Journal of Computer and Knowledge Engineering, Vol. 8, No. 2, 2025. (75-88)

Ferdowsi University of Mashhad

Journal of Computer and Knowledge Engineering

https://cke.um.ac.ir

Information and Communication Technology Association of Iran

Influence Maximization in Social Networks Using Discrete Manta-Ray Foraging Optimization Algorithm and Combination of Centrality Criteria*

Research Article

Zaynab Azizpour[1], Saeid Taghavi Afshord[2], Bagher Zarei[3], Mohammad Ali Jabraeil Jamali[4], and Shahin Akbarpour[5]

doi 10.22067/cke.2025.91839.1148

Abstract *Influence Maximization (IM) is a fundamental problem in social network analysis that seeks to identify a small set of highly influential nodes that can maximize the spread of information. Due to its NP-hard nature, finding an exact solution is computationally infeasible for large-scale networks. To address this, this paper introduces an enhanced discrete Manta-Ray Foraging Optimization (MRFO) algorithm tailored for IM. The proposed method integrates degree, closeness, and betweenness centrality measures into the fitness function and introduces a fused centrality index to improve the identification of influential nodes. To handle the discrete search space, the continuous MRFO is adapted with novel discretization mechanisms. Experimental evaluations on five real-world networks (NetScience, Email, Hamsterster, Ego-Facebook, and Pages-PublicFigure) demonstrate that the proposed method achieves higher influence spread compared to existing baseline algorithms, with average improvements of 14.63%, 12.81%, 19.03%, 15.24%, and 18.76%, respectively. These results validate the effectiveness, robustness, and practical applicability of the proposed approach for large-scale IM.*

Key Word Social networks, IM, Manta-Ray Foraging optimization algorithm, Centrality criteria.

1. INTRODUCTION

Complex networks with adjacency matrices are considered as graph G = (V, E), where V and E are edges and vertices of a graph. Each vertex in G demonstrates a user in a social network, and each edge indicates the relation between a pair of users. The size of each network is defined based on the network users, N= |V|, and available links in the network, M =|E|. The network structure is shown as n× n adjacency matrix, A = (aij), where each node can have the values {0, 1}. If user i is connected to user j, then aij =1; otherwise, aij = 0 [1]. Fig1 shows an example of social networks based on the neighbor graph.

As illustrated in Fig. 1, user relationships within a social network are determined based on the network's links and connections. These relationships significantly affect the diffusion of information across the network. Due to the large number of users and the complexity involved in identifying the most influential ones, exhaustive search methods are impractical and computationally expensive.

In recent years, social networks have gained widespread popularity, resulting in an increased impact on various aspects of society. For instance, social networks play a vital role in controlling the spread of diseases, marketing products, and conducting political campaigns such as presidential elections. A fundamental challenge in these contexts is how to effectively select influential users to maximize the spread of information or influence within the network. The IM (IM) problem addresses this challenge by seeking to identify the top k most influential nodes that can generate the largest possible spread of influence throughout the social network [2].

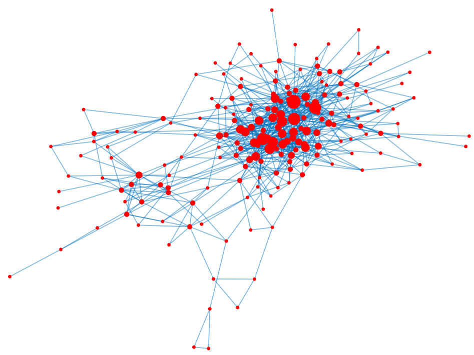


Fig 1. An example of social networks

* Manuscript received 2025 Jenuary 25, Revised 2025 July 9, Accepted 2025 July 12.

[1] Department of Computer Engineering, Shab.C., Islamic Azad University, Shabestar, Iran

[2] Corresponding author. Assistant Professor, Department of Computer Engineering, Shab.C., Islamic Azad University, Shabestar, Iran. **Email:** sa.taghavi@iau.ac.ir

[3] Corresponding author. Assistant Professor, Department of Computer Engineering, Shab.C., Islamic Azad University, Shabestar, Iran. **Email:** zarei.bager@iau.ac.ir

[4] Department of Computer Engineering, Shab.C., Islamic Azad University, Shabestar, Iran

[5] Department of Computer Engineering, Shab.C., Islamic Azad University, Shabestar, Iran

IM is a widely studied topic in the context of social networks. A social network can be modeled as a directed graph, where the users are represented as nodes and their connections as directed edges. Influence spreads throughout the network via the "word-of-mouth" effect, which captures the human-to-human transmission of information or ideas. This process can lead to either a rapid decline or exponential growth in the spread of information.

A key challenge in this domain is to estimate how many users can be influenced by a small group of highly influential individuals. The fundamental goal of IM is to identify an initial set of nodes (seed nodes) that is as small as possible while maximizing the spread of influence across the network. In other words, IM plays a crucial role in viral marketing by helping identify potential customers who can trigger widespread adoption, thereby reducing marketing costs and maximizing profit. Viral marketing leverages word-of-mouth dynamics by targeting a small group of individuals to try a product and encourage broader usage [3].

In practice, IM algorithms determine which nodes should be initially activated. Given a graph G and a parameter K, these algorithms produce an initial seed set by estimating the expected number of nodes that will be influenced through a stochastic diffusion process. The core objective in IM is to maximize the expected size of the final active set while using the smallest possible number of influential users, subject to certain constraints on the initial seed set.

The diffusion process starts with these seed nodes and aims to maximize the overall influence spread within the network. The number of nodes activated during this process determines the effectiveness of the selected seed nodes. IM is an NP-hard problem, meaning that there is no known deterministic polynomial-time algorithm to solve it optimally. Therefore, meta-heuristic optimization methods are commonly employed to find near-optimal subsets of influential users within a reasonable computational time [3,4].

IM is finding a set of influential users, (seed set S) S⊂ V consisting of $K < |V|$ nodes in a social network $G = (V, E)$, Where V is the set of nodes (users), and E is the set of (directed/undirected) edges in G (i.e. social relations between the users). The goal is to maximize K in G through the propagation of the diffusion model. The problem is described by the following equation [3]:

$$IM_M(G, K) = argmax_{e \subseteq V, |e|=k} \sigma_M(e, G) \tag{1}$$

Where σ is a function that calculates the extent of influence for a given set of nodes and represents the spread of influence by activating the set of nodes in e.

Identifying influential users within a network, particularly in large-scale social networks, is a challenging and engaging research problem. Nodes occupy different positions and play various roles within the network, and the effectiveness of influence diffusion largely depends on the underlying network topology. Certain nodes possess structural advantages that make them more effective at spreading information. For instance, central nodes often serve as key conduits for information flow, while nodes with a high number of connections (degree) significantly contribute to influence propagation. Conversely, nodes located at the periphery of the network or those forming isolated clusters may have minimal impact on overall diffusion [5].

Consequently, social importance measures such as degree, betweenness, and closeness centralities — as well as their combined usage — provide valuable insights for identifying the most influential users.

In this study, we propose a discrete method called the Centrality Measure-based Manta-Ray Foraging Optimization (CMMRFO) algorithm, which integrates multiple social importance criteria. The proposed approach has been evaluated using the Facebook dataset. CMMRFO aims to identify a minimal set of users that maximizes influence spread within the network. To achieve this, the algorithm employs a bi-objective fitness function whose weight coefficients are determined by degree, betweenness, closeness centralities, and their combination.

The main contributions of this paper are summarized as follows:

- Development of a discrete version of the MRFO algorithm for solving the discrete IM problem.
- Design of a bi-objective fitness function to simultaneously minimize the number of seed users and maximize overall influence spread.
- Incorporation of social importance measures as weight coefficients in the fitness function.
- Application of degree, betweenness, and closeness centralities, along with their fusion, to guide the search for influential users.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 describes the proposed methodology in detail. Section 4 presents the implementation details and experimental results. Finally, Section 5 concludes the paper and outlines potential future work.

2. BASIC CONCEPTS

2.1. MRFO

The MRFO algorithm is a computational method designed to solve complex optimization problems inspired by the natural foraging behavior of manta rays. In this context, the goal is to determine an optimal path that minimizes the overall cost of travel through multiple locations, analogous to a manta ray visiting several "fish cities." The MRFO algorithm identifies this optimal route through a combination of local search, evolutionary strategies, and similarity-based operations.

In practice, the underlying problem is first formulated as a mathematical optimization task. At each iteration, the algorithm generates a new candidate route for the manta ray. This candidate route is then evaluated and compared with the current best-known route. If the new route yields a lower cost, it replaces the previous best; otherwise, it is discarded. Through iterative refinement, the MRFO algorithm converges toward an optimal or near-optimal solution for the routing problem. This approach can be applied not only to manta ray path planning but also to other similar route optimization challenges.

1) Mathematical Model: The MRFO algorithm is inspired by three distinct foraging behaviors observed in manta rays: chain foraging, spiral foraging, and storm foraging. These behaviors are mathematically modeled to guide the search process toward optimal solutions, as detailed below.

2) Chain Foraging Strategy: In the chain foraging phase, manta rays detect the location of plankton and swim toward areas with higher concentrations. In the optimization analogy, these high-concentration zones correspond to promising candidate solutions. Although the global optimum is unknown, MRFO assumes that the best solution discovered so far represents the most desirable "plankton" location.

Manta rays are conceptually arranged in a head-to-tail sequence, forming a chain. Except for the leading individual, each manta ray updates its position by moving not only toward the detected food source but also relative to the preceding individual in the chain. This ensures collective information sharing and improved exploration of the search space. At every iteration, each individual's position is refined based on the best solution found up to that point. The mathematical formulation of the chain foraging behavior is provided below.

$$x_i^d(t+1) = \begin{cases} x_i^d(t) + r.\left(x_{best}^d(t) - x_i^d(t)\right) + \alpha.\left(x_{best}^d(t) - x_i^d(t)\right) & i = 1 \\ x_i^d(t) + r.\left(x_{i-1}^d(t) - x_i^d(t)\right) + \alpha.\left(x_{best}^d(t) - x_i^d(t)\right) & i = 2, \dots, N \end{cases} \quad (2)$$

$$\alpha = 2.r.\sqrt{|\log(r)|} \quad (3)$$

Where $x_i^d(t)$ is the position of the ith individual at time t in the dth dimension, r is a random vector in the range [0, 1], while α is weight factor, $x_{best}^d(t)$ is a location with the highest plankton concentrations. Fig 2. displays the behavior of food Search in two-dimensional space. The position Update of the ith individual is determined by the position $x_{i-1}(t)$ for (i-1) the individual and the position $x_{best}(t)$ of the food.

3) Storm Search Strategy: When a group of manta rays detect clusters of plankton in deeper waters, they form a long foraging chain and swim toward the food source using a spiral motion. This spiral foraging strategy resembles the approach used in the Whale Optimization Algorithm (WOA); however, in the MRFO framework, the spiral movement is incorporated specifically in the storm foraging phase. In this strategy, each manta ray moves in a spiral path toward the food source while simultaneously following the preceding individual in the chain. In this way, the manta rays align sequentially and execute a coordinated spiral search to locate and capture plankton more efficiently.

Fig 3. illustrates the storm foraging behavior in a two-dimensional space. In this phase, each individual not only follows the food targeted by his neighbor but also advances toward the food source by executing a spiral trajectory. The mathematical equations that describe this spiral motion in two-dimensional space are defined as follows:

$$\begin{cases} X_i(t+1) = X_{best} + r.\left(X_{i-1}(t) - X_i(t)\right) + e^{b\omega}.cos(2\pi\omega).\left(X_{best} - X_i(t)\right) \\ Y_i(t+1) = Y_{best} + r.\left(Y_{i-1}(t) - Y_i(t)\right) + e^{b\omega}.sin(2\pi\omega).\left(Y_{best} - Y_i(t)\right) \end{cases} \quad (4)$$

Where ω is a random number in [0, 1].

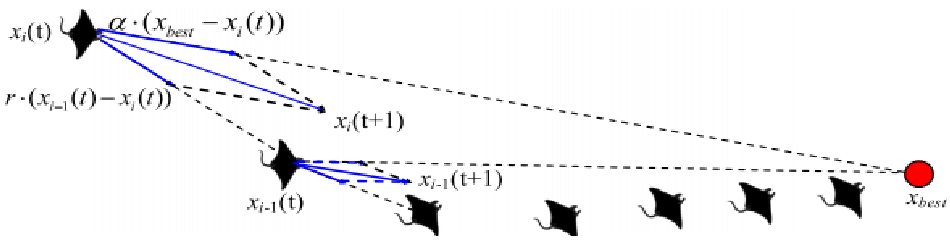


Fig. 2. The behavior of food search in two-dimensional space.

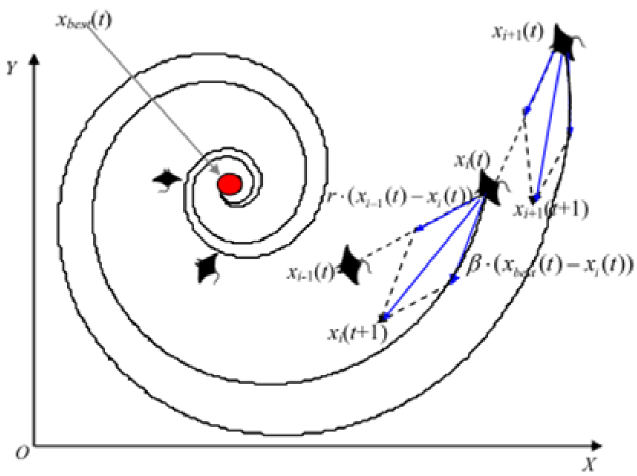


Fig 3. The behavior of storm search in two-dimensional space

This motion behavior may be extended to n-D space. For simplicity, this mathematical model of silicon search can be defined as:

$$parentx_i^d(t+1) =$$

$$\begin{cases} x_{best}^d + r.\left(x_{best}^d(t) - x_i^d(t)\right) + \beta.\left(x_{best}^d(t) - x_i^d(t)\right) & i = 1 \\ x_{best}^d + r.\left(x_{i-1}^d(t) - x_i^d(t)\right) + \beta.\left(x_{best}^d(t) - x_i^d(t)\right) & i = 2, \dots, N \end{cases}$$

$$\beta = 2e^{r_1\frac{T-t+1}{T}}.sin(2\pi r_1)$$

Where β is the weight coefficient, T is the maximum number of iterations, and $r1$ is the random number in [0, 1].

In the cyclone foraging phase, all individuals perform a randomized search using the current best-known food location as their reference point. This mechanism enhances the algorithm's exploitation capability within regions that contain promising solutions. Additionally, this strategy significantly improves the overall search process by enabling individuals to explore new regions. Specifically, each individual can be directed to search for alternative positions that deviate from the current best solution, or it can adopt a completely random position anywhere within the entire search space as a new reference. This balance between local exploitation and global exploration ensures that the MRFO algorithm maintains strong heuristic capabilities while avoiding premature convergence. The corresponding mathematical formulation for this mechanism is provided below.

$$x_{rand}^d = Lb^d + r.(Ub^d - Lb^d)$$

$$x_i^d(t+1) = \begin{cases} x_{rand}^d + r.\left(x_{rand}^d - x_i^d(t)\right) + \beta.\left(x_{rand}^d - x_i^d(t)\right) & i = 1 \\ x_{rand}^d + r.\left(x_{i-1}^d(t) - x_i^d(t)\right) + \beta.\left(x_{rand}^d(t) - x_i^d(t)\right) & i = 2, \dots, N \end{cases}$$

Where x_{rand}^d is a random position randomly produced in a search space. Lb^d and Ub^d are the lower and high limits of the search space.

4) Somersault search strategy: In this phase, the location of the food source is treated as a central axis. Each swim around this pivot point and performs somersault-like movements to discover new positions in its vicinity. Through this behavior, individuals continuously update their positions around the best solutions identified so far, enhancing local exploitation while maintaining diversity. The corresponding mathematical formulation for this behavior is presented below.

$$x_i^d(t+1) = x_i^d(t) + S.\left(r_2.x_{best}^d - r_3.x_i^d(t)\right),$$
$$i = 1, \dots, N$$

Where S is the somersault factor that determines the range of movement of Manta Rays movement and S = 2, $r2$ and $r3$ are two random values in the range [0, 1].

As shown in Equation (8), by defining the somersault range, each individual can explore a new search area bounded between its current position and its symmetric counterpart relative to the best position identified thus far. As the distance between an individual's current position and the best-known position decreases, the degree of disturbance applied to its position also diminishes. Consequently, all individuals progressively converge toward the optimal solution within the search space. Therefore, as the number of iterations increases, the somersault foraging range adaptively contracts. Fig 4. illustrates a schematic representation of the somersault foraging behavior in the MRFO algorithm.

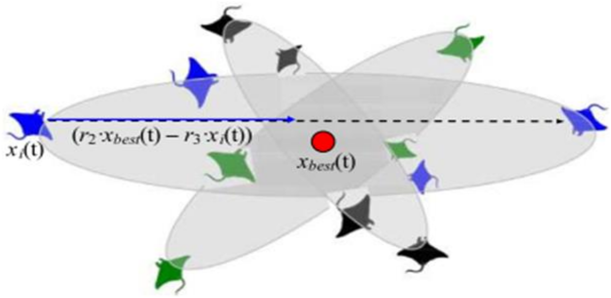


Fig. 4. Somersault foraging behavior in MRFO

Fig 5. illustrates the evolution of three individuals over 100 iterations within the search space based on the corresponding equation. The sampled points are randomly generated between each individual's current position and its symmetric position relative to . As the distance to $xbest$ decreases, the sampled points become more concentrated. This pattern ensures that densely clustered points near $xbest$enhance local exploitation, while more widely distributed points support broader exploration of the search space.

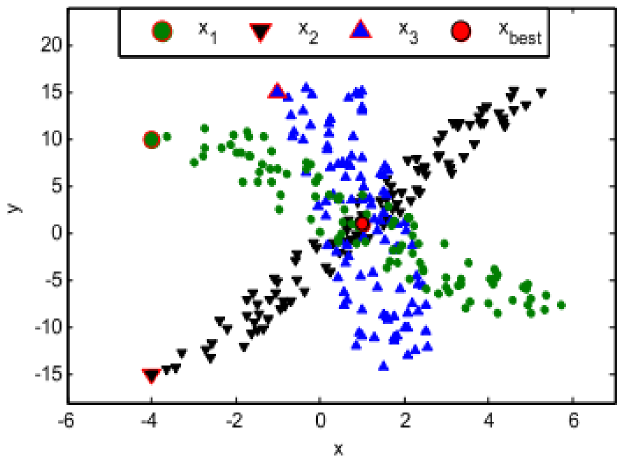


Fig. 5. The Somersault foraging behavior of three individuals in two-dimensional space

Unlike other meta-heuristic optimization methods, MRFO starts the problem by creating the initial population. Each updates their situation by considering the reference and opposite situation. The t/T value is decreased from 1/T to 1 so that heuristic search and application are performed, respectively. The best current solution is selected as a reference situation wage when t/T < U (0,1). A position is randomly created in search space, and It is selected as the reference position for the heuristic when t/T > U (0,1). It can be moved between chain search behavior and storm search behavior. Subsequently, each updates its position relative to the best position identified through the somersault search mechanism. These updates and computations continue iteratively until the predefined stopping criterion is satisfied. Ultimately, the position and fitness value of the best-performing individual are returned [25].

In the next stage, the proposed method is described in detail. This method is based on the CMMRFO algorithm, an evolutionary approach designed to solve the IM problem. In this approach, measures of user importance within the network are incorporated as key factors. The CMMRFO algorithm enhances solutions through random variations and mutations within the population, iteratively seeking better candidates. By integrating the proposed CMMRFO method with user importance metrics, the influence spread in social networks is effectively maximized, providing a more optimal solution to the IM problem.

3. LITERATURE REVIEW

In this section, the research on IM in social networks is discussed. It has more commercial applications, and so it has been more studied. One of the essential tools for identifying the influenced users is using social importance criteria. These criteria involve degree centrality, betweenness centrality, closeness centrality, and other similar criteria. The users with the maximum network influence are identified using these criteria. In the following, IM methods based on social importance criteria are studied in detail. An effective influence evaluation model based on whole valuation and variance of neighbor nodes valuation has been presented to create unreliable communication channels [6]. Then, the Moth-flame optimization algorithm has been developed to search the set of influence-maximizing nodes by using local intersection and mutation evolution above updating the conventional solution. The criterion of degree centrality was introduced by [7] for the first time and later it was used for IM. Then, using an analysis framework based on modular functions such as BC [8] was shown that a greedy natural strategy obtains a solution that can be proved to be 63% optimum for several class models. This framework suggests a reasoning approach to guarantee the performance of algorithms for this kind of influence problem in social networks. An approach based on PageRank for influence maximizing in a network to search the web has been proposed in [9]. A different approach based on simulated Annealing for IM has been suggested in [10]. This is the first SA-based algorithm for solving this problem. In addition, two heuristic methods have been proposed to accelerate the process of SA convergence, and a new method has been suggested for computation influence to accelerate this algorithm.

Some changes to the IM problem structure were presented by [11] to adjust it with particle swarm intelligent algorithms and to reach a slope in the space of the objective function. The proposed approach was tested using real and artificial data sets. The gray wolf optimization (GWO) algorithm has been considered as a particle swarm intelligence algorithm, along with page ranking and greedy algorithms were used as evaluation methods. The reason for the low performance of greedy approaches was analyzed in [12] and an efficient algorithm called degree-descending search strategy evolution (DDSE) has been proposed. Firstly, a degree-descending search strategy is suggested, which can produce a set of nodes whose influence spread can be compared to the centrality degree. An evolutionary algorithm based on DDSE has been developed that considerably improves efficiency by removing time-consuming simulations of greedy algorithms.

An improved discrete particle swarm optimization algorithm along with an advanced network topology-based strategy for influence maximizing has been proposed in [13]. In this strategy, at first, k-influenced nodes of a temporary optimal seed set are combined in an ascending order based on degree metric so that the nodes with lower degree centrality can utilize preferably the influenced neighbors. In the second step, a local greedy strategy is applied to replace the current node with the most influenced node of each node’s direct neighbor node-set of temporary seed set.

An improved greedy-based strategy called Cost-Effective Lazy Forwarding (CELF) has been created [14]. It reduces computation costs twice without damaging precision by utilizing the submodularity of objective function. Later, an optimized version called CELF++ was suggested [13], and the results showed 50% more efficiency improvement compared to CELF.

Time-sensitive centrality criteria were presented for IM in social networks by considering the diffusion value, and direct and indirect neighbors [15]. Hence, four time-sensitive centrality measures, including time-sensitive closeness centrality, time-sensitive harmonic, time-sensitive decay centrality, and time-sensitive eccentricity centrality were proposed.

Degree centrality based on various environments is used by [16] to increase its local search power. Through performing experiments, it has been specified that local search strategies based on different environments have considerable differences in improving the global search of the algorithm, and increasing the DPSO algorithm based on the degree centrality of different environments has considerable influences. Finally, the DPSO-NDC algorithm has been suggested based on the degree of centrality of the best environment with improved local search capability.

A mechanism to measure the influence index in popular social media platforms like Facebook, Twitter, and Instagram was suggested in[17]. Some sets of features determining influence on the consumers are modeled by regression approach. Infrastructure machine learning algorithms involve Ordinary Least Squares (OLS), Regression Nearest Neighbor (KNN), Support Vector Regression (SVR), and Lasso Regression models to compute cumulative scores adopted in terms of influence index. The findings show that Participation, meta-learning, and feedings are crucial in determining influencers.

An improved discrete differential evolution algorithm (IDDE) based on the network analysis has been suggested in [18]. This algorithm improves the variance of the differential evolution algorithm. After removing the objective node as an index, it gives a discrete number and discrete precision of the remaining network to evaluate the importance of the node and as a result, it presents a health function based on the network power. This method shows symmetry in two aspects. Firstly, when the number of removed objective nodes increases in a social network, global coherence decreases between the network nodes. Secondly, the range of global influence becomes small when the proposed method displays the number of objective nodes. Comparable experiments have been performed on four real-world data sets with different sizes. The results show that the IDDE algorithm outperforms the comparison algorithms.

The authors of [19] present a framework involving community detection in a social network using the Shuffled Frog Leaping algorithm (SFL). This framework aims to maximize influence spread in an independent cascade model. In this framework, various communities are identified in a social network using a community detection algorithm. Then, the SFL algorithm is applied to maximize the influence spread in these communities. The SFL algorithm is an evolutionary algorithm, searching the solutions improvement based on random changes and mutations created in the frog population. Local search strategies, including hill climbing based on lake adoption and user centrality weight, are also used to more improvement in the solutions. These strategies find the best points in the search space by using the weight of the user's centrality, and they make more improvements in the solutions by local search. Therefore, this framework, including community detection, SFL algorithm, and local search strategies, maximizes influence under the independent cascade model, and optimal solutions are provided for this problem.

A meta-heuristic approach based on multi-criteria decision-making (MCDM) has been proposed by the authors of [20] to solve the IM problem in social networks. The MCDM approach selects candidate nodes by removing less-influence nodes in the preliminary step based on the centrality criterion, and it decreases the computational cost. Afterward, an improved version of simulated Annealing (SA) with an advanced search strategy has been suggested to find an optimal solution.

An evolutionary Discrete Crow Search Algorithm (DCSA) using crow swarm intelligence has been suggested by [21] to solve effectively the IM problem. DCSA makes a new coding mechanism and discrete evolution rules. Initialization methods based on degree centrality and random walking strategy are applied to increase searchability. An IM algorithm called Weighted Artificial Bee Colony (WABC) has been proposed by [22]. It is based on a technique inspired by biology to detect the subset of users that maximizes diffusion. WABC has used ranking techniques based on classic centrality criteria. A new approach with multi-feature IM has been suggested by [23]. This approach uses the multi-feature nature of network nodes (age, gender, etc.) to consider specified groups of users. Also, it uses centrality criteria to rank the user's importance in various groups. The Discrete Bat Algorithm (DBA) has been presented by [24]. It is based on partitioning a network and increases the stability of the initial DBA. The experimental results showed that the DBA converges in each run to a specified Local Influence Estimation (LIE) value. It removes the high oscillation phenomenon of the LIE fitness value created by the main DBA. This method has been used centrality criteria for local search in the fitness function. In [29], a novel Multi-objective Cuckoo Search Algorithm (MOCSA) designed for community detection in social networks, emphasizes improved accuracy and efficiency by incorporating a strategy based on close neighbors in the objective function. In [30], a hybrid multi-objective algorithm is presented incorporating multiple optimization techniques and fuzzy clustering that outperforms existing methods in detecting overlapping communities in complex social networks. In [31], the LCD-SN algorithm enabled highly accurate and efficient community detection in social networks by leveraging local node characteristics and neighbor information without dependence on initial seed nodes. In [32], Opinion Leader Selection (OLS) as an optimization problem using bio-inspired algorithms has been formulated. It combined the African Vultures Optimization Algorithm (AVOA) and Hunger Games Search (HGS) for improved leader identification. In [33], the proposed method effectively identified influential opinion leaders in social networks using hybrid optimization algorithms and topological network analysis, achieving higher accuracy and marketing impact than existing approaches. A comparison of the previous works is shown in Table 1.

Despite extensive efforts in IM, existing methods often rely on either single centrality measures or heuristic meta-heuristics that are not fully adapted to the discrete nature of social networks. Many approaches use basic node rankings (e.g., degree or PageRank) without integrating multiple structural properties, which limits their accuracy in identifying truly influential nodes. Additionally, continuous meta-heuristic algorithms are frequently applied with minimal modification, resulting in suboptimal performance when handling the inherently discrete selection of seed nodes. Moreover, some algorithms suffer from high computational costs or slow convergence, especially on large-scale networks. Therefore, there is a clear need for a method that effectively fuses multiple centrality measures within a robust, discrete meta-heuristic framework to achieve higher influence spread while maintaining computational efficiency.

TABLE 1
The previous work's comparison

Ref	Algorithm / Method	Key Idea	Strength	Limitation
[6]	Moth-Flame Optimization	Uses whole valuation & neighbor variance to handle unreliable channels	Handles uncertainty	Limited to specific network conditions
[7][8]	Degree Centrality & Greedy	Early use of centrality; submodular function framework	Theoretical performance guarantee (63% optimal)	Greedy: high time cost, limited scalability
[9]	PageRank	PageRank-based node ranking	Intuitive for web networks	Less effective for general social graphs
[10]	Simulated Annealing	First SA-based IM with acceleration heuristics	Good exploration ability	Slow convergence in large graphs
[11]	Particle Swarm Optimization & GWO	PSO adapted to IM; GWO used for evaluation	Intelligent swarm behavior	May stagnate; lacks centrality integration
[12]	DDSE (Degree-Descending Search Evolution)	Evolutionary search avoiding costly simulation	Faster than greedy	Needs careful parameter tuning
[13]	Improved Discrete PSO	Advanced topology strategy with local greedy replacement	Good local refinement	May get stuck in local optima
[14]	CELF / CELF++	Optimized greedy selection with lazy evaluation	50% faster than CELF	Still costly for large networks
[15]	Time-sensitive Centrality	Four new time-aware centrality measures	Considers diffusion time	High computation for dynamic networks
[16]	DPSO-NDC	Local search in different environments	Improved local/global balance	Sensitive to environment selection
[17]	ML-based Influence Index	Regression models for social media	Leverages user behavior features	Not directly IM for seeding
[18]	Improved Discrete Differential Evolution	Variance-based node ranking; health function	Symmetric handling of removed nodes	Limited to certain network structures
[19]	SFL + Community Detection	Shuffled Frog Leaping with Community Detection	Uses local community structure	Relies on quality of community detection
[20]	MCDM + Improved SA	Node filtering + advanced SA	Reduces cost via pre-selection	SA still has convergence limits
[21]	Discrete Crow Search	Crow swarm intelligence; random walking	Novel coding; better exploration	Lacks robust local refinement
[22]	Weighted Artificial Bee Colony	Biological swarm inspired; uses ranking	Effective for classic criteria	Ranking alone may overlook structural synergy
[23]	Multi-Feature IM	Uses user attributes + centrality	More realistic user modeling	Needs rich attribute data
[24]	Discrete Bat Algorithm	Uses partitioning to stabilize LIE	Better stability; local search	High oscillation is removed but may converge slowly
[29]	Multi-objective Cuckoo Search Algorithm (MOCSA)	Uses cuckoo search with a neighbor-based strategy for accurate community detection.	High detection accuracy and efficiency.	May face convergence issues in very large-scale networks.
[30]	Hybrid Multi-objective Algorithm with Fuzzy Clustering	Integrates multiple optimization techniques and fuzzy clustering for overlapping community detection.	Handles overlapping communities effectively, with better performance than traditional methods.	Increased computational complexity due to the hybrid structure.
[31]	LCD-SN Algorithm	Leverages local node characteristics and neighbor data for community detection without relying on seed nodes.	High accuracy and efficiency, seed-free approach.	May require fine-tuning for networks with sparse connections.
[32]	Opinion Leader Selection (OLS) with AVOA and HGS	Formulates leader selection as an optimization problem using bio-inspired algorithms (AVOA + HGS).	More precise leader identification and robust search capability.	Algorithm performance is sensitive to parameter settings.
[33]	Hybrid Optimization for Opinion Leader Detection	Combines hybrid optimization algorithms with network topology analysis for better opinion leader identification.	Higher accuracy and marketing influence than other methods.	May need high computational resources for very dense networks.

4. THE PROPOSED METHOD

This paper presents a new method called CMMRFO for IM in social networks using the Manta Ray algorithm and a combination of centrality criteria. A weighted combination of criteria, including degree, betweenness, and closeness centralities, are used to compute the social importance of the users. Since the MRFO is continuous, it cannot be used to solve the discrete IM problem. Hence, a discrete method is proposed for the MRFO, and its discrete version is used to solve the IM problem. The details are explained in the following.

4.1. discretization of the MRFO

In meta-heuristic algorithms, discretization converts continuous values into discrete ones. The purpose is to convert continuous search space to discrete search space so that meta-heuristic algorithms can find the problem's best solution. Continuous values are converted like real numbers in search space, and the initial population as discrete values or a set of integers in most meta-heuristic algorithms for discretization. This conversion can be performed as a simple discretization, for instance, by estimating the value to the nearest discrete, or it can be performed as a complex discretization by conversion function or other methods. This conversion to discrete space helps the algorithms search for the best solution in discrete space, and in this way, efficiency improvement and the efficiency of algorithms can be increased. Since the most fundamental problems are not continuous, discretization helps the algorithms reach the optimal value of the problems [26]. In this paper, since the search space is selected among social network users, the initial population's values indicate the user's index in a social network. Hence, this problem does not involve containing values for the initial population. In addition, the severe population cannot include continuous values in the heuristic process in the meta-heuristic fragging optimization algorithm. As a result, the initial population is a limited range of user indexes in the proposed method, which is valid for the centrality threshold. They can be converted to an influence user. An integer for a random variable can be used in the heuristic step in the MRFO instead of X_{rand}^{d} which is a real random variable.

4.2. Computing the social importance of users

Increasing the users and the data volume of these users in social networks requires analyzing and extracting useful information from data. Such information can be useful for different applications like advertisements, marketing procedures of user behavior, etc. In this regard, it is one of the valuable tools in user importance criteria in social networks. These criteria investigate the user's importance in various aspects of social networks, and they are introduced as an effective tool for analyzing social network data. This study uses three criteria of centrality importance involving closeness, betweenness, and a combined criterion of the network. They are explained in detail in the following.

1) Centrality criteria: Centrality criteria are the network analysis measures used to detect the most powerful nodes. The centrality criterion quantifies direct friendship relations for a node in social networks. According to the centrality criterion, the importance of a node is determined based on its degree. Suppose G = (V, E) is a social network. V is a set of n nodes, while (n=|V|) shows the users in a network. E indicates a set of m edges between the users. (m=|E|), shows the relations between the users. Social networks are shown by a symmetric matrix A called adjacency matrix with dimensions n × n. Each entry, a_{ij} is considered the relationship between node i and node j if equal to. According to the centrality criterion, it can be computed for each user as equation 10 [27].

$$C(i) = \sum_{i=1}^{n} \sum_{j=1}^{n-1} \frac{a_{ij}}{n-1} \tag{10}$$

An ij is the relation between the user i and j in the adjacency matrix, and n is the whole number of users in a network. In this case, the value of the degree of centrality can be obtained for all users in a network, and it indicates a favor among social networks. The users having relations with many users have a high value of degree centrality. Such users are exposed to information or data diffusion in social networks. In contrast, such users with a low degree of centrality do not have so much popularity and show introverted personalities. This criterion limitation is local access to the network topology, and it uses limited local knowledge to decide about the user's importance.

2) Closeness criterion: The closeness criterion hypothesizes that the power of an individual has a reverse relation with another individual in that social network in terms of closeness with distance sum [27]. In other words, the closeness criterion for the user is obtained based on whole routes from other users in a social network. The closeness criterion is computed for each user as equation (11) [27].

$$CN(i) = \frac{1}{\sum_{i=1}^{n} \sum_{j=1}^{n-1} d_{ij}} \tag{11}$$

Where d_{ij} Shows the shortest route from node i to node j. The closeness criterion of the user shows the average distance of the user with another user in a social network as a quantitative value. The users with a high closeness criterion receive the information from each point of the network in less expected time because they are less distant from other social network users. In contrast, the user having less closeness criterion and being away from another network user receives data diffusion in networks later than expected. The limitation of this criterion is that disconnected networks do not work well, and they cover only a part of the connected network.

3) Betweenness criterion: According to betweenness criteria, the importance of a user in social networks is defined as the shortest path between other users passing through that point. In other words, the betweenness criteria for each user i is according to the shortest routes between all network users, and user i is located between them. The betweenness criterion is computed for each user as equation (12) [27].

$$B(i) = \sum_{i=1}^{n} \sum_{j=1}^{n-1} \sum_{k=1}^{n-2} \frac{\sigma_{jk}(i)}{\sigma_{jk}} \tag{12}$$

Where $\sigma_{jk}(i)$ is the shortest route from node j to k, which passes from node i. σ_{jk} shows the number of shortest routes from node j to k. In betweenness criterion data is transferred in a social network between users through the shortest routes. The individual having a high value of betweenness criteria has more control over the information in the whole network. Therefore, it can be a good alternative for selection as a user with high influence in social networks.

4) The combination criterion: Social importance criteria are those used to measure and analyze the user's importance and their communications in social networks. These criteria help us determine which entities have the maximum influence on a network and how communication is made between the users. Social importance criteria point to the influence of a user in a network. The conventional social importance criteria involve the degree of centrality. Degree Centrality measures the number of user communications. Betweenness centrality shows how a user is located in the routes between two other users. Closeness Centrality shows how a user is close to the other users in a network. The Fusion Centrality combines the effect of these three mentioned criteria simultaneously. It is defined as follows.

$$FC = w_1 * DC + w_2 * CC + w_3 * BC, \; w_1, w_2, w_3 \in [0,1] \tag{13}$$

Where FC indicates the combinations of centrality criteria, DC and w1 show the degree centrality criterion and the weight of the degree centrality criterion, respectively. Also, CC and w2 are the closeness centrality criterion and the weight of degree centrality criterion, respectively, BC and w3 refer to betweenness centrality criteria. The weight of each centrality criterion is a value between zero and one.

The classic importance criterion of centrality is determined based on the user's output and input links in a social network. The closeness criterion is specified according to the distance or required steps to reach another user in a network. At last, the combination criterion is made of these three essential criteria and combines previous criteria. To select the influenced users, the combination criterion considers three criteria, degree centrality, closeness centrality, and betweenness centrality. Other criteria investigate the user's importance in a social network in just one dimension. Using these combination criteria to analyze social network data helps users and managers make decisions about using social network data. This paper can be used as a helpful resource to study the combination of importance criteria in social networks and data analysis methods of a social network.

4.3. Solving IM

IM in social networks is detecting a set of individuals in a network with the highest potential to affect others. It is an important problem in the analysis of social networks because it can be used to spread the behaviors or particular ideas in a network or to prevent the spread of negative ideas or behaviors. Meta-heuristic algorithms are optimization algorithms used to solve classic optimization techniques. Meta-heuristic algorithms can be used to detect individuals with the highest potential to affect others in terms of IM in social networks. These algorithms are performed by producing and evaluating potential solutions and then using the best solutions to produce new ones.

1) Coding the initial population: In optimization problems, solved by meta-heuristic methods, initial population planning is an important issue. The purpose of this problem is to consider each member in the initial population as a final solution. Each member can be evaluated based on fitness function to specify the optimal solution. The MRFO is used to find k users among M users having the maximum influence in a social network. The initial population is defined as a vector of Manta rays, and elements connect them, and they constitute a solution. The initial population is defined as a vector of discrete values in this algorithm, and each entry indicates the user in the social network index. The length of this vector is equal to M (the number of users in a search space), and the value of each element shows the users' index in the social network. The value of zero in each element is related to the lack of selection of the related user, while the value of 1 shows the selection of that user. The problem space is limited according to the spread of the social network and the number of users. In other words, the users whose communication exceeds the threshold value are considered the available alternative in the problem space. The threshold value is the mean of the degree centrality for all users in the social network. So, influenced users can be selected among the users whose connection and communication are more than the mean of communication degree in the whole network. Fig 6 shows a sample of the initial solution.

	U1	U2	U3	U4	…	UM
X	5	16	0	8	…	27

Fig. 6. A sample of the initial solution in the proposed method.

According to Fig 6, it can be found that initial populations involve the element equal to the number of users in a search space, where the value of element 0 shows the selection of that user as an influenced user. If the value is non-zero, it shows the user's index in the social network. The users selected as the initial population are considered as the entry of the fitness function trying to compute the influence of users selected as influenced users.

2) Fitness function: The proposed fitness function finds the minimum number of users with maximum influence in social networks. In other words, social importance criteria are determined as an influence parameter factor in the fitness function. The fitness function has two parameters. The first parameter is related to the influence, while the second is related to the number of users in the social network. The proposed fitness function is as follows.

$$max\, f = \sum_{i=1}^{M} W_i SP_i - \sum_{i=1}^{n} W_i N_i \quad s.t. \qquad \sum_{i=1}^{n} W_i \leq 1 \qquad \sum_{i=1}^{n} N_i \geq 1 \qquad i > 0 \tag{9}$$

Where i shows the user index in the social network in equation 9; M refers to the number of users in the influenced search space; SP_i demonstrates the number of users under the influence; Ni and Wi show the number of influences and the weight of social importance, respectively. Fig 7. presents the flowchart of the proposed method. In the following, the proposed method is implemented and evaluated.

4.4. Diffusion model in the proposed methods

The information diffusion method in social networks is a form that explains how information and contents are distributed in these networks and how different individuals and users influence this information diffusion. These models are usually defined based on human behavior, social network algorithms, and individual's reactions to the information propagated by others. These models help better recognize information diffusion procedures in social networks, and better program and manage information. Information diffusion models in social networks are usually divided into threshold, cascade, trigger, and epidemic models. These models are introduced as follows.

- Threshold models: The information is directly diffused by the individuals, and each individual decides whether the information is diffused among others. This model is based on personal decision-making.
- Cascade models: The individuals are influenced by the diffused information by others, and information is diffused among others.
- Trigger models: The information is diffused automatically by the individual and without individual decision-making. This model is based on automatic processes and algorithms of social networks.

Pidemic models: The information diffusion is like the spread of an epidemic illness, which is quickly diffused in social networks. This model is based on the quick and widespread information in social networks.

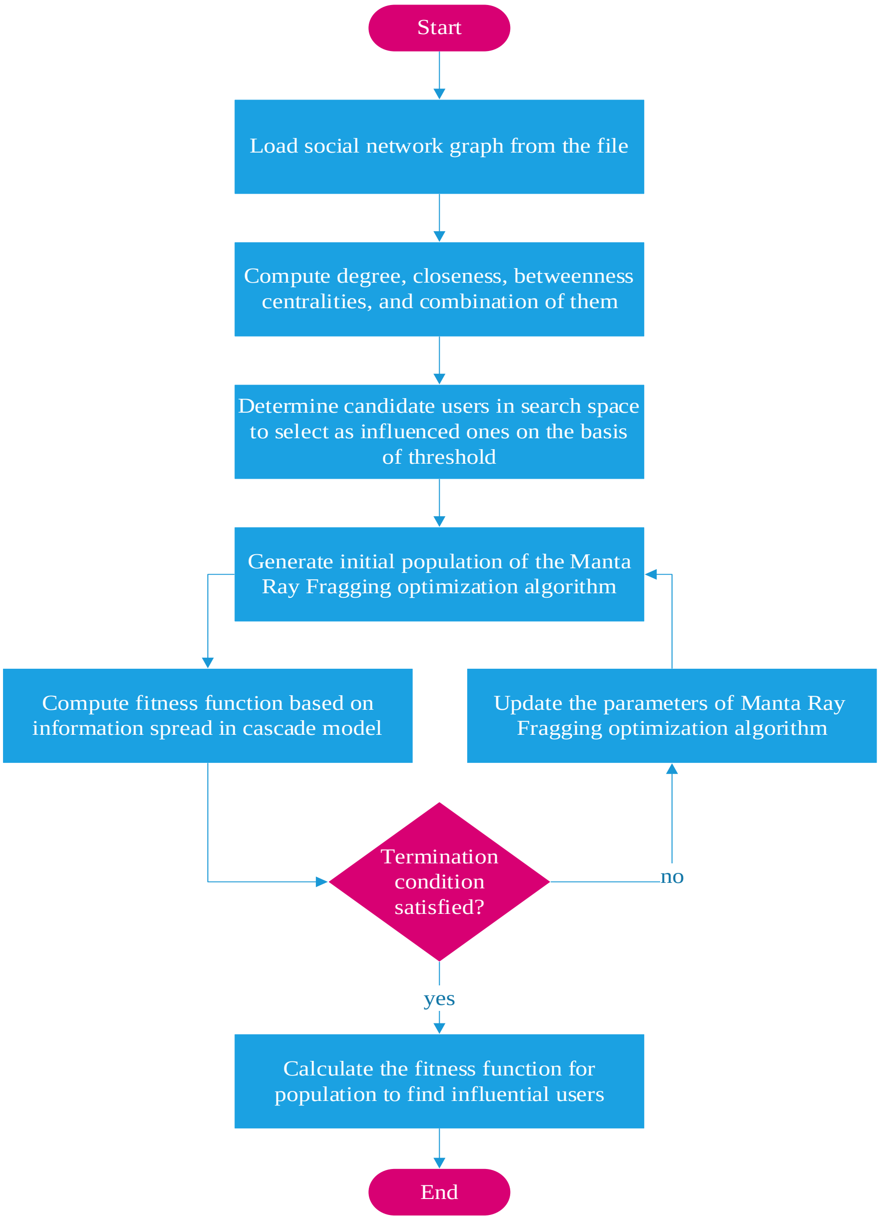


Fig. 7. Flowchart of the proposed method

The cascade diffusion model is used in the proposed model. This model shows the procedure of information transfer and content propagation in social networks. When an individual diffuses content, this content is observed by others, and some of them decide to transfer it to others. This content transfer process continues through individual social communication and is quickly transferred to other individuals on social networks as a cascade. This model shows that individuals are influenced by the contents and information diffusion by others, and this information quickly and widely spreads in social networks. This content diffusion is usually performed due to an individual's psychological and social influences, and it encourages them to transfer and diffuse information influence of others. This model is also useful for checking the effects of content and information propagation in online communities, program marketing strategies, and increasing the influence of social networks. Fig 7. presents the flowchart of the proposed method.

5. IMPLEMENTATION OF THE PROPOSED METHOD

To evaluate the effectiveness and robustness of the proposed method, extensive experiments were conducted using five well-known real-world social network datasets: NetScience, Email, Hamsterster, Ego-Facebook, and Pages-Public Figure, as referenced in [28]. These datasets differ in size, structure, and user interaction patterns, providing a diverse testing ground to verify the generalizability of the proposed approach. Specifically:

- NetScience represents a co-authorship network in scientific publications.
- Email captures email exchanges within a network.
- Hamsterster is a friendship network collected from a pet social website.
- Ego-Facebook contains ego-networks extracted from Facebook profiles.
- Pages-Public Figure includes the connections between verified Facebook pages and public figures.

Each dataset comprises a varying number of nodes (users) and edges (relationships), allowing the method's scalability and adaptability to be tested under different network topologies and densities.

In the simulation process, the proposed method first computes three well-established social importance measures—degree centrality, closeness centrality, and betweenness centrality—for each node within the network. These metrics quantify each user's potential to spread information based on their position and connectivity in the graph.

Subsequently, these centrality measures, individually and in combination, are incorporated into a customized fitness function used by the enhanced Manta-Ray Foraging Optimization (MRFO) algorithm. This fitness function aims to balance two objectives: (1) maximize the overall influence spread and (2) minimize the number of seed users, ensuring an efficient selection of influential nodes.

To initialize the MRFO algorithm, an initial population of candidate seed sets is generated. This population is strategically constrained to reduce computational overhead: candidate users are pre-selected by applying a threshold based on the mean value of a given centrality measure in the network. For example, in the degree centrality scenario, only users whose degree centrality exceeds the network average are considered as potential seeds. This pre-filtering effectively reduces the search space by excluding nodes with minimal influence potential.

During the iterative optimization, the MRFO algorithm explores this reduced solution space. In each iteration, the current population of candidate solutions is evaluated using the bi-objective fitness function. Based on the MRFO's foraging-inspired update rules, a new population is generated by refining the influential user selection to maximize influence spread while maintaining a compact seed set.

This simulation process is repeated for each scenario (degree centrality, closeness centrality, betweenness centrality, and fusion of these measures) across all datasets. The final output is a set of influential users for each network and scenario, along with quantitative results showing how much information spread is achieved relative to other baseline methods.

The parameters in the proposed CMMRFO algorithm have set based on a combination of network structural properties and algorithmic design elements that are inherently sensitive to performance. First, the initial population in the MRFO algorithm was discretized to represent user indices in the social network, with a threshold based on the average degree centrality used to filter users—this ensures that only users with above-average connectivity are considered, improving convergence toward influential nodes. Second, the fitness function incorporates two performance-sensitive parameters: the number of influenced users (spread potential) and the weight of social importance, which is calculated using a weighted combination of degree, closeness, and betweenness centrality. The weights w_1,w_2, and w_3 are bounded between 0 and 1 and directly affect the optimization outcome, making them critical to algorithm sensitivity. Finally, by using the cascade diffusion model, the influence spread is modeled realistically, and the parameterization adapts dynamically to the structure of the network, further enhancing the relevance of parameter choices to actual performance.

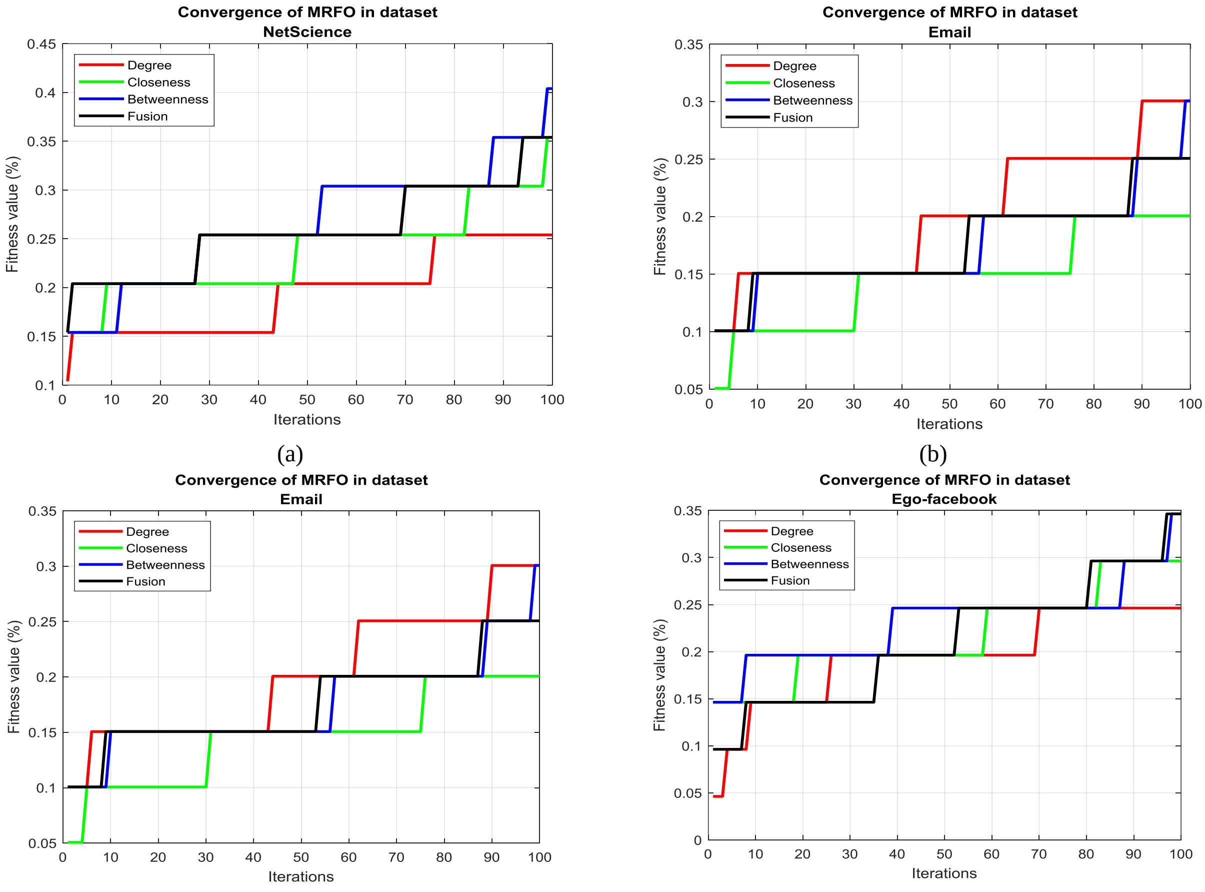
In the proposed CMMRFO algorithm, standard MRFO parameters such as the population size, maximum number of iterations, and exploration-exploitation control mechanisms are carefully set to balance search quality and computational cost. The population size determines how many candidate solutions are explored simultaneously, influencing convergence speed and diversity. The maximum iteration limit ensures that the algorithm stops after a reasonable time while allowing enough search depth. Additionally, MRFO's search operators — including Chain Foraging, Cyclone Foraging, and Somersault Foraging — control how candidate solutions update their positions. These operators are adapted to work with discrete user indices, ensuring effective exploration of the social network space. Proper tuning of these parameters helps achieve an optimal trade-off between exploration and exploitation, directly impacting the influence maximization performance. The main parameters of the proposed method are listed in Table 2.

In addition, the solution having the best fitness function value is selected as the optimal solution. The new population is evaluated based on the fitness function, and the optimal solution is selected. This procedure continues until the stop condition, which is 100 iterations, is met. The last solution related to the last iteration shows that the influenced user's selected generation represents the influencers based on a specified importance criterion. Figs 8 – 12. show a convergence diagram of the optimal point related to each social importance in each data set's MRFO algorithm.

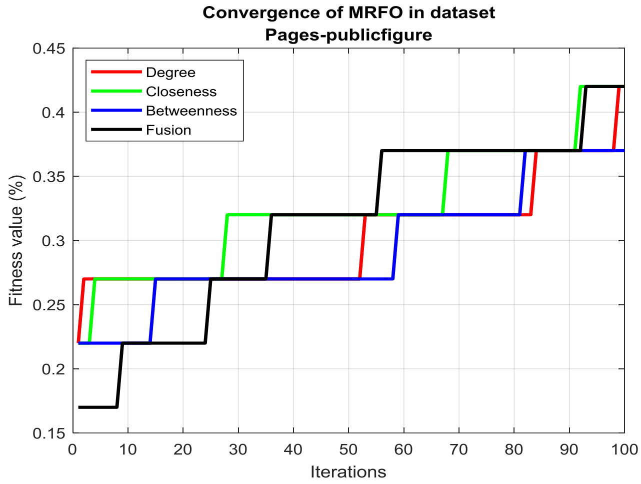
As shown in Figs 8 -12, the convergence diagram of the MRFO algorithm is drawn based on different social importance criteria in the mentioned data sets. As expected, fitness function values are increased in ascending order in each step in the Convergence diagram. Hence, it can be found that the MRFO algorithm does not fall into local traps and converges continuously toward the optimal point. Finally, influenced users are found in a network in the final solution of each scenario. In the following, the proposed method is evaluated.

TABLE 2
Parameters of the proposed method

Parameter	**Value**	**Description**
Initial Population	Users with degree centrality above mean	Limits search to influential users, improving convergence and solution quality.
Discretization Method	Continuous-to-integer mapping	Ensures valid user indices; maintains MRFO compatibility with the discrete domain.
Centrality Weights (w_1 , w_2 , w_3)	[0, 1]	Controls the importance of degree, closeness, and betweenness; balances multiple influence aspects.
Selection Threshold	Average degree centrality	Dynamically adjusts candidate pool size; aligns with network density.
Fitness Function	Spread + penalty for the number of seeds.	Encourages high influence spread with minimum seed users; balances cost and benefit.
Diffusion Model	Cascade model	Realistic simulation of influence propagation; validates optimization results.
Population Size (MRFO)	20–50	Balances exploration diversity and computational cost; affects convergence.
Maximum Iterations (MRFO)	100–500	Defines search depth; more iterations can improve solution quality.
MRFO Operators	Chain, Cyclone, Somersault Foraging	Ensure effective exploration and exploitation; adapted for discrete indices.



(a) (b) (c) (d)



(e)

Fig. 8. Convergence of MRFO based on different social importance criteria in the data set: (a) NetScience, (b) Email, (c) Hamsterster, (d) Ego-facebook, (e) Pages – publicfigure

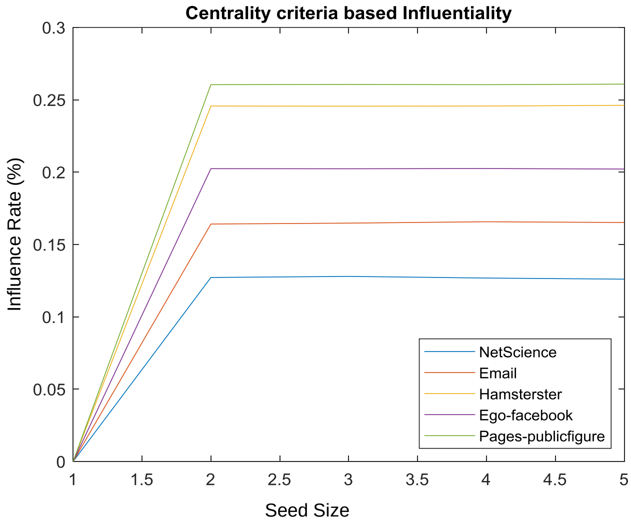


Fig. 9. The influence rate is based on the degree centrality criterion in various data sets

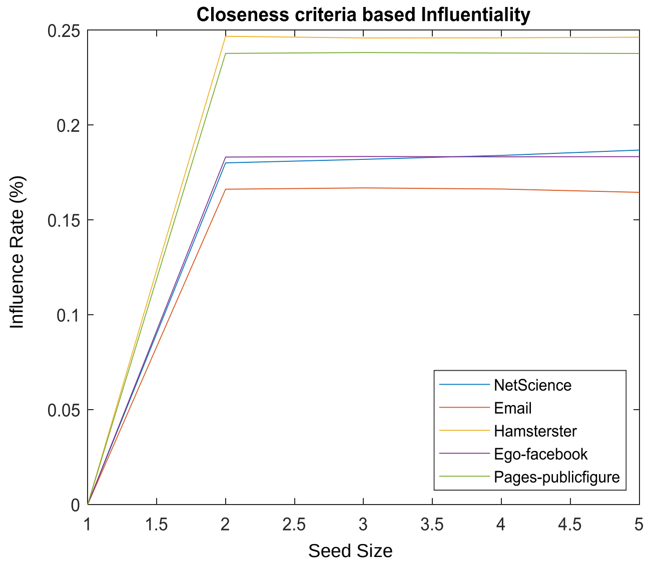


Fig. 10. The influence rate is based on the closeness centrality criterion in different data sets.

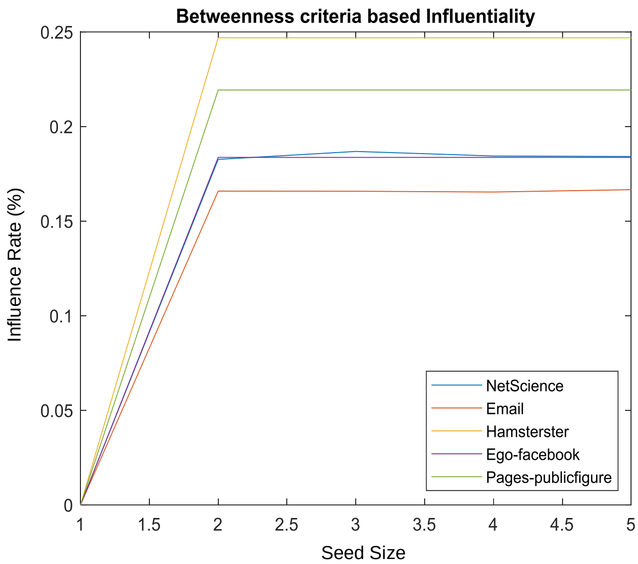


Fig. 11. The influence rate is based on betweenness centrality criteria in different data sets.

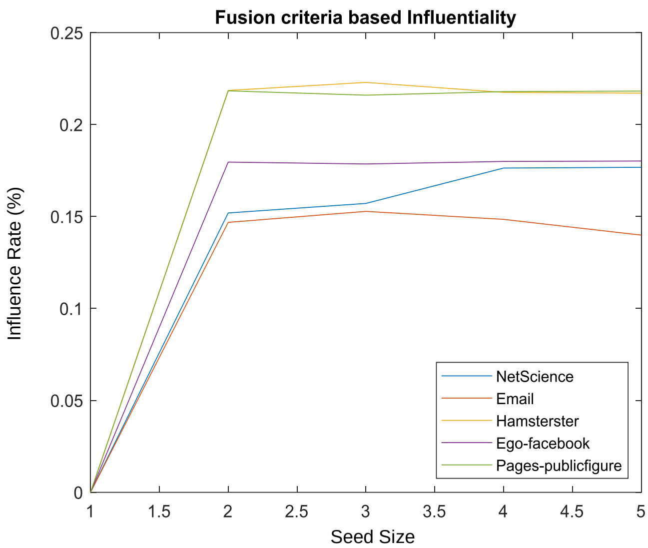


Fig 12. The influence rate is based on a combination of criteria in different data sets

5.1. Evaluation of the proposed method

After implementing the proposed method based on different scenarios, its performance is evaluated. The most conventional criterion is to evaluate the number of users influenced by users in the social network to evaluate the proposed method in terms of IM. In other words, the purpose of IM is to find the users having the most influence among other users.

Therefore, each solution that can find the most influential user can be considered optimal. In the proposed method, the problem of IM is solved by three popular importance criteria, including degree centrality, closeness centrality, betweenness centrality, and a combination of centrality criteria using the CMMRFO algorithm. The proposed method is implemented based on four different scenarios according to importance and combinational importance criteria. In the following, the number of users influenced by influenced users is inspected in each scenario. The number of users influenced by influential users is shown in Table 3. based on social importance criteria in different data sets.

According to Table 1, it can be found that the CMMRFO algorithm influences various users by each social importance criterion for each data set. On the other hand, the proposed IM finds the minimum number of users with the highest influence on the users. The results of IM are more optimal when the effect is high, and the number of influenced users is low. Therefore, the whole number of influenced users is divided by the whole number of users to compute the influence rate for each influenced user. Figs 9-12. show the influence rate of degree centrality, closeness centrality, betweenness centrality, and combination of those centralities in different data sets, respectively.

As is evident in Figs. 9 - 12, different social importance criteria obtain various influence rates. Fig 18 shows a bar graph comparing influence rates based on social importance criteria.

TABLE 3
The number of influenced users in different data sets based on social importance criteria

Datasets	Centrality	Closeness	Betweenness	Fusion criteria
NetScience	41	46	36	40
Email	128	120	124	123
Hamsterster	276	280	271	263
Ego-facebook,	620	601	621	617
Pages - publicfigure	1161	1171	1168	1166

Figes 9 to 12 illustrate how the influence rate varies with the number of selected influential users, commonly referred to as the seed size, for each centrality criterion—degree, closeness, betweenness, and their fusion. In these figures, the horizontal axis denotes the seed size, which ranges from 0 to 5, reflecting the predefined and limited number of influential users chosen by the proposed method. The vertical axis shows the influence rate, defined as the ratio of users influenced by these seeds to the total number of users in the network.

A higher influence rate indicates a more effective spread of information initiated by a small number of well-chosen seeds, demonstrating the strength of the selection strategy. As depicted, the influence rate increases rapidly for the first few seeds, highlighting that the initial influential users have the highest impact on spreading information. Subsequently, the growth slows and stabilizes, as additional seeds contribute incrementally less due to network structure saturation and overlap with already influenced nodes.

This trend across Figes 9 to 12. confirms that the proposed CMMRFO algorithm successfully prioritizes users with the most advantageous network positions for rapid and widespread diffusion. The consistency of this pattern for all centrality measures further validates the robustness and adaptability of the seed selection process in various network conditions.

Regarding the trend illustrated in the plots, it can be observed that the influence rate initially increases rapidly during the first few stages and then gradually converges to an approximately stable value. This pattern arises because, in the proportional function of the proposed CMMRFO algorithm, users with the highest centrality parameters are prioritized as influential seeds. This ensures that users with the most significant connections within the network are selected first, resulting in a steep initial rise in the number of directly influenced users. The first two influential users typically contribute to a large portion of the network being directly affected due to their high connectivity and central position within the network structure.

As the algorithm proceeds to select additional influential users, it ensures minimal overlap with the initially chosen seeds to avoid redundant influence spread. Consequently, each additional user contributes fewer new connections than the initial seeds, leading to a decrease in the growth rate of the influence rate. This behavior explains why, after selecting the first two influential users, the increase in the influence rate diminishes and gradually levels off to a near-constant value. This convergence demonstrates the efficiency and precision of the proposed CMMRFO-based selection strategy in maximizing the spread of influence with an optimal and minimal number of influential users.

Fig 13. illustrates the comparative performance of the proposed CMMRFO algorithm when applying different centrality criteria for selecting influential users. Specifically, this figure highlights how the choice of centrality measure—degree, closeness, betweenness, or their fusion—affects the spread of information within the network. The plotted influence rates demonstrate that selecting seeds based on degree centrality consistently results in a higher influence spread than using closeness, betweenness, or their combination.

This result can be explained by the inherent advantage of degree centrality: nodes with higher degrees have more direct links, allowing information to propagate rapidly to a larger portion of the network in the initial diffusion stages. In contrast, closeness and betweenness centralities, while valuable for understanding network structure, may select nodes that are strategically positioned but have fewer immediate connections, leading to a slower initial spread.

Therefore, Fig 13. validates the design decision in the CMMRFO framework to prioritize degree centrality within its hybrid selection mechanism. By doing so, the algorithm effectively balances structural awareness and computational efficiency, ensuring that the selected influential users trigger faster and broader diffusion compared to other criteria. This empirical comparison further supports the claim that the integration of social network metrics into the MRFO algorithm provides a robust solution for the influence maximization problem in large-scale networks.

As shown, Degree Centrality consistently achieves a higher influence rate than the other measures in various network datasets. This highlights that the CMMRFO algorithm, when guided by the degree centrality parameter, is more effective in identifying highly connected nodes that maximize the spread of information throughout the network.

Including this comparison was necessary to validate why the Degree Centrality parameter was given more weight in our hybrid selection mechanism within the CMMRFO framework. It supports our motivation for favoring nodes with higher direct connections to achieve faster and broader information dissemination.

The critical point is that the maximum number of users influenced by influenced users is selected based on social importance criteria and the CMMRFO algorithm. The purpose is to increase the number of users connected with these users directly and use information diffusion directly as influenced by neighbor users. The influence is maximized when many influenced users select the minimum number of users with maximum influence on other selected users, which is introduced as the optimal method. Therefore, it is necessary to compare the number of users influenced by influenced users in the proposed method and [6] by implementing the moth flame optimization algorithm without using social importance criteria and a fixed number of influenced users. The maximum number of users influenced by the CMMRFO algorithm is shown in Fig 19. and compared with the primary method.

According to Fig 14. it is evident that incorporating social importance criteria—particularly degree centrality—into the MRFO algorithm significantly improves the overall influence spread in social networks. This figure presents a comparative bar chart showing the influence rates achieved by different centrality-based strategies. Among them, the approach that relies on degree centrality consistently outperforms closeness, betweenness, and fusion criteria across multiple datasets. This improvement is attributed to the fact that users with a high degree centrality typically maintain a large number of direct connections, enabling rapid and broad diffusion of information.

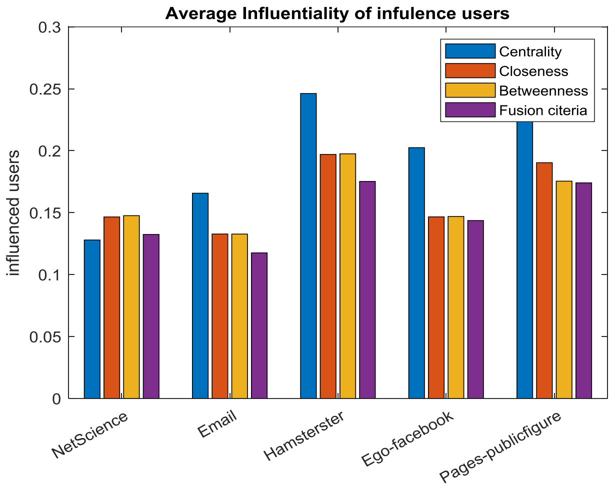


Fig. 13. Bar graph comparing influence rate based on social importance criteria.

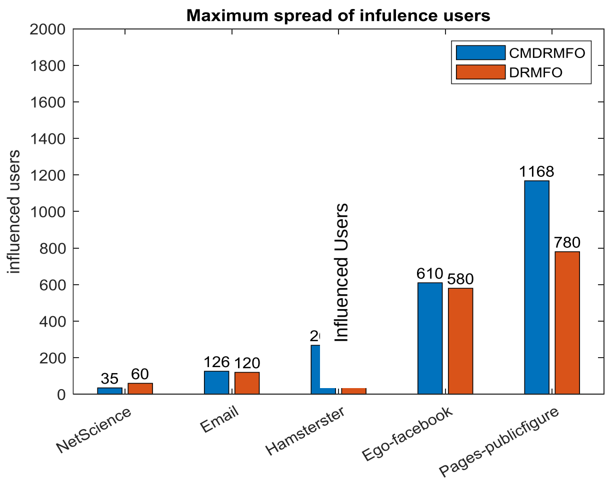


Fig. 14. The comparison of the maximum number of influence users in different data sets

Furthermore, the proposed method not only enhances the spread of influence but does so using a fixed and minimal number of seed users, highlighting its efficiency. By combining optimized fitness function parameters with centrality-based node selection, the CMMRFO algorithm ensures that influence is maximized without redundancy or overlap in the selection of seed users. This results in a higher influence rate, defined as the number of influenced users relative to the total number of seeds, compared to traditional methods. These results confirm that integrating structural characteristics of the network into the optimization process yields a more effective and scalable solution to the influence maximization problem. Fig15. compares the influence rate in the presented method and others.

According to Fig 15. it is evident that the proposed CMMRFO algorithm achieves a significantly higher influence rate compared to the existing baseline method that does not incorporate social importance criteria. This demonstrates the effectiveness of integrating the degree centrality criterion into the fitness function, as it prioritizes highly connected nodes that can spread information more rapidly and extensively. The figure illustrates that for each dataset, the influence rate increases noticeably when the centrality-based approach is used, confirming that selecting seed users based on structural importance leads to a more efficient diffusion process. Moreover, the results highlight that the improvement is most pronounced in networks with highly heterogeneous structures, where influential nodes play a critical role in connecting distant parts of the network. Overall, these findings validate that employing social importance measures, particularly degree centrality, substantially enhances the performance of influence maximization algorithms by ensuring a higher spread of information with fewer seed users.

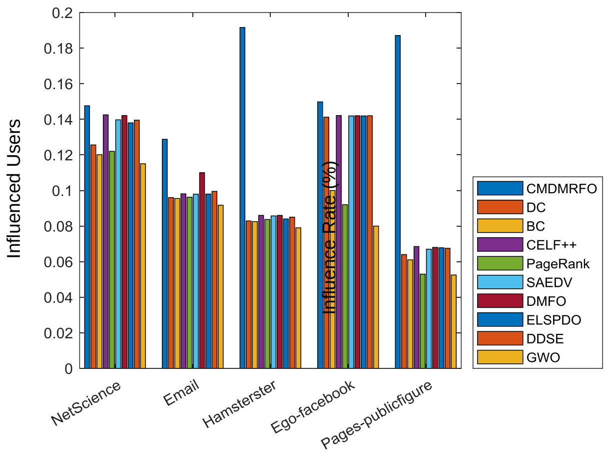


Fig. 15. The comparison of the influence rate in the proposed method and other methods

6. Conclusion and Future Work

IM is the process of selecting a small set of nodes to ensure the quickest and widest information diffusion in social networks. Detecting such nodes remains a crucial research topic with numerous practical applications. While greedy-based methods provide reliable solutions, their high computational cost due to extensive Monte Carlo simulations makes them unsuitable for large-scale networks. In contrast, structural centrality-based approaches offer an efficient alternative by leveraging the inherent properties of network graphs.

In this paper, a novel discrete version of the Manta Ray Foraging Optimization (MRFO) algorithm was developed specifically for the IM problem. The proposed method integrates three prominent centrality measures—degree, closeness, and betweenness—into a fused fitness function and applies innovative discretization techniques to handle the discrete search space effectively.

Experimental evaluations on five real-world social network datasets demonstrated that the proposed CMMRFO algorithm consistently achieves superior influence spread compared to conventional methods. Notably, it yielded average improvements of 14.63% for NetScience, 12.81% for Email, 19.03% for Hamsterster, 15.24% for Ego-Facebook, and 18.76% for Pages-PublicFigure networks. These significant improvements highlight the robustness, scalability, and practical effectiveness of the proposed approach in maximizing influence with a minimal seed set.

The key development lies in designing a bi-objective fitness function that balances two essential goals: minimizing the number of influential (seed) users and simultaneously maximizing their impact. This balance ensures not only a high influence spread but also cost-effective targeting strategies. Furthermore, the integration of fused centrality measures empowers the algorithm to exploit complementary structural information, resulting in a more accurate identification of influential nodes.

In summary, the main contributions of this work can be highlighted as follows:

- A novel discrete MRFO algorithm adapted for the discrete IM problem.
- Introduction of a fused centrality index combining degree, closeness, and betweenness to guide seed selection more effectively.
- Development of a bi-objective fitness function that optimally balances the number of seeds and influences spread.
- Comprehensive evaluation demonstrating significant influence spread improvements over baseline algorithms across various real-world networks.

These contributions together advance the state-of-the-art in influence maximization, especially for large and complex social networks.

For future research, the proposed method can be extended by integrating a community detection step alongside the centrality-based user selection. This hybrid approach would allow meta-heuristic algorithms to identify key influencers within each detected community, combining global and local structural insights. Such an extension is expected to further enhance the efficiency, precision, and adaptability of the model for large-scale and highly modular social networks.

7. References

[1] Z. Aghaee, M. M. Ghasemi, H. A. Beni, A. Bouyer, and A. Fatemi. (2021, Mar.). A Survey on Meta-Heuristic Algorithms for the IM Problem in Social Networks. *Computing*. [Online]. pp. 2437–2477. Available: https://doi.org/10.1007/s00607-021-00945-7

[2] Y. Ye, Y. Chen, and W. Han. (2022, Oct.). IM in social networks: theories, methods and challenges. *Array*. [online]. *16*, p. 100264. Available: https://doi.org/10.1016/j.array.2022.100264

[3] Z. Aghaee and S. Kianian. (2020, Apr.). IM algorithm based on reducing search space in the social networks. *SN Applied Sciences* .[online]. *2*, pp. 1-14. Available: https://doi.org/10.1007/s42452-020-03812-w

[4] P. Dey, A. Chatterjee, and S .Roy. (2019, Jul.). IM in online social network using different centrality measures as seed node of information propagation. *Sādhanā Department of Computer Science and Engineering*. [online]. *44(9)*, pp. 1-13. Available: https://doi.org/10.1007/s12046-019-1189-7

[5] L. Wang, L. Ma, C. Wang, N.-g. Xie, J. M. Koh, and K. H. Cheong. (2021, May.). Identifying influential spreaders in social networks through discrete moth-flame optimization. *IEEE Transactions on Evolutionary Computation*. [online]. *25(6)*, pp. 1091-1102. Available: http://dx.doi.org/10.1109/TEVC.2021.3081478

[6] L. C. Freeman. (1978, Sep.). Centrality in social networks: Conceptual clarification. *Social Networks*. [Online]. *1(3)*, pp. 215–239. Available: https://dx.doi.org/10.1016/0378-8733(78)90021-7

[7] D. Kempe, J. Kleinberg, and É. Tardos. (2003, Aug.). Maximizing the spread of influence through a social network. Presented at Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 137–146. [Online]. Available: https://doi.org/10.1145/956750.956769

[8] S. Brin and L. Page. (1998, Apr.). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*. [Online]. *30(1–7)*, pp. 107–117. Available: https://dx.doi.org/10.1016/S0169-7552(98)00110-X

[9] Q. Jiang, G. Song, C. Gao, Y. Wang, W. Si, and K. Xie. (2011, Aug.). Simulated annealing based influence maximization in social networks. Presented at Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence. [Online]. 25(1), pp. 127–132. Available: https://doi.org/10.1609/aaai.v25i1.7838

[10] A. Şimşek and R. Kara. (2018, Dec.). Using swarm intelligence algorithms to detect influential individuals for influence maximization in social networks. *Expert Systems with Applications*. [Online]. *114*, pp. 224–236. Available: https://doi.org/10.1016/j.eswa.2018.07.038

[11] Y. D. Navaei, M. H. Rezvani, and A. M. E. Moghaddam. (2024, Feb.). A novel neighborhood-based importance measure for social network influence maximization using NSGA-III. Presented at 2024 10th International Conference on Artificial Intelligence and Robotics (QICAR), pp. 113-118. [Online]. Available: https://doi.org/10.1109/QICAR61538.2024.10496642

[12] L. Cui, H. Hu, S. Yu, Q. Yan, Z. Ming, Z. Wen, and N. Lu. (2018, Feb.). DDSE: A novel evolutionary algorithm based on degree-descending search strategy for influence maximization in social networks. *Journal of Network and Computer Applications*. [Online]. *103*, pp. 119–130. Available: https://doi.org/10.1016/j.jnca.2017.12.003

[13] J. Tang, R. Zhang, Y. Yao, F. Yang, Z. Zhao, R. Hu, and Y. Yuan. (2019, Jan.). Identification of top-k influential nodes based on enhanced discrete particle

swarm optimization for influence maximization. *Physica A: Statistical Mechanics and Its Applications*. [Online]. 513, pp. 477–496. Available: https://doi.org/10.1016/j.physa.2018.09.040

[14] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance. (2007, Aug.). Cost-effective outbreak detection in networks. in Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 420-429. [online]. Available: https://doi.org/10.1145/1281192.1281239

[15] S. Mokhtarizadeh, B. Zamani Dehkordi, M. Mosleh, and A. Barati. (2021, Oct.). Influence maximization using time delay based harmonic centrality in social networks. *Tabriz Journal of Electrical Engineering*. [Online]. pp. 359–370.

[16] L. Han, Q. Zhou, J. Tang, X. Yang, and H. Huang. (2021, Feb.). Identifying top-k influential nodes based on discrete particle swarm optimization with local neighborhood degree centrality. *IEEE Access*. [Online]. *9*, pp. 21345–21356. Available: https://doi.org/10.1109/ACCESS.2021.3056087

[17] A. Arora, S. Bansal, C. Kandpal, R. Aswani, and Y. Dwivedi. (2019, Jul.). Measuring social media influencer index—insights from Facebook, Twitter and Instagram. *Journal of Retailing and Consumer Services*. [Online]. *49*, pp. 86–101. Available: http://dx.doi.org/10.1016/j.jretconser.2019.03.012

[18] B. Fu, J. Zhang, W. Li, M. Zhang, Y. He, and Q. Mao. (2022, Jul.). A differential evolutionary influence maximization algorithm based on network discreteness. *Symmetry*. [Online]. *14(7)*, p. 1397. Available: https://doi.org/10.3390/sym14071397

[19] B. Chatterjee, T. Bhattacharyya, K. K. Ghosh, A. Chatterjee, and R. Sarkar. (2021, Feb.). A novel meta-heuristic approach for influence maximization in social networks. *Expert Systems*. [Online]. p. 12676. Available: https://doi.org/10.1111/exsy.12676

[20] Tarun K. Biswas , A. Abbasi, Ripon, K. Chakrabortty , "An MCDM integrated adaptive simulated annealing approach for influence maximization in social networks", vol.556, pp. 27-48, (2021). Availabe: https://doi.org/10.1016/j.ins.2020.12.048

[21] H. Li, R. Zhang, Z. Zhao, X. Liu, and Y. Yuan. (2021, Nov.). Identification of top-k influential nodes based on discrete crow search algorithm optimization for influence maximization. *Applied Intelligence*. [Online]. *51(11)*, pp. 7749–7765. Available: http://doi.org/10.1007/s10489-021-02283-9

[22] R. Cantini, F. Marozzo, S. Mazza, D. Talia, and P. Trunfio. (2021, Sep.). A weighted artificial bee colony algorithm for influence maximization. *Online Social Networks and Media*. [Online]. *26*, p. 100167. Available: https://doi.org/10.1016/j.osnem.2021.100167

[23] A. Karczmarczyk, J. Jankowski, and J. Wątrobski. (2021, Jun.). Multi-criteria seed selection for targeted influence maximization within social networks. Presented at International Conference on Computational Science (ICCS 2021), pp. 454–461. [Online]. Available: https://doi.org/10.1007/978-3-030-77967-2-38

[24] L. Han, K. C. Li, A. Castiglione, J. Tang, H. Huang, and Q. Zhou. (2021, Apr.). A clique-based discrete bat algorithm for influence maximization in identifying top-k influential nodes of social networks. *Soft Computing—A Fusion of Foundations, Methodologies & Applications*. [Online]. *25(13)*, pp. 8223–8240. Available: https://doi.org/10.1007/s00500-021-05749-7

[25] W. Zhao, Z. Zhang, and L. Wang. (2020, Jan.). Manta ray foraging optimization: An effective bio-inspired optimizer for engineering applications. *Engineering Applications of Artificial Intelligence*. [Online]. *87*, p. 103300. Available: https://doi.org/10.1016/j.engappai.2019.103300

[26] S. Wang, Y. Wang, Y. Wang, and Z. Wang. (2022, Oct.). Comparison of multi-objective evolutionary algorithms applied to watershed management problem. *Journal of Environmental Management*. [Online]. *324*, p. 116255. Available: https://doi.org/10.1016/j.jenvman.2022.116255

[27] R. R. Singh. (2021, Aug.). Centrality measures: A tool to identify key actors in social networks. In *Principles of Social Networking: The New Horizon and Emerging Challenges*. [Online]. pp. 1–27. Available: https://doi.org/10.1007/978-981-16-3398-0-1

[28] R. Rossi and N. Ahmed. (2015, Jan.). The network data repository with interactive graph analytics and visualization. in Proceedings of the AAAI conference on artificial intelligence. [online]. Available: https://doi.org/10.1609/aaai.v29i1.9277

[29] S. Ghafori and F. S. Gharehchopogh. (2022, Jan.). A multiobjective Cuckoo Search Algorithm for community detection in social networks. In *Multi-objective Combinatorial Optimization Problems and Solution Methods*. [Online]. pp. 177–193. Available: https://doi.org/10.1016/B978-0-12-823799-1.00007-3

[30] A. Heydariyan, F. S. Gharehchopogh, and M. R. E. Dishabi. (2024, Jul.). A hybrid multi-objective algorithm based on slime mould algorithm and sine cosine algorithm for overlapping community detection in social networks. *Cluster Computing*. [Online]. *27(10)*, pp. 13897–13917. Available: https://doi.org/10.1007/s10586-024-04632-y

[31] J. Sheykhzadeh, B. Zarei, and F. S. Gharehchopogh. (2024, Jul.). Community detection in social networks using a local approach based on node ranking. *IEEE Access*. [Online]. *12*, pp. 92892–92905. Available: https://doi:10.1109/ACCESS.2024.3420109

[32] S. M. Aghdam, F. S. Gharehchopogh, and M. Masdari. (2024, Mar.). A hybrid approach in opinion leaders selection using African vultures optimization and hunger games search algorithms. *Social Network Analysis and Mining*. [Online]. *14(1)*, p. 60. Available: https://doi.org/10.1007/s13278-024-01228-7

[33] S. M. Aghdam, F. S. Gharehchopogh, and M. Masdari. (2025, Apr.). An opinion leader selection in social networks using hybrid amended salp swarm and improved grey wolf optimizer algorithms. *International Journal of Data Science and Analytics*. [Online]. pp. 1–26. Available: https://doi.org/10.1007/s41060-025-00760-9