Expert Systems With Applications 114 (2018) 224–236

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Using swarm intelligence algorithms to detect influential individuals for influence maximization in social networks

Aybike ŞİMŞEK[a,*], Resul KARA[b]

[a] Düzce University, Gölyaka Vocational School, Department of Computer Programming, Turkey
[b] Düzce University, Faculty of Engineering, Department of Computer Engineering, Turkey

ARTICLE INFO

Article history:
Received 11 September 2017
Revised 30 May 2018
Accepted 15 July 2018
Available online 20 July 2018

Keywords:
Social influence analysis
Influential individuals
Influence maximization
Social network
Swarm intelligence
Optimization
Large networks
Complex networks

ABSTRACT

People use online social networks to exchange information, spread ideas, learn about innovations, etc. Thus, it is important to know how information spreads through social networks. It is possible to spread information (e.g., product advertisement) to a larger number of individuals via a social network. Similarly, it is possible to minimize the spread of unwanted content (e.g., 'false news'). The key point in both cases is to identify the most influential individuals on the social network. This problem is named as Influence Maximization (IM) problem. The IM problem focuses on finding the small subset of individuals in a social environment who influence a certain group of individuals, i.e., maximize/minimize the spreading of information. Some greedy algorithms, stochastic algorithms and evolutionary optimization algorithms have been developed to find a solution to this problem. However, these methods are not at the desired level in terms of speed or solution capability. On the other hand, although many swarm intelligence algorithms that produce fast and optimal solutions can be found in the literature, these algorithms cannot be directly applied to the IM problem because no general slope is produced on the state-space surface of the IM problem's objective function. Swarm intelligence algorithms follow the general slope over the surface to converge at the global optimum. Thus, they cannot converge to the global optimum in the IM problem.

In this study, a change in the structure of the IM problem is suggested in order to tailor it to swarm intelligence algorithms and to achieve a general slope on the state-space surface of its objective function. We named this process as "reshaping". More precisely, if a social network is envisioned as a graph and individuals as nodes, reshaping means sorting the nodes in descending order (from largest to smallest) according to the metrics under consideration (i.e., metrics that give an idea about the level of influence of an individual) and renumbering the nodes according to this order. Thus, the nodes those are close to each other in terms of level of influence become closer to each other in the state-space. This creates a general slope on the state-space surface of the objective function. This simple idea paves the way for applying all swarm intelligence algorithms to this kind of problem. The proposed approach was tested with real and synthetic graphs. The experiments employed the Grey Wolf Optimizer (GWO) and Whale Optimization Algorithm (WOA) as the swarm intelligence algorithms and PageRank and Kempe et al.'s Greedy Algorithm as benchmark methods. Experimental results showed that this approach worked well.

1. Introduction

A real social network is of great benefit for information dissemination via communication among its members. The situation is the same for online social networks such as Facebook and Twitter, which transfer real social networks to the Internet. Information can be spread very quickly among users of these online platforms. This opens the door to applications like viral marketing. However, information dissemination on social networks can also be used for malicious purposes, including the spreading of 'false news' or virus-containing messages among people in a short amount of time. Influential individuals in a social network can be utilized to increase information dissemination or reduce the spread of unwanted content like gossip (Johansson, 2017).

Some individuals in a social network are more influential than others for such reasons as social status, charisma, etc. If these individuals can be identified, the spreading of information on a network can be maximized, or the source of diffused content identi-

* Corresponding author.
E-mail addresses: aybikesimsek@duzce.edu.tr (A. ŞİMŞEK), resulkara@duzce.edu.tr (R. KARA).

https://doi.org/10.1016/j.eswa.2018.07.038
0957-4174/

fied. In this respect, detection of influential individuals is the basis of both effector detection and influence maximization (IM) problems. In recent years, the IM problem has been widely studied. In contrast, the effector detection problem has received less attention. Essentially, the IM problem is a kind of effector detection problem (Tong, Li, Wu & Du, 2016). In IM, the goal is to determine which individuals will be selected as a seed set for influencing the largest number of individuals. On the other hand, the purpose of the problem of effector detection is to determine which individuals should be selected as a seed set to achieve a certain activation state (Tong et al., 2016). As a result, although there are small differences, both problems can be considered as optimization problems based on the detection of influential individuals. Kempe states that *"the influence maximization problem asks to find a k-node set of maximum influence"* (Kempe, Kleinberg & Tardos, 2003). Based on this definition, detection of influential individuals problem is referred to as IM in the remainder of the paper.

The IM is an NP-hard combinatorial optimization problem (Lappas, Terzi, Gunopulos & Mannila, 2010). Therefore, in order to solve this problem, various metrics have been used with greedy algorithms. In addition, stochastic and evolutionary methods have been applied to the IM problem (Zhang, Du & Feldman, 2017). Although, there are many swarm intelligence algorithms mentioned in the literature such as Artificial Bee Colony (ABC), Gray Wolf Optimizer (GWO), Whale Optimization Algorithm (WOA) and Ant Lion Optimizer (ALO) (Karaboga, Akay & Ozturk; Mirjalili, 2015; Mirjalili & Lewis, 2016; Mirjalili, Mirjalili & Lewis, 2014), their application to this problem is a novel approach (Gong, Yan, Shen, Ma & Cai, 2016; Wang, Gong, Song & Wang, 2017). The underlying reason is that swarm intelligence algorithms cannot be directly applied to combinatorial optimization problems. In problems such as IM, the state-space surface produced by the objective function does not have a general slope. Therefore, swarm intelligence algorithms cannot converge to an optimal solution at the global level. In this context, applying swarm intelligence algorithms to combinatorial problems such as IM would make it possible to take advantage of these algorithms. To this purpose, in this study, the IM problem was reshaped to enable swarm intelligence algorithms to be applied to these types of problems. Reshaping means sorting the nodes in descending order (from largest to smallest) according to the metrics under consideration and renumbering the nodes according to this order. By reshaping the social network graph in this way, a general slope is achieved on the state-space surface, and swarm intelligence algorithms become available for the IM problem.

In this study, first, the GWO and random selection methods were compared on the reshaped and non-reshaped graphs to show how the approach works. After that, the GWO, WOA and PageRank algorithms were compared. Since the original GWO and WOA essentially produce solutions for continuous optimization problems, small changes were made on these algorithms to make them applicable for discrete (also combinatorial) optimization problems. The effectiveness of this approach for the IM problem was demonstrated on a Facebook graph (McAuley & Leskovec, 2012).

Last but not least, the IM is not just a problem in social networks. Similar problems in very different types of networks can be handled as an IM. For example, detection of the accounts which should be immunized in a network of e-mail contacts in order to minimize the spreading of computer viruses is essentially an IM problem. In addition, detection of the patient who is the origin of the spread of an infection in a contact network of inpatients is also an IM problem. From this point of view, it is important that swarm intelligence algorithms have the capacity to solve IM problem.

In summary, the contributions made by this study include:

- paving the way for the application of swarm intelligence algorithms to the IM problem,
- demonstrating that two swarm intelligence algorithms (GWO and WOA) could be successfully applied to the IM problem,
- showing that centrality metrics can be combined and used successfully.

The remainder of this paper is organized as follows. Section 2 reviews the related works. In Section 3, the IM problem and Independent Cascade (IC) information propagation model are defined. The proposed approach is provided in Section 4. In Section 5, the experimental findings are presented. In Section 6, the statistical analysis is given and the results discussed. Section 7 concludes the paper.

2. Related works

Influence maximization is a very important topic in social network research. The IM problem is based on the detection of the most influential individuals. To this end, studies in the literature have focused on the development of metrics pointing to influential individuals and the development of algorithms using these metrics (Peng et al., 2018; Tong et al., 2016). First, the metrics used in the IM problem should be discussed. Some of the basic evolution metrics for social influence include Degree, Closeness, Eigenvector, Katz, and Betweenness Centralities (Peng et al., 2018).

Degree Centrality is calculated as the number of edges which are adjacent to a node (Borgatti, 2005; Peng et al., 2018). Directional graphs have two types of degrees: the in-degree and the out-degree. The in-degree is the number of incoming edges to a node from other nodes. The out-degree is the number of outgoing edges from a node to other nodes. The higher the degree of a node, the more it is connected to other nodes, meaning, in general, that this node can influence more nodes.

Closeness Centrality is calculated as the average shortest path from all other nodes to a node (Bergamini, Borassi, Crescenzi, Marino & Meyerhenke, 2016; Borgatti, Carley & Krackhardt, 2006). Higher values of Closeness Centrality indicate that a node has higher influence.

Betweenness Centrality is calculated as percentage of the shortest paths between any two nodes passing through a certain node (Peng et al., 2018; Rabade, Mishra & Sharma, 2014). Assuming that the information is spread via the shortest paths, it can be said that nodes with higher values of Betweenness serve as bridges for information propagation.

Katz Centrality can be thought of as the sum of the degrees of neighbors of a particular node (Peng et al., 2018; Rabade et al., 2014). The idea behind the Katz Centrality is that if a node has influential neighbors, this node is also influential (Temizsoy, Iori & Montes-Rojas, 2017).

Eigenvector Centrality is similar to Katz Centrality with respect to the basic idea and calculation (Peng et al., 2018). If the graph is strongly connected, it works well. However, actual directed networks have no large connected component and this leads to difficulties in practice (Temizsoy et al., 2017).

Apart from the metrics described above, social networks also have specific metrics. For example, to identify the importance of a Twitter user, features like Verified Account, Subscription Lists Containing User, and Age of The Profile can be used (Cossu, Labatut & Dugué, 2016). Since graph-based metrics were used in this study, user-based (nodal) metrics were not included. For more detailed information about nodal social influence metrics, the studies of Peng et al. (2018) and Cossu et al. (2016) can be examined.

However, the IM problem cannot be solved by simply selecting the individuals with the highest values of a certain metric (Borgatti, 2006). For this reason, several approaches have been developed to solve the IM problem.

One of the earliest studies on the IM problem in social networks is that of Kempe et al. (2003) and Samadi, Nagi, Semenov and Nikolaev (2018). They demonstrated that the IM problem is an NP-hard, discrete optimization problem and proposed a general greedy approach. Their approach adopts a one-by-one seeding strategy. That is, it evaluates the seed candidates one at a time and adds them to the seed set. They compared their greedy approach with Degree and Distance Centrality heuristics and found that their algorithms were more successful.

Another earlier study was conducted by Borgatti (2006). He described the IM problem as a problem of identifying the key players and evaluated the problem from two perspectives: (1) the determination of key players to ensure maximum spread of information and (2) the removal of key players from the network to divide the network into the largest number of sub-networks. For both problems, simply choosing the top *k* players did not work well. Thus, Borgatti dealt with simultaneously selecting the *k*-seed set problem (as done in this study), which is a combinatorial optimization problem. Kempe's one-by-one seeding greedy approach and Borgatti's combinatorial optimization approach have largely shaped subsequent studies. Some of the studies in the literature have adopted the greedy approach, while others have treated the problem as a combinatorial optimization problem.

Some state-of-the-art studies should also be examined. Raghavan and Zhang approached the IM problem from a different perspective and addressed it as a Weighted Target Set Selection (WTSS), which is a combinatorial optimization problem (Raghavan & Zhang). They developed a polynomial-time dynamic programming algorithm for a WTSS problem on trees. They then generalized their approach for arbitrary graphs. In a similar study, Günneç, Raghavan, and Zhang discussed the Least Cost Influence Problem (LCIP) (Günneç, Raghavan & Zhang). Their main motivation was to develop mathematical programming approaches for these kinds of problems. They developed a greedy algorithm for LCIP on trees. In addition, they developed a better dynamic programming (DP) algorithm that divides a tree into several star-graphs.

Lappas et al. introduced a problem slightly different from the IM problem and called it the *k*-Effector problem (Lappas et al., 2010). In the *k*-Effector problem, a set of nodes had been already activated by some seed nodes and it aimed to find the effector nodes that best explained the observed activation state. They showed that for arbitrary graphs, the problem was NP-hard to solve optimally and NP-hard to approximate. Their approach consisted of two stages: (1) constructing an influence tree for a given network and activation state and (2) using a dynamic-programming algorithm to select the optimal effectors on the tree. They demonstrated that their algorithms performed well on the *k*-Effector problem.

Based on the *k*-Effector problem of Lappas et al. (2010) and Tong et al. (2016) addressed the problem from two perspectives. First, they considered the distance of influence and designed an effector detection framework based on the influence distance metric. They then used the Maximum Likelihood Estimation (MLE) approach for effector detection. They formulated the effector detection problem as an optimization problem and proposed a 3-approximation algorithm.

Essentially, social networks have dynamic structure and this structure evolves over time. On this basis, Song, Li, Chen, He and Tang (2017) argued that the influencer nodes should be tracked in a dynamic network structure, and they proposed the Influential Node Tracking (INT) problem as an extension of the traditional IM problem. While IM is used to maximize common influence in a static network, INT aims to protect a set of influencer nodes having influence at the maximum level as the network continues to evolve. They developed a greedy algorithm called the Upper Bound Interchange Greedy (UBI) and its variant, UBI + for solving INT problems. Instead of constructing a seed set from the beginning, UBI uses the influencer seeds which had previously been found and changes nodes to increase the total influence. They emphasized that when working on three real-scale dynamic social networks, they achieved better results in terms of coverage and performance.

Most of the studies in the literature are based on IC and LT diffusion models, which are time-independent, i.e., in these models, the timing of the seed selection is not important. Samadi et al. noted that many practical problems are not time-independent, and they treated the selection of the seed nodes as a time-dependent problem (Samadi et al., 2018; Samadi, Nikolaev & Nagi, 2017). The introduced problem is called the Seed Activation Scheduling Problem (SASP) and it chooses the timing of seed activation under a given budget. They proved that SASP is NP-hard under the Partial Parallel Cascade (PPC) diffusion model, and developed a mixed-integer program to solve the SASP under the PPC diffusion model. With these features, the SASP has given a new dimension to the IM problem and it is a problem that needs to be studied in detail.

Liu, Cong, Zeng, Xu and Chee (2014) pointed out the importance of time in spreading the influence from one user to another, and they addressed the time-limited influence maximization problem. In order to increase the scalability of the greedy algorithm they developed, they proposed the concept of the propagation path in social networks and developed a series of new algorithms for the problem of time-limited IM.

Some studies have approached the IM problem in the context of game theory. Two of the basic studies in this area are works of Irfan and Ortiz (2014) and Molinero, Riquelme and Serna (2015). Irfan and Ortiz have proposed network influence games as a game-theoretic model of the behavior of a networked population. Also, they addressed the identification of the most influential individuals, and they developed a new approach with using pure-strategy Nash-equilibria computation. Their study is significant because it connects linear-influence games to important models in game theory. Molinero et al. have considered IM problem as a system in which the individuals are keen about perform cooperative tasks but their motivations depend on fulfillment of the tasks by other influential individuals. They have modeled this scenario as an influence game as well. Their study is also important because of it characterizes the computational complexity of several problems on influence games, which includes metrics and properties. Another game theoretic study is Ok et al.'s study (Ok, Jin, Shin & Yi, 2016). Ok et al. dealt with the Influence Diffusion Speed Maximization problem, which can be considered as part of the IM problem. They used game theory in their work and proposed a practical seeding algorithm called Practical Partitioning and Seeding (PrPaS). The authors concluded that PrPAS was superior to other algorithms based on the propagation speed over real social network topology.

Unlike other greedy approaches, an algorithm using the divide and conquer approach was developed by Song, Zhou, Wang and Xie (2015). They pointed out that finding the most effective users of social networks is an NP-hard problem, and that greedy algorithms have high computation costs. Consequently, they used the strategy of divide and conquer with parallel computing for large-scale mobile social networks. Their algorithm, the community-based greedy algorithm, involves two components: (1) dividing the large mobile social network into several communities according to information spreading and (2) selecting the communities via dynamic programming to find the most influential nodes.

Many approaches in the literature treat all nodes in a network as seed candidates. However, the number of influencers is very small compared to the total number of nodes. According to Lee and Chung (2015), the weakest point of IM is that certain users cannot be distinguished from other users. It was emphasized that it can be a good strategy to focus on maximizing the influence on certain users. In order to distinguish certain users from others, they considered the problem of IM as query processing. According to experimental results, the method they proposed was highly accurate and faster than the existing methods.

Most of the studies in the literature deal with only one social network. However, since users share information on many social networks at the same time, it is very important to conduct research on multiple networks. Zhang et al. pointed out the Least Cost Influence (LCI) problem on multiple social networks (Zhang, Nguyen, Zhang & Thai, 2016). The authors mapped a set of networks into a single network using loss and lossless mapping schemes. Working with multiple networks makes the IM problem difficult because the levels of influence of one individual in different networks may be different. Therefore, considering the total effect of an individual may involve a choice. In this respect, the IM problem on multiple networks is an interesting and difficult problem.

The IC and LT models are often mentioned in the literature. However, these models are simple models of influence diffusion in the real world. In this regard, Tong et al. have criticized diffusion models for lack of realism (Tong, Wu, Tang & Du, 2017). They stated that real-world social networks involve a great deal of uncertainness. They used the dynamic IC model and proposed the concept of adaptive seeding. They proposed a simple greedy adaptive seeding strategy and developed an efficient heuristic algorithm for better scaling. As a result of their experiments on real and synthetic datasets, they stated that the strategy they presented was superior to the basic methods.

In another study, the IM problem was considered as a structural controllability problem and a weighted maximum matching algorithm was proposed for solving the problem (Sartor, Chia, Wynter & Ruths, 2017). In their work, Sartor et al. used Kempe's algorithm as a benchmark and showed that their proposed method yielded competitive results.

Peng, Yang, Cao, Yu and Xie (2017) presented a framework to quantify social influence in mobile social networks. They stated that most evaluation models are focused on online social networks, but fail to characterize indirect influence. Subsequently, they designed the Susceptible-Infectious (SI) epidemic model to characterize the propagation dynamics of social influence. They developed an evaluation model to measure direct, indirect, and total influence using friendship entropy and interaction frequency entropy. Their algorithm sorts the nodes in descending order according to the total influence and picks the top-k node in a greedy manner. They reported that the proposed method outperformed the random and degree-based methods.

Social influence is divided into two groups: positive and negative. In this case, the social influence is considered as signed. Li et al. dealt with the IM problem on signed social networks which involve negative and positive relationships among individuals (Li et al., 2014). Their aim was to find a set of seed nodes that would provide a maximum-positive or a maximum-negative influence. They used a greedy approach for seed node selection and the polarity-related IC (IC-P) as the diffusion model. In a similar study, simulated annealing was used for seed node selection, and positive IM was investigated. The authors used Epinions, Slashdot and Wikipedia datasets from SNAP. Their method exhibited similar or better results than greedy algorithms in terms of positive influence (Li et al., 2017).

Gong et al. suggested using an evolutionary algorithm for the IM problem (Gong, Song, Duan, Ma & Shen, 2016). They proposed a memetic algorithm for maximizing community-based influence on social networks. They optimized 2-hop influence propagation to find the most influential nodes and stated that experiments on three real-life datasets (Chen, Wang & Yang, 2009; Leskovec, Kleinberg & Faloutsos, 2007; Lusseau et al., 2003) yielded better results in terms of effectiveness and efficiency.

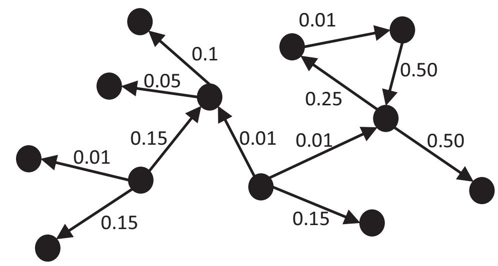

Fig. 1. Sample social graph.

Lastly, Gong et al. developed a method based on particle swarm optimization for IM in social networks (Gong et al., 2016). They redefined many aspects (update rules of particles, local search strategy, etc.) of the algorithm for the problem. The authors reported that their algorithm was effective and displayed a good performance.

As seen in the literature review, attempts to solve the IM problem have been made by using different algorithmic approaches and different metrics. In addition, the IM problem and its sub-problems have been examined for several types of networks and several different information diffusion models. Overall, the detection of the most effective k individuals in a social network can be seen as NP-hard in most cases. The existing greedy algorithms are inefficient in large networks. Moreover, greedy algorithms require heavy Monte-Carlo simulations of propagation functions (Zhou, Zhang, Zang & Guo, 2015). On the other hand, optimization algorithms generally give better results. The disadvantages are that they work more slowly. Swarm intelligence algorithms are not commonly applied on IM problems because of the previously mentioned reasons. Using swarm intelligence algorithms, which provide faster optimal results, will fill this gap in the literature.

Table 1 presents a summary of the studies that have been reviewed. Greedy approaches were adopted in most studies. Optimization algorithms were used less frequently, while the IC model was the most commonly used propagation model.

3. Problem definition

The Independent Cascade (IC) model was used to model the influence propagation. In the IC model, an individual can be found only in one of the active or inactive states. If an individual is influenced by another individual, it becomes active. An individual who becomes active can influence other individuals and cannot be passive again. Individuals who are selected as seeds are also considered as active and influenced. In the IC model, there is an influence probability of between 0 and 1, $P(e) = [0, 1]$ which is associated with an edge, $e = (u, v)$. Here, u and v are arbitrary neighbor nodes. A social network is defined as a directed graph, G = (V, E), $V = \{i_1 \cdots i_n\}$ is the set of individuals, $E = \{e_1 \cdots e_m\}$ is the set of relations between individuals, n is the number of individuals in the graph, and m is the total number of relations in the graph. A social graph example is shown in Fig. 1.

For the problem of IM, the goal is to select k seed individuals that will influence the largest number of individuals. Let $S = \{s_1 \cdots s_k\}$ be a set of seeds and $A = \{a_1 \cdots a_m\}$ be the set of individuals influenced by the k seed individuals, and also here, $S \subseteq A$. As a result, the problem can simply be defined as follows: Which k individuals should be selected as seed nodes to maximize the objective function in Eq. (1)? This problem is a combinatorial optimization problem and clearly, it cannot be solved by brute force

Table 1
Summary of the reviewed studies.

Study	Algorithm	Propagation Model
(Tong et al., 2016)	Greedy	IC
(Song et al., 2017)	Greedy	IC
(Ok et al., 2016)	Greedy	Independent Poisson Clock
(Lappas et al., 2010)	Dynamic Programming	IC
(Nguyen, & Zheng, 2013)	Greedy	IC
(Liu et al., 2014)	Greedy	Latency Aware IC
(Song et al., 2015)	Greedy	IC
(Lee & Chung, 2015)	Greedy	IC
(Zhang et al., 2016)	Greedy	LT and IC
(Tong et al., 2017)	Greedy and Heuristic	IC
(Gong et al., 2016)	Particle Swarm	IC
(Chen et al., 2009)	Greedy	IC
(Zhang et al., 2017)	Genetic Algorithm	LT
(Li et al., 2014)	Greedy	Polarity-Related IC
(Li et al., 2017)	Simulated Annealing	Polarity-Related IC
(Peng et al., 2017)	Greedy	SI
(Samadi et al., 2018)	Mixed Integer Programming	Partial Parallel Cascade (PPC)

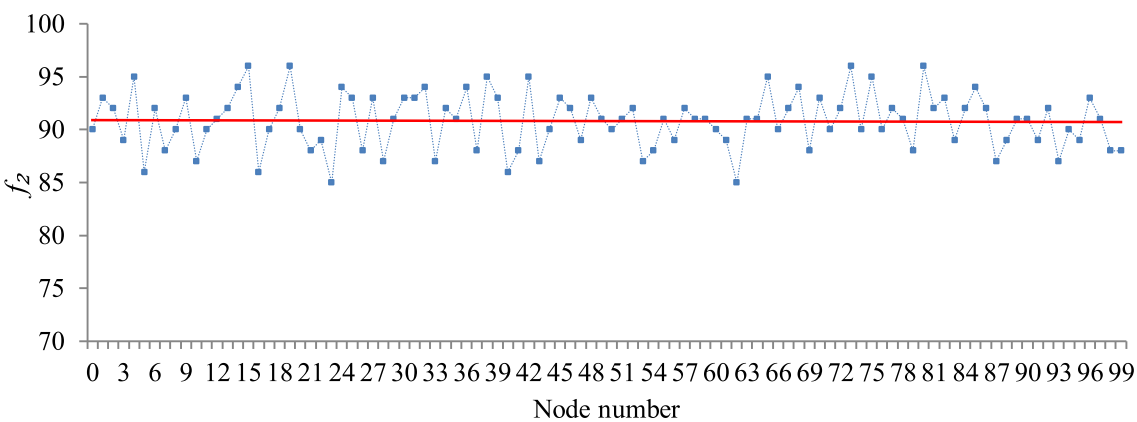

Fig. 2. Sample IM problem on Erdös–Renyi Graph – Regular Graph.

in reasonable time for large n and m values.

$$f_1(S) = |A \backslash S| \tag{1}$$

By expressing the problem as a minimization problem, the objective function can be written as

$$f_2(S) = |V| - |A \backslash S| \tag{2}$$

The purpose of Eq. (2) is to minimize the number of uninfluenced nodes in the network.

4. Tailoring the problem to swarm intelligence algorithms

4.1. Current state of the problem

First, let's examine the IM problem on a small graph and for $k = 1$. In Fig. 2, an IM problem with $k = 1$ is modeled for an Erdös-Renyi graph of 100 nodes. The horizontal axis indicates the nodes. The vertical axis represents the f_2 values obtained by selecting the corresponding node as a seed. The IC was used in the modeling. The influence probability of individuals with each other (namely, influence propagation probability, the weights of the edges in the graph) was determined as $P(e) = 0.01$ according to the uniform setting. This setting can be chosen if not enough is known about the given social network (Tong et al., 2016). The uniform setting can be used because the Erdös–Renyi graph is a randomly generated graph. As mentioned above, the IM problem is a discrete and combinatorial optimization problem. The values of the objective function in Fig. 3 are discrete values. However, even if the structure of these values is considered as a surface, there is no slope on the surface. For example, while node 63 is a good choice, node 62 can be a very bad choice. In general, a swarm intelligence algorithm tries to maximize or minimize an objective function for the different values of the variables that make up the state-space in the problem. In doing so, it assumes that the next good solution is near to the current solution. In the current problem, a swarm intelligence algorithm cannot reach the global minimum (or at least become closer) by following neighboring states. As a result, for such a problem, the jumpings of swarm intelligence algorithms such as ABC, GWO in the state-space will not go beyond being random if there are no intelligent interventions in the algorithm generation processes. In this case, swarm intelligence algorithms cannot converge globally and walk on local optimums.

4.2. Saving swarm intelligence algorithms from local optimums

Swarm intelligence algorithms can be successfully applied to multi-parameter optimization problems. However, swarm intelligence algorithms have some drawbacks including premature convergence and poor local search ability (Monson & Seppi, 2006). In recent years, hybrid approaches have been proposed to overcome these drawbacks of swarm intelligence algorithms. Some of these approaches suggest using local search methods such as Nelder-Mead and BOBYQA with swarm intelligence algorithms (Cheng, Zhan & Shu, 2016; Zahara & Kao, 2009). The Nelder-Mead method is good for the local search. However, its convergence is extremely sensitive to the selected starting point. Swarm intelligence algorithms are good for a global search. However, they also require a great deal of computation (Fan & Zahara, 2007). Similarly, BOBYQA

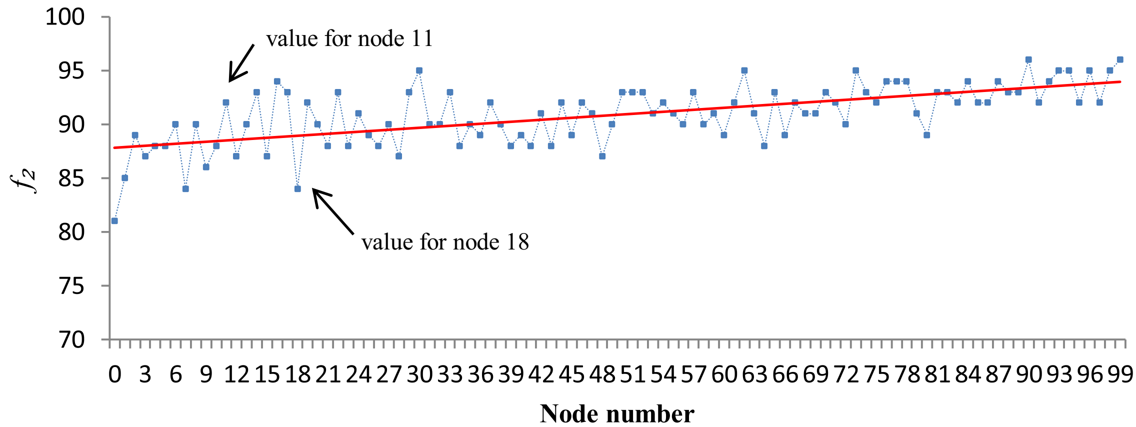

Fig. 3. Influence maximization on reordered Erdös–Renyi Graph – Outdegree Centrality.

is often used for a local search. The main objective of these hybrid approaches is to achieve global level convergence using the swarm intelligence algorithm and to conduct a local search by using methods such as Nelder–Mead and BOBYQA. However, these studies do not improve the global convergence of the swarm intelligence algorithms for the IM problem. Swarm intelligence algorithms require a general slope on the state-space surface in order to use their global search capabilities.

4.3. Proposed approach: giving a general slope to the state-space surface of the objective function in the IM problem

As mentioned previously, swarm intelligence algorithms follow the slope of the surface to move from the current state to the next good state. These algorithms assume that the next good solution is around the current state of the state-space. However, the situation is different in IM problems. What is then needed is to bring the solutions which are similar in quality closer on the surface. In the IM problem, nodes create the state-space surface. Therefore, if the nodes are sorted according to their impacts and renumbered according to this order, a slope is obtained on the state-space surface.

There are many metrics frequently used in IM problems which provide information about the level of influence of a node. Thus, the approach involves sorting the nodes in descending order according to the metrics in consideration and renumbering them according to this order.

Two of these metrics are Outdegree Centrality and Closeness Centrality. Outdegree Centrality is the number of edges from a node in a directional graph. The Outdegree Centrality C_{outdeg} for node u_j in a directed graph is $C_{outdeg}(u_j) = d_j^{out}$. Closeness Centrality is defined as the smaller average shortest paths between the node and all other nodes in a graph. Formally, $C_c(u_j) = \frac{1}{\bar{l}_{u_j}}$, where $\bar{l}_{u_j} = \frac{1}{n-1} \sum_{u_j \neq u_m} l_{j,m}$, l is the shortest path between u_j and u_m (Zafarani, Abbasi & Liu, 2014).

A greater $\mathrm{C_{outdeg}}$ or $\mathrm{C_c}$ value is an indicator that a node is more influential than the others. In this way, if the nodes are renumbered in the Erdös-Renyi graph according to $\mathrm{C_{outdeg}}$ in descending order and repeat the modeling, the result is shown in Fig. 3. Similarly, if the nodes in the graph are renumbered according to $\mathrm{C_c}$ in descending order, the result is Fig. 4. The structures of the graphs obtained in this way are exactly the same as the original graph only the nodes have been renumbered. Now, swarm intelligence algorithms can move on these new graphs. It is useful to note here that the selection of the k seed is not a simple process, such as selecting the first k seed in the reshaped graph. It cannot be said that the first 10 individuals in Fig. 4 are the 10 most influential individuals as seen in Fig. 3. There is, of course, a relationship between the calculated metric values for a node and the effect of this node, but this relationship is not linear. For example, in Fig. 3, $\mathrm{C_{outdeg}}(11) > \mathrm{C_{outdeg}}(18)$ (node 11 has a greater outdegree value than node 18), despite the fact that $f_2(11) > f_2(18)$. It should be noted that this example is a minimization example, so lower f_2 values are better. In addition, the number of individuals influenced by a seed in case $k=1$ may not be the same in case $k > 1$. For more detailed information, the study of (Kempe et al., 2003) should be examined.

4.3.1. Combining centrality metrics

The metrics may be used alone or in combination. Therefore, the nodes can be sorted using more than one metric at a time. For this it is necessary to determine the importance level of each metric. Let $\mathrm{M}=\{\mu_0 \cdots, \mu_\mathrm{N}\}$ be the metrics set, so the determined weight for the metrics $\mathrm{W}=\{\mathrm{w}_0 \cdots, \mathrm{w}_\mathrm{N}\}$ is a vector of $(0, 1)^N$. Thus, a combined metric is formed:

$$\mu_{\mathrm{combined}} = \mathrm{W} \cdot \mathrm{M}. \tag{3}$$

If the degree to which a metric is more important than another metric is not known, the same weight value can be used for all metrics. For example, let two metrics be $\mu_0 = \mathrm{C_{outdeg}}$ and $\mu_1 = \mathrm{C_c}$ and the weights specified for these metrics be $\mathrm{W}=\{0.5, 0.5\}$. The result is $\mu_{\mathrm{combined1}} = 0.5 \times \mathrm{C_{outdeg}} + 0.5 \times \mathrm{C_c}$. When the nodes are renumbered in the Erdös Renyi graph according to the $\mu_{\mathrm{combined1}}$ metric, the result is Fig. 5. Similarly, a general slope was obtained here.

4.3.2. Handling the IM problem for k = 2

To see the problem on a 3D surface, the problem for $k=2$ can be reflected on the same Erdös-Renyi graph. The original graph is considered as the regular graph, and the graphic obtained for it is shown in Fig. 6-a, the graphic obtained for the reshaped graph according to the $\mathrm{C_{outdeg}}$ metric is shown in Fig. 6-b, and the graphic obtained for the reshaped graph according to $\mathrm{C_c}$ is shown in Fig. 6-c. In Fig. 6, *blue* shows the trenches (minimums) and *brown* shows the hills (maximums). The x and y axes are node numbers. The z axis shows f_2 values. Although the node numbers appear to be shown as 0–100, the range is actually 0–99. As can be seen in the figure, the minimums in the regular graph are scattered randomly, while the minimums in the reshaped graphs are often gathered in certain regions. This shows that a slope was obtained on the state-space surface of the objective function. Now, a swarm intelligence algorithm can move on this surface.

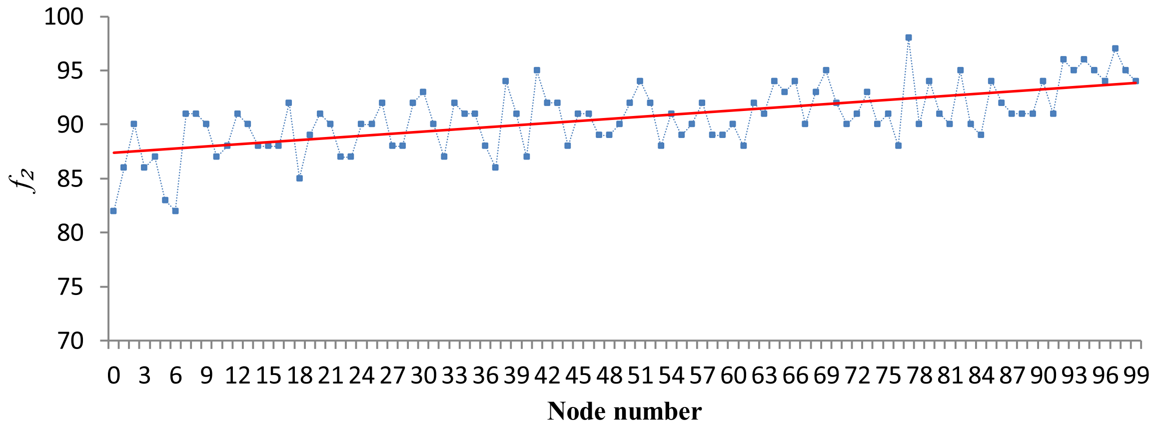

Fig. 4. Influence maximization on reordered Erdös–Renyi Graph – Closeness Centrality.

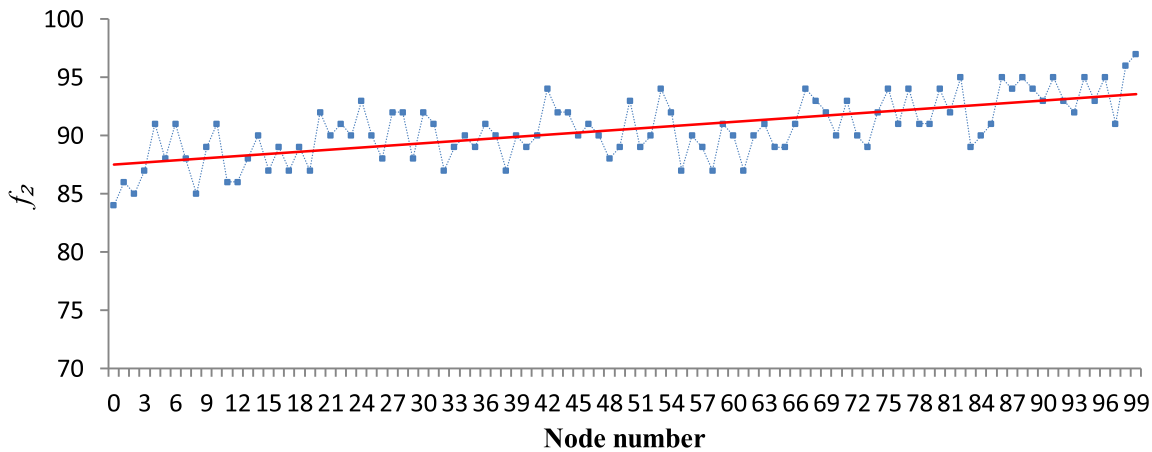

Fig. 5. Influence maximization on reordered Erdös–Renyi Graph – Combined Metric.

5. Experiments

The Facebook network was used in the experiments and was downloaded from the Standard Large Network Dataset Collection (http://snap.stanford.edu) (Leskovec & Krevl, 2014; McAuley & Leskovec, 2012). The network has 4039 nodes and 88,234 undirected edges. First, each undirected edge was replaced with two directed edges, making a total of 176,468 directed edges. A snapshot of the Facebook network is shown in Fig. 7.

There are two popular approaches for modeling influence propagation: uniform setting (US) and weighted cascade setting (WCS). As already explained, $P(e) = 0.01$ is used in the US. In WCS, the probability of influencing node v by node u is $P((u, v)) = 1/deg(v)$. Here, $deg(v)$ is the degree of node v. The WCS was chosen because the experiments were conducted on a real graph in order to obtain more realistic results (Kempe et al., 2003; Tong et al., 2016).

The experiments were carried out in two stages. In the first stage, the GWO was compared with the random selection method. Thus, it was demonstrated that the reshaping worked well. In the second stage, the GWO, WOA, PageRank and Kempe et al.'s Greedy Algorithm were compared.

5.1. Modified GWO, modified WOA, pagerank kempe et al.'s greedy algorithm

5.1.1. Modified GWO

The GWO is a meta-heuristic inspired by the hunting behavior of grey wolves (Mirjalili et al., 2014). Grey wolves have a strict hierarchy. There are four types of wolves in the hierarchy: alpha, beta, delta, and omega. The alpha wolves make decisions about hunting. However, alpha wolves follow the other wolves in some cases. Each candidate solution is a prey. The population of the alpha, beta, delta and omega agents search the solution space. In each iteration, the wolves encircle the prey. Finally, the wolves attack the prey and the best candidate is considered as the solution. For further information, the study of (Mirjalili et al., 2014) can be examined.

The original GWO was designed for continuous optimization problems. That is, the candidate solutions from the state-space are encoded as real variables. However, the state-space in the IM consists of integers. The items that the algorithm should choose are the nodes numbered as integers. Thus, the GWO was changed to produce integer solutions. To this end, integers were used in the place of the real number vectors and scalar variables in Eqs. (3.1), (3.2), (3.3) and (3.4) of Mirjalili et al. (2014). Thus, the algorithm produced integer solutions at every step. Beyond these, no further changes were made in the GWO.

5.1.2. Modified WOA

The WOA is a meta-heuristic inspired by the hunting behaviors of humpback whales (Mirjalili & Lewis, 2016). Humpback whales have a special hunting method. They create air bubbles in a helix shape around the prey and swim up toward the surface and then they force their prey to the surface. The WOA imitates this behavior of humpback whales. Like the GWO, the WOA is also designed for continuous optimization problems. Similarly, the real num-

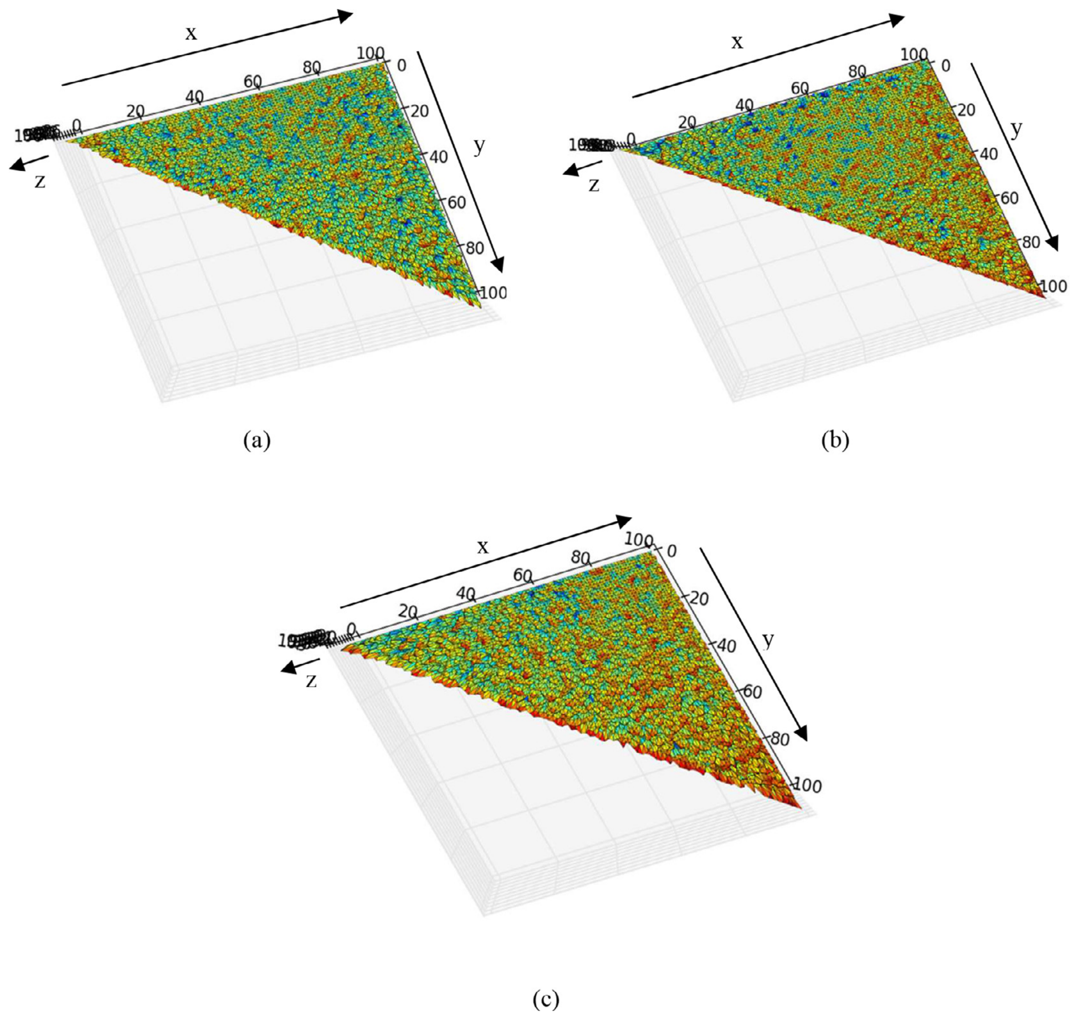

Fig. 6. Influence maximization on the Erdös–Renyi Graph: (a) regular graph, (b) reshaped graph according to Outdegree Centrality, and (c) reshaped graph according to Closeness Centrality.

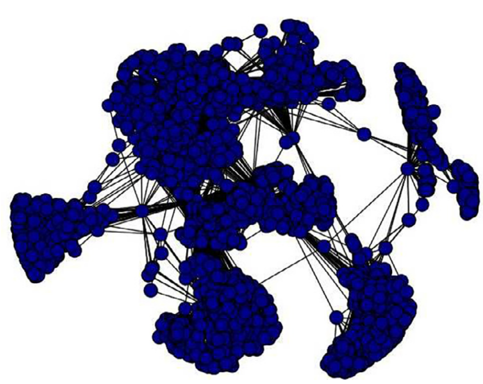

Fig. 7. Snapshot of the Facebook network.

ber vectors and scalar variables in Eqs. (2.1)–(2.8) in Mirjalili and Lewis (2016) were replaced with integers. Beyond these, no further changes were made in the WOA.

5.1.3. PageRank

PageRank is a graph-based approach for finding important nodes in a network (Page, Brin, Motwani & Winograd, 1999). It was originally designed to rank web pages for improving search results. In addition, PageRank is currently used for the detection of influential individuals on social networks (Zengin Alp & Gündüz Öğüdücü, 2018). Therefore, PageRank was chosen as a basic method for benchmarking. Since PageRank implementation can be easily found for many environments (Hagberg, Schult & Swart, 2008; Staudt, Sazonovs & Meyerhenke, 2016), details are not provided here.

5.1.4. Kempe et al.'s greedy algorithm

As we mentioned in Related Works, one of the earliest studies on the IM problem in social networks is that of Kempe et al. (2003). This algorithm is very important in the literature for influence maximization and all new algorithms. Thus, we took this greedy algorithm as the baseline for comparison. Kempe et al.'s greedy algorithm can be described as follows: Select a previously unselected node that increases the spread by the most. Do this, until the desired number of seeds to be selected. For implementation details (Kempe et al., 2003) can be examined.

5.2. GWO and random selection experiments

The GWO Python code used was downloaded from http://www.alimirjalili.com/GWO.html. The abovementioned changes were implemented on these codes and the $f_2(S)$ function in Eq. (2) was used as an objective function.

Experiments were carried out to see how the reshaping worked. Seed selection was performed both on the original graph and on the modified graphs. In order to determine whether this approach had a significant effect on the reshaped graphs, random seed selection was used as a benchmark method on both the original graph and the modified graphs.[1] In each experiment, the GWO algorithm and the random selection process were run 30 times. There were 50 iterations in each run. Eventually, the best results for the GWO and random selection were found.

The methods were applied for $k=1$, $k=2$, $k=5$, $k=10$, $k=15$, $k=20$, and $k=30$. The results obtained for each experiment are shown in Fig. 8, respectively. The vertical axis in the figures is $f_2(S)$ in Eq. (2). The horizontal axis labels Regular, OutDegree, Closeness and Combined denote “regular (original) graph”, “reshaped graph according to C_{outdeg}”, “reshaped graph according to C_c” and “reshaped graph according to $\mu_{combined1}$” respectively.

The main motivation for reshaping the problem was to obtain a slope on the state-space of the problem, so that the swarm intelligence algorithms can move on this surface. When the graphs were reshaped, the solutions that were close to each other in quality became closer to each other in the state-space as well. There was no such relationship between adjacent solutions in the original case of the problem. Moreover, the agents in a swarm intelligence algorithm look for the next best solution around the current best solution. Consequently, the expectation was that the agents of the swarm intelligence algorithm could navigate through neighbor solutions to arrive at an optimal solution. In addition, the basic assumption was that if the reshaping worked well, the swarm intelligence algorithms should give better results on the reshaped graphs than on the regular graph.

According to the experimental results, there was no significant difference between the GWO and the random selection for $k=1$ and $k=2$. In such a small graph, the problem was relatively simple for a low seed count. In the other experiments, the difference between the GWO and the random selection performance was clear. Especially when the situations for $k=15$, $k=20$, $k=30$ are examined (Fig. 8e–g), it can be seen that the reshaping process worked well. In Fig. 8e–g, on the regular graph, the GWO and the random selection method show very close results. On the other hand, the performance of the GWO was significantly better on the reshaped graphs based on the C_{outdeg}, C_c, and $\mu_{combined1}$ metrics. As the number of seeds to be selected increased, the problem became more difficult. For more difficult situations, the significant improvement of the GWO performance showed that the reshaping process worked well. As a result, the GWO algorithm achieved good results on the reshaped problem, but the behavior on the original (non-reshaped) problem did not differ from the random selection.

5.3. GWO, WOA, pagerank and Kempe et al.’s greedy algorithm experiments

In the previous section, it was shown that the reshaping worked well. This section gives the performance evaluation of the GWO, WOA, PageRank and Kempe et al.’s Greedy Algorithm on the IM problem. For the sake of brevity, we named Kempe et al.’s Greedy Algorithm as *Kempe’s*. Similar to GWO, the WOA Python code used was downloaded from http://www.alimirjalili.com/WOA.html. In addition, the $f_2(S)$ function in Eq. (2) was used as an objective function.

Experiments were carried out for $k=15$, $k=20$, and $k=30$ because the performance of the proposed approach had been determined for $k \geq 15$. The results of each experiment are shown in Fig. 9, respectively. The vertical axis in the figures is $f_2(S)$ in Eq. (2). In the experiments, the reshaped graph was used only according to C_c for the evaluation. The GWO and WOA outperformed the PageRank algorithm for all k values, and GWO and WOA gave slightly better results than Kempe et al.’s Greedy Algorithm for $k=15$ and $k=20$. Kempe et al.’s Greedy Algorithm gave better results for $k=30$. The performances of the GWO and WOA were close. At this point, we need to specify that GWO and WOA have given competitive result with Kempe et al.’s Greedy Algorithm. Swarm intelligence algorithms are promising for use in the IM problem.

Table 2
Test statistics for GWO.

	GWO2 - GWO1
Z	−2.366
Asymp. Sig. (2-tailed)	.018

6. Discussion

It can be seen from Fig. 8 that the proposed approach allows swarm intelligence algorithms to be used in the IM problem and that successful results can be obtained. At the same time, some tests were carried out to show that the results obtained were statistically significant.[2] If the proposed approach is successful, then the GWO must give a significantly better performance on the reshaped graphs than on the regular graph. Let $P^{GWO}_{reshaped}$ be the performance of the GWO on the reshaped graphs and $P^{GWO}_{regular}$ be the performance of the GWO on the regular graph. Now, the hypothesis is set and the level of significance can be determined.

$$H_0 : \left|P^{GWO}_{reshaped} - P^{GWO}_{regular}\right| \text{iszeroversus}$$

$$H_1 : \left|P^{GWO}_{reshaped} - P^{GWO}_{regular}\right| \text{ispositive}, \ \alpha = 0.05$$

The nonparametric Wilcoxon Signed Ranks Test was performed in order to observe the differences between the pre-test (before reshaping the problem) and the post-test (after reshaping the problem) application scores with a sample size of < 30. For the different numbers of seeds, $k=1$, $k=2$, $k=5$, $k=10$, $k=15$, $k=20$, and $k=30$, making up the total number of seven samples. The test statistics obtained for the GWO are shown in Table 2. In the table, GWO2 denotes the performance of the GWO on the reshaped graphs and GWO1 denotes the performance of the GWO on the regular graph. According to Wilcoxon Signed Ranks test analysis results, for 10, 15, 20 and 30 seed numbers, the GWO performance on the reshaped graphs was significantly better than its performance on the regular graph ($p=0.018$, $p < \alpha$), and the decision was to reject H_0.

On the other hand, the random selection method’s performance on the reshaped graphs and the regular graph should be nearly the same. Let $P^R_{reshaped}$ be the performance of the random selection method on the reshaped graph and $P^R_{regular}$ be the performance of the random selection method on the regular graph. Now, the hypothesis is set and the level of significance can be determined.

$$H_0 : |P^R_{reshaped} - P^R_{regular}| \text{is zero versus}$$

$$H_1 : |P^R_{reshaped} - P^R_{regular}| \text{is positive}, \ a = 0.05$$

The test statistics obtained for the random selection are shown in Table 3. In the table, Random2 represents the performance of the random selection on the reshaped graphs and Random1 denotes the performance of the random selection on the regular graph. No statistically significant difference was observed for any

[1] The main motivation of this study is to show that swarm intelligence algorithms can be used to detect influential individuals through the reshaping process.

[2] SPSS21 was used for statistical analysis.

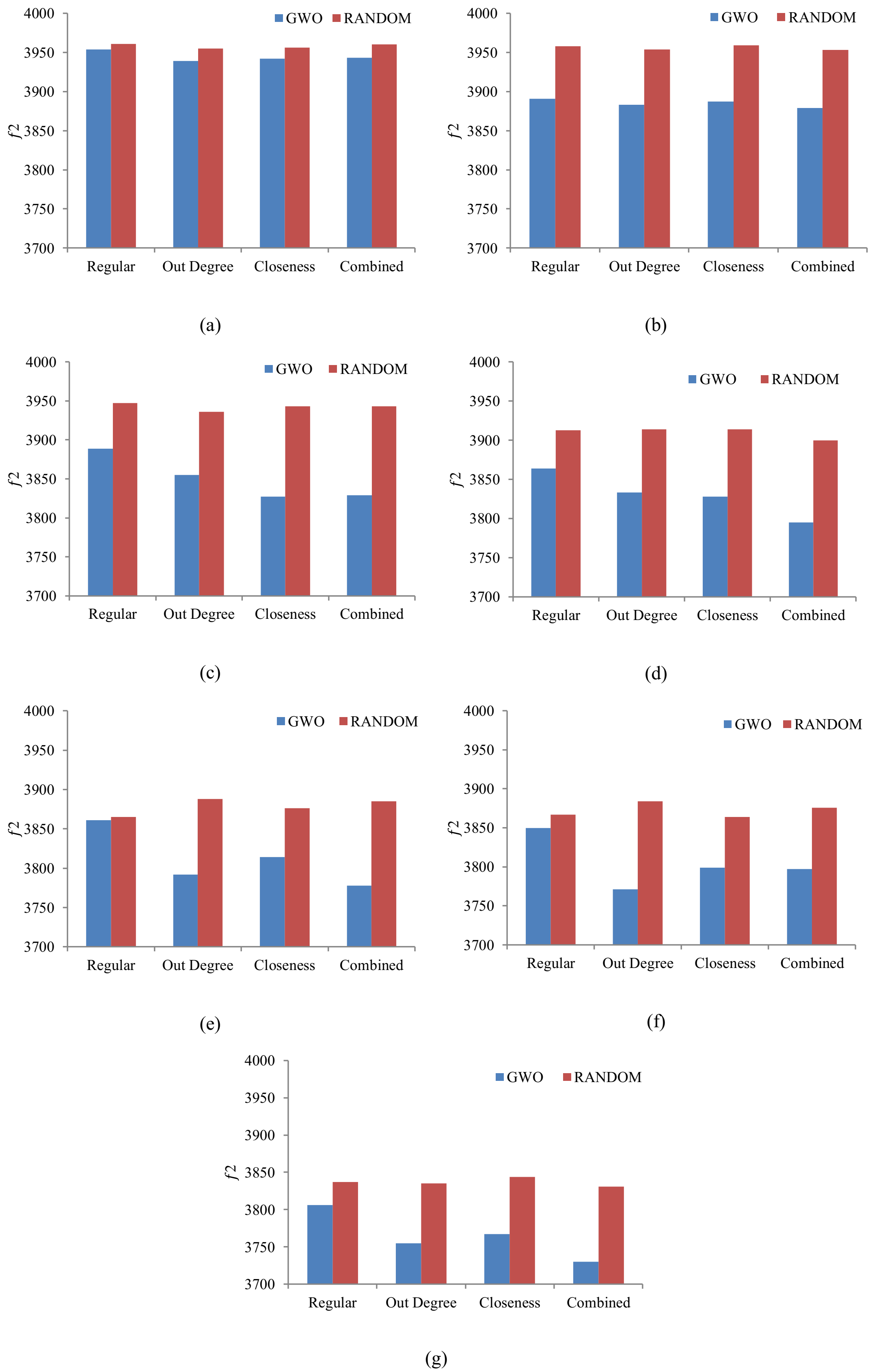

Fig. 8. Experimental results of GWO and random selection methods on the Facebook graphs for different numbers of seeds: (a) $k=1$, (b) $k=2$, (c) $k=5$, (d) $k=10$, (e) $k=15$, (f) $k=20$, and (g) $k=30$.

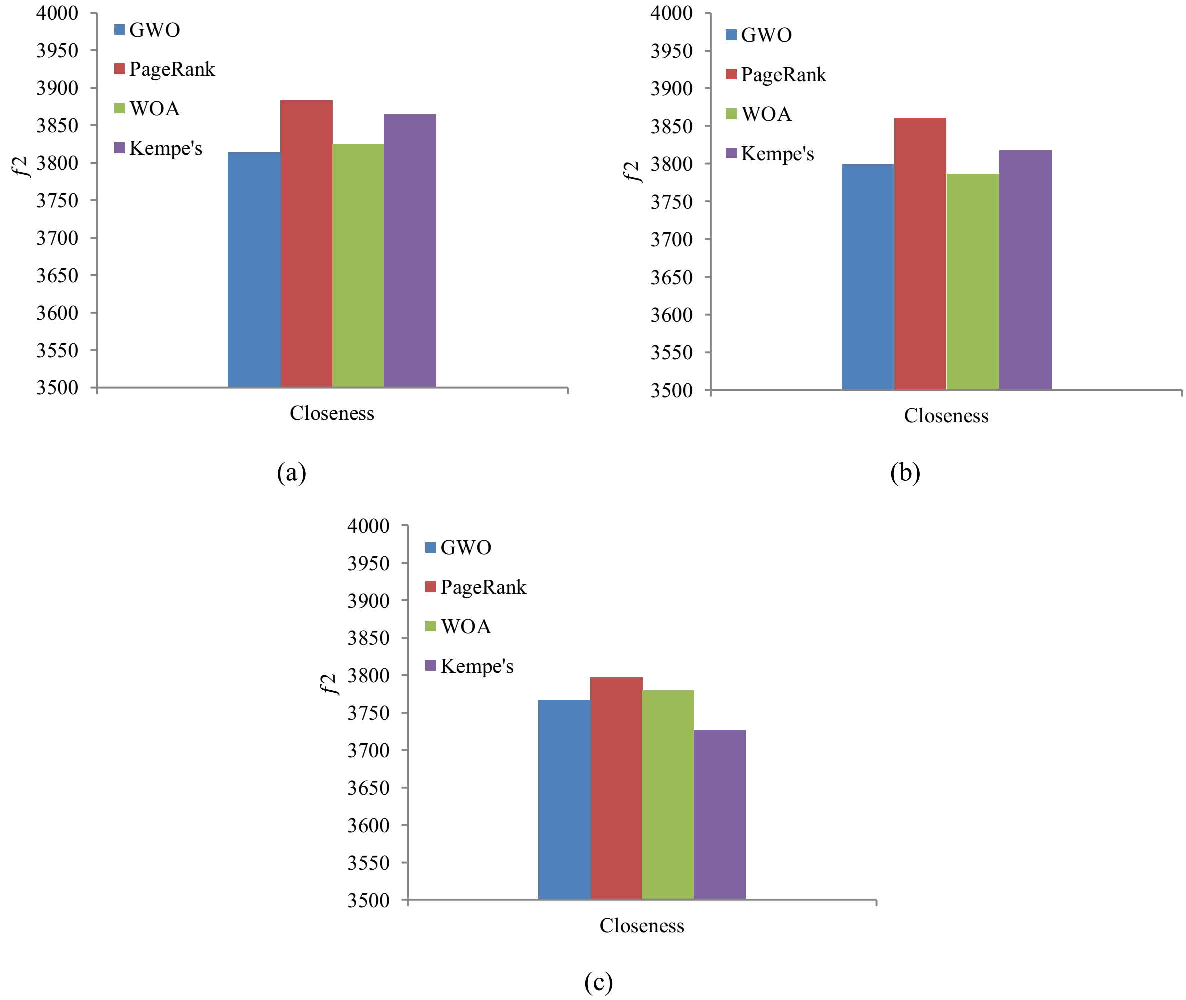

Fig. 9. Experimental results of GWO, WOA, PageRank and Kempe et al.'s Greedy Algorithm methods on the Facebook graph for different numbers of seeds: (a) $k=15$, (b) $k=20$, and (c) $k=30$.

Table 3
Test statistics for random selection.

	Random2 - Random1
Z	−.339
Asymp. Sig. (2-tailed)	.735

number of seeds between the performance of the random selection method on the reshaped graphs and its performance on the regular graph ($p=0.735$, $p>0.05$), and the decision was to accept H_0

Algorithmic complexity of the proposed approach

The first step in the re-shaping process was to calculate the metrics. The algorithmic complexity of the re-shaping process is highly dependent on the complexity of the metric calculations. For example, the algorithmic complexity of C_{outdeg} metric is $O(|V|^2)$ (Lappas et al., 2010). To calculate C_C, first, Johnson's algorithm was employed for the all-pairs shortest path problem. The algorithmic complexity of Johnson's algorithm is $O(|V|^2\log|V|+|V||E|)$. After getting the all-pairs shortest path, a nested loop was needed for calculating the average of the shortest paths for all nodes (See C_C calculation in Section 4). This adds the $O(|V|^2)$ complexity to the calculation. Finally, the algorithmic complexity of the C_C metric was obtained as $O(|V|^2\log|V|+|V|^2+|V||E|)$. The next step in the reshaping process was sorting the nodes according to the metric/metrics. If an efficient sorting algorithm like Quick Sort is chosen, the algorithmic complexity of the sorting step will be $O(|V|\log|V|)$.

Finally, the strengths and weaknesses of this approach are listed as follows:

Strengths

- The method developed the detection of the influential individuals problem via reshaping. There was no need to modify the algorithms. Thus, all swarm intelligence algorithms can be applied to the problem. In the literature, the problem has not yet been solved by reshaping.
- It is simple to implement. Reshaping the problem is a simple process involving the reordering of the nodes.

Weaknesses

- The selection of the metrics that are used for reshaping is important. This directly affects the performance of swarm intelligence algorithms. In fact, other methods in the literature have used these metrics. The selection of metrics is also important for other approaches.
- Swarm intelligence algorithms often become trapped in local optima. Minor modifications of the algorithms may be required to prevent this.

7. Conclusion

In the literature, application of swarm intelligence algorithms to the IM problem is extremely limited. Most swarm intelligence algorithms cannot be directly applied to combinatorial problems. To enable the usage of swarm intelligence algorithms for these kinds of problems, either the problem or the algorithm must be modified according to the goal.

The main contribution of this study is that it paves the way for applying all swarm intelligence algorithms to IM problems, which are combinatorial optimization problems. This study shows that swarm intelligence algorithms can be applied to the IM problem by tailoring the social network graph without changing its structure. The experimental results and the statistical analyses support this approach. Although the GWO and WOA were chosen for the evaluation, different algorithms can be applied to the IM problem.

In addition, there are some issues on this topic which are open for future research:

7.1. Determining the metrics and coefficients

It is important to decide the metrics and coefficients for a more healthy reshaping process. This study employed Outdegree Centrality and Closeness Centrality and a combined metric created as an indicator of the activity level of an individual. Moreover, there were some other metrics included such as Eigenvector Centrality and Katz Centrality. One topic of research is the degree to which a metric indicates an individual's influence level for different situations. For example, each metric can be examined separately for networks where different relationships such as a friendship (two-way relationship), a following (one-way relationship) or a trust (one-way relationship with signs) are modeled. Thus, different metrics may be preferred for different types of social networks. Another issue is the determination of the coefficients that form the combined metric. In order to form a composite metric in Eq. (3), the same coefficient was simply used for two metrics. However, a combined metric can be generated using more metrics. For this, it is necessary to determine which coefficient to choose for which metric in different situations (size of social network, type of social relationship, the number of seeds to select, the propagation model used, etc.). An expert or intelligent system can be designed to solve this problem.

7.2. Selecting a swarm intelligence algorithm for the problem

This study showed that swarm intelligence algorithms can be applied to the IM problem by tailoring the social network graph. For this purpose, the GWO and WOA were used as swarm intelligence algorithms. However, there are many swarm intelligence algorithms in the literature. An important point here is the way different algorithms will yield results for networks where different relationships such as a friendship, a following, or a trust are modeled. Different algorithms may be more appropriate for different social relationships.

7.3. Modifying an algorithm for the problem

The basic approach in this study was to adapt the problem to the algorithms; however, the algorithms may need to be modified as well. Each swarm intelligence algorithm has its own strategy for moving over the state-space surface of the problem and has solutions to get rid of local minimums or local maximums. For example, in the GWO algorithm, wolves update their positions according to the position of the best search agents. Mathematical models of the encircling and hunting behavior of the wolves are formulated by Eqs. (3.1)–(3.7) in the study by (Mirjalili et al., 2014). In these equations, linearly decreasing vectors and randomly generated scalar quantities were used. In this way, the wolves pass from one position to another on a searching space (looking for another candidate solution) and look for prey. Performance can be improved by developing a heuristic that manages the location change process. Such interventions can also be applied to other swarm intelligence algorithms.

7.4. Reshaping the problem for different network types

As mentioned in the Introduction, the IM problem is not a problem only of social networks. Similar problems in very different types of networks can be handled as IM problems. For example, detection of the accounts which should be immunized in a network of email contacts in order to minimize the spread of computer viruses is essentially an IM problem. Similarly, detection of the patient who is the origin of the spread of an infection in a contact network of inpatients is an IM problem. Because different network types have their own features, the reshaping process in these networks may vary. After reshaping is completed, swarm intelligence algorithms can be applied to IM problems in such networks.

7.5. Handling the IM problem under different information propagation models and continuous-time diffusion

There are more realistic and complex information propagation models than the IC model. Therefore, modeling of information propagation demands higher CPU power. In addition, the IC and LT models operate in discrete time. However, information does not spread in discrete time in the real world. For this reason, continuous-time propagation should be used for more realistic models (Samadi et al., 2018). This also requires higher CPU power. In this context, it is necessary to reconsider the IM problem for different continuous-time propagation models and to reshape the problem according to the new situation and to test swarm intelligence algorithms on these models as well.

Acknowledgments

We thank anonymous referees for their careful reading and their useful comments and suggestions that helped us to improve the contents and the presentation of the paper.

References

Bergamini, E., Borassi, M., Crescenzi, P., Marino, A., & Meyerhenke, H. (2016). Computing Top- k closeness centrality faster in unweighted graphs. In *2016 proceedings of the eighteenth workshop on algorithm engineering and experiments (ALENEX)* (pp. 68–80). PA: Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9781611974317.6.

Borgatti, S. P. (2005). Centrality and network flow. *Social Networks, 27*(1), 55–71. https://doi.org/10.1016/j.socnet.2004.11.008.

Borgatti, S. P. (2006). Identifying sets of key players in a social network. *Computational and Mathematical Organization Theory, 12*(1), 21–34. https://doi.org/10.1007/s10588-006-7084-x.

Borgatti, S. P., Carley, K. M., & Krackhardt, D. (2006). On the robustness of centrality measures under conditions of imperfect data. *Social Networks, 28*(2), 124–136. https://doi.org/10.1016/j.socnet.2005.05.001.

Chen, W., Wang, Y., & Yang, S. (2009). Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD international conference on knowledge discovery and data mining - KDD '09* (p. 199). ACM Press. https://doi.org/10.1145/1557019.1557047.

Cheng, S., Zhan, H., & Shu, Z. (2016). An innovative hybrid multi-objective particle swarm optimization with or without constraints handling. *Applied Soft Computing, 47*, 370–388. https://doi.org/10.1016/j.asoc.2016.06.012.

Cossu, J.-V., Labatut, V., & Dugué, N. (2016). A review of features for the discrimination of twitter users: Application to the prediction of offline influence. *Social Network Analysis and Mining, 6*(1), 25. https://doi.org/10.1007/s13278-016-0329-x.

Fan, S.-K. S., & Zahara, E. (2007). A hybrid simplex search and particle swarm optimization for unconstrained optimization. *European Journal of Operational Research, 181*(2), 527–548. https://doi.org/10.1016/j.ejor.2006.06.034.

Gong, M., Song, C., Duan, C., Ma, L., & Shen, B. (2016a). An efficient memetic algorithm for influence maximization in social networks. *IEEE Computational Intelligence Magazine, 11*(3), 22–33. https://doi.org/10.1109/MCI.2016.2572538.

Gong, M., Yan, J., Shen, B., Ma, L., & Cai, Q. (2016b). Influence maximization in social networks based on discrete particle swarm optimization. *Information Sciences, 367*(368), 600–614. https://doi.org/10.1016/j.ins.2016.07.012.

Günneç, D., Raghavan, S., & Zhang, R. (2016). Tailored incentives and least cost influence maximization on social networks. *The Robert H. Smith school of business and institute for systems research*. Maryland, USA, Tech. Rep.: University of Maryland.

Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring network structure, dynamics, and function using NetworkX. In *7th python in science conference (SciPy2008)* (pp. 11–15).

Irfan, M. T., & Ortiz, L. E. (2014). On influence, stable behavior, and the most influential individuals in networks: A game-theoretic approach. *Artificial Intelligence, 215*, 79–119. https://doi.org/10.1016/j.artint.2014.06.004.

Johansson, T. (2017). Gossip spread in social network Models. *Physica A: Statistical Mechanics and Its Applications, 471*, 126–134. https://doi.org/10.1016/j.physa.2016.11.132.

Karaboga, D., Akay, B., & Ozturk, C. (2007). Artificial Bee Colony (ABC) optimization algorithm for training feed-forward neural networks. *Modeling decisions for artificial intelligence* (pp. 318–329). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-73729-2_30

Kempe, D., Kleinberg, J., & Tardos, É. (2003). Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '03* (p. 137). ACM Press. https://doi.org/10.1145/956750.956769.

Lappas, T., Terzi, E., Gunopulos, D., & Mannila, H. (2010). Finding effectors in social networks. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '10* (p. 1059). ACM Press. https://doi.org/10.1145/1835804.1835937.

Lee, J.-R., & Chung, C.-W. (2015). A query approach for influence maximization on specific users in social networks. *IEEE Transactions on Knowledge and Data Engineering, 27*(2), 340–353. https://doi.org/10.1109/TKDE.2014.2330833.

Leskovec, J., Kleinberg, J., & Faloutsos, C. (2007). Graph evolution: Densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data, 1*(1), 1–41. https://doi.org/10.1145/1217299.1217301.

Leskovec, J., & Krevl, A. (2014, Jun 1). [Online]. Available: https://snap.stanford.edu/data

Li, D., Wang, C., Zhang, S., Zhou, G., Chu, D., & Wu, C. (2017). Positive influence maximization in signed social networks based on simulated annealing. *Neurocomputing, 260*, 69–78. https://doi.org/10.1016/j.neucom.2017.03.003.

Li, D., Xu, Z.-M., Chakraborty, N., Gupta, A., Sycara, K., & Li, S. (2014). Polarity related influence maximization in signed social networks. *PLoS ONE, 9*(7), e102199. https://doi.org/10.1371/journal.pone.0102199.

Liu, B., Cong, G., Zeng, Y., Xu, D., & Chee, Y. M. (2014). Influence spreading path and its application to the time constrained social influence maximization problem and beyond. *IEEE Transactions on Knowledge and Data Engineering, 26*(8), 1904–1917. https://doi.org/10.1109/TKDE.2013.106.

Lusseau, D., Schneider, K., Boisseau, O. J., Haase, P., Slooten, E., & Dawson, S. M. (2003). The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. *Behavioral Ecology and Sociobiology, 54*(4), 396–405. https://doi.org/10.1007/s00265-003-0651-y.

McAuley, J., & Leskovec, J. (2012). Learning to discover social circles in ego networks. In *NIPS'12 proceedings of the 25th international conference on neural information processing systems* (pp. 539–547).

Mirjalili, S. (2015). The ant lion optimizer. *Advances in Engineering Software, 83*, 80–98. https://doi.org/10.1016/j.advengsoft.2015.01.010.

Mirjalili, S., & Lewis, A. (2016). The whale optimization algorithm. *Advances in Engineering Software, 95*, 51–67. https://doi.org/10.1016/j.advengsoft.2016.01.008.

Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey wolf optimizer. *Advances in Engineering Software, 69*, 46–61. https://doi.org/10.1016/j.advengsoft.2013.12.007.

Molinero, X., Riquelme, F., & Serna, M. (2015). Cooperation through social influence. *European Journal of Operational Research, 242*(3), 960–974. https://doi.org/10.1016/j.ejor.2014.11.006.

Monson, C. K., & Seppi, K. D. (2006). Adaptive diversity in PSO. In *Proceedings of the 8th annual conference on Genetic and evolutionary computation - GECCO '06* (p. 59). ACM Press. https://doi.org/10.1145/1143997.1144006.

Nguyen, H., & Zheng, R. (2013). On budgeted influence maximization in social networks. *IEEE Journal on Selected Areas in Communications, 31*(6), 1084–1094. https://doi.org/10.1109/JSAC.2013.130610.

Ok, J., Jin, Y., Shin, J., & Yi, Y. (2016). On maximizing diffusion speed over social networks with strategic Users. *IEEE/ACM Transactions on Networking, 24*(6), 3798–3811. https://doi.org/10.1109/TNET.2016.2556719.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). *The PageRank citation ranking: bringing order to the web*. San Francisco, USA, Tech. Rep.: InfoLab, Stanford University.

Peng, S., Yang, A., Cao, L., Yu, S., & Xie, D. (2017). Social influence modeling using information theory in mobile social networks. *Information Sciences, 379*, 146–159. https://doi.org/10.1016/j.ins.2016.08.023.

Peng, S., Zhou, Y., Cao, L., Yu, S., Niu, J., & Jia, W. (2018). Influence analysis in social networks: A survey. *Journal of Network and Computer Applications, 106*, 17–32. https://doi.org/10.1016/j.jnca.2018.01.005.

Wang, Qixiang, Gong, M., Song, Chao, & Wang, Shanfeng (2017). Discrete particle swarm optimization based influence maximization in complex networks. In *2017 IEEE congress on evolutionary computation (CEC)* (pp. 488–494). IEEE. https://doi.org/10.1109/CEC.2017.7969351.

Rabade, R., Mishra, N., & Sharma, S.. *Survey of influential user identification techniques in online social networks* https://doi.org/10.1007/978-3-319-01778-5_37.

Raghavan, S., & Zhang, R. (2015). Weighted target set selection on social networks. *The Robert H. smith school of business and institute for systems research*. Maryland, USA, Tech. Rep.: University of Maryland.

Samadi, M., Nagi, R., Semenov, A., & Nikolaev, A. (2018). Seed activation scheduling for influence maximization in social networks. *Omega, 77*, 96–114. https://doi.org/10.1016/j.omega.2017.06.002.

Samadi, M., Nikolaev, A., & Nagi, R. (2017). The temporal aspects of the evidence-based influence maximization on social networks. *Optimization Methods and Software, 32*(2), 290–311. https://doi.org/10.1080/10556788.2016.1214957.

Sartor, G., Chia, Y. K., Wynter, L., & Ruths, J. (2017). A weighted maximum matching algorithm for influence maximization and structural controllability 1 1supported by the international design centre (Grant IDG31300103). *IFAC-PapersOnLine, 50*(1), 14447–14453. https://doi.org/10.1016/j.ifacol.2017.08.2288.

Song, G., Li, Y., Chen, X., He, X., & Tang, J. (2017). Influential node tracking on dynamic social network: An interchange greedy approach. *IEEE Transactions on Knowledge and Data Engineering, 29*(2), 359–372. https://doi.org/10.1109/TKDE.2016.2620141.

Song, G., Zhou, X., Wang, Y., & Xie, K. (2015). Influence maximization on large-scale mobile social network: A divide-and-conquer method. *IEEE Transactions on Parallel and Distributed Systems, 26*(5), 1379–1392. https://doi.org/10.1109/TPDS.2014.2320515.

Staudt, C. L., Sazonovs, A., & Meyerhenke, H. (2016). NetworKit: A tool suite for large-scale complex network analysis. *Network Science, 4*(04), 508–530. https://doi.org/10.1017/nws.2016.20.

Temizsoy, A., Iori, G., & Montes-Rojas, G. (2017). Network centrality and funding rates in the e-MID interbank market. *Journal of Financial Stability, 33*, 346–365. https://doi.org/10.1016/j.jfs.2016.11.003.

Tong, G. A., Li, S., Wu, W., & Du, D.-Z. (2016). Effector detection in social networks. *IEEE Transactions on Computational Social Systems, 3*(4), 151–163. https://doi.org/10.1109/TCSS.2016.2627811.

Tong, G., Wu, W., Tang, S., & Du, D.-Z. (2017). Adaptive influence maximization in dynamic social networks. *IEEE/ACM Transactions on Networking, 25*(1), 112–125. https://doi.org/10.1109/TNET.2016.2563397.

Zafarani, R., Abbasi, M. A., & Liu, H. (2014). *Social media Mining: An introduction*. Cambridge University Press.

Zahara, E., & Kao, Y.-T. (2009). Hybrid Nelder–Mead simplex search and particle swarm optimization for constrained engineering design problems. *Expert Systems with Applications, 36*(2), 3880–3886. https://doi.org/10.1016/j.eswa.2008.02.039.

Zengin Alp, Z., & Gündüz Öğüdücü, Ş. (2018). Identifying topical influencers on twitter based on user behavior and network topology. *Knowledge-Based Systems, 141*, 211–221. https://doi.org/10.1016/j.knosys.2017.11.021.

Zhang, H., Nguyen, D. T., Zhang, H., & Thai, M. T. (2016). Least cost influence maximization across multiple social networks. *IEEE/ACM Transactions on Networking, 24*(2), 929–939. https://doi.org/10.1109/TNET.2015.2394793.

Zhang, K., Du, H., & Feldman, M. W. (2017). Maximizing influence in a social network: Improved results using a genetic algorithm. *Physica A: Statistical Mechanics and Its Applications, 478*, 20–30. https://doi.org/10.1016/j.physa.2017.02.067.

Zhou, C., Zhang, P., Zang, W., & Guo, L. (2015). On the upper bounds of spread for greedy algorithms in social network influence maximization. *IEEE Transactions on Knowledge and Data Engineering, 27*(10), 2770–2783. https://doi.org/10.1109/TKDE.2015.2419659.