DOI: 10.3745/KIPSTD.2010.17D.4.311

전략적 중요도를 고려한 연관규칙의 발견: WARM

최 덕 원†

요 약

본 논문은 가중치를 고려한 연관규칙탐사 알고리즘(WARM)을 제시한다. 각 전략적 요소항목에 가중치를 부여하는 것과, 각 전략요소 항목별로 원시 자료값을 정규화하는 것이 이 논문에서 제시하는 알고리즘의 중요한 내용을 구성하고 있다. 본 논문은 TSAA 알고리즘을 확장 발전시킨 연구로서 전략적 중요도를 반영하는 항목으로는 각 품목의 이익기여도, 마케팅 가치, 고객만족도 등을 사용하였다. 한 대형할인점의 실제 거래자료를 사용하여 알고리즘의 성능을 검사하였으며, Apriori, TSAA 및 WARM의 세 가지 알고리즘을 사용한 탐사결과를 비교 분석하였다. 분석의 결과 세 가지 알고리즘은 연관분석 행태에 있어서 각각 독특한 탐사행태를 보이는 것으로 나타났다.

키워드: 이행적 연관규칙 탐사, Apriori, RSAA, TSAA, 장바구니 분석

Association Rule Discovery Considering Strategic Importance: WARM

Doug Won Choi†

ABSTRACT

This paper presents a weight adjusted association rule mining algorithm (WARM). Assigning weights to each strategic factor and normalizing raw scores within each strategic factor are the key ideas of the presented algorithm. It is an extension of the earlier algorithm TSAA (transitive support association Apriori) and strategic importance is reflected by considering factors such as profit, marketing value, and customer satisfaction of each item.

Performance analysis based on a real world database has been made and comparison of the mining outcomes obtained from three association rule mining algorithms (Apriori, TSAA, and WARM) is provided. The result indicates that each algorithm gives distinct and characteristic behavior in association rule mining.

Keywords: Association Rule Mining, Apriori, RSAA, TSAA, Market Basket Analysis

1. Introduction

Association rule mining is widely used to find interesting relationships among the items of a market basket. Apriori algorithm is generally used to find Boolean associations [1, 7]. RSAA(relative support association Apriori) is a modification to find the association rules among the infrequent data items [5]. Both Apriori and RSAA algorithms make a brute force search of the association rules to test if they meet the condition for 'minimum support' or 'minimum relative support.'

Oftentimes the important semantic relationships among the market basket items are ignored since the associations with low support have to be dropped out from further considerations by most algorithms. For example, Apriori discovers frequent itemsets solely by counting the frequency of occurrence of the items in the database. There is no semantics associated with it. RSAA tries to find association rules with relatively low support [5]. Both algorithms assume that each item has equal importance in its value.

However, in the real world, certain items are more valuable than others even though they show up less frequently in the market basket. For example, the sale of a TV set returns much greater profit than a toaster oven. In this regard, a new data mining technique will be useful to find the association rules that reflect the semantic aspects or strategic importance. This paper presents a new algorithm WARM (weight adjusted association rule mining) which is an extension of the earlier algorithm TSAA [15]. While TSAA considers the semantic importance between items (e.g, TV set versus toaster-oven), WARM considers relative and strategic importance of various aspects of each item. For instance, an item may have different strategic values

† 정 회 원 : 성균관대학교 시스템경영공학과 교수
논문접수 : 2010년 6월 11일
수 정 일 : 1차 2010년 7월 22일
심사완료 : 2010년 7월 23일

pertaining to its contributions to profit, marketing value, and customer satisfaction.

2. Association rule mining and its variations

2.1 Association rule

Basic definition about association rule is briefly summarized here for refreshments of the readers' memory. Let $\tau = \{i_1, i_2, \cdots, i_m\}$ be a set of items and D a database of transactions where each transaction T is a set of items such that $T \subseteq \tau$. An association rule is an implication of the form $A \Rightarrow B$, where $A \subset \tau$, $B \subset \tau$, and $A \cap B = \phi$. The rule $A \Rightarrow B$ holds in the transaction set D with support s, where s is the percentage of transaction in D that contains both A and B. The rule $A \Rightarrow B$ has confidence c in the transaction set D if c is the percentage of transaction in D which contains A and it also contains B.

$$s = \text{Support}(A \Rightarrow B) = P(A \cup B)$$
$$c = \text{Confidence}(A \Rightarrow B) = P(B|A)$$

2.2 Algorithms for association rule mining

Association rules can be classified into many types such as Boolean, quantitative, single vs. multi-dimensional, and single-level vs. multi-level association rules, etc. [2, 3, 7]. 'Market basket analysis' belongs in the Boolean association rule category and Apriori is a typical algorithm of market basket analysis. Many variations to the Apriori algorithm have been attempted to improve the algorithm performance. They include AprioriTid, MSApriori (Multiple Support Apriori), MaxOccur (Maximum Occurrence), and RSAA (Relative Support Apriori algorithm) [2, 4, 7]. However, they differ only in minor features such as different ordering of the data elements, or placing more importance on scarce data items, etc.

Recent studies on association rule mining address various topics such as multiple support values, recurrent cyclic association rule, spatial association rule, time limited sequential association rule, and intermittent and simultaneous sequential pattern [10, 12]. Surveys on these researches indicate that no transitive association rule has been addressed nor explored so far.

2.2.1 Apriori

Since the Apriori algorithm [7] is the very foundational work of association rule mining, we assume the readers are friendly with the algorithm.

2.2.2 AprioriTid

AprioriTid uses Apriori-gen function to generate the candidate itemsets. It differs from Apriori in that it does not need to rescan the database to calculate the support level of the candidate itemsets. It scans the database only once to determine the 1-frequent itemsets and then uses only those frequent itemsets to generate the next stage candidate itemsets [12].

2.2.3 MSApriori

Some items occur more often than others in the transaction database. MSApriori(Multiple Support Apriori) is an improved version of Apriori which considers the frequency of occurrence of individual item. In fact, this algorithm assigns 'minimum item support (MIS)' level to each item in order to take into account the frequency of occurrence of the individual item [8, 9].

2.2.4 MaxOccur (Maximum Occurrence)

Earlier association rule mining algorithms did not consider recurrence issue. MaxOccur defines the association rules that consider the recurrent items when generating the k-candidate itemsets. The algorithm begins with finding the 1-frequent itemsets. It then examines the frequency of recurrent items in the frequent itemset. The process continues to generate k-itemsets until it finds the predefined 'Maximum Occurrence' item. The infrequent items that fail to meet the support level are removed from further processing [12].

2.2.5 RSAA (Relative Support Apriori)

RSAA considers the relative frequency of occurrence [5, 6, 12] and uses two support values in the mining process. The first one (min_sup_1) is the user defined 'minimum support' which is used to find the frequent itemsets. The items that do not meet this value are grouped into infrequent itemset. The second value is the 'relative support (min_sup_2)' which is used to evaluate the infrequent itemset and to see if they meet the relative support. 'Min_sup_1' must be greater than 'min_sup_2.' RSAA can also be applied to where there is a hierarchical relationship among the data items

2.2.6 TSAA (Transitive Support Association Apriori)

TSAA [15] is similar to RSAA in the sense that it uses minimum relative support as the second cut-off value. However, it uses the maximum ratio of 'k-itemset frequency over individual member item frequency' in determining additional k-frequent itemsets. This way of ratio calculation ($Rsup_j$) allows the k-candidate itemsets which have very low frequency of occurrence to be considered for inclusion into the transitive association itemset if any single member of the k-candidate itemset also has a very low frequency of occurrence. For example, if a k-itemset contains TV, the frequency of occurrence of this k-itemset is likely to be very low (say it is 'R'). However, since the frequency of occurrence of the single item 'TV' will also be very low (say it is 'r,' and usually R<r). Therefore, the relative ratio $Rsup_j$ = R/r can be large enough to pass the second cut-off level (i.e., minimum relative support). In this regard, we can say that TSAA considers strategic importance of infrequent itemsets in determining the transitive association relationship.

3. Weight Adjusted Association Rule Mining: WARM

Researches on association rule mining mostly used different weights or support levels based on the characteristics of the individual item [13, 14]. However, these approaches have weakness in that the rule mining outcome is very sensitive to how the user places weights upon the importance of each item characteristics. In this paper, we adopted to use only measurable item properties in order to exclude the subjective variations in assigning weights on each item characteristics. The weights on each item characteristics should also be determined in such a way that they are not overridden by support level and other characteristics. The item characteristics should be selected to best reflect the strategic importance of the organization. Examples of strategically important characteristics may include profitability, strategic marketing value, and customer satisfaction. This paper introduces a new algorithm WARM (weight adjusted association rule mining) which considers these strategic aspects.

3.1 Definitions

Let I^n={ I_1^n, I_2^n, I_3^n, ⋯ } denote the item set consisting of n-itemsets. For example, I^1={(i), (j), (k), ⋯}, I^2={(i, j), (j, k), (k, l), ⋯}, I^3={(i, j, k), (k, l, m), ⋯} and the subscript t of I_t^n indicates the t^{th} element of I^n. For instance, if n=3 in this example, I_1^3 = (i, j, k), I_2^3 = (k, l, m).

The following symbols represent the strategic factors used in this paper and these are derived through interviews and literature survey.

$\Pi(I_t^n)$: normalized profitability of t^{th} itemset I_t^n, n=1, 2, 3, ⋯
$\mu(I_t^n)$: normalized strategic marketing value of I_t^n
$\chi(I_t^n)$: normalized customer satisfaction for I_t^n
$W(I_t^n)$: strategic importance adjusted aggregate weight for item set I_t^n
ω_x : relative weight of strategic factor x.
$P(I_t^n)$: profitability of itemset I_t^n
$V(I_t^n)$: strategic marketing value of itemset I_t^n
$C(I_t^n)$: customer satisfaction of itemset I_t^n
$S(I_t^n)$: normalized frequency of occurrence
R_n : minimum aggregate weight (threshold) for n-itemset

Let P_i , P_j be the raw profitability of itemset i, j and I^3={(i, j, k), (k, l, m)}. Then I_1^3 = (i, j, k), I_2^3 = (k, l, m) and we define $P(I_1^3) = P_i + P_j + P_k$ for the sake of computational simplicity. Now we know that $P(I_2^3) = P_k+P_l+P_m$. And the relationship between normalized profitability $\Pi(I_t^n)$ and raw profitability $P(I_t^3)$ is defined as follows.

$$\Pi(I_t^n) = P(I_t^n) \;/\; \Sigma_j\, P(I_j^n), \quad j=1..n$$

Let v_i , v_j be the raw strategic marketing value of itemset i, j. In the same token, we can define the relationship between normalized strategic marketing value $\mu(I_t^n)$ and raw strategic marketing value $V(I_t^n)$ as follows.

$$\mu\ (I_t^n) = V(I_t^n) \;/\; \Sigma_j\, V(I_j^n)$$

And say that C_i , C_j are the raw value of customer satisfaction of itemset i, j. Then the relationship between normalized customer satisfaction $\chi(I_t^n)$ and raw customer satisfaction $C(I_t^n)$ can be defined as follows.

$$\chi(I_t^n) = C\ (I_t^n) \;/\; \Sigma_j\, C(I_j^n)$$

$S(I_t^n)$ is the normalized frequency of occurrence of itemset I (i.e., normalized support). If itemset I_t^n appears f_t times in the database D, then

$$S(I_t^n) = f_t \;/\; \Sigma_j\ (f_j)$$

Let ω_k be the weight placed on item characteristics $k\in\{\pi, \mu, \chi, S\}$ and $W(I_t^n)$ be the aggregate weight for item set I_t^n, then $W(I_t^n)$ is defined as follows.

$$W(I_t^n) = \omega_\pi(I_t^n)+\omega_\mu(I_t^n)+\omega_\chi(I_t^n)+\omega_s(I_t^n)$$

where $\omega_\pi + \omega_\mu + \omega_\chi + \omega_s = 1$

Now the above formula shows that $W(I_t^n)$ is the aggregation of three strategic factors (profitability, strategic marketing value, and customer satisfaction) and one support value each of which is weight adjusted by the respective strategic importance.

Lastly, we need to define the minimum aggregate weight (threshold) R_n that is required of an 'n-candidate itemset' to be an 'n-frequent itemset.' Let $|I^n|$ denote the cardinality of n-candidate itemset. In this paper, for the sake of simplicity, we define R_n as follows.

$$R_n = 1 \;/\; |I^n|$$

But it can be freely modified to fit the needs of the user.

Suppose $W(I_t^n)$ for a certain n-candidate itemset is 0.15 and R_n is 0.1, then this itemset becomes n-frequent itemset since it exceeds the threshold support level R_n.

WARM follows basically the same process as TSAA in association rule mining. The different features of WARM are as the following. Minimum aggregate weight (threshold) R_n is used instead of 'minimum relative support.' Aggregate weight $W(I_t^n)$ of n-candidate itemset is used instead of itemset frequency count. Those candidate itemsets which have $W(I_t^n) \geq R_n$ become n-frequent itemsets.

3.2 Sample application

An example of the WARM procedure may best explain the algorithm. A market basket database has ten instances of transactions out of seven products as shown in <Table 1>. And a survey produced the strategic value of each item as shown in <Table 2>. Calculating the normalized values of the strategic factors of each item according to the definitions of section 3.1 we obtain <Table 3>.

(Fig. 1) shows the frequency count of each 1-candidate itemset and the derivation of its normalized support. Suppose we use the relative weight of each strategic factor as follows.

$$\omega_s=0.4,\ \omega_\pi=0.3,\ \omega_\chi=0.2,\ \omega_\mu=0.1$$

Then we can calculate the aggregate weight of

〈Table 1〉 Transaction database

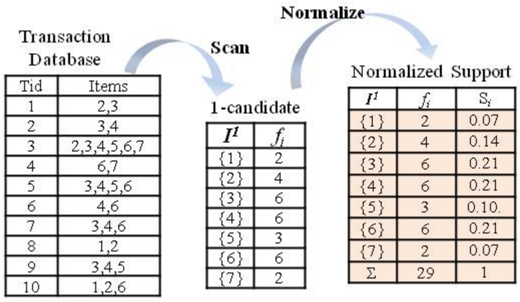
Tid	Items	Tid	Items
1	2, 3	6	4, 6
2	3, 4	7	3, 4, 6
3	2, 3, 4, 5, 6, 7	8	1, 2
4	6, 7	9	3, 4, 5
5	3, 4, 5, 6	10	1, 2, 6

〈Table 2〉 Strategic values

Item No	Profitability(P)	Customer satisfaction(C)	Marketing value(V)
1	10	50	50
2	20	70	30
3	30	60	100
4	40	30	60
5	50	90	20
6	60	60	50
7	70	80	70

〈Table 3〉 Normalized values of strategic factors

I^1	P($)	Π	C(%)	χ	V(%)	μ
1	10	0.04	50	0.11	50	0.13
2	20	0.07	70	0.16	30	0.08
3	30	0.11	60	0.14	100	0.26
4	40	0.14	30	0.07	60	0.16
5	50	0.18	90	0.20	20	0.05
6	60	0.21	60	0.14	50	0.13
7	70	0.25	80	0.18	70	0.19
Σ	280	1.00	440	1.00	380	1.00



(Fig. 1) 1-candidate itemset and normalized support

1-candidate itemsets using the formula $W(I_t^n) = \omega_\pi(I_t^n)+\omega_\mu(I_t^n)+\omega_\chi(I_t^n)+\omega_s(I_t^n)$ and the result is as shown in <Table 4>.

Since there are seven 1-candidate itemsets, $|I^1|=7$ and minimum aggregate weight is $R_1= 1/7$. Applying this cut-off value to <Table 6> we get 1-frequent itemset consisting of {3}, {4}, {6}, and {7}. The next step is to generate 2-candidate itemsets using these four itemsets and we obtain I^2={(3,4), (3,6), (3,7), (4,6), (4,7), (6,7)}.

Suppose we use the same relative weight as before, then we obtain the aggregate weight of 2-candidate itemset I^2 as in <Table 6>. Now we have six 2-candidate itemsets. So $|I^2|=6$, $R_2=1/6$, and applying this cut-off value, we obtain 2-frequent itemsets and 2-infrequent itemsets as in (Fig 2).

With the three 2-frequent itemsets we can generate 3-candidate itemset. In our example we have only one itemset (3, 4, 6). Since this is the only 3-candidate itemset, we don't need any further processing and accept this as the 3-frequent itemset after checking its support value. On the other hand, we apply the same procedure to the three

〈Table 4〉 Aggregate weight calculation of 1-candidate itemset I^1

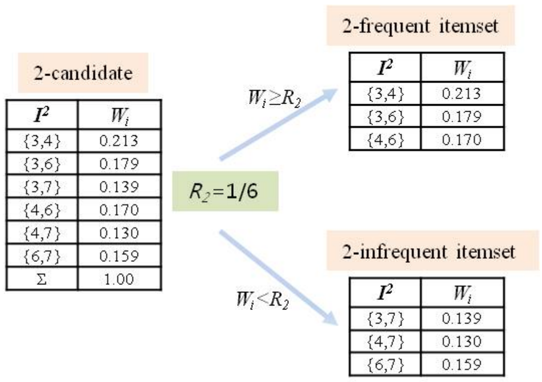
I^1	S_t	Π_t	χ_t	μ_t	W_t
	ω_s=0.4	ω_π=0.3	ω_χ=0.2	ω_μ=0.1	
(1)	0.07	0.04	0.11	0.13	0.075
(2)	0.14	0.07	0.16	0.08	0.117
(3)	0.21	0.11	0.14	0.26	0.171
(4)	0.21	0.14	0.07	0.16	0.156
(5)	0.10	0.18	0.20	0.05	0.139
(6)	0.21	0.21	0.14	0.13	0.188
(7)	0.07	0.25	0.18	0.19	0.158
Σ	1.00	1.00	1.00	1.00	1.00

〈Table 5〉 Norpmalized values of strategic factors for I^2

I^2	P($)	Π	C(%)	χ	V(%)	μ
(3,4)	70	0.12	90	0.13	160	0.19
(3,6)	90	0.15	120	0.18	150	0.18
(3,7)	100	0.17	140	0.13	170	0.20
(4,6)	100	0.17	90	0.13	110	0.13
(4,7)	110	0.18	110	0.16	130	0.16
(6,7)	130	0.21	140	0.20	120	0.14
Σ	600	1.00	690	1.00	840	1.00

〈Table 6〉 Aggregate weight calculation of 2-candidate itemset I^2

I^2	S_t	Π_t	χ_t	μ_t	W_t
	ω_S=0.4	ω_π=0.3	ω_χ=0.2	ω_μ=0.1	
(3,4)	0.33	0.12	0.13	0.19	0.213
(3,6)	0.20	0.15	0.18	0.18	0.179
(3,7)	0.07	0.17	0.20	0.20	0.139
(4,6)	0.20	0.17	0.13	0.13	0.170
(4,7)	0.07	0.18	0.16	0.16	0.130
(6,7)	0.13	0.21	0.20	0.14	0.159
Σ	1.00	1.00	1.00	1.00	1.00



(Fig. 2) 2-candidate itemset divided into two groups

2-infrequent itemsets and generate 3-candidate itemsets. <Table 7> shows the list of 3-candidate itemsets and their aggregate weight derivations.

Here $|I^3|$=3 and R_3=1/3, so we can accept (3,4,7) and (3,6,7) as 3-frequent itemsets. We observe that these last two itemsets might have been thrown away with other algorithms.

Using the same transaction database of <Table 1>, we tried to compare the outcomes of three different algorithms, i.e., Apriori, TSAA, and WARM. <Table 8> shows that each algorithm produces distinct outcomes.

In order to verify the validity of the algorithm, we performed a second data set test using a real world database which was collected from a local large-scale discount store (E-Mart). It contained 730 transaction records of 20 items. A data cleaning and transformation was performed in the data preparation stage. <Table 9> is the list of items described in terms of their categories. The algorithm was implemented using Microsoft Excel and the result showed distinct and characteristic behavior in association rule mining as compared to Apriori and TSAA.

<Table 10> is a summary of the comparative analysis of three algorithm performances. The distinct itemsets of the table indicate that WARM consistently discovers more frequent itemsets than the other algorithms although the result may vary depending on how we set the minimum support and minimum relative support. Both WARM and TSAA produce 11 additional 3-frequent itemsets. However, there was no identical itemsets (overlapping) produced by the two algorithms. This is a good indication that the data mining behaviors of the two algorithms are quite different and unique.

The advantages of WARM algorithm are as follows. We don't need to set minimum support or minimum relative support level. Various strategic factors may be considered depending on the user's needs and problem situation. The computational procedure is so simple that it can be easily implemented into any kind of spread sheet program.

〈Table 7〉 Aggregate weight calculation of 3-candidate itemset I^3

I^3	S_t	Π_t	χ_t	μ_t	W_t
	ω_S=0.4	ω_π=0.3	ω_χ=0.2	ω_μ=0.1	
(3,4,7)	0.56	0.30	0.31	0.36	0.41
(3,6,7)	0.33	0.34	0.37	0.35	0.34
(4,6,7)	0.11	0.36	0.31	0.29	0.25
Σ	1.00	1.00	1.00	1.00	1.00

〈Table 8〉 Comparison of three algorithm outcomes

In	WARM	TSAA	Apriori
	Weights 0.4, 0.3, 0.2, 0.1	MinSup 40 MinRSup 0.5	MinSup 40
1-frequent itemset	(3), (4), (6), (7)	(2), (3), (4), (6)	(2), (3), (4), (6)
2-frequent itemset	(2,4), (3,6), (4,6)	(2,6), (3,6)	(3,4), (4,6)
3-frequent itemset	(3,4,6)	(3,4,6)	(3,4,6)
Additional 3-itemset	(3,4,7), (3,6,7)	(2,3,6)	

〈Table 9〉 List of items

Item ID	Description	Item ID	Description
1	Baby food	11	Sauce
2	Bread	12	Liquor
3	Cooked meal	13	Stock farm product
4	Noodle	14	Fish
5	Frozen food	15	Vegetable
6	Milk & beverage	16	Dried fish
7	Coffee & tea	17	Gimchie
8	Confectionery	18	Egg
9	Canned food	19	Rice cake
10	Ham & fish cake	20	Fruit

〈Table 10〉 Comparison of results

		WARM	TSAA	Apriroi
No. of itemsets	1-frequent	9	11	11
	2-frequent	13	4	4
	3-frequent	10	5	5
	4-frequent	9	5	5
	Additional 3-itemset	11	11	-
Common itemsets	1-frequent	{4} {6} {8} {15} {16} {20}		
	2-frequent	{4,6} {6,8} {6,15} {6,20}		
	3-frequent	{4,6,8} {4,6,15} {4,6,20} {6,8,15} {6,15,20}		
	4-frequent	{4,6,8,15} {4,6,8,20} {4,6,15,20} {6,8,15,20}		
	Additional 3-itemset	-		

〈Table 10〉의 계속

		WARM	TSAA	Apriroi
Distinct itemsets	1-frequent	{7} {13} {14} {15}	{2} {10} {11} {12} {18}	
	2-frequent	{4,8} {4,13} {4,15} {4,20} {6,13} {8,13} {13,15}{13,20} {15,20}	-	-
	3-frequent	{4,6,13} {4,15,20} {6,8,20} {6,13,15} {6,13,20}	-	-
	4-frequent	{4,6,13,15} {4,6,13,20} {4,13,15,20} {6,8,13,15} {6,13,15,20}	{4,8,15,20}	
	Additional 3-itemset	All	All	-

4. Conclusion

WARM is a novel approach to association rule mining since it considers semantic and strategic features in addition to the frequency of occurrence. It can be applied to marketing campaigns, sales promotion, electronic commerce, customer relationship management, and many more.

WARM considers more strategic factors than other algorithms for each and every item. The sample strategic factors used in this paper are just a few of the possible choices. In realistic application, users can pick as many factors as they want in their algorithm construction. Normalized weight adjustment in WARM is especially important because it works as a systematic mechanism to reduce the overriding effect of one strategic factor over another. For example, if item A gives $10 profit and item B $10,000, then item B will have 1000 times more weight than item A. And this effect will be even more aggravated if customer satisfaction can take values only in the range between 1 and 10. Since WARM uses normalization within factor and relative weight between factors, it provides a double-folded subjective impact reduction mechanism.

WARM consistently produced more frequent itemsets than other algorithms. More careful experiment and analysis is necessary to find out whether the mining result contains too many unimportant itemsets and to find out how we can reduce meaningless outcomes. The computational complexity issue, the reduction of database access time, and other computational efficiency related issues were not of primary concern in this research. Instead, it focused on the semantic and strategic aspects of association rule mining.

References

[1] P. Adrians, D. Zantige, *Data Mining*, Addison Wesley, 1996

[2] R.Agrawal, T.Imielinski, A.Swami, “Mining Association Rules between Sets of Items in Large Database,” Proceedings of ACM SIGMOD, Vol.22 No.2 207, 1993

[3] R. Agrawal, R. Srikant, “Fast Algorithms for Mining Association Rules,” Proceedings of VLDB conference, Vol.20, 487, 1994

[4] Y. Bastide, G. Stumme, L. Lakhal, “Generating a Condensed Representation for Association Rules,” Journal of intelligent information systems, Vol.24 No.1 26-60, 2005

[5] Dan S. Ha, Bu H. Hwang, “Association rule mining technique for scarce and semantic items using relative support,” Journal of the Korea information science society, Vol.28 577-586, 2001

[6] Bum S. Ha, *Association rule mining technique for scarce and semantic data items*, master’s thesis, Chunnam University, 2001

[7] J. Han, M. Kamber, *Data Mining, Concepts and Techniques*, Morgan Kaufmann, 2006

[8] B. Liu, W. Hsu, Y. Ma, Mining Association Rules with Multiple Minimum Supports,” Proceedings of the ACM SIGKDD(KDD-99) 337-341, 1999

[9] B. Liu, Y. Ma, C.K. Wong, P. S. Y u, “Scoring the data using association rules, Applied intelligence,” Vol.18 No.2, 2003

[10] Jong Sun Park, Jean H. Yi, “A study on sequential pattern mining algorithm,” Statistics Research, Vol.11 56-73, 2003

[11] Jong S. Park, Won K. Yu, Kee H. Hong, “Association rule mining and its applications,” Journal of the Korea information science society, Vol.16, 1998

[12] Im Y. Song, *Image data mining with frequent itemsets and scarce and semantic items*, master’s thesis, Seoul Industrial University, 2004

[13] C. H. Cai, W. C. Ada, Cheng, C. H. and Kwong, W. W., “Mining Association Rules with Weighted Items,” Proc. of 1998 Intl. Database Engineering and Applications Symposium (IDEAS'98), Cardiff, Wales, U.K., pp.68-77, 1998.

[14] S. Yue, E. Tsang, D. Yeung, D. Shi, “Mining Fuzzy Association Rules with Weighted Items,” IEEE International Conference on Systems, Man and Cybernetics, v.3, pp.1906-1911, 2000.

[15] D. W. Choi, Y.J. Hyun, “Transitive Association Rule Discovery by Considering Strategic Importance,” The 10th IEEE International Conference on Computer and Information Technology (CIT-2010), pp.1654-1659, 2010.

최 덕 원

e-mail : dougch01@paran.com

1973년 서울대학교 공과대학 산업공학과(학사)

1975년 한국과학원 산업공학과(공학석사)

1994년 Dep't of Computer and Information Science, Temple University(공학박사)

1978년~1986년 성균관대학교 산업공학과 조교수

1994년 ~ 현 재 성균관대학교 시스템경영공학과 교수

관심분야 : 정보시스템경영, 데이터마이닝, BPM, 지식경영, 지능정보시스템