*International Journal of Innovative Computing, Information and Control*
Volume **9**, Number **4**, April **2013**

ICIC International ©2013 ISSN 1349-4198
pp. 1679–1690

# ASSOCIATION RULE MINING BASED ON TOPOLOGY FOR ATTRIBUTES OF MULTI-VALUED INFORMATION SYSTEMS

Yan Zhang[1], Zheng Pei[1] and Peng Shi[2,3]

[1]School of Mathematics and Computer Engineering
Xihua University
Chengdu 610039, P. R. China
pqyz@263.net

[2]School of Engineering and Science
Victoria University
Melbourne, Vic 8001, Australia
Peng.Shi@vu.edu.au

[3]School of Electrical and Electronic Engineering
The University of Adelaide
Adelaide, SA 5005, Australia

Received January 2012; revised May 2012

Abstract. *Association rules mining aims to extract associations and causal structures among sets of frequent items or attributes in a large database. In practice, interesting association rules satisfy predefined minimum support and minimum confidence thresholds. In this paper, we propose a new method to generate association rules which is focused on not only minimum support and minimum confidence thresholds but the shortest length among templates as well. The method is started by a transformation of a multi-valued information system into a two-valued information system. Then, we obtain a binary relation on attributes of the two-valued information system and deduce a topology for the attributes based on the binary relation. Formally, we present two kinds of lattice of the topology for the attributes, i.e., the lattice of the topology and the quotient lattice of the topology which is deduced by the support of subset of attributes. Finally, a new association rules mining method is proposed in the quotient lattice of the topology. Compared with existing association rules mining methods, three contributions of our method were achieved as: (1) all templates of association rules are embedded in the quotient lattice of the topology for attributes; (2) templates with minimum support are shown in the quotient lattice, and association rules with confidence 1 can be mined from equivalent classes of the quotient lattice; (3) association rules with minimum support, confidence 1 and the shortest length among templates can be extracted from the quotient lattice. Examples show that our method is an alternative approach for association rules mining.*
**Keywords:** Knowledge discovery in databases, Association rules, Topology, Lattice

1. **Introduction.** Data mining is very important and necessary in information processing due to data's abundant. The aim of data mining is to extract non-trivia, implicit, previously unknown and potentially useful information from large databases, such as scientific data [4, 5, 18, 27], network data [7, 11, 17, 19, 20], and marketing transaction data [2, 13, 14, 22, 23]. Data mining can be categorized into several models, including association rules, clustering and classification. Among these models, association rule mining is the most popular method, which aims to extract associations and causal structures among sets of frequent items or attributes in a database, and is widely applied to scientific and industrial problems.

Formally, association rules are extracted from a two-valued information system $\mathscr{A} = (U, A)$ (where $U$ is a non-empty finite set of objects, $A$ is a non-empty finite set of attributes, and for any $a_i \in A$, there is mapping $a_i : U \longrightarrow \{0, 1\}$), they are considered interesting if it satisfies predefined minimum support and minimum confidence thresholds [1]. For any fixed $\mathscr{A} = (U, A)$, association rules are formalized as follows [21]: Let a template $T = \{a_{i_1}, a_{i_2}, \cdots, a_{i_k}\} \subseteq A$, the support $support_{\mathscr{A}}(T) \subseteq U$ of $T$ be the number of objects which satisfy $T$. An association rule generated from $T$ is

$$\phi \equiv P \to Q \equiv \bigwedge_{a_{i_l} \in P} a_{i_l} \to \bigwedge_{a_{i_r} \in Q} a_{i_r}, \tag{1}$$

where $P \cup Q = T$ and $P \cap Q = \emptyset$, $P$ and $Q$ are antecedent and consequent, respectively. The confidence of the association rule $\phi$ is defined by

$$confidence_{\mathscr{A}}(\phi) = \frac{support_{\mathscr{A}}(a_{i_1} \wedge a_{i_2} \wedge \cdots \wedge a_{i_k})}{support_{\mathscr{A}}(\bigwedge_{a_{i_l} \in P} a_{i_l})}. \tag{2}$$

There are two basic steps used in existing methods to generate association rules (in which, $s$ and $c$ are thresholds of support and confidence, respectively) [21]:

1. Generate as many templates $T = \{a_{i_1}, a_{i_2}, \cdots, a_{i_k}\}$ as possible such that $support_{\mathscr{A}}$ $(T) \geq s$ and $support_{\mathscr{A}}(T \wedge a_m) < s$ for any $a_m \neq T$;
2. Search for a partition $\{P, Q\}$ of $T$ such that
   (a) $support_{\mathscr{A}}(P) \leq \frac{support_{\mathscr{A}}(T)}{c}$;
   (b) $P$ has the shortest length among templates satisfying (a).

In practice, interesting associations rules are more than enough; therefore, it is desirable to find methods reducing them under no loosing information, and many improved association rule mining methods have been proposed, such as in [3], the COGAR framework is introduced to efficiently support constrained generalized association rule mining, the opportunistic confidence constraint is proposed to discriminate between significant and redundant rules. In [8], the approach based on soft set theory is presented to mine regular association rules and maximal association rules from transactional data-sets. In [9], a change and connection mining algorithm are used to discover a time delay between the quantitative changes in the data of two temporal information systems and for generating the association rules of changes from their connected decision table. In [12], the particle swarm optimization algorithm is presented to determine the threshold values of support and confidence, the method improves the quality of association rule mining. In [16], a form of the directed item-sets graph is used to store the information of frequent item-sets of transaction databases, the mining algorithm of maximal frequent item-sets and association rule based on the graph is developed. In [26], a strategy is defined by combining data mining and statistical measurement techniques, including redundancy analysis, sampling and multivariate statistical analysis, which is used to discard the non-significant rules. In [29], an evolutionary method for directly mining interesting association rules is developed, in which, whether a rule is interesting or not is decided by its relation to the keywords, and semantic and statistical methods are introduced to measure such relation.

Different from the above mentioned association rule mining method, the proposed method pays more attention to generate association rules in which $P$ has "the shortest length among templates". To do this, we firstly present a multi-valued information system which is transformed into a two-valued information system, obtain a binary relation on attributes of the two-valued information system and deduce a topology for the attributes based on the binary relation, all of these are also discussed in our previous work [24]. Then, we discuss two kinds of lattice of the topology for the attributes, the lattice of the topology and the quotient lattice of the topology. Formally, the quotient lattice of the topology is deduced by the support of subset of attributes. Finally, a new association rules mining method is proposed in the quotient lattice of the topology. We compare our new method with [20], which is widely used in existing association rules mining methods. The results demonstrate the advantages and effectiveness of our method.

2. **Topology for Attributes of Multi-valued Information Systems.** Multi-valued information systems can be transformed into two-valued information systems by adding attributes, formally, let $\mathscr{A} = (U, A)$ be a multi-valued information systems, i.e., for any $a_i$, $V_{a_i} = \{v_{i1}, v_{i2}, \cdots, v_{ij_i}\}$. Accordingly, we can obtain a new set of attributes as follows:

$$A' = \{a_{11}, a_{12}, \cdots, a_{1j_1}, a_{21}, a_{22}, \cdots, a_{2j_2}, \cdots, a_{n1}, a_{n2}, \cdots, a_{nj_n}\}, \tag{3}$$

in which, $a_{ij'_i}$ ($i \in \{1, 2, \cdots, n\}, j'_i \in \{1, 2, \cdots, j_i\}$) is equal to $v_{ij'_i}$ of $a_i \in A$, then $\mathscr{A}' = (U, A')$ is a two-valued information system, i.e., for any $a_{ij'_i} \in A'$, $a_{ij'_i} : U \longrightarrow \{0, 1\}$ means that in $\mathscr{A} = (U, A)$, object $u$ has the value $v_{ij'_i}$ or not, i.e., $a_{ij'_i}(u) = 1$ or $a_{ij'_i}(u) = 0$, respectively.

**Example 2.1.** *Considering the multi-valued information system $\mathscr{A} = (U, A)$, in which, $U = \{u_1, u_2, \cdots, u_{18}\}$ with 18 objects and $A = \{a_1, a_2, a_3, a_4, a_5\}$ (shown in Table 1).*

Table 1. The information system $\mathscr{A} = (\{u_1, u_2, \cdots, u_{18}\}, \{a_1, a_2, a_3, a_4, a_5\})$

| $A \backslash U$ | $u_1$ | $u_2$ | $u_3$ | $u_4$ | $u_5$ | $u_6$ | $u_7$ | $u_8$ | $u_9$ | $u_{10}$ | $u_{11}$ | $u_{12}$ | $u_{13}$ | $u_{14}$ | $u_{15}$ | $u_{16}$ | $u_{17}$ | $u_{18}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $a_1$ | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| $a_2$ | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| $a_3$ | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 |
| $a_4$ | 2 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 |
| $a_5$ | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 2 |

*According to Table 1, we obtain new $A' = \{a_{11}(= 0), a_{12}(= 1), a_{21}(= 1), a_{22}(= 2), a_{23}(= 3), a_{31}(= 1), a_{32}(= 2), a_{41}(= 0), a_{42}(= 1), a_{43}(= 2), a_{51}(= 1), a_{52}(= 2)\}$, and the new two-valued information system $\mathscr{A}' = (U, A')$ is shown in Table 2.*

Table 2. The two-valued information system $\mathscr{A}' = (U, A')$

| $A' \backslash U$ | $u_1$ | $u_2$ | $u_3$ | $u_4$ | $u_5$ | $u_6$ | $u_7$ | $u_8$ | $u_9$ | $u_{10}$ | $u_{11}$ | $u_{12}$ | $u_{13}$ | $u_{14}$ | $u_{15}$ | $u_{16}$ | $u_{17}$ | $u_{18}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $a_{11}$ | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| $a_{12}$ | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| $a_{21}$ | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $a_{22}$ | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $a_{23}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $a_{31}$ | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| $a_{32}$ | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| $a_{41}$ | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 |
| $a_{42}$ | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $a_{43}$ | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| $a_{51}$ | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 |
| $a_{52}$ | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |

Formally, we have the following notions based on a two-valued information system.

**Definition 2.1.** *Let* $\mathscr{A} = (U, A)$ *be a two-valued information system.* $\forall a_i \in A$*, denotes*

$$A_{a_i} = \{a_j \in A | \forall u' \in \{u \in U | a_i(u) = 1\}, a_j(u') = 1\}, \tag{4}$$

*then a binary relation on* $A$ *induced by* $A_{a_i}$ $(a_i \in A)$ *is defined as follows,* $\forall a_i, a_j \in A$*,*

$$R_A(a_i, a_j) = \begin{cases} 1, & \text{if} \quad a_j \in A_{a_i}, \\ 0, & \text{if} \quad a_j \notin A_{a_i}. \end{cases} \tag{5}$$

**Example 2.2.** *Continues Example 2.1. Based on Table 2 and (4), we can obtain* $A_{a_{11}} = \{a_{11}\}$*,* $A_{a_{12}} = \{a_{12}\}$*,* $A_{a_{21}} = \{a_{21}\}$*,* $A_{a_{22}} = \{a_{22}\}$*,* $A_{a_{23}} = \{a_{11}, a_{23}, a_{31}, a_{41}, a_{52}\}$*,* $A_{a_{31}} = \{a_{31}\}$*,* $A_{a_{32}} = \{a_{32}\}$*,* $A_{a_{41}} = \{a_{41}\}$*,* $A_{a_{42}} = \{a_{32}, a_{42}, a_{51}\}$*,* $A_{a_{43}} = \{a_{11}, a_{43}, a_{52}\}$*,* $A_{a_{51}} = \{a_{51}\}$ *and* $A_{a_{52}} = \{a_{52}\}$*. The binary relation on* $A$ *induced by the above mentioned* $A_{a_{ij}}$ *is shown in Table 3.*

TABLE 3. The binary relation on $A = \{a_{11}, a_{12}, \cdots, a_{51}, a_{52}\}$

| $A' \backslash A'$ | $a_{11}$ | $a_{12}$ | $a_{21}$ | $a_{22}$ | $a_{23}$ | $a_{31}$ | $a_{32}$ | $a_{41}$ | $a_{42}$ | $a_{43}$ | $a_{51}$ | $a_{52}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $a_{11}$ | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $a_{12}$ | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $a_{21}$ | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $a_{22}$ | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $a_{23}$ | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| $a_{31}$ | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| $a_{32}$ | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| $a_{41}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| $a_{42}$ | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| $a_{43}$ | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| $a_{51}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| $a_{52}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

**Proposition 2.1.** [24] *Let* $\mathscr{A} = (U, A)$ *be a two-valued information system. The binary relation* $R_A$ *on* $A$ *is decided by (5), then*

1. $a_j \in A_{a_i}$ *implies* $A_{a_j} \subseteq A_{a_i}$*;*
2. $R_A$ *is a reflexive and transitive relation on* $A$*.*

**Definition 2.2.** *Let* $\mathscr{A} = (U, A)$ *be a two-valued information system and the binary relation* $R_A$ *on* $A$ *decided by (5).* $(A, R_A)$ *is called an approximation space of* $\mathscr{A} = (U, A)$*. For any* $A_k \subseteq A$*,* $\overline{R_A}(A_k)$ *and* $\underline{R_A}(A_k)$*, which are called upper approximation and lower approximation of* $A_k$ *about* $(A, \overline{R_A})$*, respectively, are defined as follows:*

$$\overline{R_A}(A_k) = \{a_i \in A | A_k \cap A_{a_i} \neq \emptyset\}, \quad \underline{R_A}(A_k) = \{a_i \in A | A_{a_i} \subseteq A_k\}, \tag{6}$$

*where,* $A_{a_i}$ *is defined by (4).*

In Example 2.2, we notice that $R_A$ is not symmetrical relation on $A$, i.e., it is not necessary that $R_A$ is an equivalence relation on $A$. If $R_A$ is an equivalence relation, then for any $a_i \in A$, $A_{a_i}$ is the equivalent class $[a_i]_{R_A}$, and $\forall A_k \subseteq A$, $\overline{R_A}(A_k) = \{a_i \in A | [a_i]_{R_A} \cap A_k \neq \emptyset\}$, $\underline{R_A}(A_k) = \{a_i \in A | [a_i]_{R_A} \subseteq A_k\}$, which are Pawlak's upper approximation and lower approximation.

**Theorem 2.1.** [24] *Let* $\mathscr{A} = (U, A)$ *be a two-valued information system.* $T_{R_A} = \{\underline{R_A}(A_k) | A_k \subseteq A\}$ *is a topology for* $A$*,* $(A, T_{R_A})$ *is a topological space.* $T_{R_A}$ *is called the topology for* $A$ *induced by the approximation space* $(A, R_A)$*.*

**Theorem 2.2.** *Let $\mathscr{A} = (U, A)$ be a two-valued information system. Then,*

1. *for any $a_i \in A$, $\underline{R_A}(A_{a_i}) = A_{a_i}$, i.e., $A_{a_i} \in T_{R_A}$;*
2. *$\mathcal{B}_A = \{A_{a_i} | a_i \in A\}$ is a basis for the topology $T_{R_A}$.*

**Proof:** 1) According to (6), for any $a_j \in \underline{R_A}(A_{a_i})$, we have $a_j \in A_{a_j} \subseteq A_{a_i}$, hence, $\underline{R_A}(A_{a_i}) \subseteq A_{a_i}$. On the other hand, according to Proposition 2.1, for any $a_j \in A_{a_i}$ implies $A_{a_j} \subseteq A_{a_i}$, this means $a_j \in \underline{R_A}(A_{a_i})$, i.e., $A_{a_i} \subseteq \underline{R_A}(A_{a_i})$.

2) For any $a_i \in A$, due to $a_i \in A_{a_i}$, hence, $\bigcup_{a_i \in A} A_{a_i} = A$. On the other hand, $\forall A_{a_i}$, $A_{a_j} \in \mathcal{B}_A$, assume $a_k \in A_{a_i} \cap A_{a_j}$, then $a_k \in A_{a_i}$ and $a_k \in A_{a_j}$, according to Proposition 2.1, $A_{a_k} \subseteq A_{a_i}$ and $A_{a_k} \subseteq A_{a_j}$, i.e., $a_k \in A_{a_k} \subseteq A_{a_i} \cap A_{a_j}$ holds, according to [10] (Chapter 1 Theorem 11), $\mathcal{B}_A = \{A_{a_i} | a_i \in A\}$ is a basis for $T_{R_A}$.

According to Theorem 2.2, the topology $T_{R_A}$ for $A$ can be expressed by

$$T_{R_A} = \left\{ \bigcup_{A_{a_i} \in \mathcal{B}'_A} A_{a_i} | \forall \mathcal{B}'_A \subseteq \mathcal{B}_A \right\}. \tag{7}$$

Because any multi-valued information system $(U, A)$ can be transformed into its two-valued information system $(U, A')$ by (3), we still denote the topology $T_{R_A}$ for $A$ of $(U, A)$ without confusion by $T_{R_{A'}}$ for $A'$ of $(U, A')$.

**Example 2.3.** *Continues Example 2.2. According to Theorem 2.2, we know that $\mathcal{B}_A = \{A_{a_{11}}, A_{a_{12}}, \cdots, A_{a_{52}}\}$ is a basis for $T_{R_A}$ for $A$ of Table 1. According to (7), theoretically, we have $|T_{R_A}| \leq 2^{12} = 4096$, i.e., the number of subsets of $\mathcal{B}_A$. In this example, e.g., due to $A_{a_{23}} = \{a_{11}, a_{23}, a_{31}, a_{41}, a_{52}\}$, hence, $A_{a_{11}} \cup A_{a_{23}} = A_{a_{23}} \cup A_{a_{31}} = \cdots = A_{a_{23}}$, hence, $|T_{R_A}|$ is less than 4096, we finally generate $T_{R_A}$ and $|T_{R_A}| = 879$ without including $\emptyset$.*

## 3. Lattice Structures of the Topology $T_{R_A}$ for $A$ of $(U, A)$.

In this section, we discuss two kinds of lattice structure of the topology $T_{R_A}$ for $A$ of $(U, A)$. One is directly constructed on $T_{R_A}$, the other is constructed on a quotient set of $T_{R_A}$.

### 3.1. The lattice of the topology $T_{R_A}$.

$T_{R_A}$ for $A$ of $(U, A)$ is generated by the basis $\mathcal{B}_A = \{A_{a_i} | a_i \in A\}$, i.e., for any $T_1 \in T_{R_A}$, $T_1 = \bigcup_{a_i \in A}^{A_{a_i} \subseteq T_1} A_{a_i}$. $T_{R_A}$ is an poset by $\forall T_1, T_2 \in T_{R_A}$, $T_1 \leq T_2 \Longleftrightarrow T_1 \subseteq T_2$. On the poset $(T_{R_A}, \leq)$, for any $T_1, T_2 \in T_{R_A}$, we define

$$T_1 \wedge T_2 = \left( \bigcup_{a_i \in A}^{A_{a_i} \subseteq T_1} A_{a_i} \right) \cap \left( \bigcup_{a_j \in A}^{A_{a_j} \subseteq T_2} A_{a_j} \right), \tag{8}$$

$$T_1 \vee T_2 = \left( \bigcup_{a_i \in A}^{A_{a_i} \subseteq T_1} A_{a_i} \right) \cup \left( \bigcup_{a_j \in A}^{A_{a_j} \subseteq T_2} A_{a_j} \right). \tag{9}$$

For any $a_k \in T_1 \wedge T_2$, we have $a_k \in \bigcup_{a_i \in A}^{A_{a_i} \subseteq T_1} A_{a_i}$ and $a_k \in \bigcup_{a_j \in A}^{A_{a_j} \subseteq T_2} A_{a_j}$, hence, there exists $A_{a_i} \subseteq T_1$ and $A_{a_j} \subseteq T_2$ such that $a_k \in A_{a_i}$ and $a_k \in A_{a_j}$, according to Proposition 2.1, $A_{a_k} \subseteq A_{a_i}$ and $A_{a_k} \subseteq A_{a_j}$, hence,

$$A_{a_k} \subseteq \left( \bigcup_{a_i \in A}^{A_{a_i} \subseteq T_1} A_{a_i} \right) \cap \left( \bigcup_{a_j \in A}^{A_{a_j} \subseteq T_2} A_{a_j} \right) = T_1 \wedge T_2,$$

this means $T_1 \wedge T_2 = \bigcup_{a_k \in T_1 \wedge T_2} A_{a_k} \in T_{R_A}$. Similarly, we can prove $T_1 \vee T_2 \in T_{R_A}$. Formally, (8) and (9) can be generalized to any subset $\mathcal{T} \subseteq T_{R_A}$, i.e.,

$$\wedge \mathcal{T} = \bigwedge_{T_i \in \mathcal{T}} T_i = \bigcap_{T_i \in \mathcal{T}} \left( \bigcup_{a_i \in A}^{A_{a_i} \subseteq T_i} A_{a_i} \right);$$

$$\vee \mathcal{T} = \bigvee_{T_i \in \mathcal{T}} T_i = \bigcup_{T_i \in \mathcal{T}} \left( \bigcup_{a_i \in A}^{A_{a_i} \subseteq T_i} A_{a_i} \right).$$

**Theorem 3.1.** $(T_{R_A}, \wedge, \vee)$ *is a complete lattice, in which,* $\wedge$ *and* $\vee$ *are defined by (8) and (9),* $A$ *and* $\emptyset$ *are the greatest and least elements, respectively.*

**Example 3.1.** *Continues Example 2.3.* $(T_{R_A}, \wedge, \vee)$ *generated by the basis* $\mathcal{B}_A = \{A_{a_{11}}, A_{a_{12}}, \cdots, A_{a_{52}}\}$ *is a complete lattice with 880 elements, in which,* $A = \{a_{11}, a_{12}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{41}, a_{42}, a_{43}, a_{51}, a_{52}\}$ *and* $\emptyset$ *are the greatest and least elements, respectively.* $A_{a_{11}} = \{a_{11}\}$, $A_{a_{12}} = \{a_{12}\}$, $A_{a_{21}} = \{a_{21}\}$, $A_{a_{22}} = \{a_{22}\}$, $A_{a_{31}} = \{a_{31}\}$, $A_{a_{32}} = \{a_{32}\}$, $A_{a_{41}} = \{a_{41}\}$, $A_{a_{51}} = \{a_{51}\}$ *and* $A_{a_{52}} = \{a_{52}\}$ *are all atoms of* $(T_{R_A}, \wedge, \vee)$, $A - A_{a_{11}}$, $A - A_{a_{12}}$, $A - A_{a_{21}}$, $A - A_{a_{22}}$, $A - A_{a_{31}}$, $A - A_{a_{32}}$, $A - A_{a_{41}}$, $A - A_{a_{51}}$ *and* $A - A_{a_{52}}$ *are all molecules.*

3.2. **The quotient lattice of the topology $T_{R_A}$.** For any $T \in T_{R_A}$, we denote

$$T^{\downarrow} = \{u \in U | \forall a_i \in T, a_i(u) = 1\}. \tag{10}$$

As a special case, we denote $\emptyset^{\downarrow} = U$. In Example 2.1, according to Table 3, we have $A^{\downarrow}_{a_{12}} = (\{a_{12}\})^{\downarrow} = \{u \in U | a_{12}(u) = 1\} = \{u_5, u_7, u_{18}\}$ and $T^{\downarrow} = (A_{a_{12}} \cup A_{a_{22}})^{\downarrow} = (\{a_{12}, a_{22}\})^{\downarrow} = \{u \in U | a_{12}(u) = a_{22}(u) = 1\} = \{u_5, u_{18}\}$.

**Proposition 3.1.** *For any* $T_1, T_2 \in T_{R_A}$, *1) if* $T_1 \subseteq T_2$, *then* $T_1^{\downarrow} \supseteq T_2^{\downarrow}$; *2)* $(T_1 \cup T_2)^{\downarrow} = T_1^{\downarrow} \cap T_2^{\downarrow}$.

For any $T_1, T_2 \in T_{R_A}$, we define a binary relation on $T_{R_A}$ as follows:

$$T_1 \sim_{\downarrow} T_2 \text{ if and only if } T_1^{\downarrow} = T_2^{\downarrow}, \tag{11}$$

it is obvious that $\sim_{\downarrow}$ is an equivalence relation on $T_{R_A}$, denote quotient set (all equivalent classes of $T_{R_A}$) of $T_{R_A}$ decided by $\sim_{\downarrow}$ as $T_{R_A}/\sim_{\downarrow} = \{[T] | T \in T_{R_A}\}$, in which, for any $T' \in [T]$, $T'^{\downarrow} = T^{\downarrow}$.

**Example 3.2.** *Continues Example 2.3. There are 880 elements in* $T_{R_A}$. *According to (11), we obtain 46 equivalent classes of* $T_{R_A}$, *in which, there are 760 elements in* $[\{a_{11}, a_{12}\}] = [A]$ *such that* $\{a_{11}, a_{12}\}^{\downarrow} = A^{\downarrow} = \emptyset$, *9 elements in* $[\{a_{11}, a_{21}, a_{31}\}]$ *such that* $\{a_{11}, a_{21}, a_{31}\}^{\downarrow} = \{u_1\}$, *15 elements in* $[\{a_{11}, a_{22}, a_{31}\}]$ *such that* $\{a_{11}, a_{22}, a_{31}\}^{\downarrow} = \{u_2, u_3, u_4, u_8, u_9, u_{10}, u_{13}, u_{15}, u_{16}, u_{17}\}$, *and 16 elements in* $[\{a_{12}, a_{31}\}]$ *such that* $\{a_{12}, a_{31}\}^{\downarrow} = \{u_{18}\}$, *all equivalent classes of* $T_{R_A}$ *are shown in Table 4.*

For any $[T] \in T_{R_A}/\sim_{\downarrow}$ and $|[T]| > 1$, $([T], \subseteq)$ is a poset, e.g., $[\{a_{11}, a_{32}, a_{42}, a_{51}\}]$ such that $(\{a_{11}, a_{32}, a_{42}, a_{51}\})^{\downarrow} = \{u_6\}$, $[\{a_{11}, a_{32}, a_{42}, a_{51}\}] = \{\{a_{11}, a_{32}, a_{42}, a_{51}\}, \{a_{11}, a_{21}, a_{32}\}, \{a_{11}, a_{21}, a_{32}, a_{42}, a_{51}\}, \{a_{11}, a_{21}, a_{51}\}, \{a_{11}, a_{32}, a_{51}\}, \{a_{11}, a_{21}, a_{32}, a_{51}\}\}$, in which, $\{a_{11}, a_{21}, a_{32}\} \subseteq \{a_{11}, a_{21}, a_{32}, a_{51}\} \subseteq \{a_{11}, a_{21}, a_{32}, a_{42}, a_{51}\}$, and $\{a_{11}, a_{21}, a_{32}, a_{42}, a_{51}\}$ is the greatest element.

**Proposition 3.2.** *For any* $[T] \in T_{R_A}/\sim_{\downarrow}$ *and* $T', T'' \in [T]$, *1)* $T' \cup T'' \in [T]$; *2) there exists the greatest element in* $[T]$, *denoted by* $\cup[T]$.

TABLE 4. All equivalent classes of $T_{R_A}$

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $[A]$ | $[A_{a_{11}}]$ | $[A_{a_{12}}]$ | $[A_{a_{21}}]$ | $[A_{a_{22}}]$ | $[A_{a_{23}}]$ | $[A_{a_{31}}]$ | $[A_{a_{32}}]$ | $[A_{a_{41}}]$ | $[A_{a_{42}}]$ | $[A_{a_{43}}]$ | $[A_{a_{51}}]$ | $[A_{a_{52}}]$ | $[\{a_{11}, a_{21}\}]$ |
| 760 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 |
| $[\{a_{11}, a_{22}\}]$ | $[\{a_{11}, a_{31}\}]$ | $[\{a_{11}, a_{32}\}]$ | $[\{a_{11}, a_{41}\}]$ | $[\{a_{11}, a_{51}\}]$ | $[\{a_{11}, a_{52}\}]$ | $[\{a_{12}, a_{21}\}]$ | $[\{a_{12}, a_{22}\}]$ | | | | | | |
| 1 | 1 | 1 | 1 | 1 | 1 | 5 | 1 | | | | | | |
| $[\{a_{12}, a_{31}\}]$ | $[\{a_{12}, a_{32}\}]$ | $[\{a_{21}, a_{32}\}]$ | $[\{a_{22}, a_{31}\}]$ | $[\{a_{41}, a_{52}\}]$ | $[\{a_{22}, a_{32}\}]$ | $[\{a_{22}, a_{41}\}]$ | $[\{a_{22}, a_{51}\}]$ | | | | | | |
| 16 | 4 | 4 | 2 | 1 | 1 | 1 | 1 | | | | | | |
| $[\{a_{22}, a_{52}\}]$ | $[\{a_{31}, a_{41}\}]$ | $[\{a_{31}, a_{52}\}]$ | $[\{a_{12}, a_{22}, a_{32}\}]$ | $[\{a_{11}, a_{32}, a_{41}\}]$ | $[\{a_{11}, a_{41}, a_{52}\}]$ | | | | | | | | |
| 1 | 1 | 1 | 6 | 9 | 1 | | | | | | | | |
| $[\{a_{11}, a_{32}, a_{42}, a_{51}\}]$ | $[\{a_{11}, a_{21}, a_{31}\}]$ | $[\{a_{11}, a_{22}, a_{31}\}]$ | $[\{a_{22}, a_{41}, a_{52}\}]$ | $[\{a_{31}, a_{41}, a_{52}\}]$ | | | | | | | | | |
| 6 | 9 | 15 | 1 | 1 | | | | | | | | | |
| $[\{a_{11}, a_{22}, a_{32}\}]$ | $[\{a_{11}, a_{22}, a_{41}\}]$ | $[\{a_{11}, a_{22}, a_{43}, a_{52}\}]$ | $[\{a_{11}, a_{31}, a_{41}\}]$ | $[\{a_{11}, a_{31}, a_{52}\}]$ | | | | | | | | | |
| 6 | 1 | 3 | 1 | 1 | | | | | | | | | |

In fact, according to (10), $T'^{\downarrow} = \{u \in U | \forall a_i \in T', a_i(u) = 1\} = T''^{\downarrow} = \{u \in U | \forall a_j \in T'', a_j(u) = 1\} = \{u \in U | \forall a_k \in T' \cup T'', a_k(u) = 1\} = (T' \cup T'')^{\downarrow}$, i.e., operator $\cup$ is closed in $[T]$ and there exists the greatest element in $[T]$. For any $[T_1], [T_2] \in T_{R_A}/\sim_{\downarrow}$, define $[T_1] \leq [T_2]$ if and only if $T_1^{\downarrow} \subseteq T_2^{\downarrow}$, it is obvious that $(T_{R_A}/\sim_{\downarrow}, \leq)$ is a poset. For any $[T_1], [T_2] \in T_{R_A}/\sim_{\downarrow}$, we define

$$[T_1] \vee [T_2] = [(\cup[T_1]) \cap (\cup[T_2])], \quad [T_1] \wedge [T_2] = [T_1 \cup T_2]. \tag{12}$$

Due to for any $T_1, T_2 \in T_{R_A}$,

$$(\cup[T_1]) \cap (\cup[T_2]) = \bigcup_{a_i \in (\cup[T_1]) \cap (\cup[T_2])} A_{a_i} \in T_{R_A}, \quad T_1 \cup T_2 = \bigcup_{a_j \in T_1 \cup T_2} A_{a_j} \in T_{R_A},$$

this means $[(\cup[T_1]) \cap (\cup[T_2])] \in T_{R_A}/\sim_{\downarrow}$ and $[T_1 \cup T_2] \in T_{R_A}/\sim_{\downarrow}$.

**Proposition 3.3.** *For any $T' \in [T_1]$ and $T'' \in [T_2]$, $[T_1 \cup T_2] = [T' \cup T'']$.*

**Proof:** According to (10) and (11), for any $T' \in [T_1]$ and $T'' \in [T_2]$, $\{u \in U | \forall a_i \in T_1, a_i(u) = 1\} = T_1^{\downarrow} = T'^{\downarrow} = \{u \in U | \forall a_j \in T', a_j(u) = 1\}$ and $\{u \in U | \forall a_k \in T_2, a_k(u) = 1\} = T_2^{\downarrow} = T''^{\downarrow} = \{u \in U | \forall a_l \in T'', a_l(u) = 1\}$, according to Property 3.3, we have $(T_1 \cup T_2)^{\downarrow} = T_1^{\downarrow} \cap T_2^{\downarrow} = T'^{\downarrow} \cap T''^{\downarrow} = (T' \cup T'')^{\downarrow}$, i.e., $T_1 \cup T_2 \in [T' \cup T'']$.

**Theorem 3.2.** *$(T_{R_A}/\sim_{\downarrow}, \wedge, \vee)$ is a bounded lattice, called quotient lattice of $T_{R_A}$, in which, the greatest and least elements are $[\emptyset]$ and $[A]$, respectively.*

**Example 3.3.** *Continues Example 3.2. According to (10), for any $[T] \in T_{R_A}/\sim_{\downarrow}$, we can obtain each $[T]^{\downarrow} = T^{\downarrow}$ (shown in Table 5).*

4. **Association Rules Mining from $\boldsymbol{T_{R_A}/\sim_{\downarrow}}$.** In quotient lattice $(T_{R_A}/\sim_{\downarrow}, \wedge, \vee)$ of $T_{R_A}$, we denote $\mathcal{T}_s \subset T_{R_A}/\sim_{\downarrow}$ such that for any $[T]_s \in \mathcal{T}_s$, 1) $[T]_s \in T_{R_A}/\sim_{\downarrow}$; 2) $1 \leq s \leq |U|$; 3) $|[T]_s^{\downarrow}| \geq s$; 4) for any $[T'] \in T_{R_A}/\sim_{\downarrow}$, if $[T'] < [T]_s$, then $|[T']^{\downarrow}| < s$. We propose the following method to generate an association rule $\phi$:

1. Generate templates: For any $[T_1] \in T_{R_A}/\sim_{\downarrow}$ such that there exists $[T]_s \in \mathcal{T}_s$ and $[T]_s \leq [T_1]$ in the quotient lattice $(T_{R_A}/\sim_{\downarrow}, \wedge, \vee)$ of $T_{R_A}$, then $T \in [T_1]$ is a template;
2. Search a partition: For any $T \in [T_1]$, if $P \in [T_1]$ is a minimum element, then $T = \{P, Q\ (= T - P)\}$.

Formally, the partition $T = \{P, T - P\}$ generates the association rule $\phi \equiv P \longrightarrow Q$, where $Q = T - P$ and the confidence of $\phi$ is

$$confidence_{\mathscr{A}}(\phi) = \frac{support_{\mathscr{A}}(T)}{support_{\mathscr{A}}(P)} = \frac{|T^{\downarrow}|}{|P^{\downarrow}|} = 1. \tag{13}$$

TABLE 5. All $[T]^{\downarrow} = T^{\downarrow}$ of $T_{R_A}/\sim_{\downarrow}$

| | | | | | | |
|---|---|---|---|---|---|---|
| $[A_{a_{11}}]$ | $[A_{a_{12}}]$ | $[A_{a_{21}}]$ | $[A_{a_{23}}]$ | | | |
| $\{u_1,u_2,u_3,u_4,u_6,u_8,u_9,u_{10},u_{11},u_{12},u_{13},u_{14},u_{15},u_{16},u_{17}\}$ | $\{u_5,u_7,u_{18}\}$ | $\{u_1,u_6,u_7\}$ | $\{u_{11}\}$ | | | |
| $[A_{a_{22}}]$ | $[A_{a_{32}}]$ | $[A_{a_{42}}]$ | | | | |
| $\{u_2,u_3,u_4,u_5,u_8,u_9,u_{10},u_{12},u_{13},u_{14},u_{15},u_{16},u_{17},u_{18}\}$ | $\{u_5,u_6,u_7,u_{12},u_{14}\}$ | $\{u_5,u_6,u_7\}$ | | | | |
| $[A_{a_{31}}]$ | $[A_{a_{43}}]$ | $[\{a_{11},a_{21}\}]$ | $[\{a_{12},a_{21}\}]$ | $[\{a_{12},a_{31}\}]$ | | |
| $\{u_1,u_2,u_3,u_4,u_8,u_9,u_{10},u_{11},u_{13},u_{15},u_{16},u_{17},u_{18}\}$ | $\{u_{14}\}$ | $\{u_1,u_6\}$ | $\{u_7\}$ | $\{u_{18}\}$ | | |
| $[A_{a_{41}}]$ | $[\{a_{12},a_{22},a_{32}\}]$ | $[A_{a_{52}}]$ | | | | |
| $\{u_2,u_3,u_4,u_8,u_9,u_{10},u_{11},u_{12},u_{13},u_{15},u_{16},u_{17},u_{18}\}$ | $\{u_5\}$ | $\{u_1,u_{11},u_{12},u_{14},u_{18}\}$ | | | | |
| $[A_{a_{51}}]$ | $[\{a_{11},a_{21},a_{31}\}]$ | $[\{a_{11},a_{32},a_{41}\}]$ | $[\{a_{12},a_{22}\}]$ | | | |
| $\{u_2,u_3,u_4,u_5,u_6,u_7,u_8,u_9,u_{10},u_{13},u_{15},u_{16},u_{17}\}$ | $\{u_1\}$ | $\{u_{12}\}$ | $\{u_5,u_{18}\}$ | | | |
| $[\{a_{11},a_{22}\}]$ | $[\{a_{11},a_{31}\}]$ | | | | | |
| $\{u_2,u_3,u_4,u_9,u_{10},u_{12},u_{13},u_{14},u_{15},u_{16},u_{17}\}$ | $\{u_1,u_2,u_3,u_4,u_8,u_9,u_{10},u_{11},u_{13},u_{15},u_{16},u_{17}\}$ | | | | | |
| $[\{a_{11},a_{22},a_{32}\}]$ | $[\{a_{12},a_{32}\}]$ | $[\{a_{21},a_{32}\}]$ | $[\{a_{11},a_{32}\}]$ | $[\{a_{22},a_{32}\}]$ | $[\{a_{31},a_{52}\}]$ | $[\{a_{41},a_{52}\}]$ |
| $\{u_{12},u_{14}\}$ | $\{u_5,u_7\}$ | $\{u_6,u_7\}$ | $\{u_6,u_{12},u_{14}\}$ | $\{u_5,u_{12},u_{14}\}$ | $\{u_1,u_{11},u_{18}\}$ | $\{u_{11},u_{12},u_{18}\}$ |
| $[\{a_{11},a_{41}\}]$ | $[\{a_{22},a_{31}\}]$ | | | | | |
| $\{u_2,u_3,u_4,u_8,u_9,u_{10},u_{11},u_{12},u_{13},u_{15},u_{16},u_{17}\}$ | $\{u_2,u_3,u_4,u_8,u_9,u_{10},u_{13},u_{15},u_{16},u_{17},u_{18}\}$ | | | | | |
| $[\{a_{11},a_{51}\}]$ | $[\{a_{11},a_{52}\}]$ | $[\{a_{22},a_{52}\}]$ | $[\{a_{22},a_{41},a_{52}\}]$ | | | |
| $\{u_2,u_3,u_4,u_6,u_8,u_9,u_{10},u_{13},u_{15},u_{16},u_{17}\}$ | $\{u_1,u_{11},u_{12},u_{14}\}$ | $\{u_{12},u_{14},u_{18}\}$ | $\{u_{12},u_{18}\}$ | | | |
| $[\{a_{22},a_{41}\}]$ | $[\{a_{11},a_{22},a_{41}\}]$ | | | | | |
| $\{u_2,u_3,u_4,u_8,u_9,u_{10},u_{12},u_{13},u_{15},u_{16},u_{17},u_{18}\}$ | $\{u_2,u_3,u_4,u_8,u_9,u_{10},u_{12},u_{13},u_{15},u_{16},u_{17}\}$ | | | | | |
| $[\{a_{22},a_{51}\}]$ | $[\{a_{11},a_{22},a_{43},a_{52}\}]$ | $[\{a_{11},a_{32},a_{42},a_{51}\}]$ | | | | |
| $\{u_2,u_3,u_4,u_5,u_8,u_9,u_{10},u_{13},u_{15},u_{16},u_{17}\}$ | $\{u_{14}\}$ | $\{u_6\}$ | | | | |
| $[\{a_{31},a_{41}\}]$ | $[\{a_{11},a_{31},a_{41}\}]$ | | | | | |
| $\{u_2,u_3,u_4,u_8,u_9,u_{10},u_{11},u_{13},u_{15},u_{16},u_{17},u_{18}\}$ | $\{u_2,u_3,u_4,u_8,u_9,u_{10},u_{11},u_{13},u_{15},u_{16},u_{17}\}$ | | | | | |
| $[\{a_{11},a_{22},a_{31}\}]$ | $[\{a_{11},a_{31},a_{52}\}]$ | $[\{a_{11},a_{41},a_{52}\}]$ | $[\{a_{31},a_{41},a_{52}\}]$ | | | |
| $\{u_2,u_3,u_4,u_8,u_9,u_{10},u_{13},u_{15},u_{16},u_{17}\}$ | $\{u_1,u_{11}\}$ | $\{u_{11},u_{12}\}$ | $\{u_{11},u_{18}\}$ | | | |

Accordingly, the extracted association rule $\phi$ satisfies: 1) the template of $\phi$ is $support_{\mathscr{A}}(T) \geq s$; 2) the confidence of $\phi$ is $c \leq \frac{support_{\mathscr{A}}(T)}{support_{\mathscr{A}}(P)} = confidence_{\mathscr{A}}(\phi) = 1$; 3) the antecedent $P$ of $\phi$ has the shortest length among the template of $\phi$.

**Example 4.1.** *In Table 5, if we fix $s = 10$, then we obtain all $[T] \in T_{R_A}/\sim_{\downarrow}$ such that $|[T]^{\downarrow}| \geq 10$, i.e., $\{[A_{a_{11}}], [A_{a_{22}}], [A_{a_{31}}], [A_{a_{41}}], [A_{a_{51}}], [\{a_{11},a_{22}\}], [\{a_{11},a_{31}\}], [\{a_{11},a_{41}\}], [\{a_{22},a_{41}\}], [\{a_{31},a_{41}\}], [\{a_{22},a_{51}\}], [\{a_{11},a_{22},a_{31}\}], [\{a_{11},a_{51}\}], [\{a_{22}, a_{31}\}], [\{a_{11},a_{22},a_{41}\}], [\{a_{11},a_{31},a_{41}\}]\}$, in which, $\mathcal{T}_{10} = \{[\{a_{11},a_{22},a_{31}\}]\}$. Due to $[\{a_{22},a_{31}\}] = \{\{a_{22},a_{31}\},\{a_{22}, a_{31},a_{41}\}\}$, we have a template $T = a_{22} \wedge a_{31} \wedge a_{41}$ and the association rule*

$$\phi_1 \equiv \{a_{22},a_{31}\} \longrightarrow \{a_{41}\} \equiv (a_2 = 2) \wedge (a_3 = 1) \longrightarrow (a_4 = 0)$$

*with $support_{\mathscr{A}}(T) = |T^{\downarrow}| = |[\{a_{22},a_{31}\}]^{\downarrow}| = 11$ and $confidence_{\mathscr{A}}(\phi_1) = \frac{|\{a_{22},a_{31},a_{41}\}^{\downarrow}|}{|\{a_{22},a_{31}\}^{\downarrow}|} = 1$. Due to $[\{a_{11},a_{22},a_{31}\}] = \{\{a_{11},a_{22},a_{31}\},\{a_{11},a_{22},a_{51}\},\{a_{11},a_{31},a_{51}\},\{a_{11},a_{41},a_{51}\},\{a_{11}, a_{22},a_{31},a_{41}\},\{a_{11},a_{22},a_{31},a_{51}\},\{a_{11},a_{22},a_{41},a_{51}\},\{a_{11},a_{31},a_{41},a_{51}\},\{a_{11},a_{22},a_{31},a_{41},a_{51}\}, \{a_{22},a_{31},a_{51}\},\{a_{22},a_{41},a_{51}\},\{a_{22},a_{31},a_{41},a_{51}\},\{a_{31},a_{51}\},\{a_{31},a_{41},a_{51}\},\{a_{41},a_{51}\}\}$, minimum elements of $[\{a_{11},a_{22},a_{31}\}]$ are $\{\{a_{31},a_{51}\},\{a_{41},a_{51}\},\{a_{11},a_{22},a_{31}\},\{a_{11},a_{22},a_{51}\}\}$, for any $T \in [\{a_{11},a_{22},a_{31}\}]$ such that $T$ is not a minimum element, $T$ is a template such that $support_{\mathscr{A}}\ (T) = |[\{a_{11},a_{22},a_{31}\}]^{\downarrow}| = 10$, e.g., for template $\{a_{11},a_{22},a_{41},a_{51}\}$, we generate the association rule $\phi_2 \equiv a_{11} \wedge a_{22} \wedge a_{51} \longrightarrow a_{41}$. For template $\{a_{11},a_{22},a_{31},a_{41},a_{51}\}$, we generate association rules $\phi_3 \equiv a_{11} \wedge a_{22} \wedge a_{31} \longrightarrow a_{41} \wedge a_{51}$, $\phi_4 \equiv a_{11} \wedge a_{22} \wedge a_{51} \longrightarrow a_{31} \wedge a_{41}$, $\phi_5 \equiv a_{31} \wedge a_{51} \longrightarrow a_{11} \wedge a_{22} \wedge a_{41}$ and $\phi_6 \equiv a_{41} \wedge a_{51} \longrightarrow a_{11} \wedge a_{22} \wedge a_{31}$.*

5. **Example Analysis.** In this section, we continue Example 2.1 to compare our new method with that based on frequent closed itemsets [20]:

1. Discover all frequent closed itemsets in the information system, i.e., itemsets that are closed and have support greater or equal to *minsupport* (threshold $s$ of support);
2. Derive all frequent itemsets from the frequent closed itemsets found in phase 1. This phase consists in generating all subsets of the maximal frequent closed itemsets and deriving their support from the supports of frequent closed itemsets;
3. For each frequent itemset $T \subseteq A$ found in phase 2, generate all associate rules that can be derived from $T$ and have confidence greater or equal to *minconfidence* (threshold $c$ of confidence).

The itemset $T \subseteq A$ is said to be frequent if the support of $T$ is at least $s$, and denote the set of frequent itemsets $L = \{T \subseteq A | support(T) \geq s\}$. The set of maximal frequent itemsets is defined as $M = \{T \in L | \nexists T' \in L, T \subset T'\}$. To define frequent closed itemset, one needs Galois connection $(\uparrow, \downarrow)$ between the power set of $U$ and power set of $A$, i.e., for each $U_1 \subseteq U$, $U_1^{\uparrow} = \{a_i \in A | \forall u \in U_1, a_i(u) = 1\}$ and $\downarrow$ is defined by (10), then the itemset $I \subseteq A$ is a closed itemset if and only if $I = (I^{\downarrow})^{\uparrow}$. The closed itemset $I$ is said to be frequent if the support of $I$ is at least $s$, and denote the set of all frequent closed itemsets $FC = \{I \subseteq A | I = (I^{\downarrow})^{\uparrow}$ and $support(I) \geq s\}$. The set of maximal frequent closed itemsets is defined as $MC = \{I \in FC | \nexists I' \in FC, I \subset I'\}$. Evidently, all maximal frequent itemsets are also maximal frequent closed itemsets, i.e., $M = MC$. From the formal concept analysis point of view, for each $I \in FC$, $(I^{\downarrow}, I)$ is a formal concept of $\mathscr{A} = (U, A)$ and $FC$ is the set of intensions of all formal concepts of $\mathscr{A} = (U, A)$ such that $support(I) \geq s$.

**Example 5.1.** *Let the two-valued information system be shown in Table 2. To mine associate rules from Table 2, the method based on frequent closed itemsets firstly generates all frequent closed itemsets, i.e., the closure function* Gen-Closure *(using the closure operator $\downarrow\uparrow$ to generators and their support) is applied to each generator in $FCC_i$ (generators of size $i = 1, 2, \cdots, |A|$), determining the candidate closed itemsets and their support. Next, the set of candidate closed itemsets obtained is pruned: closed itemsets with sufficient support value are inserted in the set of frequent closed itemsets $FC_i$ (generator of the frequent closed itemset). Finally, generators in the set $FCC_{i+1}$ (containing all $i+1$-generators that will be used to construct the set of candidate frequent closed itemsets at iteration $i+1$) are determined by applying the function* Gen-Generator *(it returns the the set $FCC_{i+1}$) to the generators of frequent closed itemsets in $FC_i$. This process takes place until $FCC_{i+1}$ is empty. Then, all frequent closed itemsets have been produced and their support is known. During iterations of generating all frequent closed itemsets, one pass over the two-valued information system is necessary, in order to construct the set of candidate frequent closed itemsets (closures of generators) and count their support, e.g., for $i = 1$ in Table 2, i.e., 1-itemsets generators $FCC_1 = \{a_{11}, a_{12}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{41}, a_{42}, a_{43}, a_{51}, a_{52}\}$, by applying* Gen-Closure *on $FCC_1$, all candidate closed itemsets are shown in Table 6.*

Table 6. Candidate closed itemsets ($CCI$) of $FCC_1$

| $FCC_1$ | $a_{11}$ | $a_{12}$ | $a_{21}$ | $a_{22}$ | $a_{23}$ | $a_{31}$ |
|---|---|---|---|---|---|---|
| $CCI$ | $\{a_{11}\}$ | $\{a_{12}\}$ | $\{a_{21}\}$ | $\{a_{22}\}$ | $\{a_{11}, a_{23}, a_{31}, a_{41}, a_{52}\}$ | $\{a_{31}\}$ |
| $FCC_1$ | $a_{32}$ | $a_{41}$ | $a_{42}$ | $a_{43}$ | $a_{51}$ | $a_{52}$ |
| $CCI$ | $\{a_{32}\}$ | $\{a_{41}\}$ | $\{a_{32}, a_{42}, a_{51}\}$ | $\{a_{11}, a_{43}, a_{52}\}$ | $\{a_{51}\}$ | $\{a_{52}\}$ |

*Assuming that* minsupport *is* $s = 4$, *according to Table 5, we obtain* $FC_1 = \{\{a_{11}\}, \{a_{22}\}, \{a_{31}\}, \{a_{32}\}, \{a_{41}\}, \{a_{51}\}, \{a_{52}\}\}$. *By applying* Gen-Generator *on* $FC_1$, *we obtain* $FCC_2 = \{\{a_{11}, a_{22}\}, \{a_{11}, a_{31}\}, \{a_{11}, a_{32}\}, \{a_{11}, a_{41}\}, \{a_{11}, a_{51}\}, \{a_{11}, a_{52}\}, \{a_{22}, a_{31}\}, \{a_{22}, a_{32}\}, \{a_{22}, a_{41}\}, \{a_{22}, a_{51}\}, \{a_{22}, a_{52}\}, \{a_{31}, a_{32}\}, \{a_{31}, a_{41}\}, \{a_{31}, a_{51}\}, \{a_{31}, a_{52}\}, \{a_{32}, a_{41}\}, \{a_{32}, a_{51}\}, \{a_{32}, a_{52}\}, \{a_{41}, a_{51}\}, \{a_{41}, a_{52}\}, \{a_{51}, a_{52}\}\}$. *This process takes place until* $FCC_{i+1}$ *is empty. Finally, we obtain all frequent closed itemsets* $FC = \{\{a_{11}\}, \{a_{22}\}, \{a_{31}\}, \{a_{32}\}, \{a_{41}\}, \{a_{51}\}, \{a_{52}\}, \{a_{11}, a_{22}\}, \{a_{11}, a_{31}\}, \{a_{11}, a_{41}\}, \{a_{11}, a_{51}\}, \{a_{11}, a_{52}\}, \{a_{22}, a_{41}\}, \{a_{22}, a_{51}\}, \{a_{31}, a_{41}\}, \{a_{11}, a_{22}, a_{41}\}, \{a_{11}, a_{31}, a_{41}\}, \{a_{22}, a_{31}, a_{41}\}, \{a_{11}, a_{22}, a_{31}, a_{41}, a_{51}\}\}$.

*Based on frequent closed itemsets* $FC = \bigcup_{i=1}^{|A|} FC_i$, *frequent itemsets* $L = \bigcup_k L_k$ *can be derived, i.e., we put each frequent closed itemset* $I \in FC$ *in the set of frequent itemsets* $L_{|I|}$ *corresponding to the size of* $I$ *and determine the size* $k$ *of the largest frequent itemsets. Then we construct all sets* $L_i$, *starting from* $L_k$ *down to* $L_1$, *the set* $L_{i-1}$ *is completed using itemsets in* $L_i$. *For each i-itemset* $I$ *in* $L_i$, *all* $(i-1)$*-subsets of* $I$ *are generated. All subsets that are not present in* $L_{i-1}$ *are added to the end of* $L_{i-1}$ *with support value equal to the support of* $c$. *This process takes place until* $L_1$ *has been completed. According to above mentioned* $FC$, *frequent itemsets* $L = \bigcup_k L_k$ *is shown in Table 7.*

TABLE 7. Frequent itemsets $L = \bigcup_k L_k$ based on $FC$

| | |
|---|---|
| $L_1$ | $\{\{a_{11}\}, \{a_{22}\}, \{a_{31}\}, \{a_{32}\}, \{a_{41}\}, \{a_{51}\}, \{a_{52}\}\}$ |
| $L_2$ | $\{\{a_{11}, a_{22}\}, \{a_{11}, a_{31}\}, \{a_{11}, a_{41}\}, \{a_{11}, a_{51}\}, \{a_{11}, a_{52}\}, \{a_{22}, a_{41}\}, \{a_{22}, a_{51}\}, \{a_{31}, a_{41}\}\}$ |
| $L_3$ | $\{\{a_{11}, a_{31}, a_{41}\}, \{a_{11}, a_{22}, a_{41}\}, \{a_{22}, a_{31}, a_{41}\}\}$ |
| $L_5$ | $\{\{a_{11}, a_{22}, a_{31}, a_{41}, a_{51}\}\}$ |

*For* $\{a_{11}, a_{22}\} \in L_2$, *1-subsets are* $\{a_{11}\} \in L_1$ *and* $\{a_{22}\} \in L_1$, *there is no 1-subsets added to the end of* $L_1$. *For* $\{a_{22}, a_{31}, a_{41}\} \in L_3$, *2-subset* $\{a_{22}, a_{31}\}$ *is not present in* $L_2$ *and* $support(\{a_{22}, a_{31}\}) = support(\{a_{22}, a_{31}, a_{41}\}) = 11$, *hence,* $\{a_{22}, a_{31}\}$ *is added to the end of* $L_2$. *For* $\{a_{11}, a_{22}, a_{31}, a_{41}, a_{51}\} \in L_5$, *4-subsets are not present in* $L_4 = \emptyset$ *and their support value are equal to the support of* $\{a_{11}, a_{22}, a_{31}, a_{41}, a_{51}\}$, *hence,* $L_4 = \{\{a_{11}, a_{22}, a_{31}, a_{41}\}, \{a_{11}, a_{22}, a_{31}, a_{51}\}, \{a_{11}, a_{22}, a_{41}, a_{51}\}, \{a_{11}, a_{31}, a_{41}, a_{51}\}, \{a_{22}, a_{31}, a_{41}, a_{51}\}\}$. *Based on all frequent itemsets and their support, the problem of generating valid association rules can be solved in a straightforward manner, i.e., for every frequent itemset* $I_1$, *all subsets* $I_2$ *of* $I_1$ *are derived and the ratio* $\frac{support(I_1)}{support(I_2)}$ *is computed, if it is at least* minconfidence $c$, *then the rule* $I_2 \longrightarrow (I_1 - I_2)$ *is generated, e.g., according to* $L_4$ *and* $L_5$, *we can generate* $a_{11} \wedge a_{31} \wedge a_{41} \wedge a_{51} \longrightarrow a_{22}$ *with support 10 and* confidence *1.*

**Example 5.2.** *Let the two-valued information system be shown in Table 2. To mine associate rules from Table 2, the method proposed in this paper is as follows: 1) Select subset* $B$ *of the basis* $\mathcal{B}_A = \{A_{a_i} | a_i \in A\}$ *for the topology of Table 2 such that for each* $A_{a_j} \in B$, $support(A_{a_j}) \geq s$, *Assuming* $s = 4$, *according to Table 5, we obtain* $B = \{\{a_{11}\}, \{a_{22}\}, \{a_{31}\}, \{a_{32}\}, \{a_{41}\}, \{a_{51}\}, \{a_{52}\}\}$; *2) Basis* $B$ *is used to generate the topology such that their support is at least* minsupport, *in the example, the topology is* $L = \bigcup_k L_k$ *of Example 5.1; 3) Generate the quotient lattice* $L/\sim_\downarrow$ *of the topology by using operator* $\downarrow$, *in this example,* $L/\sim_\downarrow = \{[\{a_{11}\}], [\{a_{22}\}], [\{a_{31}\}], [\{a_{32}\}], [\{a_{41}\}], [\{a_{51}\}], [\{a_{52}\}], [\{a_{11}, a_{22}\}], [\{a_{11}, a_{31}\}], [\{a_{11}, a_{41}\}], [\{a_{11}, a_{51}\}], [\{a_{11}, a_{52}\}], [\{a_{22}, a_{41}\}], [\{a_{22}, a_{51}\}], [\{a_{31}, a_{41}\}], [\{a_{11}, a_{31}, a_{41}\}], [\{a_{11}, a_{22}, a_{41}\}], [\{a_{22}, a_{31}, a_{41}\}], [\{a_{11}, a_{22}, a_{31}, a_{41}, a_{51}\}]\}$; *4) Generate valid association rules in* $L/\sim_\downarrow$, *i.e., for every equivalent classe* $[I] \in L/\sim_\downarrow$, *we can extract valid association rules such that their support is at least* minsupport, confidence *1 and the shortest length antecedent that is a minimum element in* $[I]$, *e.g., in equivalent classe* $[\{a_{11}, a_{22}, a_{31}, a_{41}, a_{51}\}]$, *we can extract association rules*

*"$a_{41} \wedge a_{51} \longrightarrow a_{11} \wedge a_{22} \wedge a_{31}$" with support 10,* confidence *1 and the shortest length antecedent $a_{41} \wedge a_{51}$ that is a minimum element in* $[\{a_{11}, a_{22}, a_{31}, a_{41}, a_{51}\}]$.

Compared with Examples 5.1 and 5.2, differences between them are: 1) In Example 5.1, all frequent closed itemsets are generated by using the closure operator $\downarrow\uparrow$ to generators and their support. In the method proposed in this paper (Example 5.2), all frequent closed itemsets are included in equivalent classes of the quotient lattice $L/\sim_{\downarrow}$ of the topology for attributes $A$, i.e., all frequent closed itemsets are the greatest elements of equivalent classes of the quotient lattice $L/\sim_{\downarrow}$ such that their support at least *minsupport*, e.g., $\{a_{22}, a_{31}, a_{41}\}$ is a frequent closed itemsets with support 11 by the closure operator $\downarrow\uparrow$, on the other hand, $\{a_{22}, a_{31}, a_{41}\}$ is the greatest element of equivalent class $[\{a_{22}, a_{31}, a_{41}\}]$; 2) In Example 5.1, all frequent itemsets are generated by all $i$-subsets of a frequent closed itemsets $I$ ($i \leq |I|$) with their support values equal to the support of $I$. In Example 5.2, all frequent itemsets are generated by basis Basis $B$ with their support at least *minsupport*, and they are equivalent classes by the equivalence relation $\sim_{\downarrow}$, e.g., frequent itemsets $[\{a_{22}, a_{31}, a_{41}\}] = \{\{a_{22}, a_{31}\}, \{a_{22}, a_{31}, a_{41}\}\}$; 3) In Example 5.1, interesting association rules satisfy their support and confidence at least *minsupport* and *minconfidence*. In Example 5.2, interesting association rules satisfy at least *minsupport* and *minconfidence* as well as the shortest length of their antecedent.

6. **Conclusions.** In this paper, associate rules mining is finished in the quotient lattice $(T_{R_A}/\sim_{\downarrow}, \wedge, \vee)$ of $T_{R_A}$, its advantages are that 1) the relation among attributes is explained by the basis $\mathcal{B}_A$; 2) the topology $T_{R_A}$ is generated by the basis $\mathcal{B}_A$; 3) templates are found in the quotient lattice. Extracted associate rules satisfy at least *minsupport* $c$, confidence 1 and the shortest length of antecedents.

**Acknowledgements.** This work is partially supported by the National Key Basic Research Program, China (2012CB215202), the 111 Project (B12018), the National Natural Science Foundation of China (61174058, 60974052 and 61134001), Sichun Key Laboratory of Intelligent Network Information Processing (SGXZD1002-10) and Sichuan Key Technology Research and Development Program (2012GZ0019).

**REFERENCES**

[1] R. Agrawal, T. Imielinski and A. Swami, Mining association rules between sets of items in large databases, *Proc. of ACM SIGMOD*, pp.207-216, 1993.
[2] J. K. Bae and J. Kim, Integration of heterogeneous models to predict consumer behavior, *Expert Systems with Applications*, vol.37, pp.1821-1826, 2010.
[3] E. Baralis, L. Cagliero, T. Cerquitelli and P. Garza, Generalized association rule mining with constraints, *Information Sciences*, vol.194, pp.68-84, 2012.
[4] Y. L. Chen and C. H. Weng, Mining association rules from imprecise ordinal data, *Fuzzy Sets and Systems*, vol.159, pp.460-474, 2008.
[5] Y. Chen and C. Weng, Mining fuzzy association rules from questionnaire data, *Knowledge-Based Systems*, vol.22, no.1, pp.46-56, 2009.
[6] B. Ganter and R. Wille, *Formal Concept Analysis: Mathematical Foundations*, Springer, Berlin, 1999.
[7] P. Y. Hsu, Y. L. Chen and C. C. Ling, Algorithms for mining association rules in bag databases, *Information Sciences*, vol.166, no.1, pp.31-47, 2004.
[8] T. Herawan and M. M. Deris, A soft set approach for association rules mining, *Knowledge-Based Systems*, vol.24, pp.186-195, 2011.
[9] K. Kaneiwa, A rough set approach to multiple dataset analysis, *Applied Soft Computing*, vol.11, pp.2538-2547, 2011.
[10] J. L. Kelley, *General Topology*, Springer-Verlag, 1955.
[11] R. Kosala, H. Blockeel, M. Bruynooghe and J. Van den Bussche, Information extraction from structured documents using, *Data & Knowledge Engineering*, vol.58, pp.129-158, 2006.

[12] R. J. Kuo, C. M. Chao and Y. T. Chiu, Application of particle swarm optimization to association rule mining, *Applied Soft Computing*, vol.11, pp.326-336, 2011.

[13] C. Lee, B. Song and Y. Park, Design of convergent product concepts based on functionality: An association rule mining and decision tree approach, *Expert Systems with Applications*, vol.39, pp.9534-9542, 2012.

[14] A. H. Lim, C. S. Lee and M. Raman, Hybrid genetic algorithm and association rules for mining workflow best practices, *Expert Systems with Applications*, 2012.

[15] T. Y. Lin, X. Hu and E. Louie, A fast association rule algorithm based on bitmap and granular computing, *Proc. of the 12th IEEE Internat. Conf. on Fuzzy Systems*, pp.678-683, 2003.

[16] X. B. Liu, K. Zhai and W. Pedrycz, An improved association rules mining method, *Expert Systems with Applications*, vol.39, pp.1362-1374, 2012.

[17] B. Liu, R. Grossman and Z. Yanhong, Mining web pages for data records, *IEEE Intelligent Systsystems*, vol.19, no.6, pp.49-55, 2004.

[18] D. Men and Z. Pei, Extracting linguistic rules from data sets using fuzzy logic and genetic algorithms, *Neurocomputing*, vol.78, no.1, pp.48-54, 2012.

[19] V. Nebot and R. Berlanga, Finding association rules in semantic web data, *Knowledge-Based Systems*, vol.25, pp.51-62, 2012.

[20] M. J. Zaki and C. J. Hsiao, Efficient algorithms for mining closed itemsets and their lattice structure, *IEEE Transactions on Knowledge and Data Engineering*, vol.17, no.4, pp.462-478, 2005.

[21] Z. Pawlak and A. Skowron, Rough sets and Boolean reasoning, *Information Sciences*, vol.177, no.1, pp.41-73, 2007.

[22] Z. Pei, Y. Xu, D. Duan and K. Qin, Extracting complex linguistic data summaries from personnel database via simple linguistic aggregations, *Information Sciences*, vol.179, pp.2325-2332, 2009.

[23] Z. Pei, G. Resconi, A. J. Van Der Wal, K. Qin and Y. Xu, Interpreting and extracting fuzzy decision rules from fuzzy information system and their inference, *Information Sciences*, vol.176, pp.1869-1897, 2006.

[24] Z. Pei and K. Qin, Topological space for attributes set of a formal context, *RSKT2007, LNAI4481*, pp.460-466, Canada, 2007.

[25] P. Salgado, Rule generation for hierarchical collaborative fuzzy system, *Applied Mathematical Modelling*, vol.32, pp.1159-1178, 2008.

[26] I. N. M. Shaharanee, F. Hadzic and T. S. Dillon, Interestingness measures for association rules based on statistical validity, *Knowledge-Based Systems*, vol.24, pp.386-392, 2011.

[27] F. Q. Shi, S. Q. Sun and J. Xu, Employing rough sets and association rule mining in KANSEI knowledge extraction, *Information Sciences*, vol.196, pp.118-128, 2012.

[28] R. Wille, Restructuring the lattice theory: An approach based on hierarchies of concepts, in *Ordered Sets*, I. Rival (ed.), Reidel, Dordrecht, Boston, 1982.

[29] G. Yang, S. Mabu, K. Shimada and K. Hirasawa, A novel evolutionary method to search interesting association rules by keywords, *Expert Systems with Applications*, vol.38, pp.13378-13385, 2011.