was registered into MySQL and the association rule was extracted from the transaction data.

The value which induces other items from the association rule which were extracted and transaction data is computed.

A. Result

1) Ranking of value

When we used the association rule for 4 and 5 as the minimum count number, the items with low-ranking profits were raised to a higher rank (Tables II and III). Examination of the association rule for items whose ranking improved revealed significantly, and so more than the number in the part number of the rule head is that there are items in the body, and a relatively high degree of confidence was found.

2) Relationship of the association rule relevant to items

Shown in Tables IV and V of the item within the top five from Table II, the relevant results of the association rules. Table 4 indicates the average value of confidence and the number of the rule that has items in the head, and Table5 indicates the average value of the confidence and the number of the rule that have items in the body. When the minimum numbers are 6 or 7, there are few multi-association rules, and no significant change of ranking was observed. The change of the profits relevant to other purchase is presented in Fig. 1.

TABLE II: THE VALUE WHICH INDUCES OTHER ITEMS USING THE RULE OF 0.5% OF THE MINIMUM CONFIDENCE AND MINIMUM COUNT 4

Rank	Original Rank	Item-name	Value which induces other items
1	17	Netherlands Village Salt-free Popcorn	370516
2	204	Dark Cheddar cheese	303283
3	27	Netherlands Village Raisin Cookie	229337
4	53	Hilltop Conditioning Shampoo	216882
5	839	Bees Wax Candle	211199
6	215	Hilltop MG Acetaminophen	184025
7	104	Mars Low-fat French Fries	143848
8	939	Toilet Cleaner	129321
9	10	Helmond Green Pepper	122882
10	9	Carrington Roast Turkey	120298
11	2	Better Chicken soup	118497
12	1	Great Taste Black Bread	117139
13	6	Special Enlarged Peanut Butter	113474
14	7	Milk Cottage Cheese	112075
15	4	Rye Bread	110748
16	43	Fresh Lima beans	109858
17	206	Mars Freezing Cheese Pizza	109754
18	29	Monde Medal Rice Pudding	109685
19	161	Scissors	108584
20	5	Netherlands Village Cookie	108388

TABLE III: THE VALUE WHICH INDUCES OTHER ITEMS USING THE RULE OF 0.5% OF THE MINIMUM CONFIDENCE AND MINIMUM COUNT 5

Rank	Original Rank	Item-name	Value which induces other items
1	17	Netherlands Village Salt-free Popcorn	164361
2	204	Dark Cheddar cheese	157099
3	3	High mountain New Potato	127410
4	1	Great Taste Black Bread	115359
5	2	Better Chicken soup	111219
6	27	Netherlands Village Raisin Cookie	111172
7	4	Rye Bread	107846
8	5	Netherlands Village Cookie	107838
9	6	Special enlarged peanut butter	107432
10	7	Milk Cottage Cheese	107072
11	10	Helmond Green Pepper	106428
12	9	Carrington Roast Turkey	106295
13	8	Strawberry Drink	106096
14	11	Hilltop Mint Mouthwash	104804
15	12	Fort Williams Barbecue	103514
16	13	Carlsson Sliced cheese	103079
17	53	Hilltop Conditioning Shampoo	103055
18	14	Great taste Rye bread	102621
19	15	Foot Long Hot Dog	101744
20	16	Mars Pickled Plum	101661

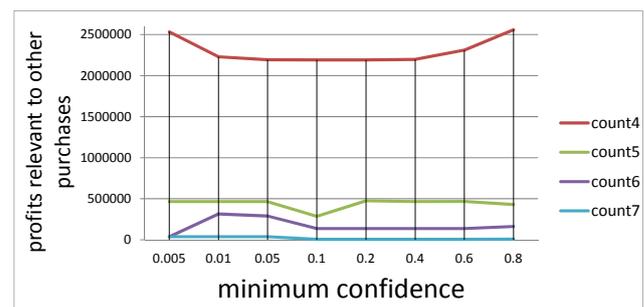


Fig. 1. The situation of the profits relevant to other purchase by the minimum confidence and the minimum count number

B. Considerations

It is thought that it is whether the reason out of which the difference of the minimum count number came to the result of 4 and 5 and the result of 6 and 7, and the multi-association rule are related. On the confidence of the characteristic, confidence of a multi-association rule tends to increase more than the confidence of one-on-one association rule. Moreover, Tables IV and V indicate that the simulation is reflected in the theoretical reasoning of the proposed technique that makes the item profits of a head standard.

Next, we describe the influence of a change in minimum confidence. A change of the profits relevant to other purchase seldom occurred (Fig. 1). Intuitively, it is considered that the low confidence association rule contribute to high profits relevant to other purchase, though the result was not.

A possible reason is that the percentage of the number of occurrences of head and body of the item are balanced, and the profits relevant to other purchase have been canceled (Fig. 1). In addition, extremely high only the profits relevant to other purchase when the number 4 count minimum, the comparison of Table II and III, the value which induces other items even Netherlands Village Salt-free Popcorn is the same item is very different. Tables VI and VII were developed in an effort to examine the association rule relevant to this item. The number for the multi-association rule in the minimum count number 4 is comparable to the minimum count number 5, and it increases by about five times. The multi-association rule with high confidence is repeatedly reflected in the calculation, and so it is thought that high value which other items was calculated.

Finally, we think that the items on top of new ranking are effective in real situations where they are loss leaders in daily cross-selling.

TABLE IV: THE AVERAGE VALUE OF THE CONFIDENCE AND THE NUMBER OF THE RULE WHICH HAS ITEMS IN THE HEAD

Item name	Average confidence	Number
Netherlands Village Salt-free Popcorn	0.69	283
Dark Cheddar cheese	0.7	167
Netherlands Village Raisin Cookie	0.71	189
Hilltop Conditioning Shampoo	0.69	177
Bees Wax Candle	0.81	151

TABLE V: THE AVERAGE VALUE OF THE CONFIDENCE AND THE NUMBER OF THE RULE WHICH HAS ITEMS IN THE BODY

Item name	Average confidence	Number
Netherlands Village Salt-free Popcorn	0.54	146
Dark Cheddar cheese	0.56	87
Netherlands Village Raisin Cookie	0.51	103
Hilltop Conditioning Shampoo	0.52	96
Bees Wax Candle	0.35	119

TABLE VI: THE AVERAGE VALUE OF THE CONFIDENCE AND THE NUMBER OF THE RULE WHICH HAS NETHERLANDS VILLAGE SALT-FREE POPCORN IN THE HEAD

		Average confidence	Number
All Association Rule	count5	0.6	60
	count4	0.69	283
One-on-One Association Rule	count5	0.05	8
	count4	0.04	22
Multi Association Rule	count5	0.69	52
	count4	0.74	261

TABLE VII: THE AVERAGE VALUE OF THE CONFIDENCE AND THE NUMBER OF THE RULE WHICH HAS NETHERLANDS VILLAGE SALT-FREE POPCORN IN THE BODY

		Average confidence	Number
All Association Rule	count5	0.43	36
	count4	0.54	146
One-on-One Association Rule	count5	0.06	8
	count4	0.04	22
Multi Association Rule	count5	0.53	28
	count4	0.62	124

V. CONCLUSION

This study, calculates the potential value that induces other purchases, using the association rule. This method is simply defined and computed; moreover presentation by ranking is simple enough to suggest a marketing strategy for sales promotion and advertising. The result of the simulation suggests using the association rule for determining items with high potential to induce other purchases. Moreover, the sales section can influence customer purchases by adjusting the minimum support.

For example, the minimum support is set low and it is made to make many association rules reflect where they want to unearth the low item of the sales gross profit from the first. How to raise the degree of minimum support can be considered where they would like to unearth useful items.

Finally, since the multi-association rule with high confidence contributes significantly to purchase related profits, it is important to consider the multi-association rule when determining what items have high potential to induce other purchases.

REFERENCES

- [1] Rukesh Agrawal, Tomasz Imielinski, and Arun Swami. Mining association rules between sets of items in large databases. in *Proc. the 1993 ACM SIGMOD Conference Washington DC*, pp. 207-216, 1993.
- [2] Tom Brijs, Gilbert Swinnen, Koen Vanhoof, and Geert Wets, "Building an association rules framework to improve product assortment decisions," *Data mining and Knowledge Discovery*, vol. 8, pp. 7-23, 2004.
- [3] F. Tao, Fionn Murtagh, and Mohsen Farid, "Weighted association rule mining using weighted support and significance framework," in *Proc. the 9th SIGKDD Conference*, 2003.
- [4] C. H. Cai, W. C. Fu, C. H. Cheng, and W. W. Kwong, "Mining association rules with weighted items," *International Database Engineering and Application Symposium*, pp. 68-77, 1998.
- [5] Kazumitsu Yagi, Shingo Okano, and Yasuhiko Morimoto. Mining of important items from bipartite-graph of association rules. in Japanese DEWS2006 6A-i7, (2006)
- [6] A. Agrawal, T. Imielinski, and A. Swami, "Trans. Datamining Mining: A Performance," *IEEE Trans. on Knowledge and Data Engineering*, vol. 5, vol. 6, pp.914-925, 1993.



Tatsuya Mori is with Tokyo University of Science, Japan.