서명된 속성 소셜 네트워크에서의 Absolute-Fair Maximal Balanced Cliques 탐색

양예선[1], 펭소니[1], 박두순[1,*], 이혜정[2]
[1]순천향대학교 소프트웨어융합학과
[2]순천향대학교 AI • SW 교육원
yangyxuan621@gmail.com, peng.sony61@gmail.com, parkds@sch.ac.kr, hjlee100@sch.ac.kr

Absolute-Fair Maximal Balanced Cliques Detection in Signed Attributed Social Network

Yixuan Yang[1], Sony Peng[1], Doo-Soon Park[1,*], HyeJung Lee[2]
[1]Dept. of Software Convergence, Soonchunhyang University
[2] Institute for Artificial Intelligence and Software, Soonchunhyang University

Abstract

Community detection is a hot topic in social network analysis, and many existing studies use graph theory analysis methods to detect communities. This paper focuses on detecting absolute fair maximal balanced cliques in signed attributed social networks, which can satisfy ensuring the fairness of complex networks and break the bottleneck of the ″information cocoon ″.

1. Introduction

There are complex interpersonal relationships between people in the real world, which can form a complex network. A social network is usually represented as a graph structure composed of multiple nodes (that is, modeled as a graph). Each node can be regarded as an individual member or organization. A social network graph can effectively illustrate the social relationship between nodes.

In the study of complex networks, community detection, which aims to discover the cohesive subgraph structure in the network, is a fundamental problem in graph analysis and has received much attention in recent years. As a basic cohesive subgraph model, among them, clique is widely used to reveal the dense community structure of graphs.[1]

Mining cliques is widely used in graph analysis, and the existing related works include protein structure detection, community enumeration, etc [2]. It can also solve some key problems in social networks, such as mining overlapping communities in social networks, Das et al. [3] utilized a parallel graph algorithm to enumerate maximal cliques of big-scale biological networks; Cheng et al. [4] first proposed an efficient partition-based approach to deal with the problem of processing large graphs with limited memory for maximal cliques enumeration algorithms.

In fact, in a social network graph, the topological structure between nodes can be divided into two types weighted edges and unweighted edges. Among them, a social network containing edges with weights(positive/negative) is called signed social Network, there have been many works devoted to the study of signed social networks and to analyzing the balance of the network/subnetwork. It is worth noting that in addition to the topology information between nodes, the nodes themselves also contain a lot of attribute information, this kind of network is called attribute networks. At present, there are also a lot of research works on the mining community based on attribute networks [5][6]. However, these works either require a high correlation of attributes in communities or aim to find communities that satisfy certain attribute constraints. None of them consider the fairness of attributes in the community. Recently, the concept of "fairness" has attracted a lot of attention in machine learning to achieve trustworthy AI [7][8]. Inspired by this, several new studies have emerged by combining fairness with clique detection in attribute graphs [9].

*corresponding author: Doo-Soon Park
Acknowledgements: This research was supported by the National Research Foundation of Korea (No. NRF-2022R1A2C1005921) and BK21 FOUR (Fostering Outstanding Universities for Research) (No.5199990914048).

In our work, we intend to consider “fairness” and “balance” together in the social networks. We devote ourselves to mining the maximal balanced cliques and the fairness in each clique is absolute fair.

Detecting balanced cliques can break through the existing "information cocoon" and other problems, meanwhile, considering fairness can constrain some attributes of the cliques so that only the very large factions formed by certain nodes we are interested in are obtained. An online social network where each user has an attribute indicating his/her gender. For instance, if we wish to find a community with the same number of men and women, through the detection of attribute communities, gender bias can be overcome [10]. Or we want to find some communities of scientists from different disciplines in a collaborative network, each node has an attribute representing his/her research topic, and the fair group at this time can consider the fairness of different research topics. Combined with the maximal balanced cliques of the signed social network and absolute fair cliques, we can break through the information cocoon of the fair communities and obtain more meaningful data.

2. Problem Statement

Before presenting our problem addressed in this paper, in order to improve the understanding, several fundamental definitions in signed attributed social networks and absolute fair maximal balanced cliques will be provided.

Definition 1 (Signed Attributed Social Network) A Signed Attributed Social Network can be represented as G=(V, E, W, A), with V indicating the set of nodes, $V = \{v_1, v_2, v_3, \ldots, v_n\}$, E indicating the set of edges $E = \{e_{ij}|i,j \in V\}$, W indicating the set of edges’ weight (we use “+” to represent positive weight, “-” to represent negative weight), $W = \{W^+ \cup W^-\}$, A indicating the set of attributes on nodes, $A = \{v.val|v \in V\}$.

Definition 2 (Absolute Fair Maximal Clique)[9] Given an attributed social network G=(V, E, A), a clique C of G is an absolute fair maximal clique of G if **(1)** (Fairness) the number of nodes for each $a_i \in A$ is the same; **(2)** (Maximal) there is no clique $C' \supset C$ satisfying **(1)**.

Definition 3 (Maximal Balanced Clique) Given a signed network G=(V, E, W), a maximal balanced clique MBC is a maximal subgraph of G, MBC is a complete subgraph which $\forall(u,v) \in MBC \rightarrow (u,v) \in W^+ \cup W^-$, and MBC can be divided into 2 subcliques C_1 and C_2, **(3)** (Balanced) $\forall u,v \in C_1$ or $u,v \in C_2 \rightarrow (u,v) \in E^+$, and $\forall u \in C_1, v \in C_2$ or $u \in C_2, v \in C_1 \rightarrow (u,v) \in E^-$.

Definition 4 (Absolute Fair Maximal Balanced Clique) Given a signed Attributed Social Network G=(V, E, W, A), a pair of cliques are C_1 and C_2, C_1 and C_2 satisfying **Definition 2** and **Definition 3**, the $C = C_1 \cup C_2$ is an Absolute Fair Maximal Balanced Clique.

Problem Definition: Given a signed attributed social network G, our research aims to enumerate all absolute fair maximal balanced cliques from G. It is formulated as follows,

$$\Omega = AFMBC(G) \tag{1}$$

Ω is the cliques set which contains absolute fair maximal balanced cliques.

3. Proposed Approach

In this section, we will propose the approach of detecting absolute fairness maximal balanced cliques in the Signed Attribute Social Network.

In our previous works, we have proved the equivalence of the maximal cliques with the equi-concepts in the concept lattice by Formal Concept Analysis (FCA), we used the subgraphs with positive and negative inputs respectively as the formal context to get the maximal balanced cliques.

In our paper, we apply it for absolute fair cliques, so we also consider the attribute of nodes in the maximal balanced cliques, and make sure the number of nodes with attribute in each subclique of maximal balanced cliques.

The process of the approach is described as a pseudo-code as follows.

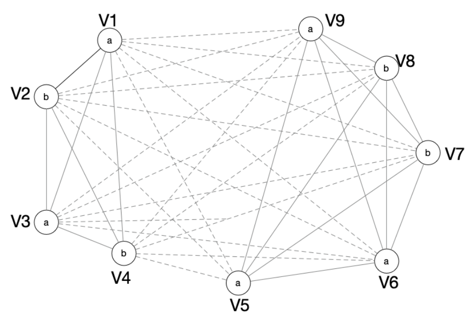
Algorithm 1: The algorithm of detecting AFMBC

```
Input G=(V, E, W, A) as the formal context
Cocept Lattice CL1, CL2 ← FCA(G(W+)), FCA(G(W-))
Equi-Concepts EC1, EC2 ← concepts in CL1, CL2 &&
                    concept.intent == concept.extent
for ec1, ec2 in EC1, EC2:
    Filter the ec1, ec2 that satisfy definition 2
    If ec1, ec2 :
        Ω ← pair (ec1, ec2)
    Else:
        prune equi-concepts that do not satisfy the
        definition 2 according to prune redundant
        nodes, obtain pairs (ec1, ec2).
        Ω ← pair (ec1, ec2)
Return Ω
```

To understand the process more clearly, we will give an example to illustrate.

Example 1:

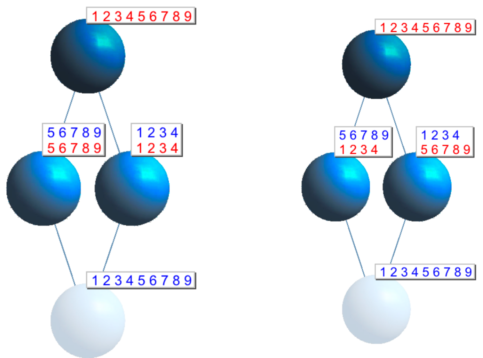
According to this graph shown as Figure 1, we present the formal context as an adjacency matrix in Table 1.



(Figure 1) A Signed Attribute Social Network g.

<Table 1> Formal Context for Signed Attribute Social Network g.

1	1	1	1	-1	-1	-1	-1	-1
1	1	1	1	-1	-1	-1	-1	-1
1	1	1	1	-1	-1	-1	-1	-1
1	1	1	1	-1	-1	-1	-1	-1
-1	-1	-1	-1	1	1	1	1	1
-1	-1	-1	-1	1	1	1	1	1
-1	-1	-1	-1	1	1	1	1	1
-1	-1	-1	-1	1	1	1	1	1
-1	-1	-1	-1	1	1	1	1	1



(Figure 2) The Concept Lattices of Positive and Negative Subgraphs.

Obviously, from the two concept lattices in Figure 2, we can get the maximal balance clique as: {(1,2,3,4), (5,6,7,8,9)}. But the clique (5,6,7,8,9) does not satisfy fairness, so we have to prune and derive new pseudo-equi-concepts: (5,6,7,8), (5,7,8,9) and (6,7,8,9). In the final step, we obtain 3 absolute fair maximal balanced cliques: {(1,2,3,4), (5,6,7,8}, {(1,2,3,4), (5,7,8,9)} and {(1,2,3,4), (6,7,8,9)}.

4. Conclusion

This paper mainly studies the problem of detecting absolute fair maximal balanced cliques from signed attributed social networks. We proposed an approach to detecting absolute fair maximal balanced cliques. First, to obtain all maximal balanced cliques, we enumerate them using the FCA method. Then, to guarantee the fairness of the detected maximal cliques, we filter and derive fairness-compliant subcliques. Through Example 1, it can be proved that our method is feasible.

Reference

[1] Yixuan Y., Fei H, Doo-Soon P., et al. Modelling Prevention and Control Strategies for COVID-19 Propagation with Patient Contact Networks[J]. Human-centric Computing and Information Sciences, 2021, 11.

[2] Yixuan Y., Fei H, BeiBei P., et al. Dynamic Maximal Cliques Detection and Evolution Management in Social Internet of Things: A Formal Concept Analysis Approach, in IEEE Transactions on Network Science and Engineering, doi: 10.1109/TNSE.2021.3067939.

[3] Das A, Sanei-Mehri S V, Tirthapura S. Shared-memory parallel maximal clique enumeration[C]//2018 IEEE 25th International Conference on High Performance Computing (HiPC). IEEE, 2018: 62-71.

[4] Cheng J, Zhu L, Ke Y, et al. Fast algorithms for maximal clique enumeration with limited memory[C]//Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. 2012: 1240-1248.

[5] Falih I., Grozavu N., Kanawati R., et al. Community detection in attributed network[C]//Companion Proceedings of the The Web Conference 2018. 2018: 1299-1306.

[6] Chunaev P. Community detection in node-attributed social networks: a survey[J]. Computer Science Review, 2020, 37: 100286.

[7] Mehrabi N., Morstatter F., Saxena N., et al. A survey on bias and fairness in machine learning[J]. ACM Computing Surveys (CSUR), 2021, 54(6): 1-35.

[8] Vincent-Lancrin, Stéphan, and Reyer Van der Vlies. "Trustworthy artificial intelligence (AI) in education: Promises and challenges." (2020).

[9] Fei H., WenQing H., Doo-Soon P., Absolute Fair Maximal Cliques Detection in Attributed Social Networks, The 16th International Conference on Multimedia and Ubiquitous Engineering.

[10] Minjia P., Rong-Hua L., Qi Z, et al. Fairness-aware Maximal. Clique Enumeration[J]. arXiv preprint arXiv:2107.10025, 2021.