Combined Mining: Analyzing Object and Pattern Relations for Discovering Actionable Complex Patterns *

Longbing Cao
Advanced Analytics Institute, University of Technology Sydney, Australia

December 31, 2012

Abstract

Combined mining is a technique for analyzing object relations and pattern relations, and for extracting and constructing actionable complex knowledge (patterns or exceptions) in complex situations. Although combined patterns can be built within a single method, such as combined sequential patterns by aggregating relevant frequent sequences, this knowledge is composed of multiple constituent components (the left hand side) from multiple data sources which are represented by different feature spaces, or identified by diverse modeling methods. In some cases, this knowledge is also associated with certain impact (influence, action or conclusion, on the right hand side). This paper presents a high-level picture of combined mining and the combined patterns from the perspective of object and pattern relation analysis. Several fundamental aspects of combined pattern mining are discussed, including feature interaction, pattern interaction, pattern dynamics, pattern impact, pattern relation, pattern structure, pattern paradigm, pattern formation criteria, and pattern presentation (in terms of pattern ontology and pattern dynamic charts). We also briefly illustrate the concepts and discuss how they can be applied to mining complex data for complex knowledge in either a multi-feature, multi-source, or multi-method scenario.

1 Introduction

In this paper, we introduce the concept of combined (pattern) mining. Combined mining is proposed for handling the complexity of employing multi-feature sets, multi-information sources, constraints, multi-methods and multi-models in data mining, and for analyzing complex relations between objects or descriptors (attributes, sources, methods, constraints, labels and impacts) or between identified patterns during the learning process. Combined patterns may be formed through the analysis of the internal relations between objects or pattern constituents obtained by a single method on a single dataset, for instance, combined sequential

*This work is sponsored in part by two Australian Research Council Discovery Grants (DP1096218 and DP130102691) and an ARC Linkage Grant (LP100200774).

patterns formed from analyzing the relations within a discovered sequential pattern space.

With the exception of object and pattern relation analysis, which is a very new topic in the data mining community, many approaches and algorithms are available in the literature on other aspects of the above combinations. The main contribution of combined mining is that it enables the extraction, discovery, construction and induction of knowledge which consists not simply of discriminant objects but also of interactions and relations between objects, as well as their impact. This is called *actionable complex patterns*, because they reflect pattern elements and relations, which form certain pattern structures and dynamics, and indicate decision-making actions.

Combined mining provides an overall solution for meeting the challenge of mining complex knowledge in complex data [8]. It also substantially builds upon other individual approaches such as conceptual inductive learning [22, 23] and inference, generalization, aggregation and summarization [20, 45], in order to integrate them with data-driven knowledge discovery from complex environments. Specifically, pattern relation analysis augments the following areas: knowledge representation and reasoning [28], inductive learning [22], semantic and ontological engineering [30], pattern theory [16], and pattern language [1].

This paper will not discuss specific combined mining techniques, which are available from our other references [7, 9, 10, 39, 40, 41, 42, 43, 44, 45]; rather, it is intended to present an abstract high level picture of combined mining by addressing some very important issues in combined mining and general data mining and machine learning. This includes concepts, combined pattern formation, presentation and applications, and evaluation. In particular, the paper examines concepts, mechanism design and the representation of combined pattern formation criteria, pattern relations, paradigms, and structures. These aspects have not been discussed in the literature to date, yet they expand the potential of combined mining into a much bigger space, including non-structural patterns such as the conditional probability-based pattern merger. This paper represents an abstraction of the existing specific designs and methods. It is expected that readers will acquire a high level understanding of combined mining through reading this paper and refer to specific papers for details. We aim to motivate researchers to consider fundamental issues in data mining, including object relation analysis, pattern structure, pattern relation, paradigm, ontology and evolution. This may enable readers to access the great potential of using combined mining for other complex problem-solving.

It is worthwhile clarifying the difference between combined mining and other relevant techniques and purposes. First, the actionability of patterns [8, 11] and how to discover actionable patterns [26, 34, 27] are beyond the scope of this paper, even though we argue that combined mining intends to deliver actionable results. Readers who are interested in these topics can find many references, such as those on domain driven data mining [4, 8, 5] and action rule discovery [26, 27, 34]. Second, combined mining tackles a range of different scenarios from multiple sets of features [18] to multiple sources and multiple methods as necessary for problem-solving [7]. Third, typical inductive learning techniques such as inductive learning for rule generalization [35, 22] are designed to generalize symbolic descriptions from examples and observations, with or without exceptions, discover alternative hypotheses, and handle meta-values. While combined mining can be used for conceptual learning, it mainly tackles the complexity of pattern discovery environments. Finally, combined mining could be treated as a hybrid data mining method if it were to be applied for multi-method combined pattern mining to generate many different combinations of different data mining approaches [17, 24, 25, 27]. The combination of classification with association rule mining, for instance, produces an associative classifier which may be built on an unordered dataset to predict online shopping fraud.

In addition, while ontology and context may need to be incorporated or considered in combined pattern representation, presentation and learning, the purpose of combined mining is not the same as that of ontology mining and context-dependent knowledge discovery. As opposed to pattern theory and pattern language, which ambitiously aims to provide a general mathematical and algebraic framework for representing and inferring patterns as structures and architectures regulated by rules and combinatory operations, combined mining focuses on discovering and extracting more meaningful patterns from complex data.

The remainder of the paper is organized as follows. Section 2 introduces the relevant background concepts and related work, especially the need to discover and induce complex but actionable knowledge. Basic concepts of combined patterns are introduced in Section 3. In Section 4, we discuss the key issues in combined pattern formation. Pattern presentation is discussed in Section 5. Applications and an evaluation of combined mining, including case studies, are outlined in Section 6. Section 7 discusses the challenges and prospects of combined mining. The paper concludes in Section 8.

2 Background

In this section, we discuss our conceptual understanding of the word ‘pattern’ and discuss the trend of mining complex data for complex knowledge.

2.1 Pattern As a General Concept

What do we really mean by a pattern? Can we define it in strictly logic terms? Such questions have been studied by colleagues for decades, such as the general pattern theory [16] and pattern language [1]. From the very high level perspective, all kinds of knowledge from data are to be mined for patterns. More specifically, in this paper, ‘pattern’ covers both patternable (i.e. the pattern we usually refer to pattern mining) and non-patternable (i.e. an exception or unusual pattern) findings from data. Hereafter ‘pattern’ covers ‘outlier’ and ‘exception’.

A pattern is a combination of relevant *descriptors* (attributes, called internal elements in this paper) associated with certain *relations* (for instance, frequency, classifier or probabilistic distribution) and *constraints*. In the existing pattern discovery and exception mining, a resultant pattern is an individual outcome that has one of the pattern structures detailed below:

Type I: {Antecedent}, a combination of attributes, in which a pattern is composed of a collection of internal elements in the underlying problem. Typical examples include frequent pattern mining, association rules and clustering. Unsupervised learning usually delivers different combinations of underlying variables.

Type II: $\{\{Antecedent\}\ \{connective\}\ \{Consequent\}\}$, or $\{\{Premise\}\ \{connective\}$ $\{Conclusion\}\}$, in which a series of attributes are connected with one another to form the antecedent (or premise) and are then associated in terms of certain *connectives* with (or lead to) an additional *consequent* (or *conclusion*, *Impact*). Supervised learning such as classification usually delivers outcomes associated with supervised indicators (e.g., class labels). The combination of unsupervised learning with supervised learning [38] also results in this type of deliverable, such as the frequent pattern-based classifier and classification rules. Emerging discussions on high utility pattern mining also fall into this category.

Type III: $\{\{Antecedent\}|\_\{condition\}\ \{connective\}\ \{Consequent\}\}$, in which the occurrence of an antecedent connects consequent results from certain *conditions*. A condition may be an exception or exclusion of some attributes, a constraint, or a certain context. For instance, a mobile preference pattern $\{\{business\ managers\}|\_\{between\ 20\ and\ 35\ years\ old\}\ \{\{more\ likely\}$ $\{touse\ iphone\ rather\ than\ Blackberry\}\}$.

We call the source data and pattern outcomes *Type I* if they consist of internal elements only; *Type II* data and patterns include additional external conclusions or impacts. If a source or pattern is context, condition or constraint dependent, then it falls into Type III.

Type II and Type III patterns are clearly much more informative and actionable [4, 10, 8] than Type I patterns, because they consist of external information (the impact indicator and/or condition) which is in addition to the internal descriptors. While it is often costly or even impossible to obtain the external information, domain and background knowledge driven or partial label based semi-supervised learning is highly valued for learning Type II and Type III patterns on Type I Data. This paper will mainly discuss pattern combination aspects, pattern structures, relations and paradigms for these three types of data and patterns.

As we will discuss with regard to pattern representation (Section 4) and presentation (Section 5), a pattern in this sense is a knowledge element, but not all knowledge elements form patterns. We will discuss the applicability of knowledge representation and reasoning to pattern representation and inference (see Section 6.1).

2.2 Trend of Discovering and Inducing Complex but Actionable Knowledge

Mining complex data for complex knowledge has been discussed as a major challenge for next-generation data mining [8] such as domain driven data mining for actionable knowledge discovery [4, 5, 8] and action rule mining [26, 34, 27]. Another trend is to discover interesting patterns, and further induce meaningful concepts [22] from the learned patterns into logic-style deliverables indicating decision-making actions [8]. With increases in data size, complexity, value recognition and strategic use, there is a clear need to mine complex big data for complex but actionable knowledge.

The identified individual patterns of existing approaches confront, but are not limited to, the key issues that patterns are not sufficiently informative and are often not actionable.

From the information perspective, the resultant patterns often have simple structures, and/or are simple combinations of pattern elements. They often ignore dependency between features and objects, overlook coupling relations between attributes, attribute values, objects and patterns, and neglect the impact of patterns. Consequently, although the findings are interesting, they are not sufficiently informative to support decision-making actions (we call the patterns actionable if they can inform or support decision-making actions).

An example in mining debt-related taxpayer lodgment and payment behaviors shows that the use of decision tree and rain forest leads to interesting rules that indicate by whom and in what circumstances, a taxpayer's debt is collectable. However, some of the corresponding debts are never recoverable. This indicates that the rules are not actionable, because the debts identified by the patterns cannot be collected. This example shows the difference between pattern interestingness and pattern actionability [8, 4, 11].

There may be many and varying reasons for this [4, 5, 11]. In this paper, we attribute the weakness of knowledge actionability of findings resulting from existing methods to the oversight or limited consideration of the following aspects: feature interaction, pattern interaction, pattern dynamics, pattern impact, pattern relation, pattern structure, selection criteria, and pattern presentation. These aspects raise some fundamental issues in mining complex data for complex but actionable knowledge. We will discuss them briefly in the following sections by introducing the methodology of combined mining.

3 Combined Pattern Concepts

3.1 Preliminary

Combined patterns are formed by one or more of the following pattern elements:

- Data sources $\mathscr{D}$, which are the multiple data feeds that arise from the features of the patterns mined
- Feature sets $\mathscr{F}$, which are extracted or constructed from data sources, and form the constituent components of the left hand side of patterns
- Modeling methods $\mathscr{R}_1$, which correspond to certain objective functions, and generate the constituent components of the patterns on either one or more data or feature sets, interchangeably. $\mathscr{R}_1$ also represents a set of coupling relations that associate the features
- Pattern impact(s) $\mathscr{I}$, in some cases (Type II and Type III patterns below), one to many impacts are associated with the constituent components
- Impact coupling relations $\mathscr{R}_2$, which capture the relations between patterns and impacts.

The engagement of the above elements and their instantiations in respective combined patterns are based on the specific problem and its context, as well as the analytical and business objectives. Consequently, combined pattern mining is a process of producing actionable knowledge (patterns or exceptions)

1) with multiple constituent components forming the pattern antecedent on the left hand side of the patterns, extracted from multiple data sources, represented by different feature spaces, or identified by diverse modeling methods,

2) for Type II patterns, consisting of pattern consequent on the right hand side of the patterns, namely pattern impact(s),

3) for Type III patterns, incorporating context (or condition, constraint) on the pattern antecedent,

4) a pattern mining method set (also representing a coupling relation set 1 $\mathscr{R}_1$) which associates multiple components with one another,

5) for Type II and Type III patterns, a coupling relation set 2 $\mathscr{R}_2$ which associates multiple patterns with respective pattern impact(s).

Definition 1 (Type I Combined Pattern) *Type I combined patterns $\mathscr{P}_1$:*

$$\mathscr{P}_1 : \mathscr{R}_1(\mathscr{F}(\mathscr{D})) \tag{1}$$

where patterns $\mathscr{P}_1$ are identified through data mining methods $\mathscr{R}_1$ deployed on features $\mathscr{F}$ from dataset $\mathscr{D}$. Any pattern instance p_i ($p_i \in \mathscr{P}_1$) is also called an atomic pattern*.*

Definition 2 (Type II Combined Pattern) *Type II combined patterns $\mathscr{P}_2$:*

$$\mathscr{P}_2 : \mathscr{R}_2(\mathscr{R}_1(\mathscr{F}(\mathscr{D})), \mathscr{I}) \to \mathscr{I} \tag{2}$$

where patterns $\mathscr{P}_2$ are identified through data mining methods $\mathscr{R}_1$ deployed on features $\mathscr{F}$ from dataset $\mathscr{D}$, the patterns are also associated with impact $\mathscr{I}$ through relations $\mathscr{R}_2$.

We see that Type I combined patterns are special cases of Type II patterns.

$$\mathscr{P}_2 : \mathscr{R}_2(\mathscr{P}_1) \to \mathscr{I} \tag{3}$$

Definition 3 (Type III Combined Pattern) *Type III combined patterns $\mathscr{P}_3$:*

$$\mathscr{P}_3 : \mathscr{R}_2(\mathscr{R}_1(\mathscr{F}(\mathscr{D})), \mathscr{I})|_{\mathscr{C}} \to \mathscr{I} \tag{4}$$

where patterns $\mathscr{P}_3$ are identified in the same way as Type II patterns $\mathscr{P}_2$ but under condition $\mathscr{C}$.

We see that Type II patterns are generalized Type III combined patterns as they do not consider the condition of pattern existence.

$$\mathscr{P}_3 : \mathscr{C}(\mathscr{P}_2) \to \mathscr{I} \tag{5}$$

Accordingly, the atomic Type I combined patterns appear as the left hand side of the atomic Type II and Type III patterns.

3.2 Pattern Combination Aspects

The word "combined" in combined pattern mining refers to either one or more of the following combination approaches.

- Combinations of multiple sets of the same pattern elements: In this case, a *single combined pattern* consists of different components of the same pattern element, for instance, attributes in different feature sets or data sources.

  Example 1 *Students (c_1) living in rural suburbs (demographic data) as well as (c_2) with a low level of subject engagement activities (learning data) are more likely to (c_3) fail a subject (impact).*

- Combinations of atomic patterns: Here a *compound combined pattern* includes two to many atomic patterns identified by either the same method or several methods, or several atomic patterns from the same method but associated with different impact levels or types.

  Example 2 *p_1: a student who frequently accesses library resources (atomic frequent pattern 1 on library data) is more likely to pass a subject (impact), and p_2: a student who has frequent access to university online websites (atomic frequent pattern 2 on online data) is also more likely to pass a subject (impact). p_3: A student who frequently accesses the library as well as online resources has an even higher likelihood of passing the subject.*

As a result, a combined pattern appears as either a single pattern or a collection of atomic patterns. Example 1 is a single Type II pattern that combines features from two data sources. Example 2 illustrates a compound combined pattern with the combination of Type II atomic patterns. We will further discuss the pattern structures in Section 4.4.

The combination of elements of data sources $\mathscr{D}$, feature sets $\mathscr{F}$, modeling methods $\mathscr{R}_1$, pattern impact(s) $\mathscr{I}$, and impact coupling relations $\mathscr{R}_2$ may take many different forms. The main aspects considered for the merger of atomic patterns include: Feature Interaction, Pattern Interaction, Pattern Dynamics, and Pattern Impact.

- Feature interaction: although features are generally treated as being independent from one another in classic statistics, data mining and machine learning analysis, interactions are actually embedded between features in terms of certain relations. Different types of dependency determine various interaction forms, and result in diverse pattern outlets.

  Example 3 *In educational data analysis, a student's cultural background (such as birth country and language) is highly associated with the suburb in which s/he lives. Such a correlation may need to be considered in learning their relations with academic performance rather than treating them independently and equally.*

- Pattern interaction: refers to the interactions between pattern constituents for single combined patterns or between atomic patterns. The interaction is embodied through some form of relation, whether syntactic, semantic and/or mathematical.

  Example 4 *In Example 1, the pattern consists of two constituents c_1 of demographical characteristics and c_2 about learning behaviors. They come from different data sources, which capture two aspects of students. In the pattern, they are of a conjunction relation; further c_1 and c_2 form a causal coupling with c_3 as the effect, represented by $(c_1 \wedge c_2) \rightarrow c_3$. In Example 2, atomic pattern p_3 is a merger of patterns p_1 and p_2, denoted as $p_3 = p_1 \wedge p_2$, with the conjunction relation between p_1 and p_2.*

- Pattern dynamics: refers to the change of pattern elements, structures and/or pattern relations, which may be caused by the underlying elements, namely: data sources, features and/or impacts.

  Example 5 *Example 3 illustrates a pattern that indicates a relation between a student's cultural background and the suburb in which they live. We also* *find that a student moving from his/her birth country to live in another country may be indicative of a high risk of failing fundamental subjects. This pattern can be treated as a derivative on top of Example 3, which indicates the change of pattern constituent from a birth country to a different country of residence.*

- Pattern impact: refers to the influence of a pattern associated, for instance with certain business outcomes or concerns, represented as a label, risk level or dollar value, or a nominal influence defined for certain purposes, such as utility.

  Example 6 *In Example 2, failing or passing a subject is an impact triggered by or associated with the corresponding learning behaviors.*

- Pattern context: refers to the environment in which a pattern is discovered or extracted, representing where or under what condition or constraint a pattern exists or occurs.

  Example 7 *In Example 1, living in rural suburbs can be treated as a context of the pattern.*

  Pattern context plays an important role in pattern combination. In Section 4, context (including condition and constraint) may be involved in every aspect of pattern formation criteria, relation, paradigm and structure.

4 Combined Pattern Formation

4.1 Pattern Formation Criteria

In Section 3.2, we discussed the main factors of pattern combinations, including feature interaction, pattern interaction, pattern dynamics and pattern impact. These factors form the foundation of combined pattern formation as well as pattern evolution. In the selection of combined patterns, we need to consider these factors.

The value of discovered and constructed combined patterns can be evaluated in terms of two perspectives: technical significance and business impact [11]. We will discuss the business impact of a combined pattern in Section 6.

The technical significance of a pattern is equivalent to the so-called 'interestingness'. To reflect the relations between patterns (see Section 4.2), the technical criteria can be specified in terms of the following three key perspectives: pattern similarity, pattern dissimilarity, and pattern dependence.

- Pattern similarity: this is to analyze whether two or more atomic patterns share enough similarity to form a pair or cluster of similar patterns. The similarity measures may be further specified in terms of the feature set, interaction, relation, distance, density, shape, structure or impact between constituent patterns.

  Example 8 *In an online banking business, p_4: frequent testing of different customer accounts where all attempts fail within a very short time period is highly likely to indicate an ID takeover fraud. In other cases, p_5: a customer frequently testing different accounts where the last try succeeds within a very short time period is also highly likely to indicate an ID takeover fraud. Here, patterns p_4 and p_5 form a combined pattern through sharing a common* *structure, i.e. frequently testing different accounts within a very short time period, and ID takeover fraud.*

- Pattern dissimilarity: refers to the difference between constituent patterns which are comparable. While we focus on difference, the comparability of atomic patterns is also very important, in order to connect them into a combined pattern. The comparability may lie in the aspects of the feature set, interaction, relation, distance, density, shape, structure or impact between patterns.

  Example 9 *Taking Example 8 in online banking fraud detection, if p_6: consists of p_4 associated with no dollar loss, while p_7: is composed of p_5 with high dollar loss, then these two new patterns form a combined pattern connected through a common structure, i.e. different, frequently tested accounts within a very short time period and ID takeover fraud, but they are highly dissimilar to each other since they lead to very different impacts.*

- Pattern dependence: indicates that the constituent patterns share some dependent relations in the aspects of the feature set, interaction, relation, structure, or impact.

  Example 10 *In an online banking business, if p_5: a customer frequently tests different accounts but with the last try succeeds within a very short time period, and another pattern p_8 consists of a further condition that the IP address of the customer falls in a list of suspicious countries on top of p_5, then p_5 and p_8 form a combined pattern, and p_8 is dependent on p_5.*

The similarity, dissimilarity and dependence may be embodied through explicit or implicit relations. Explicit relations may appear as structural or semantic operators, which can be obtained from domain knowledge. Implicit relations appear in mathematical terms - particularly statistical functions - which are usually learned from the data. In addition, patterns may be similar, dissimilar or dependent under certain conditions or within certain contexts, which forms conditional similarity, dissimilarity and dependence. This needs to be considered in pattern relation, paradigm and structure formation, learning, representation and presentation.

4.2 Pattern Relations

Pattern relations refer to the couplings between patterns and between pattern elements (constituents). There are different aspects to the evaluation of pattern relations; for instance, in terms of different relations and structures, such as structural relation, semantic relation, dependency relation, and probability relation. In addition, the coupling relations between patterns can also be explored from temporal, inferential and party-based aspects [31]. The proper representation, reasoning and checking of such relations may be beyond that of basic Boolean operations [15].

Below is listed a few possible relations (and the corresponding connectives) between patterns, or between pattern constituents, by considering multiples of the above aspects:

- *Serial coupling*: The constituent patterns are with sequential order represented by operator ',', e.g., $\{a, b, \cdots, n\}$.
- *Causal coupling*: There is causality relation between the constituent patterns represented by operator '$\rightarrow$', e.g., $\{a \rightarrow d\}$.

- *Synchronous coupling*: The constituent patterns occur in concurrency, represented by operator '$\|$', e.g., $\{b\|c\|, \cdots, \|k\}$.
- *Conjunction coupling*: The constituent patterns take place together, represented by operator '$\wedge$', e.g., $a \wedge b$.
- *Disjunction coupling*: At least one of the constituent patterns must happen, represented by operator '$\vee$', e.g., $d \vee k$.
- *Exclusive coupling*: Different patterns occur on a mutually exclusive basis, represented by operator '$\oplus$', e.g., $\{d\oplus, \cdots, \oplus j\}$.
- *Dependent coupling*: Some patterns have required dependents such as prefix or postfix components, represented by operator '$\Rightarrow$', e.g., $\{c \Rightarrow (b \vee m)\}$.

Taking the causal sequential relation as an example, there may be many different relation combinations that exist between sequential patterns or sequential elements; for instance:

- *Positive enabling* relation, represented by $a \to b$, in which element (or sub-pattern) a positively enables (followed by) b;
- *Negative enabling* relation, represented by $\neg a \to b$, in which element (or sub-pattern) a negatively enables b, or the non-occurrence of a positively enables b;
- *And split* relation, represented by $a \to (b \wedge c)$, where elements b and c must be conducted once a has occurred;
- *Or split* relation, represented by $a \to (b \vee c)$, where either b or c happens once a has occurred;
- *And join* relation, represented by $(a \wedge b) \to c$, where c happens only if both a and b have been conducted;
- *Or join* relation, represented by $(a \vee b) \to c$, where c happens only if either a or b has been conducted.

Considering the pattern relations in combined pattern mining discussed above, many novel structures of patterns can be extracted or constructed (see Section 4.4). Below, we present some examples of combined patterns constructed by the above pattern relations.

Example 11 *Figure 1 illustrates four types of complex compound patterns in high utility sequence analysis (where $--$ represents a pattern coupling relation). For instance, (a) shows the* and split *hierarchical relation between a, b, c, d, indicating that a sequential element a associated with element b and then c has low utility level u_l; while a associated with element b and then d has high utility level u_2. The elements b, c and d connecting with impacts u_1 and u_2 form a complex* and split *hierarchical relation.*

4.3 Combined Pattern Paradigms

The combination of the aforementioned pattern elements and combination factors in terms of specific selection criteria and pattern coupling relations will contribute to different pattern paradigms. Below, we discuss three combined pattern paradigms: similar combined patterns, dissimilar combined patterns, and dependent combined patterns.

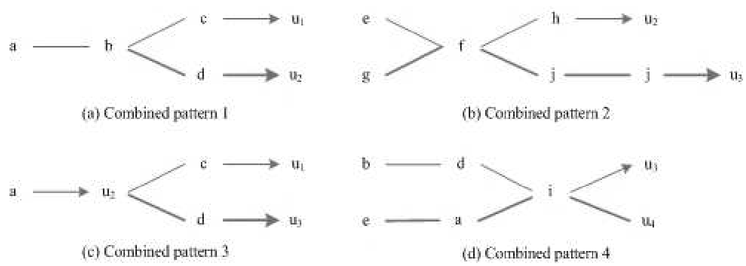

Figure 1: Complex Combined Patterns

Scenario 1 *(Similar Combined Patterns) The constituent patterns in a combined pattern share some similarity in feature, interaction, relation, structure or impact. As discussed above, similarity is measured through distance, density, shape, structure or relation that is specific to certain pattern mining methods.*

Typical similar combined patterns include the combinations of frequent patterns, high utility patterns, clusters, and classes.

Scenario 2 *(Dissimilar Combined Patterns) The constituent patterns in a combined pattern have some dissimilarity in feature, interaction, relation, structure or impact, as measured by distance, density, shape, structure or relation specific to certain pattern mining methods.*

Typical dissimilar combined patterns are contrast-based combined patterns, in which two atomic patterns are associated with opposite impacts (labels).

Scenario 3 *(Dependent Combined Patterns) Also called* Conditional Combined Patterns, *in which one pattern forms the precondition of another. The condition may come from the aspect of feature, interaction, relation, structure or impact.*

Three types of dependent combined patterns are *incremental patterns*, *decremental patterns*, and *conditional probability patterns*.

Instance 1 *(Incremental Patterns) Also called* Prefix Combined Patterns, *in which any two neighboring atomic patterns in the combined collection form an incremental relation, namely pattern $i+1$ sharing some incremental part of features, pattern elements, structures or impacts on top of pattern i.*

Example 12 *The constituent pattern 1 shows that a student living in a rural suburb has a high risk of failing a subject. Constituent pattern 2 indicates that if the student lives in a rural suburb, and the suburb is a low socio-economic area, then s/he has a very high likelihood of failing the subject.*

Instance 2 *(Decremental Patterns) Also called* Postfix Combined Patterns, *where the constituent pattern i consists of an additional part of features, pattern elements, structures or impacts compared to pattern $i+1$.*

In Example 12, the constituent patterns 2 and 1 form a decremental partnership.

In incremental and decremental patterns, some atomic patterns serve as the underlying patterns; the immediate neighboring patterns are the derivative form of them. For example, in [9], we specify the *underlying-derivative* combined patterns.

Incremental and decremental patterns are more about pattern structure dependence. Readers may refer to [9] to access details about how incremental and decremental frequent patterns and frequent sequences are formed in social security data. In practice, such structural relations are often hard to detect in large data; this is even challenging which structural relations appear to be implicit. This is particularly difficult if we do not have the hypothesis of what kind of incremental or decremental relations exist in the data. For patterns appearing in trees, graphs or other unstructured formats, it would be much more difficult to learn and extract such relations.

In addition, conditional probability patterns cater for implicit dependency between atomic patterns when a probabilistic model fits a dataset.

Instance 3 *(Conditional Probability Patterns) The constituent patterns form a conditional probability relation in terms of features, elements, interaction, structure or impact.*

The conditional probability relation may be embodied through certain statistical relations and functions. For instance, a chain of states may affect one another, which can be modeled according to the Markov assumption.

4.4 Combined Pattern Structures

The extraction and construction of combined patterns contribute to different types of patterns and pattern structures. The structures of combined patterns depend on many aspects, including pattern mining methods, feature interaction, pattern interaction, pattern dynamics, pattern impact, combination factors, and pattern selection criteria. Different combinations of these aspects will lead to a variety of combined pattern structures and novel pattern types. Within the universal space of combined patterns, a feasible direction for creating combined patterns is to combine atomic patterns in terms of one to two aspects by following the three pattern combination paradigms discussed in Section 4.3. Below, we discuss a few scenarios.

Scenario 4 *(Single Combined Patterns) A single combined pattern consists of different elements within one pattern presentation. The pattern is a mixture of components such as features, elements, impacts and relations from different sources or methods.*

For instance, an associative classifier generates single combined associations with labels attached.

Scenario 5 *(Compound Combined Patterns) A compound combined pattern is composed of a collection of patterns connected through particular similar or dissimilar relations or functions in terms of features, elements, structures, impacts or other aspects.*

Incremental patterns and decremental patterns are compound combined patterns. Based on the structure of the compound, we may have basic compound patterns and complex compound patterns.

Instance 4 *(Basic Compound Patterns) A basic compound pattern consists of a set of atomic patterns connected by one simple combination relation, function or strategy.*

For instance, a set of RFID purchase transaction associations has the same level of customer value.

Instance 5 *(Complex Compound Patterns) A complex compound pattern consists of a set of atomic patterns connected by more than one combination relation, function or strategy. The constituent patterns may form a complex structure, for instance, a hierarchical relation between them.*

Based on the number of atomic patterns in a compound pattern, we have pair compound patterns and cluster compound patterns, for short pair patterns and cluster patterns.

Instance 6 *(Pair Patterns) A pair combined pattern consists of two constituent patterns, which share some similarity, dissimilarity and dependence in feature, interaction, relation, structure or impact.*

Based on the similarity and dependence between the constituents, different types of pair patterns may be extracted or derived: similar pairs in which the sub-patterns share some similarity, dissimilar pairs in which they appear very differently but have some dependent relation. For similar pairs, incremental and decremental pairs can be constructed with a dependent coupling relation between constituents. Dissimilar pairs may present as parallel pairs in which the atomic patterns are positively or negatively connected in terms of *And* relation, or contrast pairs with the components negatively or exclusively related to an impact (including label).

Example 13 *Given behavior elements a,b,c,d, the impact i, a parallel similar pair is formed as $\{a,b,c||a,b,d\}$, where a,b,c and a,b,d are two sub-patterns; a contrast dissimilar pair is as $\{a,b,c \rightarrow i||a,b,d \rightarrow \neg i\}$, in which a,b,c is associated with positive impact while a,b,d with negative impact.*

Instance 7 *(Cluster Patterns) A cluster combined pattern consists of more than two constituent patterns, which share some similarity, dissimilarity and dependence in feature, interaction, relation, structure or impact.*

Similar to pair patterns, atomic patterns may be connected in terms of different pattern relations, and many appear as an incremental cluster or decremental cluster.

Both pair patterns and cluster patterns may take the form of basic or complex compound patterns. In the real world, atomic patterns may be connected in very complicated relations and form hybrid patterns.

Instance 8 *(Hybrid Patterns) A hybrid combined pattern consists of more than two constituent patterns, which are connected in terms of complex pattern relations such as exclusive and precedence relations, and structures with components linked together in terms of structural relations such as the And Split, Or Split, And Join and Or Join.*

Figure 1 illustrates a few hybrid combined patterns.

Considering the influence of compound patterns and the negation of pattern elements, we can have *positive combined patterns* and *negative combined patterns*, as well as *single impact combined patterns* and *multiple impact combined patterns*.

Instance 9 *(Positive Combined Patterns) A positive combined pattern only consists of constituent patterns connected in a positive relation, in aspects of similarity and dependence in feature, interaction, relation, structure or impact.*

For instance, $\{a,b,c \rightarrow i\}$ is a positive compound pattern with a single impact.

Instance 10 *(Negative Combined Patterns) A negative combined pattern consists of at least one negation relation, on either the pattern elements or pattern impact side. Negative relations on the dissimilarity and dependence in feature, interaction, relation, structure or impact exist in negative patterns.*

For instance, $\{a, b, d \rightarrow \neg i\}$ is a negative combined pattern with a single impact.

5 Combined Pattern Presentation

The representation and presentation of combined patterns are very much dependent on the relations and structures of pattern combinations. For syntactic pattern relations and structures, tools in ontology engineering [30] can be used, with temporal logic [2] to represent the logic relations between elements. This is also applicable for patterns with explicit structures and relations. Those patterns with implicit relations and structures are more suitable for probabilistic relations and formulas. Below, we introduce pattern ontology and the pattern dynamic chart as two of the presentation tools.

5.1 Pattern Ontology

Pattern ontology is motivated by the frequently used concept of design ontology pattern (ODP) [46] for designing reusable and high-quality templates and software in the architecture field [1]. In the data mining community, ontology has been used in different situations: ontology for representing business problems and data mining input and results, and mining ontology data. In this paper, we bring the ODP concept for combined pattern mining to represent patterns in terms of ontology and to infer pattern relations during pattern dynamics. We can then define a pattern ontology to enable the representation and reasoning of patterns. In pattern ontology language, a combined pattern is a pattern instance; each constituent pattern can be represented as an ontological item or element; the relations between constituents can be represented in terms of ontological relations and operators in the first-order temporal logic.

Ontological tools including items, operators, classes, relations, instances and axioms are used to represent pattern elements and pattern relations in combined patterns. Axioms can be defined to represent pattern relations. For instance, $\{p_1 \rightarrow p_2\}$ represents that pattern p_1 causes pattern p_2, $\{p_1 \wedge p_2\}$ represents that patterns p_1 and p_2 take place together, $\{p_1 \vee p_2\}$ represents that patterns p_1 or p_2 exist. In impact-oriented hybrid combined patterns, we may see that multiple patterns are associated with the same impact (as shown in [9, 7, 41, 42]).

The concept of pattern ontology can be very valuable in representing and reasoning about combined patterns. It can also play an important role in converting patterns into knowledge that can bridge the gap between data modelers and business analysts. For instance, business rules may be relatively easily created on top of ontologically represented patterns.

5.2 Pattern Dynamic Chart

Because of the intrinsic relations between constituent patterns, it would simplify the user's life if combined patterns could be presented in an aligned way. This motivated us to develop a pattern chart to engage the combinational aspects of a combined pattern, including constituents, relations, and impacts. If other dimensions such as pattern dynamics and pattern change can be embodied in the chart, it will help users to understand the evolution of the family of combined patterns and their impact change. This will help the understanding of underlying business, and assist in detecting and predicting the crucial changes of the underlying business for early or online intervention.

To give an example, the attachment of impact with cluster patterns may make the analysis and understanding very complicated. Cluster patterns may be used to explore pattern evolution and influence dynamics by converting them into a *pattern impact relation chart*, (or *pattern evolution chart*), in which the horizontal axis represents the pattern growth, and the vertical axis represents the pattern influence dynamics. Figure 2 illustrates the evolution of a cluster pattern consisting of eight elements and five impact levels. In [9, 7, 42], we introduce a pattern dynamic chart to present a set of behavior changes and their associated impact on different performance metrics such as confidence, lift, contribution and the impact of each behavior that contributes to overpayments in social security.

6 Applications and Evaluation

In this section, we discuss the process and application of deploying combined mining and the evaluation of identified and constructed combined patterns.

6.1 Deployment of Combined Mining

While it is difficult, if not impossible (and perhaps even pointless), to provide general solutions for deploying combined mining in a general sense, here we summarize some of our experiences in generating combined patterns. Combined patterns may be identified through a one-off or multi-step process. A typical multi-step combined mining process works as follows:

1) Business understanding to identify the corresponding pattern combination factors available in the particular problems and the aspects to be considered for specific business goals;
2) Factor/aspect analysis to select/develop proper metrics to evaluate the necessity and contribution of each factor and their respective roles in terms of business goals, and to identify the factors and aspects essential for problem-solving;
3) Atomic patterns are discovered in each factor/aspect; they may be sorted in terms of relevant metrics.
4) The relations between and within the atomic patterns are analyzed;
5) Associated atomic patterns are combined (merged) according to relations by pattern merging methods and merging purposes (often dependent on business problems and objectives);
6) The merged patterns may be further converted and/or presented into deliverables that satisfy specific purposes.

We further discuss strategies and issues in deploying combined mining, based on our elementary experiences in the real world [7, 8, 9, 10, 39, 40, 41, 42, 43, 44, 45].

- Aspect: Several combination aspects, from source, feature, pattern, method and dynamics to impact, have been discussed in Section 3.2, along with their need to be evaluated in terms of the underlying business goals, and their roles and importance in solution seeking and feasibility in the practical sense. If many factors are involved, they need to be prioritized and only the really relevant and essential factors may be considered. For instance, the coupling relationship between features (i.e. relation learning [33, 32] and coupled object analysis [6, 29]) may be computationally too costly to be considered in feature interaction.
- Framework: As discussed in [10], there may be multiple different frameworks to support combined mining. The most cost-effective framework is selected for a given problem, by considering the feasibility, effectiveness and efficiency of implementing it in problem-solving.
- Representation and reasoning: This may involve representation and reasoning [14, 3] of combined patterns, pattern relations and dynamics. Besides the rule-based logic representation [28, 15] of pattern constituents and their relationships, other methods including ontology and semantic web-based representation and reasoning [30], model-based reasoning [28], graph-based representation and reasoning [12] and statistics-based representation and inference [37, 13, 19] may be very useful in describing combined patterns from specific domains.
- Relation: While both basic logic relationships and other types of relationships, as discussed in Section 4.2, may exist in a pattern space, the most relevant relations between patterns rely mainly on the data type and business goals. One may focus on those relations that are of particular interest and extractable while ignoring those that are weak or not representable.
- Paradigm: Several types of combined pattern paradigms are discussed in Section 4.3. They can be used to guide the construction of combined patterns for a specific business problem. One to several paradigms may exist in the data, which can then be further represented in terms of corresponding pattern structures discussed in Section 4.4. While not all introduced paradigms and structures may exist in a specific dataset, specific combined pattern paradigms and structures may be identified and designed for a target system.
- Evaluation: As we will discuss in Section 6.2, both the technical significance and business value of a combined pattern needs to be assessed. Ideally, a combined pattern is actionable if it is not only technically interesting but also of great business value.
- Presentation: Limited research is available on designing proper mechanisms to present combined patterns. From the user perspective, combined patterns may be converted into business rules [4, 7] in terms of certain representation specifications, with the consideration of business impact and pattern dynamics. Richer pattern presentation language and tools are essential in delivering business-friendly patterns.
- Computational complexity: We have not discussed this issue in this paper, although it is crucial in a case study. Often, strategies and designs in catering for multiple sources, features, methods, constraints, relations and impacts

may be compromised to the point where the discovery and delivery of actionable combined patterns are are not possible. Typically, only one to two aspects may be catered for in the process.

- Conflict resolution: Tradeoff between comprehensive design and acceptable performance may be necessary so that the deliverables can be produced within an affordable scope. This may also be reflected in pattern evaluation and selection: an appropriate selection strategy is essential if none of the patterns satisfy both technical and business performance [11].

6.2 Evaluation

Thanks to the objectives of combined mining of delivering actionable patterns for smarter decision-making, the evaluation of combined patterns becomes a critical issue. As we discussed in [11], the actionability is the key concept we use for measuring the value of a combined pattern for both technical significance and business performance. Actionability refers to the quality and power of *actionable knowledge discovery and delivery* (AKD) outcomes for effective decision-making and problem-solving. The power to work is an optimal outcome and objective from AKD through the best integration of six core dimensions: Problem, Data, Environment, Model, Decision and Optimization [4]. As a result, AKD can be viewed as a six dimension-based optimization process [11, 8]:

$$AKD ::= optimization(problem, data, environment, model, decision) \quad (6)$$

Actionability computing [4] is thus likely to become an interesting topic to explore. We need to evaluate and analyze actionability on problem, data, environment, model, decision and optimization [4]. In practice, actionability may be interpreted in varying terms: for instance, autonomy, deliverability and transferability, dependability, explainability and interpretability, impact, repeatability, semantics and understandability, and trust. The resultant actionable knowledge (patterns) can lead to effective actions for better results (decision, answer, conclusion, etc.) [4].

Often it is necessary to prioritize the main objectives of deliverables, rather than checking every aspect of actionability as discussed above. There could be a conflict of interest between a high expectation of technical significance and high satisfaction of business performance [11]. In this case, domain knowledge and business goals will play an important role in gaining a balanced output.

6.3 Case Studies

First, we present an example of mining multi-source combined associations. In practice, factor analysis, atomic pattern discovery, pattern relation analysis and pattern merger need to consider the pragmatic aspects of an underlying problem, including available sources, attributes, methods, constraints, evaluation metrics, business goals and expectations. The above combined mining process may be specified accordingly. For instance, in [7], we discuss a multi-source combined pattern mining process in Fig.1 in Section IV.C, which firstly analyzes and identifies a data source to discover patterns, The identified patterns are then used to guide the feature selection and pattern mining on each of the other sources. Pattern relations and domain knowledge are then involved to analyze the relations between patterns from different sources. A pattern merger method is called to combine

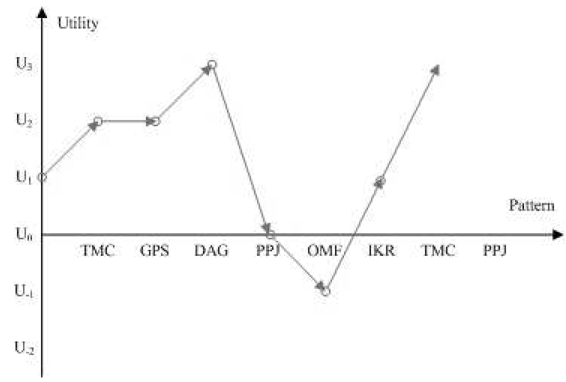

Figure 2: Pattern Evolution Chart

atomic patterns from different sources into cluster or pair associations. Examples of cluster patterns and pair patterns produced by this approach are demonstrated in [9] in social security data. In [10], a few general architectures for mining complex knowledge in complex data are introduced.

Second, we present an example of mining combined sequential patterns. For instance, in [9], we illustrate combined frequent activity patterns through the analysis of the relations between identified frequent sequential activities. We find *impact-contrasted activity patterns* with atomic patterns

$$\begin{cases} p \to T, \ \ in \ \ data \ \ set \ \ D_1 \\ p \to \neg T, \ \ in \ \ data \ \ set \ \ D_2 \end{cases} \tag{7}$$

We also find *impact-reversed activity patterns*, such as

$$\begin{cases} p \to T, \ \ in \ \ data \ \ set \ \ D_1 \\ p \wedge q \to \neg T, \ \ in \ \ data \ \ set \ \ D_1 \end{cases} \tag{8}$$

Finally, the combined mining approach can be used to identify multi-levels of high utility sequential patterns by extending high utility sequence mining [36]. Following the incremental frequent activity patterns discussed in [9], we may find an incremental cluster sequential pattern, with different levels of utility associated with the atomic patterns.

$$\begin{cases} TMC \to U_1 \\ TMC, GPS \to U_2 \\ TMC, GPS, DAG \to U_2 \\ TMC, GPS, DAG, PPJ \to U_3 \\ TMC, GPS, DAG, PPJ, OMF \to U_0 \\ TMC, GPS, DAG, PPJ, OMF, IKR \to U_{-1} \\ TMC, GPS, DAG, PPJ, OMF, IKR, TMC \to U_1 \\ TMC, GPS, DAG, PPJ, OMF, IKR, TMC, PPJ \to U_3 \end{cases}, \tag{9}$$

The identified multiple levels of high utility patterns can then be presented in terms of the pattern evolution chart. Figure 2 shows the pattern evolution and its impact (here, utility) change from element *TMC* to *PPJ* (representing different behaviors).

7 Challenge and Prospect

Combined mining and the approaches of object relation and pattern relation analysis have great potential in tackling big data for smarter decisions. Besides the challenges and prospects raised for handling particular data, we here discuss some of the issues and opportunities surrounding combined mining and pattern relation analysis.

- Pattern structure: Apart from the structures we have discussed in this paper and references, the main challenges lie in the automatic understanding and extraction of hidden structures in patterns, especially a large group of identified patterns. For instance, while it may be easy to talk about incremental and decremental patterns within frequent patterns, extracting such relations in tree-based, graph-based and model-based systems is not obvious. In these cases, structures are present either in implicit forms or in terms of sophisticated presentation formats.
- Pattern relation: This is the most interesting part of analyzing combined patterns, however, it is very challenging. Pattern relations are assumed to be more varied than in knowledge representation, which may involve syntactic, semantic and mathematical relations between patterns. Besides relation representation, many other research issues emerge, such as relation reasoning and inference, relation learning, relation discovery, and relation summarization and presentation.
- Pattern ontology: Pattern ontology is a very promising issue to be further explored. It is brand new in the data mining community, so before any firm outcomes are available, techniques in ontological engineering, semantic web and software design patterns may be used to represent and present combined patterns and their relations. It could become a very interesting and practical field in developing effective pattern ontology languages for formally representing, modeling, and reasoning about pattern structures and relations. While such a pattern ontology algebra (language) could be built on top of existing knowledge representation techniques, it is important to work out the differences and the need to develop the corresponding tools for capturing not only the syntactic but also the semantic and mathematical aspects between patterns.
- Pattern combination: The combination of patterns from pattern constituents and atomic patterns relies on many aspects, such as data characteristics analysis, combination aspect analysis, relation extraction, decomposition of pattern elements, and summarization and aggregation of patterns into pattern families. Domain knowledge, meta-knowledge, domain expert and the above aspects contribute to the merger of meaningful combined patterns. Data miners can summarize and aggregate patterns into combined patterns. It is also important to engage domain experts to evaluate the combination methods and the subsequent deliverables.
- Process automation: Combined mining follows the general data mining process, with additional focus on tackling challenges emerging in the underlying problem. Similar to a typical data mining task, many procedures such as business understanding, data understanding, feature analysis, and evaluation and deployment are not easily automated. However, the key procedures

of pattern abstraction, summarization, aggregation, generalization and inference could be automated if the underlying problem is clear and the relevant domain knowledge is clearly represented in the modeling. With the support of pattern ontology, it may be more manageable to guide the inference, summarization and aggregation into combined patterns by generating pattern configuration specifications and delivering ontologically represented pattern sets.

Besides pattern discovery and learning theory in data mining and machine learning, the problem solving of object and pattern relation learning has strong connections with other techniques examining other purposes, such as the pattern theory [16] for representing and inferring patterns, inductive learning for inducing, configuring and generalizing knowledge, knowledge engineering and ontological engineering for representing and reasoning about knowledge, and the constraint theory and other disciplines of algebra, geometry, statistics and statistical communications.

8 Conclusions

The analysis of object relations and pattern relations and structures is a very important issue in data mining and machine learning. Limited research on this issue has been conducted. The deliverables from current pattern mining are mainly individual patterns, which are often not informative and not actionable. This is because of the lack of pattern dimension analysis, including feature interaction, pattern interaction, pattern dynamics, pattern impact, pattern relation, pattern structure, selection criteria, and pattern presentation. Taking these pattern dimensions into consideration, combined mining is a technique to identify, extract and construct complex patterns, which appear as either single patterns or compound patterns with constituents from different dimensions (elements, features, relations, interactions, structures, constraints, and impacts) linked by proper connectives for various actionable semantics.

This paper presents a high level picture of combined mining, and discusses many novel aspects of pattern relation analysis and combined patterns. Pattern combination dimensions, pattern combination criteria, pattern relations, pattern structures and pattern paradigms, which are important for constructing combined patterns and for discovering actionable knowledge in complex data, are discussed. Pattern ontology and the pattern dynamic chart are also introduced to present combined patterns.

Combined pattern and combined mining present a general paradigm with great potential for identifying and producing informative and actionable patterns. One can ‘project’ one’s own problem or application onto the proposed framework of combined pattern space while selecting one’s own collection of ‘good’ patterns.

We are working on a deep understanding of pattern relations, from the perspectives of similarity, dissimilarity and dependence between patterns, to develop more complex but actionable knowledge. Pattern representation, inference and ontology are some of the key issues in our further study towards creating a pattern descriptive language.

9 Acknowledgment

We appreciate the anonymous reviewers for their very valuable comments and suggestions.

References

[1] Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I. and Angel, S. *A Pattern Language*, Oxford University Press, 1977.

[2] Baier, C. and Joost, P. K. *Principles of Model Checking*, Cambridge, MA: The MIT Press, 2008.

[3] Brachman, R.J. and Levesque,H.J. *Knowledge Representation and Reasoning*, Morgan Kaufmann, 2004.

[4] Cao, L. Actionable Knowledge discovery and delivery. *WIREs Data Mining and Knowledge Discovery 2,* 2, 149-163, 2012.

[5] Cao, L. Domain driven data mining: challenges and prospects. *IEEE Trans. on Knowledge and Data Engineering 22,* 6, 755-769, 2010.

[6] Cao, L., Ou, Y. and Yu, P.S. Coupled behavior analysis with applications, *IEEE Trans. Knowl. Data Eng.*, 24(8): 1378-1392, 2012.

[7] Cao, L., Zhang, H., Zhao, Y., Luo, D. and Zhang, C. Combined mining: discovering informative knowledge in complex data, IEEE Trans. SMC Part B, 41(3): 699 - 712, 2011.

[8] Cao, L., Yu, P.S., Zhao, Y. and Zhang, C. *Domain Driven Data Mining*, Springer, 2010.

[9] Cao, L., Zhao, Y. and Zhang, C. Mining impact-targeted activity patterns in imbalanced data. *IEEE Transactions on Knowledge and Data Engineering 20,* 8, 1053-1066, 2008.

[10] Cao, L., Zhao, Y., Zhang, H., Luo, D. and Zhang, C. Flexible Frameworks for Actionable Knowledge Discovery. *IEEE Transactions on Knowledge and Data Engineering,* 22, 9, 1299-1312, 2010.

[11] Cao, L., Luo, D. and Zhang, C. Knowledge actionability: satisfying technical and business interestingness, *International Journal of Business Intelligence and Data Mining*, 2(4): 496-514, 2007.

[12] Chein, M. and Mugnier, M. *Graph-based Knowledge Representation: Computational Foundations of Conceptual Graphs*, Springer, 2009.

[13] Cox, D. R. *Principles of Statistical Inference*, Cambridge University Press, 2006.

[14] Fagin, R., Halpern, J.Y., Moses, Y. and Vardi, M. *Reasoning About Knowledge*, MIT Press, 1995.

[15] Givant, S. and Halmos, P. *Introduction to Boolean Algebras*, Springer, 2009.

[16] Grenander, U. and Miller M. *Pattern Theory: From Representation to Inference*, Oxford University Press, 2007.

[17] Lent, B., Swami, A. N. and Widom, J. Clustering association rules. In *ICDE97*, 220-231, 1997.

[18] Lesh, N., Zaki, M. J. and Ogihara, M. Mining features for sequence classification. In *KDD99*, 342-346, 1999.

[19] Liese, F. and Miescke, K.J. *Statistical Decision Theory: Estimation, Testing, and Selection*, Springer, 2008.

[20] Liu, B., Hsu, W. and Ma, Y. Pruning and summarizing the discovered associations. In *KDD99*, 125-134, 1999.

[21] Luger, G. *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*, Addison-Wesley, 2002.

[22] Michalski, R.S. A theory and methodology of inductive learning in *Machine Learning: An Artificial Intelligence Approach*, R. S. Michalski, J. Carbonell and T. Mitchell (Eds.), pp. 83-134, 1983.

[23] Mooney, R.J. Integrating Abduction and Induction in Machine Learning, in *Abduction and Induction*, P. Flach and A. Kakas (Eds.), pp.181-191, Kluwer Academic Publishers, 2000.

[24] Ozgur, A. Tan, P.-N., and Kumar, V. Rba: an integrated framework for regression based on association rules. *SDM'04*, 210-221, 2004.

[25] Plasse, M., Niang, N., Saporta, G., Villeminot, A. and Leblond, L. Combined use of association rules mining and clustering methods to find relevant links between binary rare attributes in a large data set. *Computational Statistics & Data Analysis 52*, 596-613, 2007.

[26] Ras, Z.W. and Wieczorkowska, A. Action-rules: how to increase profit of a company, in *Proceedings of PKDD'00*, LNAI 1910, 587-59, Springer, 2000.

[27] Ras, Z. and Dardzinska, A. Action Rules Discovery Based on Tree Classifiers and Meta-actions, *ISMIS09*, 66-75, 2009.

[28] Russell, S.J. and Norvig, P. *Artificial Intelligence: A Modern Approach*, Prentice Hall, 2010.

[29] Song, Y., Cao, L. et al. Coupled behavior analysis for capturing coupling relationships in group-based market manipulations. *KDD 2012*, 976-984, 2012.

[30] Staab, S. and Studer, R. *Handbook on Ontologies*, 2nd ed. Berlin, DE: Springer Verlag, 2009.

[31] Wang, C., Cao, L. Formalization and Verification of Group Behavior Interactions. In *UTS AAI Technical Report*, 2011.

[32] Wang, C., Cao, L. et al. Coupled nominal similarity in unsupervised learning, *CIKM 2011*, 973-978, 2011.

[33] Wang, C., Wang, M., She, Z. and Cao, L. CD: A Coupled Discretization Algorithm. *PAKDD*, 407-418, 2012.

[34] Wang, K., Jiang, Y. and Tuzhilin, A. Mining Actionable Patterns by Role Models, *ICDE06*, 16-25, 2006.

[35] Wojtusiak, J., Michalski, R. S., Kaufman, K. and Pietrzykowski, J. Multitype Pattern Discovery Via AQ21: A Brief Description of the Method and Its Novel Features, in *Reports of the Machine Learning and Inference Laboratory, MLI 06-2*, George Mason University, Fairfax, VA, June, 2006.

[36] Yin, J., Zheng, Z. and Cao, L. USpan: An Efficient Algorithm for Mining High Utility Sequential Patterns, *KDD 2012*, 660-668, 2012.

[37] Young, G.A. and Smith, R.L. *Essentials of Statistical Inference*, CUP, 2005.

[38] Yu, H. Yang, J. and Han, J. Classifying large data sets using SVM with hierarchical clusters, In *KDD03*, 306-315, 2003.

[39] Zhang, H., Zhao, Y., Cao, L., and Zhang, C. Combined association rule mining, In *PAKDD08*, 1069-1074, 2008.

[40] Zhang H, Zhao Y, Cao L et al. Customer activity sequence classification for debt prevention in social security. *Journal of Computer Science and Technology*, 24(6): 1000-1009, 2009.

[41] Zhao, Y., Zhang, H., Figueiredo, F., Cao, L., and Zhang, C. Mining for combined association rules on multiple datasets. In *DDDM07*, 18-23, 2007.

[42] Zhao, Y., Zhang, H., Cao, L., Zhang, C., and Bohlscheid, H. Combined pattern mining: from learned rules to actionable knowledge. *AI08*, 393-403, 2008.

[43] Zhao, Y., Zhang, H., Cao, L., Zhang, C., and Bohlscheid, H. Efficient mining of event-oriented negative sequential rules. *WI08*, 336-342, 2008.

[44] Zhao, Y., Zhang, H., Wu, S., Pei J., Cao, L., Zhang, C., and Bohlscheid, H. Debt detection in social security by sequence classification using both positive and negative patterns, *ECML-PKDD*, LNAI 5782, 648-663, 2009.

[45] Zhao, Y., Zhang, C., and Cao, L. (eds.) *Post-Mining of Association Rules: Techniques for Effective Knowledge Extraction*. Information Science Reference, 2009.

[46] http://ontologydesignpatterns.org