Fairness-Aware Maximal Cliques Identification in Attributed Social Networks With Concept-Cognitive Learning

Min Tao, Fei Hao, Ling Wei, Huilai Zhi, Sergei O. Kuznetsov, and Geyong Min, *Member, IEEE*

***Abstract*—Attributed social networks are pervasive in real life and play a crucial role in shaping various aspects of society. These networks not only capture the connections between individuals but also encompass the associated attributes and characteristics. Analyzing and understanding these attributes provide insights into social behaviors, information diffusion patterns, and the formation of influential communities. Consequently, we propose a novel algorithm for detecting fairness-aware maximal cliques in the attributed social networks. We extract the concept lattice of attributed social networks and quantify these concepts using the concept stability and fairness measures defined in this article. By utilizing the proposed fairness-aware distance, we identify fairness-aware maximal cliques within attributed social networks. The effectiveness of the algorithm is then validated using five real-world network datasets. Experimental results fully demonstrate the effectiveness and scalability of our approach in identifying key structures, analyzing attribute networks, and promoting the development of responsible computational systems.**

***Index Terms*—Attributed social networks, cliques, fairness, formal concept analysis (FCA), stability.**

I. Introduction

In reality, people's daily communication and information transmission are increasingly inseparable from online social networks (OSNs), and these social networks generate massive social big data every day. Thus, the research on social network analysis has become a hot topic. Researchers from the area of social computing usually model OSNs, explore their topological relationships, analyze network influence, and so forth [1], [2]. From the communication point of view, the nature of communication in OSNs, communication behavior and privacy, is also widely investigated [1], [3]. However, due to the large size of these OSNs, it will be difficult work to perform social network analysis directly on the entire social network where the network structures barely change [4]. Thus, many researchers focus on mining key structures within social networks to obtain social intelligence and computational social services. Among these key structures, the clique is the most well-known subgraph structure since many real-life problems can be modeled as clique structures. For example, in a social network, nodes are different users, and edges represent users' mutual acquaintances. Then, the clique structure represents a group of people who know each other. In addition, the clique structure also has many applications in other fields, such as the discovery of abnormal trading groups in financial networks [5], epidemic risk monitoring [6], and protein structure mining [7]. However, most existing research works primarily focus on mining structures in the graph while ignoring the attribute information of the nodes themselves, thus paying insufficient attention to the internal fairness of these structures [8], [9]. As a consequence, they could lead to discrimination toward certain populations when exploited in human-centric computational social applications.

As the main information representation model of various ubiquitous networks, graphs are widely used in network modeling, analysis, and mining. Based on this model, a variety of topological structure mining methods have been proposed (such as maximal clique enumeration [10], k-clique community detection [11], [12], and k-clique mining [13]). These clique structures have huge implications in many disciplines such as sociology, biological sciences, and chemistry potential and application value. However, most of the existing research works mainly focus on mining the structures in the graph, while ignoring the attribute information of the nodes themselves, resulting in insufficient attention to the internal fairness of these structures [9]. Therefore, when they are used in anthropocentric computing social applications, there may be issues of discrimination against certain groups.

Meanwhile, along with the widespread use of responsible artificial intelligence systems in various fields, there has been a

Manuscript received 18 March 2024; revised 11 July 2024; accepted 11 August 2024. This work was supported in part by the Ministry of Education Humanities and Social Sciences Research Youth Fund Project under Grant 22YJCZH046 and in part by the National Natural Science Foundation of China under Grant 62477029 and Grant 12171392. The work of Sergei O. Kuznetsov was supported by the Basic Research Program and computational resources of HPC facilities at HSE University. *(Corresponding author: Fei Hao.)*

Min Tao and Fei Hao are with the School of Computer Science, Shaanxi Normal University, Xi'an 710119, China (e-mail: feehao@gmail.com).

Ling Wei is with the School of Mathematics, Northwest University, Xi'an 710100, China.

Huilai Zhi is with the Faculty of Mathematics and Computer Science, Quanzhou Normal University, Quanzhou 362000, China.

Sergei O. Kuznetsov is with the School of Data Analysis and Artificial Intelligence, National Research University Higher School of Economics, Moscow 109028, Russia.

Geyong Min is with the Department of Computer Science, University of Exeter, Exeter EX4 4QF, U.K.

Digital Object Identifier 10.1109/TCSS.2024.3445721

Fig. 1. Picture of two brides. The left is the American bride, and the right is the northern Indian bride.

growing interest and emphasis on the concept of "fairness" in both academia and industries. Although intelligent algorithms are believed to avoid subjective judgments and make decisions more objective and fair, real-world research has shown that this is not entirely accurate [14]. The reason for this is that many machine learning models rely on training datasets, and if there is sexism, racial discrimination, or other factors that may lead to inequity, machine learning models may reflect this inequity. For instance, if mammograms of white women are used for training, but breast cancer data are distributed differently between black and white women, it is unfair that breast cancer prediction algorithms are less accurate for black women [15]. Similarly, it is unjust that African–American offenders are more likely to be judged as potentially higher risk in the U.S. justice system's COMPAS system, making it harder for them to be released [16]. Surprisingly, these inequities still exist in some widely known areas [17]. For example, deep neural network learning models trained on the widely known ImageNet dataset mistakenly identify the image on the left in Fig. 1 as a bride while identifying the image on the right as performance art or clothing. This is because over 45% of the data in the ImageNet dataset come from the United States, which has only 4% of the world's population, while data from China and India, two of the world's most populated countries, make up only 3% [18]. Therefore, the concept of fairness is of crucial significance in practical applications.

Therefore, it is necessary to consider the fairness of structures when mining networks for implementing various responsible computational social systems. Notably, the cohesiveness of structures can be characterized by stability, while structural equity can be measured by fairness [19]. Therefore, this article adopts stability (measure the cohesiveness of structures in the network) and fairness (evaluate the difference with the number of attributes for nodes of structures in the network) as metrics for mining fairness-aware maximal cliques.

The stability of the formal concept was introduced in [20] and can be used to measure the cohesiveness of structures in the network [21]. This definition is very beneficial for our research. In this study, our focus is on identifying the fairness-aware maximal cliques in attributed social networks. More specifically, we propose a new method for identification by quantitatively defining both stability and fairness. To assess the effectiveness of our approach, we validate and evaluate it using real network datasets. Furthermore, we conduct case study to demonstrate the practical applications of the issues addressed in this research. The main contributions of this article are summarized as follows.

1) *Quantitative definitions of structure stability and fairness*: In concept-cognitive learning (CCL), we quantify the mining demands in different scenarios and introduce the concept stability and fairness. Specifically, we use stability to describe the degree of connection within the structure and fairness to measure the fairness level among various attributes within the structures.
2) *An effective fairness-aware key structures identification algorithm*: Considering the particularity of networks, we propose an interestingness-based key structure mining algorithm. The aim is to enhance the utilization of knowledge concealed in the concept lattice more effectively. The algorithm is suitable for mining fairness-aware maximal cliques in the networks. Experimental results show that the algorithm can effectively identify fairness-aware maximal cliques in attributed social networks.
3) *Comprehensive evaluation*: Additionally, we conduct extensive experimental validation and evaluation on five real network datasets. Experimental results demonstrate that the approach accomplishes fairness-aware maximal cliques recognition tasks within a unified CCL framework.

The subsequent structure of this article is as follows. Section II describes the related work in detail. Section III provides preliminary knowledge and a clear definition of the research problem addressed in this article. The approach for obtaining fairness-aware maximal cliques from attributed social networks is presented in Section IV. Section V presents the experimental results and analysis. Finally, Section VI concludes this article and proposes future research prospects.

II. Related Work

In this section, we present a review of related research on CCL and fairness-aware structure mining.

A. Concept Cognition Measures

An important approach in CCL is formal concept analysis (FCA), which regards concepts as fundamental cognitive units and can effectively simulate real-world entities. As a powerful computational intelligence method, FCA vividly embodies the generalization and specialization relationship between concepts through the Hasse diagrams; hence it is widely used in various fields such as data mining, machine learning, and software engineering. Since FCA only considers positive attributes (objects), it supports binary decisions of acceptance and rejection only. Consequently, many researchers have shifted their focus onto concept learning that combines three-way concept analysis (3WCA) and FCA [22], [23]. Hao et al. proposed an incremental three-way concept lattice construction algorithm for knowledge discovery from social networks, solving the problem of dynamically growing formal context [24]. Chunduri et al. developed a parallel algorithm for concept generation and three-way concept lattice construction in large datasets [25].

However, recent researches have discovered that noisy contexts can impair the cognitive quality of concepts [26]. Therefore, how to improve the legibility of concepts and select useful ones has become a concern. In response, numerous researchers have focused on concept interestingness, such as stability and separation. Among them, concept stability is a prominent index employed to evaluate the degree of concept connection and has been proven to be NP-complete [27]. Concerning the calculation of concept stability, Mouakher et al. proposed an algorithm called DFSP, which calculates the stability by minimizing the search space and intelligently computing the generator [28]. However, this approach does not consider the implicit knowledge held within the relationship between concepts. To address this shortcoming, Gao et al. improved the DFSP by proposing a new algorithm for concept stability calculation, known as CICal [29], which is specially tailored for social network scenarios and plays a critical role in this article.

B. Fairness-Aware Structures Detection

1) Attributed Graph Mining: Our work considers the attribute information of nodes, which pertains to attributed graph mining. There have been many studies on attributed graph mining, such as Li et al. proposed an embedding-based model to discover communities in attributed graphs [30]. Fang et al. developed an index structure called CL-tree to support efficient attributed community search while studying the attributed community search problem [31]. Xie et al. considered both influence and node attributes and evaluated the community cohesion by defining a new scoring function to discover the attributed pkd-truss community [32].

2) Fairness-Aware Data Mining: Roughly speaking, fairness-aware data mining over graphs requires that the algorithm should not yield discriminatory predictions or decisions against individuals from any specific sensitive subgroup.

In addition, Pizzuti and Socievole proposed an attributed graph community mining algorithm that considered both node similarity and structural connectivity to mine fairness-aware cohesive subgraphs in attributed graphs [33]. Zhang et al. [34] and Pan et al. [35] first considered fairness in graph mining and proposed weak and strong fairness clique models as fairness representations of mining results. Recently, Hao et al. introduced fairness into the clique model and proposed the AFCMiner algorithm to extract fairness cliques from networks [36].

In contrast to other works on attributed graph mining, this article simultaneously characterizes the stability and fairness of conceptual structures. From this assessment, we identify these essential fairness-aware maximal cliques that we aim to mine. To obtain these structures, we employ the FCA approach to mine the attributed social network, which results in the acquisition of target structures.

III. Problem Formulation

In this section, we introduce the preliminary knowledge about maximal clique, CCL theory. Additionally, we formally describe the problem addressed in this study.

We first introduce the maximal clique, a structure of significant interest in contemporary network mining research. Next, we provide a comprehensive overview of the principles and applications of the CCL theory, which is a critical theoretical foundation for identifying key structures from social networks. Last, we provide a formal description of the research problem addressed in this article, clarifying our research goals and specific objectives.

A. Maximal Clique

Maximal clique is a crucial aspect of network analysis and has been widely applied in various fields. As the most fundamental cohesive subgraph model, maximal clique is commonly used to expose the dense cluster structure of the network. This structure is frequently utilized in overlapping community mining, protein structure identification, and abnormal transaction identification in financial networks [36], [37].

In this article, we focus on the structure obtained from the attributed social network, which is a special type of maximal clique—the fairness maximal clique. Consequently, the maximal clique is defined as follows.

Definition 1 (Maximal Clique) [38]: If we denote a network by G=(V, E), where V is the set of nodes and E is the set of edges, a subset $C \subset V$ is considered a clique in G if any two nodes $v_i, v_j \in C$ are connected by an edge $(v_i, v_j) \in E$.

If it is impossible to find a node v_0 from $V \backslash C$ such that $C \cup \{v_0\}$ is still a clique, then we call C a maximal clique.

B. CCL

In CCL, FCA is a classic intelligent computing technology. It uses induction to summarize relationships between objects and attributes, which are expressed in the form of concepts. These concepts form a hierarchy with clear partial order relationships, greatly facilitating data understanding and mining [39].

Definition 2 (Formal Context) [40]: The basic structure of FCA is the triple FC=(O, A, I) called *(formal) context*, where O and A denote the sets of objects and attributes, respectively. The incidence binary relation $I \subseteq O \times A$ between O and A defines whether object $o \in O$ has attribute $a \in A$. A formal context is usually represented by a binary object-attribute matrix or by a cross table.

Derivation operations or *primes* are defined for a context (O, A, I) as follows. For $X \subseteq O$ and $Y \subseteq A$, one has the following:

$$X' = \{a \in A | \forall o \in X, (o, a) \in I\} \tag{1}$$

$$Y' = \{o \in O | \forall a \in Y, (o, a) \in I\}. \tag{2}$$

For a subset of objects X, set X' is the set of all common attributes of objects from X; for a subset of attributes Y, set Y' is the set of all objects having all attributes from Y.

Now *(formal) concept* is defined as follows.

Definition 3 (Concept) [41]: A pair (X, Y) where $X \subseteq O$ and $Y \subseteq A$ is a *(formal) concept* if $X' = Y$ and $Y' = X$. If (X, Y) is a formal concept, X is called *extent* and Y is called *intent* of the concept.

There is a natural partial order $\leq$ on the set of formal concepts of a context defined in the following way:

$$(A, B) \leq (C, D) \text{ if } A \subseteq C (\Longleftrightarrow B \supseteq D). \tag{3}$$

The partial order on concepts makes an algebraic lattice—called *concept lattice*—so that every two concepts have a supremum (join) and infimum (meet) w.r.t. $\leq$.

If there is no (X, Y) such that $(A, B) < (X, Y) < (C, D)$, then (A, B) is called the lower neighbor of (C, D) and (C, D) is called the upper neighbor of (A, B), denoted by $(A, B) \prec (C, D)$. Relation $\prec$ is called the *covering relation* of the concept lattice.

Definition 4 (Generator) [26]: For a formal context FC=(O, A, R) and a formal concept (A, B) of FC, subset $P \subseteq A$ ($Q \subseteq B$) is an extent (intent) generator of concept (A, B) if $P' = B$ ($Q' = A$).

Definition 5 (Stability) [26]: Let $\text{Gen}(A, B)$ denotes the number of all extent generators of the concept (A, B), then *(intentional) stability* $\sigma(A, B)$ the concept (A, B) is defined as follows:

$$\sigma(A, B) = \frac{|\{P \subseteq A \mid P' = B\}|}{2^{|A|}} = \frac{\text{Gen}}{2^{|A|}}. \tag{4}$$

C. Problem Description

This article mainly adopts the CCL theory to identify fairness-aware maximal cliques in IoT networks.

First, the concept of fairness is presented. Basically, fairness can usually be divided into several aspects such as group fairness, individual fairness, counterfactual fairness, degree-related fairness, and application-specific fairness. However, these concepts of fairness often cannot be satisfied simultaneously and sometimes even conflict with each other [42]. For our application scenario, we believe that the selected structure is regarded as the fairest structure when the number of nodes corresponding to different attribute values is equal; and the greater the difference, the more unfair the structure is. Then, a formalism of the problem addressed in this article can be presented as follows.

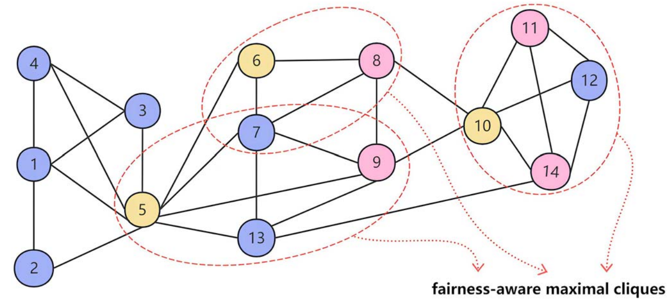

Fig. 2. Results for the fairness-aware maximal cliques of network G.

1) Fairness-Aware Maximal Cliques Identification: Given an attributed social network $G = (V, E, M)$, where V represents a set of nodes, E represents a set of edges that capture the relationships between nodes, and M denotes an attribute set that records self-attribute information for the nodes. The aim of the identification task is to obtain all fairness-aware maximal cliques in this attribute network.

The problem of identifying maximal cliques has been proven to be NP-hard [43]. Moreover, fairness considerations make the obtained cliques more complex.

To better illustrate our problem, here is a corresponding case.

Example 1: Fig. 2 shows the topology of a simple attributed social network G, in which the color of the node represents the attribute value of the node. In this network, the attributes of the nodes have three different values. After structures identification, all fairness-aware maximal cliques can be extracted.

As can be seen from Fig. 2, for such an attributed social network containing 14 nodes, the final fairness-aware maximal cliques that can be obtained are node sets 6,7,8,5,7,9,13, and 10,11,12,14. It is worth noting that the maximum clique 1,3,4,5 does not meet our requirement that the difference in the number of nodes corresponding to any two attributes is not greater than 1, so it is not the result we need. It also needs to be noted that we also require that each attribute value corresponds to at least one node in the fairness maximal clique. We will describe the adopted mining method in detail in subsequent sections.

IV. CCL-Based Fairness-Aware Maximal Cliques Detection

In this section, we propose a fairness-aware maximal cliques detection approach based on CCL theory. To better present the approach proposed in this article, it is elaborated with the following parts: 1) formalize the network and construct the corresponding concept lattice; 2) give the quantification approach of concept stability and fairness, as well as the formula for calculating the importance of stability and fairness; and 3) propose the fairness-aware maximal cliques detection algorithm in this article.

A. Overview of the Proposed Approach

Fig. 3 shows the algorithm framework proposed in this article for identifying fairness-aware maximal cliques in attributed social networks based on CCL. This framework is used to detect

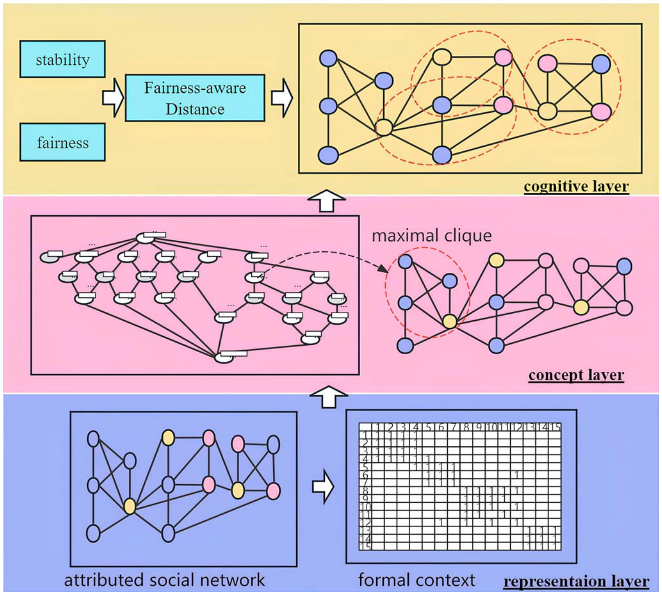

Fig. 3. Framework for the fairness-aware key structures identification algorithm.

all of the fairness-aware maximal cliques existing in attributed social networks. It is divided into the following three layers: the representation layer, the concept layer, and the cognitive layer. The function of each layer is introduced as follows.

In the representation layer, we represent the network as a formal context as input for FCA. Then, concepts are extracted from the formal context through FCA, and the concept lattice is also generated. In the cognitive layer, knowledge hidden in the concepts is mainly learned through the assessment of concept stability and fairness that facilitate the selection of the fairness-aware maximal cliques by fairness-aware distance calculation.

B. Formal Context Construction

In this article, the network is formalized as an undirected graph that is a classic mathematical problem. Therefore, a network can be represented as a graph $G = (V, E)$, where V is the set of vertices in the graph, representing individuals in the network, and E is the set of edges in the network, representing the relationship between individuals in the network. For FCA, the input is the formal context FC=(O, A, R). If each vertex in the graph G is regarded as an object and an attribute of FC at the same time, then the formal context can be represented as $FC = (V, V, R)$, where R is the binary relationship between vertices, which happens to be very similar to the adjacency matrix represented by the graph. Based on this idea, the formal context of networks can be constructed. Intuitively, we provide the modified adjacency matrix which is equivalent to the formal context $FC = (V, V, R)$.

Definition 6 (Modified Adjacency Matrix) [12]: Let $G = (V, E)$ be a network with the vertex set $V = \{v_1, v_2, ..., v_n\}$. The rule for constructing the modified adjacency matrix M with size $n \times n$ is as follows:

$$M = \begin{cases} m_{ij} = 1, & \text{if} \quad i = j \\ m_{ij} = 1, & \text{if} \quad (v_i, v_j) \in E \\ m_{ij} = 0, & \text{otherwise.} \end{cases} \tag{5}$$

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1	1	1	1	1									
2	1	1			1									
3	1		1	1	1									
4	1		1	1	1									
5	1	1	1	1	1	1	1		1				1	
6					1	1	1	1						
7					1	1	1	1	1				1	
8						1	1	1	1	1				
9					1		1	1	1	1			1	
10								1	1	1	1	1		1
11										1	1	1		1
12										1	1	1		1
13					1		1		1				1	1
14										1	1	1	1	1

Fig. 4. Modified adjacency matrix of network G.

The modified adjacency matrix is equivalent to the formal context $FC = (V, V, R)$ corresponding to the network. After completing the construction of the formal context, the corresponding concept lattice can be generated by using the FCA method in Section III-B.

Example 2: According to the above definition, a simple network G, containing 14 nodes and 28 edges (see Fig. 2), can be transformed into a modified adjacency matrix, namely the formal context, as shown in Fig. 4 (note: we only represent the one value in the matrix). Taking this formal context as the input of FCA, the corresponding concept lattice can be obtained, as shown in Fig. 5. Each node in the lattice represents a concept, and the numbers within the nodes indicate concept numbers.

To differentiate between the extent and intent of each concept, we denote the intent by adding the letter "a" in front of the numbers. For instance, concept 23 represents an equiconcept that was defined in [12] (that is, the extent and intent are the same for a symmetry formal context), thus its extent is $\{7, 8, 9\}$, and its intent is also $\{7, 8, 9\}$.

In the concept lattice of Fig. 5, we have marked the equiconcepts and the fairness equiconcepts. The fairness equiconcepts correspond to the fairness-aware maximal cliques we want to find, which are the results we want to obtain from the network.

C. Fairness-Aware Key Structures Detection

We utilize formal concepts to sample attributed social networks, with the generated subgraphs representing the union of concept extents and intents. To quantify the stability and fairness of concept and our structure selection criteria, this section presents the relevant quantitative indicators.

1) Concept Stability: For network G, its formal context is $FC = (V, V, R)$. For a formal concept (A, B) of FC, we

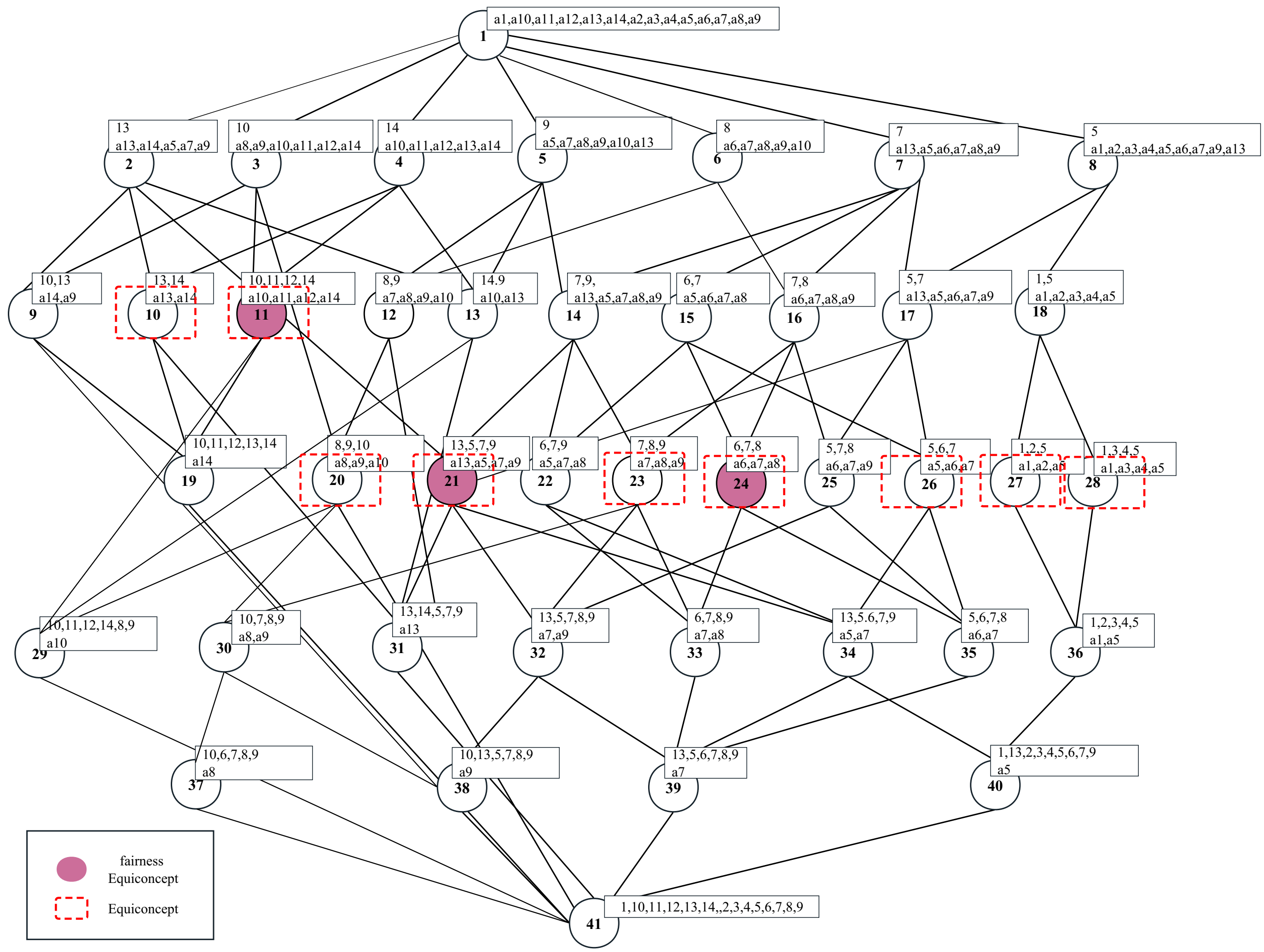

Fig. 5. Concept lattice of network G. Equiconcept is equivalent to the maximal clique in the network, and fairness equiconcept is the fairness-aware maximal clique.

consider its (intentional) stability

$$\sigma(A, B) = \frac{|\{P \subseteq A \mid P' = B\}|}{2^{|A|}} = \frac{|\mathrm{Gen}(A, B)|}{2^{|A|}}$$

according to Definition 5.

However, during the process of calculating the generators, the edges connecting the concept itself and the nodes outside the concept can affect the number of generators, resulting in a reduction in the final number. That is, the influence on the edges between concepts and external vertices in the original definition may lead to different stability for two concepts with the same internal structure, which is inconsistent with our expectations. In this article, we pay more attention to the internal stability of the obtained structure; hence, this article makes some improvements to the definition of generator.

Definition 7 (Cluster-Oriented Generator): For a network G, its formal context is FC=(V, V, R). For a formal concept (A, B) in the formal context FC, a subset $P \subseteq A$ is called a cluster-oriented generator of the concept (A, B) if $P' = B$ is satisfied in the new formal context FC_1=$(A \cup B, A \cup B, R_1)$. Note that $R_1 = \{\{v_1, v_2\}|v_1, v_2 \in A \cup B\} \subseteq R$ and R_1 is the binary relationship between nodes in $A \cup B$.

The following Example illustrates the advantages of our proposed cluster-oriented generator.

Example 3: As shown in Fig. 2, the internal structures of the two concepts $\{\{5, 7, 9, 13\}, \{5, 7, 9, 13\}\}$ and $\{\{10, 11, 12, 14\}, \{10, 11, 12, 14\}\}$ are identical. Based on Definition 4, the former has nine generators, while the latter has 13 generators, resulting in stability degrees of $(9/16) = 0.5625$ and $(13/16) = 0.8125$, respectively. However, both concepts can exhibit new stability degrees of $(15/16) = 0.9375$ by utilizing our proposed cluster-oriented generator, indicating that our improved approach more accurately reflects stability within the concept.

The calculations conducted using our proposed cluster-oriented generator can better capture the internal stability of concepts. A higher stability indicates a closer connection between the internal nodes of the corresponding subattributed social network, while lower stability suggests looser connections between internal nodes of the corresponding subattributed

Algorithm 1: CalStabilty: Calculation of Stability

Input:
The Concept lattice L
Output:
Set of stability for each concept $\mathbb{K}$;

```
Initialize 𝕂=∅
begin
for ∀(A, B) ∈ L
    begin
        CGen=0
        for ∀A* ⊂ A
            if (A*)′=B
            CGen+=1
        ϕ=CGen/2^|A|
        𝕂=𝕂 ∪ ϕ
    end
return 𝕂
end
```

Algorithm 2: CalFairness: Calculation of Concept Fairness

Input:
The concept lattice L
The attributed social network G=(V, E, M)
Output:
Set of fariness for each concept $\mathbb{C}$;

```
Initialize ℂ=∅,ℂ′=∅
begin
calculate Unfairness Index set ℂ′ according to Definition 8
max = max(ℂ′)
for ∀c ∈ ℂ′
    begin
        φ=2^c/2^max
        ℂ=ℂ ∪ φ
    end
return ℂ, max
end
```

social network. We provide the pseudocode for the stability calculation, as shown in Algorithm 1, as follows.

2) Concept Fairness: For a given attributed social network $G = (V, E, M)$, where V represents the vertex set, E represents the edge set, and M represents the attribute set, after using the FCA method, the concept set C=$\{(A_1, B_1), (A_2, B_2), ..., (A_n, B_n)\}$ can be obtained (n is the number of concepts obtained from G). Therefore, we define the fairness index of the concept as follows.

Definition 8 (Fairness Index): For the concept (A_i, B_i), a corresponding subattributed social network $G_1 = (A_i \cup B_i, E_1, M)$ can be constructed, where $A_i \cup B_i$ is the node set, $E_1 = \{\{v_1, v_2\}|v_1 \in A_i \cup B_i, v_2 \in A_i \cup B_i\}$, M is the attribute set in attributed social network G. Then, the fairness index $idx(A_i, B_i)$ of the concept (A_i, B_i) is as follows:

$$idx(A_i, B_i) = \sum_{1 \leq j < k \leq |M|} |a_j - a_k|, \qquad i \neq j. \tag{6}$$

Here, a_j denotes the number of nodes in subattributed social network G' that has the attribute $m_j \in M$. If the number of nodes corresponding to each attribute is equal in the subattributed social network G', the idx is 0, indicating that the subattributed social network is absolute fair. The larger the value of idx, the lower the fairness among attributes in the subattributed social network. Next, we consider mapping it to the 0–1 interval.

Definition 9 (Fairness): For the concept set $C = \{(A_1, B_1), (A_2, B_2), ..., (A_n, B_n)\}$, the corresponding fairness index set is $I = \{idx(A_1, B_1), idx(A_2, B_2), ..., idx(A_n, B_n)\}$. We can map each fairness index to the range [0, 1] by the following formula:

$$\varphi(A_i, B_i) = \frac{2^{idx(A_i, B_i)}}{2^{\max(I)}}. \tag{7}$$

After normalization, it is observed that the closer the idx_i is to 1, the more unfairness the structure is; the closer the idx_i is to 0, the fairer the structure is. This quantification is a good description of the degree of fairness of the structure. Therefore, the fairness is an effective measure of fairness within a given attributed social network. Lower scores indicate a more even distribution of attributions among nodes in the subgraph, while higher scores reflect the distribution of attributions among nodes is more different. We present the pseudocode for concept fairness calculation, as shown in Algorithm 2, as follows.

Example 4: According to Example 2, 41 concepts can be extracted from the attributed social network G (see Fig. 5). Following the definitions of stability and fairness given earlier, the stability and fairness of each concept can be calculated.

Finally, we give the definition of fairness-aware distance for concept selection.

3) Concept Fairness-Aware Distance: For concept selection, this study uses the following formula for distance calculation to determine the fairness-aware maximal cliques we ultimately need to mine.

Definition 10 (Fairness-Aware Distance): For a concept (A, B), whose stability is $\sigma(A, B)$ and fairness is $\varphi(A, B)$, then the fairness-aware distance $\ell(A, B)$ of the concept (A, B) is as follows:

$$\begin{aligned}\ell(A, B) &= (1 - \sigma(A, B)) * \varphi(A, B)\\ &= \left(1 - \frac{|\mathrm{Gen}(A, B)|}{2^{|A|}}\right) * \frac{2^{idx(A_i, B_i)}}{2^{\max(I)}}.\end{aligned} \tag{8}$$

After providing the definition of fairness-aware distance, we can get the following conclusion.

For the concept set C=$\{(A_1, B_1), (A_2, B_2), ..., (A_n, B_n)\}$, the corresponding fairness index set is $I = \{idx(A_1, B_1), idx(A_2, B_2), ..., idx(A_n, B_n)\}$. The fairness-aware distance of the fairness equiconcept must be less than $(1/2^{\max(I)})$.

Proof 1: In an attributed social network $G = (V, E, M)$, where V, E, and M represent the node set, edge set, and attribute set, respectively, we consider a fairness equiconcept (B, B) in G, where $B \in V$. The number of cluster-oriented generators, denoted as $|\mathrm{Gen}(B, B)|$, is $2^{|B|} - 1$, because only the

TABLE I
DEGREES OF CONCEPTS

Concept	Clique	$\phi(C)$	$\varphi(C)$	$\ell(C)$	Result
10	13,14	0.75	0.003906	0.000976563	F
11	10,11,12,14	0.9375	0.003906	0.000244141	T
20	8,9,10	0.875	0.0625	0.0078125	F
21	5,7,9,13	0.9375	0.003906	0.000244141	T
23	7,8,9	0.875	0.0625	0.0078125	F
24	6,7,8	0.875	0.000244	3.05176E-05	T
26	5,6,7	0.875	0.0625	0.0078125	F
27	1,2,5	0.875	0.0625	0.0078125	F
28	1,3,4,5	0.9375	1.0	0.0625	F

empty set does not satisfy Definition 4. The fairness equiconcept requires that the number of nodes $|B|$, must be greater than or equal to the number of attributes $|M|$. Furthermore, the $idx(B,B)$ of the fairness equiconcept (B,B) must be either 0 or $|M|-1$. So $idx(B,B)<|B|$, therefore

$$\begin{aligned}\ell(B,B) &= (1-\sigma(B,B)) * \varphi(B,B)\\ &= \left(1-\frac{|\mathrm{Gen}(B,B)|}{2^{|B|}}\right) * \frac{2^{idx(B,B)}}{2^{\max(I)}}\\ &= \frac{1}{2^{|B|}} * \frac{2^{idx(B,B)}}{2^{\max(I)}}\\ &\leq \frac{1}{2^{|B|}} * \frac{2^{|B|}}{2^{\max(I)}} = \frac{1}{2^{\max(I)}}.\end{aligned}$$

Note: "I" carries the same meaning as defined in Definition 9. □

Example 5: Building upon Example 4, we can utilize Definitions 5 and 9 to compute the stability and fairness of each formal concept. We then calculate the fairness-aware distances for these concepts using Definition 10. Subsequently, we screened out all fairness equiconcepts based on our method and thus discovered all fairness-aware maximal cliques. Due to space constraints, only the corresponding calculation results for equiconcepts among all concepts are provided in Table I [note: the fairness index set I is the fairness index of all concepts, and $\max(I)=12$].

D. Algorithm

Based on the above discussions, our detection algorithm for fairness-aware maximal cliques is presented in Algorithm 3.

The working procedure of Algorithm 3 is as follows. Initially, the attributed network is taken as an input, and based on the method in Section IV-B, the formal context and concept lattice L are constructed to extract all concepts (line 3). Subsequently, the stability and fairness sets of all concepts are calculated separately (lines 4 and 5). Next, by incorporating the stability, fairness, and fairness-aware distances for all concepts are calculated (line 8). The resulting fairness-aware distance is compared with the theoretical threshold, and the final fairness maximal clique is selected (lines 9 and 10).

1) Time Complexity Analysis: We provide a detailed analysis of the time complexity. First, in constructing the concept lattice and detecting the maximal cliques, the time complexity is $O(|V|^3)$[36], where $|V|$ is the number of nodes in network G.

Algorithm 3: Fairness-aware maximal cliques Detection Algorithm

Input:
The attributed social network G=(V,E,M)
Output:
Fairness maximal clique set S for the network G
1: Initialize S=$\emptyset$
2: **begin**
3: Generate lattice L via FCA in Section IV-B
4: $\mathbb{K}$ = $CalStability(L)$
5: $\mathbb{F}, max$ = $CalFairness(L,G)$
6: **for** i **in** $len(L)$
7: **begin**
8: ℓ_i = $(1-K_i)*F_i$
9: **if**$(\ell_i \leq \frac{1}{2^{max}})$
10: S=$S\cup L_i$
11: **end**
12: **return** S
13: **end**

We use γ to denote the number of maximal cliques and k to denote the number of nodes in maximal clique. The complexity of calculating the stability of the maximal clique depends on the number of its generators, i.e., $O(2^k)$. Furthermore, the complexity of calculating fairness for the maximal clique is $O(k^2)$. Therefore, with these analyses, the complexity of our algorithm is $O(|V|^3+\gamma*2^k+\gamma*k^2)$. Considering that k is usually a small number in real-world datasets (the larger the maximal clique in the network, the lower the probability of its existence), we consider $O(2^k)$ complexity to be acceptable. Therefore, the complexity of our algorithm is $O(|V|^3)$.

V. EXPERIMENTS

In this section, five real network data are used to conduct relevant experiments and evaluations based on the method proposed in this article.

A. Datasets

In this section, we evaluate our proposed method using five real datasets. The statistics of these datasets are presented in Table II. *Dataset I* (Forum)[1] is a Facebook-like forum network which was attained from the same online community as the OSN. *Dataset II* (OClinks)[2] originate from an online community for students at the University of California, Irvine.

The dataset represents interpersonal relationships, where each edge signifies an interaction between two individuals. *Dataset III* (Grid)[3] is a biological networks about biogrid-worm's links. *Dataset IV* (Dmela)[4] represents a network of protein–protein interactions. *Dataset V* (CaHelp)[5] is a Collaboration network of arXiv high energy physics.

[1] https://networkrepository.com/fb-forum.php
[2] http://opsahl.co.uk/tnet/datasets/OClinks_w.dl
[3] https://networkrepository.com/bio-grid-worm.php
[4] https://networkrepository.com/bio-dmela.php
[5] https://networkrepository.com/ca-HepPh.php

TABLE II
INFORMATION OF DATASETS

	Dataset	Nodes	Edges	Description
I	Forum	0.9k	7.1k	Dynamic Networks
II	OClinks	1.9k	20.3k	Social Network
III	Grid	3.5k	13.1k	Biological Networks
IV	Dmela	7.4k	25.6k	Biological Networks
V	CaHelp	12k	118.6k	Collaboration Networks

Before analysis, the datasets undergo preprocessing. For each node, the weights of its connected edges are aggregated to derive the node's value. Subsequently, the node is assigned an attribute label based on this value. In dataset forum, According to the user's activity level, users are divided into three categories (active, general, and inactive), thereby assigning attribute values to each node. As attribute information is unavailable in other datasets, we assign random attribute values to the nodes. Following the completion of attributed network processing, dataset forum contains three attribute values, while other datasets comprise four attribute values.

All experiments are performed on a PC with Inter Core i-11357G7, 2.40 GHz CPUs, and 16 GB RAM. The source code of this article is released at Github: https://github.com/Sinocifeng/fairCliques.

B. Comparison Algorithm

In this section, we briefly introduce several comparison algorithms used in this article. The fairness-aware maximal clique we define corresponds to the relative fairness-aware maximal clique in [34], so we choose the RFCRefine and RFCAlter algorithms for comparison. In addition, the results mined by the AFCMiner method are very similar to ours, and we also chose this method for comparison.

1) *RFCRefine [34]:* The fundamental idea behind the RFCRefine algorithm is to initially identify weak fair maximal cliques with an adequate number of nodes across the entire network, followed by the enumeration of all relative fair maximal cliques within these weak fair maximal cliques.
2) *RFCAlter [34]:* The fundamental concept of the RFCAlter algorithm is to circumvent the repetitive enumeration of weakly fair cliques employed by RFCRefine and instead utilize the attribute selection search method to directly identify all relative fair cliques within the network.
3) *AFCMiner+ [36]:* The AFCMiner algorithm is based on FCA and relies on a newly defined concept—the attributed equiconcept to directly obtain fair maximal cliques. However, the fair maximal cliques mined by the AFCMiner algorithm require the same number of nodes corresponding to each attribute, which is more stringent than our restriction. Therefore, we extended the node based on this method to relax the requirements and meet our results, so we labeled it as AFCMiner+.

TABLE III
COMPARISON OF F1 SCORES OF FAIRNESS-AWARE MAXIMAL CLIQUES DETECTION ALGORITHMS ON THE TESTED ATTRIBUTED NETWORKS

Dataset	RFCRefine	RFCAlter	AFCMiner+	Our Approach
Forum	0.974	0.923	0.900	**1.0**
OClinks	0.772	0.759	0.730	**0.785**
Grid	0.629	0.671	**0.711**	0.596
Dmela	0.684	0.718	0.649	**0.743**
CaHelp	0.910	0.917	0.913	**0.927**

Note: The bold numbers represent the maximum F1 scores for each dataset.

4) *Our Approach:* Our approach is based on FCA, and by characterizing the stability and fairness of the formal concepts in the concept lattice, we can discover the fairness-aware maximal cliques directly.

C. Experimental Results

In this section, we conduct a series of experimental evaluations on the proposed method. In experiment 1, we validate the computational results of the proposed method regarding fairness-aware maximal cliques about five datasets. For experiment 2, we create four subgraphs for datasets II and V by randomly selecting 20%–80% of edges and assessing the execution times of all proposed algorithms.

1) Performance Evaluation on Fairness-Aware Maximal Cliques Mining: In this experiment, we utilized F1-measure to assess the effectiveness of identifying fairness-aware maximal cliques in the networks using our approach. The F1-measure is defined as follows:

$$F1 = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

in which precision represents the ratio of correct detections to the total number of detection results, and recall denotes the ratio of correct detections to the number of ground-truth structures. According to this formula, it is evident that a higher F1 score corresponds to a more accurate prediction of the fairness-aware maximal cliques.

Table III shows the experimental results of our fairness-aware maximal cliques on five datasets. Obviously, apart from the Grid dataset, our method achieves the highest scores across the remaining four datasets; in the Grid, the F1-score of our method is relatively low, mainly due to the small number of fairness-aware maximal cliques in this dataset. In fact, in other datasets, our approach shows better performance than the other three methods. Notably, in the Forum dataset, we successfully identified all fairness-aware maximal cliques. The underperformance of AFCMiner+ may be because of the algorithm's strict requirement for an equal number of nodes corresponding to each attribute within the fair maximal clique. This constraint leads the AFCMiner algorithm to ignore some potential during the retrieval of attributed equiconcepts.

In addition, the performance of RFCRefine and RFCAlter may be due to the loss of some possible results during the weak

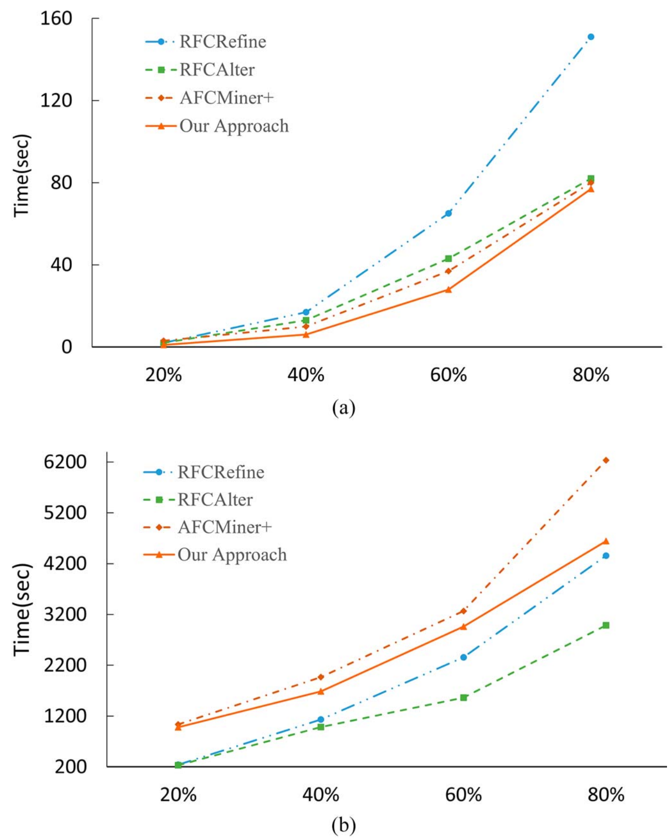

Fig. 6. Scalability of RFCRefine, RFCAlter, AFCMiner+, and Our approach.

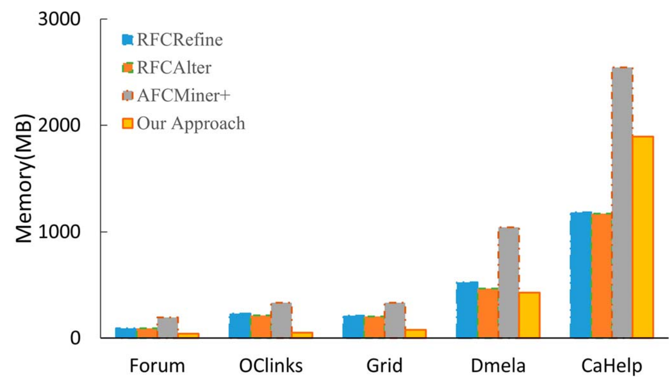

Fig. 7. Memory cost for RFCRefine, RFCAlter, AFCMiner+, and Our approach.

fairness clique search and attribute selection search. Furthermore, we also explored fair maximum clique mining based on classic maximum clique mining algorithms (such as the BK algorithm [44], PC algorithm [45], and greedy algorithm). Experimental results show that although the BK and PC algorithms can find all fairness-aware maximal cliques in the network in theory, stack overflow occurs on other datasets except Forum and OClinks. Although the greedy algorithm has a shorter execution time, the number of fairness-aware maximal cliques it mines is far less than the four methods introduced in this article, so the experimental results based on these three classic methods are not presented in our article.

2) Scalability Test: Due to the relatively low number of fairness-aware maximal cliques in the Forum, Grid, and Dmela datasets, to assess the scalability and operational efficiency of the proposed algorithm, we generated four subgraphs for the OClinks and CaHelp datasets by randomly selecting 20%–80% of edges and evaluated the runtime of these algorithms.

Fig. 6(a) and 6(b) shows the runtime on the OClinks and CaHelp. Obviously, in the OClinks, our approach performs better than the other algorithms; while in the CaHelp, the runtime of our approach shows a steady increase as the graph size increases. It should be noted that in the CaHelp, the runtime of our method is higher than RFCRefine and RFCAlter, which is reasonable. As the number of edges increases, more connections and relationships are generated in the graph, which expands the size and complexity of the graph, so we need to spend more time building the concept lattice. The RFCRefine and RFCAlter algorithms perform better than our method in runtime. The main reason is that the pruning process in these two algorithms can significantly reduce the network size and decrease the computational complexity.

Despite longer runtimes in large-scale networks, our method excels in discovering more fairness-aware maximal cliques. AFCMiner+ performs poorly, mainly because its stringent fairness definition requires additional node scaling to meet our fairness criteria, which takes extra time.

3) Memory Test: We also examined the memory usage of the four methods during their operation.

As depicted in Fig. 7, we compared the memory consumption of these methods across five test datasets. Our method demonstrates lower memory usage when handling smaller datasets. However, on the CaHelp dataset, both RFCRefine and RFCAlter methods show reduced memory usage due to their effective pruning capabilities with larger network sizes, thereby significantly decreasing memory requirements. In contrast, the AFCMiner+ method exhibits higher memory consumption, as expected, because it needs to record extensive node expansion information after identifying the strictly fairness-aware maximal cliques.

4) Absolute Fairness-Aware Maximal Clique Number: This article stipulates that the difference in the number of nodes corresponding to any two attributes within the identified fairness-aware maximal clique should not exceed 1. In particular, when the number of nodes corresponding to all attributes is equal, we refer to the fairness-aware maximal clique as an absolute fairness-aware maximal clique. We also investigate the ability of these four methods to identify such absolute fairness-aware maximal cliques.

According to Table IV, we compared the number of absolute fairness-aware maximal cliques identified for five datasets. "Num" denotes the count of absolute fairness-aware maximal cliques in each network. The results show that our approach mined more absolute fairness-aware maximal cliques across all five datasets compared to RFCRefine and RFCAlter. This indicates our substantial advantage in absolute

TABLE IV
COMPARISON OF NUMBER OF ABSOLUTE FAIRNESS-AWARE MAXIMAL CLIQUES DETECTION ALGORITHMS

Dataset	Num	RFCRefine	RFCAlter	AFCMiner+	Our Approach
Forum	20	19	18	18	**20**
OClinks	244	204	203	**227**	213
Grid	24	15	15	**20**	16
Dmela	13	12	12	**13**	**13**
CaHelp	301	267	251	**289**	277

Note: The bold numbers represent the maximum number of absolute fairness-aware maximal cliques for each dataset.

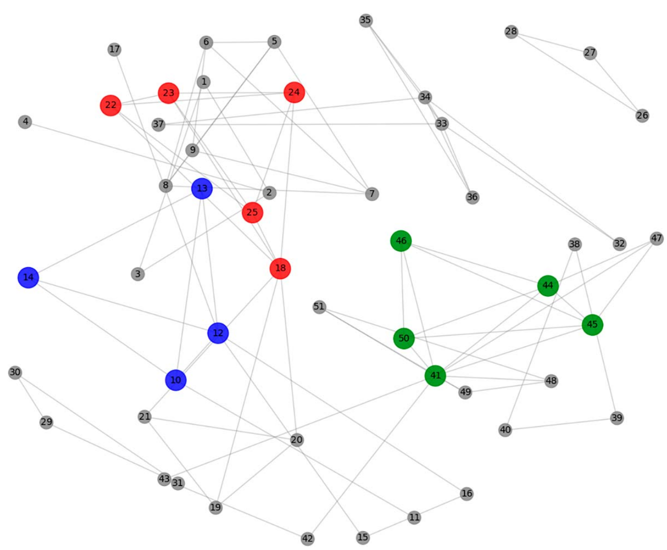

Fig. 8. Visualization of partial nodes in the networkScience.

fairness-aware maximal clique mining. However, when benchmarked against AFCMiner+, our method exhibits slightly inferior performance on datasets except Forum. This variance primarily stems from AFCMiner's deliberate focus on mining absolute fairness-aware maximal cliques, leveraging attributed equiconcept for efficient extraction of absolute fairness-aware maximal cliques.

D. Case Study

This section aims to illustrate the potential application of the proposed method through a case study conducted on an attributed network dataset.

We have selected "networkScience"[6]—a network of coauthorships in the area of network science. The network consists of 1461 nodes and 2742 edges. With the help of this network topology, we randomly assign four attribute values to nodes in the network as labeled data to distinguish the research profiles of different scholars.

By applying our method, we identified 21 fairness-aware maximal cliques in this network. Due to the large scale of the network, Fig. 8 only shows the visualization results of the detected fairness-aware maximal cliques in a subnetwork composed of partial nodes from the network. The red, blue, and green nodes in Fig. 8 represent nodes in three distinct fairness-aware maximal cliques. This implies that we can select such groups to review the project, which can effectively ensure the diversity and fairness of the professional fields within the review team. Therefore, during the project review process, we can obtain diversified review opinions from different backgrounds and professional fields. Such a team can examine all aspects of the project more comprehensively and improve the objectivity and quality of the review.

Let us consider the fairness-aware maximal clique 41, 44, 45, 46, 50 composed of green nodes, representing experts from computer science, environmental studies, economics, law, and sociology, respectively. Such a combination of review experts offers several advantages.

1) *Fairness*: The review process covers multiple disciplines and professional fields, preventing any single field from dominating the review results and ensuring fairness.
2) *Objectivity*: These experts have established professional trust and interactions, facilitating a fair assessment of the project's quality and innovation.
3) *Comprehensiveness*: The expert team spans various fields, enabling a thorough examination of the project's strengths and weaknesses from diverse angles (e.g., technical feasibility, economic sustainability, environmental impact, legal compliance, and social implications).

This expert team enhances the review process by providing a comprehensive evaluation, thereby bolstering the objectivity, fairness, and overall quality of the review.

VI. CONCLUSION

In this article, we propose a method based on the FCA framework to identify fairness-aware maximal cliques while considering fairness in attributed social networks. Initially, we utilize the network representation of FCA to extract concepts in the network structure. Subsequently, we introduced cluster-oriented generators to enhance stability. Additionally, we introduce evaluation metrics to evaluate the fairness of concepts. By synergizing the interestingness of these two concepts, we effectively uncover fairness-aware maximal cliques in attributed social networks. We verify the effectiveness and practicality of our method through experiments on five real attributed networks. Future research avenues involve exploring the integration of graph pruning techniques with multidimensional attribute data in attribute networks, as well as delving into more efficient methods for identifying fairness-aware maximal cliques.

[6]http://konect.cc/networks/dimacs10-netscience/

REFERENCES

[1] O. Severiukhina, K. Bochenina, S. Kesarev, and A. Boukhanovsky, "Parallel data-driven modeling of information spread in social networks," in *Proc. 18th Int. Conf. Comput. Sci. (ICCS)*, Wuxi, China. New York, NY, USA: Springer-Verlag, 2018, pp. 247–259.
[2] C. Chen, Z. Jiang, and J. Ma, "Privacy protection for marginal-sensitive community individuals against adversarial community detection attacks," *IEEE Trans. Comput. Social Syst.*, vol. 11, no. 1, pp. 782–794, Feb. 2024.
[3] Y. Yang, S. Li, and P. Zhang, "Data-driven accident consequence assessment on urban gas pipeline network based on machine learning," *Rel. Eng. Syst. Saf.*, vol. 219, 2022, Art. no. 108216.

[4] B. Viswanath, A. Mislove, M. Cha, and K. P. Gummadi, "On the evolution of user interaction in Facebook," in *Proc. 2nd ACM Workshop Online Social Netw.*, 2009, pp. 37–42.

[5] F.-B. Shi, X.-Q. Sun, H.-W. Shen, and X.-Q. Cheng, "Detect colluded stock manipulation via clique in trading network," *Physica A Statist. Mech. Appl.*, vol. 513, pp. 565–571, Jan. 2019.

[6] L. D. Valdez, L. Vassallo, and L. A. Braunstein, "Epidemic control in networks with cliques," *Phys. Rev. E*, vol. 107, no. 5, 2023, Art. no. 054304.

[7] S. Roscoe, M. Khatri, A. Voshall, S. Batra, S. Kaur, and J. Deogun, "Formal concept analysis applications in bioinformatics," *ACM Comput. Surv.*, vol. 55, no. 8, pp. 1–40, 2022.

[8] H. Yu, R. Ma, J. Chao, and F. Zhang, "An overlapping community detection approach based on deepwalk and improved label propagation," *IEEE Trans. Comput. Social Syst.*, vol. 10, no. 1, pp. 311–321, Feb. 2023.

[9] R. Missaoui, A. Messaoudi, M. H. Ibrahim, and T. Abdessalem, "Detecting communities in complex networks using formal concept analysis," in *Advances in Knowledge Discovery and Management*, vol. 9, New York, NY, USA: Springer-Verlag, 2022, pp. 77–105.

[10] X. Li, R. Zhou, L. Chen, Y. Zhang, C. Liu, Q. He, and Y. Yang, "Finding a summary for all maximal cliques," in *Proc. IEEE 37th Int. Conf. Data Eng. (ICDE)*, Piscataway, NJ, USA: IEEE Press, 2021, pp. 1344–1355.

[11] A. Alsini, A. Datta, and D. Q. Huynh, "On utilizing communities detected from social networks in hashtag recommendation," *IEEE Trans. Comput. Social Syst.*, vol. 7, no. 4, pp. 971–982, Aug. 2020.

[12] F. Hao, G. Min, Z. Pei, D.-S. Park, and L. T. Yang, "k-clique community detection in social networks based on formal concept analysis," *IEEE Syst. J.*, vol. 11, no. 1, pp. 250–259, Mar. 2017.

[13] F. Hao, D.-S. Park, G. Min, Y.-S. Jeong, and J.-H. Park, "k-cliques mining in dynamic social networks based on triadic formal concept analysis," *Neurocomputing*, vol. 209, pp. 57–66, Oct. 2016.

[14] A. Angerschmid, J. Zhou, K. Theuermann, F. Chen, and A. Holzinger, "Fairness and explanation in AI-informed decision making," *Mach. Learn. Knowl. Extraction*, vol. 4, no. 2, pp. 556–579, 2022.

[15] V. F. Grabinski and O. W. Brawley, "Disparities in breast cancer," *Obstetrics Gynecol. Clinics*, vol. 49, no. 1, pp. 149–165, 2022.

[16] N. A. Saxena, K. Huang, E. Defilippis, G. Radanovic, and Y. Liu, "How do fairness definitions fare?: Examining public attitudes towards algorithmic definitions of fairness," in *Proc. AAAI/ACM Conf.*, 2019, pp. 99–106.

[17] A. Rahmattalabi et al., "Exploring algorithmic fairness in robust graph covering problems," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–12.

[18] J. Zou and L. Schiebinger, "AI can be sexist and racist—It's time to make it fair," *Nature*, vol. 559, pp. 324–326, Jul. 2018.

[19] A. Mouakher and A. Ko, "Efficient assessment of formal concept stability in the Galois lattice," *Int. J. General Syst.*, vol. 51, no. 8, pp. 791–821, 2022.

[20] S. Kuznetsov, "Stability as an estimate of the degree of substantiation of hypotheses derived on the basis of operational, similarity," *Nauchno Tekhnicheskaya Seriya*, vol. 2, no. 12, pp. 21–29, 1990.

[21] F. Hao, J. Gao, Y. Lin, Y. Wu, and J. Shang, "Concept stability entropy: A novel group cohesion measure in social networks," *IEEE Trans. Emerg. Topics Comput.*, early access, Sep. 20, 2023.

[22] J. Li, C. Huang, J. Qi, Y. Qian, and W. Liu, "Three-way cognitive concept learning via multi-granularity," *Inf. Sci.*, vol. 378, pp. 244–263, 2017.

[23] J. Qi, T. Qian, and L. Wei, "The connections between three-way and classical concept lattices," *Knowl.-Based Syst.*, vol. 91, pp. 143–151, Jan. 2016.

[24] F. Hao, Y. Yang, G. Min, and V. Loia, "Incremental construction of three-way concept lattice for knowledge discovery in social networks," *Inf. Sci.*, vol. 578, pp. 257–280, Nov. 2021.

[25] R. K. Chunduri and A. K. Cherukuri, "Distributed three-way formal concept analysis for large formal contexts," *J. Parallel Distributed Comput.*, vol. 171, pp. 141–156, Jan. 2023.

[26] M. Klimushkin, S. Obiedkov, and C. Roth, "Approaches to the selection of relevant concepts in the case of noisy data," in *Proc. 8th Int. Conf. Formal Concept Anal.* (ICFCA), Agadir, Morocco. New York, NY, USA: Springer-Verlag, 2010, pp. 255–266.

[27] M. A. Babin and S. O. Kuznetsov, "Approximating concept stability," in *Proc. 10th Int. Conf. Formal Concept Anal.* (ICFCA), Leuven, Belgium. New York, NY, USA: Springer-Verlag, 2012, pp. 7–15.

[28] A. Mouakher, and S. B. Yahia, "On the efficient stability computation for the selection of interesting formal concepts," in *Information Sciences*, vol. 472, pp. 15–34, 2019.

[29] J. Gao, F. Hao, Z. Pei, and G. Min, "Learning concept interestingness for identifying key structures from social networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 4, pp. 3220–3232, Oct.–Dec. 2021.

[30] Y. Li, C. Sha, X. Huang, and Y. Zhang, "Community detection in attributed graphs: An embedding approach," in *Proc. AAAI Conf. Artif. Intell.*, 2018, vol. 32, no. 1, pp. 338–345.

[31] Y. Fang, R. Cheng, S. Luo, and J. Hu, "Effective community search for large attributed graphs," *Proc. VLDB Endowment*, vol. 9, no. 12, pp. 1233–1244, 2016.

[32] X. Xie, M. Song, C. Liu, J. Zhang, and J. Li, "Effective influential community search on attributed graph," *Neurocomputing*, vol. 444, pp. 111–125, Jul. 2021.

[33] C. Pizzuti and A. Socievole, "A genetic algorithm for community detection in attributed graphs," in *Proc. 21st Int. Conf. Appl. Evol. Comput.* (EvoApplications), Parma, Italy. New York, NY, USA: Springer-Verlag, 2018, pp. 159–170.

[34] Q. Zhang, R.-H. Li, M. Pan, Y. Dai, Q. Tian, and G. Wang, "Fairness-aware maximal clique in large graphs: Concepts and algorithms," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 11, pp. 11368–11387, Nov. 2023.

[35] M. Pan, R.-H. Li, Q. Zhang, Y. Dai, Q. Tian, and G. Wang, "Fairness-aware maximal clique enumeration," in *Proc. IEEE 38th Int. Conf. Data Eng. (ICDE)*, Piscataway, NJ, USA: IEEE Press, 2022, pp. 259–271.

[36] F. Hao, Y. Yang, J. Shang, and D.-S. Park, "AFCMiner: Finding absolute fair cliques from attributed social networks for responsible computational social systems," *IEEE Trans. Comput. Social Syst.*, vol. 10, no. 6, pp. 3000–3011, Dec. 2023.

[37] X. Li, R. Zhou, L. Chen, C. Liu, Q. He, and Y. Yang, "One set to cover all maximal cliques approximately," in *Proc. Int. Conf. Manage. Data*, 2022, pp. 2006–2019.

[38] T. Yu and M. Liu, "A memory efficient maximal clique enumeration method for sparse graphs with a parallel implementation," *Parallel Comput.*, vol. 87, pp. 46–59, Sep. 2019.

[39] P. Vilakone, M.-P. Hong, and D.-S. Park, "Divide the FCA network graph into the various community based on the k-clique methods," in *Proc. Adv. Comput. Sci. Ubiquitous Comput. (CSA-CUTE),* New York, NY, USA: Springer-Verlag, 2021, pp. 113–119.

[40] F. Hao, D.-S. Park, and Z. Pei, "When social computing meets soft computing: Opportunities and insights," *Human-Centric Comput. Inf. Sci.*, vol. 8, pp. 1–18, 2018.

[41] Y. Yang, F. Hao, B. Pang, G. Min, and Y. Wu, "Dynamic maximal cliques detection and evolution management in social internet of things: A formal concept analysis approach," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 3, pp. 1020–1032, May/Jun. 2022.

[42] Y. Dong, J. Ma, S. Wang, C. Chen, and J. Li, "Fairness in graph mining: A survey," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 10, pp. 10583–10602, Oct. 2023.

[43] M. Liquière, "Some links between formal concept analysis and graph mining," *Mining Graph Data*, vol. 12, pp. 227–252, 2006.

[44] C. Bron and J. Kerbosch, "Algorithm 457: Finding all cliques of an undirected graph," *Commun. ACM*, vol. 16, no. 9, pp. 575–577, 1973.

[45] Y. Xu, J. Cheng, A. W.-C. Fu, and Y. Bu, "Distributed maximal clique computation," in *Proc. IEEE Int. Congr. Big Data*, Piscataway, NJ, USA: IEEE Press, 2014, pp. 160–167.

Min Tao received the B.S. degree in software engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2020. Currently, he is working toward the master's degree in software engineering with Shaanxi Normal University, Xi'an, China.

His research interests include social computing, graph data mining, and formal concept analysis.

Fei Hao received the Ph.D. degree in computer science and engineering from Soonchunhyang University, Asan, South Korea, in 2016.

Since 2016, he has been with Shaanxi Normal University, Xi'an, China, where he is currently an Associate Professor. From 2020 to 2022, he was the Marie Sklodowska-Curie Fellow with the University of Exeter, Exeter, U.K. His research interests include social computing, ubiquitous computing, big data analytics, knowledge graph, and edge intelligence.

Ling Wei was born in 1972. She received the Ph.D. degree in mathematics from Xi'an Jiaotong University, Xi'an, China, in 2005.

Currently, she is a Professor with Northwest University, Xi'an, China. Her research interests include formal concept analysis, rough set theory, three-way decisions, and granular computing.

Huilai Zhi received the M.Sc. degree from Xihua University, Chengdu, China, in 2007, and the Ph.D. degree from Shanghai University, Shanghai, China, in 2010, both in science.

Currently, he is a Professor with Quanzhou Normal University, Quanzhou, China. His research interests include knowledge space theory, granular computing, and formal concept analysis.

Sergei O. Kuznetsov received the Ph.D. degree in theoretical computer science from Moscow Institute for Physics and Technology, Moscow, USSR, in 1990, and the Doctor of Science degree in machine learning models based on concept lattices from the Computing Center of Russian Academy of Science, Moscow, Russia, in 2002.

Currently, he is a Full Professor with HSE University, Moscow, Russia, being the Head of the School for Data Analysis and Artificial Intelligence, International Laboratory for Intelligent Systems, and an Academic Supervisor of the master's program in data science.

Geyong Min (Member, IEEE) received the B.Sc. degree in computer science from Huazhong University of Science and Technology, Wuhan, China, in 1995, and the Ph.D. degree in computing science from the University of Glasgow, Glasgow, U.K., in 2003.

Currently, he is a Professor in high performance computing and networking with the Department of Computer Science, College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, U.K. His research interests include computer networks, wireless communications, parallel and distributed computing, ubiquitous computing, multimedia systems, modeling, and performance engineering.