Alternative Rule Reasoning: Association Rule Tree Reasoning with a Constraining Rule Ascertained using a Reasoning Framework in 2D Interestingness Area

Peera Liewlom

Abstract—**Reasoning task is necessary for scientific work. We proposed a method to ascertained association rules from a reasoning framework for rule reasoning tasks. The proposed method does not require defining the values of measures, minimum support, and minimum confidence, which introduce the following reasoning limitations in the conventional method: 1) difficult to define the reasonable values of both measures and 2) difficult to describe the reasoning relation among the determined rules. The reasoning framework is used to derive reasoning relations as sequential relations of the terms related to the domain. A sequential relation can be represented as the relations among a pair of rules between a general rule and its extended rule (i.e., a specific rule). The framework begins from the domain rule, i.e., the first general rule or domain rule acting as the constraining rule (new constraint), and then sequentially determines all the specific rules. The concept behind the reasoning to determine specific rules is that the rule extension should be beneficial (rather than incurring losses). The benefit of rule extension can be represented through the slope of interestingness in a 2D interestingness area, where the X-axis is the support value and the Y-axis is the confidence value. The increasing confidence of the rule extending on the Y-axis should be greater than the decreasing support on the X-axis. The result of the proposed method is an association rule tree branching from the domain rule, i.e., the root node acted as the constraining rule. We discovered this from a breast cancer dataset. The domain is "Class=recurrent-events." This tree has only 25 rules with sequential relations of 15 terms related to the domain from 286 records of dataset. Finally, these results are analyzed to verify the usability of the reasonableness.**

Index Terms—**Association Rule, Association Rule Tree, Data Mining, Rule Reasoning**

I. Introduction

ONE of the popular techniques for reasoning task is the association rule mining technique [1], [2]. This technique is a descriptive data mining approach to demonstrate the relationship among itemsets in a dataset with two interestingness values: support and confidence. With the association rules formed as the left-hand side {LHS} ⇒ right-hand side {RHS}, it is easy to comprehend that the itemset on the {LHS} is the antecedent, and the itemset on the {RHS} is the consequent. Several studies have applied the association rule discovery technique to various reasoning tasks. For example, some studies have organized association rules reasoning to perform root cause analysis with knowledge tools [3]–[5]. The association rule discovery technique has also been applied with case-based reasoning [3] and expert systems [4]. Another study [5] organized a frequent itemset to the 4M1E cause system with an ontology. Some studies have organized related [6] or sequential [7] events to determine the antecedent in cause-and-effect analysis. However, all reasoning tasks that have employed the traditional association rule discovery technique used statistical interestingness measures as the minimum support (minSup) and minimum confidence (minConf), limiting their reasoning abilities.

Note that minSup and minConf are defined by the user. These measures imply that the rules having higher support and confidence values than the defined values are the interesting rules and the rules having lower support and confidence values than the defined values are the uninteresting rules. This concept provides the first reasoning limitation, i.e., what are reasonable defined values? The second reasoning limitation is related to how we describe the reasoning relation of interesting rules such that unrelated rules are uninteresting and can thus be pruned by the reasoning task.

The first limitation of minSup and minConf is explained by the following example. If we are interested in rules for reasoning the cause of accident, the rules are selected in the form Antecedent ⇒ Consequence or {LHS} ⇒ {RHS} from all rules ascertained using the Apriori technique [1], [2]. All rules have the same Consequence {An Accident} for an Antecedent reasoning or analyzing the causes in {LHS}. The rule {Fast Driving} ⇒ {An Accident} may have low support and confidence values because many people drive fast without having an accident. This rule may be pruned by high values of minSup and/or minConf or by constructing a classifier that requires rules with high confidence; however, this rule is important for reasoning.

For reasoning the cause of an accident by the inverse of the rule, i.e., {An Accident} ⇒ {Fast Driving}, most accidents may occur when people are driving fast, which results in a higher confidence value than the first form. However, both rules have the same support value, which

Manuscript received January 04, 2021; revised June 5, 2021.
P. Liewlom is a lecturer in the Department of Computer and Information Science, Kasetsart University, Chalermphrakiat Sakon Nakhon Province Campus, Sakon Nakhon Province, 47000, Thailand (corresponding author e-mail: peera.li@ku.th).

may be pruned by minSup before discovering this important relation.

In addition, if both forms pass the minSup, the next step is to analyze the items that occur with {Fast Driving} causing {An Accident}. When we add these items to the {LHS} in the first form, the opportunity to pass the minConf is greater than the inverse form that adds these items to the {RHS}. This opportunity is described by the anti-monotone principle [8] of the support of the increasing itemsets, more members in itemset give less support. By calculating the confidence, the support of these itemsets on {LHS} is divided to calculate the confidence value that generates more opportunities to pass the minConf. Moreover, the confidence value of a rule in the form {Antecedent} ⇒ {Consequence} is equal to the conditional probability Pr(Consequence|Antecedent) of Bayes' law that provides more information than the other form. Therefore, we focus on the first form for a reasoning task.

Some studies use multiple supports to retrieve frequent itemsets from multi-level datasets [9] or single-level datasets [10]. With these techniques, the user defines the multiple support thresholds. Thresholds are difficult to determine when working with unfamiliar datasets.

Some association rule techniques [11] have proved this first limitation by selecting terms from the domain expert's knowledge. Thus, this technique provides high efficiency; however, it may produce different results depending on different experts. In response, this study proposes another approach using the "domain rule" as the constraint for ascertaining the related terms of sequential association rules.

The second limitation of minSup and minConf is explained by the following example. Consider the rules {the dawn}⇒ {sun rising} and {the dawn, the rooster crowing} ⇒ {sun rising}. Both rules give high confidence suitable for constructing a high accuracy classifier; however the term (or item) {the rooster crowing} of the second rule is not required for reasoning tasks because {the rooster crowing} has no direct rational relationship with the {sun rising}.

These limitations can be proved by various techniques. A statistical technique has been applied to identify self-sufficient itemsets by filtering uninteresting itemsets [12]. Top-K association rules are discovered without the minSup [13]. The association using the implication of propositional logic that is supported by its contrapositive [14] is also used for filtering uninteresting rules without the threshold.

However, we propose another approach that can filter uninteresting rules without the threshold and explain the sequential relations among all rules. The sequential relation supports rule chaining, which is important for reasoning tasks, to clearly describe the order of Antecedent Itemsets related to the Domain Itemset of interest. Therefore, we require a new method or measure for reasoning the sequential relations among rules.

Some rules types discovered using the association rule discovery technique are suitable for reasoning when the consequent {RHS} of all rules has only a single class as a member, e.g., class association rules (CARs) [15]. Thus, the antecedents {LHS} of all rules are for reasoning the {RHS} with minSup and minConf; however, it remains difficult to define reasonable minSup and minConf values.

Note that CARs also have the second limitation. The objective of CARs is to perform classification based on association rules (CBA) [15], where most confidence rules provide the best accuracy. This led to selecting rules with high confidence when constructing a classifier; thus, some interesting rules with low confidence may be pruned. In addition, some rules with high support and confidence values may be important for the classifier but not required for reasoning tasks.

Some data mining techniques have been applied to identify sequential patterns from sequential or temporal data [16]–[18]. However, these techniques are unsuitable to determine sequential patterns from the general data for the reasoning tasks.

Previous studies [5], [7], [19], [20] have applied association rules for knowledge discovery, e.g., reasoning cause-and-effect analysis as a fishbone diagram. These fishbone diagrams [5], [7], [19], [20] are used to categorize and describe the details of the causes regarding the effect. A fishbone diagram is similar to a tree reasoning for cause-and-effect analysis. From these studies, it can be assumed that association rules can be applied to a tree structure with a sequential relation for reasoning tasks.

The knowledge graph created by [21] is a useful technique for reasoning the relation between the {LHS} and {RHS} terms. This technique selects all terms with 20% top scores of TF–IDF values to create the knowledge graph. Thus, this technique still exhibits the first limitation of a user-defined value to prune uninteresting terms or rules, replacing the minSup and minConf user-defined problems by the %top scores of user-defined problem. Moreover, this technique represents the rule {} ⇒ {term} as the low information rule for exclusion, while the proposed technique in this study uses this type of rule as the domain rule to ascertain all sequential relations with this rule.

In addition to existing reasoning tools, i.e., statistical inductive rule reasoning, forward chaining, backward chaining, and categorizing cause details, we require alternative rule reasoning or rule chaining because rules with the same consequence {RHS}, similar to CARs are not suitable for backward or forward chaining. For example, assume A, B, and C are itemsets. If we have two rules A ⇒ B and B ⇒ C, we can assume that rule A ⇒ C. However, if we have two rules A ⇒ C and B ⇒ C, we can only assume that we have two rules. In the second case, we only know that A and B are related to C; thus, we require the reasoning of the relation of A and B (or among item members in {LHS}). For example, the rule {Fast Driving, Drunk Driving} ⇒ {An Accident} is explained with "Fast Driving" and "Drunk Driving" because both are related to "An Accident." However, this requires more explanation, e.g., which main term is related to the domain {An Accident} or which composited terms are related to the main term boost the relation level to the domain {An Accident}. This example is one of general scientific reasoning.

Thus, we propose a method to ascertain association rules using a reasoning framework that is not require defining minSup and minConf. The proposed method guarantees the determination of important rules from a reasoning perspective defined by the framework. The rules and sequential relations are chained and represented in a tree structure to explain the sequential relations of terms (or item members in {LHS}) that are related to the domain defined as the domain rule at the root node or the constraining rule.

Each relation in the tree is a reasoning relation for a pair of main term and composited terms that can be represented as the relation between a general rule and a specific rule. The specific rule extends one term (or item) in the {LHS} from the general rule. The characteristics of rule extension are as follows: 1) increasing the rate of the confidence value from the general rule to the specific rule and 2) decreasing the rate of the support value from the general rule to the specific rule, both of which are described by the anti-monotone principle [8] and are validated in cases wherein rules have the same {RHS}.

Both the above-mentioned characteristics are used to form the concept of the reasoning framework to generate reasoning relations, i.e., "the rule extending should gain profit of the Increasing Rate of the Confidence over the cost loss of the Decreasing Rate of the Support," which can be measured using the slope of interestingness. This measure was developed from the profitability concept in our previous paper [22], which was used to construct an itemsets tree. Herein, a new reasoning measure is defined as the formal definition for the rule tree in the interestingness area.

The remainder of this paper is organized as follows. Related works, definitions, and the proposed method are detailed. Thereafter, the experiment to test the proposed method using the compact breast cancer dataset [23] is presented in section V. The association rules tree can explain the sequential relations between items (or terms) that are related to {Class=recurrent-events} (or the domain). Finally, the results, discussion, and conclusions are provided.

II. Related Works

The main type of association rule is used for descriptive data mining, and the main tool used to determine the association rules comprises various interestingness measures, e.g., support and confidence [1], [2], lift [24], and conviction [25].

The above measures are statistical measures. The association rules must pass the minimum values of these measures as defined by the user. In addition, some reasonable rules may be pruned or filtered, and some unreasonable rules may be ascertained if these measures have high values.

Rules with high values for these measures are profitable for CBA [15], and more rules will improve efficiency of CBA classifier. Note that CBA uses rules determined from CARs.

Rules generated from the minSup and minConf may include a large number of rules that need to be pruned again using additional algorithms or measures, which is referred to as the optimization of association rule mining [26]. Previous studies have used genetic algorithms and fitness functions [26], [27], and other studies have used the ant colony optimization with a pheromone value [28]. In addition, some studies have used the harmonic mean [29] statistical function or the enhanced confidence factor [30] to prune the weak rules. Note that all optimization tasks are based on user-defined minSup and minConf values.

However, when these rules are used for reasoning tasks, we must know which rules are reasonable for selection. Some rules with high minSup and minConf values may be unreasonable and some rules with the low values may be reasonable. Therefore, we require a new measure or reasoning framework to prune unreasonable rules.

Pruning is the main task to discover interesting rules based on interestingness measures. Therefore, we can use a reasoning framework to prune unreasonable rules. The remaining rules should be described with reasonableness, e.g., forward chaining, backward chaining, and rule chaining.

Chain is one of the properties of an intrinsic order graph [31]. This paper applies this property for rules with an ordered property that can represent an ordered graph. We call the ordered relation of these rules as rule chaining.

Rule chaining can be represented by a knowledge graph [21] or a rule analysis diagram [32]. Nevertheless, a domain knowledge for the reasoning task is still required. This study proposes the domain rule and reasoning framework as the domain knowledge to present the single root node with a sequential relation of the knowledge graph for the reasoning task.

Rule chaining for reasonableness with one root node has been applied in many studies [5], [7], [19], [20] with the fishbone diagram [33]. In these studies, rules are chained by the main causes at the core branch of the fishbone diagram. All rules are then categorized in relation to these main causes, similar to a taxonomy tree of the main causes. Thus, the association rules are the reasonableness by rule chaining, similar to a tree structure.

In our previous paper [22], we proposed the slope of interestingness measure to represent an itemset tree. However, all itemsets were generated using traditional measures as minSup. As a result, this measure generated many individual trees with unreasonableness.

In another paper [34], we applied the concept of the profitability-of-interestingness measure [22] for the CBA classification task [15]. The association rules (i.e., CARs) were pruned (filtered) by the profitability-of-interestingness measure, following which the CBA classifiers were constructed. The performance of CBA with these pruned rules was as good as a CBA with unpruned rules; however, a problem of defining a reasonable minSup value remained.

Herein, we focus on sequential relations to solve the problems identified in our previous papers [22], [34]. The sequential relation can be determined using an algorithm or framework. We propose a new reasoning framework to prune unreasonable rules without defining the minSup and minConf. In addition, rule chaining is formed as the sequential relations in a single tree structure. Thus, all rules with low support exhibit reasonableness as long as they are related to the rule chaining. We also provide an advanced definition for the slope of interestingness that is suitable for the proposed reasoning framework to plot association rules tree in a 2D interestingness area.

III. Definitions

Here, we describe related definitions.

A. Dataset

Given a dataset as the data mining context Ɗ (O,I,R), O is a set of objects containing attributes with a value or items in a set I and R is a binary relation on O × I. For example, an object x, $x \in O$, has a relation to item y, $y \in I$. This relation is called xRy.

B. Association rule

The association rule comprises the itemset on the left-hand side of the rule ({LHS}) and the itemset on the right-hand side of the rule ({RHS}) in the form

$\{LHS\} \Rightarrow \{RHS\}$, where $\{LHS\} \subset I \wedge \{RHS\} \subset I \wedge \{LHS\} \cup \{RHS\} \subset I \wedge \{LHS\} \cap \{RHS\} = \varnothing$.

The rule represents the quantitative associations in dataset Ɗ between the antecedent {LHS} and the consequent {RHS} measured by the rule support and confidence described by definitions C and D.

The antecedent is related to objects x_a, $\forall x_a \in X_A \wedge X_A \subset O$. Each object has relations to all m items of y_j in {LHS}, for j = 1 to m $\wedge \forall y_j \in \{LHS\}$ then $\forall y_j$ [$x_a R y_j$].

The consequent is related to objects x_b, $\forall x_b \in X_B \wedge X_B \subset O$. Each object has relations to all p items of y_k in $\{LHS\} \cup \{RHS\}$, for k = 1 to p $\wedge$ p = m + n $\wedge \forall y_k \in \{LHS\} \cup \{RHS\}$ then $\forall y_k$ [$x_b R y_k$].

The symbol ⇒ represents the associative relation in a traditional association rule [1], [2].

C. The rule support

The rule support or S(Rule) that is discovered from dataset Ɗ is the ratio of the number of objects with all item members in the given rule to the number of all objects in the given dataset Ɗ.

The given rule $\{LHS\} \Rightarrow \{RHS\}$ has item members y_k, $\forall y_k \in \{LHS\} \cup \{RHS\}$. The number of related objects is $|X_B|$, $x_b \in X_B \wedge \forall y_k$ [$x_b R y_k$]. The rule support is calculated as follows [1], [2]:

$S(Rule) = |X_B| / |O|$, where $x_b \in X_B \wedge X_B \subset O \wedge y_k \in \{LHS\} \cup \{RHS\} \wedge \forall y_k [x_b R y_k]$.

D. The confidence

The confidence or C(Rule) that is discovered from dataset Ɗ is the ratio of the number of objects in Ɗ related to the antecedent and consequent to the number of objects in Ɗ related to the antecedent. Confidence is calculated as follows [1], [2]:

$C(Rule) = |X_B| / |X_A|$, where $x_b \in X_B \wedge X_B \subset O \wedge y_k \in \{LHS\} \cup \{RHS\} \wedge \forall y_k [x_b R y_k] \wedge x_a \in X_A \wedge X_A \subset O \wedge y_j \in \{LHS\} \wedge \forall y_j [x_a R y_j]$.

E. A pair of general and specific rules

A pair of general and specific rules (or a pair of rule extensions) comprises rules with the same {RHS} but has different one item-member in {LHS}. Here, the shorter rule has a k-itemset in the {LHS} (the general rule), and the longer rule has a (k+1)-itemset (the specific rule). The k-itemset is an itemset with k item members. {LHS} of the general rule (or $\{LHS\}_{g\text{-rule}}$), is covered by {LHS} of the specific rule (or $\{LHS\}_{s\text{-rule}}$). We can say that this pair of rules has the strict partial order relation on {LHS} of both the rules, $\{LHS\}_{g\text{-rule}} < \{LHS\}_{s\text{-rule}} \wedge |\{LHS\}_{g\text{-rule}}| = k \wedge |\{LHS\}_{s\text{-rule}}| = k + 1$.

Because $\{LHS\}_{g\text{-rule}}$ is covered by $\{LHS\}_{s\text{-rule}}$, $\{LHS\}_{g\text{-rule}} \cup \{RHS\}$ is then covered by $\{LHS\}_{s\text{-rule}} \cup \{RHS\}$. Also, the g-rule is covered by s-rule, or g-rule < s-rule.

This pair can be written in the form pair(g-rule, s-rule), when g-rule and s-rule are the general and specific rules, respectively. For example, pair (A, B) implies that A is a general rule of B or B is a specific rule of A.

Given a pair of general and specific rules with the same {RHS}, the rule extension RE = pair(g-rule, s-rule), where g-rule generated from k-itemset in {LHS} and is formed as $\{LHS_g\} \Rightarrow \{RHS\}$, s-rule generated from (k+1)-itemset in {LHS} is formed as $\{LHS_s\} \Rightarrow \{RHS\}$, $\{LHS_g\} \cup \{i\} = \{LHS_s\}$, $i \in I \wedge i \notin \{LHS_g\} \cup \{RHS\}$, and k is an integer.

Note that k is an integer beginning from 1 because the domain rule is defined by G using k = 0.

F. The sequential relation of rules

The rule extensions of the pairs of rules in E can be a sequential relation, if these pairs can be chained. For example the relations of pairs (A, B) and (B, C) by rule extension based on definition E can be represented together as the single sequential relation in the order of A <B <C (abbreviated as Seq(A, B, C)). This implies that A is a general rule of B by rule extension and B is a general rule of C by rule extension; however, A is not a general rule of C by rule extension because of the strict partial order relation.

Thus the sequential relation of n rules can be represented in the following form.

Seq(r1, r2,.., r(n − 1), rn), where r1 < r2 < … < r(n − 1) < rn, and n is the number of rules in a sequential relation.

For rule extensions, the sequential relation can be represented as the order of items in the {LHS} of the rule related to I_{domain} because $\{LHS\}_{g\text{-rule}}$ is covered by $\{LHS\}_{s\text{-rule}}$. To consider the rule $\{Item_1, Item_2, Item_3\} \Rightarrow I_{domain}$, the {LHS} of this rule has three members, $|\{LHS\}| = 3$. Thus, this rule has the order position number 3 of all the three rules in the same sequential relation. The general rules before the rule at the order position 3 are the rule at the order position number 2, $\{Item_1, Item_2\} \Rightarrow I_{domain}$, and the general rule at the order position number 1, $\{Item_1\} \Rightarrow I_{domain}$. Therefore, the positions of items in {LHS} are very important to orderly present the sequential relation of all related rules. In this paper, we use {LHS} with the order position $\{LHS_{ord}\}$ to present the sequential relation of rules.

G. The domain rule

The domain rule is the most general rule of any pair of rule extensions. All rules have the same {RHS} called the domain itemset, I_{domain}, $I_{domain} \subset I$, i.e., the itemset of domain items acting as domain knowledge, and we need to know all related items with sequential relations that can be represented in the antecedent itemset {LHS} of all rule extensions. This special rule only has domain items on {RHS} and an empty set on {LHS}. This rule considers the question: "is there any term extending relatively to the domain and what is it?" The answers are the terms subsequently added to all discovered specific rules.

The domain rule is the first rule at the root node of the association rule tree (defined as J) plotting in an interestingness area (defined as H). The rule must take the following form.

$\{\} \Rightarrow I_{domain}$ or $\varnothing \Rightarrow I_{domain}$, where I_{domain} ={domain items} $\neq \varnothing \wedge I_{domain} \subset I$.

The support of the domain rule is calculated using definition C as follows.

Support of domain rule

$S(\varnothing \Rightarrow I_{domain}) = |X_{DR}| / |O|$, where $x_{dr} \in X_{DR} \wedge X_{DR} \subset O \wedge y_d \in \varnothing \cup I_{domain} \wedge \forall y_d [\ x_{dr} R y_d\ ]$.

The confidence of the domain rule is calculated using definition D as follows.

Confidence of the domain rule

$C(\varnothing \Rightarrow I_{domain}) = |X_{DR}| / |X_A|$, where $x_{dr} \in X_{DR} \wedge X_{DR} \subset O \wedge y_d \in \varnothing \cup I_{domain} \wedge \forall y_d [\ x_{dr} R y_d\ ] \wedge x_a \in X_A \wedge X_A \subset O \wedge y_i = \varnothing \wedge \forall y_i [\ x_a R y_i\ ]$.

Because all objects x_a have $\varnothing$, the number of objects with $\varnothing$ is $|O|$, $|X_A| = |O|$. The confidence of the domain rule can be revised as follows.

$C(\varnothing \Rightarrow I_{domain}) = |X_{DR}| / |O|$

Thus, the support and confidence are equal in this special rule.

Note that the domain rule is not the interesting rule for association rule discovery. However, this rule type exists and is useful for reasoning tasks, such as "sequential rule constraint or the constraining rule" described herein. The reasoning framework could start from a reasonable rule without defining the minSup and minConf such that the support and confidence values of the domain rule are usable. If the domain rule does not exist, the most general rules are generated from two itemsets, {LHS} and I_{domain} in {RHS}. This case generates many rules with various support and confidence values, and some rules may not exist with the limitations of minSup and minConf, as described in the previous section.

H. The interestingness area

The interestingness area is the area of the coordinate grid obtained by the X and Y axes. Both axes are represented by the interestingness values of association rules, where the X-axis represents the support value and the Y-axis represents the confidence value. Thus, all rules with both values can be plotted in this area.

I. The slope of interestingness

The slope of interestingness measures the strength of the relationships (strong or weak) between rules in a pair of general and specific rules. This measure considers the changes in two interestingness values between the rules, i.e., an increasing rate of confidence and a decreasing rate of support.

From the general characteristics of any pair of rule extensions, the support of a longer rule is less than or equal to that of a shorter rule according to the anti-monotone principle [8]; however, the confidence may not be related to the anti-monotone principle. Thus, the good relations of the rule extension should exhibit greater confidence in a longer rule. Nevertheless, defining the values of the Increasing Rate of the Confidence and Decreasing Rate of the Support required to optimally discover the reasonable rules is difficult.

To address these problems, we should consider the support and confidence values together. One of the solutions is that the specific rule should gain profit of the increasing rate of the confidence over the cost loss of the decreasing rate of the support ratio from the general rule, or sentence A. This solution is transformed into Equations (1)–(9). The increasing rate of confidence is denoted $(+)R_C$, and the decreasing rate of support is denoted $(-)R_S$, where $(+)R$ represents an increasing rate, and $(-)R$ represents a decreasing rate, C is confidence, and S is support. First, the sentence A is transformed into Equation (1) as follows.

$$(+)R_C - (-)R_S \geq 0, \quad \text{when } (+)R_C \text{ is the increasing rate of Confidence, and } (-)R_S \text{ is the decrasing rate of Support} \qquad (1)$$

$$R_C = \frac{C(s-rule) - C(g-rule)}{C(g-rule)}, \quad \text{when } s-rule \text{ is the Specific Rule, and } g-rule \text{ is the General Rule} \qquad (2)$$

$$R_S = \frac{S(s-rule) - S(g-rule)}{S(g-rule)}, \quad \text{when } s-rule \text{ is the Specific Rule, and } g-rule \text{ is the General Rule} \qquad (3)$$

Equations (1)–(3) can be revised into Equations (4)–(9).

$$(+)\frac{C(s-rule)-C(g-rule)}{C(g-rule)} - (-)\frac{S(s-rule)-S(g-rule)}{S(g-rule)} \geq 0 \qquad (4)$$

$$\frac{C(s-rule)-C(g-rule)}{C(g-rule)} - \frac{S(g-rule)-S(s-rule)}{S(g-rule)} \geq 0 \qquad (5)$$

$$\frac{C(s-rule)-C(g-rule)}{C(g-rule)} \geq \frac{S(g-rule)-S(s-rule)}{S(g-rule)} \qquad (6)$$

$$\frac{C(s-rule)-C(g-rule)}{S(g-rule)-S(s-rule)} \geq \frac{C(g-rule)}{S(g-rule)} \qquad (7)$$

$$\frac{C(s-rule)-C(g-rule)}{(-)(S(s-rule)-S(g-rule))} \geq \frac{C(g-rule)}{S(g-rule)} \qquad (8)$$

$$\frac{(+)\Delta C}{(-)\Delta S} - \frac{C(g-rule)}{S(g-rule)} \geq 0, \quad \text{let } \Delta C = C(s-rule) - C(g-rule) \text{ and } \Delta S = S(s-rule) - S(g-rule) \qquad (9)$$

From Equation (9), the problem of the zero divisor by ΔS is avoided by mapping all values of the $((+)\Delta C)/((-)\Delta S)$ to slopes with the angle value in two-dimensional plotting. Here, the angles range from 0° to 90°. In general cases, we can use the function arctan for mapping the $((+)\Delta C)/((-)\Delta S)$ to the slope with the angle value. However, in case of the zero divisor in the function arctan, this is a maximum slope having angle of 90° in the vertical direction, this case is usable by plotting the pair of rule extending with the angle of 90° on the interestingness area that undefined by the angle from direct calculating the arctan of the $((+)\Delta C)/((-)\Delta S)$ with the zero divisor. Note that the (C(g-rule)/S(g-rule)) component can be transformed into the slope of the general rule node in the interestingness area, where the X and Y axes represent support and confidence, respectively. This slope is used as a base slope to compare the change in angle based on the rule extension. The $((+)\Delta C)/((-)\Delta S)$ component represents increasing confidence per decreasing support. This component can be transformed into the slope of a change related to the rule extension, i.e., the slope of the rule extension $Slope_{RE}$. Then, the slope of rule extension (or pair(g-rule,s-rule)) is compared with a base slope $Slope_{B}$, which is defined using Equations (10) and (11).

$$Slope\ of\ Interestingness\ of\ pair(g-rule, s-rule)$$

, or $SI(g-rule, s-rule)$

$$= arctan\left(\frac{(+)\Delta C}{(-)\Delta S}\right) - arctan\left(\frac{C(g-rule)}{S(g-rule)}\right),$$

$$if\ \Delta S = 0\ then\ arctan\left(\frac{(+)\Delta C}{(-)\Delta S}\right) = 90\ by\ plotting$$

$$the\ rule\ extension\ on\ the\ interestingness\ area \tag{10}$$

$$SI(g-rule, s-rule) = (-)Slope_{RE} - Slope_{B}$$
$$where\ arctan\left(\frac{(+)\Delta C}{(-)\Delta S}\right) = (-)Slope_{RE}$$
$$and\ arctan\left(\frac{C(g-rule)}{S(g-rule)}\right) = Slope_{B} \tag{11}$$

Note that the base slope always yields a positive number, and the slope of the rule extension typically yields a negative number because the support of the specific rule generally has less value than the support of the general rule. However, the sign (−) of $Slope_{RE}$ changes the negative value to a positive value that can be compared with the base slope. Therefore, the sign (−) acts as a mirror of the slope by the vertical plane. Thus, the range of angles in the vertical direction is 0° to 90° (Fig. 1).

It is easy to apply the Slope of Interestingness to the interestingness area. The specific rule node giving more $slope_{RE}$ in the vertical direction from the general rule node has better interestingness than the rule node giving less $slope_{RE}$, where the greater slope in the vertical direction represents more interestingness. Here, the largest usable slope is 90° (a right angle) when plotting in the interestingness area (Fig. 1).

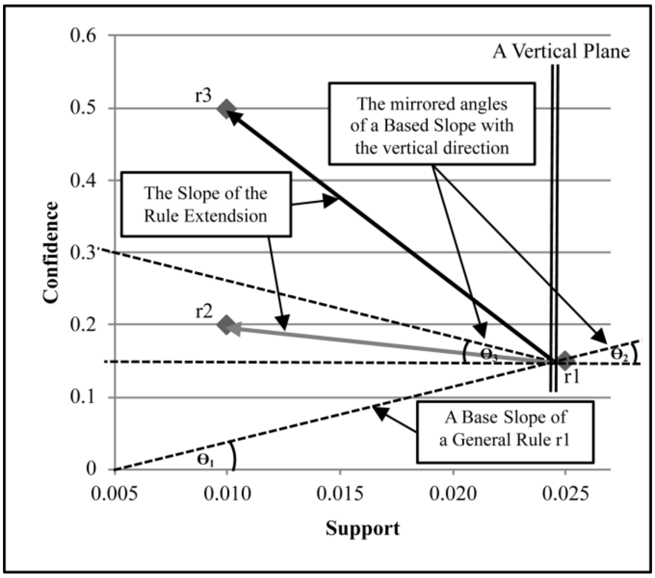


Fig. 1. Slope of Interestingness

As shown in Fig. 1, angles Θ_1, Θ_2, and Θ_3 are equal, where angle Θ_1 is the angle of a based slope from the general rule, and angles Θ_2 and Θ_3 are mirror angles of the based slope on the vertical plane. Note that both mirror angles (Θ_2, Θ_3) are equal in the vertical direction. Here, the largest usable slope is 90° (a right angle) when plotting the rule extension in the interestingness area.

The Slope of Interestingness is easy to explain with this area. Here, r1 is the general rule of r2 and r3. The relation of the pair (r1, r2) yields a smaller slope compared with the base slope in the vertical direction. Then, the relation of pair (r1, r2) has less interestingness compared to its general rule. In contrast, the relation of pair (r1, r3) yields a larger slope compared with the base slope in the vertical direction. Then, the relation of pair (r1, r3) has more interestingness compared with its general rule.

In addition, we should complete the Slope of Interestingness according to the implication of the relation between a general rule and a specific rule. Thus, we rewrite Equations (9) and (11) as Equation (12) as follows:

$$SI(g-rule, s-rule)$$
$$= (-)Slope_{RE} - Slope_{B}, \begin{cases} If\ SI \geq 0, SI\ means\ Strong \\ Otherwise, SI\ means\ Weak \end{cases} \tag{12}$$

J. The association rule tree

The association rule tree comprises nodes and edges. Each node presents the association rule as definition B. having the support as definition C. and the confidence as definition D. Thus, all nodes can plot in the interestingness area as definition H. This rule tree represents all rules with a sequential relation as in definition F, i.e., the most general rule of all sequential relations is the same rule. Thus the tree has a single root node, determined by the domain rule according to definition G. Other nodes are rule extensions as defined by definition E. These rules extend in an orderly manner from the domain rule. All pairs of rule extensions are measured according to the Slope of Interestingness (definition I); however, only pairs with “strong relations” are edges represented in the tree.

IV. Proposed Association Rules Tree with Reasoning Framework in 2D Interestingness Area

The proposed framework is for orderly describing the sequential relation of terms (or items) related to the domain. The framework based on the rule extensions is the rule chaining that is a different reasoning framework from forward or backward chaining.

Our reasoning framework is used to determine terms subsequently added to the domain as a scientific investigation. Here, the general question of the investigation is: “how are terms related to the domain?” First, we investigate the natural representation of the domain using natural characteristics, the confidence, and support. Second, we investigate the sequential relations as the reasonability of each term related to the domain by the rational measure, and we then select reasonable terms. Finally, we obtain all sequential terms related to these reasonable terms that boost the reasonability of the domain.

The natural representation of the domain and terms is represented as the association rules. Here, the domain is represented as the domain rule having a natural representation with support and confidence. The terms and related terms are represented as the rule nodes of a plotted plotting in the 2D interestingness area. Note that all nodes extending in the tree are measured for reasonability by the reasoning concept that is developed to the measure, Slope of Interestingness, defined as I. For the benefit of sequential relation reasoning, the two itemsets having the same members but different order are not the same. For instance, {A,B,C} is different from {B,C,A}. Thus, the rule {A,B,C} ⇒ {D} is different from {B,C,A} ⇒ {D}.

We brief some characteristics of our framework in Table I.

TABLE I
SOME CHARACTERISTICS OF THE PROPOSED FRAMEWORK

Subject	Substantial Representation
The objective of the reasoning framework is to describe the main and composite terms in the cause set related to the domain problem	Representation as the Association Rules Tree. The tree can orderly represent the main and composite terms through the rule chaining of the pairs (g-rules, s-rule) related to I_{domain}, defined as G.
Knowledge discovery	The Association Rule Tree Reasoning with a Constraining Rule Ascertained using a Reasoning Framework in 2D Interestingness Area.
Reasoning task	Rule Chaining through Rule Extensions**.**
Rule reasoning	Using Association Rules as the Cause-and-Effect Analysis. The rules are in the following form: if antecedent then consequent.
Causes	Items in {LHS} or the Antecedent.
Sequential relation among items in causes	Represent as the position order of items in $\{LHS_{ord}\}$
Domain or effect	I_{domain}, defined as G. or the Consequent.
Domain problem	Domain Rule acting as the Constraining Rule defined as G.
Rule chaining	The sequential relation of rules is the order relation comprising the pairs of the general and specific rules defined as F.
Sequential relation measure	Slope of the Interestingness plotting on 2D Interestingness Area.

Therefore, we have developed a method to determine association rules from a reasoning framework wherein it is not necessary to define minSup and MinConf, unlike the traditional method.

The process of the proposed method is described as follows.

Step 1. Define the domain itemset, I_{domain}, that is needed to find the related terms. The support (definition C) and confidence (definition D) are selected to define the interestingness area (definition H).

Step 2. The domain rule (definition G), i.e., {} ⇒ I_{domain} or ∅ ⇒ I_{domain}, plots this rule coordinate (support, confidence) in the interestingness area as the root node of the association rule tree. This rule can calculate support and confidence using the original method [1], [2] (definition C and D). The root node has a 0-itemset in the LHS of the rule, i.e., k = 0. Prior to measuring pairs of rule extensions (definition E) using the Slope of Interestingness (definition I), the status of the first node is referred to as "ready to grow."

Step 3. We verify all nodes with the status "ready to grow" before plotting the nodes in the interestingness area, and if these nodes have the maximum confidence value, or confidence = 1, the status of these nodes is set to "sprouted." Thereafter, if we cannot generate the (k + 1)-itemset from the nodes with the "ready to grow" status, we set the status of these nodes to "sprouted."

Step 4. From all "ready to grow" rule nodes with a k-itemset in the $\{LHS_{ord}\}$ of the rule, or $\{LHS_{ord}\}_k$, we generate specific rule nodes with a (k + 1)-itemset in the $\{LHS_{ord}\}$ of the rules (or $\{LHS_{ord}\}_{k+1}$) by adding an item y in itemset I of the dataset as follow.

$\{LHS_{ord}\}_k \cup \{y\} = \{LHS_{ord}\}_{k+1}$, $y \in I \wedge y \notin \{LHS\}_k \cup I_{domain}$, the position of y in $\{LHS_{ord}\}_{k+1}$ is k+1.

All specific rules are in the form $\{LHS_{ord}\}_{k+1} \Rightarrow I_{domain}$. Here, we calculate support and confidence values of these specific rules tested using the slope of interestingness in step 5 before plotting these nodes in the Interestingness Area.

Step 5. All pairs of each "ready to grow" rule node and its specific rules are measured the relation by the Slope of Interestingness of pair(A,B), where A represents the "ready to grow" rule and B corresponds to each specific rule of A. If the relation of any pair of rules extensions is "strong" as defined by (12), then plot these specific rule nodes in the interestingness area and draw an edge between the specific and general rules of these pairs. Next, we set the status of these specific rule nodes to "Next generate" and set the general rule of these pairs with the "ready to grow" status to the "sprouted" status. The other nodes with "weak" relation paired with this general rule are pruned out.

When we complete this step for all the nodes with the "ready to grow" status, we set the status of the "Next generate" nodes as "ready to grow" and k = k + 1.

Step 6. Here, we verify the stopping conditions. If no nodes with the "ready to grow" status are present, we stop the plotting; otherwise, we repeat step 3. Finally, we get all the rules with the "sprouted" status and Association Rule Tree in the Interestingness Area.

All steps are written to Pseudo Code as follows:

Association Rule Tree Plotting

Define I_{domain} as the {RHS} of all rules.

Define Interestingness Area as a two axes area: axis X = support value and axis Y = confidence value.

Define Constraining Rule as ∅ ⇒ I_{domain}.

//Start with the Constraining Rule as root node of the tree
k = 0. // |{LHS of the Constraining Rule}|

Plot_Node(support(Constraining Rule), confidence(Constraining Rule)) in Interestingness Area.

Ready-to-Grow ← Constraining Rule.
Sprouted-Nodes = ∅.

While Ready-to-Grow not empty

```
{  Next-Gen = ∅.
   for each rule r in Ready-to-Grow
   {  Specific-Rules = ∅.
      if Confidence(r) < 1
      then Specific-Rules ⟵ Generate_Specific_Rules(r).
      If Specific-Rules not empty
      then
      {  for each s-rule in Specific-Rules
         {  if SI(r,s-rule) not Strong
            then delete s-rule from Specific-Rules.
            else
            {  Plot_Node(support(s-rule),confidence(s-rule))
               in Interestingness Area.

               Draw_Edge(r,s-rule) in Interestingness Area.

               Next-Gen ⟵ s-rule.
            }
         }
      }
   }
   Move all rules r from Ready-to-Grow to Sprouted-Nodes.
   Ready-to-Grow = ∅.

   If Next-Gen ≠ ∅
   then Move all s-rule from Next-Gen to Ready-to-Grow.

   k = k + 1.
}
Return Sprouted-Nodes and Association Rule Tree Plotting
in Interestingness Area.
```

Generate_Specific_Rules(r)

```
Gen-S = ∅.
rset = {LHS of r} ∪ I_domain.
For each item y in I and y ∉ rset
{  {LHS_ord of s-rule} = {LHS_ord of r}.
// before adding y: |{LHS_ord of s-rule}| = k.

   adding y to position k+1 of {LHS_ord of s-rule}.
//after adding y: |{LHS_ord of s-rule}| = k+1

   if ∀y_k [ xRy_k ], x ∈ X, X ⊆ O, X≠∅, y_k ∈ I and ∀y_k ∈
{LHS of s-rule} ∪ I_domain.
   then
   {  s-rule = {LHS_ord of s-rule} ⇒ I_domain.
      Gen-S ⟵ s-rule.
   }
}
Return all s-rule in Gen-S.
```

The association rule tree consists of all sequential relations of nodes related to the root node, which is the domain rule. Each sequential relation of rules is present in the form Seq(r1, r2, r3, …, r(n − 1), rn) defined as definition F, where r1 is the general rule of r2, r2 is the general rule of r3, and so on until r(n − 1) is the general rule of rn.

To consider the detail of ri, i is an integer from 1 to n. Given i = k and k = |{LHS}| of the rule, the members in {LHS} of ri are yj, j=1 to i. For example, the Seq(r1,r2,r3), r1 is $\{y1\} \Rightarrow I_{domain}$ and r2 is $\{y1,y2\} \Rightarrow I_{domain}$ and r3 is $\{y1,y2,y3\} \Rightarrow I_{domain}$. We can see that the main term (or item) is y1, the composite terms are y2 and y3 orderly, and all terms are related to I_{domain}.

V. EXPERIMENT

The proposed method was applied to a breast cancer dataset [23] comprising 286 records with 10 attributes, including the class attribute. Here, we selected the domain as a class item {Class=recurrent-events} to discover sequential terms related to this domain. All terms can be represented by CARs discovered using WEKA software [23], and all the {RHS} of all CARs are {Class=recurrent-events}. We only select the CARs that extend from the domain rule orderly using the proposed method. All rules can be represented by an association rule tree on a 2D interestingness area.

However, we conducted an experiment that focused on explaining the dataset problems that the proposed method is suitable to solve. The problems associated with rules distribution of imbalance classes of the target dataset are as follows. While there are 286 records in the dataset, only 85 records involve {Class=recurrent-events}. Consequently, the rules with {Class=no-recurrent-events} are excluded. It is observed from the results that this dataset has clearly described both limitations of minSup and minConf values.

Thus, our experimental process comprises three steps:

First, we show the distribution of these CARs retrieved from dataset with imbalance classes on the 2D Interestingness Area described by definition H. Then, we define the rule distribution problem for this dataset and explain why our proposed method is suitable to solve the problems to ascertain the sequential rules with domain {Class=recurrent-events}.

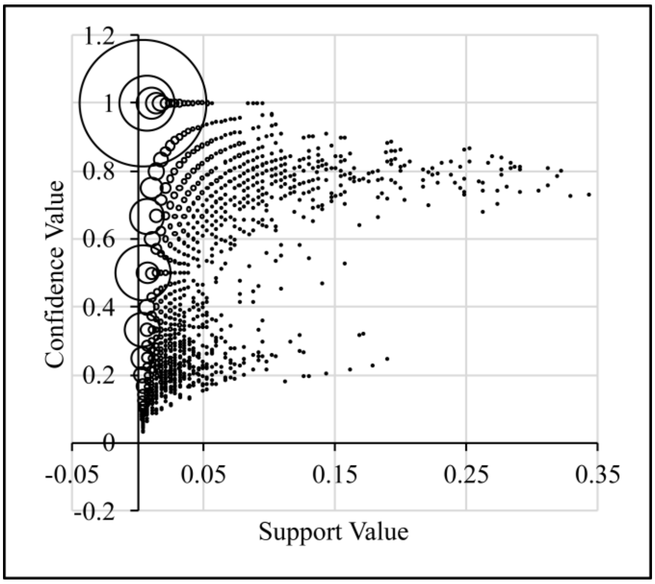
Second, we show the results obtained using the proposed method on this dataset for the reasoning tasks.

Finally, we analyze the proposed method for reasoning the sequential relations related to domain {Class=recurrent-events} compared with the traditional method exhibiting the positive CBA classifier that attempts to classify the class {Class=recurrent-events} first. For this comparison, we use CARs determined using WEKA software [23] to construct models for the two methods. However, the association rule discovery is the descriptive data mining technique. The new technique proposed in this study is for describing the sequential relations among items (or terms) for reasoning and the main and composite terms related to the domain. Thus, the analysis of the positive CBA classifier focuses on how the items (or terms) work on the CBA classifier, signifying that it is not necessary to set various test datasets and use all records of dataset as the training dataset for analyzing the relations of terms only.

VI. RESULTS

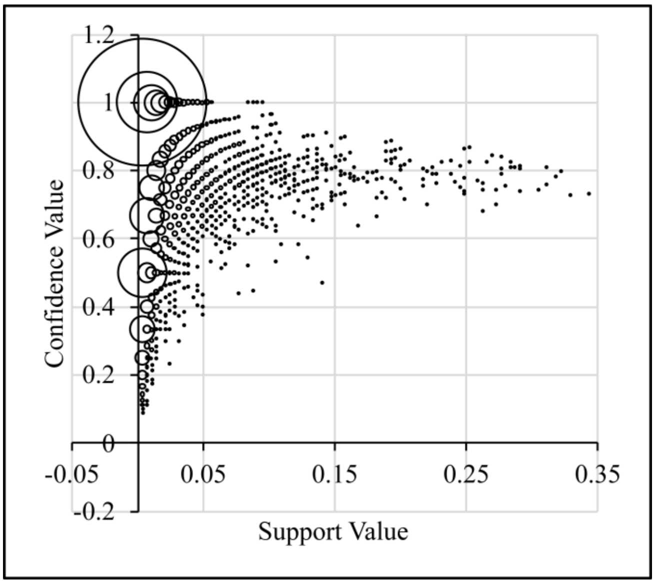
The breast cancer dataset [23] is a class imbalanced dataset with 10 attributes, 201 records with class item {Class=no-recurrent-events}, and only 85 records with class item {Class=recurrent-events}. The number of terms or items is 51 (exclude class items).

In the first step of the experiment, we obtain the distributions of CARs discovered in the target dataset. CARs are association rules with a class item on {RHS}. The class item on {RHS} can be considered a domain itemset where {Class=recurrent-events} is a positive class and {Class=no-recurrent-events} is a negative class. All {LHS} on rules can be considered as terms related to the domain. The distributions of both negative and positive CARs are shown in Figs. 2 to 4. These distributions are analyzed to define the problems that the proposed method can appropriately solve.



Note: The center of each bubble indicates the support and confidence of the rules. The size of the bubble represents the number of rules at the same coordinate.

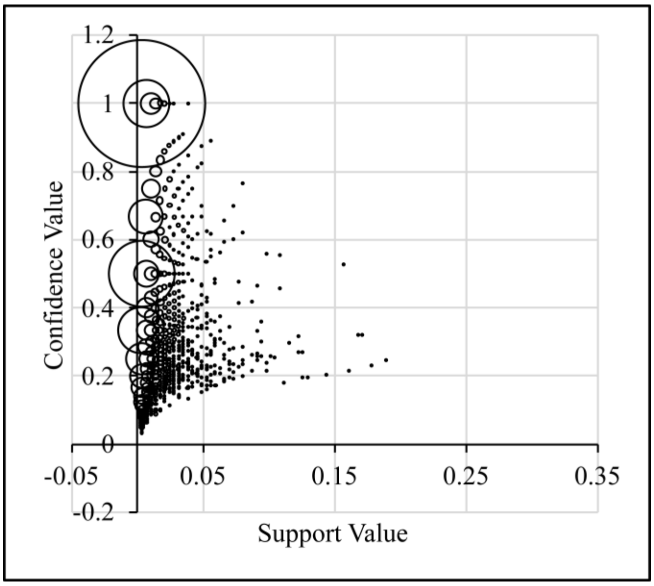
Fig. 2. Distribution of all 64,743 association rules with classes (using MS Excel)



Note: The center of each bubble indicates the support and confidence of the rules. The size of the bubble represents the number of rules at the same coordinate.

Fig. 3. Distribution of all 39,640 association rules with class itemset {Class=no-recurrent-events} (using MS Excel)

As shown in Fig. 2, a total of 64,743 CARs generated by WEKA were identified in the breast cancer dataset. The radius of each point is the number of rules represented at that point. As shown in Figs. 3 and 4, 39,640 CARs with Class item {Class=no-recurrent-events} and 25,103 CARs with Class item {Class=recurrent-events}, respectively, were found. This dataset has an unusual distribution of CARs. Most rules with {Class=no-recurrent-events} are plotted in the upper area, whereas rules with {Class=recurrent-events} are plotted in the lower-left area.



Note: The center of each bubble indicates the support and confidence of the rules. The size of the bubble represents the number of rules at the same coordinate.

Fig. 4. Distribution of all 25,103 association rules with class itemset {Class=recurrent-events} (using MS Excel)

The unusual distribution of the dataset led to determining the association rules with {Class=no-recurrent-events} at high minSup and minConf values. At low minSup and minConf values, it is difficult to prune the uninteresting rules owing to the limitations described in a previous section. The results represented in Figs. 2–4 can be tested using various minSup and minConf values with different boundary lines. For example, with boundary lines drawn at minSup = 0.1 and minConf = 0.5, the number of rules with class {Class=no-recurrent-events} is 406 and the number of rules with {Class=recurrent-events} is only 11. Note that all rules with {Class=recurrent-events} have confidence values less than 0.5 and are not usable because the same {LHS} with another class has a higher confidence value. If minConf is less than 0.5, it is difficult to explain the exclusion of rules with other class, e.g., rule {eat food A} ⇒ {sick} has the confidence 0.4 or {eat food A} ⇒ {no sick} has the confidence 0.6, which imply that {eat food A} may cause the sick, but the more confidence rule may be used to conclude that {eat food A} is not the cause of the sick. However, the proposed method can solve these reasoning problems only using the selected domain {Class=recurrent-events}, and minSup and minConf do not need to be defined. The proposed method is suitable for this case because some scientific reasoning requires positive reasoning, e.g., a dolphin is a mammal, and it is not a fish. If you define a mammal as something that is not a fish with fins or a tail, i.e., negative reasoning, this negative reasoning can lead to a dolphin being defined as a fish. In case of the breast cancer dataset, the proposed method started from a domain rule with support and confidence values that were less than 0.5. These low values are reasonable because of this general rule.

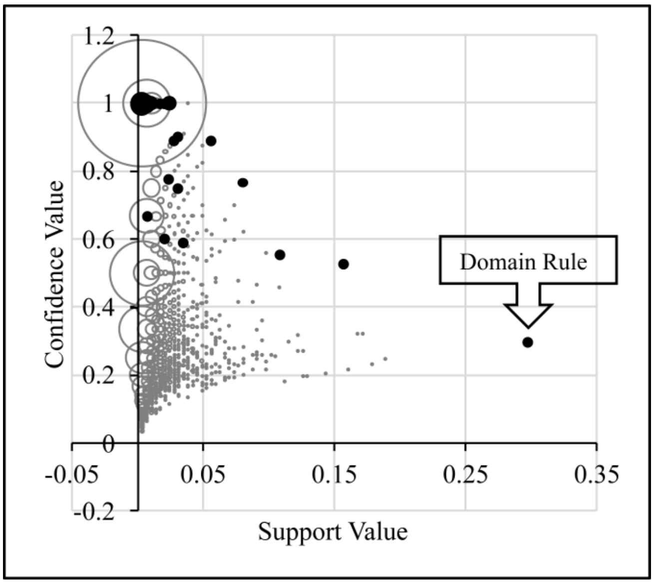
The second step is to show the results of the proposed method. Here, we selected the domain as a class item {Class=recurrent-events} to discover sequential terms related to this domain. The association rule tree generated from the breast cancer dataset comprised 26 rules, the domain rule was the root node of the tree, six rules with 1-itemset in {LHS}, 14 rules with 2-itemset in {LHS}, and five rules with 3-itemset in {LHS}. Note that the code for nodes begins with "N0," i.e., the Domain rule. We assign codes for nodes from N1 to N6 to rules with 1-itemset in the {LHS}, all nodes increase from N0. In addition, the remaining codes for nodes use the "-" sign followed by a number to describe the branch of node. For example, node N2-1 is the first branch of node N2, which is a second branch of N0. All rules are shown in Table II.

TABLE II
CODES FOR NODES OF ASSOCIATION RULE TREE GENERATED FROM BREAST CANCER DATASET

Node Code	Association Rules	Support	Confidence
N0	{∅} ⟹ {Class=recurrence-events}	0.297	0.297
N1	{deg-malig=3} ⟹ {Class=recurrence-events}	0.157	0.529
N2	{node-caps=yes} ⟹ {Class=recurrence-events}	0.108	0.554
N3	{inv-nodes=6-8} ⟹ {Class=recurrence-events}	0.035	0.588
N4	{inv-nodes=9-11} ⟹ {Class=recurrence-events}	0.021	0.600
N5	{inv-nodes=12-14} ⟹ {Class=recurrence-events}	0.007	0.667
N6	{inv-nodes=24-26} ⟹ {Class=recurrence-events}	0.003	1.000
N2-1	{node-caps=yes, deg-malig=3} ⟹ {Class=recurrence-events}	0.080	0.767
N2-2	{node-caps=yes, irradiate=yes} ⟹ {Class=recurrence-events}	0.056	0.889
N3-1	{inv-nodes=6-8,breast=left} ⟹ {Class=recurrence-events}	0.031	0.750
N3-2	{inv-nodes=6-8,deg-malig=3} ⟹ {Class=recurrence-events}	0.031	0.900
N3-3	{inv-nodes=6-8,irradiate=yes} ⟹ {Class=recurrence-events}	0.024	0.778
N3-4	{inv-nodes=6-8,breast-quad=right_low} ⟹ {Class=recurrence-events}	0.010	1.000
N4-1	{inv-nodes=9-11,irradiate=no} ⟹ {Class=recurrence-events}	0.010	1.000
N4-2	{inv-nodes=9-11,age=30-39} ⟹ {Class=recurrence-events}	0.007	1.000
N5-1	{inv-nodes=12-14,node-caps=yes} ⟹ {Class=recurrence-events}	0.007	1.000
N5-2	{inv-nodes=12-14,breast=left} ⟹ {Class=recurrence-events}	0.007	1.000
N5-3	{inv-nodes=12-14,breast-quad=left-up} ⟹ {Class=recurrence-events}	0.003	1.000
N5-4	{inv-nodes=12-14,menopause=ge40} ⟹ {Class=recurrence-events}	0.003	1.000
N5-5	{inv-nodes=12-14,tumor-size=25-29} ⟹ {Class=recurrence-events}	0.003	1.000
N5-6	{inv-nodes=12-14,tumor-size=30-34} ⟹ {Class=recurrence-events}	0.003	1.000
N3-1-1	{inv-nodes=6-8,breast=left, deg-malig=3} ⟹ {Class=recurrence-events}	0.028	0.889
N3-1-2	{inv-nodes=6-8,breast=left, irradiate=yes} ⟹ {Class=recurrence-events}	0.024	1.000
N3-3-1	{inv-nodes=6-8,irradiate=yes, menopause=ge40} ⟹ {Class=recurrence-events}	0.017	1.000
N3-3-2	{inv-nodes=6-8,irradiate=yes, deg-malig=3} ⟹ {Class=recurrence-events}	0.021	1.000
N3-3-3	{inv-nodes=6-8,irradiate=yes, breast=left} ⟹ {Class=recurrence-events}	0.024	1.000

As shown in Table II, all 26 rules have an extremely low support because the ratio of records with {Class=recurrent-events} was only 0.297. Therefore, the common association rule discovery technique (i.e., providing a user-defined minSup) typically fails to discover these rules. For example, with minSup = 0.1, only two rules were determined. However, the proposed method found 25 rules, excluding the first rule, which was set as the domain rule. Although the support values of some rules (e.g., N6 and N5-3 to N5-6 were extremely low (S = 0.003); however, these rules were acceptable if there reasonableness was evident according to the Slope of Interestingness measure.

To consider rule N3 with S = 0.034965, the main composition at the left side of the rule is the item "inv-nodes=6-8," which is a composition of the other nine rules with high confidence [0.75–1.00]. If we define minSup = 0.035, these 10 rules are not determined using the traditional CARs technique. These rules are reasonable to describe the sequential relation of item "inv-nodes=6-8" related to the recurrent-events class, which can be determined by the proposed reasoning framework.



Note: The center of each bubble indicates the support and confidence of the rules. The size of the bubble represents the number of rules at the same coordinate.

Fig. 5. Distribution of all 26 rules with Class itemset {Class=recurrent-events} discovered using the reasoning framework and domain rule constraint, represented as black points, overlaying on 25,103 association rules with class itemset {Class=recurrent-events} (using MS Excel)

Figure 5 shows all the rules listed in Table II. Here, the black bubbles plotting Interesting Areas compared to all rules with {class=recurrent-events} from Fig. 4. discovered using WEKA software. The domain rule is the rightmost node of the graph and acts as a sequential rule constraint or the constraining rule. The other 25 rules have a sequential relation related to the domain rule as an association rules tree, as shown in Figures 6 to 8 where we separated the tree into three groups of branches to simplify the complex tree with 26 nodes. The first group included branches that satisfy the stopping conditions in the 1-itemset in the {LHS} of rules (k = 1). The second group included branches that satisfy the stopping conditions in the 2-itemset in the {LHS} of the rules (k = 2), and the third group included the remaining branches.

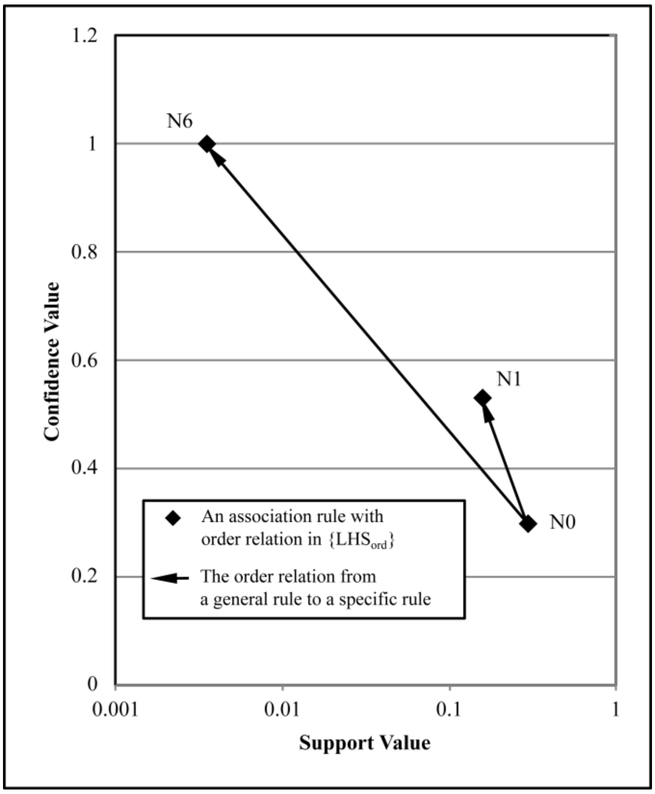


Fig. 6. First group of branches of association rule tree from breast cancer dataset (created using Excel), X-axis is a log scale

As shown in Figure 6, the first group of branches of association rule tree ascertained in the breast cancer dataset has only two rules that do not extend any new branch (or rule extension). The 1-itemsets in {LHS} of N1 and N6 act as the main terms related to the domain, i.e., {Class=recurrence-events}. Here, N6, {inv-nodes=24-26} ⇒ {Class=recurrence-events}, exhibits high confidence, which clearly explains the main cause (or Term) relate to the effect (or Domain). However, the N1 case, {deg-malig=3} ⇒ {Class=recurrence-events}, exhibits medium confidence and high support values, which has the potential to grow new branches; however, why did this not happen? To consider other rules with the term {deg-malig=3} as the composited terms, all four rules (N2-1, N3-2, N3-1-1, and N3-3-2) exhibit high confidence values. Here, we assume item "deg-malig=3" demonstrates sequential relations other than the main relation; however, all sequential relations together act as the main term related to the domain. This characteristic is the cause behind N1 having no branches of a sequential relation.

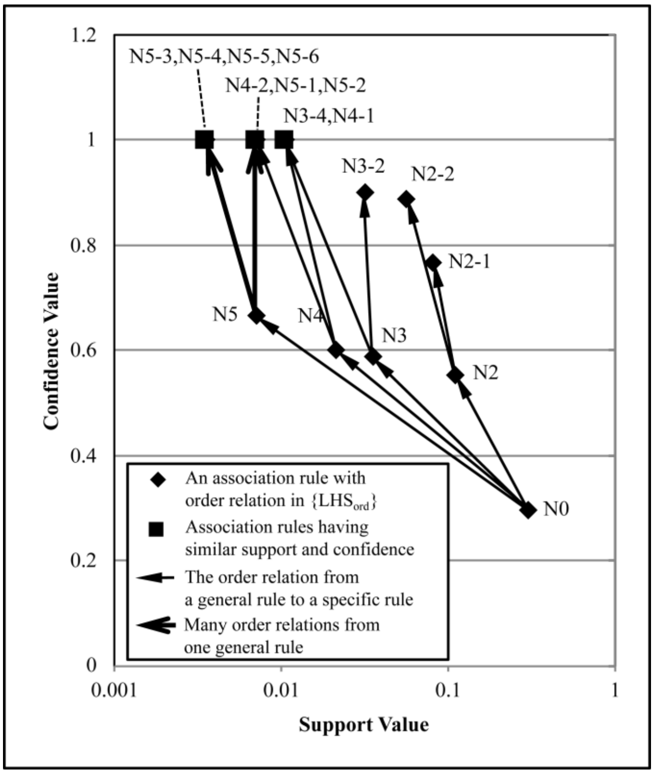


Fig. 7. Second group of the branches of association rule tree from the breast cancer dataset (created using Excel), X-axis is a log scale

As shown in Figure 7, this part of the tree includes branches that satisfied the stopping conditions in the 2-itemset in the {LHS} of the rules. Here, the blue points indicate that each point has only one node, and the black points indicate that each point has many nodes that overlay at the same coordinate. The branches of N2 ({node-caps=yes } ⇒ {Class=recurrence-events}) have two nodes, i.e., N2-1 ({node-caps=yes, deg-malig=3} ⇒ {Class=recurrence-events}) and N2-2 ({node-caps=yes, irradiate=yes} ⇒ {Class=recurrence-events}). This explains why item "node-cap=yes" acted as the main term when items "deg-malig=3" and "irradiate=yes" acted as the composited terms related to the domain, i.e., {Class=recurrence-events}.

For nodes (N3, N4, and N5), all cause itemsets with these nodes were grouped in the "inv-nodes" attribute, and these nodes act as the main terms related to the domain. Note that these nodes had medium confidence values (0.58–0.67). The remaining 12 nodes with an item (or term) extended from the {LHS} of these rules boosted the high confidence (≥ 0.9) in the domain.

As shown in Figure 8, this part of the tree has branches that satisfied the stopping condition in the 3-itemset in the {LHS} of the rules. Here, the red point indicates that this point had only one node growing from many tree paths. The main term is only the "inv-nodes=6-8" item. The first order of the sequential relations extending from the main term is the term extending in N3-1 and N3-3, i.e., "breast=left" and "irradiate=yes." The second order of the sequential relations extending from its first order is the term extending in N3-1-1, N3-1-2, N3-1-3, N3-3-1, and N3-3-2, or "deg-malig=3," "menopause=ge40," "irradiate=yes," and "breast=left." Note that some nodes have the same item members, i.e., N3-1-2 and N3-3-3. These are the same nodes as the alias name codes to highlight the order in the sequential relations related to the domain.

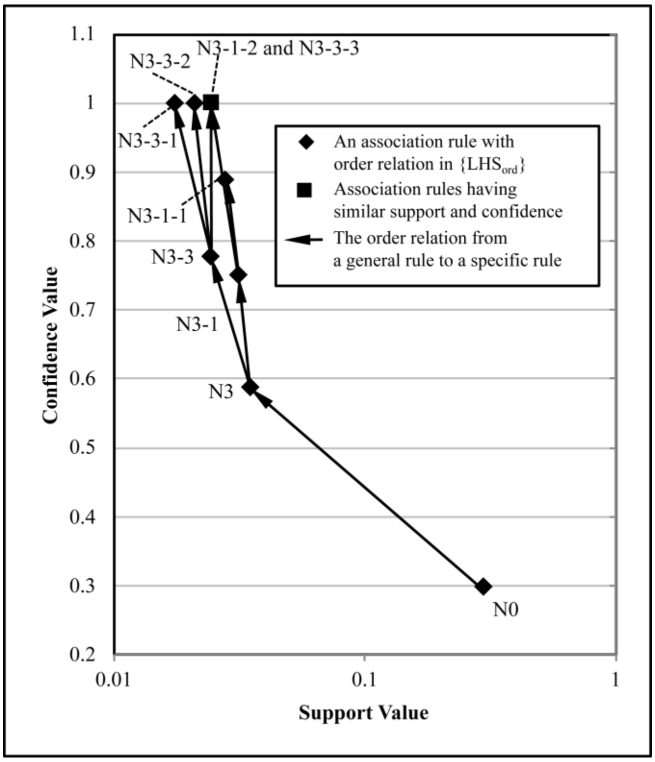


Fig. 8. Third group of the branches of association rule tree from the breast cancer dataset (created using Excel), X-axis is a log scale

In the third group, the first order of the sequential relations is the term extending with the 2-itemsets in the {LHS} of the rules that increased the confidence values from 0.59 to the range [0.75–0.78]. The second order of the sequential relations is the term extending with the 3-itemset in the {LHS} of the rules that increase the confidence values from [0.75–0.78] to [0.89–1].

The last step is to compare the proposed method with other methods. However, the proposed method identifies sequential relations among rules from a nontemporal dataset for a reasoning task; thus, it is difficult determine comparable methods. In a case where the domain item is class, the CBA classifier, which is used to classify this domain class first (the other class is the default class), can be used for a reasoning task to define the rules related to the domain. We refer to this type of CBA classifier for a reasoning task as the positive CBA classifier because we focus on the positive class {Class=recurrent-events}. However, this comparison is not undertaken to determine the better method. A rule-based classifier requires more rules for an enhanced performance, whereas a reasoning task focuses on the explanation of interesting patterns. Herein, the interesting patterns are the sequential relations between terms and domain.

The graphs shown in Figs. 5 to 8 demonstrate the ability to reason the sequential relation between terms and domain using the association rule tree. In case the domain is a class, we can consider the usability of the association rule tree in classification tasks to compare with the positive CBA classifier. However, we must be careful to define appropriate conditions for comparison. First, we used the CBA with the CARs technique [15] to analyze the relations among items (or terms) for reasoning the item causes related to the domain because we can transform the rules with the domain in the {RHS} to CARs if the domain is class. Second, we compared classifiers constructed using only CARs with {Class=recurrence-events}, and we either used the other class as the default class, or we use the positive CBA classifier. Finally, we focused on the quality of rules for reasoning the relation among rules in the classifiers, that we avoid the effect of various test datasets by using all records of dataset as train dataset for analyzing only the relations among items (or terms).

The positive CARs discovered from WEKA with minSup = 0 and minConf = 0 are 25,103 rules. The best CBA classifier consists of 47 rules with 31 terms. This classifier gives 78 true positive records and 0 false positive record, with accuracy of 0.976. In contrast, if we need the best coverage CBA classifier, this classifier consists of 170 rules with 32 terms, of which 31 terms are same as the best CBA classifier plus term {inv-nodes=9-11}. The best coverage CBA classifier gives 85 all true positive records, but gives the high false positive records of 144, with an accuracy of 0.497. We separate 170 rules, first 47 best rules and the others, to show the reasoning problem as Table III.

From Table III, Groups I and II consist of almost identical rule groups, but the result is very different. This problem is described because the same terms can perform the different {LHS} of rules, the relation of terms in {LHS}. Thus we need systematic framework to analyze the relation. The analysis is difficult because some necessary rules are excluded by the construction of CBA classifier. The problem is defined by the terms and their frequency from the best CBA classifier as shown in Table IV.

TABLE III

THE CHARACTERISTICS OF CARS GROUPS FROM THE BEST COVERAGE CBA CLASSIFIER OF THE BREAST CANCER DATASET

Characteristics	Group I: First best 47 rules	Group II: The remain 123 rules that follow Group I
The number of rules	47	123
The number of related items (or terms)	31	32 (same as Group I plus {inv-nodes=9-11})
The number of true positive records	78	7
The number of false positive records	0	144
Accuracy	0.976	0.497
Precision	1.000	0.371
Coverage	0.918	1.000
F-measure	0.957	0.541

From Table IV, we present the difficulty of analysis by two examples. Example I: The term {deg-malig=3} is presented in many rules but the rule {deg-malig=3} ⇒ {Class=recurrent-events} is excluded because of the low confidence although it has high support. Thus, the term {deg-malig=3} is considered as the only composite term. However, the analysis of the main term can be described by our new technique. The rule with {LHS} having one term should not to be excluded although the low confidence if that term is important for reasoning.

TABLE IV
TERMS AND THEIR FREQUENCY FROM 47 RULES OF THE BEST POSITIVE CBA CLASSIFIER FROM THE BREAST CANCER DATASET

Terms (or items)	Counts
irradiat=no	21
breast=left	20
deg-malig=3	**16**
menopause=premeno	16
inv-nodes=0-2	15
breast=right	14
irradiat=yes	13
menopause=ge40	13
breast-quad=left_low	11
breast-quad=left_up	10
node-caps=yes	**7**
age=50-59	7
age=40-49	6
deg-malig=1	6
deg-malig=2	6
tumor-size=30-34	6
tumor-size=25-29	5
node-caps=no	5
tumor-size=20-24	5
age=30-39	4
age=60-69	4
breast-quad=right_up	4
inv-nodes=3-5	4
tumor-size=15-19	4
tumor-size=35-39	3
tumor-size=40-44	3
inv-nodes=6-8	**2**
breast-quad=right_low	1
breast-quad=central	1
menopause=lt40	1
tumor-size=50-54	1

Example II: the term {irradiat=no} is presented in 21 rules and the opposite term {irradiat=yes} is presented in 13 rules. Both terms are presented in 34 rules from 47 rules of the best positive CBA classifier. The number of related composite terms of {irradiat=no} is 27 terms. The number of related composite terms of {irradiat=yes} is 19. The analysis of related terms of the term (irradiat=no} and the opposite term is ambiguous because both terms have 18 same related terms. However, our new technique proves this problem of ambiguity by the sequential relation.

To prove the problem of example I, we ascertained only 25 rules with (Class=recurrence-events}. except the domain rule working as constraining rule shown as Table II. These rules are represented by only 15 terms (or items). Six rules, N1 to N6, represent the main terms. The main terms are {deg-malig=3}, {inv-nodes=6-8}, {inv-nodes=9-11}, {inv-nodes=12-14}, {inv-nodes=24-26}, and {node-caps=yes}. The nine remaining terms are the composite terms. The composite term are {age=30-39}, {breast=left}, {breast-quad=left_up}, {breast-quad=right_low}, {irradiat=yes}, {irradiat=no}, {menopause=ge40}, {tumor-size=25-29}, and {tumor-size=30-34}.

To prove the problem of example II, we found the sequential relations. Term {irradiat=no} is the composite term relate to the main term {inv-nodes=9-11}. Term {irradiat=yes} is the composite term relate to the main term {inv-nodes=6-8}.

To compare terms represented into rules between our 25 rules and 47 rules from the best positive CBA classifier. We found only one rule is the same rule in both rule group because of the different objectives of both techniques. However, the number of same terms represented in both rules group is high: 12 from 15 terms, represented as filled gray color cell in Table IV. We found 3 main terms (represented as bold and underline text) and all nine composite terms of our rules working on 21 rules from 47 rules. The main terms is {deg-malig=3}, {inv-nodes=6-8}, and {node-caps=yes}. Twenty one rules from the best positive CBA classifier have {deg-malig=3} or {inv-nodes=6-8} or {node-caps=yes}. These rules found 22 related composite terms including all nine composite terms found by our technique, and each rule has at least one term of these 9 composite terms. We can use the association rule tree discovered by reasoning framework to describe the sequential relation of 21 rules from 47 rules of the best positive CBA classifier. The association rule tree can be excluded from the non-related main terms and its branches (N4, N5, and N6 and its branches). The remaining tree consist of three main terms and nine composite terms that sufficient to describe the hidden sequential relation in 21 of 47 rules. This comparative analysis show ability of our technique of reasoning the sequential relation hidden in association rules.

VII. DISCUSSION

The proposed method differs from previous methods. The proposed method ascertains a small number of rules, similar to the optimization of association rule mining [26]–[30]. However, these studies use various techniques such as using algorithms or measures to repeatedly prune the output rules already pruned through traditional measures [1], [2]. The proposed method prunes rules using a reasoning framework adapted directly from the literature [1], [2] for reasoning tasks; however, the proposed method uses the domain rule to define the objective of the reasoning task.

The proposed method is an alternative method that does not define the minSup and minConf and differs from techniques [12]–[14] that do not define the objective of the reasoning task by the constraining rule, i.e., the domain rule.

This rule is a new constraint described herein.

The proposed tree reasoning method differs from other tree reasoning techniques, e.g., the fishbone diagram [5], [7], [19], [20]. The proposed method employs a tree with sequential relations (not just the categories of causes), which are usually defined by the user or 4MIE and are detailed by the rules. In addition, the rules obtained using the proposed method are determined using a reasoning framework, whereas the rules obtained through a fishbone diagram are determined using the traditional technique with minSup and minConf.

The proposed tree reasoning method differs from knowledge graph [21] created from 20% top score of TF–IDF. Our tree has only one root node constraining all sequential relation using reasoning framework.

The proposed method fixes the problems of previous studies [22], [34]. For example, rules with reasonableness are not pruned by minSup or minConf. Additionally, the {RHS} of the rules is the {Domain}, which can be applied by any itemset (not only {Class}, which is limited to using CARs for classification tasks [22], [34]. Thus, the proposed method can be used for both classification and reasoning tasks. In the literature [22], the more than one itemset trees from CARs are generated so that we can obtain different numbers of trees at various minimum supports. In another study [34], CARs with the profitability-of-interestingness measure for CBA yields a small number of rulesets but performs similar to the traditional CARs. However, at various minSup values, this technique outputs different rulesets. Additionally, the rulesets [34] can be transformed into trees with many root nodes because the true root node of the tree is pruned or cut off by minSup, and the remaining root nodes are the root nodes of branches. However, using a single dataset, the proposed method works effectively using one relatively expanding tree related to the domain with reasonableness.

VIII. CONCLUSIONS

These results demonstrate the ease of creating and the sequential relations of terms related to the domain using the association rule tree. This tree is suitable to define the main terms and composite terms to describe sequential relations in reasoning tasks.

Herein, we have described the determination of association rules using a reasoning framework with a domain rule constraint, i.e., a constraining rule that solves both the limitations of the traditional technique, i.e., reasonable values for measures and the reasoning relation among discovered rules. With the proposed reasoning method, minSup and minConf need not be defined to prune unreasonable rules, unlike the traditional association rules discovery technique [1], [2]. The rules are chained by a tree that is beneficial for reasoning the sequential relation of terms as the sequential antecedence and domain as the consequence.

This method is an alternative to ascertain the association rules with sequential relations by the reasoning framework in the 2D interestingness area. In the future, we can apply the proposed method to other reasoning frameworks.

REFERENCES

[1] R. Agrawal, T. Imieliński, and A. Swami, “Mining association rules between sets of items in large databases,” in *Acm sigmod record*, 1993, vol. 22, pp. 207–216.

[2] R. Agrawal and R. Srikant, “Fast algorithms for mining association rules,” in *Proc. 20th int. conf. very large data bases, VLDB*, 1994, vol. 1215, pp. 487–499.

[3] R. Chougule, D. Rajpathak, and P. Bandyopadhyay, “An integrated framework for effective service and repair in the automotive domain: An application of association mining and case-based-reasoning,” *Computers in Industry*, vol. 62, no. 7, pp. 742–754, 2011.

[4] C. K. H. Lee, K. L. Choy, G. T. Ho, K.-S. Chin, K. M. Y. Law, and Y. K. Tse, “A hybrid OLAP-association rule mining based quality management system for extracting defect patterns in the garment industry,” *Expert Systems with Applications*, vol. 40, no. 7, pp. 2435–2446, 2013.

[5] Z. Xu, Y. Dang, and P. Munro, “Knowledge-driven intelligent quality problem-solving system in the automotive industry,” *Advanced Engineering Informatics*, vol. 38, pp. 441–457, 2018.

[6] C.-W. Cheng, C.-C. Lin, and S.-S. Leu, “Use of association rules to explore cause–effect relationships in occupational accidents in the Taiwan construction industry,” *Safety science*, vol. 48, no. 4, pp. 436–444, 2010.

[7] L. Cheng, Z. Y. Schon, R. M. Arnaldo Valdés, V. F. Gómez Comendador, and F. J. Sáez Nieto, “A Case Study of Fishbone Sequential Diagram Application and ADREP Taxonomy Codification in Conventional ATM Incident Investigation,” *Symmetry*, vol. 11, no. 4, p. 491, 2019.

[8] P.-N. Tan, M. Steinbach, and V. Kumar, “Association analysis: basic concepts and algorithms,” in *Introduction to Data mining*, Addison-Wesley Boston, MA, 2005.

[9] J. Han and Y. Fu, “Discovery of multiple-level association rules from large databases,” in *VLDB*, 1995, vol. 95, pp. 420–431.

[10] B. Liu, W. Hsu, and Y. Ma, “Mining association rules with multiple minimum supports,” in *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, 1999, pp. 337–341.

[11] J. Annapurna and A. K. Cherukuri, “Exploring attributes with domain knowledge in formal concept analysis,” *Journal of computing and information technology*, vol. 21, no. 2, pp. 109–123, 2013.

[12] G. I. Webb, “Self-sufficient itemsets: An approach to screening potentially interesting associations between items,” *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 4, no. 1, pp. 1–20, 2010.

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

[14] A. T. H. Sim, M. Indrawan, and B. Srinivasan, “A Threshold Free Implication Rule Mining,” in *Proceedings of the International MultiConference of Engineers and Computer Scientists*, 2008, vol. 1.

[15] B. Liu, W. Hsu, and Y. Ma, “Integrating classification and association rule mining.,” in *KDD*, 1998, vol. 98, pp. 80–86.

[16] R. Agrawal and R. Srikant, “Mining sequential patterns,” in *Proceedings of the eleventh international conference on data engineering*, 1995, pp. 3–14.

[17] R. Srikant and R. Agrawal, “Mining sequential patterns: Generalizations and performance improvements,” in *International Conference on Extending Database Technology*, 1996, pp. 1–17.

[18] J. Pei, J. Han, and W. Wang, “Mining sequential patterns with constraints in large databases,” in *Proceedings of the eleventh international conference on Information and knowledge management*, 2002, pp. 18–25.

[19] T. Djatna and I. M. Alitu, “An application of association rule mining in total productive maintenance strategy: an analysis and modelling in wooden door manufacturing industry,” *Procedia Manufacturing*, vol. 4, pp. 336–343, 2015.

[20] G. Casola, C. Siegmund, M. Mattern, and H. Sugiyama, “Data mining algorithm for pre-processing biopharmaceutical drug product manufacturing records,” *Computers & Chemical Engineering*, vol. 124, pp. 253–269, 2019.

[21] H.-J. Kim, J.-W. Baek, and K. Chung, “Optimization of Associative Knowledge Graph Using TF-IDF Based Ranking Score,” *Applied Sciences*, vol. 10, no. 13, p. 4590, 2020.

[22] P. Liewlom, “Representation of Class Association Rules (CARs) with Itemsets Tree Plot: Case study of Breast Cancer Dataset,” in *NCIT2017: The 9th National Conference on Information Technology*, Thailand, 2017, pp. 99–104, [Online]. Available: https://ncit2017.ict.mahidol.ac.th/.

[23] G. Holmes, A. Donkin, and I. H. Witten, “Weka: A machine learning workbench,” 1994.

[24] M. J. Berry and G. Linoff, “Data Mining Techniques: For Marketing, Sales, and Customer Support,” 1997.

[25] S. Brin, R. Motwani, J. D. Ullman, and S. Tsur, “Dynamic itemset counting and implication rules for market basket data,” in *Proceedings of the 1997 ACM SIGMOD international conference on Management of data*, 1997, pp. 255–264.

[26] M. Saggar, A. K. Agrawal, and A. Lad, “Optimization of association rule mining using improved genetic algorithms,” in *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)*, 2004, vol. 4, pp. 3725–3729.

[27] P. Sonar and U. Bhosle, “Optimization of association rule mining for mammogram classification,” *International Journal of Image Processing*, vol. 11, no. 3, pp. 67–85, 2017.

[28] B. Patel, V. K. Chaudhari, R. K. Karan, and Y. K. Rana, “Optimization of association rule mining apriori algorithm using ACO,” *International Journal of Soft Computing and Engineering*, vol. 1, no. 1, pp. 24–26, 2011.

[29] J. Alwidian, B. Hammo, and N. Obeid, “Enhanced CBA algorithm based on apriori optimization and statistical ranking measure,” in *Proceeding of 28th International Business Information Management Association (IBIMA) conference on Vision*, 2016, pp. 4291–4306.

[30] Y. Wang and T. Murata, “Association Rule Mining with Data Item including Independency based on Enhanced Confidence Factor,” in *Proceedings of the International MultiConference of Engineers and Computer Scientists*, 2017, vol. 1.

[31] L. González, “Edges, chains, shadows, neighbors and subgraphs in the intrinsic order graph,” *IAENG International Journal of Applied Mathematics*, vol. 42, no. 1, pp. 66–73, 2012.

[32] D. Li, D. Yang, and J. Zhang, “ARB: Knowledge Discovery and Disease Diagnosis on Thyroid Disease Diagnosis integrating Association Rule with Bagging Algorithm,” *Engineering Letters*, vol. 28, no. 2, pp390-399, 2020.

[33] K. Ishikawa, *Introduction to quality control*. Productivity Press, 1990.

[34] P. Liewlom, “Class-Association-Rules Pruning by the Profitability-of-Interestingness Measure: Case Study of an Imbalanced Class Ratio in a Breast Cancer Dataset,” in *2020 4th International Conference on Computer, Software and Modeling. 17-19 July 2020, Rome Italy.* (accepted) *Journal of Advances in Information Technology (JAIT),* vol. 12, no. 3, 2021.