Neurocomputing 410 (2020) 41–50

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Using hybrid normalization technique and state transition algorithm to VIKOR method for influence maximization problem

Xiaojun Zhou [a,b], Rundong Zhang [a,*], Ke Yang [a], Chunhua Yang [a], Tingwen Huang [c]

[a] School of Automation, Central South University, Changsha 410083, China
[b] The Peng Cheng Laboratory, Shenzhen 518000, China
[c] Texas A & M University at Qatar, PO Box 23874, Doha, Qatar

ARTICLE INFO

Article history:
Received 26 February 2019
Revised 23 April 2020
Accepted 24 May 2020
Available online 1 June 2020
Communicated by Y. Yuan

Keywords:
Influence maximization
Normalization technique
Multiple criteria decision making
State transition algorithm

ABSTRACT

Influence maximization problem is the procedure of attempting to identify a group of K nodes in a social network in order to maximize the dissemination of influence under certain influence models. Based on state transition algorithm (STA) and a multiple criteria decision making (MCDM) method called Vise Kriterijumska Optimizacija I Kompromisno Resenje (VIKOR), a novel hybrid approach has been proposed to cope with the influence maximization problem in this paper. Firstly, an intelligent optimization paradigm called STA is introduced to obtain the most appropriate weights that are used to integrate the criteria of each alternative in the VIKOR method. Then, a hybrid normalization technique has been presented to allow the process of aggregating criterion with numerical and comparable data properly in this method. Several typical networks have been used to testify the effectiveness of proposed method and technique. Compared with other approaches, experimental results show that our approach can solve the influence maximization problem more effectively.

1. Introduction

The influence maximization problem is the procedure of aiming to select "important nodes" from a social network to maximize the expected number of influenced users [1]. Measuring the influential nodes and maximizing influence in social networks have received significant attention from the domain of network science in recent years. And that has the considerable significance for controlling the outbreak of epidemics, enhancing the effect of advertisements, optimizing information dissemination and identifying the influential individual in social networks [2]. For example, finding the most influential one who is easily to diffuse infectious disease in social network can help government adopt measures to control the outbreak of epidemics. At present, the measure of influential nodes mainly locates nodes that play an important role in information dissemination through the network topology such as neighborhood, location and path [3].

But current literatures mainly address influence maximization problem from two main perspectives which involve many disadvantages. The first perspective is focusing on novel centrality measures to identify influential nodes. Kandhway et al. compared the performance of various centrality measures (pagerank, degree, closeness and betweenness) in maximizing the spread of a message in the optimal control framework [4]. Kundu et al. proposed a centrality measure for independent cascade model based on diffusion probability (or propagation probability) and degree centrality [5]. Deng et al. considered four different types of centrality measurement methods and added a modification coefficient to evaluate the edge probability to solve the influence maximization [6]. But these literatures aim at concentrating the current state of research on centrality measures for social networks. Partial centrality measures have been considered and complete information is hard to find [7]. Another perspective is using ruled-based and heuristic algorithms to settle the influence maximization. Yang et al. proposed a multi-objective discrete particle swarm optimization algorithm for influence maximization-cost minimization problem [8]. Tsai et al. proposed an algorithm called search economics for influence maximization (SEIM) which is motivated by the concept of return on investment to design its search strategies [9]. However, traditional rule-based and heuristic algorithms did not focus on finding the "key-player" in connected social graph [10].

Adopting MCDM method to solve the influence maximization problem can avoid above-mentioned disadvantages. Because

* Corresponding author at: School of Automation, Central South University, Changsha 410083, China.

E-mail addresses: michael.x.zhou@csu.edu.cn (X. Zhou), zrd1018@csu.edu.cn (R. Zhang).

https://doi.org/10.1016/j.neucom.2020.05.084
0925-2312/

plenty of centrality measures such as degree centrality, betweenness centrality, eigenvector centrality, closeness centrality, page rank centrality can be considered and the "key-players" can be found by MCDM method. There are scholars trying to solve the influence maximization problem by MCDM method [11–13]. The MCDM method aims to identify a set of ordered alternatives involving multiple independent criteria. In fact, every node in social network with many centrality measures is the same as alternative with criteria in MCDM method. Nodes can be considered as alternatives and centrality measures can be seen as criteria. It can be concluded that conflicting and non-commensurable criteria must exist in MCDM circumstance. Although the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) has been adopted by Mehrdad et al. to solve the influence maximization problem [13], there is still plenty of room for improvement. TOPSIS has poor performance in solving influence maximization problem. Integrating unit inconsistency data is not reasonable and data with small order of magnitude can be neglected in integrating process. The local optimal weights in TOPSIS are obtained.

In this paper, another typical MCDM method called Vise Kriterijumska Optimizacija I Kompromisno Resenje (VIKOR) has been employed to solve the influence maximization problem. VIKOR was developed by Opricovic in 1998 [14]. VIKOR method focuses on selecting and ranking from a set of feasible alternatives. It determines compromise solution for a problem with conflicting criteria to help the decision maker reaching a final course of action. The optimal solution is so difficult to obtain that the compromise solution is often chosen instead. Using a particular measure of "closeness", the best alternative can be chosen as the one which is nearest to the positive ideal solution. The VIKOR method can balance the maximum group effectiveness of the "numerousness" and the minimum individual sacrifice of the "antagonist" which can also been seen as its major advantage [15]. VIKOR has been selected to settle the influence maximization problem in this study. But there still exist two main challenges have not been settled.

The first challenge is that unit inconsistency and magnitude mismatch problem commonly exist when integrating different centrality measures in VIKOR. Non-commensurable data exist in data mining commonly. It is necessary to normalize the unit and magnitude of different centrality measures into the same quantity. Because it is unreasonable to integrate unit inconsistency data with different physical meaning. Also, it is difficult to operate data with different orders of magnitude because data with small order of magnitude are easily to be neglected in integrating process [16]. A hybrid normalization technique has been proposed in this study to resolve this issue.

Another challenge is that the weights in VIKOR method are usually obtained from experience. The significance level of each criterion can be represented by corresponding weight. The traditional VIKOR method assumes that the weights of the criteria are determined previously by decision makers (DMs) or experts [17]. However, obtaining weights according to the cognition and strategic ability of the DMs or experts are hard to be assumed precisely [18]. Obtaining the weights in VIKOR has been modelled as an optimization problem which can be solved by optimization algorithm in this study. Recently, in order to solve various optimization problems, many optimization algorithms such as genetic algorithm (GA), differential evolutionary algorithm (DE), particle swarm optimization (PSO) have been developed. A novel intelligent optimization algorithm named state transition algorithm (STA) has been employed successfully in various complicated optimization problems, found to be robust and rapidly as compared to traditional optimization methods and exhibited excellent applicability [19–21]. In this study, STA has been adopted to obtain the optimal weights in VIKOR method. Three types of case studies have been presented to demonstrate the excellent performance of the proposed method in this paper. The major novelties and contributions in this paper are briefly listed as below:

(i) This study settles the influence maximization problem with VIKOR method firstly. Employing VIKOR method to solve the influence maximization problem is efficient.
(ii) This study proposes a hybrid normalization technique to handle the unit inconsistency and magnitude mismatch problem in data preprocessing. Data with inconsonant unit are hard to integrate and data with small order of magnitude will be neglected under preprocessing.
(iii) The weights in traditional VIKOR method are obtained optionally and groundless. Obtaining the weights in VIKOR method is an optimization problem. In this study, the weights in VIKOR method are obtained by STA, which has good performance both in speediness and accuracy under many optimization circumstances.

The remainder of this paper is organized as follows. The next Section 2 provides a description of influence maximization problem. Utilization of VIKOR as MCDM method is presented in Section 3. In Section 4, the normalization technique and STA this paper proposed are described. Section 5 provides an illustrative case study to show the applicability of the extended VIKOR method. Section 6 concludes and suggests future research directions for improvement.

2. Description of influence maximization problem

The definition of the influence maximization problem will be described in this section.

2.1. Influence maximization

Influence maximization was formally defined by Kemple et al. in 2003 [22]. In the study of this problem, a social network is modelled to represent as a graph $G = (V, E)$, where V is the group of

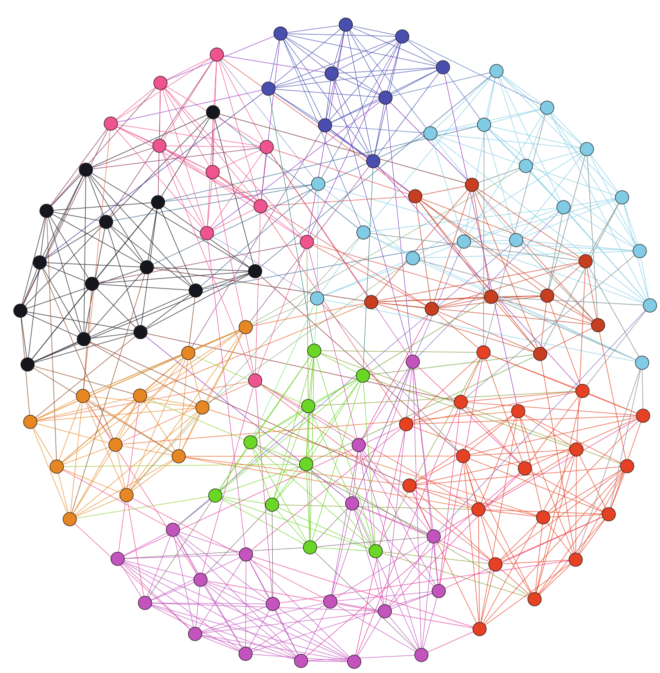


Fig. 1. Sample social network for influence maximization problem.

nodes in G and E is the group of edges in G [1]. Fig. 1 shows the sample social network. In this example, users that are denoted as nodes in social network can be symbolized as V, the social relationships between users that are denoted as connecting lines can be represented as E. Due to the structure of network, some nodes can have intrinsically higher influence than others at a global level. The influence maximization problem aims to obtain a K-sized group of users to spread influence rapidly and diffusely. The nodes are evaluated according to their centrality measures and "important nodes" can be identified by ranking the nodes in social network. Centrality measure is an important fields of social network research which has been extensively studied [23]. The concepts of centrality measures have been presented as follows:

2.1.1. Degree centrality

Degree centrality (D_C) has been widely accepted as the simplest and most popular measure. In social networks, people transmit information to and exert influence on others who they acquaint with. The one who acquaints with more people are more important than others, in this sense, having higher degree. As for the person j in social network, degree centrality of the one can be stated as follows:

$$Degree \quad D_C(j) = \sum \delta(i,j) \tag{1}$$

If person i is connected to person j, then $\delta(i,j) = 1$ and otherwise $\delta(i,j) = 0$. But in many actual social networks, person i acquaint with person j but opposite is not. For person i, person j is out-degree centrality (O-D_C) and person i is in-degree centrality (I-D_C) for person j [24]. From Fig. 2, person 5 is out-degree centrality and person 1 is in-degree centrality for person 4.

2.1.2. Eigenvector centrality

If a person is friendly with many people, it is probable to be contagious because the one can effect with others easily. This index that is calculated by using vectors and eigenvalues is rather more complicated than the others. It is updated repeatedly for the original value which is supposed as the centrality of eigenvalue vector concerning all the vertices. The formula can be stated as follows:

$$x_i = \frac{1}{\lambda}\sum_{i\in M(i)} x_j = \frac{1}{\lambda}\sum_{j=1}^{N} A_{ij}x_j; \tag{2}$$

where N is the amount of vertices, $M(i)$ represents the group of neighbors for vertex i, A is the adjacency matrix, λ represents the eigenvalue of matrix A. The eigenvalue of matrix A can be counted by the formula as follows:

$$det(A - \lambda I) = 0 \tag{3}$$

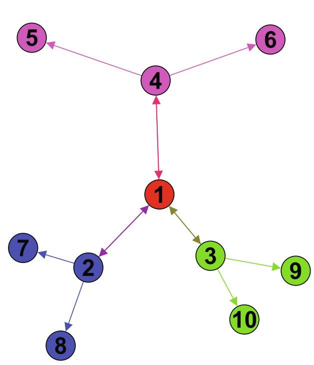


Fig. 2. Degree centrality.

The eigenvalue centrality (E_C) of the one in social network is assumed as the first eigenvector of the adjacency matrix in value [25].

2.1.3. Closeness centrality

The connections between people determine the effect and message dispersion in a social network. A person may be more important if the one can contact with more other people directly or go through shortest roundabout paths than other people. The closeness centrality (C_C) index of person i can be described as the reverse gross of the shortest distance of person i from all others in the network [26]. The formula can be stated as follows:

$$Closeness \quad C_C(i) = \frac{1}{\sum_{v_j\in V\wedge i\neq j} d_{ij}}; \tag{4}$$

where d_{ij} denotes the shortest path from person i to person j in length.

2.1.4. Betweenness centrality

When one acquaints with more others in the social network, the one can be seen as a pivotal person because of the flow can be controlled between them. One of the shortest path would exist between two people if they are chosen in a connected diagram. The total number of short paths between person i and j is noted as g_{ij}. The number of short paths between person i and j by way of person a to all the short paths between person i and j ratio is the betweenness centrality (B_C) of person a which can be described as follows [27]:

$$Betweenness \quad B_{Cij}(k) = \frac{g_{ij}(k)}{g_{ij}} \tag{5}$$

Fig. 3 shows that person 3 has the highest betweenness centrality, which means the one is more important than others.

2.1.5. PageRank centrality

PageRank (PR) can be used as a centrality based on the web, mainly by counting citations or hyperlinks to the targeted web page. This can be used as measure of importance or quality for the webpage. PageRank centrality extends this idea by not counting links from all pages equally and normalizing the number of links on a page [28]. Assuming page A has pages $T_1, T_2, \ldots, T_n$ which link to it. $PR(A)$ is defined as the number of links existing with page A. The parameter damping factor which can be denoted as d is set within the range of $[0,1]$. The PageRank index of page A can be formulated as follows:

$$PR(A) = 1 + d\left(\sum_{i=1}^{n}\frac{PR(T_i)}{C(T_i)} - 1\right) \tag{6}$$

The centralities mentioned above are normally used indexes which are applied to identify the most important people in influence maximization problem. Using different weights for these criteria will generate different ranking orders in this problem. The question of which one is the most significant in social network can be answered by weighting and combining multiple measures in VIKOR method.

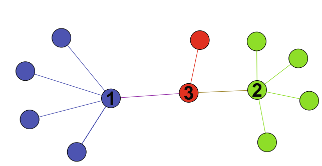


Fig. 3. Betweenness centrality.

2.2. Multiple criteria decision making

Generating alternatives, establishing criteria, evaluating alternatives, assessing criteria weights and application of a ranking method comprise the MCDM procedure [29]. The evaluation of alternatives according to different criteria should be performed depending on the objective of the influence maximization problem. The alternatives should be evaluated according to each criterion from series of established criteria. In fact, the influence maximization problem can be settled by MCDM method. The nodes and centrality measures in influence maximization problem can be seen as the alternatives and criteria in MCDM method respectively. Finding the most influential node is the target of the influence maximization problem, which is the same as the aim of MCDM method. Obtaining the most suitable weights for representing the significance of each criterion is an essential work. It is arduous for a decision maker (DM) to assign precise weights for all criteria. Most MCDM problems could be described by way of the following sets:

- A set of m alternatives called $A = \{A_1, A_2, \ldots, A_m\}$.
- A set of n criteria called $C = \{C_1, C_2, \ldots, C_n\}$.
- A set of performance ratings of $A_i (i = 1, 2, \ldots, m)$ on criteria $C_j (j = 1, 2, \ldots, n)$ called $F = \{f_{ij} | i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n\}$.
- A set of weights $w_j (j = 1, 2, \ldots, n)$ corresponding to f_{ij} for each alternative A_i.

The MCDM method is trying to rank the alternatives with various criteria. Different MCDM methods rank alternative A_i by integrating corresponding performance ratings f_{ij} with weights w_j in diverse ways. In this paper, the alternatives can be denoted as the nodes in social network and the criteria $C_1, C_2, \ldots, C_n$ can represent the above mentioned centrality measures respectively. So n equals to 6 and m can be identified according to the size of social network. The VIKOR method ranks nodes in influence maximization problem by integrating various centrality measures.

3. Utilization of VIKOR method as MCDM method

The VIKOR method was designed for multiple criteria optimization of decision making systems. VIKOR is a typical MCDM method. In this section, The VIKOR was taken to deal with the influence maximization problem. The weights which were obtained intuitively in previous VIKOR method for each criterion is vital and hard to identify. Determining the weights is a significant optimization problem which can be solved by intelligent optimization algorithm.

3.1. VIKOR method

The VIKOR method that consists of defining positive and negative ideal targets to determine the relative distance of each alternative has been developed to solve the MCDM problem [30]. A weighted compromise ranking is obtained to determine the importance of the m alternatives A_i available after each relative distance is calculated. It focuses on ranking a set of alternatives with conflicting criteria and determining compromise solutions for a problem, which can help the decision makers to reach a feasible solution closest to the ideal one. Inspired from the $L_p - metric$ which is used as an aggregating function in VIKOR method, the compromise function of multiple criteria measures can be established. The measure L_{pi} represents the distance between the alternative A_i and the ideal solution. f_{ij} represents each one of the various alternatives A_i measuring against the criteria C_j. f_j^* and f_j^- represent the best and worst values of evaluations of alternatives by the jth criteria respectively. w_j denotes the weight of the jth criteria. The VIKOR method comes into being by the form of $L_p - metric$ which can be formulated as follows:

$$L_{pi} = \left\{ \sum_{j=1}^{n} \left[w_j \frac{(f_j^* - f_{ij})}{(f_j^* - f_j^-)} \right]^p \right\}^{1/p} \quad (7)$$
$$1 \leqslant p \leqslant \infty; \quad i = 1, 2, \ldots m$$

The VIKOR method deploys $L_{1,i}$ (as S_i) and $L_{\infty,i}$ (as R_i) to signify the measure of ranking. VIKOR is mainly based on conventional theory of compromise programming. The compromise ranking algorithm is composed of the enumerative steps [31,32]:

Step 1: Determine the best f_j^* and the worst f_j^- values of evaluations of alternatives A_i by each criteria C_j. For a benefit criterion that will be:
$f_j^* = \max_i f_{ij}$ and $f_j^- = \min_i f_{ij}$.
If it is a cost criterion that is to be:
$f_j^* = \min_i f_{ij}$ and $f_j^- = \max_i f_{ij}$
Step 2: Compute S_i and R_i using Eqs. (8) and (9) respectively. S_i and R_i represent the effectiveness measure and the sacrificing measure for the alternatives respectively.

$$S_i = \sum_{j=1}^{n} \left[w_j \frac{(f_j^* - f_{ij})}{(f_j^* - f_j^-)} \right] \quad (8)$$

$$R_i = max_j \left[w_j \frac{(f_j^* - f_{ij})}{(f_j^* - f_j^-)} \right] \quad (9)$$

where w_j is the weight of criteria C_j.
Step 3: Compute Q_i using Eq. (10) for alternative A_i.

$$Q_i = v \frac{(S_i - S^*)}{(S^- - S^*)} + (1 - v) \frac{(R_i - R^*)}{(R^- - R^*)} \quad (10)$$

where $S^* = \min_i S_i, S^* = \max_i S_i,\ R^* = \min_i R_i, S^* = \max_i R_i, v$ is the weight for the decision making strategy of the maximum group effectiveness and $(1 - v)$ is the weight of the individual sacrifice. This study set v at 0.3 according to the experimental result.
Step 4: Form three ranking lists by ranking the alternatives according to ascending order of the values S, R and Q. Meanwhile, the lower value will be accepted because it denotes a better alternative.
Step 5: The compromise alternative ranking of the best by the minimum of Q is labeled as the alternative A' if the following two conditions are satisfied.

- Condition One. Acceptable advantage: $Q(A) - Q(A') \geqslant 1/(m-1)$. where, A is the alternative ranking secondly by the minimum of Q.
- Condition Two. Acceptable stability: The best rank by S or/ and R must be alternative A'.

Step 6: If one of the conditions in Step 5 is no satisfied, a group of compromise solutions are proposed which include:

- Alternatives $A', A, \ldots, A^n$ if only condition one is not satisfied. The closeness of the alternative A^n ranked nth by Q is determined by $Q(A^n) - Q(A') \geqslant 1/(m-1)$.
- Alternatives A' and A if only condition two is not satisfied
  Step 7: Select the compromise alternative $Q(A^i) (1 \leqslant i \leqslant m)$ as the best solution with the minimum value of Q_i with above conditions satisfied.

VIKOR provides a particularly effective tool in MCDM situations where human vague judgments including preferences are not involved. The index Q obtained by Eq. (10) can be denoted as Q_v. Q_v that denotes the realistic result can be used to represented as the ranking scores obtained by VIKOR. The relationship between ranking result and weights can be denoted as follows:

$$Q_v = f_{VIKOR}(w) \tag{11}$$

where $w = [w_1, \ldots, w_j, \ldots, w_n]$ and $w_j \in [0,1], \sum_{j=1}^{n} w_j = 1$. It is obvious to see that the ranking result obtained by VIKOR is confirmed after the weights are determined. Determining the weights in VIKOR is significant because unreasonable weights can obtain unfeasible ranking results. The weights in traditional VIKOR method are obtained optionally and groundless. Finding suitable weights in VIKOR is an optimization problem. This study aims to settle this optimization problem.

3.2. Optimization of the weights in VIKOR method

The weights in VIKOR method that were obtained intuitively in previous literatures are hard to be identified. Finding suitable weights in VIKOR is an optimization problem which aims to get the maximal correlation coefficient between ideal and reality ranking [13]. The formula of optimization problem can be stated as follows:

$$\max \quad \rho_{Q_v,Q_q} = \frac{(Q_v - \bar{Q_v})(Q_q - \bar{Q_q})}{\sigma_{Q_v,Q_q}} \tag{12}$$

where $Q_v = f_{VIKOR}(w)$ denotes the score of VIKOR method which can be seen as realistic result and Q_q denotes the partial ideal score. ρ represents the correlation coefficient between the two scores.

4. The proposed method

The proposed technique and method in this study are introduced in this section. Taking the normalization technique to normalize non-commensurable criteria is an intractable work in MCDM problem. The STA can obtain the optimal weights for criteria rapidly and accurately.

4.1. Normalization technique for preprocessing

Vector normalization technique is commonly used in previous literatures to normalize the decision matrix [33]. It is unreasonable to use vector normalization technique when the data with different orders of magnitudes. Because the data will exist great differences after taking the vector normalization technique. A hybrid normalization technique is proposed in this study to replace the vector normalization technique for solving the different orders of magnitude problem. The hybrid normalization technique can be described as below.

Decimal scale normalization technique is based on the movement of decimal point of the performance ratings. The decimal point numbers are moved depends on the maximum absolute performance ratings. The decimal scale normalization can be described as follows:

$$r = \frac{f}{10^m} \tag{13}$$

where m is the smallest integer satisfying $\max|r| < 1$. After taking the decimal scale normalization, the data can be with the same order of magnitude. But the maximum data may equal to 1. Global normalization performs a linear alteration on the original data. That can be formulated as follows:

$$r'_{ij} = \frac{r_{ij} - \min(\min_i(r_{ij}))}{\max(\max_i r_{ij}) - \min(\min_i r_{ij})}$$

After taking the global normalization, the data can be normalized within the range of $[0,1]$. The Gaussian normalization can be described as follows:

$$r = \frac{r'_i - \bar{r}'}{\sqrt{S^2}} \tag{14}$$

where

$$\bar{r} = \frac{\sum_{i=1}^{m} r'_{ij}}{m} \quad \& \quad S^2 = \sum_{i=1}^{m}\left(r'_{ij} - \bar{r}'\right)^2 \tag{15}$$

The Gaussian normalization method essentially normalizes the rating distribution of the data to a Gaussian distribution. Integrating different centrality measures with unit inconsistency problem and magnitude mismatch problem in VIKOR method is unreasonable. Because unit inconsistency data have different physical meaning. Also, operating data with different orders of magnitude is easily to neglect data with small order of magnitude. In this study, the hybrid normalization technique is proposed to settle this challenge.

4.2. Optimization of weights in VIKOR using STA

Inspired by the modern control theory and representation of state space, Zhou proposed the novel intelligent optimal algorithm, called state transition algorithm [34,35]. STA has been used to solve so many optimization problems [36–39]. The state is represented as a solution and the process of updating the current solution can be considered as state transition. Generally speaking, the standard form of generation of solution in STA can be formulated as follows:

$$\begin{cases} \mathbf{x}_{i+1} = A_i \mathbf{x}_i + B_i \mathbf{u}_i; \\ y_{i+1} = f(\mathbf{x}_{i+1}), \end{cases} \tag{16}$$

where $\mathbf{x} \in \mathbb{R}^n$. $\mathbf{x}$ is the current solution of the optimization problem named the state, A_i and B_i are state transition matrices which are transformation operators with proper dimension; u_i represents a history state or a function of $\mathbf{x}_i$; $f(\cdot)$ indicates the evaluation function, called fitness function [40].

There exist four transformation operators in STA to generate the candidate solutions which will be chosen in the next iteration. The four transformation operators are designed as follows:

- Rotation transformation operator.

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha \frac{1}{n\|\mathbf{x}_i\|_2} R_r \mathbf{x}_i, \tag{17}$$

where α is positive rotation factor; $R_r \in \mathbb{R}^{n\times n}$ is a matrix whose elements are selected randomly within the interval of $[-1,1]$. $\|\cdot\|_2$ is the 2-norm of a vector. The rotation transformation aims to search in a hypersphere with a maximal radius [41].

- Translation transformation operator.

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \beta R_t \frac{\mathbf{x}_i - \mathbf{x}_{i-1}}{\|\mathbf{x}_i - \mathbf{x}_{i-1}\|_2}, \tag{18}$$

where β is positive translation factor; $R_t \in \mathbb{R}$ is a random matrix whose elements between -1 and 1. The formula of searching along a line from $\mathbf{x}_{i-1}$ to $\mathbf{x}_i$ at the starting point $\mathbf{x}_i$ with the length β has been contained in the translation transformation [42].

- Expansion transformation operator.

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \gamma R_e \mathbf{x}_i \tag{19}$$

where γ is positive expansion factor; $R_e \in \mathbb{R}^{n\times n}$ is a diagonal matrix obeying the Gaussian distribution whose elements are chosen randomly. Expanding the elements in $\mathbf{x}$ to the extent of $[-\infty, +\infty]$ is contained in the expansion transformation so that searching the entire space can be guaranteed [43].

- Axesion transformation operator.

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \delta R_a \mathbf{x}_i \tag{20}$$

where δ is positive axesion factor; $R_a \in \mathbb{R}^{n\times n}$ is a diagonal matrix obeying the Gaussian distribution whose elements are selected randomly and there are only one random position with a non-zero value. Searching along the axes is included in the axesion transformation to strengthen the single dimensional search [44].

For better understanding of detailed programs of STA in MATLAB, the toolbox up to date can be downloaded via the following web page link: https://www.mathworks.com/matlabcentral/fileexchange/52498-state-transition-algorithm. STA is employed to cope with the optimization problem of obtaining the optimal weights which can be stated as follows:

a) Fitness function.

The output of VIKOR is obtained by calculation of each criteria on the basis of network and respective weights. The correlation coefficient of scores is considered as the fitness function, which can be acquired by the social skills questionnaire denoted as $Q_{questionnaire}(Q_q)$ and corresponding calculated scores denoted as $Q_{vikor}(Q_v)$ based on VIKOR method [13]. The correlation coefficient has be formulated in Eq. (12).

b) Constraint.

Every solution in each population is a n-dimensional variable which represents weights of various criteria in VIKOR method. In order to keep the binary opposition between any two criteria and satisfy the constraint condition strictly, linear normalization has been adopted before employing the weights into VIKOR. The linear normalization can be described as follows:

$$w'_j = \frac{w_j}{\sum_{j=1}^{n} w_j}; \tag{21}$$

the constraint on the aforementioned weights is formulated as follows:

$$w'_j \in [0,1], j = 1, 2, \ldots, n \quad \& \quad \sum_{j=1}^{n} w'_j = 1 \tag{22}$$

STA generates weights randomly which do not obey the constraint condition at the beginning. After taking the linear normalization, the constraint condition will be satisfied and relative relation can be kept strictly so that the weights can be adopted into VIKOR method.

The challenge that integrating different indexes by VIKOR method with uncertain weights can be settled by adopting STA. The optimal weights can be identified by this novel optimization algorithm.

5. Experiment results and analysis

There typical networks are used to implement the proposed method in this section. The description of network and parameters setting of STA are included in this section. Finally, the result analysis is represented in detail.

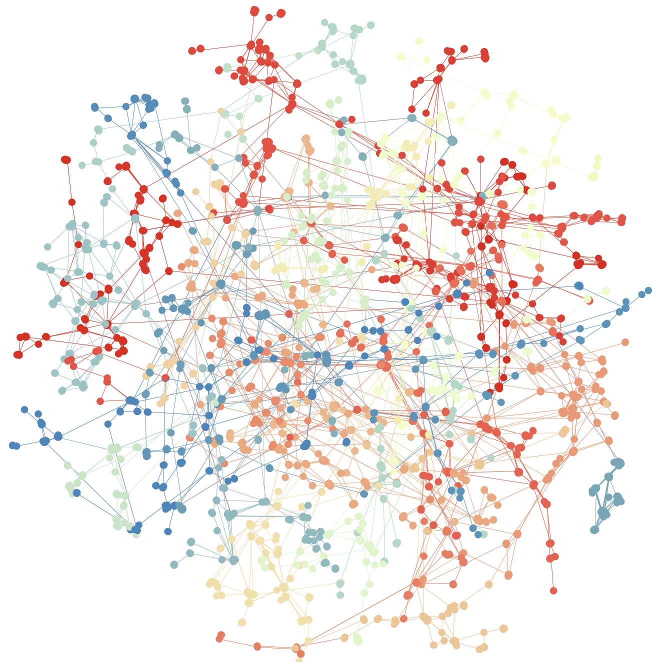


Fig. 4. Abrar student network.

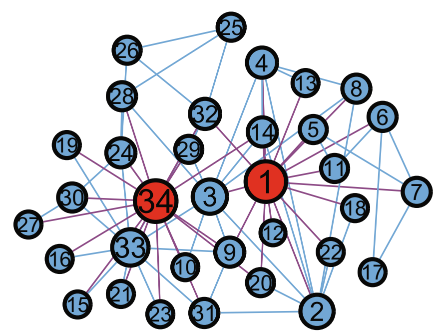


Fig. 5. Zachary's karate club network.

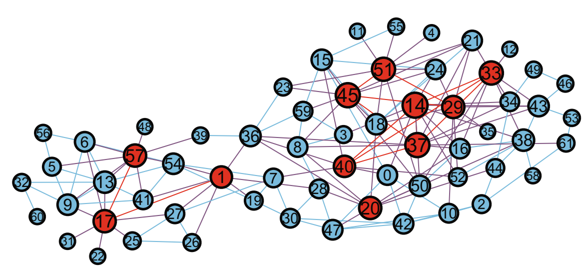


Fig. 6. Bottlenose Dolphin network.

5.1. The dataset

5.1.1. Abrar student network

A widely known medium size social network called the Abrar dataset has been extensively studied [45]. The directed links between individual a and b can be established if a has the cellphone number of b. Inspired by the repeated random subsampling (RRSS), 41 students are randomly selected from the total number of students (about 25%) in the network to complete a social skills questionnaire. Filling in the social skill questionnaire

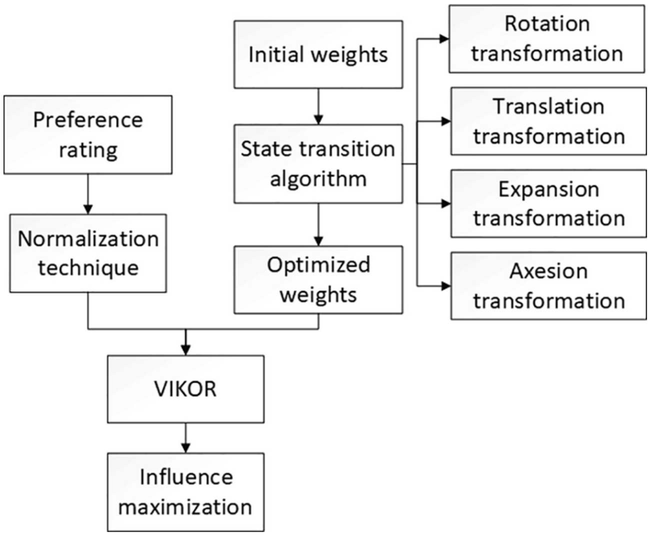


Fig. 7. The framework of the proposed method.

is performed by the selected samples aiming to yields the social skill score of each student in the sample. In 1992, Inderbitzen and Foster had designed the questionnaire which is employed to implement this work [46]. Fig. 4 shows the Abrar student network which is a complex and multi-nodes network.

5.1.2. Zachary's karate club

Zachary's karate club network is extensively used as an exemplar for evaluating measures of social network structure[47]. The network describes a friendship network among 34 members of a karate club during 2 years at an American university in the early 1970s. Inspired by the cross-validation technique, 34 members with the highest potential in the club have been selected. The "key-player" in this club should be identified. Fig. 5 depicts the structure of selected members in zachary's karate club. The amount of communities can not be seen directly in the network. The centrality measures can be obtained from this network.

5.1.3. Bottlenose Dolphin network

Bottlenose dolphin network is a well-known network and it is often used for analysis of the social network. The data were obtained by Lusseau who had observed the interaction patterns among individuals over 7 years [48]. Fig. 6 shows the connected graph of the bottlenose dolphin network. We can see from the dolphin network that two connected communities are contained in the dolphin network. And the most influential node in each community is not straight-forward to see. The dolphin network is a typical denser network with more equal degree distributions. Greater divergence of results among single-criterion methods of identifying central nodes will be expected. Based on observations of Lusseau, node 37 is the most influence node because removing it decreases the relation between two group of dolphins noticeably and returning it recovers the previous closeness relation again.

5.2. Parameter setting of STA

The input parameters referring to weights of criteria could be selected randomly so that an initial population can be generated and the optimal weight for each criteria can be detected. Achieving the maximum correlation coefficient calculated by Q_q and Q_v is the optimization target. The optimal parameters setting of STA are enumerated in Table 3. The maximum number of iterations is abbreviated to MaxIter. SE is the abbreviation of search enforcement which represents the number of candidate solution in each iteration. Various termination can be applied in STA to terminate the process appropriately. In this study, if the change of correlation coefficient between two iterations has reached the specified accuracy ε, STA terminates and the best solution will be obtained [49,50]. The MaxIter and SE are set in order to keep the compared experiment with literature [13] fairly.

5.3. Result analysis

The results of the experiment with three typical networks are shown in this part. Fig. 7 shows the framework of this paper. The proposed method in this study can be understood distinctly. Table 1 is obtained by Abrar student network. And the parameter analysis of v in VIKOR-STA* method based on Abrar student network has been shown in Fig. 8. We can see that the correlation coefficient reaches the maximum value when v equals to 0.3. Tables 2 and 4 are generated over Zachary's Karate Club and Bottlenose Dolphin

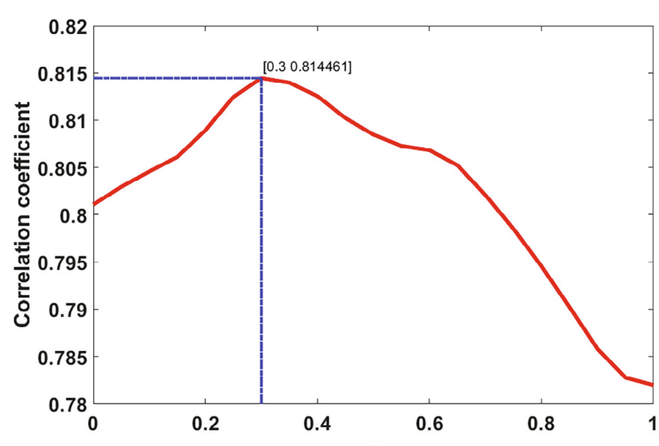


Fig. 8. Resulted correlation coefficient varying with v in VIKOR based on Abrar student network.

Table 1
Resulted best solutions on Abrar student network.

Result	Criteria	VIKOR-STA*	VIKOR-STA	TOPSIS-STA	TOPSIS-GA[13]	VIKOR	TOPSIS
Optimized Weights	$I\text{-}D_C$	0.312992	0.21517	0.27012	0.321927	0.1667	0.1667
	$O\text{-}D_C$	0.393278	0.27517	0.542484	0.412728	0.1667	0.1667
	E_C	0	0.09670	0	0.00952	0.1667	0.1667
	C_C	0.0341794	0.19981	0	0.003089	0.1667	0.1667
	B_C	0	0	0	0.012475	0.1667	0.1667
	PR	0.259546	0.21325	0.187396	0.240261	0.1667	0.1667
Fitness	ρ	0.812037	0.76519	0.65214	0.51956	0.4924	0.4663
Time	Second	1.263	1.48	1.155	9.8	1.2958	1.2957

Table 2
Resulted best solutions on Zachary's Karate Club.

Result	Criteria	VIKOR-STA*	VIKOR-STA	TOPSIS-STA	TOPSIS-GA	VIKOR	TOPSIS
	$I\text{-}D_C$	0	0	0.315234	0.25608	0.1667	0.1667
	$O\text{-}D_C$	0.203222	0.227715	0.362119	0.542007	0.1667	0.1667
Optimized	E_C	0.205778	0.227715	0	0	0.1667	0.1667
Weights	C_C	0.18497	0.104105	0.238071	0	0.1667	0.1667
	B_C	0.201265	0.214728	0	0	0.1667	0.1667
	PR	0.204765	0.225737	0.187396	0.298843	0.1667	0.1667
Fitness	ρ	0.881045	0.840191	0.770301	0.51956	0.4301	0.3913
Time	Second	1.856	2.375	1.723	5.633	1.2823	1.2873

Table 3
Parameters of STA.

Parameter	MaxIter	SE	α_{min}	α_{max}	β	γ	δ	ε
Value	200	20	0	1	1	1	1	10^{-6}

Table 4
Resulted best solutions on Bottlenose Dolphin Network.

Result	Criteria	VIKOR-STA*	VIKOR-STA	TOPSIS-STA	TOPSIS-GA	VIKOR	TOPSIS
	$I\text{-}D_C$	0.258448	0.122057	0.163614	0.168579	0.1667	0.1667
	$O\text{-}D_C$	0.188257	0.22788	0.471346	0.434769	0.1667	0.1667
Optimized	E_C	0	0	0	0	0.1667	0.1667
Weights	C_C	0.209808	0.166525	0	0.03913	0.1667	0.1667
	B_C	0.153121	0.194948	0.118564	0.208374	0.1667	0.1667
	PR	0.190366	0.2886	0.246476	0.149148	0.1667	0.1667
Fitness	ρ	0.83085	0.80125	0.7046495	0.700864	0.6344	0.5128
Time	Second	1.854	2.079	1.723	4.717	1.2859	1.2563

Network respectively. They show that VIKOR-STA* merging the hybrid normalization technique and STA with VIKOR has the highest correlation coefficient with acceptable time complexity. Results show that the proposed method has a better performance than existing ones. Not only does VIKOR method perform better than TOPSIS, but also STA shows better capability than GA in solving influence maximization problem. The VIKOR and TOPSIS methods with average weights are included. The compared results verify the effectiveness of adopting optimization algorithm to obtain optimal weights in VIKOR and TOPSIS method. From the compared experiments, we can see that the weights in VIKOR and TOPSIS method have a significant impact on the ranking results. Adopting STA to obtain the optimal weights in VIKOR method can exhibit a better performance. Also, merging the hybrid normalization technique can increase the correlation coefficient and reduce the time complexity, which indicates the effectiveness of taking the hybrid normalization technique. As we can see, the degree centrality and PageRank centrality are always more or equal important than other centrality measures from Tables 1, 2 and 4. The degree centrality contains in-degree centrality and out-degree centrality. The degree centrality and page-rank centrality represent nodes' connection in social network. The results indicate that the two centrality measures may be more important than others.

It is reasonable to come to a conclusion that VIKOR-STA* is effective because of its high correlation coefficient and low time complexity. There are two major reasons for this. First, the hybrid normalization technique can settle the problem of integrating unit inconsistency and magnitude mismatch data in VIKOR method. Second, STA can find the optimal weights in VIKOR rapidly and accurately. So, we can make a summary that the proposed method in this study can solve the influence maximization problem effectively and performs more excellent than existing ones.

6. Conclusion

Influence maximization problem is hard to be dealt with in social network because it is difficult to evaluate which person is the "key-player" or "most influential one" by centrality measures in the network. In this paper, we settle the influence maximization problem with MCDM method call VIKOR. Non-commensurable data exist in data mining commonly. The unit inconsistency and magnitude mismatch problem commonly exist when integrating different centrality measures in VIKOR. A hybrid normalization technique has been proposed to normalize the unit and magnitude of different centrality measures into the same quantity in the preprocessing. Results have shown that the hybrid normalization technique has a good performance in solving this problem. Getting the weights in VIKOR method is a significant work which can be considered as an optimization problem. STA is a novel intelligent optimization algorithm which can be used to solve this optimization problem. Experimental results over three typical networks have shown that STA outperforms GA in getting weights in VIKOR. And comparative analysis shows that VIKOR exhibits more excellent performance than TOPSIS in dealing with influence maximization problem. In brief, the technique and method this paper proposed are efficient in dealing with this influence maximization problem. According to these results, further study can be focused on developing a clean model with the least amount of centrality measures for the problem. Also, some machine learning techniques can be introduced to further improve the efficiency and effectiveness of the proposed method.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Xiaojun Zhou: Investigation, Conceptualization, Writing - review & editing. **Rundong Zhang:** Writing - review & editing, Software. **Ke Yang:** Writing - review & editing, Validation. **Chunhua Yang:** Writing - review & editing, Supervision. **Tingwen Huang:** Writing - review & editing, Supervision.

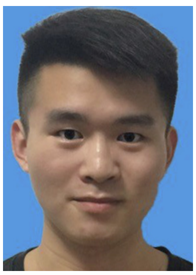
Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant Nos. 61873285, 61860206014), the 111 Project (Grant No. B17048), the Innovation-Driven Plan in Central South University (Grant No. 2018CX012) and Hunan Provincial Natural Science Foundation of China (Grant No. 2018JJ3683).

References

[1] Y. Li, J. Fan, Y. Wang, K. Tan, Influence maximization on social graphs: a survey, IEEE Trans. Knowl. Data Eng. 30 (10) (2018) 1852–1872.
[2] Z. Wang, C. Du, J. Fan, Y. Xing, Ranking influential nodes in social networks based on node position and neighborhood, Neurocomputing 260 (2017) 466–477.
[3] R. Alhajj, N. Memon, Introduction to the second issue of social network analysis and mining journal: scientific computing for social network analysis and dynamicity, Social Network Anal. Min. 1 (2) (2011) 73–74.
[4] K. Kandhway, J. Kuri, Using node centrality and optimal control to maximize information diffusion in social networks, IEEE Trans. Syst. Man Cybern. Syst. 47 (7) (2017) 1099–1110.
[5] S. Kundu, C. A. Murthy, S. K. Pal, A new centrality measure for influence maximization in social networks, in: Pattern Recognition and Machine Intelligence, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 242–247.
[6] X. Deng, Y. Dou, T. Lv, Q.V.H. Nguyen, A novel centrality cascading based edge parameter evaluation method for robust influence maximization, IEEE Access 5 (2017) 22119–22131.
[7] A. Landherr, B. Friedl, J. Heidemann, A critical review of centrality measures in social networks, Business Inf. Syst. Eng. 2 (6) (2010) 371–385.
[8] J. Yang, J. Liu, Influence maximization-cost minimization in social networks based on a multiobjective discrete particle swarm optimization algorithm, IEEE Access 6 (2018) 2320–2329.
[9] C.-W. Tsai, S.-J. Liu, Seim: search economics for influence maximization in online social networks, Fut. Gen. Comput. Syst. 93 (2019) 1055–1064.
[10] Y. Meng, Y. Yi, F. Xiong, C. Pei, Tone hop approach for dynamic influence maximization problem, Phys. A Stat. Mech. Appl. 515 (2019) 575–586.
[11] I. Mesgari, M.A.M.A. Kermani, R. Hanneman, A. Aliahmadi, Identifying key nodes in social networks using multi-criteria decision-making tools, Math. Technol. Networks (2015) 137–150.
[12] M. Jalayer, M. Azheian, M.A.M.A. Kermani, A hybrid algorithm based on community detection and multi attribute decision making for influence maximization, Comput. Ind. Eng. 120 (2018) 234–250.
[13] M.A.M.A. Kermani, A. Badiee, A. Aliahmadi, M. Ghazanfari, H. Kalantari, Introducing a procedure for developing a novel centrality measure (sociability centrality) for social networks using topsis method and genetic algorithm, Comput. Human Behav. 56 (2016) 295–305.
[14] S. Jendo, W.M. Paczkowski, Multicriteria discrete optimization of large scale truss systems, Struct. optim. 6 (4) (1993) 238–249.
[15] S. Opricovic, G.-H. Tzeng, Compromise solution by mcdm methods: a comparative analysis of vikor and topsis, Eur. J. Oper. Res. 156 (2) (2004) 445–455.
[16] M. Injadat, F. Salo, A.B. Nassif, Data mining techniques in social media: a survey, Neurocomputing 214 (2016) 654–670.
[17] E. Zavadskas, Z. Turskis, Multiple criteria decision making (mcdm) methods in economics: an overview, Technol. Econ. Dev. Econ. 17 (2011) 397–427.
[18] M. Tavana, R.K. Mavi, F.J. Santos-Arteaga, E.R. Doust, An extended vikor method using stochastic data and subjective judgments, Comput. Ind. Eng. 97 (2016) 240–247.
[19] F. Zhang, C. Yang, X. Zhou, W. Gui, Fractional-order pid controller tuning using continuous state transition algorithm, Neural Comput. Appl. 29 (10) (2018) 795–804.
[20] X. Zhou, C. Yang, W. Gui, A statistical study on parameter selection of operators in continuous state transition algorithm, IEEE Trans. Cybern. (2018) 1–9.
[21] J. Han, C. Yang, X. Zhou, W. Gui, A two-stage state transition algorithm for constrained engineering optimization problems, Int. J. Control Autom. Syst. 16 (2) (2018) 522–534.
[22] H. Wu, W. Liu, K. Yue, W. Huang, K. Yang, Maximizing the spread of competitive influence in a social network oriented to viral marketing, International Conference on Web-Age Information Management (2015) 516–519.
[23] M. Kimura, K. Saito, K. Ohara, H. Motoda, Efficient analysis of node influence based on sir model over huge complex networks, 2014 International Conference on Data Science and Advanced Analytics (DSAA), 2014, pp. 216–222..
[24] M. Li, J. Wang, H. Wang, Y. Pan, Essential proteins discovery from weighted protein interaction networks, Bioinf. Res. Appl. (2010) 89–100.
[25] M. Fraschini, A. Hillebrand, M. Demuru, L. Didaci, G.L. Marcialis, An eeg-based biometric system using eigenvector centrality in resting state brain networks, IEEE Signal Process. Lett. 22 (6) (2015) 666–670.
[26] C. Salavati, A. Abdollahpouri, Z. Manbari, Ranking nodes in complex networks based on local structure and improving closeness centrality, Neurocomputing (2018) 131–134.
[27] M.H. Chehreghani, An efficient algorithm for approximate betweenness centrality computation, Comput. J. 57 (9) (2014) 1371–1382.
[28] F. Lamberti, A. Sanna, C. Demartini, A relation-based page rank algorithm for semantic web search engines, IEEE Trans. Knowl. Data Eng. 21 (1) (2009) 123–136.
[29] H. Liao, Z. Xu, X. Zeng, Hesitant fuzzy linguistic vikor method and its application in qualitative multiple criteria decision making, IEEE Trans. Fuzzy Syst. 23 (5) (2015) 1343–1355.
[30] V.P. Muddineni, A.K. Bonala, S.R. Sandepudi, Enhanced weighting factor selection for predictive torque control of induction motor drive based on vikor method, IET Electr. Power Appl. 10 (9) (2016) 877–888.
[31] A. Shemshadi, H. Shirazi, M. Toreihi, M. Tarokh, A fuzzy vikor method for supplier selection based on entropy measure for objective weighting, Expert Syst. Appl. 38 (10) (2011) 12160–12167.
[32] B. Sennaroglu, G.V. Celebi, A military airport location selection by ahp integrated promethee and vikor methods, Transp. Res. D Transp. Environ. 59 (2018) 160–173.
[33] M. Behzadian, S.K. Otaghsara, M. Yazdani, J. Ignatius, A state-of the-art survey of topsis applications, Expert Syst. Appl. 39 (17) (2012) 13051–13069.
[34] X. Zhou, C. Yang, W. Gui, State transition algorithm, J. Ind. Manage. Optim. 8 (4) (2012) 1039–1056.
[35] X. Zhou, C. Yang, W. Gui, Nonlinear system identification and control using state transition algorithm, Appl. Math. Comput. 226 (2014) 169–179.
[36] Z. Huang, C. Yang, X. Zhou, W. Gui, A novel cognitively inspired state transition algorithm for solving the linear bi-level programming problem, Cogn. Comput. 10 (5) (2018) 816–826.
[37] M. Huang, X. Zhou, T. Huang, C. Yang, W. Gui, Dynamic optimization based on state transition algorithm for copper removal process, Neural Comput. Appl. 31 (7) (2019) 2827–2839.
[38] X. Zhou, R. Zhang, X. Wang, T. Huang, C. Yang, Kernel intuitionistic fuzzy c-means and state transition algorithm for clustering problem, Soft Comput. doi:10.1007/s00500-020-04879-8..
[39] X. Zhou, Q. Wang, R. Zhang, C. Yang, A hybrid feature selection method for production condition recognition in froth flotation with noisy labels, Miner. Eng. doi:10.1016/j.mineng.2020.106201..
[40] X. Zhou, J. Long, C. Xu, G. Jia, An external archive-based constrained state transition algorithm for optimal power dispatch, Complexity (2019), https://doi.org/10.1155/2019/4727168.
[41] X. Zhou, K. Yang, Y. Xie, C. Yang, T. Huang, A novel modularity-based discrete state transition algorithm for community detection in networks, Neurocomputing 334 (2019) 89–99.
[42] Q. Wang, M. Huang, X. Zhou, Feature selection in froth flotation for production condition recognition, IFAC-PapersOnLine 51 (21) (2018) 123–128.
[43] Z. Huang, C. Yang, X. Zhou, T. Huang, A hybrid feature selection method based on binary state transition algorithm and relieff, IEEE J. Biomed. Health Inf. 23 (5) (2018) 1888–1898.
[44] X. Zhou, J. Zhou, C. Yang, W. Gui, Set-point tracking and multi-objective optimization-based pid control for the goethite process, IEEE Access 6 (2018) 36683–36698.
[45] K. Lewis, J. Kaufman, M. Gonzalez, A. Wimmer, N. Christakis, Tastes, ties, and time: a new social network dataset using facebook.com, Social Networks 30 (4) (2008) 330–342.
[46] H.M. Inderbitzen, S.L. Foster, The teenage inventory of social skills: development, reliability, and validity, Psychol. Assess. 4 (4) (1992) 451–459.
[47] W.W. Zachary, An information flow model for conflict and fission in small groups, J. Anthropol. Res. 33 (4) (1977) 452–473.
[48] N.M.E.J. Lusseau D, Identifying the role that animals play in their social networks, Proc. Roy. Soc. B Biol. Sci. 271 (6) (2004) 477–481..
[49] F. Zhang, C. Yang, X. Zhou, H. Zhu, Fractional order fuzzy pid optimal control in copper removal process of zinc hydrometallurgy, Hydrometallurgy 178 (2018) 60–76.
[50] X. Zhou, P. Shi, C.-C. Lim, C. Yang, W. Gui, A dynamic state transition algorithm with application to sensor network localization, Neurocomputing 273 (2018) 237–250.

Xiaojun Zhou received his Bachelor's degree in Automation in 2009 from Central South University, Changsha, China and received the PhD degree in Applied Mathematics in 2014 from Federation University Australia. He is currently an Associate Professor at Central South University, Changsha, China. His main interests include modeling, optimization, and control of complex system, artificial intelligence and machine learning, optimization theory and algorithms, duality theory and global optimization with applications. E-mail: michael.x.zhou@csu.edu.cn.



Rundong Zhang received the B.S. degree in automation in 2017 from Wuhan University of Technology, Wuhan, China. He is currently pursuing the M.S. degree from Central South University, Changsha, China. His current research interests include multiple criteria decision making, state transition algorithm, and influence maximization. E-mail: zrd1018@csu.edu.cn.

Ke Yang received the B.S. degree in electrical engineering and automation from Hunan Institute of Engineering, Xiangtan, China, in 2017, where he is currently pursuing the M.S. degree from Central South University, Changsha, China. His current research interests include community detection, state transition algorithm, and data mining. E-mail: yang_ke@csu.edu.cn.

Chunhua Yang received her M.Eng. in Automatic Control Engineering and her Ph.D. in Control Science and Engineering from Central South University, China in 1988 and 2002 respectively, and was with the Electrical Engineering Department, Katholieke Universiteit Leuven, Belgium from 1999 to 2001. She is currently a full professor in the School of Information Science & Engineering, Central South University. Her research interests include modeling and optimal control of complex industrial process, intelligent control system, and fault-tolerant computing of real-time systems. E-mail: ychh@csu.edu.cn.

Tingwen Huang received the B.S. degree in mathematics from Southwest Normal University, Chongqing, China, in 1990, the M.S. degree in applied mathematics from Sichuan University, Chengdu, China, in 1993, and the Ph. D. degree in mathematics from Texas A & M University, College Station, TX, USA, in 2002. He was a Lecturer with Jiangsu University, Zhenjiang, China, from 1994 to 1998, and a Visiting Assistant Professor with Texas A & M University, in 2003. He was an Assistant Professor from 2003 to 2009, an Associate Professor from 2009 to 2013, and has been a Professor since 2013, with the Texas A & M University at Qatar, Doha, Qatar. His current research interests include neural networks, complex networks, chaos and dynamics of systems, and operator semigroups and their applications. Dr. Huang serves as an Associate Editor for the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS and the IEEE TRANSACTIONS ON CYBERNETICS. E-mail: Tingwen.huang@qutar.tamu.edu