This full-text paper was peer-reviewed and accepted to be presented at the IEEE WiSPNET 2017 conference.

Ranking of Association Rules Toward Smart Decision for Smart City

Subrata Datta,[1] Kalyani Mali[2] and Pratyay Roy[3]
[1]Dept. of Information Technology, NITMAS, Kolkata, India
[2]Dept. of Computer Science & Engg., Kalyani University, Kalyani, India
[3]Dept. of Computer Science & Engg., NITMAS, Kolkata, India
[1]Subrataju2008@gmail.com [2]Kalyanimali1992@gmail.com [3]pratyayroy@outlook.com

***Abstract*—Ranking of association rules based on their significance in knowledge discovery has become an important issue in data mining. Traditional mining algorithms may often generate a huge number of rules including less or non-significant ones. Users expect is to deal with the most significant association rules to get a smart decision. Ranking express the level of significance which may reduce the confusion in decision making. This paper introduces *Gravity* as a measure of rule significance and henceforth ranks the association rules. Experimental analysis shows the effectiveness of the proposed technique.**

***Index Terms*—Association rule, gravity, trust, dissociation, ranking, smart city.**

I. Introduction

Application of association rule includes the domain of market basket analysis, bioinformatics, learning system, intrusion detection, crowdsourcing etc. The latest addition to this list is smart city. The concept of a smart city is associated with the proper Information and Communication Technology (ICT) enhanced governance [1]. Voluminous data may be generated from the components of a smart city like mobility, environment, healthcare, transportation, education, energy etc. To facilitate such services knowledge discovery is needed. Association rule mining is one of the key players in knowledge discovery for making smart decision.

Existence of an association rule may depend on user defined threshold values where users have no prior idea about what should be the appropriate threshold values. So there is a chance to generate huge number of association rules. Dealing with a huge number of rules is the primary problem in decision making. How much an association rule is trustworthy in decision making is also a matter of discussion. Evolution of association rule with high dissociation is a crucial problem in data mining. Conceptually association rule with high dissociation makes no sense as because of dissociation is opposite of association. So the measurement of significance of association rule has become an important task in data mining. This paper introduces *gravity* as a measure of rule significance where trustworthiness, accuracy, sensitivity, reliability, novelty and certainty have been considered. Ranking is necessary for indication of the level of significance of association rules.

The rest of the paper is organized as follows. Section II refers to the background. Dissociation and trustworthiness are discussed in Section III. Components of rule significance are discussed in Section IV. Rule ranking procedure is discussed in Section V. Experimental results are presented in Section VI. Finally Section VII concludes the study.

II. Background

A. Association Rule

Formal statement of the problem corresponding to frequent pattern mining and association rule mining [2] is as follows: Let $I = \{i_1, i_2, i_3, \ldots, i_m\}$ be a set of items and database $DB = \{T_1, T_2, T_3, \ldots, T_n\}$ be a set of transactions where each transaction T is a set of items such that $T \subseteq I$. *Support(s)* is an objective measure used for frequent pattern mining which indicates the frequency of occurrence of an item or a set of items in the database *DB*. Mathematically *support(s)* of an itemset X may be defined as follows.

$$s(X) = \frac{\text{Frequency of occurrence of } X \text{ in } DB}{\text{Total number of transactions in } DB}. \quad (1)$$

Definition 1 (Frequent Itemset)**.** An itemset having support $s \geq minsup$ is considered as frequent itemset where *minsup* is known as user defined threshold parameter.

Association rule is expressed in the form of $X \rightarrow Y$ where itemsets $X, Y \subset I$ and $X \cap Y = \emptyset$. *Confidence(c)* is an objective measure used for association rule mining from frequent patterns which indicates the conditional probability of Y given that X has occurred. Mathematically *confidence (c)* of $X \rightarrow Y$ may be defined as follows.

$$c\,(X \rightarrow Y) = P\left(\frac{Y}{X}\right) = \frac{s(XY)}{s(X)}. \quad (2)$$

Definition 2 (Association Rule)**.** A rule having confidence $c \geq minconf$ is considered as an association rule where *minconf* is known as user defined threshold parameter.

B. Ranking of Association Rule

Ranking of association rules have been performed under various observations by different scholars. One of the pioneer methods [3] works under multiple criteria, but no focus on dissociation. An allowable amount of dissociation has been considered in *Rank Index* [4]. Few earlier works [5], [6] are remarkable ones. *RANWAR* [7] is a popular technique for ranking but it is more suitable in bioinformatics. Top-k association rule mining algorithm [8] mines top ranked rules,

978-1-5090-4442-9/17/$31.00 ©2017 IEEE

TABLE I
A 2×2 CONTINGENCY TABLE.

	Y	$\bar{Y}$	$\sum$
X	f_{11}	f_{10}	f_{1+}
$\bar{X}$	f_{01}	f_{00}	f_{0+}
$\sum$	f_{+1}	f_{+0}	

but limited to the user defined number k. Measurement and analysis of rule significance [9] has been stated by different scholars based on their observations. Measures for trustworthy association rule [10] are also a relevant research work.

C. Our Contribution

Key contributions of this work are as follows.

(1) Introduction of gravity as the measure of rule significance.
(2) Ranking of association rules based on their significance.
(3) Higher significant rules possess lower dissociation.
(4) Higher significant rules are more trustworthy.

III. Dissociation and Trustworthiness

A. Dissociation

Dissociation [11] holds the opposite meaning of association. Two patterns having equal support may possess different dissociation. Association rule with low dissociation is desirable. A typical 2×2 contingency table for binary variables X and Y is shown in Table I. The cells of the table refer to the frequency count denoted by f_{ij} where $i, j \in 0, 1$. Presence and absence of an item is indicated by 1 and 0 respectively. The symbol '$\sum$' refers to row/column summation.

Definition 3 (Dissociation)**.** The probability of complementary appearance for any two itemset in a single transaction is called dissociation between them.

Refer to Table II, for any two itemset say X and Y,

$$\text{Dissociation}\,(d) = P\left(X\bar{Y} + \bar{X}Y\right). \tag{3}$$

Example 1. Tables II and III shows an artificial grocery dataset [12] and some of the extracted association rules respectively. Though the generated rules have 100% confidence but also possess high dissociation (60%/80%) which is a contradictory matter. This scenario clearly depicts the limitation of traditional mining algorithm regarding significant rule evolution.

B. Trustworthiness

Trustworthiness is the measurement of acceptability of an association rule to a user regarding synthesis of decision. It has been observed in many cases that the association rules with high confidence also contain high dissociation which is a contradictory matter. The scenario raises the problem regarding viability of the rule. To solve the problem we have introduced trustworthiness. Objective measure *trust* [13] is the backbone of trustworthiness. Trustworthiness (tw) is a precision like measurement and it shows the deviation of trust with respect to dissociation.

TABLE II
AN ARTIFICIAL GROCERY DATASET.

Tid	Items
1	WHB, CC, WH
2	WHB, CC, WH
3	WB, CC, WH
4	WB, CC, WH
5	W, CC, WH

TABLE III
RULES WITH HIGH DISSOCIATION.

Sl. No.	Rule descrption	Confidence (%)	Dissociation (%)
1	WHB→CC	100	60
2	WB→CC	100	60
3	WB→WH	100	60
4	W→CC	100	80
5	W→WH	100	80

Definition 4 (Trust)**.** The probability of occurrence of any two itemsets through the maximum acceptability of association via directs (co-occurrence) or indirect (association against dissociation) way with respect to the maximum probability between antecedent and consequent accompanied by the probability of null transactions.

Refer to Table I, for any two itemsets, say X and Y,

$$\text{Trust}(t) = \frac{\max[P\left(XY\right), 1 - \left\{P\left(X\bar{Y}\right) + P\left(\bar{X}Y\right)\right\}]}{\left[\max\left\{P\left(X\right), P\left(Y\right)\right\} + P\left(\bar{X}\bar{Y}\right)\right]}. \tag{4}$$

Range of trust is [0, 1]. Trust is a symmetric objective measure that accounts dissociation and null transactions.

Definition 5 (Trustworthiness)**.** Trustworthiness is the deviation of trust with respect to dissociation.

For any two itemsets say X and Y, measurement of trustworthiness is as follows.

$$\begin{aligned}\text{Trustworthiness}\,(tw) &= \frac{\text{Trust}}{\text{Trust} + \text{Dissociation}}\\ \implies \qquad tw &= \frac{t}{t+d}.\end{aligned} \tag{5}$$

Lemma 1. Range of trustworthiness (tw) is [0, 1].

Proof. Range of trust $(t) = [0, 1]$.
⇒ If $d = 0, tw = 1$.
⇒ When $d \neq 0$, then $0 < tw < 1$, as $t + d \geq t$.
⇒ When $t = 0$, then $tw = 0$. □

IV. Components of Rule Significance

The generalization form of a contingency table (ex. Table I) is known as *confusion matrix* where f_{11}, f_{10}, f_{01} and f_{00} indicate true positive (TP), false negative (FN), false positive (FP) and true negative (TN) respectively. The major components of rule significance measure are discussed below.

Definition 6 (Accuracy)**.** Accuracy of a rule $R = X \to Y$ is known as precision in information retrieval.

$$\begin{aligned} Acc\,(R) &= \frac{TP}{TP+FP} = \frac{f_{11}}{f_{11}+f_{01}} \\ &= \frac{f_{11}}{f_{+1}} = \frac{\frac{f_{11}}{N}}{\frac{f_{+1}}{N}} = \frac{P\,(f_{11})}{P\,(f_{+1})} \\ &= P\,(Y|X)\,. \end{aligned} \tag{6}$$

Definition 7 (Reliability)**.** Reliability of a rule $R = X \to Y$ indicates the prediction of positive cases.

$$\begin{aligned} NegRel\,(R) &= \frac{TN}{TN+FN} = \frac{f_{00}}{f_{00}+f_{10}} \\ &= \frac{f_{00}}{f_{+0}} = \frac{\frac{f_{00}}{N}}{\frac{f_{+0}}{N}} = \frac{P\,(f_{00})}{P\,(f_{+0})} \\ &= P(\bar{Y}|\bar{X}). \end{aligned} \tag{7}$$

Definition 8 (Sensitivity)**.** Sensitivity of a rule $R = X \to Y$ is equivalent to recall in information retrieval.

$$\begin{aligned} Sens\,(R) &= \frac{TP}{TP+FN} = \frac{f_{11}}{f_{11}+f_{10}} \\ &= \frac{f_{11}}{f_{1+}} = \frac{\frac{f_{11}}{N}}{\frac{f_{1+}}{N}} = \frac{P(f_{11})}{P(f_{1+})} \\ &= P(X|Y). \end{aligned} \tag{8}$$

Definition 9 (Specificity)**.** Specificity of a rule $R = X \to Y$ is defined as the conditional probability that X is false given that Y is false. It is equivalent to the recall of negative cases in information retrieval.

$$\begin{aligned} Spec\,(R) &= \frac{TN}{TN+FP} = \frac{f_{00}}{f_{00}+f_{01}} \\ &= \frac{f_{00}}{f_{0+}} = \frac{\frac{f_{00}}{N}}{\frac{f_{0+}}{N}} = \frac{P(f_{00})}{P(f_{0+})} \\ &= P(\bar{X}|\bar{Y}). \end{aligned} \tag{9}$$

Definition 10 (Novelty)**.** Novelty refers the closeness between antecedent and consequent part of a rule $R = X \to Y$.

$$Nov\,(R) = P\,(f_{11}) - P\,(f_{1+}) * P(f_{+1}). \tag{10}$$

Novelty is relative to accuracy, negative reliability, sensitivity and specificity.

Definition 11 (Relative Accuracy)**.**

$$RAcc\,(X \to Y) = P\,(Y|X) - P(Y).$$

Definition 12 (Relative Negative Reliability)**.**

$$RNegRel\,(X \to Y) = P\left(\bar{Y}|\bar{X}\right) - P(\bar{Y}).$$

Definition 13 (Relative Sensitivity)**.**

$$RSens\,(X \to Y) = P\left(\bar{X}|\bar{Y}\right) - P(X).$$

Definition 14 (Relative Specificity)**.**

$$RSpec\,(X \to Y) = P\left(\bar{X}|\bar{Y}\right) - P(\bar{X}).$$

Weighted Relative Accuracy (WRAcc) [14] is a compact measure for relative concept of accuracy, negative reliability, sensitivity and specificity.

Definition 15 (WRAcc)**.** WRAcc of a rule $R = X \to Y$ is defined as the gain factor of an association rule which trades off between generality and relative accuracy.

$$WRAcc\,(R) = P\,(X) * (P\,(Y \mid X) - P\,(Y)). \tag{11}$$

Lemma 2. $WRAcc\,(R) = Nov(R)$.

Proof.

$$\begin{aligned} WRAcc(R : X \to Y) &= P\,(X) * (P\,(Y|X) - P\,(Y)) \\ &= P\,(X) * P\,(Y|X) - P\,(X) * P(Y) \\ &= P\,(XY) - P\,(X) * P(Y) \\ &= Nov(R). \end{aligned}$$ □

Lemma 3. $WRAcc(R) = WRSens(R)$.

Proof.

$$\begin{aligned} WRAcc\,(R : X \to Y) &= P\,(X) * (P\,(Y|X) - P\,(Y)) \\ &= P\,(X) * P\,(Y|X) - P\,(X) * P(Y) \\ &= P\,(XY) - P\,(X) * P\,(Y) \\ &= P\,(Y) * (P\,(X|Y) - P\,(X)) \\ &= WRSens(R). \end{aligned}$$ □

Lemma 4. $WRAcc\,(R) = WRSpec(R)$.

Proof.

$$\begin{aligned} WRAcc\,(R : X \to Y) &= P\,(X) * (P\,(Y|X) - P\,(Y)) \\ &= P\,(X) * P\,(Y|X) - P\,(X) * P(Y) \\ &= P\,(XY) - P\,(X) * P\,(Y) \\ &= \left(1 - P\left(\bar{Y}X\right) - P\left(Y\bar{X}\right) - P\left(\bar{X}\bar{Y}\right)\right) - \left(1 - P\left(\bar{X}\right)\right) \\ &\quad \left(1 - P\left(\bar{Y}\right)\right) \\ &= \left(1 - P\left(\bar{Y}\right) - P\left(\bar{X}\right) + P\left(\bar{X}\bar{Y}\right)\right) - (1 - P\left(\bar{Y}\right) \\ &\quad - P\left(\bar{X}\right) + P\left(\bar{Y}\right) * P\left(\bar{X}\right)) \\ &= P\left(\bar{X}\bar{Y}\right) - P\left(\bar{Y}\right) * P(\bar{X}) \\ &= P\left(\bar{Y}\right)\left(P\left(\bar{X}|\bar{Y}\right) - P\left(\bar{X}\right)\right) \\ &= WRSpec(R). \end{aligned}$$ □

Lemma 5. $WRAcc\,(R : X \to Y) = WRNegRel(R)$.

Proof.

$$\begin{aligned} WRAcc\,(R : X \to Y) &= WRSpec(R) \\ &= P\left(\bar{Y}\right)\left(P\left(\bar{X}|\bar{Y}\right) - P(\bar{X})\right) \\ &= P\left(\bar{X}\bar{Y}\right) - P\left(\bar{X}\right) P(\bar{Y}) \\ &= P\left(\bar{X}\right)\left(P\left(\bar{Y}|\bar{X}\right) - P\left(\bar{Y}\right)\right) \\ &= WRNegRel(R). \end{aligned}$$ □

Definition 16 (Certainty Factor)**.** Certainty factor (CF) is an objective measure used to judge the accuracy, truthfulness and reliability of a prediction.

This full-text paper was peer-reviewed and accepted to be presented at the IEEE WiSPNET 2017 conference.

Certainty factor [15] is neither a probability nor a truth value.

$$CF(X \rightarrow Y) = \max\left(\frac{P(Y|X) - P(Y)}{1 - P(Y)}, \frac{P(Y|X) - P(X)}{1 - P(X)}\right) \quad (12)$$

V. Rule Ranking Procedure

A. Enumeration of Gravity

Gravity expresses the amount of significance of an association rule.

Definition 17 (Gravity)**.** Gravity is a compact measure under trustworthiness, accuracy, sensitivity, specificity, novelty, certainty and dissociation used for the measurement of rule significance.

The mathematical structure of gravity (G) for a rule ($R : X \rightarrow Y$), is given below.

$$\text{Gravity } (R) = tw + WRAcc + CF. \quad (13)$$

Gravity provides numerical value corresponding to rule significance. Thus association rules with higher gravity are the most desirable for making smart decision in smart city [16].

Lemma 6. Range of gravity is $[-\infty, +\infty]$.

Proof. Range of $tw = [0, 1]$.
Range of WRAcc=$[-\infty, +\infty]$.
Range of $CF = [-1, 1]$.
So Range of gravity is $[-\infty, +\infty]$. □

B. Gravity Based Ranking

Higher the gravity betters the rank. If rules have equal gravity then tie breaking [17] is done.

Tie Breaking: Let R_i and R_j are two rules having equal gravity then R_i precedes R_j if

1. Confidence of R_i is larger than that of R_j.
2. Confidence of R_i and R_j are the identical, but the support of R_i is larger than that of R_j.
3. The confidence and support of R_i and R_j are the identical, but R_i contains less number of attributes in its antecedent than that of R_j.

The proposed algorithm of ranking is furnished below.

Algorithmic View

Input: Transactional database DB, minsupp, minconf.
Output: Association rule with rank.

Scan DB and find set of all frequent 1-itemset L_1 applying minsupp.
$L = L_1$ $/* L =$ Set of all frequent itemset */
For $(k = 2; L_{k-1} \neq 0; k++)$ {
$C_k \leftarrow L_{k-1} \bowtie L_1 /* C_k$ is candidate set */
$\forall i \in C_k$ { /* i is any itemset */
$s \leftarrow support(DB, i)$
if $s \geq minsupp$ then
$L_k \leftarrow L_k \cup i$ /* itemset i is added to L_k */
end if
}
$L = L \cup L_k$
}
$AR \leftarrow \emptyset$ /*AR= Association rule*/
For each i in L {
$\forall X, Y \quad (i = X \cup Y)$ {
$c \leftarrow confidence(X, Y)$
if $c \geq minconf$ then
$AR \leftarrow AR \cup \{X \rightarrow Y\}$
$AR.\ Count++$
$t \leftarrow trust(X \rightarrow Y)$
$d \leftarrow dissociation(X \rightarrow Y)$
$tw \leftarrow trustworthiness(X \rightarrow Y)$
$WRAcc \leftarrow weighted\_relative\_accuracy (X \rightarrow Y)$
$CF \leftarrow certainty\_factor(X \rightarrow Y)$
G$\leftarrow gravity(X \rightarrow Y)$
l$\leftarrow antecedent_length(X \rightarrow Y)$
end if
}
}
For (j=1; j<= AR. Count; J++) {
$Rank \leftarrow sort\ (AR)\ in\ G,\ c,\ s,\ \frac{1}{l}$
Return AR
}

VI. Experimental Results

We have conducted our experiments on Intel Core 2 Duo processor with 4 GB RAM under the platform of Windows 7 and coding has been done using Python 3.5. Experiments have been done on different datasets shown in Table IV. Association rules extracted from Chess[1] dataset are ranked by gravity and shown in Table V. The scenario clearly depicts that the rules having high rank also have high trustworthiness (tw) but low dissociation (d) which proves the effectiveness of gravity. Table VI shows some of the rules extracted from Extended Bakery[2] dataset of size 20 K and Grocery[3] dataset which are ranked by gravity. Fig. 1 shows that the top ranked rules have low dissociation which refers to high association. This indicates that, if the rule has strong association then it will be awarded with high rank. A comparative study between ranking by gravity and ranking by confidence based on the rule

[1] http://fimi.ua.ac.be/data/chess.dat.
[2] https://wiki.csc.calpoly.edu/datasets/wiki/apriori.
[3] https://github.com/stedy/Machine-Learning-with-R-datasets/blob/master/groceries.csv.

TABLE IV
DATASETS.

Dataset	# transactions	# item
Extended bakery	1,000	50
Extended bakery	5,000	50
Extended bakery	20,000	50
Grocery	9,835	169
Chess	3,196	36

TABLE V
HIGHER GRAVITY, CERTAINTY FACTOR, TRUSTWORTHINESS & LOWER DISSOCIATION REFERS TO HIGHER RANK.

Chess dataset
minsupp = 0.92, minconf = 0.97

Rank	Rule	G	CF	tw	d
1	{66}→{60}	1.9738	1	0.9599	0.04005
2	{58,66}→{60}	1.9735	1	0.9596	0.04036
3	{29,66}→{60}	1.9713	1	0.9574	0.04255
4	{29,58,66}→{60}	1.9709	1	0.9571	0.04286
5	{52,66}→{60}	1.9703	1	0.9565	0.04349

TABLE VI
ILLUSTRATION OF FIVE TOP RANKED RULES USING GRAVITY.

Rank	Extended bakery (20 K) minsupp = 0.01 minconf = 0.25	Grocery dataset minsupp = 0.01 minconf = 0.25
1	{23,41,43}→{24,40}	{citrus fruit, root vegetables}→{other vegetables}
2	{24,40}→{23,41,43}	{tropical fruit, root vegetables}→{other vegetables}
3	{23,24,43}→{40,41}	{tropical fruit, curd}→{yogurt}
4	{40,41}→{23,24,43}	{tropical fruit, whole milk, root vegetables}→{other vegetables}
5	{24,40,43}→{23,41}	{butter, whipped/sour cream}→{whole milk}

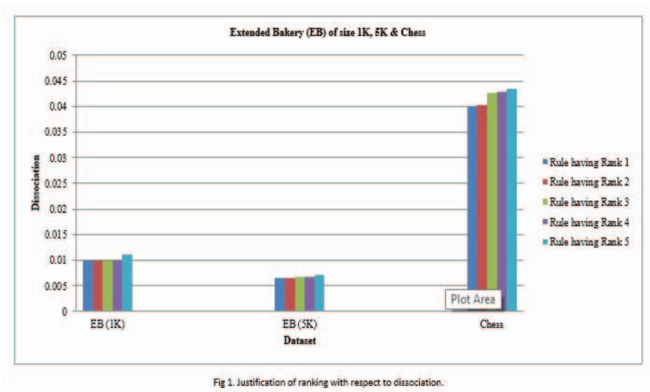

Fig. 1. Justification of ranking with respect to dissociation.

quality measures such as average certainty factor (CF), average trustworthiness (tw), average length (l) and max length (l) is furnished in Fig. 2. Length (l) of a rule refers to the number of items associated with it. Thus average length and max length refers to average number of items and maximum number of items in a rule respectively. Presence of average dissociation in top ranking rules under gravity and confidence is furnished in Fig. 3.

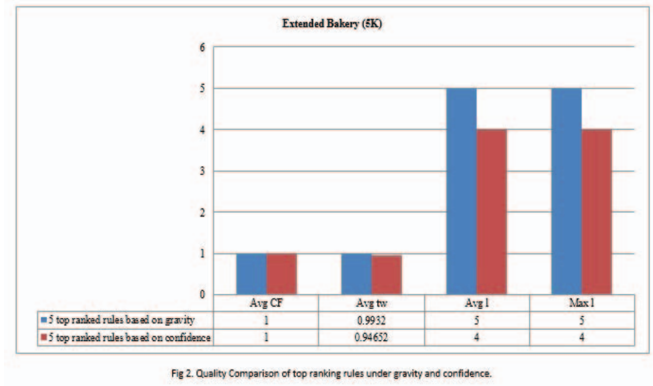

Fig. 2. Quality comparison of top ranking rules under gravity and confidence.

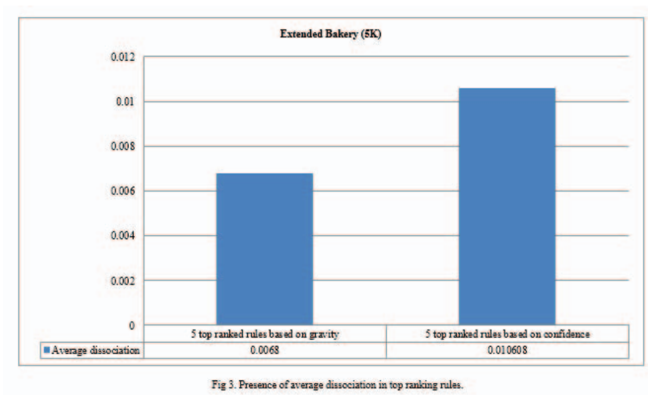

Fig. 3. Presence of average dissociation in top ranking rules.

VII. CONCLUSION

Smart governance in a smart city requires smart decision in business intelligence. Retrieval of proper information from the large transactional databases in a smart city may help in getting smart decision. Significant association rules include proper information. Ranking expresses the level of significance of association rules based on their gravity. Gravity is a measure used for the computation of rule significance. Gravity supports association devoid of dissociation. An association rule may be lucrative to a citizen in the smart city if it is trustworthy and informative. Ranking using gravity just declares the most probable lucrative association rules. Since the main purpose of the paper is to establish a new idea for rule ranking, we have not concentrate on the efficiency of the proposed algorithm. Future efforts should try to develop computationally more efficient algorithm for ranking of association rules based on gravity.

Acknowledgements: This work was supported by the Kalyani University DST PURSC (No: SR/59/2-23/2013(a) dated 29.10.2014).

REFERENCES

[1] Z. Khan, A. Anjum, and S. L. Kiani, "Cloud based big data analytics for smart future cities," in *IEEE/ACM 6th International Conference on Utility and Cloud Computing*, 2013, pp. 381–386.

[2] R. Agarwal, T. Imielinski, and A. Swami, "Mining association rules between sets of items in large databases," in *ACM SIGMOD'93*, Washington DC, 1993, pp. 207–216.

[3] M. C. Chen, "Ranking discovered rules from data mining with multiple criteria by data envelopment analysis," *Expert Systems with Applications*, 2007, vol. 33, no. 4, pp. 1110–1116.

[4] S. Datta and S. Bose, "Mining and ranking association rules in support, confidence, correlation and dissociation framework," in *Proc. of FICTA, AISC*, Springer, NIT Durgapur, 2015, vol. 404, pp. 141–152.

[5] M. Toloo, B. Sohrabi, and S. Nalchigar, "A new method for ranking discovered rules from data mining by DEA," *Expert Systems with Applications*, 2009, vol. 36, no. 4, pp. 8503–8508.

[6] S. Bouker, "Ranking and selecting association rules based on dominance relationship," in *IEEE 24th International Conference on Tools with Artificial Intelligence*, 2012, vol. 1, pp. 658–665.

[7] S. Mallik, A. Mukhopadhyay, and U. Maulik, "RANWAR: rank-based weighted association rule mining from gene expression and methylation data," *IEEE Transactions on Nanobioscience*, 2015, vol. 14, no. 1.

[8] P. Fournier-Viger, C. W. Wu, and V.S. Tseng, "Mining top-k association rules," in *Canadian Conference on Artificial Intelligence*, Springer Berlin Heidelberg, 2012, pp. 61–73.

[9] I. N. M Shaharanee, F. Hadzic, and T. S. Dillon, "Interestingness measures for association rules based on statistical validity," *Knowledge-Based Systems*, 2011, vol. 24, no. 3, pp. 386–392.

[10] C. Weiß, "Statistical mining of interesting association rules," *Statistics and Computing*, 2008, vol 18, no. 2, pp. 185–194.

[11] S. Datta and S. Bose, "Discovering association rules partially devoid of dissociation by weighted confidence," in *2nd IEEE International Conference on Recent Trends in Information Systems*, Jadavpur University, Kolkata, 2015, pp. 138–143.

[12] J. R. Nayak and D.J. Cook, "Approximate association rule mining," in *Proc. of Floria Artificial Intelligence Research Symposium*, 2001.

[13] S. Datta and K. Mali, "Trust: a new objective measure for symmetric association rule mining in account of dissociation and null transaction," in *IEEE International Conference on Advanced Computing*, Madras Institute of Technology, Chennai, 2017, in press.

[14] N. Lavrac, P. Flach, and B. Zupan, "Rule evaluation measures: A unifying view," in *International Conference on Inductive Logic Programming*, Springer Berlin Heidelberg, 1999, pp. 174–185.

[15] D. H. Glass, "Confirmation measures of association rule interestingness," *Knowledge-Based Systems*, 2013, vol. 44, pp. 65–77.

[16] E. A. Nuaimi, H. A. Neyadi, N. Mohamed, and J. Al-Jaroodi, "Applications of big data to smart cities," *Journal of Internet Services and Applications*, 2015, vol. 6, no. 1, pp. 1–15.

[17] S. Wedyan, "Review and comparison of associative classification data mining approaches," *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 2014, vol. 8, no. 1, pp. 34–45.