Research on Parallel Attribute Exploration Algorithm Based on Unrelated Attribute and Intent Sets

Daojun Han[1,2], Wan Chen[1,2], and Xianyu Zuo[1,2](✉)

[1] Henan Key Laboratory of Big Data Analysis and Processing, Henan University, Kaifeng, Henan, China

[2] School of Computer and Information Engineering, Henan University, Kaifeng, Henan, China

http://cs.henu.edu.cn/

Abstract. The attribute exploration algorithm is an important knowledge discovery tool for obtaining the stem basis and intent of a given formal context in formal concept analysis. However, when the scale of the formal context is large, the traditional attribute exploration algorithm and the improved attribute exploration algorithm still traverse the attribute sets in the lexicographical order, which leads to the calculation process of the algorithm is too time-consuming. It seriously hinders the promotion and application of the current big data era. The time-consuming bottleneck mainly exists in the link of "traversing attribute sets in the lexicographic order". To solve this problem, first, we construct a prefix dictionary tree in the inverse linear order of the attribute sets in the inter-layer cardinality order and use the irrelevant definition. Second, we propose and prove three theorems by the above definition and formal concept analysis. Third, according to these theorems, we put forward a new parallel attribute exploration algorithm based on attribute and intent sets. In the process of calculating intent and pseudo-intent sets, the algorithm skips the computing process of attribute sets which are intent sets and neither intent nor pseudo-intent sets with the help of these proposed theorems and reduces search space of the algorithm and the scale of implication calculation, so as to reduce the time complexity. Experimental results show that the worst time complexity of the algorithm is $O(M^2 \times \max(P, G))$. Compared with the improved algorithm, this algorithm has obvious time performance advantages.

Keywords: Formal concept analysis · Stem base · Intent · Parallel attribute exploration · Lexicographic order · Knowledge discovery

This work was supported by the Scientific and technological project of Henan Province (Grant No. 202102310340), Foundation of University Young Key Teacher of Henan Province (Grant No. 2019GGJS040, 2020GGJS027), and Key scientific research projects of colleges and universities in Henan Province (Grant No. 21A110005).

D. Jiang and H. Song (Eds.): SIMUtools 2021, LNICST 424, pp. 577–599, 2022.
https://doi.org/10.1007/978-3-030-97124-3_45

1 Introduction

Formal Concept Analysis (abbrev. FCA) [1] was proposed by Professor Wille in Germany in the early 1980s. It uses Galois connection to obtain common attributes and possessions of objects from a given formal context. These objects with common attributes are then summarized as formal concepts or concepts for short. According to the generalization-instancing relationship between formal concepts, concepts of the formal context are formed into a complete lattice, namely the Concept Lattice [2–4].

As the kernel data structure of FCA, concept lattice has attracted wide attention from an abundance of researchers. It has been widely used in data mining [5–8], rule extraction [9,10], knowledge discovery [11,12], access control [13,14], granular computing [15], and conceptual cognitive learning [16–18] and other fields [19–22].

Among them, for a given formal context, the attribute exploration algorithm [23] obtains the implication relationship between attributes in the formal context through interactive inquiry with the domain expert. This kind of implication relationship can express the knowledge contained in the formal context and is an important form of knowledge representation. Different from other existing methods of obtaining knowledge based on known data for statistics or induction, attribute exploration can actively interact with the domain expert to explore the implication relationship between attribute sets. The exploration process can be generalized as the following steps: firstly, we obtain attribute sets and its' lexicographical relationship according to the given formal context; secondly, according to the lexicographical order, it gradually asks the domain expert whether the implication relationship between attribute sets is valid; third, it is up to for the domain expert to judge whether these problems are true; fourth, if the implication relationship is not valid, the domain expert provides a counterexample from the formal context that refutes the implication relationship and adds it to the current formal context. If it is valid, then it adds the implication to the candidate stem base; fifth, we should find the next attribute set to be interacted with the domain expert according to the lexicographical order; finally, the algorithm can explore the implication relations between all attribute sets according to the lexicographical order, so as to it can acquire the stem base of the given formal context.

At present, the attribute exploration algorithm has been widely used in many fields. Borchmann [24] started from the exploration of classic attribute exploration algorithm, and retained the completeness, non-redundancy of classic attribute exploration algorithm, and asked the domain expert about the characteristics of the smallest number of implications, and proposed a general attribute exploration framework. The various variants of attribute exploration are regarded as an instance under this framework. Jäschke and Rudolph [25] proposed a method to support attribute exploration through Web information retrieval, and extended attribute exploration to the Web domain, and improved query efficiency. Potoniec and Rudolph [26] combined attribute exploration with machine learning to make the development and optimization of domain-specific ontology easier. Hanika and Zumbragel [27] proposed a collaborative concept exploration algorithm, which established a consortial expert based on a well-formed query to give multiple domain experts the initial ability to handle false acceptance implications. Codocedo and Baixeries [28] sampled the formal context through the attribute exploration algorithm and proposed a general method to convert the irreducible pattern structure to the representation formal context, which reduces the complexity of mining pattern structure.

However, the process is too time-consuming due to the completeness of the attribute exploration algorithm when calculating the implication. Ryssel [29,30] used appropriate premises to improve the efficiency of attribute exploration, but this scheme improves efficiency while also causing the problem of not being able to obtain the minimum cardinality implication. Zhao [31] used the definition of correlation between attribute sets and the stem base to improve the traditional attribute exploration algorithm to a certain extent. On the basis of Zhao Xiaoxiang's improved attribute exploration algorithm, Shen [32] further proposed an irrelevant definition to further improve the algorithm after studying the attribute exploration algorithm between the two departments [35], but the algorithm still traverses attribute sets in lexicographic order, and the complexity is still high.

In fact, any algorithm that calculates the pseudo-intent of the formal context in lexicographical order has shown that it cannot obtain the stem base of the formal context [33] with the polynomial-time complexity. Therefore, for theoretically difficult problems, a popular method is to explore the possibility of parallelization of known algorithms. Kriegel [34] proposed a parallel attribute exploration algorithm, which reduces the overall time-consuming of the attribute exploration algorithm, but the algorithm still traverses too many attribute sets in lexicographic order, so it is still very time-consuming.

Different from the aforementioned schemes, we propose a new algorithm which focuses on finding less time-consuming algorithms as a whole, but also considers avoiding the inherent redundant calculation processes of the algorithm. Therefore, this paper proposes a new parallel attribute exploration under unrelated attribute and intent sets algorithm (PAEUIS), which constructs a prefix dictionary tree according to the attribute sets of the formal context in the inter-layer cardinality order and the inverse linear order within the layer, and skips the attribute sets in the layer that are obviously intent sets and not related to the stem base, which achieves the pruning effect in the prefix dictionary tree layer, and reduces the search space of the algorithm, and calculates the implication expressions in parallel for the remaining pseudo-intents, and reduces the overall consumption of the algorithm time. Our approach achieves new state-of-the-art performances on randomly generated datasets in the field of attribute exploration.

2 Preliminaries

This section briefly illustrates some related fundamental notions of FCA [1,5,23,24] used in this article as follows:

Definition 1. *A formal context $K = (G, M, I)$ is composed of two sets G and M and the relationship I between them, where G is the set of objects, M is an attribute set, and the relationship between them is $I \subseteq G \times M$. $(G, M) \in I$ or (gIm) means that the object g has an attribute m. We use $(g, m) \notin I$ to indicate that the object g does not possess the attribute m.*

Definition 2. *Let $K = (G, M, I)$ be a formal context, then define the following operations in the object subset $A \subseteq G$, and the attribute subset $B \subseteq M$:*

$$A^I = \{m \in B \mid \forall g \in A, (g, m) \in I\}$$
$$B^I = \{g \in A \mid \forall m \in B, (g, m) \in I\}$$

A^I refers to a collection of attributes shared by all objects in A, and B^I represents a collection of objects shared by all attributes in B. If for any $g \in G,\ m \in M$ there is, the formal context is said to be regular.

Definition 3. *Let $K = (G, M, I)$ be a formal context, then in the object subset $A \subseteq G$, and the attribute subset $B \subseteq M$, if $A,\ B$ satisfies $A^I = B,\ B^I = A$, then we call the two-tuple $(A,\ B)$ is a formal concept, abbreviated as a concept. Among them, A is the extent of the concept (A, B), and B is the intent of the concept $(A,\ B)$.*

Definition 4. *Suppose $K = (G, M, I)$ is a formal context, then the concepts on the formal context $K = (G, M, I)$ has the following basic properties $(\forall\ A, A_1, A_2 \subseteq G,\ \forall\ B, B_1, B_2 \subseteq M)$:*

Property 1. $A_1 \subseteq A_2 \rightarrow {A_1}^I \supseteq {A_2}^I,\ B_1 \subseteq B_2 \rightarrow {B_1}^I \supseteq {B_2}^I$;

Property 2. $A \subseteq A^{II},\ B \subseteq B^{II}$;

Property 3. $A = A^{III},\ B = B^{III}$;

Property 4. $A \subseteq B^I \Leftrightarrow B \subseteq A^I$

Property 5. *Both (A^{II}, A^I) and (B^I, B^{II}) are concepts.*

Definition 5. *Suppose $K = (G, M, I)$ is a formal context, then $P \subseteq M$ is called the pseudo-intent in the formal context K and satisfies the following conditions:*

1. *$P^{II} \neq P$;*
2. *For every pseudo-intent $Q \subsetneq P$ there is $Q^{II} \subseteq P$.*

Definition 6. *Suppose $K = (G, M, I)$ is a formal context. Let $B_1, B_2 \subseteq M$, then any formula of the form $B_1 \rightarrow B_2$ is called an implication formula. If there exists ${B_1}^I \subseteq {B_2}^I$ is valid, then the implication formula ${B_1}^I \subseteq {B_2}^I$ is valid in the formal context K.*

Definition 7. *Suppose* $K = (G, M, I)$ *is a formal context, and the set of implication* $Imp(K) = \{P \rightarrow P^{II} \backslash P \mid P \text{ is a pseudo-intent of } K\}$ *is the Duquenne-Guigues Base on the formal context, referred to as the stem base.*

Definition 8. *Suppose* $K = (G, M, I)$ *is a formal context, and the set of implication* $Imp(K)$ *has the implication expression of* $C \rightarrow D \in Imp(K)$*. If there exists* $C \nsubseteq T$ *or* $D \subseteq T$ *for the attribute set* $T \subseteq M$*, then* T *is related to* $C \rightarrow D$*. If the attribute set* T *is related to all the implication expressions in* $Imp(K)$*, then it is said that* T *is related to* $Imp(K)$*.*

Definition 9. *Suppose* $K = (G, M, I)$ *is a formal context, the attribute set* $M = m_1, \cdots, m_n$*, if the basic linear order given on* M *is* $(m_1 < m2 < \cdots < m_n)$*, then for any* $B_1, B_2 \subseteq M$ *there exists* $B_1 < B_2$ *if and only if* $m_i \in B_2 - B_1$ *exists and* $B_1 \cap \{m_1, \cdots, m_{i-1}\} = B_2 \cap \{m_1, \cdots, m_{i-1}\}$*.*

Definition 10. *Suppose* $K = (G, M, I)$ *is a formal context, the attribute set* $M = m_1, \cdots, m_n$*, if the basic linear order given on* M *is* $(m_1 < m2 < \cdots < m_n)$*, then the lexicographical relationship* $<$ *of the attribute set in the formal context* K *is the linear order relationship of* 2^M*. According to the lexicographical relationship, the attribute set in the formal context can be tested one by one whether the attribute set is an intent or pseudo-intent.*

Definition 11. *Suppose* $K = (G, M, I)$ *is a formal context, and the set of implication* $Imp(K)$ *has the implication form* $C \rightarrow D \in Imp(K)$*. If there exists* $C \subseteq T$ *and* $D \nsubseteq T$ *for the attribute set* $T \subseteq M$*, then* T *and* $C \rightarrow D$ *are not related. If the attribute set* T *is not related to any one of the implication expression in* $Imp(K)$*, then it is said that* T *is not related to* $Imp(K)$*.*

Definition 12. *Suppose* $K = (G, M, I)$ *is a formal context, and* $Imp(K)$ *is the stem base of* K*. As for any one implication* $C \rightarrow D \in Imp(K)$*, if the attribute set* T *is not related to the implication* $C \rightarrow D$*, then* T *is neither an intent nor pseudo-intent in the formal context* K*.*

3 Research on Parallel Attribute Exploration Algorithm Based on Unrelated Attribute and Intent Sets

For a given formal context, the implication relationship between attribute sets, that is, the knowledge implied by the formal context can be represented by the inclusion relationship between attribute sets. Attribute exploration algorithm is a tool to obtain these inclusion relationships in FCA, and can find the intent and stem base of a given formal context.

The famous NextClosure algorithm proposed by Ganter can be used to obtain the intent and stem base of a given formal context. The mathematical idea behind the algorithm is mainly to traverse attribute sets in lexicographical order, and actively asks the domain expert whether the implication between attributes is valid, so as to calculate all the intent and pseudo-intent sets of a given formal context. The implications obtained from the pseudo-intent attribute sets constitute the stem base of the given formal context. One of the great advantages of this algorithm is that the next (pseudo) intent is uniquely determined, but it may require backtracking to find it. In addition, the algorithm is essentially in linear order, that is, it cannot be parallelized. In this paper, we find that linear order is not used, and layer the attribute sets in a set cardinal order, and hierarchically parallelize to determine whether the attribute sets are intents or pseudo-intents, and in the process of exploration, the intent sets and irrelevant pseudo-intent sets are skipped the implication judgment of the attribute sets can avoid a lot of redundant calculation, and then quickly calculate the stem base of a given formal context. Besides, since in the implementation of multiple threads, there is no communication between different threads, that is, the calculation of the implication formula of the attribute sets in each cardinality layer can be calculated in parallel.

3.1 Theoretical Basis

To facilitate the following explanation, this article first sets definitions as follows.

Definition 13. *Let $K = (G, M, I)$ be a formal context, attribute set $B, B^+ \subseteq M$, if $B < B^+$ and the interval $<B, B^+>$ is an empty set, then it is said that B^+ is only greater than B, which is recorded as $B^+ \succ B$.*

Definition 14. *Let $K = (G, M, I)$ be a formal context, the attribute set $B, D \subseteq M$ and $B < N$, $N \not\supset B$. If there is $T \supset B$ for any attribute set $T \in <B, N>$, then it is said that N is non-trivial and only greater than B, which is recorded as $B \precnsim N$.*

Definition 15. *Let $K = (G, M, I)$ be a formal context and a Stem base $Imp(K)$ on K, any implication $C \to f(g(C)) - C \in Imp(K)$. If there is an attribute set B^+, $N \in M$ satisfies $B^+ > B$ and B^+ is not related to the Stem base $Imp(K)$, $B \precnsim N$ and N is related to $Imp(K)$. Then in the interval $<B, \min(f(g(B)), \; N)>$, there is neither an intent nor pseudo-intent.*

Definition 16. *Suppose $K = (G, M, I)$ is a formal context, and the attribute set $M = m_1, \cdots, m_n$, if the basic reverse linear order given on M is $(m_1 > m2 > \cdots > m_n)$, then the lexicographic relationship $>$ of the attribute set is an inverse linear order relationship of 2^M in the formal context. According to the reverse lexicographical relationship, the attribute set in the formal context can be tested one by one whether the attribute set is an intent or pseudo-intent.*

Definition 17. *Let $K = (G, M, I)$ be a formal context, the attribute set $M = m_1, \cdots, m_n$, if the basic reverse linear order given on M is $(m_1 > m2 > \cdots > m_n)$, according to the reverse linear order, all the attribute sets of the formal context K can be constructed as a prefix dictionary tree (the attribute sets of the parent node is actually contained in the attribute sets of the child node and the sibling node follows the reverse linear order).*

Based on the above definitions, the following findings in this paper can be used as a theoretical basis for further improving the attribute exploration algorithm.

Theorem 1. *Let $K = (G, M, I)$ be a formal context, any one implication $C \to D \in Imp(K)$. If attribute set $T = C \cap D$, then T is an intent.*

Proof. Since $C \to D \in Imp(K)$, so by Definition 7, we know that $D = C^{II} - C$, then $T = C \cap D = C \cap \{C^{II} - C\} = C^{II}$. That is, $T = C^{II}$. From Definition 7, we can see that $T^{II} = C^{II^{II}} = C^{II} = T$.

Theorem 1 shows that if the attribute sets only composed of the antecedents and consequents of any one implication in the stem base, then these attribute sets must be an intent. Because in the process of attribute exploration, the stem base is only obtained by the implication of the pseudo-intent sets, so the attribute sets that satisfy Theorem 1 do not need to calculate the corresponding implications.

Theorem 2. *Let $K = (G, M, I)$ be a formal context, any one implication $C \to D \in Imp(K)$. According to Definition 16, the reverse linear order of the sub-nodes of C in the dictionary prefix tree is $C_1 > C_2 > \cdots > C_n$, if $D \subseteq <C_n, C_1>$, then the node between the largest child node C_1 to D (including the D node) is not related to $Imp(K)$, so it does not exist pseudo-intent and intent in the interval $<D, C_1>$.*

Proof. According to Definition 17, the sub-nodes of C can be set as $(C_1, C_2, \cdots, C_n) \subset C$ in the dictionary prefix tree. If $D \subseteq <C_n, C_1>$, so for any attribute set $T \subseteq <D, C_1>$ contains at least one antecedent and does not include the consequent of the implication $C \to D \in Imp(K)$, then it can be known that T is neither an intent nor pseudo-intent from Definition 12.

Theorem 2 shows that, in the prefix dictionary tree, if the attribute sets of this layer are contained in the interval between any one known implication subsequent and the largest child node of the implication antecedent in the stem base, then there are neither an intent nor pseudo-intent attribute sets in this interval. This provides a theoretical basis for skipping the calculation of these attribute sets when parallelly judging these attributes are intents or pseudo-intents in the inner layer of the prefix dictionary tree.

Theorem 3. *Suppose $K = (G, M, I)$ is a formal context, if $D = \emptyset$ for any one implication $C \to D$, then the implication $C \to D$ is must be valid in the formal context K, and the attribute set C must be an intent.*

Proof. For any one implication $C \to D$ and $D = \emptyset$, the implication $C \to D = C \to \emptyset$. And in the formal context $K = (G, M, I)$, there is $C^I \subseteq G = \emptyset^I$, so any one implication $C \to \emptyset$ is valid in the formal context K. And because the implication $C \to \emptyset$ holds in the formal context K, so $\emptyset = C^{II} - C$ namely $C^{II} = C$, so C must be an intent.

Theorem 3 shows that when judging implications in the formal context, if the consequents of any one implication is an empty set, then the implication must be valid in the given formal context, and the antecedents of the implication must be an intent. Because, the stem base is only obtained by implications of the pseudo-intent sets in the process of attribute exploration, so the implications that satisfy Theorem 3 do not need to be judged whether it is valid in the given formal context.

3.2 Algorithm

Based on the above definitions and theorems, according to the recursive definition of pseudo-intent, we designed a parallel attribute exploration algorithm based on pseudo-intent and intent sets (PAEUIS) which arranges attribute sets in a hierarchical cardinal order and skips irrelevant attribute and intent sets and only considers pseudo-intent sets. The algorithm description is shown in Algorithm 1.

At the beginning of the PAEUIS algorithm, the given formal context is known, the attribute set and the cardinality between attribute sets are known, and the stem base is an empty set. The algorithm first starts to obtain the empty attribute set and attribute sets for single attributes, then uses Definition 12 and Theorem 1 to skip implications that do not require the attribute set as the antecedent in the cardinal order of the attribute set between interlamination calculation. The specific steps of the algorithm are as follows:

Steps 2–9 in the PAEUIS algorithm, the algorithm judges whether these attributes of the single point layer are intents in parallel according to Definition 3. If the attribute set currently determined is an intent, then this attribute set is not required further calculations; otherwise, the attribute set currently determined must be a pseudo-intent. The corresponding implications are directly calculated for these pseudo-intent attribute sets and add them to the set of implication, and then wait for the end of parallel processes of all the attribute sets in the first layer, so far the calculation of the first layer attribute sets of the algorithm is completed. Step 12 of the algorithm means that after the implications of all pseudo-intent sets in the current layer are calculated, the algorithm needs to enter the next layer through the cardinal number auto-increment method to judge and calculate the implication for the attribute sets of the next layer. In steps 13–16 of the algorithm, according to Theorem 2, it judges parallelly whether the attribute sets in the current layer are contained in the interval of any one known implication subsequent and the largest child node of the antecedent, thereby skipping these attribute intervals. According to Definitions 11 and 12, Steps 17–19 are to judge the irrelevant relationship between the attribute sets in the current layer and the obtained implication set in parallel, so as to find the attribute sets in this layer that are not related to the stem base, and skip the judgment that takes these attribute sets as antecedent. In steps 20–25, according to Theorem 1, the algorithm judges in parallel whether the attribute set in the current layer is composed of any one known implication antecedent and subsequent and determines whether the attribute sets in the current layer in

Algorithm 1. PAEUIS

Input Formal context: $K(G, M, I)$.
Output Stem base: $Imp(K)$, Intent set: $C_{(}K)$.
1: $k = 1,\ B[k] = \{\emptyset\} \cup M[k],\ Imp(K) = \emptyset$
2: **for** *all* b *in* $B[k]$ *in parallel* **do**
3: **if** $b^{II} == b$ **then**
4: b is an intent, $C(K) = C(K) \cup \{b\}$, and skip the implication of attribute set b.
5: **else**
6: $Imp(K) = Imp(K) \cup \{b \rightarrow b^{II} - b\}$
7: **end if**
8: Wait for termination of all parallel processes in this layer.
9: **end for**
10: $Imp_{k-1}(K) = Imp(K),\ C_{k-1}(K) = C_k(K)$
11: **while** $k < |M|$ **do**
12: $k = k + 1$
13: **for** *all* b *in* $B[k]$ *in parallel* **do**
14: **if** $b \subseteq <F, D_1>$, F is the consequent of any one implication $D \rightarrow F \in Imp_{k-1}(K)$, D_1 is the largest child node of D **then**
15: b is neither an intent nor pseudo-intent, skip the calculation of attribute set b.
16: **end if**
17: **if** b is unrelated to $Imp_{k-1}(K)$ **then**
18: b is neither an intent nor pseudo-intent, skip the calculation of attribute set b.
19: **end if**
20: **if** b is consisted of antecedent and consequent of any one implication $Imp_{k-1}(K)[i]$ in $Imp_{k-1}(K)$ **then**
21: b is an intent, $C_k(K) = C_{k-1}(K) \cup \{b\}$, skip the calculation of attribute set b.
22: **end if**
23: **if** $b^{II} - b == \emptyset$ **then**
24: b is an intent, $C_k(K) = C_{k-1}(K) \cup \{b\}$, skip the calculation of attribute set b.
25: **end if**
26: $Imp_k(K) = Imp_{k-1}(K) \cup \{b \rightarrow b^{II} - b\}$
27: Wait for termination of all parallel processes in this layer.
28: **end for**
29: **end while**
30: $Imp(K) = Imp_k(K),\ C(K) = C_k(K)$
31: **return** $Imp(K),\ C(K)$.

parallel are intents or not by Definition 3 and Theorem 3, thus skipping the judgment of this kind of attribute set as the antecedent implication. After these two judgments, the attribute sets that algorithm needs to calculate implications are all pseudo-intents sets. Step 26 is to calculate the corresponding implications for the pseudo-intent sets, and add the implications to the set of implication.

Step 27 is to wait for the end of all parallel processes in the current layer. Step 31 of the final algorithm directly outputs the set of implication Imp(K), namely the stem base and intent set of the given formal context.

Suppose the scale of the formal context is $G \times M$. In the PAEUIS algorithm, all attribute sets need to be divided into M layers, so the time complexity of the entire process layering is $O(M)$. When the algorithm calculates the first-layer attributes in parallel, it needs to traverse the formal context once for the first-layer attribute b in parallel and to calculate b^{II} once, so for the first-layer attribute sets, the time complexity is $O(G \times M)$. In each of the following layers, the irrelevance of any attribute set b to the implications in stem base is related to the scale of the stem base, because the scale of stem base is related to the scale of the formal context, but the relationship is not very clear. Therefore the scale of stem base can be set as P, then the complexity of calculating a single attribute set b unrelated to stem base in each layer is $O(M \times P)$. Similarly, for any one attribute set b in each layer, the time complexity of computing b is whether composed of any one implication antecedent and subsequent in stem base is $O(M \times P)$. Finally, for each layer of pseudo-intent set implication parallel calculation, it is still necessary to traverse the formal context once, that is, the time complexity is $O(G \times M)$. Therefore, the worst time complexity of the PAEUIS algorithm is $O(M \times \max(M \times P,\ M \times G))$. Although the time complexity of this algorithm is related to the scale P of stem base, compared with the best time complexity $O(M \times M \times G \times P \times P)$ of the AEUS algorithm, the time complexity of the algorithm in this paper is significantly less than that of the AEUS algorithm.

4 A Case Study of Attribute Exploration

This section uses an example to illustrate the running process of Shen Xiajiong's attribute exploration algorithm (AEUS) and the PAEUIS algorithm proposed in this article, and focuses on the comparison of the number of implications required in the attribute exploration process.

Example 1. Given a formal context $K = (G, M, I)$ as shown in Table 1, where $G = \{1, 2, 3, 4\}$, $M = \{a, b, c, d, e, f, g, h, i\}$ and the basic linear order on M is $a < b < c < d < e < f < g < h < i$.

4.1 An Example of AEUS Algorithm

The initial state of the formal context K is shown in Table 2, $K_0 = \emptyset$, intent set $C_0(K) = \emptyset$, stem base $Imp_0 = \emptyset$, and the first attribute set $B_0 = \emptyset$ according to the lexicographic order. Specific steps are as follows:

Table 1. The formal context of K

	a	b	c	d	e	f	g	h	i
1	0	0	1	1	1	1	1	0	0
2	0	0	0	0	0	0	1	0	1
3	0	0	1	1	1	0	1	0	0
4	1	1	1	1	0	0	0	1	0

Table 2. The formal context of K_0

a	b	c	d	e	f	g	h	i

1. In the formal context K_0, $\emptyset^{II} = \{abcdefghi\}$, we ask the domain expert whether the implication $\emptyset \to \emptyset^{II} - \emptyset = \emptyset \to \{abcdefghi\}$ is valid in the formal context K? In the formal context K, $\emptyset^I = \{1,2,3,4\},\ \{abcdefghi\}^I = \emptyset$. And because $\{1,2,3,4\} \nsubseteq \emptyset$, the implication $\emptyset \to \{abcdefghi\}$ does not hold in the formal context K. We take a counterexample 1 from the formal context K to refute the implication $\emptyset \to \{abcdefghi\}$, and add it in the formal context K_0 to obtain the formal context K_1 as shown in Table 3, where $C_1(K) = C_0(K),\ B_1 = B_0,\ Imp_1(K) = Imp_0(K)$.

Table 3. The formal context of K_1

	a	b	c	d	e	f	g	h	i
1	0	0	1	1	1	1	1	0	0

2. In the formal context K_1, $\emptyset^{II} = \{cdefg\}$, we ask the domain expert whether the implication $\emptyset \to \emptyset^{II} - \emptyset = \emptyset \to \{cdefg\}$ holds in the formal context K? In K, $\emptyset^I = \{1,2,3,4\},\ \{g\}^I = \{1,2,3,4\}$. And since $\{1,2,3,4\} \nsubseteq \{1,2,3\}$, the implication $\emptyset \to \{g\}$ does not hold in K. We take a counterexample 2 from the formal context K to refute the implication $\emptyset \to \{abcdefghi\}$, and add it in the formal context K_1 to obtain the formal context K_2 as shown in Table 4, where $C_2(K) = C_1(K),\ B_2 = B_1,\ Imp_2(K) = Imp_1(K)$.
3. In the formal context K_2, $\emptyset^{II} = \{g\}$, we ask the domain expert whether the implication $\emptyset \to \emptyset^{II} - \emptyset = \emptyset \to \{g\}$ holds in K? In K, $\emptyset^I = \{1,2,3,4\},\ \{g\}^I = \{1,2,3,4\}$. And since $\{1,2,3,4\} \nsubseteq \{1,2,3\}$, the implication $\emptyset \to \{g\}$ does not hold in K. We take a counterexample 4 from the formal context K to refute the implication $\emptyset \to \{g\}$ and add it in the formal context K_2 to get the formal context K_3 as shown in Table 5, where $C_3(K) = C_2(K),\ B_3 = B_2,\ Imp_3(K) = Imp_2(K)$.
4. In the formal context K_3, $\emptyset^{II} = \{\emptyset\}$, we ask the expert whether the implication $\emptyset \to \emptyset^{II} - \emptyset = \emptyset \to \{\emptyset\}$ holds in K? In K, $\emptyset^I = \{1,2,3,4\},\ \{\emptyset\}^I =$

Table 4. The formal context of K_2

	a	b	c	d	e	f	g	h	i
1	0	0	1	1	1	1	1	0	0
2	0	0	0	0	0	0	1	0	1

Table 5. The formal context of K_3

	a	b	c	d	e	f	g	h	i
1	0	0	1	1	1	1	1	0	0
2	0	0	0	0	0	0	1	0	1
4	1	1	1	1	0	0	0	1	0

$\{1,2,3,4\}$. And since $\{1,2,3,4\} == \{1,2,3,4\}$, the implication $\emptyset \to \{\emptyset\}$ holds in K. And because $\emptyset^{II} = \emptyset$, $\emptyset$ is an intent, and add it in C_3. So $C_4(K) = C_3(K) \cup \emptyset,\ K_4 = K_3,\ Imp_4(K) = Imp_3(K)$, then we need to calculate B_4.

5. According to the lexicographic order, we know that the next attribute set of $\emptyset$ is $\emptyset^+ = \{i\}$. Since $\emptyset \subseteq \{i\}$, $\{i\}$ is related to $Imp_4(K)$, $B_4 = \{i\}$. In the formal context K_4, $\{i\}^{II} = \{gi\}$, we ask the domain expert whether the implication $\{i\} \to \{i\}^{II} - \{i\} = \{i\} \to \{g\}$ holds in the formal context K? In the formal context K, $\{i\}^I = \{2\}$, $\{g\}^I = \{1,2,3\}$. And because $\{2\} \subseteq \{1,2,3\}$, the implication $\{i\} \to \{g\}$ holds in K. And because $\{i\}^{II} = \{gi\} \neq \{i\}$, $\{i\}$ is a pseudo-intent, and add $\{i\} \to \{g\}$ in Imp_4. So $Imp_5(K) = Imp_4(K) \cup \{i \to g\},\ K_5 = K_4,\ C_5(K) = C_4(K)$, then we need to calculate B_5.
6. According to the lexicographic order, we know that the next attribute set of $\{i\}$ is $\{i\}^+ = \{h\}$, because $i \nsubseteq \{h\}$ so $\{h\}$ is related to $Imp_5(K)$, $B_5 = \{h\}$. In the formal context K_5, $\{h\}^{II} = \{abcdh\}$, we ask the domain expert whether the implication $\{h\} \to \{h\}^{II} - \{h\} = \{h\} \to \{abcd\}$ is valid in the formal context K? In the formal context K, $\{h\}^I = \{4\}$, $\{abcd\}^I = \{4\}$. Since $\{4\} == \{4\}$, the implication $\{h\} \to \{abcd\}$ is valid in K. And because $\{h\}^{II} = \{abcdh\} \neq \{h\}$, $\{h\}$ is a pseudo-intent, and add $\{h\} \to \{abcd\}$ in Imp_5. So $Imp_6(K) = Imp_5(K) \cup \{h \to abcd\},\ K_6 = K_5,\ C_6(K) = C_5(K)$, then we need to calculate B_6.
7. According to the lexicographic order, we know that the next attribute set of $\{h\}$ is $\{h\}^+ = \{hi\}$, because $\{i \cup h\} \subseteq \{hi\}$ and $\{g\} \nsubseteq \{hi\}$, $\{abcd\} \nsubseteq \{hi\}$, $\{hi\}$ is not related to $Imp_6(K)$, then we calculate $T = \{B_5\}^{II} = \{h\}^{II} = \{abcdh\}$. According to Definition 15, $N = \{g\}$, because $T > N$, $B_6 = N = \{g\}$. In the formal context K_6, $\{g\}^{II} = \{g\}$, we ask the domain expert whether the implication $\{g\} \to \{g\}^{II} - \{g\} = \{g\} \to \emptyset$ holds in the formal context K? In K, $\{g\}^I = \{1,2,3\}$, $\emptyset^I = \{1,2,3,4\}$. And since $\{1,2,3\} \subseteq \{1,2,3,4\}$, the implication $\{g\} \to \emptyset$ holds in K. And because $\{g\}^{II} = \{g\} == \{g\}$, $\{g\}$ is an intent and add it in C_6. So $C_7(K) = C_6(K) \cup \{g\},\ K_7 = K_6,\ Imp_7(K) = Imp_6(K)$, then we need to calculate B_7.

8. According to the lexicographic order, we know that the next attribute set of $\{g\}$ is $\{g\}^+ = \{gi\}$, since $\{g \cup \emptyset\} \subseteq \{gi\}$ and $\{h\} \nsubseteq \{gi\}$, $\{gi\}$ is related to $Imp_7(K)$, $B_7 = \{gi\}$. In the formal context K_7, $\{gi\}^{II} = \{gi\}$, we ask the domain expert whether the implication $\{gi\} \rightarrow \{gi\}^{II} - \{gi\} = \{gi\} \rightarrow \emptyset$ holds in K? In K, $\{gi\}^I = \{2\}$, $\emptyset^I = \{1, 2, 3, 4\}$. And since $\{2\} \subseteq \{1, 2, 3, 4\}$, the implication $\{gi\} \rightarrow \emptyset$ holds in K. And because $\{gi\}^{II} = \{gi\} == \{gi\}$, $\{gi\}$ is an intent, and add it in C_7. So $C_8(K) = C_7(K) \cup \{gi\}$, $K_8 = K_7$, $Imp_8(K) = Imp_7(K)$, then we need to calculate B_8.
9. According to the lexicographic order, we know that the next attribute set of $\{gi\}$ is $\{gi\}^+ = \{gh\}$, because $\{h\} \subseteq \{gh\}$ and $\{adcd\} \nsubseteq \{gh\}$, $\{gh\}$ is not related to $Imp_8(K)$, then we need to calculate $T = \{abcdgh\}$. According to Definition 15, $N = \{f\}$, because $T > N$, $B_8 = N = \{f\}$. In the formal context K_8, $\{f\}^{II} = \{cdefg\}$, we ask the domain expert whether the implication $\{f\} \rightarrow \{f\}^{II} - \{f\} = \{f\} \rightarrow \{cdeg\}$ in the formal context K? In K, $\{f\}^I = \{1\}$, $\{cdeg\}^I = \{1, 3\}$. And because $\{1\} \subseteq \{1, 3\}$, the implication $\{f\} \rightarrow \{cdeg\}$ is valid in K. And because $\{f\}^{II} = \{cdefg\} \neq \{f\}$, $\{f\}$ is a pseudo-intent, and add $\{f\} \rightarrow \{cdeg\}$ in Imp_8. So $Imp_9(K) = Imp_8(K) \cup \{f \rightarrow cdeg\}$, $K_9 = K_8$, $C_9(K) = C_8(K)$, then we need to calculate B_9.
10. According to the lexicographic order, we know that the next attribute set of $\{f\}$ is $\{f\}^+ = \{fi\}$, because $\{f\} \subseteq \{fi\}$ and $\{cdeg\} \nsubseteq \{fi\}$, $\{fi\}$ is not relevant to $Imp_9(K)$, then we need to calculate $T = \{cdefg\}$. According to Definition 15, $N = \{e\}$, because $T > N$, $B_9 = N = \{e\}$. In the formal context K_9, $\{e\}^{II} = \{cdefg\}$, we ask the domain expert whether the implication $\{e\} \rightarrow \{e\}^{II} - \{e\} = \{e\} \rightarrow \{cdfg\}$ is valid in the formal context K? In the formal context K, $\{e\}^I = \{1, 3\}$, $\{cdfg\}^I = \{1\}$. And since $\{1, 3\} \nsubseteq \{1\}$, the implication $\{e\} \rightarrow \{cdfg\}$ does not hold in K. We take a counterexample 3 from the formal context K to refute the implication $\{e\} \rightarrow \{cdfg\}$ and add it in the formal context K_9 to get the formal context K_{10} as shown in Table 6, where $C_{10}(K) = C_9(K)$, $B_{10} = B_9$, $Imp_{10}(K) = Imp_9(K)$.

Table 6. The formal context of K_{10}

	a	b	c	d	e	f	g	h	i
1	0	0	1	1	1	1	1	0	0
2	0	0	0	0	0	0	1	0	1
4	1	1	1	1	0	0	0	1	0
3	0	0	1	1	1	0	1	0	0

11. In the formal context K_{10}, $\{e\}^{II} = \{cdeg\}$, we ask the domain expert whether the implication $\{e\} \rightarrow \{e\}^{II} - \{e\} = \{e\} \rightarrow \{cdg\}$ is valid in the formal context K? In K, $\{e\}^I = \{1, 3\}$, $\{cdg\}^I = \{1, 3\}$. Since $\{1, 3\} == \{1, 3\}$, the implication $\{e\} \rightarrow \{cdg\}$ is valid in K. And because

$\{e\}^{II} = \{cdeg\} \neq \{e\}$, $\{e\}$ is a pseudo-intent, and add $\{e\} \rightarrow \{cdg\}$ in Imp_{10}. So $Imp_{11}(K) = Imp_{10}(K) \cup \{e \rightarrow cdg\}$, $K_{11} = K_{10}$, $C_{11}(K) = C_{10}(K)$, then we need to calculate B_{11}.

12. The subsequent steps of the algorithm AEUS are shown in Table 7.

Table 7. The subsequent calculation process of the AEUS algorithm

Step	B_i	B_i^+	$Imp_{i+1}(K)$	$C_{i+1}(K)$	Whether B_i^+ is not related to $Imp_{i+1}(K)$?	T	N	B_{i+1}	Whether $B_{i+1} \rightarrow \{B_{i+1}\}^{II} - \{B_{i+1}\}$ is valid in the formal context K?
(12)	$B_{10} = \{e\}$	$B_{10}^+ = \{ei\}$	$Imp_{11}(K) = Imp_{10} \cup \{e \rightarrow cdg\}$	$C_{11}(K) = C_{10}(K)$	$\{ei\}$ is not related to $Imp_{11}(K)$	$\{cdeg\}$	$\{d\}$	$\{d\}$	$\{d\} \rightarrow \{c\}$ is valid in K
(13)	$B_{11} = \{d\}$	$B_{11}^+ = \{di\}$	$Imp_{12}(K) = Imp_{11} \cup \{d \rightarrow c\}$	$C_{12}(K) = C_{11}(K)$	$\{di\}$ is not related to $Imp_{12}(K)$	$\{cdg\}$	$\{c\}$	$\{c\}$	$\{c\} \rightarrow \{d\}$ is valid in K
(14)	$B_{12} = \{c\}$	$B_{12}^+ = \{ci\}$	$Imp_{13}(K) = Imp_{12} \cup \{c \rightarrow d\}$	$C_{13}(K) = C_{12}(K)$	$\{ci\}$ is not related to $Imp_{13}(K)$	$\{cdg\}$	$\{cd\}$	$\{cd\}$	$\{cd\} \rightarrow \emptyset$ is valid in K
(15)	$B_{13} = \{cd\}$	$B_{13}^+ = \{cdi\}$	$Imp_{14}(K) = Imp_{13}$	$C_{14}(K) = C_{13}(K) \cup \{cd\}$	$\{cdi\}$ is not related to $Imp_{14}(K)$	$\{cdgi\}$	$\{cdg\}$	$\{cdg\}$	$\{cdg\} \rightarrow \emptyset$ is valid in K
(16)	$B_{14} = \{cdg\}$	$B_{14}^+ = \{cdgi\}$	$Imp_{15}(K) = Imp_{14}$	$C_{15}(K) = C_{14}(K) \cup \{cdg\}$	$\{cdgi\}$ is not related to $Imp_{15}(K)$	$\{cdeg\}$	$\{b\}$	$\{cdeg\}$	$\{cdeg\} \rightarrow \emptyset$ is valid in K
(17)	$B_{15} = \{cdeg\}$	$B_{15}^+ = \{cdegi\}$	$Imp_{16}(K) = Imp_{15}$	$C_{16}(K) = C_{15}(K) \cup \{cdeg\}$	$\{cdegi\}$ is related to $Imp_{16}(K)$			$\{cdegi\}$	$\{cdegi\} \rightarrow \{abfh\}$ is valid in K
(18)	$B_{16} = \{cdegi\}$	$B_{16}^+ = \{cdegh\}$	$Imp_{17}(K) = Imp_{16} \cup \{cdegi \rightarrow abfh\}$	$C_{17}(K) = C_{16}(K)$	$\{cdegh\}$ is not related to $Imp_{17}(K)$	$\{cdefg\}$	$\{b\}$	$\{cdefg\}$	$\{cdefg\} \rightarrow \emptyset$ is valid in K
(19)	$B_{17} = \{cdefg\}$	$B_{17}^+ = \{cdefgi\}$	$Imp_{18}(K) = Imp_{17}(K)$	$C_{18}(K) = C_{17}(K) \cup \{cdefg\}$	$\{cdefgi\}$ is not related to $Imp_{18}(K)$	$\{abcdefgh\}$	$\{b\}$	$\{b\}$	$\{b\} \rightarrow \{acdh\}$ is valid in K
(20)	$B_{18} = \{b\}$	$B_{18}^+ = \{bi\}$	$Imp_{19}(K) = Imp_{18} \cup \{b \rightarrow acdh\}$	$C_{19}(K) = C_{18}(K)$	$\{bi\}$ is not related to $Imp_{19}(K)$	$\{abcdh\}$	$\{a\}$	$\{a\}$	$\{a\} \rightarrow \{bcdh\}$ is valid in K
(21)	$B_{19} = \{a\}$	$B_{19}^+ = \{ai\}$	$Imp_{20}(K) = Imp_{19} \cup \{a \rightarrow bcdh\}$	$C_{20}(K) = C_{19}(K)$	$\{ai\}$ is not related to $Imp_{20}(K)$	$\{abcdh\}$		$\{abcdh\}$	$\{abcdh\} \rightarrow \emptyset$ is valid in K
(22)	$B_{20} = \{abcdh\}$	$B_{20}^+ = \{abcdhi\}$	$Imp_{21}(K) = Imp_{20}(K)$	$C_{21}(K) = C_{20}(K) \cup \{abcdh\}$	$\{abcdhi\}$ is not related to $Imp_{21}(K)$	$\{abcdegh\}$		$\{abcdegh\}$	$\{abcdegh\} \rightarrow \emptyset$ is valid in K
(23)	$B_{21} = \{abcdegh\}$	$B_{21}^+ = \{abcdeghi\}$	$Imp_{22}(K) = Imp_{21}(K)$	$C_{22}(K) = C_{21}(K) \cup \{abcdegh\}$	$\{abcdeghi\}$ is not related to $Imp_{22}(K)$	$\{abcdefghi\}$		$\{abcdefghi\}$	$\{abcdefghi\} \rightarrow \emptyset$ is valid in K

After the calculation of the above process, we can see that the intent set of this formal context set is $\{\emptyset, \{g\}, \{gi\}, \{cd\}, \{cdeg\}, \{cdefg\}, \{abcdh\}, \{abcdefghi\}\}$ and the stem base is $Imp(K) = \{i \rightarrow g, h \rightarrow abcd, f \rightarrow cdeg, e \rightarrow cdg, d \rightarrow c, c \rightarrow d, cdegi \rightarrow abfh, b \rightarrow acdh, a \rightarrow bcdh, abcdegh \rightarrow fi\}$. And from the above process, it can be seen that the AEUS algorithm skips a part of attribute sets based on the linear order, but the algorithm still calculates too many implications of attribute sets, and there are many repeated calculations when the subsequent is the empty set, and the existence of the redundant calculation of attribute sets which are composed of the antecedent and subsequent in the obtained implication sets in these implications.

4.2 Example Process of PAEUIS Algorithm

According to Definition 16, the inverse linear order of the attribute set of the formal context K shown in Table 1 is $a > b > c > d > e > f > g > h > i$, and according to the inverse linear order all the obtained attribute sets constitute a 9-layer prefix dictionary tree in the inter-layer attribute sets cardinality order and the inner-layer attribute sets reverse linear order. The initial state of the formal context K as shown in Table 1, the implication set $Imp(K) = \emptyset$, the intent set $C(K) = \emptyset$.

1. Then we need to calculate the first-layer attribute set of the formal context K, and the first-level attribute set is $B[1] = \{\emptyset \cup M[1]\} = \{\emptyset, a, b, c, d, e, f, g, h, i\}$. The parallel calculation of the attribute set b in $B[1]$ is $\emptyset^{II} =$

$\emptyset$, $\{a\}^{II} = \{abcdh\} \neq \{a\}$, $\{b\}^{II} = \{abcdh\} \neq \{b\}$, $\{c\}^{II} = \{cd\} \neq \{c\}$, $\{d\}^{II} = \{cd\} \neq \{d\}$, $\{e\}^{II} = \{cdeg\} \neq \{e\}$, $\{f\}^{II} = \{cdefg\} \neq \{f\}$, $\{g\}^{II} = \{g\}$, $\{h\}^{II} = \{abcdh\} \neq \{h\}$, $\{i\}^{II} = \{gi\} \neq \{i\}$. So $C(K) = \{\emptyset,\ g\}$, $Imp(K) = \{a \to bcdh,\ b \to acdh,\ c \to d,\ d \to c,\ e \to cdg,\ f \to cdeg,\ h \to abcd,\ i \to g\}$.

2. Then we need to calculate the second-layer attribute set of the formal context K. First, according to Theorem 2, Theorem 1, and $Imp(K)$, we can skip the attribute set range of this layer is all child nodes of a, all child nodes of b, and all child nodes of c except a node cd, all child nodes of d, all child nodes of e, all child nodes of f, all child nodes of h, and all child nodes of i. Then, the remaining attribute set, namely $\{gh, gi\}$, it is judged parallelly the correlation with $Imp(K)$ based on Definition 12. It can be seen that $\{gh\}$ is not related to $Imp(K)$, and according to Theorem 1, we can know $\{gi\}$ is an intent. So $C_2(K) = C(K) \cup \{cd, gi\},\ Imp_2(K) = Imp(K)$.
3. Then we need to calculate the third-layer attribute set of the formal context K. First, according to Theorem 2 and $Imp_2(K)$, we can skip the attribute set range of this layer is all sub-nodes of a, all sub-nodes of b, and all sub-nodes of d, all sub-nodes of e, all sub-nodes of f. Then the remaining attribute set, namely $\{T \cup \{ghi\} |\ T \leq cde\}$, where $T \leq cde$ are elements of c's child nodes in the reverse linear order that are less than or equal to cde. According to Definition 12, the correlation between this attribute set and $Imp_2(K)$ is judged in parallel. It can be seen that $\{cde, cdf, T|\ T \leq cdh\}, ghi\}$ is not related to $Imp_2(K)$, and $\{cdg\}$ is not an intent by Definition 3. So $C_3(K) = C_2(K) \cup \{cdg\},\ Imp_3(K) = Imp_2(K)$.
4. Then we need to calculate the fourth-layer attribute set of the formal context K. First, according to Theorem 2 and $Imp_3(K)$, we can skip the attribute set range of this layer is all sub-nodes of a, all sub-nodes of b, and all sub-nodes of d, all sub-nodes of e, all sub-nodes of f. Then the remaining attribute set, namely $\{T|\ T \leq cdef\}$, where $T \leq cdef$ are the elements of c in the reverse linear order which are less than or equal to $cdef$. According to Definition 12, then we need to judge the correlation between this attribute set and $Imp_3(K)$ in parallel. So we know $\{cdef, T|\ T \leq cdeh\}$ is not related to $Imp_3(K)$, and $\{cdeg\}$ is an intent by Theorem 1. Therefore $C_4(K) = C_3(K) \cup \{cdeg\},\ Imp_4(K) = Imp_3(K)$.
5. Then we need to calculate the fifth-layer attribute set of the formal context K. First, according to Theorem 2, Theorem 1, and $Imp_4(K)$, we can skip the attribute set range of this layer is all sub-nodes of a except a node $abcdh$, all sub-nodes of b, all sub-nodes of d, all sub-nodes of e, all sub-nodes of f. Then the remaining attribute set i.e. $\{T|\ T \leq cdefg\}$, where $T \leq cdefg$ are the elements of c in the reverse linear order which are less than or equal to $cdefg$, and $\{cdefg\}$ is an intent by Theorem 1. According to Definition 12, we need to judge the correlation between this attribute set and $Imp_4(K)$ in parallel. Obviously, $\{cdegi\}$ is a pseudo-intent attribute set and the rest attribute sets are not related to $Imp_3(K)$. So $C_5(K) = C_4(K) \cup \{abcdh, cdefg\},\ Imp_5(K) = Imp_4(K) \cup \{cdegi \to abfh\}$.

6. Then we need to calculate the sixth-layer attribute set of the formal context K. First, according to Theorem 2 and $Imp_5(K)$, we can skip the attribute set range of this layer is all sub-nodes of b and all sub-nodes of d. Then the remaining attribute set i.e. $\{T|\ T \leq abcdef \cup T \leq cdefgh\}$, where $T \leq abcdef$ are the elements of a in the reverse linear order which is less than or equal to $abcdef$, $T \leq cdefgh$ are the elements of c in the reverse linear order which are less than or equal to $cdefgh$. According to Definition 12, the correlation between this attribute set and $Imp_5(K)$ is judged in parallel. It can be seen that this attribute set is not related to $Imp_5(K)$. So $C_6(K) = C_5(K),\ Imp_6(K) = Imp_5(K)$.
7. Then we need to calculate the seventh-layer attribute set of the formal context K. First, according to Theorem 2, Theorem 1, and $Imp_6(K)$, we can skip the attribute set range of this layer is all sub-nodes of b. Then the remaining attribute set i.e. $\{T \cup cdefghi|\ T \leq abcdefg\}$, where $T \leq abcdefg$ are the elements of a in the reverse linear order which are less than or equal to $abcdefg$. According to Definition 12, we need to judge the correlation between this attribute set and $Imp_6(K)$ in parallel and it is judged parallelly whether the attribute set exists intent set by Definition 3. It can be seen that $\{abcdegh\}$ is a pseudo-intent and the rest attribute sets are not related to $Imp_6(K)$. So $C_7(K) = C_6(K) \cup \{abcdegh\},\ Imp_7(K) = Imp_6(K) \cup \{abcdegh \to fi\}$.
8. Then we need to calculate the eighth-layer attribute set of the formal context K. First, according to Theorem 2, Theorem 1, and $Imp_7(K)$, we can skip the attribute set range of this layer is all sub-nodes of a and b. Obviously, it can be seen that this attribute set is not related to $Imp_7(K)$. So $C_8(K) = C_7(K) \cup \{abcdegh\},\ Imp_8(K) = Imp_7(K)$.
9. For the ninth-layer attribute set of the formal context K, that is, we need the attribute set $\{abcdefghi\}$ to calculate. According to Theorem 1, $\{abcdefghi\}$ is an intent. So $C_9(K) = C_8(K),\ \{abcdefghi\},\ Imp_9(K) = Imp_8(K)$.

After the calculation of the above process, we know that the intent set of this formal context is $\{\emptyset, \{g\}, \{gi\}, \{cd\}, \{cdeg\}, \{cdefg\}, \{abcdh\}, \{abcd\ \ efghi\}\}$ and the stem base of this formal context is $Imp(K) = \{i \to g, h \to abcd, f \to cdeg, e \to cdg, d \to c, c \to d, cdegi \to abfh, b \to acdh, a \to bcdh, abcdegh \to fi\}$. Compared with the AEUS algorithm, we know the PAEUIS algorithm is correct with the stem base. And in the PAEUIS algorithm, because all attribute sets of the formal context are constructed a prefix dictionary tree with the inter-layer attribute set in the cardinality order and the inner attribute sets in the inverse linear order, it can guarantee the completeness of the PAEUIS algorithm. And after we skip attribute sets that are not related to the stem base, then intent sets are eliminated by the definition of intent, and then all pseudo-intent sets are obtained, so the non-redundancy of the algorithm can be guaranteed.

By comparing the two processes, we can see that when calculating implications of the AEUS algorithm, it needs to judge whether the implication is valid in the formal context, and because there are many implications whose consequent is $\emptyset$, it causes the redundancy calculation of $\emptyset^I$, besides, there are many redundant calculations in which the implication antecedent is the attribute set composed of the antecedent and consequent in acquired implication set, and because the

AEUS algorithm is still serial, the search space of the algorithm is very large when M is getting larger. However, the PAEUIS algorithm does not use linear order, the attribute set of the formal context is calculated according to the inter-layer cardinality order and the inverse linear order within the layer is used to calculate the attribute set. So, many attribute set intervals that are not related to the stem base are skipped in the calculation process for the inner-layer attribute sets. In addition, the PAEUIS algorithm can skip some calculations that are obviously intent sets through Theorem 1. Then, the PAEUIS algorithm uses the parallel computing implication method to save the running time of the algorithm for the remaining pseudo-intent attribute sets in the inner-layer.

5 Experiments and Analysis

5.1 Experimental Setting

In order to assess the effectiveness of the new improved algorithm proposed in this paper, we use the random function in Java language MATH library to simulate and generate a set of test data as formal context. The algorithm proposed in this paper (PAEUIS) is compared with parallel attribute exploration (PAE) [34], Shen Xiajiong's improved unrelated attribute exploration algorithm [32] (AEUS), and the traditional attribute exploration algorithm [1] (TAE). All of above algorithms were performed on a machine with 3.3 GHz CPU, 16 GB main memory and the Windows 10 operating system. Note that all of above algorithms were implemented in IDEA and JDK9. The experimental design is divided into three aspects. The first aspect is to change the form context scale and observe the change of the stem base quantity. The second aspect is to change the form context scale and observe the change of intent quantity. Third, we change the formal context scale and observe the time-consuming changes of the above algorithms.

In the first set of experiments, we set the formal context with the same number of objects as 20, and the number of attributes range from 0 to 50 at an interval of 10 to test. The purpose of the test is to fix the number of objects to change the number of attributes, and observe the change of the number of stem base. The test results are shown in Fig. 1.

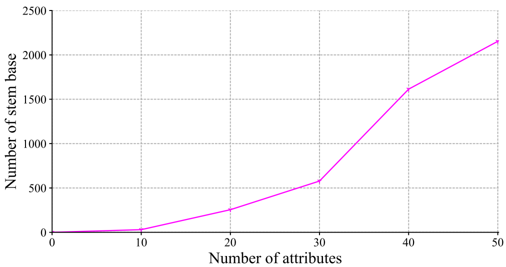

Fig. 1. The number of stem base with attributes change.

In the second set of experiments, we set the formal context with the same number of attributes as 20, and the number of objects range from 0 to 100 at an interval of 20 to test. The purpose of the test is to fix the number of attributes to change the number of objects, and observe the change of the number of stem base. The test results are shown in Fig. 2.

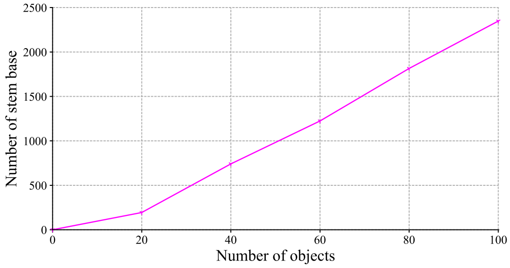

Fig. 2. The number of stem base with objects change.

In the third set of experiments, we set the formal context with the same number of objects as 20, and the number of attributes range from 0 to 50 at an interval of 10 to test. The purpose of the test is to fix the number of objects to change the number of attributes, and observe the change of the number of intent. The test results are shown in Fig. 3.

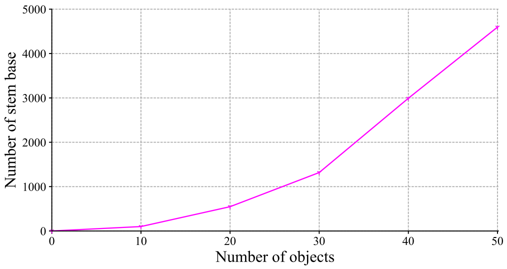

Fig. 3. The number of intent with attributes change.

In the fourth set of experiments, we set the formal context with the same number of attributes as 20, and the number of objects range from 0 to 100 at an interval of 20 to test. The purpose of the test is to fix the number of attributes to change the number of objects, and observe the change of the number of intent. The test results are shown in Fig. 4.

In the fifth set of experiments, we set the formal context with the same number of objects as 20, and the number of attributes range from 0 to 50 at an interval of 10 to test. The purpose of the test is to fix the number of objects to

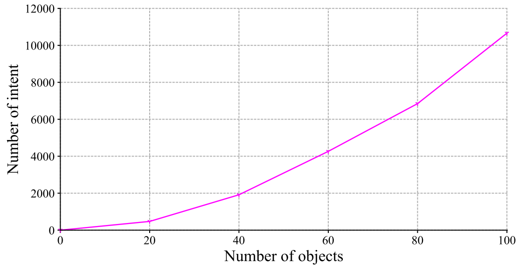

Fig. 4. The number of intent with objects change.

change the number of attributes, and observe the change of the time-consuming of above algorithms. The test results are shown in Fig. 5.

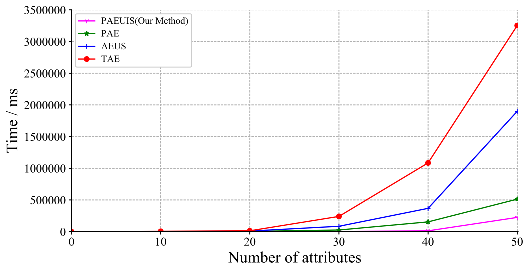

Fig. 5. The consuming-time of attributes change.

In the sixth set of experiments, we set the formal context with the same number of attributes as 20, and the number of objects range from 0 to 50 at an interval of 20 to test. The purpose of the test is to fix the number of attributes to change the number of objects, and observe the change of the number of above algorithms. The test results are shown in Fig. 6.

In the seventh experiment, we set an equal number of objects and attributes in the formal context. The number of objects range from 0 to 60 and were tested at an interval of 10. The purpose of the test is to change the scale of the formal context and observe the optimization efficiency of the PAEUIS algorithm. The test results are shown in Fig. 7.

5.2 Experimental Analysis

According to the first and second sets of experiments, it can be seen that the number of stem base will increase with the increase of the number of objects and attributes under the given formal context. At the same time, we compared the stem base obtained by PAE, AEUS and TAE algorithms, and found the results were consistent obtained by these four algorithms. Therefore, these two sets of

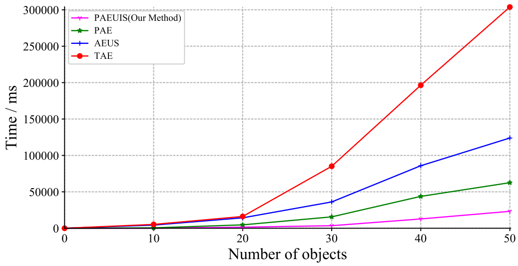

Fig. 6. The consuming-time of objects change.

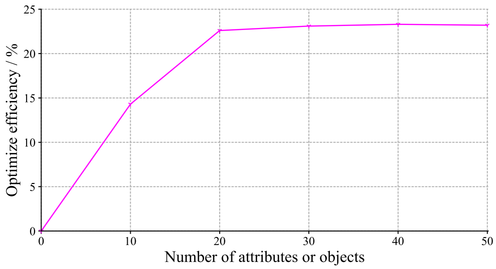

Fig. 7. The optimize efficiency of attributes or objects change.

experiments can be proved the correctness of the calculation of the stem base by the proposed PAEUIS algorithm.

According to the third and fourth sets of experiments, it can be seen that the number of stem base will increase with the increase of the number of objects and attributes under the given formal context. At the same time, we compared the intent obtained by PAE, AEUS and TAE algorithms, and found the results were consistent obtained by these four algorithms. Therefore, these two sets of experiments can be proved the correctness of the intent calculated by the PAEUIS algorithm proposed in this paper.

According to the fifth and sixth sets of experiments, it can be seen that the algorithm proposed in this paper takes less time than the three algorithms compared no matter how the number of objects is fixed and changed the number of attributes or the number of attributes is fixed and changed the number of objects. Moreover, the more the number of objects or attributes is, the more time would be saved by the algorithm proposed in this paper. Among them, in the fifth set of experiment, it can be seen that the PAEUIS algorithm only takes 18.9% time of AEUS algorithm when the number of attributes is 30. In the sixth set of experiment, it can be seen that the PAEUIS algorithm only takes 15.1% time of AEUS algorithm when the number of objects is 30. Furthermore, the seventh set of experiment shows that the PAEUIS algorithm can effectively reduce the scale of the implications calculated by the attribute exploration algorithm, and the number of implications reduced is about 23% compared with AEUS from an overall perspective.

The experimental results show that the proposed PAEUIS algorithm can effectively further reduce the time complexity of the attribute exploration algorithm.

6 Conclusion and Further Research

Based on the traditional attribute exploration algorithm and the improved attribute exploration algorithm under the unrelated set, first, this paper makes further research on the problem of high time complexity of the algorithm. This paper hopes that the algorithm skips some attribute sets that are neither intent nor pseudo-intent and certain intent to reduce the time complexity of the algorithm. Second, this article further summarizes three theorems based on the irrelevant definition of the AEUS algorithm, and makes a rigorous mathematical demonstration of the proposed theorems. Finally, a further improved attribute exploration algorithm (PAEUIS) is given based on these theorems. This algorithm constructs a prefix dictionary tree by taking all the attribute sets of the formal context according to the inter-layer cardinal order and the inverse linear order within the layer. With the help of the above theorems, it automatically jumps the attribute sets that are not related to the stem base and that are obviously intent sets in parallel when calculating the implications of the attribute set in attribute exploration. So the search space of the algorithm is further reduced, and then the implications of the pseudo-intent set are calculated in parallel, reducing the calculation scale of implication. Because attribute exploration can semi-automatically assist users in acquiring unknown knowledge, the next step is to combine attribute exploration with deep learning to further develop an automatic learning system using attribute exploration.

References

1. Ganter, B.: Formal Concept Analysis: Mathematical Foundations, 1st edn. Springer, Heidelberg (1999)
2. Li, J., et al.: Concept lattice theory and method and their research prospect. Pattern Recogn. Artif. Intell. **33**(7), 619–642 (2020)
3. Zhang, L., Zhang, H., Shen, X., Yin, L.: A bottom-up algorithm of vertical assembling concept lattices. Int. J. Data Min. Bioinform. **7**, 229–244 (2013)
4. Yang, J., Shen, X., Chen, W., Ge, Q., Zhang, L., Chen, H.: A model study on collaborative learning and exploration of RBAC roles. Wirel. Commun. Mob. Comput. (2021). https://doi.org/10.1155/2021/5549109
5. Shemis, E., Mohammed, A.: A comprehensive review on updating concept lattices and its application in updating association rules. Wiley Interdisc. Rev. Data Min. Knowl. Discov. **11**(2), e1401 (2021)
6. Shen, X., Yang, J., Zhang, L., Yang, G.: An interactive role learning and discovery model for multi-department RBAC building based on attribute exploration. J. Ambient Intell. Humaniz. Comput. (2020). https://doi.org/10.1007/s12652-020-02634-3

7. Zhang, L., Huang, Z., Liu, W., Guo, Z., Zhang, Z.: Weather radar echo prediction method based on convolution neural network and Long Short-Term memory networks for sustainable e-Agriculture. J. Clean. Prod. (2021). https://doi.org/10.1016/j.jclepro.2021.126776
8. Zhang, L., Xu, C., Gao, Y., Han, Y., Du, X., Tian, Z.: Improved Dota2 lineup recommendation model based on a bidirectional LSTM. Tsinghua Sci. Technol. **25**(6), 712–720 (2020)
9. Wei, L., Liu, L., Qi, J., Qian, T.: Rules acquisition of formal decision contexts based on three-way concept lattices. Inf. Sci. **516**, 529–544 (2020). https://doi.org/10.1016/j.ins.2019.12.024
10. Zhang, L., Tang, S., Lv, L.: An finite iterative algorithm for sloving periodic Sylvester bimatrix equations. J. Frankl. Inst. **357**(15), 10757–10772 (2020)
11. Hua, M., Zhen, Z., Xiaoqing, L.: A method for mining decision rules based on decision formal context. J. Hebei Univ. (Nat. Sci. Ed.) **9**(01), 1–6 (2021)
12. Ling, W., Siyu, Z.: Granules and knowledge structure in three-way concept analysis. J. Northwest Univ. (Nat. Sci. Ed.) **50**(04), 537–545 (2020)
13. Xueyan, Z.: Research on role updating technology based on concept lattices. Inner Mongolia University of Science and Technology (2020)
14. Jingyu, W., Junqing, L., Yuesheng, T.: Design of role based access control for triadic concept analysis. J. Harbin Univ. Sci. Technol. **25**(02), 31–37 (2020)
15. Tang, J., Wei, L., Ren, R., Zhao, S.: Granule description using possible attribute analysis. J. Shandong Univ. (Nat. Sci. Ed.) **56**(01), 75–82 (2021)
16. Mi, Y., Liu, W., Shi, Y., Li, J.: Semi-supervised concept learning by concept-cognitive learning and concept space. IEEE Trans. Knowl. Data Eng. (99), 11 (2020)
17. Huilai, Z., Li, Y.: Knowledge representation based on concept cluster. J. Northwest Univ. (Nat. Sci. Ed.) **50**(04), 529–536 (2020)
18. Yan, E., Yu, C., Lu, L., Hong, W., Tang, C.: Incremental concept cognitive learning based on three-way partial order structure. Knowl.-Based Syst. **220**(23), 106898 (2021). https://doi.org/10.1016/j.knosys.2021.106898
19. Li, J.H., He, J.J., Wu, W.Z.: Optimization of class-attribute block in multi-granularity formal concept analysis. J. Shandong Univ. (Nat. Sci.) **55**(05), 1–12 (2020)
20. Zhou, Y., Li, J.: Skill reduction and assessment in formal context. Comput. Sci. Explor. 1–10 (2021)
21. Lei, Z., Yu, H., Qiang, G.: A privacy protection scheme for IoT big data based on time and frequency limitation. Wirel. Commun. Mob. Comput. (3), 1–10 (2021)
22. Lv, L., et al.: Contract and Lyapunov optimization based load scheduling and energy management for UAV charging stations. IEEE Trans. Green Commun. Netw. https://doi.org/10.1109/TGCN.2021.3085561
23. Ganter, B., Obiedkov, S.: Conceptual Exploration. 1st edn. Springer, Heidelberg (2016)
24. Borchmann, D.: A general form of attribute exploration. ArXiv. vol. abs/1202.4824 (2012)
25. Jäschke, R., Rudolph, S.: Attribute exploration on the web. Computer Science (2013)
26. Potoniec, J., Rudolph, S., Lawrynowicz, A.: Towards combining machine learning with attribute exploration for ontology refinement. In: International Semantic Web Conference (2014)
27. Hanika, T., Zumbragel, J.: Towards collaborative conceptual exploration. ArXiv. vol. abs/1712.08858 (2018)

28. Codocedo, V., Baixeries, J., Kaytoue, M., Napoli, A.: Sampling representation contexts with attribute exploration. In: Cristea, D., Le Ber, F., Sertkaya, B. (eds.) ICFCA 2019. LNCS (LNAI), vol. 11511, pp. 307–314. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-21462-3_20
29. Reppe, H.: Attribute exploration using implications with proper premises. In: Eklund, P., Haemmerlé, O. (eds.) ICCS-ConceptStruct 2008. LNCS (LNAI), vol. 5113, pp. 161–174. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-70596-3_11
30. Ryssel, U., Distel, F., Borchmann, D.: Fast algorithms for implication bases and attribute exploration using proper premises. Ann. Math. Artif. Intell. **70**(1-2), 25–35 (2013)
31. Xiaoxiang, Z., Ping, Q.: On attribute exploration algorithms. J. Front. Comput. Sci. Technol. (2009)
32. Xiajiong, S., Jiyong, Y., Lei, Z.: Attribute discovery algorithm based on unrelated attribute sets. Computer Science (2020)
33. Distel, F.: Hardness of enumerating pseudo-intents in the lectic order. In: Kwuida, L., Sertkaya, B. (eds.) ICFCA 2010. LNCS (LNAI), vol. 5986, pp. 124–137. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-11928-6_9
34. Kriegel, F.: Parallel Attribute Exploration. In: Haemmerlé, O., Stapleton, G., Faron Zucker, C. (eds.) ICCS 2016. LNCS (LNAI), vol. 9717, pp. 91–106. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40985-6_8
35. Shen, X., Yang, J., Zhang, L., Yang, G.: An interactive role learning and discovery model for multi-department RBAC building based on attribute exploration. J. Ambient Intell. Humaniz. Comput. (2020)