

Detection of Maximal Balance Clique Using Three-way Concept Lattice

Yixuan Yang¹, Doo-Soon Park¹, Fei Hao^{2,3}, Sony Peng¹, Hyejung Lee⁴, and Min-Pyo Hong⁵

Abstract

In the era marked by information inundation, social network analysis is the most important part of big data analysis, with clique detection being a key technology in social network mining. Also, detecting maximal balance clique in signed networks with positive and negative relationships is essential. In this paper, we present two algorithms. The first one is an algorithm, MCDA1, that detects the maximal balance clique using the improved three-way concept lattice algorithm and object-induced three-way concept lattice (OE-concept). The second one is an improved formal concept analysis algorithm, MCDA2, that improves the efficiency of memory. Additionally, we tested the execution time of our proposed method with four real-world datasets.

Keywords

Formal Concept Analysis, Maximal Balanced Clique, Signed Networks, Three-Way Concept

1. Introduction

In response to the demands of the Fourth Industrial Revolution era, many researchers are directing their attention to social network structures analysis in computer science, with the cliques emerging as a very popular research direction in social network structure mining. Vilakone et al. [1,2] used clique detection and standardized cumulative gains to solve some problems in the movie recommendation system. However, detecting the maximal clique and balanced cliques in social networks, are not only applicable to social media software, but also can be applied to other domains, such as the Internet of Things (IoT), healthcare, biology, and more.

Meanwhile, maximal clique detection can be used to detect the structure of proteins, which can prevent or control diseases. For this reason, we can find a group of the most suitable medical specialists for a target patient to provide the best medical treatment through remote consultation.

In many existing works, cliques are typically detected in undirected and unweight graphs [3,4]. However, some social networks contain positive and negative relationships between users, and are called signed social networks. Significantly, there are some studies about mining cliques from the social

※ This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Manuscript received November 22, 2021; first revision March 14, 2022; second revision September 7, 2022; accepted October 1, 2022.

* **Corresponding Author:** Doo-Soon Park (parkds@sch.ac.kr)

¹ Dept. of Software Convergence, Soonchunhyang University, Asan, Korea (yangyuan621@gmail.com, parkds@sch.ac.kr, peng.sony61@gmail.com)

² School of Computer Science, Shaanxi Normal University, Xi'an, China (feehao@gmail.com)

³ Dept. of Computer Science, University of Exeter, UK (feehao@gmail.com)

⁴ Dept. of Innovation and Convergence, Hoseo University, Cheonan, Korea

⁵ Dept. of Computer Sciences and Engineering, Soonchunhyang University, Asan, Korea (hmp4321@naver.com)

Currently affiliation for author, Hyejung Lee, is the Institute for Artificial Intelligence and Software, Soonchunhyang University, Asan, Korea (hjlee100@sch.ac.kr)

network [5,6], while several studies have been conducted using signed social networks [7-9]. However, few studies have used formal concept analysis (FCA) to do the maximal balanced clique detection in the signed social networks.

To solve the maximal balanced clique detection problem (MBCP), it is necessary to propose some more efficient methods; therefore, this paper focuses on designing the modified three-way concept algorithms, which are more suitable for the signed social networks.

The organization of our paper is described as follows. We introduce the research background and our motivation in Section 1, with the related works and some terminologies introduced in Section 2, and then the maximal balanced clique detection algorithms for signed social networks are presented in Section 3. To show its validness and efficiency, we have applied it to some real datasets and synthetic datasets in Section 4; meanwhile, we also compare the experimental results with other methods. Finally, the conclusion and future works are provided in Section 5.

2. Basic Notions

In this section, we introduce several basic concepts of the signed networks, maximal balanced cliques, and three-way concept lattice.

A signed social network G is characterized as a graph structure, and the graph composed of nodes and edges. The nodes represent the users in social network, and edges have weights, which are positive and negative values and represent positive and negative relationships, respectively. For example, as shown in Fig. 1, this is a signed social network, and there are five nodes in a signed social network G , between these nodes, the solid lines represent the positive edges, and the dotted lines represent the negative edges.

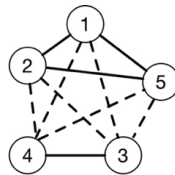


Fig. 1. Signed social network G .

To explain the definitions of the maximal balanced network [7], FCA [10], and three-way concept [11,12], we present here an example that lets us easily understand the original FCA and three-way concept.

Definition 1 (Operator \leq and \geq [11]). Let $F = (O, A, I)$ be a formal context. For any $X, M \subseteq O, Y, N \subseteq A$, a pair of three-way operators are defined mathematically below:

$$X^{\leq} = (X^{\uparrow}, X^{\uparrow*}) \tag{1}$$

$$Y^{\leq} = (Y^{\downarrow}, Y^{\downarrow*}) \tag{2}$$

$$(Y, N)^{\geq} = \{x \in O \mid x \in Y^{\downarrow} \text{ and } x \in N^{\downarrow*}\} = Y^{\downarrow} \cap N^{\downarrow*} \tag{3}$$

$$(X, M)^{\geq} = \{a \in A \mid a \in X^{\uparrow} \text{ and } a \in M^{\uparrow*}\} = X^{\uparrow} \cap M^{\uparrow*} \tag{4}$$

Based on Definition 1, three-way concepts are defined in detail as follows.

Definition 2 (Three-way concepts [11]). The three-way concepts including two parts: AE-concept and OE-concept.

- AE-concept: Let $F = (O, A, I)$ be a formal context. For any $X, M \subseteq O, Y \subseteq A, ((X, M), Y)$ is called an attribute-induced three-way concept (AE-concept), if and only if $(X, M)^\geq = Y$ and $Y^\leq = (X, M)$, where (X, M) is called the extent and Y is called the intent of $((X, M), Y)$. The set of all AE-concepts of $F = (O, A, I)$ can form a lattice by some partial order structure.
- OE-concept: Let $F = (O, A, I)$ be a formal context. For any $X \subseteq O, Y, N \subseteq A, (X, (Y, N))$ is called an object-induced three-way concept (OE-concept), if and only if $X^\leq = (Y, N)$ and $(Y, N)^\geq = X$, where X is called the extent and (Y, N) is called the intent of $(X, (Y, N))$. The set of all OE-concepts of $F = (O, A, I)$ can form a lattice by some partial order structure, and this lattice is called OE lattice.

Example 1. Table 1 is a formal context F . Table 2 is a supplemental formal context F^c for Table 1. Among them, O represents the object of the form context, and A represents the attribute corresponding to object O . Through the traditional FCA method mentioned above, using the formal context of Tables 1 and 2 and the operations in Definition 1 above, we can obtain all object sets with the same attributes (i.e., concepts, which are represented by nodes in Fig. 2), and form a concept lattice of the formal context F^c , which is represented as Fig. 2.

Table 1. Formal context F

$O \times A$	a_1	a_2	a_3	a_4	a_5
o_1	1	1	0	1	1
o_2	1	1	1	0	0
o_3	0	0	0	0	1
o_4	1	1	1	0	0

Table 2. Supplemental formal context F^c

$O \times A$	a_1	a_2	a_3	a_4	a_5
o_1	0	0	1	0	0
o_2	0	0	0	1	1
o_3	1	1	1	1	0
o_4	0	0	0	1	1

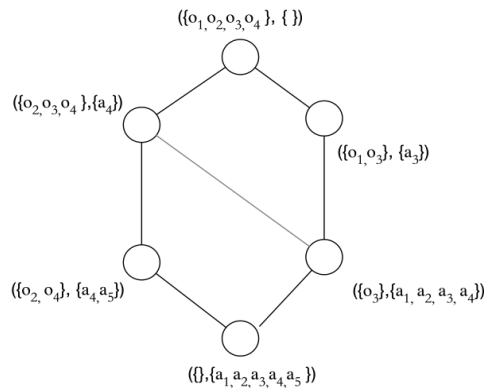


Fig. 2. Concept lattice of the supplement context F^c .

Similar to the acquisition process in Fig. 2, the concept lattice of the formal context F can be obtained. Based on the formal contexts F and F^c, the OE-concept lattice is obtained by Definition 2, which is shown in Fig. 3.

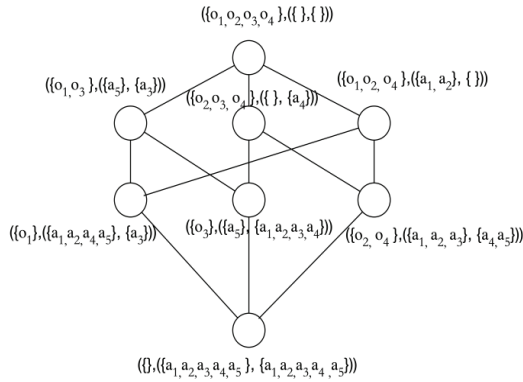


Fig. 3. OE lattice of contexts F and F^c.

3. Proposed Maximal Cliques Detection Approach

This section proposes two novel methods for maximal balanced clique detection in the signed networks. The first method obtains the modified OE lattice through the existing three-concept lattice algorithm using the improved three-way formal context and modified formal context. We have detected all maximal balanced cliques in the OE lattice. The second method is to optimize the first method to improve the detection of the maximal balanced cliques using the traditional FCA along with the modified formal context.

In the signed social networks, we applied the formal context and supplemental formal context for representing the positive and negative relationships between nodes, which is denoted as modified formal context.

Definition 3 (Modified formal context). If a signed network *g* has nodes {v₁, v₂, ..., v_n}, each node v_i represents an object and its attribute, and edge e = (v_i, v_j) is the relationship between the objects and attributes, I or I^c. If e ∈ E⁺, notes I(v_i, v_j) = 1, if e ∈ E⁻, notes I^c(v_i, v_j) = 1. Then, if F = (O, A, I) and F^c = (O, A, I^c) are modified formal contexts, they can be represented as two adjacency matrixes containing only 0 and 1.

$$M(F) = \begin{cases} m_{ij} = 1, & \text{if } (v_i, v_j) \in E^+ \text{ and } i \neq j \\ m_{ij} = 1, & \text{if } i = j \\ m_{ij} = 0, & \text{if } (v_i, v_j) \notin E^+ \text{ and } i \neq j \end{cases} \quad (5)$$

$$M(F^c) = \begin{cases} m_{ij} = 1, & \text{if } (v_i, v_j) \in E^- \text{ and } i \neq j \\ m_{ij} = 1, & \text{if } i = j \\ m_{ij} = 0, & \text{if } (v_i, v_j) \notin E^- \text{ and } i \neq j \end{cases} \quad (6)$$

Definition 4 (Modified OE and AE). Taking the modified formal context and supplemental formal context for generating OE-concept and AE-concept in a three-way concept algorithm, we can get a modified OE lattice (denoted MOE) and a modified AE lattice (denoted MAE).

Maximal cliques detection approach 1 (MCDA1): The modified formal context $F = (O, A, I)$ and supplemental formal context $F^c = (O, A, I^c)$ are used to obtain modified three-way concept lattices, MOE and MAE. When there is a concept in MOE and $X = Y$, i.e., $(X, (X, N))$, X and N form the maximal balanced clique represented by $\{X, N\}$, and $(X, (X, N))$ is the maximal balanced OE-concept.

Theorem 1. Let MOE be the OE-concept lattice of modified formal contexts F and F^c . For any $(X, (Y, N)) \in \text{MOE}$, if $X = Y$, it is denoted as $(X, (X, N))$, $\{X, N\}$ is the maximal balanced clique, such that $\forall u, v \in X$ or $u, v \in N \rightarrow (u, v) \in E^+$, and $\forall u \in X, v \in N$ or $u \in N, v \in X \rightarrow (u, v) \in E^-$.

Proof. There is $(X, Y) \in \text{FC}(F)$ and $(M, N) \in \text{FC}(F^c)$, and $(X, (Y, N)) \in \text{MOE}$.

- 1) If $X = Y$, we can get $(X, X) = (X, Y) \in \text{FC}(F)$, while (X, X) is an equi-concept in $\text{FC}(F)$ and (X, X) is a maximal clique and $\forall u, v \in X, (u, v) \in E^+$.
- 2) Because of $(X, (Y, N)) \in \text{MOE}$, if $X = Y$, we can get $(X, (Y, N)) = (X, (X, N)) \in \text{MOE}$. According to Definition 4, for a modified OE-concept $(X, (X, N))$, $(X, N)^{\geq} = X^{\downarrow} \cap N^{\downarrow*} = X$. Based on the definitions of concept, $X^{\downarrow} = Y = X$, $N^{\downarrow*} = M$, then $X^{\downarrow} \cap N^{\downarrow*} = X \cap M = X$. Then, $X \subseteq M$. Because of $(M, N) \in \text{FC}(F^c)$, $\forall u \in M, v \in N$ or $u \in N, v \in M \rightarrow (u, v) \in E^-$, we can get that $\forall u \in X \subseteq M, v \in N$ or $u \in N, v \in X \subseteq M \rightarrow (u, v) \in E^-$.

The working process of MCDA1 is described in Algorithm 1. In Algorithm 1, the sets of maximal balanced cliques and modified OE-concept are initialized in Line 1. Line 3 constructs formal contexts F and F^c for the signed social network G . The Construct_MFC is given by Algorithm 2. Line 4 follows the rules of the traditional three-way concept algorithm according to the modified contexts F and F^c to get the modified OE lattice. Lines 5–8 use Theorem 1 to detect the maximal balanced cliques from MOE. Finally, set the MBC returns from Line 9.

Algorithm 1 The detection of Maximal Balanced Cliques algorithm 1 ($G = (V, E^+, E^-)$)

```

1: Initialize the set of MBC, MOE
   //MBC is the set of Maximal Balanced Cliques
   //MOE is the set of Modified OE lattice
2: begin
3:  $F, F^c \leftarrow \text{Construct\_MFC}(G)$  // Construct the formal context  $F$  and  $F^c$  for signed social network  $G$ , Construct_MFC is Algorithm 2
4:  $\text{MOE} \leftarrow \text{Three-way concept}(V, F, F^c)$  // Using the traditional Three-way concept algorithm with modified context  $F$  and  $F^c$  to get a modified OE lattice MOE
5: for each  $(x, (y, n)) \in \text{MOE}$ 
6:   if  $x = y$  :
7:     if  $\text{length}(x) > 1$  and  $\text{length}(y) > 1$  :
8:        $\text{MBC} \leftarrow (x, n)$ 
9: return  $\text{MBC}$ 
10: end

```

Algorithm 2 Construct_MFC ($G = (V, E^+, E^-)$)

```

1: Initialize the matrix of  $F, F^c$  is full of 0
2: begin
3: for each  $(u_i, u_j) \in (V, V)$  :
4:   if  $i = j$  :
5:      $F(u_i, u_j) \leftarrow 1$ 
6:      $F^c(u_i, u_j) \leftarrow 1$ 
7:   if  $(u_i, u_j) \in E^+$  and  $i \neq j$  :
8:      $F(u_i, u_j) \leftarrow 1$ 
9:   if  $(u_i, u_j) \in E^-$  and  $i \neq j$  :
10:     $F^c(u_i, u_j) \leftarrow 1$ 
11: return  $F, F^c$ 
12: end

```

In Line 1 of Algorithm 2, we initialized the matrices of both contexts F and F^c to 0. Lines 3–10 construct two matrices, the three-way formal context F , and the supplemental formal context F^c , according to Definition 12. Line 11 returns the three-way contexts F and F^c .

Here, we present an example that lets us easily understand how to use an MCDA1 for the maximal balanced clique detection in the signed social networks.

Example 2. Fig. 4 shows a signed social network graph, which includes 11 nodes, eight positive edges, and 13 negative edges. The modified OE lattice of the signed social network G is obtained using Algorithms 1 and 2, meanwhile, Fig. 5 lists all OE-concepts of the modified OE lattice. The extents and intents of each concept are separated by “#”. According to Theorem 1, it is easy to find four suitable OE-concepts in the following list: $(3, 4, 6) \# ((3, 4, 6), (2, 10, 11))$; $(5, 8) \# ((5, 8), (7, 9))$; $(2, 10, 11) \# ((2, 10, 11), (3, 4, 6))$; $(7, 9) \# ((7, 9), (5, 8))$. Moreover, we obtained two maximal balanced cliques: $\{\{3, 4, 6\}, \{2, 10, 11\}\}$; $\{\{5, 8\}, \{7, 9\}\}$.

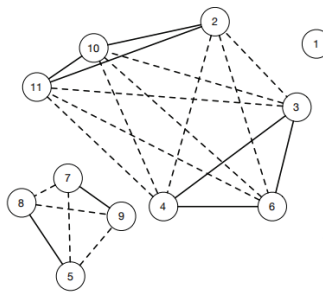


Fig. 4. The signed social network G .

```
(3, 4, 6)#((3, 4, 6), (2, 10, 11))
(5, 9)#(( ), (5, 9))
(2, 6, 10, 11)#(( ), (6, ))
(5, 7, 9)#(( ), (5, ))
(3, 11)#(( ), (3, 11))
(9, )#((7, 9), (5, 8, 9))
(8, 9)#(( ), (8, 9))
(5, 7, 8)#(( ), (7, ))
(11, )#((2, 10, 11), (3, 4, 6, 11))
(2, 4, 10, 11)#(( ), (4, ))
(5, 8)#((5, 8), (7, 9))
(1, )#((1, ), (1, ))
(3, 10)#(( ), (3, 10))
(6, 10)#(( ), (6, 10))
(3, )#((3, 4, 6), (2, 3, 10, 11))
(5, )#((5, 8), (5, 7, 9))
(7, )#((7, 9), (5, 7, 8))
(4, 10)#(( ), (4, 10))
(2, 6)#(( ), (2, 6))
(6, 11)#(( ), (6, 11))
(8, )#((5, 8), (7, 8, 9))
(7, 8, 9)#(( ), (8, ))
(10, )#((2, 10, 11), (3, 4, 6, 10))
(2, 3)#(( ), (2, 3))
(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11)#(( ), ( ))
(2, 10, 11)#((2, 10, 11), (3, 4, 6))
(4, 11)#(( ), (4, 11))
(7, 9)#((7, 9), (5, 8))
(5, 8, 9)#(( ), (9, ))
(2, )#((2, 10, 11), (2, 3, 4, 6))
(5, 7)#(( ), (5, 7))
(4, )#((3, 4, 6), (2, 4, 10, 11))
( )#((1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11), (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11))
(6, )#((3, 4, 6), (2, 6, 10, 11))
(7, 8)#(( ), (7, 8))
(2, 4)#(( ), (2, 4))
```

Fig. 5. Modified OE-concepts list.

In Fig. 6, the purple and orange parts are two maximal balanced cliques with vertex label numbers consistent with the above results.

Due to the three-way concept method being a NP-completed problem, so the MCDA1 also suffered from a high time complexity problem because if there are massive nodes in a signed social network, this

method cannot quickly obtain the modified OE lattice, making it unable to quickly obtain the maximal balanced cliques. Therefore, we analyzed the characteristics of OE-concepts based on the three-way concept method, and proposed another detection algorithm of maximal balanced cliques that can be quickly obtained, starting directly from the traditional concept analysis with the formal context and supplemental formal context, which is described as MCDA2.

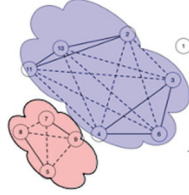


Fig. 6. Maximal balanced cliques in signed social network G .

Maximal cliques detection approach 2 (MCDA2). The modified formal context $F = (O, A, I)$ and supplement context $F^c = (O, A, I^c)$ are used as two input formal contexts for the traditional FCA method. According to the input contexts and formal context analysis, we obtained two modified formal concept lattices, $\Omega(F)$ and $\Omega(F^c)$.

Principle 1. Compare with each concept in $\Omega(F)$ and $\Omega(F^c)$. When there are two concepts (X, Y) and (C, D) in the $\Omega(F)$ and $X = Y$ and $C = D$, they are equi-concepts. Concurrently, if there is (X, D) in the $\Omega(F^c)$, X is the extent of (X, Y) and (X, D) , D is the intent of (C, D) and (X, D) , then X and D form a maximal balanced clique, which is denoted as $\{X, D\}$.

Algorithm 3 The detection of Maximal Balanced Cliques algorithm 2 ($G = (V, E^+, E^-)$)

```

1: Initialize the set of MBC, EC,  $\Omega(F)$ ,  $\Omega(F^c)$ 
   //MBC is the set of Maximal Balanced Cliques
   //EC is the set of equiconcept
   // $\Omega(F)$ ,  $\Omega(F^c)$  are the lattices of formal context  $F$  and  $F^c$ 
2: begin
3:  $F, F^c \leftarrow \text{Construct\_MFC}(G)$  // Construct the formal context  $F$  and  $F^c$  for signed social network  $G$ , Construct\_MFC is Algorithm 2
4:  $\Omega(F) \leftarrow \text{Formal concept analysis}(F)$  // Construct the lattice of formal context  $F$  by a traditional formal concept analysis method
5:  $\Omega(F^c) \leftarrow \text{Formal concept analysis}(F^c)$  // Construct the lattice of formal context  $F^c$  by a traditional formal concept analysis method
6: for each concept  $(x, y)$  in  $\Omega(F)$  :
7:   if  $x == y$  :
8:      $EC \leftarrow (x, y)$ 
9:   for each  $(x, y) \in EC$  :
10:    for each  $(m, n) \in \Omega(F^c)$  :
11:      if  $x == m$  or  $x == n$  :
12:         $MBC \leftarrow (x, n)$ 
13: return  $MBC$ 
14: end

```

Algorithm 3 explains the above processes. In Algorithm 3, Line 1 initializes the sets of maximal balanced cliques, equi-concept, concept lattice $\Omega(F)$, $\Omega(F^c)$. Line 3 constructs the formal contexts F and F^c for the signed social network G , while the construct-MFC is given in Algorithm 2. Lines 4–5 obtain

the concept lattice of formal contexts F and F^c . Lines 6–8 obtain the set of equi-concept. Lines 9–12 detect maximal balanced cliques through Approach 2 based on Principle 1. Finally, the set of maximal balanced clique will return in Line 9.

Example 3. Fig. 7 shows the concept lattice of contexts F and F^c . If we follow Approach 2, we can get four equi-concepts (the number of extent or intent is more than one): $(3, 4, 6)\#(3, 4, 6)$; $(2, 10, 11)\#(2, 10, 11)$, $(5, 8)\#(5, 8)$, $(7, 9)\#(7, 9)$. We used the previous four equi-concepts to obtain the four concepts in lattice $\Omega(F^c)$, and followed the rules in Approach 2: $((3, 4, 6) \# (2, 10, 11))$; $((5, 8) \# (7, 9))$; $((2, 10, 11) \# (3, 4, 6))$; $((7, 9) \# (5, 8))$. Then, we obtained two maximal balanced cliques: $\{\{3, 4, 6\}, \{2, 10, 11\}\}$; $\{\{5, 8\}, \{7, 9\}\}$. As shown in Fig. 6, the purple and orange parts are two maximal balanced cliques with vertex labels numbers consistent with the above results.

① The all concepts in $\Omega(F)$:

$(3, 4, 6)\#(3, 4, 6)$
 $(1.)\#(1.)$
 $(2, 10, 11)\#(2, 10, 11)$
 $(5, 8)\#(5, 8)$
 $(7, 9)\#(7, 9)$
 $()\#(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11)$
 $(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11)\#()$

② The all concepts in $\Omega(F^c)$:

$(2, 10, 11)\#(3, 4, 6)$	$(4, 10)\#(4, 10)$
$(5, 9)\#(5, 9)$	$(2, 6)\#(2, 6)$
$(6.)\#(2, 6, 10, 11)$	$(6, 11)\#(6, 11)$
$(5.)\#(5, 7, 9)$	$(7, 8, 9)\#(8.)$
$(3, 11)\#(3, 11)$	$(8.)\#(7, 8, 9)$
$(5, 8, 9)\#(9.)$	$(3, 4, 6, 10)\#(10.)$
$(11.)\#(3, 4, 6, 11)$	$(2, 3)\#(2, 3)$
$(8, 9)\#(8, 9)$	$(3, 4, 6)\#(2, 10, 11)$
$(7.)\#(5, 7, 8)$	$(4, 11)\#(4, 11)$
$(3, 4, 6, 11)\#(11.)$	$(5, 8)\#(7, 9)$
$(4.)\#(2, 4, 10, 11)$	$(9.)\#(5, 8, 9)$
$(3.)\#(2, 3, 10, 11)$	$(10.)\#(3, 4, 6, 10)$
$(7, 9)\#(5, 8)$	$(2, 3, 4, 6)\#(2.)$
$(2.)\#(2, 3, 4, 6)$	$(5, 7)\#(5, 7)$
$(1.)\#(1.)$	$(2, 4, 10, 11)\#(4.)$
$(3, 10)\#(3, 10)$	$(2, 6, 10, 11)\#(6.)$
$(6, 10)\#(6, 10)$	$(7, 8)\#(7, 8)$
$(2, 3, 10, 11)\#(3.)$	$(2, 4)\#(2, 4)$
$(5, 7, 9)\#(5.)$	$()\#(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11)$
$(5, 7, 8)\#(7.)$	$(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11)\#()$

Fig. 7 Maximal Balanced Cliques in Signed Social Network G.

4. Experiment and Performance Analysis

This section presents experimental datasets and execution times. The experimental environment consists of the macOS Catalina operating system and two Intel 2.3 GHz Quad-Core Intel Core i5, RAM 16G, and uses the programming language Python 3.8. In this section, we used two approaches proposed in Section 3 to perform experiments.

The modified three-way concept lattice method can deal with symmetric and asymmetric formal contexts, but the modified FCA can only handle the symmetric formal context of maximal balanced detection. Therefore, we first tested approach 1 (modified three-way concept) using three different real-life asymmetric datasets. We used large symmetric real-life datasets to compare approach 1 to approach 2 (modified FCA).

4.1 Time Efficiency Analysis

4.1.1 Asymmetric dataset

In this part, we tested the run-time of three asymmetric datasets by MCDA1. There are three real-life datasets (Table 3), and the first dataset is 11 nodes from AdjWordNet [13], which involves synonyms and antonyms of certain words. The second dataset is a network of intra-organization [14], which is a dataset of a consulting company (46 employees) and records the requests a frequent quantity of information or advice among employees. The third dataset is Freeman's EIES networks [15], which includes some personal relationships among 46 researchers.

Table 3. Asymmetric dataset

	Number of nodes	Number of edges
AdjWordNet	11	15
Intra	46	879
EIES	46	695

As shown in Fig. 8, comparing the run-time among different datasets, the first dataset is less than that of the other two datasets. Thus, although we can see that MCDA1 is highly sensitive to the number of nodes and edges, it has the advantage of being able to compute the asymmetric formal context that is not limited to social network datasets.

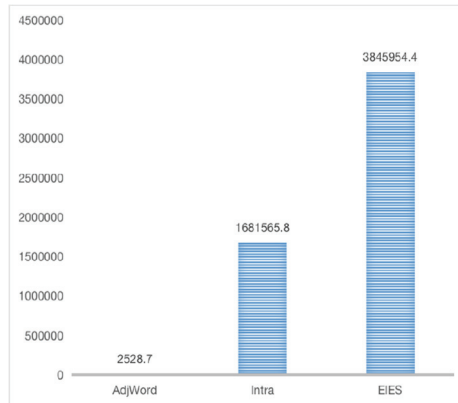


Fig. 8. Running time of modified three-way concept (unit: ms).

4.1.2 Symmetric dataset

In this part, we tested the run-times of MCDA1 and MCDA2 by using the Slashdot dataset [16]. Slashdot is known as a specific user community, whose website has the ability to allow users to submit the latest news on the site and for editors to evaluate the content. Additionally, the Slashdot dataset entails friends and relationships between Slashdot users. Slashdot includes 77,350 nodes and 516,575 edges, but in this paper, we take only the first 2,000 nodes of the dataset and experiment in groups of different sizes.

We transform the relationships of a dataset into undirected relationships without considering the positive or negative relationship in a single direction to obtain a symmetric matrix that simulates an undirected weighted social network (Table 4). The edges number of each dataset is counted in Fig. 9.

Fig. 10 shows the comparison of the run-time on datasets 1–7 with the MCDA1, MCDA2 and FCA-based methods (which was proposed in [17]). Note that the algorithm only works for detecting a k-balance clique in the signed social network to make this algorithm comparable with our algorithms, and we made a small improvement on the algorithm. The blue line shows the run-time by the MCDA1, the red line shows the run-time by the MCDA2, and the green line shows the run-time by the FCA-based method. According to Fig. 11, it is observed that the execution time of the MCDA2 method significantly outperforms the other two methods.

Table 4. Symmetric dataset

	Dataset													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Number of nodes	100	200	300	400	500	600	700	800	900	1,000	1,300	1,500	1,700	2,000

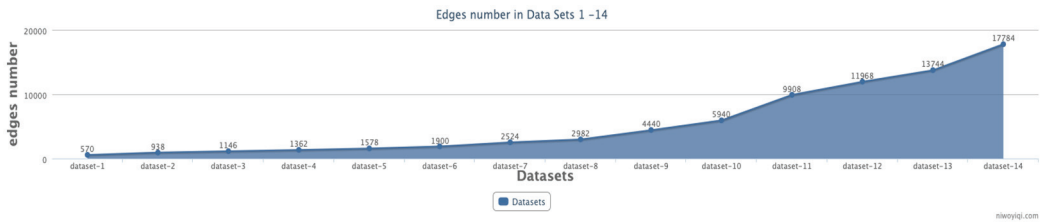


Fig. 9. Edges number of datasets 1–14.

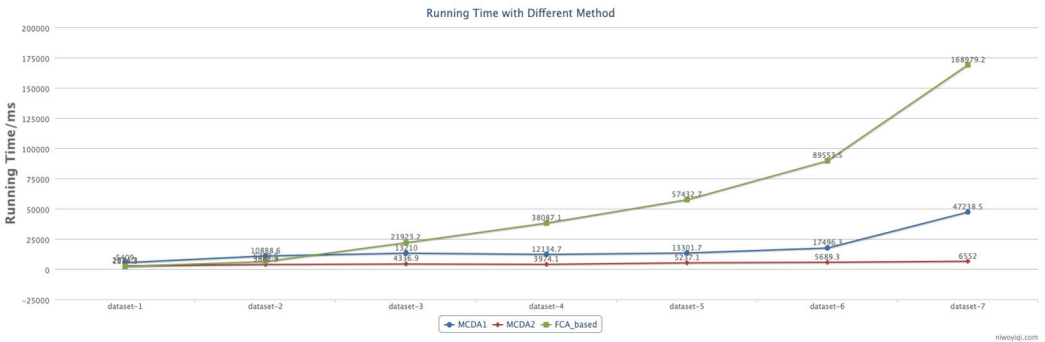


Fig. 10. Execution time of MCDA1, MCDA2, and FCA-based (datasets 1–7).

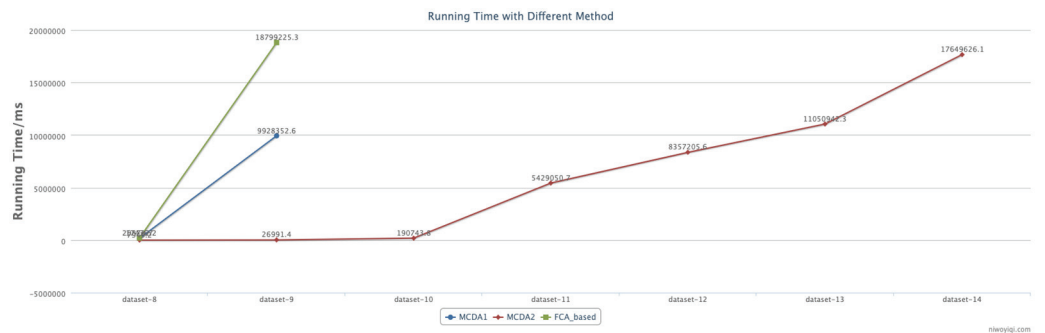


Fig. 11. Execution time of MCDA1, MCDA2, and FCA-based (datasets 8–14).

4.2 Case Study of Detecting Synonyms in Vocabulary

In Section 4.1, we used three datasets to compare the time efficiency of the approaches and demonstrate the utility of our methods. However, we not only focused on the time efficiency of our method, but also cared about the utility. Hence, in the case study, we applied the AdjWordNet dataset to our method, and the results illustrate the utility of our method. AdjWordNet is downloaded from WordNet (www://wordnet.priceton.edu). This dataset contains 117,000 synonyms and antonyms of certain words. Among them, the synonyms are linked by positive edges, while the antonyms are linked by negative edges, without the presences of edges between unrelated words.

In our case study, to observe the flow and effect of our method more clearly, we selected 25 words randomly from AdjWordNet. The graph of the 25 words is shown in Fig. 12 and Table 5.

Fig. 12 shows the 25 words and their relationships, and among them, the orange solid line shows that the two words are synonyms, while the blue dotted line represents that the two words are antonyms, with the words without a line bearing no relationship. Table 5 contains the vocabulary list of the nodes in Fig. 12.

Table 5. Twenty-five random words of AdjWordNet data

raw	rough	rude	relaxing	refined
interior	smooth	assumed	false	intimate
uneasy	ungratified	restful	reposeful	unsatisfied
outer	existent	veridical	actual	sour
participating	undynamic	active	undermentioned	treasured

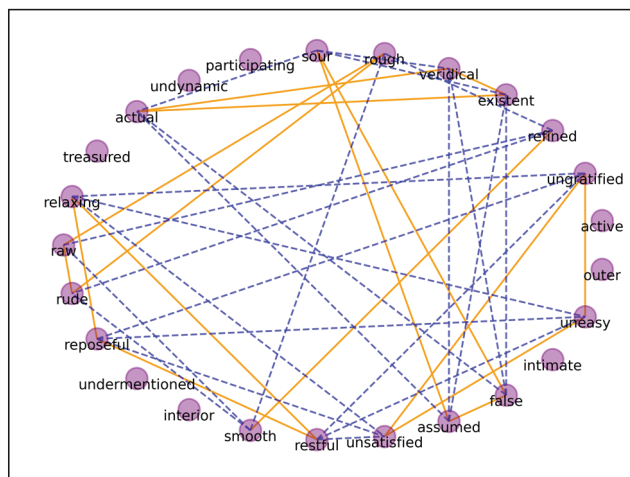


Fig. 12. Twenty-five words in AdjWordNet.

The results which are obtained by our method is shown in Table 6. In this table, the words in C_1 and C_2 are denoted as synonyms for each line, respectively and each column C_1 and C_2 represents a group of two antonyms.

We visualized the three maximal balanced cliques in Table 6, as shown in Fig. 13, which denotes the following: the green nodes are the first maximal balanced cliques, the orange nodes are the second maximal balanced cliques, the gray nodes are the third maximal balanced cliques, and the remaining purple nodes are the vocabulary with no synonyms and antonyms in this dataset.

Table 6. Result of maximal balanced cliques in AdjWordNet

Maximal balanced clique	C ₁	C ₂
1st	refined, smooth	raw, rough, rude
2nd	assumed, false, sour	existent, veridical, actual
3rd	relaxing, restful, reposeful	uneasy, ungratified, unsatisfied

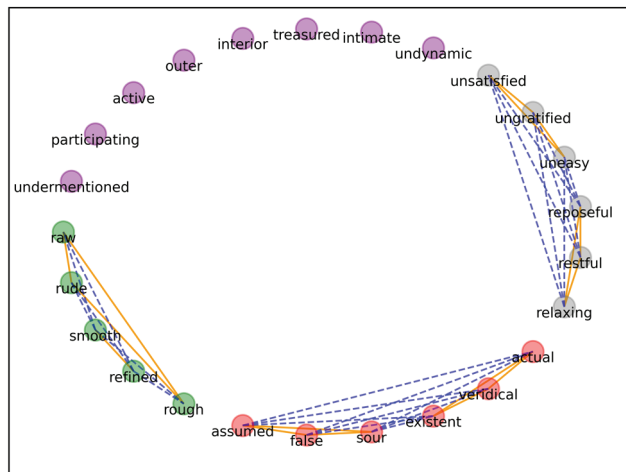


Fig. 13. Node visualization of three maximal balanced cliques.

5. Conclusion

In this paper, we proposed an improved three-way concept algorithm MCDA1 and an improved formal concept algorithm MCDA2 for detecting the maximal balanced clique in signed social networks. To demonstrate the feasibility of the proposed method, in the performance evaluation, four real-life signed social networks were used to detect maximal balanced cliques, and synonyms and antonyms were used on the AdjWorldNet dataset. Also, we found that the two methods provide both advantages and disadvantages. Accordingly, MCDA1 method has a wider application range on asymmetric datasets, whereas the MCDA2 method has a faster computation speed and higher efficiency. This proposed method can be applied in detecting synonyms and antonyms in a vocabulary list, detecting social network user groups, and analyzing the protein molecular structure. Moreover, it can be applied to more scenarios in the future. Going forward, we plan to solve the problem of long run-times in the MCDA1.

Acknowledgement

This research was supported by the National Research Foundation of Korea (No. NRF-2020RIA2B5B01002134) and the BK21 Fostering Outstanding Universities for Research (No. 5199990914048), European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 840922, the National Natural Science Foundation of China (No. 61702317), and the Fundamental Research Funds for the Central Universities (No. GK202103080).

References

- [1] P. Vilakone, X. Xinchang, and D. S. Park, "Movie recommendation system based on users' personal information and movies rated using the method of k-clique and normalized discounted cumulative gain," *Journal of Information Processing Systems*, vol. 16, no. 2, pp. 494-507, 2020.
- [2] P. Vilakone, K. Xinchang, and D. S. Park, "Personalized movie recommendation system combining data mining with the k-clique method," *Journal of Information Processing Systems*, vol. 15, no. 5, pp. 1141-1155, 2019.
- [3] L. Yuan, L. Qin, X. Lin, L. Chang, and W. Zhang, "Diversified top-k clique search," *The VLDB Journal*, vol. 25, pp. 171-196, 2016.
- [4] 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 Systems Journal*, vol. 11, no. 1, pp. 250-259, 2017.
- [5] S. Kumar, F. Spezzano, V. S. Subrahmanian, and C. Faloutsos, "Edge weight prediction in weighted signed networks," in *Proceedings of 2016 IEEE 16th International Conference on Data Mining (ICDM)*, Barcelona, Spain, 2016, pp. 221-230.
- [6] J. Kunegis, A. Lommatzsch, and C. Bauckhage, "The Slashdot Zoo: mining a social network with negative edges," in *Proceedings of the 18th International Conference on World Wide Web*, Madrid, Spain, 2009, pp. 741-750.
- [7] Z. Chen, L. Yuan, X. Lin, L. Qin, and J. Yang, "Efficient maximal balanced clique enumeration in signed networks," in *Proceedings of the Web Conference 2020*, Taipei, Taiwan, 2020, pp. 339-349.
- [8] J. Salminen, M. Hopf, S. A. Chowdhury, S. G. Jung, H. Almerexhi, and B. J. Jansen, "Developing an online hate classifier for multiple social media platforms," *Human-centric Computing and Information Sciences*, vol. 10, article no. 1, 2020. <https://doi.org/10.1186/s13673-019-0205-6>
- [9] S. Ahmad, M. Z. Asghar, F. M. Alotaibi, and I. Awan, "Detection and classification of social media-based extremist affiliations using sentiment analysis techniques," *Human-centric Computing and Information Sciences*, vol. 9, article no. 24, 2019. <https://doi.org/10.1186/s13673-019-0185-6>
- [10] Y. Yang, F. Hao, B. Pang, K. Qin, Z. Pei, and B. Li, "A quick algorithm on generating concept lattice for attribute-incremental streaming data," in *Proceedings of 2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; and IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, Zhangjiajie, China, 2019, pp. 2811-2816.
- [11] S. Yang, Y. Lu, X. Jia, and W. Li, "Constructing three-way concept lattice based on the composite of classical lattices," *International Journal of Approximate Reasoning*, vol. 121, pp. 174-186, 2020.
- [12] F. Hao, Y. Yang, G. Min, and V. Loia, "Incremental construction of three-way concept lattice for knowledge discovery in social networks," *Information Sciences*, vol. 578, pp. 257-280, 2021.
- [13] G. A. Miller, *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press, 1998.
- [14] R. L. Cross and A. Parker, *The Hidden Power of Social Networks: Understanding How Work Really Gets Done in Organizations*. Boston, MA: Harvard Business School Publishing, 2004.
- [15] S. C. Freeman and L. C. Freeman, "The networkers network: a study of the impact of a new communications medium on sociometric structure," School of Social Sciences, University of California, CA, 1979.
- [16] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Signed networks in social media," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Atlanta, GA, 2010, pp. 1361-1370.
- [17] F. Hao, S. S. Yau, G. Min, and L. T. Yang, "Detecting k-balanced trusted cliques in signed social networks," *IEEE Internet Computing*, vol. 18, no. 2, pp. 24-31, 2014.



Yixuan Yang <https://orcid.org/0000-0002-9334-566X>

She received her B.Sc. degree in Software Engineering from Taiyuan University of Technology, China (2017), and her M.Sc. degree in Software Engineering from Shaanxi Normal University, China (2020). She is currently pursuing her Ph.D. degree at Soonchunhyang University, South Korea. Her research interests include social computing, community detection and formal concept analysis.



Doo-Soon Park <https://orcid.org/0000-0002-2776-8832>

He received his Ph.D. in Computer Science from Korea University in 1988. Currently, he is a professor in the Department of Computer Software Engineering at Soonchunhyang University, South Korea. He is Dean of Graduate School at Soonchunhyang University, Director of BK21 FOUR Well-Life Big Data at Soonchunhyang University. He was President of KIPS from 2015 to 2015. He was editor in chief of JIPS from 2009 to 2012. He has served as an organizing committee member of international conferences including, BIC 2021, MUE 2021, WorldIT 2021, CSA 2020, BIC 2019, FutureTech 2019, WorldIT 2019, FutureTech 2018, BIC 2018. His research interests include social computing, big data analysis. He is a member of ACM, KIPS, and KMS.



Fei Hao <https://orcid.org/0000-0001-5288-5523>

He received the B.Sc. degree in Information and Computing Science and the M.Sc. degree in Computer Software and Theory from Xihua University, China, in 2005 and 2008, respectively, and the Ph.D. degree in Computer Science and Engineering from Soonchunhyang University, South Korea, in 2016. Since 2016, he has been with Shaanxi Normal University, Xi'an, China, where he is an Associate Professor. He is currently taking a Marie Skłodowska-Curie Individual Fellowship at the University of Exeter, Exeter, United Kingdom. His research interests include social computing, ubiquitous computing, big data analysis and processing and mobile cloud computing.



Sony Feng <https://orcid.org/0000-0003-3847-9662>

She received the bachelor's degree in ITE department from Royal University of Phnom Penh (2018), Cambodia. Currently, she is a master's student at Soonchunhyang University, South Korea. Her areas of interest include data mining, mobile artificial intelligence, and cloud computing.



Hyejung Lee <https://orcid.org/0000-0001-8359-1608>

She received her B.S., M.S., and Ph.D. in computer science engineering from Soonchunhyang University, South Korea, in 1997, 2000, and 2006, respectively. She is currently an assistant professor with the Department of Innovation and Convergence, Hoseo University (HSU), Asan, South Korea. Currently, her research interests contain data mining, parallel processing, big data processing, and computer education.



Min-Pyo Hong <https://orcid.org/0000-0002-0759-1777>

He received B.S. degree in computer science and statistics from Chungnam University in 1982, M.S. degree in Business Administration from Chungnam University in 1990. His research interests include SNA and Recommendation systems.