Evaluating Centrality-Based Seed Node Strategies for Influence Diffusion in OSNs: A Study across SCC, WCC and Full Networks using SIR, LT and IC Diffusion Models

Halima Baabcha[1, 3*], Meriem Laifa[1, 3] and Samir Akhrouf [2,3]
[1] Department of Computer Science, University Mohammed El Bachir El Ibrahimi, Bordj Bou Arreridj , Algeria
[2] LIAM Laboratory, University Mohamed Boudiaf , M’sila, , Algeria
[3] Department of Computer Science, University Mohamed Boudiaf, M'sila, Algeria
Email: halima.baabcha@univ-bba.dz, meriem.laifa@univ-bba.dz, samir.akhrouf@univ-msila.dz
*Corresponding author

Keywords: centrality measures, influential nodes, influence diffusion, SIR, LTM, ICM, SCC, WCC and viral marketing

Received: May 11, 2025

This study investigates influence diffusion in online social networks (OSNs) through a comprehensive analysis of centrality measures and diffusion models using the Higgs Twitter dataset. We model OSNs as directed graphs, focusing on strongly connected components (SCCs) and weakly connected components (WCCs). Seven centrality measures (out-degree, in-degree, betweenness, closeness, eigenvector, PageRank, and Katz centrality) are calculated to identify key influential nodes. The top-ranked nodes are then subjected to influence diffusion simulations using three models: Linear Threshold (LT), Independent Cascade (IC), and Susceptible-Infected-Recovered (SIR) across three types of activity networks with different structural characteristics. Our findings reveal significant variations in centrality performance depending on network topology and diffusion dynamics. This methodology integrates structural network analysis with dynamic diffusion modeling to evaluate the effectiveness of influence spread. The experimental results show that out-degree and betweenness centralities are most effective for influence propagation, with the SIR model supporting sustained diffusion. The experimental results reveal that out-degree and betweenness centralities are the most effective measures for influence propagation, with out-degree being particularly impactful for initiating diffusion. The SIR model demonstrated superior efficacy for sustained influence spread, aligning more closely with real-world influence dynamics. Additionally, analyzing influence propagation within WCCs enables more computationally efficient identification of key influencers, without significant loss in accuracy. This work offers actionable insights for influence modeling and provides a practical methodology for selecting centrality measures tailored to specific diffusion scenarios. It explores the influence diffusion across different platforms, enabling researchers to assess and compare user impact by offering a detailed examination of network structures, key node significance and influence diffusion.

Povzetek: Študija na usmerjenih grafih združi več centralnostnih mer z difuzijskimi modeli kot metodo za izbiro ključnih vplivnežev in prilagojeno oceno širjenja vpliva glede na topologijo in dinamiko.

1 Introduction

In the modern digital era, OSNs have evolved into dynamic ecosystems that facilitate seamless communication, large-scale information exchange, and interactive social engagement. With billions of users actively participating daily, OSNs generate massive volumes of data, providing invaluable opportunities for researchers to investigate user behavior [1], social influence analysis [2], and the complex mechanisms of information diffusion [3]. These analyses are crucial for a variety of applications, such as enhancing public health interventions [4], mitigating the spread of misinformation [5], detecting communities [6], and optimizing viral marketing strategies [7].

Essentially, Viral Marketing (VM) offers an effective advertising strategy for commercial companies via OSNs, where companies try to promote their services and products through word-of-mouth propagation among friends or followers.

A major challenge in VM is identifying key users who wield significant influence within networks. These key nodes have a significant impact in determining the efficiency of information propagation, making them valuable for both accelerating and inhibiting the spread of messages. Various centrality metrics including degree centrality, PageRank, closeness, eigenvector, and Katz centrality are widely used to identify critical nodes that maximize influence propagation [8][9]. Each centrality measure presents unique methodologies for ranking users based on their potential influence. The efficiency of these measures depends significantly based on the nature of the interaction and the underlying network structure.

However, despite their widespread adoption, the comparative effectiveness of centrality measures remains an open research question, with diverse interpretations and applications across different network structures. This complexity arises because centrality metrics can yield differing node importance rankings, depending on network topology.

This study advances the discourse on influence maximization by providing a extensive evaluation of centrality measures in the context of OSNs. By elucidating the relationship between network structure, node importance, and information diffusion, our research lays a foundation for future investigations into evolving user behaviors, temporal dynamics, and adaptive strategies for influence propagation in ever-changing digital landscapes. Furthermore, this study serves as a valuable starting point for new researchers in this area, offering a structured framework that can be extended through the develop of hybrid centrality measures or novel diffusion models to better modeling the complexity of real-world social networks.

To provide clear objectives, this study is guided by the following research questions :

RQ1: Which centrality measures (in-degree, out-degree, betweenness, closeness, eigenvector, PageRank, and Katz) perform best for maximizing influence under different diffusion models (SIR, IC, and LT) ?

RQ2: How do structural properties of networks particularly SCCs,WCCs, and full networks affect the effectiveness of diffusion?

RQ3: How consistent are the rankings of top-k influential nodes across SCC, WCC, and full networks when evaluated with different centrality measures?

RQ4: How do the diffusion models (LT, IC, and SIR) differ in terms of coverage, speed, and sustainability of influence spread when applied to centrality-based seed selection?

The rest of this paper is organized as follows: Section 2 presents a literature review, while Section 3 introduce key definitions and models used in this study. The proposed methodology is detailed in Section 4. Results and discussion are presented in section 5. Section 6 discusses the implications of our study. Finally, this paper is concluding in Section 7.

2 Literature review

Influence Maximization (IM) is a significant problem in complex networks and a key step in solving this problem is identifying influential nodes that can maximize the information or influence propagation. Centrality measures are widely used to identify the influential nodes in complex networks. Here, we provide a brief literature review on the influential node's detection using centrality measures, and a comparison of information diffusion using different models and centrality measures as seed nodes.

For a more detailed review, we refer the reader to our previous work [10]. In particular, our earlier work presented a survey of social influence analysis in OSNs, with a particular emphasis on viral marketing applications. That study reviewed the most recent and important methods for influence modeling, influence maximization, and influential node detection. It aimed to guide novice researchers by synthesizing leading works and identifying open challenges in this area. However, was primarily conceptual and did not provide empirical validation of centrality measures or diffusion models. The present paper addresses this gap by moving beyond a survey to a systematic experimental evaluation of seven centrality measures under three diffusion models (LT, IC, and SIR), tested across SCC, WCC, and full network structures. This transition from theoretical synthesis to quantitative comparison represents a key extension of our prior work.

Authors in [11] present a study of social networks, models of OSNs and different models used for identiying influential users in OSN. They summarized the common concepts and terms used in describing OSN, such as connectivity, diameter, strongly connected components (SCC) and weakly connected components (WCC). They reviewed principal centrality measures such as in-degree, out-degree, eigenvector, closeness and betweenness centrality then highlighted the role of centrality for the detection of influential nodes. Furthermore, they discussed and analyzed different information diffusion models, such as Linear Threshold (LT), Independent cascade (IC) and Susceptible-Infected-Recovered (SIR) models with detailing their mechanisms of node activation and implications for the identification of influential nodes. Then, they reviewed some significant influential node identification techniques.

Additionally, Singh et al. [12] presented a comprehensive survey about information diffusion in social networks. They discussed various metrics and methods to analyze influence and its propagation models. They started with various concepts and centrality measures such as degree, closeness and betweenness centrality, density and degree distribution etc, which were important for understanding the structural roles of actors within social networks and their potential impact on information propagation. Further, authors discussed quantitative and qualitative analysis of social influence and influence maximization. Finally, they discussed various influence propagation models, particularly the IC, LT models, epidemic models and so on with their special variants.

Arrami et al. [13] presented a study whose primary goal was to review different works that aimed to select opinion leaders in OSNs, contributing significantly to understanding their impact on consumer behavior and marketing strategies. They explored various opinion leader detection methods, then classified them into two categories: centrality techniques and maximization techniques. Furthermore, they provided the advantages of the centrality measures resulting from the analysis of OSNs and their limits.

Another study focused specifically on identifying and analyzing the measures, approaches, and models used to measure user influence on OSNs through a comprehensive literature review [14]. Authors conducted a literature review of 25 studies published between 2014 and 2020, identifying 21 influence metrics, 4 algorithms, and 8 diffusion models. They categorized approaches into centrality-based, diffusion-based, walk-based, link-based, and popularity-based measures, and also examined Twitter-specific metrics such as Retweet Impact and FollowerRank. The study highlighted that LT/IC models and greedy algorithms are inefficient for large-scale networks, while heuristic methods improve scalability. Although this work offered a broad theoretical overview and emphasized challenges, it did not provide empirical comparisons of centrality measures. Our work complements this by experimentally evaluating seven centrality measures under three diffusion models (LT, IC, and SIR) across SCC, WCC, and full networks.

On the other hand, Singh [15] conducted a review of centrality measures applied to social network data analysis and key nodes identification. The study began with examining the fundamental notion of centrality measures with a particular attention to their roles in social network analysis. The author then provided a concise overview of computational algorithms for these measures, followed by an exploration of various research directions related to centrality metrics. Additionally, this review summarized several applications of differents centrality measures for analyzing real-world OSNs. While the study highlights the importance of centrality in ranking nodes based on their significance in various applications, it lacks a practical component in the study of influence diffusion models and the effectiveness of centrality measures in large-scale or sparse networks.

Regarding information diffusion models, Yujie [16], presented a comprehensive survey about several classic information diffusion models in OSN, like an explanatory model: the Suspected-Infected (SI) model, the Suspected-Infected-Suspected (SIS) model, and the Suspected-Infected-Recovered-Suspected (SIRS) model; predictive model: IC, LT models, emphasizing the critical role these networks play as data dissemination platforms. Then, they discussed some applications ofthose information propagation models in different social networks. Furthermore, the study highlights the practical applications of these models across various social networks, illustrating their importance in managing information propagation and mitigating misinformation.

The absence of a practical component in the study of influence diffusion models and the effectiveness of centrality measures in different network structures presents significant weaknesses. Additionally, the relationship between centrality measures and information diffusion models is not adequately specified, potentially leading to confusion about how these approaches can complement each other in identifying effective influencers. Furthermore, these works do not clarify the implications of different network structures on the applicability of these models, which could impact the accuracy of influencer identification and engagement strategies.

Existing studies often focus on either (i) ranking influential nodes using centrality measures without validating their actual spreading potential [11][14] or (ii) analyzing information diffusion without systematically comparing different seed selection strategies [17]. Additionally, many evaluations have been conducted on one type of network structure without considering variations in SCC and WCC. To address these gaps, this study provides a comprehensive evaluation of centrality measures as a seed selection method to maximize influence diffusion. We analyze how different centrality-based strategies perform in replies, mentions, and retweet networks and compare their effectiveness in maximizing influence spread under IC, LT, and SIR models. By combining structural network analysis with diffusion simulations, this study offers valuable insights into the relative strengths and weaknesses of different centrality measures in OSN.

3 Background

3.1 Definitions

An OSN can be modeled as a weighted graph $G = (V, E, w)$, where V is a set of nodes $(v_1, v_2, ..., v_i)$ representing users, E is a set of edges $(e_1, e_2, ..., e_j)$ representing social interactions, and w denotes the edge weights. In this network representation, V corresponds to actors (individuals), while E captures the relationships between them, such as replies, mentions, and retweets. For instance, if an edge e_j exists between nodes v_m and v_n, it signifies a connection or association between the corresponding actors.

The primary goal of this study is to analyze information diffusion in OSNs. To achieve this, we examined various types of graphs, including SCC, WCC, and complete complex networks. Addressing this issue requires reviewing several definitions of influence metrics within OSNs. This study specifically focuses on the following definitions:

3.1.1 Influential users

Top-k users are the k users that are capable of generating maximum influence and widest information propagation to their connected users in an OSN, and can be characterized as a person who has the power to affect people, actions, or events. The top k nodes in an OSN are crucial for understanding the structure and evolution of the network and can be used to detect influential users, optimize information diffusion, and improve the overall performance of the network [18].

3.1.2 Strongly connected components (SCC)

This is a maximal subset of nodes in a directed graph such that every node is reachable from every other node within that subset. This means that for any two nodes v_i and v_j in the SCC, there exists a directed path from v_j to v_i and a directed path from v_i to v_j. SCCs are crucial for understanding the structure and connectivity of directed graphs, with applications in several fields such as OSN analysis [19].

3.1.3 Weakly connected components (WCC)

Is a maximal set of nodes $C \subseteq VC$ such that for every pair of nodes there is a path between v_i and v_j if the edge direction is ignored. Equivalently, if a directed graph is transformed into an undirected graph by treating every directed edge as an undirected edge, the connected components of this undirected version are WCCs [20].

3.2 Centrality measures

In this section, the most important metrics are introduced. In graph theory, centrality is defined as a measure of the importance of a given node in a graph. The problem of finding the most influential users in OSNs is, in the end, a measure of importance. Centrality has attracted the most attention in the early years of influencer identification [21]. For clarity, all the abbreviations used in this paper are listed in Table 1.

3.2.1 Degree centrality

This ranks users with more connections higher in terms of centrality; in other words, it measures the total number of connections a user has with other users. However, it does not indicate the frequency of communication and is often a local maximum for network measures [22]. In an undirected graph, the degree centrality C_d for user v_i in an undirected graph is:

$$C_d(v_i) = d_i \qquad (1)$$

Where d_i is the degree (number of adjacent edges) of node v_i. In a directed graph, there are two types of degree centrality.

In-degree centrality: It refers to the number of edges pointing inwards at a node, usually refers to the popularity of a user [22]. We formulate:

$$C_d(v_i) = d_i^{in} \qquad (2)$$

Out-degree centrality: The out-degree of a node refers to the number of edges that lead it out. We formulate:

$$C_d(v_i) = d_i^{out} \qquad (3)$$

And consider users with more connections to be more important users [23]. Users with a large number of connections are linked to other users in the graph. Such a node is considered important if it has many neighbors because it has a higher likelihood of intercepting or capturing whatever flows are resources or information through the network [22].

3.2.2 Betweenness centrality

Betweenness centrality [22] indicates the capability of a user for faster transfer of information through the graph. As is the case for edges with high betweenness, users with high betweenness occupy critical positions in the graph structure and are therefore able to play critical roles [24]. This is often enabled by the large amount of flow carried by users that occupy a position at the interface of tightly knit groups. Betweenness centrality measures the amount of network flow that a given node controls, while this measure gives the volume by traffic that passes through a node in the network [23] [25]. However, this metric can measure the possibility of infection by the traffic that passes through a node and controls the diffusion of infection to other nodes in the network. The betweenness centrality of node v is given by :

$$B(v) = \sum_{u \neq v \neq w} \frac{\sigma_{u,w}(v)}{\sigma_{u,w}} \qquad (4)$$

Where $\sigma u, w (v)$ refers to the total number of shortest paths connecting u and w that pass-through v and $\sigma u, w$ is the total number of shortest paths from u to w.

3.2.3 Closeness centrality

This is a measure of the importance of a node in the graph. It captures the average distance between one node and every other node in a graph. The fact behind Closeness centrality indicates that the more central the user, the faster it can reach other nodes in the graph [26]. A user with a higher closeness centrality is more prominent within the network, as they can reach all other nodes more quickly. This metric is computed using the following formula:

$$C_c(u) = (N - 1) / \Sigma\, d(u, v) \qquad (5)$$

Where $d(u, v)$ is the shortest path distance between the nodes u and v. In closeness centrality, the intuition is that the more central the nodes, the more quickly they can reach other nodes. Formally, these nodes should have a smaller average shortest path length than other nodes.

3.2.4 Eigenvector centrality

Is a measure of a node's influence in a network, based on the idea that a node's importance depends on its neighbors' importance. It is defined for both directed and undirected graphs [27]. This measure is calculated using the following equation:

$$C_e(v_i) = \frac{1}{\lambda} \sum_{j=1}^{n} A_{j,i} C_e(v_j) \qquad (6)$$

We can use the adjacency matrix A of the graph. Let $Ce(v_i)$ denote the eigenvector centrality of node v_i. We want the centrality of v_i to be a function of its neighbors' centralities. It is posited that this is proportional to the summation of the centralities. In the adjacency matrix A of the network, Ai,j indicates the connection (and possibly the strength of that connection) between nodes i and j, and λ is a constant (the leading eigenvalue of A). Nodes connected to other highly central nodes received higher scores. Eigenvector centrality captures the number of connections a node has and crucially, the influence of these connections. This principle underlies other centrality measures such as PageRank, which adopts the same concept for ranking web pages [25]. Previous research has used eigenvector centrality to detect influential nodes in a graph [28][29].

3.2.5 Katz centrality

Calculates the influence of users (nodes) by considering all network paths. This is a measure of node influence in a network that extends the concept of eigenvector centrality by incorporating both the number of immediate neighbors and the importance of more distant neighbors. Unlike pure eigenvector centrality, it assigns a small amount of centrality to each node as a baseline, and then propagates influence through the network; attenuating contributions by a factor of α (often called the attenuation or damping factor). The Katz centrality considers all network links [30]. The main difference between Katz and closeness centralities is that the former assigns a certain minimum score to every user in the network [31]. While, Katz has high computational complexity of centrality, which limits its application to large networks [30].
Formally, for a node *i*, its Katz centrality *xi* is given by:

$$C_{katz}(v_i) = \alpha \sum_{j=1}^{n} A_{j,i} C_{katz}(v_j) + \beta \qquad (7)$$

Where:

- $A_{i,j}$ is the *(i,j)* entry of the adjacency matrix *A*.
- α is a constant selected such that $\alpha < \frac{1}{\lambda_{max}}$, where $\lambda\ max$ is the largest eigenvalue of *A*. This condition ensures convergence.
- β is a constant "bias" or the baseline score received by each node.

In essence, Katz centrality not only counts direct connections but also indirect connections of every length, weighting longer paths by decreasing the power of α. A node is considered more central if it has many neighbors and/or it is connected to other central nodes [23].

3.2.6 Page rank centrality

Is an algorithm proposed by Page and Brin, the co-founders of Google, to rank web pages [32]. It is a widely known measure for ranking web pages based on their importance in the web, which is considered as a graph. It counts the quality of the links to a node to determine the prominence of the node. This measure can be calculated using the following equation.

$$PR(i) = \frac{(1-d)}{n} + d * \sum_{j \in \Gamma(i)} \frac{PR(j)}{C(j)} \qquad (8)$$

Where *PR(i)* and *PR(j)* represent the PageRank values of nodes *i* and *j*, C(j) represents the out degree of the node, *d* is the damping factor (usually 0.85), *n* is the total number of web pages, and *Γ(i)* is the set of neighboring nodes of node *i* [33].

3.3 Information diffusion models

Information diffusion models are mathematical and computational frameworks used to describe and analyze how information, ideas, behaviors, or innovations spread through a network of individuals or entities. These models are widely used in OSNs analysis, epidemiology, marketing, and communication studies to understand the mechanisms governing influence propagation. Typically, these models consider factors such as influence probability, network structure, and individual adoption thresholds. The most common categories include the IC, LT, and SIR models. These models can help researchers and practitioners predict trends, optimize viral marketing campaigns, and mitigate misinformation spread in networks [34][35]. Information diffusion models typically involve three primary components: the sender, who initiates the process; the receiver, who receives the information; and the medium, which facilitates the transmission of information. The diffusion process can be influenced by differents parameters such as user characteristics, network structure, and external factors, which can affect the spread of information [36]. For more details on this, check [17]. An overview of these models is presented in the reminder of this section.

3.3.1 Susceptible-infected-recovered (SIR) model

Is a well-known approache for evaluating centrality measures in OSNs analysis. The epidemiological model simulates the diffusion of a virus within a network, classifying nodes into three distinct categories [10] (as illustrated in Figure 1):

- *Susceptible (S) nodes: I*s unaware of the information diffusion within the graph. These nodes are uninfected but not immune, and can be contaminated by neighboring infected nodes. Initially, all nodes were considered susceptible, except for the source node.
- *Infected (I) nodes*: An infected node has acquired and is aware of the information spreading through the network, and it actively shares this information with its neighboring nodes. After a specific time period, the infected node moves to the recovered state, with this transition governed by the infection probability β at each time step and the recovery probability λdt over a time interval dt. The average duration a node remains infected is denoted by *D*.
- *Recovered (R) nodes:* The recovered nodes lose interest in the information and no longer spread it. They also become immune to further infection. At the end of the process, only susceptible and recovered nodes remain in the network.

The dynamics of the *SIR* model are governed by a set of ordinary differential equations that describe the transitions between these states. The total population in the network is denoted by *N*, and at any time *t*, the sum of susceptible, infected, and recovered nodes equals N. The SIR system can be described using the following ordinary differential equations:

$$\frac{dS}{dt} = \frac{\beta IS}{N}$$

$$\frac{dI}{dt} = \frac{\beta IS}{N - \lambda I} \quad \textbf{1.1}$$

$$\frac{dR}{dt} = \lambda I$$

$$S(t) + I(t) + R(t) = N \quad (9)$$

The SIR model extends beyond traditional epidemiological contexts to the analysis of the information spread in social networks, providing insights into the influence exerted by various nodes based on their centrality measures [23].

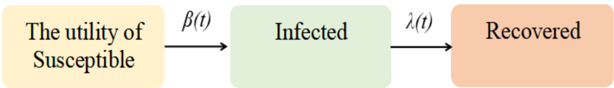

Figure 1: Illustration of SIR model.

3.3.2 Linear threshold model (LTM)

In this model, every edge e(u, v) carries a weight W(u, v), with the total weight of all incoming edges to node v being less than or equal to 1, and each node *v* is also associated with a threshold θ_v. The LT model starts with some active nodes, with all other inactive nodes, and a random choice of thresholds θ. The LTM samples the value of *v* for each user *v* uniformly at a random probability from [0,1]. At step 0, nodes in the seed set S are marked as active, while all other nodes are initialized as inactive. The model then proceeds iteratively, updating each node's status. In step *t*, all nodes that were active in step *t-1* stay active, and any inactive node *v* at step *t-1* may become active. The influence spread of the seed set *S* under the LT model, denoted as *σ(S)*, represents the expected number of nodes that become active when *S* is activated initially. For more details on this model, check [10] [37].

3.3.3 Independent cascade model (ICM)

In ICM, a probability *p (u,v)* is associated with each edge *e(u,v)* where *u* and *v* are two nodes in the graph. *p (u,v)* is the probability that *u* succeeds in activating *v*. In this model, node *v* is independently activated by each of its incoming neighbors by introducing an influence probability *p (u,v)* to each edge *e(u,v)*. Based on the influence probabilities and given a seed set S at time step 0, a diffusion instance of the IC model unfolds in discrete steps. Each active node *u* at step *t* activates each of its outgoing neighbors *v* which is inactive in step *t-1* with probability *p(u,v)*. The activation process can be considered as flipping a coin with head probability *p (u,v)*: If the result is heads, then *v* is activated; otherwise, *v* remains inactive. The diffusion instance terminates when no additional nodes are activated. The influence spread of seed set *S* under IC is the expected number of activated nodes when *S* is the initial active node set and the above stochastic activation process is applied. For more details on this model, check [10] [38].

4 Methodology

Our methodology enables an in-depth analysis of how centrality-based node selection affects influence spread across networks. By systematically partitioning the network into its WCC and SCC, we can assess how structural properties impact influence propagation. As illustrated in Figure 2, the proposed methodology consists of three main steps.

Table 1: Abbreviations

Node Id	Value
E	Set of edges between users
V	Set of users in the graph
w	Weights associated with edge in E
Indg	In-degree centrality
Outdg	Out-degree centrality
Cl	Closeness centrality
Bt	Betweenness centrality
Ev	Eigenvector centrality
Pr	Pagerank
Kz	Katz centrality
WCC	Weakly connected components
SCC	Strongly connected components
SIR	Suspected - Infected - Recovered
LTM	Linear threshold model
ICM	Independent cascade model
S	Seed nodes
G (V, E)	Graph
N	Total number of nodes in the graph

4.1 Step 1: Initialization and SCC, WCC identification

In this initial step, we decomposed the graph based on the WCC and SCC for each network. For SCC, we used Tarjan's [39][40] algorithm which is a Depth-First Search (DFS)-based approach used to find SCCs in a directed graph. It was introduced by Robert Tarjan in 1972 [41] and is known for its efficiency with a time complexity of *O (V+E)*, where *V* is the number of vertices, and *E* is the number of edges. The algorithm is based on DFS traversal and utilizes a low link value to track the smallest reachable node from a given node. It employs a stack-based approach to store nodes in the current SCC and detects cycles and back edges, which helps to identify SCCs efficiently [42] [43]. Detecting SCCs in a directed graph is essential for understanding information diffusion, as SCCs represent regions where information can circulate freely among nodes.

One of the most significant aspects of SCC detection is its ability to identify self-contained clusters, in which nodes mutually reinforce the information propagation. For example, in OSNs, a strongly connected community ensures that a message (e.g., news, a viral post, or misinformation) continues circulating within the group, making it highly resilient to external removal [44]. Conversely, nodes outside an SCC might receive information, but do not contribute to its continued spread. This is particularly relevant for influence maximization [45][46][47], where marketers, advertisers, or political campaigns seek to target key SCCs to ensure that information cascades efficiently within a community before attempting to expand outward [34].

For efficient WCC detection, we are used the Breadth-First Search (BFS) method [48], which allows the analysis of how information can spread across different parts of a network even when direct connections are absent. It is an effective method for detecting WCCs in directed graphs. To find WCCs using BFS, the directed graph is first treated as undirected by considering all edges as bidirectional. The algorithm then initializes a set of unvisited nodes and iterates through each node in the network. Each unvisited node, BFS is performed to explore all reachable nodes by systematically traversing neighbors, marking them as visited, and grouping them into a component. This process is repeated until all nodes are visited, resulting in the identification of all WCCs in the graph. The BFS approach operates with a time complexity of $O(V + E)$, making it suitable for large-scale networks. For instance, on OSNs, WCCs can reveal clusters of users who, although not directly connected, can still influence each other through intermediaries.

This understanding is crucial for modeling diffusion processes, designing effective communication strategies, and controlling the spread of information or misinformation. Studies have shown that the structure of WCCs significantly affects the efficiency and reach of network information dissemination. [49] Analyzed influnece diffusion and consensus dynamics in weakly connected component in directed graphs and provided insights into how information spreads in such structures. Therefore, analyzing WCCs provides valuable insights into the structural mechanisms that underpin information diffusion in directed graphs.

4.2 Step 2: Influence calculation and top-k nodes selection

Once the network components (SCCs, WCCs, and the full network) are identified, we determine the top ten most influential seed nodes using multiple centrality measures. These measures have been extensively cited in the literature as strong indicators of influence maximization. We used seven centrality measures: in-degree, out-degree, betweenness, closeness, pagerank, eigenvector, and Katz centrality, each highlighting the different structural aspects of a node's influence.

For each network structure type (SCC, WCC, and full network), we identified nodes with the highest centrality values to serve as seed nodes in the influence diffusion models (SIR, LT, and IC). Consider (S_{Indg}), (S_{Outdg}), (S_{Cl}), (S_{Bt}), (S_{Ev}), (S_{Pr}), and (S_{Kz}) represent nodes with the top centrality values. For the selection of seed nodes, we studied the three networks separately for each type of structure (SCC, WCC, and the full network). By selecting the top central nodes in the SCCs, WCCs, and the full network, We selected the top-10 nodes ($k = 10$) as seeds, following common practice in influence maximization research where small, fixed seed sets are widely adopted to balance diffusion effectiveness with computational feasibility. In addition, in viral marketing contexts, companies typically target only a limited number of highly influential users due to budgetary and operational costs, since engaging fewer but strategically chosen seeds is more cost-effective and still sufficient to trigger large-scale diffusion. Therefore, our choice of $k = 10$ aligns with both prior studies and practical considerations [50][51][51][53]and centrality value were used in their raw form for ranking nodes. Since each centrality measure was applied independently, only the relative order of nodes was necessary for seed selection, and no normalization was performed. we performed a comparative analysis of the influence spread in different network structures.

4.3 Step 3: Influence diffusion simulation

To evaluate the effectiveness of centrality-based seed selection in influence propagation, we utilized the Higgs Twitter dataset[1], which captures user interactions surrounding the discovery of a Higgs boson-like particle on July 4, 2012. The dataset includes retweets, replies, mentions between July 1 and July 7, 2012, and consists of three directional networks representing different types of user activities, with anonymized user IDs maintained across all layers. This structure enables large-scale network studies, where one layer represents the social structure while the other three encode distinct user dynamics. We focused on the Higgs Twitter dataset, but it's important to mention that the dataset contains three separate graph representations (retweet networks, reply networks, and mention networks). Each of the three graphs has its own structural characteristics, which include size, density and degree-distribution. Thus, it also provides the different testbeds in a single dataset and allows us to evaluate the performance of the centrality measures and diffusion models on different networks. Table 2 outlines the dataset's characteristics. The study was implemented using Google Collaboratory, leveraging libraries such as NetworkX, iGraph, NumPy, Pandas, Matplotlib, and NDlib for the propagation process.

To simulate information diffusion using the selected seed nodes across the three types of networks, we employed the three most widely used diffusion models [54]: LT, IC, and SIR. Each model represents a distinct mechanism by which individuals adopt and spread information, ensuring more comprehensive understanding diffusion of influence. These models are particularly relevant for viral marketing analysis, as they capture different dynamics of information spread. The selected seed nodes were designated as source nodes in these models.

The diffusion process was analyzed separately for each types of network, allowing us to compare the influence maximization strategies across different network topologies. Key performance metrics include the final reachability (number of influenced nodes), speed of information propagation, and influence retention within the network substructures. Influence spread was measured over the course of the diffusion process and summarized at its termination. For the LT and IC models, we recorded both the number of activated nodes and the number of inactivated nodes across time steps, reporting the final state once propagation completed. For the SIR model, we tracked the temporal evolution of susceptible, infected, and recovered nodes, with the final counts at the last time step serving as the evaluation metrics. For each centrality measure and diffusion model, we performed 50 independent runs, and the reported results correspond to the averages across these runs. Tables (6-8) present the average influence spread measured under each diffusion model for each network structure, with values representing the final node states after diffusion terminates.

To simulate the spread of influence among selected key users identified using centrality measures, the parameter values for each model were carefully chosen to ensure a realistic representation of information propagation. For the SIR model, the infection probability (β) was set to 0.5, and the recovery rate (γ) to 0.05, simulating an epidemic-like diffusion process where nodes stop transmitting influence once they recover [55][56][57]. This configuration reflects short-term influence dynamics, where information spread is temporary.

The LT model was configured with a threshold of 0.1, meaning that at least 10% of a node's neighbors must be active before the node itself becomes active. This choice represents a gradual accumulation of influence, where activation requires reinforcement from multiple sources, making it suitable for modeling social behaviors like opinion formation and product adoption [58]. For the IC model, an infection probability of 0.9 was used to simulate a highly influential propagation process, where selected influencers have a strong likelihood of activating their neighbors. Since IC model lacks a recovery mechanism, this setup is ideal for modeling permanent adoption scenarios, such as viral marketing. To ensure consistency across the models, we conducted 50 iterations for each simulation, allowing for a comprehensive assessment of influence diffusion patterns under different network structures. These configurations enabled a comparative analysis of how centrality-based influencers drive information spread across diverse diffusion models. This comprehensive evaluation helps to determine which model best utilizes the network structure for optimal influence dissemination. By integrating SCC and WCC identification, top-k seed nodes selection, and multi model diffusion simulation, this approach provides a robust framework for analyzing and optimizing the information propagation.

1.2 5 Results and discussion

5.1 Results

Our experiments' results are categorized into two major sections: (i) the ranking of the most influential nodes using centrality measures in various network structures (Tables 3-5), and (ii) the diffusion performance of these measures in three different diffusion models (Tables 6-8).

The top-10 nodes for each centrality measure for the reply, mention, and retweet networks are presented in Tables (3-5), respectively, and detecting equivalence across the Full, SCC, and WCC structures demonstrates how node rankings can vary with network representation.

The influence spread results using the SIR, LT, and IC models for the same networks are summarized in Tables (6-8), respectively. These tables show the average number of activated nodes produced by different centrality measures, across the Full, SCC, and WCC structures, which allows for a direct comparison of their effectiveness at driving diffusion. To ensure the robustness of our findings, all results were averaged over 50 independent runs for each diffusion model and centrality measure [59][60].

Table 3, presents the top ten key users selected by each centrality in the different graphs structures on Twitter reply dataset. The equivalence of the top ten nodes between the Full, SCC, and WCC networks shows that certain measures remain highly stable, while others vary significantly. Eigenvector centrality was the most consistent, with 100% equivalence in WCC and full network and 90% equivalence in SCC and WCC/full network, meaning that all the top ten nodes remain the same in both networks. Similarly, in-degree, eigenvector, betweenness, pagerank, and Katz exhibit 100% equivalence in the WCC and full network, indicating that the core influential nodes remained unchanged regardless of the inclusion of WCCs. Eigenvector centrality remains stable because it identifies nodes that are connected to other influential nodes, making their rankings less sensitive to changes in network structure. Similarly, in-degree, betweenness, and pagerank rely on direct connections or shortest paths, which remain largely unchanged between WCC and the full network. This confirms that WCC retains the essential structure of the network, meaning additional weakly connected nodes do not significantly impact rankings.

The equivalence between SCC and WCC/Full is low in most centrality measures, with in-degree, closeness, pagerank, and Katz at 10%, out-degree and betweenness at 40%. This suggests that SCC functions as a structurally independent core, and expanding to the full network significantly changes the role of key nodes. The low equivalence percentages indicate that central nodes in SCC are not necessarily influential in the broader network. The higher out-degree equivalence (40%) suggests that some information-spreading nodes retain their importance, but in general, SCC nodes play distinct roles compared to their WCC/full network.

Similarly, Table 4 shows that the most consistent measures eigenvector, betweenness, pagerank, and Katz show 100% equivalence in the full network and WCC, meaning all top ten nodes remained the same. This occurs because these measures account for both direct and indirect influence, and the removal of WCC does not significantly affect major hubs.

Similarly, out-degree centrality exhibits 100% equivalence between the WCC and full networks, confirming that the core influential nodes remain unchanged regardless of the inclusion of WCC. Since WCC retains all major hubs, the rankings of the most central nodes remain consistent.

In contrast, the equivalence between SCC and WCC/full varies: it is high for betweenness (90%), eigenvector and closeness (80%), out-degree (90%) and lower for Katz and pagerank (40%-60%). This suggests that the SCC functions as an independent structural core, and expanding to the full network significantly alters the role of key nodes. The lower equivalence for Katz and pagerank indicates that the introduction of weakly connected nodes redistributes influence across the network. While the SCC captures a dense, highly connected subset, the WCC and full network introduce additional connections that shift centrality rankings.

Finally, Table 5 shows that eigenvector centrality showed 100% equivalence across networks. This stability occurs because eigenvector centrality is based on recursive influence meaning that if a node is central in WCC, it remains central when weakly connected nodes are added to form the full network. Additionally, other measures such as out-degree, betweenness, in-degree, and Katz also exhibit stability between WCC and the full network. This suggests that the core influential nodes remain largely unchanged regardless of whether WCC are included.

The equivalence between SCC and WCC/full networks is low (40%) for in-degree and Katz Centrality, indicating that SCC functions as a structurally independent core. This happens because SCC consists of a highly interconnected subset of nodes that operate differently from the full and WCC networks. The low equivalence (40%) means that when the network expands beyond SCC, the importance of key nodes shifts significantly.

This suggests that SCC acts as an independent core structure where central nodes play different roles compared to the broader network.

The SIR model reveals that out-degree, betweenness and pagerank outperform other measures in facilitating the spread. Despite its theoretical efficiency [61], closeness centrality results in minimal diffusion. This occurs because out-degree centrality identifies nodes with the most direct connections, allowing them to infect many others immediately, making it highly effective for rapid diffusion. Katz centrality also performs well as it considers both direct and indirect influence, making it a strong measure for long-range impact. However, closeness centrality, which prioritizes nodes with the shortest paths to all others, is ineffective in sparse networks, where long distances between nodes hinder rapid information propagation; it is worth noting that the performance of closeness centrality can vary depending on the specific network and the parameters used in the SIR model [62].

Similarly, in the LTM, out-degree centrality leads to the highest activation in full and WCC network, followed by Indg centrality and Bt centrality in SCC network, while the other centrality performs poorly. The LTM activates nodes based on the influence of their neighbors, which explains why out-degree centrality is the most effective nodes with many outgoing links can spread influence more easily. In-degree centrality also contributes because nodes that receive influence from many sources are more likely to activate.

Finally, the IC model produces even more restrictive results, yet out-degree centrality remains the most effective. The IC mode introduces probabilistic activation, meaning nodes influence their neighbors with a certain probability. Since nodes with high out-degree have more neighbors, they increase the chances of spreading activation, even in a probabilistic setting. (See Table 6, Table 7, and Table 8 for detailed results).

5.2 Discussion

From the experimental results, it can be seen that out-degree and betweenness centralities are the most effective measures for influence propagation. Out-degree centrality, in particular, is strong in initiating the diffusion process due to its capacity to activate many neighbors [62]. This is because out-degree centrality measures the number of edges originating from a node, indicating its potential to spread influence to other nodes in the network [63]. Betweenness centrality, on the other hand, plays a critical role in sustaining influence over time by bridging different parts of the network. This is because betweenness centrality measures the fraction of shortest paths between all node pairs that pass-through a given node, indicating its ability to connect different parts of the network and facilitate the spread of influence [64][65].

The SIR model demonstrated superior efficacy for sustained diffusion, reflecting a more realistic simulation of influence propagation compared to LT and IC models, which are more appropriate for large networks into account the influence propagation over time and the role of network structure in facilitating or hindering this process [14][66][67].

Overall, our reserch draw attentions to the necessity of considering network structure and the dynamics of

Influence propagation when designing marketing campaigns or evaluating the influence value of a node in a network. The most well-known and frequently utilized metrics in social network analysis are centrality measures, nevertheless, their application in the identification of the influencers depends on the network's characteristics. A key contribution of this study lies in demonstrating that SCC/WCC decomposition yields new insights that are not captured when only the full network is analyzed. WCC analysis provides computational efficiency without significantly altering results, making it a practical alternative for large-scale networks. SCC analysis, on the other hand, reveals structurally distinct influencer roles, which are obscured at the full-network level. This dual

perspective extends state-of-the-art centrality-based influence maximization approaches by linking performance not only to the choice of centrality measure and diffusion model but also to the underlying structural representation of the network.
Unlike prior influence maximization studies that apply centrality measures directly on full networks [68][69], or analyze these measures only at a theoretical level without specification of network structure [11][12][14][15], our work introduces a decomposition-based evaluation across SCCs and WCCs. This design provides two key insights: (i) performance differences of centrality measures are topology-dependent, with measures such as out-degree dominating in SCCs while betweenness gains importance in WCCs, and (ii) WCC-based analysis offers significant computational efficiency without substantially sacrificing diffusion performance. These insights, which are not captured when analyzing only full networks, highlight the value of incorporating SCC/WCC decomposition in centrality-based influence maximization.
Our approach combining multi-model diffusion, centrality comparison, and structural analysis (SCC/WCC/Full) offers a more complete understanding of influence propagation. These findings provide a foundation for designing hybrid influence maximization frameworks and strategies that are both effective and computationally efficient.

Table 2: Summary statistics of used datasets.

	Full graph		**SCC graph**		**WCC graph**	
	Nodes	**Egdes**	**Nodes**	**Egdes**	**Nodes**	**Edges**
Reply	3891	32523	322	708	12839	14944
Mention	116408	150818	1801	7069	91606	132068
Retweet	256491	328132	984	3850	223833	308596

6 Implications

The findings of this study provide valuable theoretical and practical implications for OSN analysis, influence maximization, and information diffusion strategies. By systematically comparing different centrality measures across various network structures and diffusion models, this study offers a comprehensive perspective on how influence propagates in OSNs. The integration of structural network considerations with diffusion modeling contributes to a more robust methodology for influence analysis.

From a practical perspective, social media analysts and digital marketers can apply these insights to refine influence-spreading strategies. Identifying key seed users based on diffusion models that align with specific campaign objectives can enhance brand visibility and engagement in viral marketing. A notable contribution of this study is its focus on network component structures, particularly the role of SCCs and WCCs in shaping diffusion outcomes. Unlike conventional approaches that treat networks as uniform entities, our findings demonstrate that strategies targeting SCCs can improve local influence retention, whereas those focusing on WCCs enable broader dissemination across loosely connected communities. These insights can inform the development of more effective influence maximization algorithms that integrate both local and global centrality measures.
Another important finding from our comparative analysis is the necessity of selecting diffusion models that best suit the network's characteristics and the intended spread of information. This study is among the first to provide a detailed, quantitative evaluation of how local centrality measures, such as out-degree centrality, play a dominant role in the LT model, whereas global measures like pagerank, and betweenness centrality become more relevant in the SIR model. Additionally, out-degree centrality emerges as the most effective measure for initiating the diffusion process due to its direct quantification of an individual’s capacity to spread information. Conversely, betweenness centrality plays a crucial role in sustaining influence by acting as a bridge between different network segments.

Among the diffusion models assessed, the SIR model is particularly effective due to its balanced treatment of infection and recovery, supporting sustained diffusion compared to LT or IC models. This makes it well-suited for simulating real-world information dissemination where messages need to spread continuously until fading.

Finally, this study highlights the practical utility of WCC-based analysis as a computationally efficient alternative to full-network evaluation. Since centrality rankings for measures such as in-degree, PageRank, Katz, and eigenvector remain stable in WCC, analyzing the full network is often unnecessary. WCC retains the essential structural features while excluding weakly or entirely disconnected nodes, leading to reduced computational complexity without compromising accuracy. This refinement is particularly valuable for large-scale OSN applications where rapid decision-making is critical for example, identifying and activating key Twitter accounts during the first minutes of a breaking news event, selecting top YouTube influencers for short-lived marketing campaigns, or monitoring Facebook groups during emergencies to support timely information dissemination.

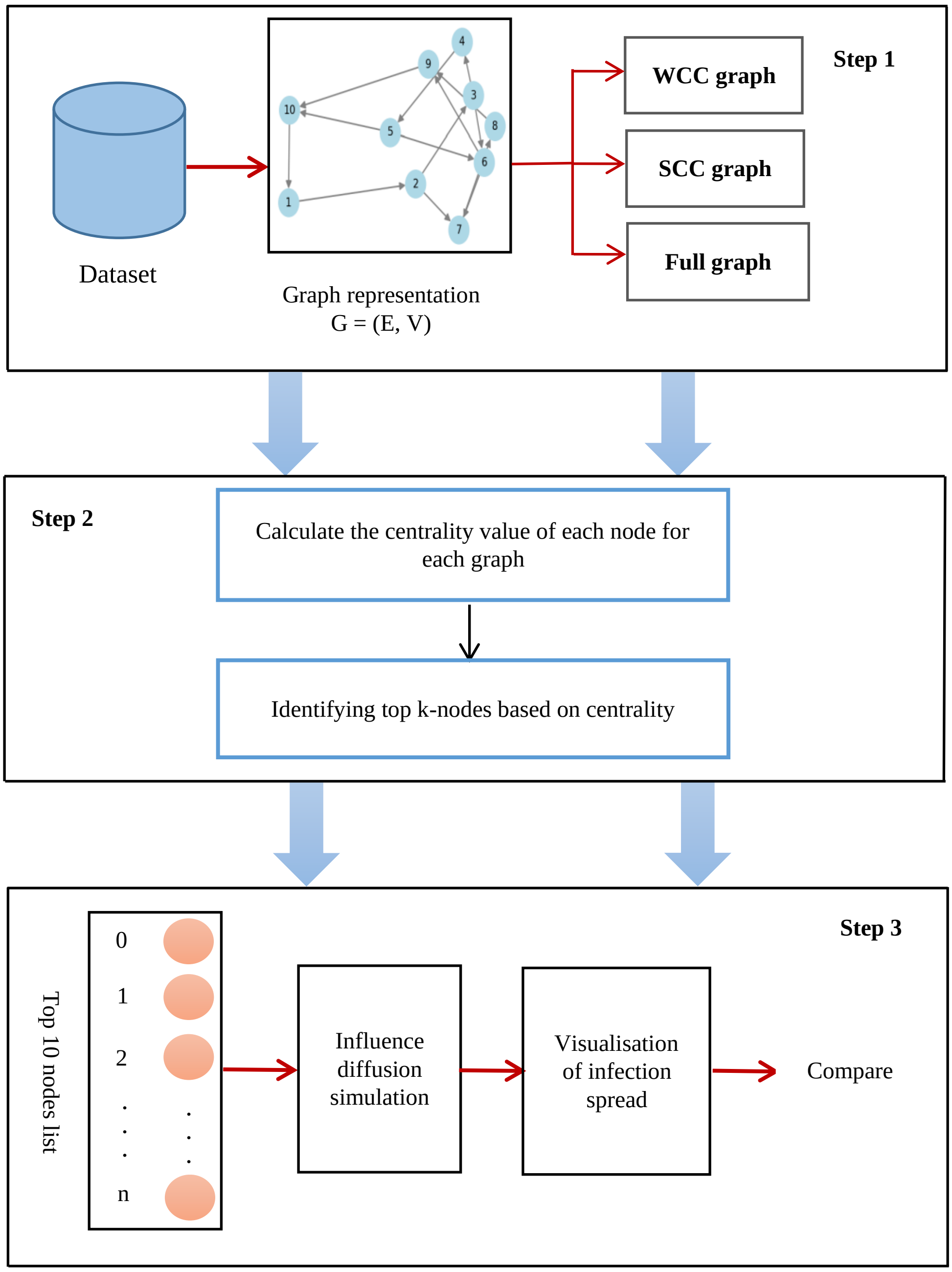

Figure 2 : Workflow of the methodology

Table 3: Top 10 key users identified by various centrality measures in reply dataset.

	Ranking	Indg		Outdg		Ev		Bt		Cl		Pr		Kz	
		Node Id	**Value**	**Node Id**	**Value**	**Node Id**	**Value**	**Node Id**	**Value**	**Node Id**	**Value**	**Node Id**	**Value**	**Node Id**	**Value**
SCC network	1	677	23	9021	24	9021	1.0000	67382	44556.25	67382	0.3116	677	0.0438	9021	0.0068
	2	9021	21	67382	15	50218	0.5113	35376	32084.41	98204	0.2833	9021	0.0215	677	0.0065
	3	67382	14	52908	14	50244	0.4649	12751	29038.85	13808	0.2803	67382	0.0186	52908	0.0052
	4	6241	13	6241	13	8855	0.4465	9021	27318.42	35376	0.2755	36436	0.0182	6241	0.0052
	5	13808	13	98204	10	80429	0.3863	13808	27098.21	12751	0.2750	13808	0.0180	13808	0.0050
	6	52908	12	9964	9	80426	0.3552	6940	23586.11	677	0.2736	52908	0.0154	67382	0.0050
	7	98204	11	33833	9	9036	0.3157	69883	21405.54	6241	0.2677	6241	0.0136	12751	0.0047
	8	12751	11	12751	8	8989	0.2996	69891	21252.83	214092	0.2618	5137	0.0128	3604	0.0046
	9	3604	9	42172	8	71888	0.2989	6241	20867.89	89805	0.2603	3604	0.0122	9964	0.0045
	10	9964	8	3604	1	50277	0.2983	98204	19752.06	42172	0.2597	12751	0.0109	5137	0.0044
WCC network	1	677	1206	9021	35	9021	1.0000	13808	1.4e+06	88	0.2462	677	0.0648	677	0.0078
	2	88	1071	16695	33	50218	0.5113	677	1.2e+06	677	0.2338	88	0.0254	88	0.0067
	3	220	470	433454	32	50244	0.4649	36436	1.2e+06	27311	0.2249	10836	0.0122	220	0.0032
	4	3549	218	6241	26	8855	0.4465	67382	1.2e+06	76803	0.2221	220	0.0110	3549	0.0014
	5	317	168	113517	25	80429	0.3863	52908	1.2e+06	96775	0.2204	10844	0.0104	317	0.0011
	6	349	142	67382	24	80426	0.3552	42177	1.1e+06	3998	0.2172	10867	0.0079	349	0.0008
	7	1988	105	52908	20	9036	0.3157	52882	1.1e+06	98204	0.2171	201222	0.0079	3369	0.0007
	8	7690	96	72466	18	8989	0.2996	201222	1.1e+06	103986	0.2170	207364	0.0079	7690	0.0007
	9	3369	92	269152	17	103181	0.2989	12751	6.9e+05	163525	0.2167	152385	0.0079	1988	0.0006
	10	16460	79	20971	17	71888	0.2989	6940	6.4e+05	184519	0.2161	237807	0.0079	16460	0.0006
Full network	1	677	1206	9021	35	9021	1.0000	13808	1.4e+06	124554	1.000	677	0.0243	677	0.0026
	2	88	1071	16695	33	50218	0.5113	677	1.2e+06	286277	1.000	88	0.0095	88	0.0022
	3	220	470	433454	32	50444	0.4649	36436	1.2e+06	274148	1.000	10836	0.0046	220	0.0010
	4	3549	218	359985	31	8855	0.4465	67382	1.2e+06	274149	1.000	220	0.0041	3549	0.0004
	5	317	168	6241	26	80429	0.3863	52908	1.2e+06	279630	1.000	10844	0.0039	317	0.0004
	6	349	142	113517	25	80426	0.3552	42177	1.1e+06	179783	1.000	10867	0.0029	349	0.0002
	7	1988	105	67382	24	9036	0.3157	52882	1.1e+06	188380	1.000	201222	0.0029	3369	0.0002
	8	7690	96	52908	20	8989	0.2996	201222	1.1e+06	52504	1.000	207364	0.0029	7690	0.0002
	9	3369	92	72466	18	103181	0.2989	12751	6.9e+05	201087	1.000	152385	0.0029	1988	0.0002
	10	16460	79	269152	17	71888	0.2989	6940	6.4e+05	191708	1.000	237807	0.0029	16460	0.0002

Table 4: Top 10 key users identified by various centrality measures in mention dataset.

	Ranking	Indg		Outdg		Ev		Bt		Cl		Pr		Kz	
		Node Id	**Value**	**Node Id**	**Value**	**Node Id**	**Value**	**Node Id**	**Value**	**Node Id**	**Value**	**Node Id**	**Value**	**Node Id**	**Value**
SCC network	1	88	462	89805	129	88	1.0000	13808	728765.58	88	0.4807	88	0.0975	88	0.0083
	2	3998	137	1276	42	3998	0.5429	12751	613890.54	3998	0.4059	3998	0.0457	677	0.0027
	3	677	130	9021	42	13808	0.5284	89805	559561.26	89805	0.3992	52087	0.0406	3998	0.0021
	4	13808	119	26158	41	12751	0.2706	64911	551493.61	13808	0.3874	13808	0.0288	13808	0.0019
	5	1988	81	6241	35	67382	0.2605	88	520146.06	2417	0.3784	64911	0.0244	2417	0.0016
	6	2417	68	67382	32	5226	0.2366	6940	492184.47	5226	0.3742	677	0.0149	9021	0.0015
	7	5226	63	492	32	677	0.2214	67382	465101.31	1276	0.3742	3604	0.0092	1988	0.0013
	8	9021	55	12751	31	64911	0.2676	35376	424761.87	12751	0.3735	2417	0.0086	12751	0.0011
	9	12751	54	20385	30	11991	0.2115	110903	284255.80	26158	0.3684	12751	0.0049	3604	0.0011
	10	35376	51	4665	28	52087	0.1855	9021	263752.60	6241	0.3671	9021	0.0048	6241	0.0011
WCC network	1	88	11953	89805	169	88	1.0000	88	5.1e+07	88	0.3553	13813	0.0769	88	0.0089
	2	677	3906	26158	57	3998	0.5429	64911	5.0e+07	2417	0.3120	88	0.0466	677	0.0029
	3	2417	2533	1276	49	13808	0.5284	13808	4.2e+07	3998	0.3114	4741	0.0170	2417	0.0020
	4	59195	1601	9021	44	13813	0.5217	89805	2.6e+07	13813	0.3075	3998	0.0130	59195	0.0014
	5	3998	1587	6241	43	12751	0.2706	12751	2.6e+07	89805	0.3030	677	0.0125	7533	0.0012
	6	7533	1528	492	40	67382	0.2605	6940	2.3e+07	13808	0.2994	3369	0.0102	383	0.0010
	7	383	1357	149922	39	5226	0.2366	110903	2.2e+07	1276	0.2987	2417	0.0097	3998	0.0008
	8	1988	1189	67382	38	677	0.2214	67382	2.2e+07	677	0.2984	52087	0.0087	3369	0.0007
	9	13813	1066	4665	38	4259	0.2178	35376	1.9e+07	5226	0.2953	59195	0.0082	1988	0.0007
	10	519	805	12751	37	64911	0.2176	3998	1.9e+07	26158	0.2910	13808	0.0082	13813	0.0006

Full network	1	88	11953	89805	169	88	1.0000	88	5.1e+07	107757	1.000	13813	0.0642	88	0.0070
	2	677	3906	26158	57	3998	0.5429	64911	5.0e+07	124554	1.000	88	0.0389	677	0.0023
	3	2417	2533	1276	49	13808	0.5284	13808	4.2e+07	286277	1.000	4741	0.0142	2417	0.0016
	4	59195	1601	9021	44	13813	0.5217	89805	2.6e+07	284372	1.000	3998	0.0108	59195	0.0011
	5	3998	1587	8241	43	12751	0.2706	12751	2.6e+07	274148	1.000	677	0.0104	7533	0.0010
	6	7533	1528	492	40	67382	0.2605	6940	2.3e+07	274149	1.000	3369	0.0085	383	0.0008
	7	383	1357	149922	39	5226	0.2366	110903	2.2e+07	111667	1.000	2417	0.0081	3998	0.0006
	8	1988	1189	67382	38	677	0.2214	67382	2.2e+07	158380	1.000	52087	0.0073	3398	0.0005
	9	13813	1066	4665	38	4259	0.2178	35376	1.9e+07	52504	1.000	59195	0.0069	1988	0.0005
	10	519	805	12751	37	64911	0.2176	3998	1.9e+07	185346	1.000	13808	0.0068	13813	0.0005

Table 5: Top 10 key users identified by various centrality measures in retweet dataset.

	Ranking	Indg		Outdg		Ev		Bt		Cl		Pr		Kz	
		Node Id	Value	Node Id	Value	Node Id	Value	Node Id	Value	Node Id	Value	Node Id	Value	Node Id	Value
SCC network	1	88	209	53508	35	88	1.0000	64911	205916.69	88	0.4701	88	0.0583	88	0.0076
	2	3998	105	64911	34	3998	0.7022	6940	174494.42	677	0.4142	2342	0.0377	1988	0.0038
	3	677	95	27705	31	11991	0.6915	35376	162656.19	1988	0.4041	64911	0.0269	677	0.0033
	4	13808	71	492	26	42172	0.6521	28951	154796.13	27705	0.3949	3998	0.0212	5226	0.0026
	5	1988	69	75798	25	64911	0.5449	88	137928.53	6940	0.3938	39420	0.0167	349	0.0025
	6	2417	66	182906	25	13808	0.4732	103447	114613.31	3998	0.3905	169287	0.0161	9964	0.0021
	7	5226	47	39885	24	39885	0.4648	3547	114184.74	5226	0.3867	134095	0.0161	6940	0.0021
	8	9021	46	103447	23	56968	0.4268	511	99540.44	53508	0.3857	28951	0.0157	13808	0.0020
	9	12751	42	4509	22	3547	0.4199	9021	85390.09	50305	0.3842	13808	0.0156	3998	0.0019
	10	35376	38	35376	22	110903	0.4018	1988	81233.94	64911	0.3829	42172	0.0147	519	0.0019
WCC network	1	88	14060	38535	134	88	1.0000	64911	5.3e+07	88	0.3234	88	0.0275	88	0.0038
	2	14454	9160	181190	84	3998	0.7022	88	4.4e+07	677	0.2971	2342	0.0158	14454	0.0019
	3	677	5613	81405	66	11991	0.6915	35376	2.9e+07	38535	0.2957	64911	0.0098	677	0.0015
	4	1988	4335	64911	49	42172	0.6521	28951	2.7e+07	1988	0.2907	39420	0.0079	1988	0.0010
	5	349	2802	54301	49	64911	0.5449	6940	2.5e+07	14454	0.2849	14454	0.0077	283	0.0007
	6	283	2039	27705	48	13808	0.4732	103447	2.5e+07	1276	0.2832	677	0.0075	349	0.0006
	7	3571	1980	53508	42	39885	0.4648	3547	2.0e+07	6940	0.2831	2567	0.0069	68278	0.0006
	8	6948	1959	232850	41	56968	0.4268	677	1.8e+07	3998	0.2810	134095	0.0067	6948	0.0005
	9	14572	1692	492	38	3547	0.4199	39885	1.7e+07	349	0.2793	169287	0.0067	3571	0.0005
	10	68278	1689	52204	38	110903	0.4018	1988	1.4e+07	5226	0.2771	1988	0.0062	3549	0.0005
Full network	1	88	14060	38535	134	88	1.0000	64911	5.3e+07	326123	1.000	88	0.0253	88	0.0034
	2	14454	6190	181190	84	3998	0.7022	88	4.4e+07	326124	1.000	2342	0.0146	14454	0.0017
	3	677	5613	81405	66	11991	0.6915	35376	2.9e+07	81236	1.000	64911	0.0091	677	0.0013
	4	1988	4335	64911	49	42172	0.6521	28951	2.7e+07	57763	1.000	39420	0.0091	1988	0.0009
	5	349	2802	54301	49	64911	0.5449	6940	2.5e+07	172959	1.000	14454	0.0071	283	0.0006
	6	283	2039	27705	48	13808	0.4732	103447	2.5e+07	283287	1.000	677	0.0069	349	0.0006
	7	3571	1980	53508	42	39885	0.4648	3547	2.0e+07	263179	1.000	2567	0.0064	68278	0.0005
	8	6948	1959	232850	41	56968	0.4268	677	1.8e+07	293265	1.000	134095	0.0062	6948	0.0005
	9	14572	1692	492	38	3547	0.4199	39885	1.7e+07	321295	1.000	169287	0.0062	3571	0.0004
	10	68278	1689	52204	38	110903	0.4018	1988	1.4e+07	430115	1.000	1988	0.0057	3549	0.0004

Table 6: Results of the influence propagation of various centrality measures in a reply network under IC, LT and SIR model.

		LTM			ICM			SIR		
		SCC	WCC	Full	SCC	WCC	Full	SCC	WCC	Full
Reply network	Indg	100	46	46	10	10	10	208	408	410
	Outdg	102	219	233	10	27	25	242	528	647
	Cl	62	59	10	10	10	10	211	466	10
	Ev	52	48	48	10	10	10	190	460	416
	Bt	73	64	64	10	11	10	219	497	494
	Pr	93	10	10	10	10	10	200	374	343
	Kz	99	46	46	10	10	10	224	365	505

Table 7: Results of the influence propagation of various centrality measures in a mention network under IC, LT and SIR model.

		LTM			ICM			SIR		
		SCC	**WCC**	**Full**	**SCC**	**WCC**	**Full**	**SCC**	**WCC**	**Full**
Mention network	**Indg**	317	10	10	14	10	10	1094	1872	1848
	Outdg	333	209	209	35	54	45	1096	1927	1911
	Cl	290	146	10	25	28	10	1117	1854	11
	Ev	287	42	42	10	10	14	1140	1904	1814
	Bt	350	112	112	34	39	30	1101	1756	1922
	Pr	334	14	14	13	10	10	1121	1886	1909
	Kz	332	10	10	18	10	10	1155	1858	1874

Table 8: Results of the influence propagation of various centrality measures in a retweet network under IC, LT and SIR model.

		LTM			ICM			SIR		
		SCC	**WCC**	**Full**	**SCC**	**WCC**	**Full**	**SCC**	**WCC**	**Full**
Retweet network	**Indg**	227	10	10	10	10	10	608	1234	1678
	Outdg	193	184	184	21	62	50	638	1635	1635
	Cl	217	18	10	16	28	10	619	10	10
	Ev	138	27	27	10	10	11	657	1671	1671
	Bt	219	45	45	13	17	14	642	1664	1664
	Pr	143	13	13	10	10	10	569	1727	1727
	Kz	227	10	10	10	10	10	625	1296	1296

7 Conclusion

This study presented a comprehensive analysis of influential node identification in OSNs using seven centrality measures, including in-degree, out-degree, betweenness, closeness, eigenvector, pagerank and Katz centrality. It aimed to evaluate how influence propagates across different network structures, such as SCC and WCC and discern the effectiveness of various diffusion models, specifically the LT, IC, and SIR models. The primary objective of this study was to deepen the understanding of influence maximization in OSNs by systematically evaluating centrality measures in relation to network structures and influence diffusion dynamics.

Our analysis revealed that out-degree and betweenness centralities were the most effective measures for influence propagation, with out-degree being particularly strong in initiating the diffusion process due to its capacity to activate many neighbors, while betweenness played a critical role in sustaining influence over time by bridging different parts of the network. Furthermore, the SIR model demonstrated superior efficacy for sustained diffusion, reflecting a more realistic simulation of influence dissemination compared to LT and IC models, which are more appropriate for activation-based scenarios. We also found that utilizing WCC can optimized influence campaigns, enabling marketers to enhance efficiency while retaining key influencers.

However, the study acknowledged limitations, including reliance on established centrality measures and diffusion models that may not fully accounted for the complexity of user interactions in rapidly evolving digital landscapes. In future work, we aim to propose a hybrid model that combines the strengths of local and global centrality measures with the dynamics of different diffusion models based on our findings. This hybrid approach will incorporate temporal changes and real-time feedback mechanisms, ensuring adaptability to the evolving nature of OSNs.

References

[1] Safari, R. M., Rahmani, A. M., & Alizadeh, S. H. (2019). User behavior mining on social media: a systematic literature review. *Multimedia Tools and Applications*,*78*(23),33747-33804. https://doi.org/10.1007/s11042-019-08046-6

[2] 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

[3] Yan, Z., Zhou, X., Ren, J., Zhang, Q., & Du, R. (2023). Identifying underlying influential factors in information diffusion process on social media platform: A hybrid approach of data mining and time series regression. *Information Processing & Management*,*60*(5),103438. https://doi.org/10.1016/j.ipm.2023.103438

[4] Shen, B., Guan, T., Ma, J., Yang, L., & Liu, Y. (2021). Social network research hotspots and trends in public health: A bibliometric and visual analysis. *Public Healt in Practice*, *2*, 100155. https://doi.org/10.1016/j.puhip.2021.100155

[5] Ni, P., Zhu, J., Gao, Y., & Wang, G. (2024). Minimizing the misinformation concern over social networks. *Information Processing & Management*, *61*(1),103562. https://doi.org/10.1016/j.ipm.2023.103562

[6] Zheng, H., Zhao, H., & Ahmadi, G. (2024). Towards improving community detection in complex networks using influential nodes. *Journal of Complex Networks*, *12*(1),cnae001. https://doi.org/10.1093/comnet/cnae001

[7] Wang, W., & Street, W. N. (2018). Modeling and maximizing influence diffusion in social networks for viral marketing. *Applied network science*, *3*, 1-26. https://doi.org/10.1007/s41109-018-0062-7

[8] Medjahed, F., Molina, E., & Tejada, J. (2025). Effectiveness of Centrality Measures for Competitive Influence Diffusion in Social Networks. *Mathematics (2227-7390)*,*13*(2). https://doi.org/10.3390/math13020292

[9] Meshcheryakova, N., & Shvydun, S. (2024). A comparative analysis of centrality measures in complex networks. *Automation and Remote Control*, *85*(8), 685-695. https://doi.org/10.1134/S0005117924700127

[10] Baabcha, H., Laifa, M., & Akhrouf, S. (2022). Social Influence Analysis in Online Social Networks for Viral Marketing: A Survey. In *International Conference on Managing Business Through Web Analytics* (pp. 143-166). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-06971-0_11

[11] Sarkar, D., Kole, D. K., & Jana, P. (2016). Survey of influential nodes identification in online social networks. *International Journal of Virtual Communities and Social Networking (IJVCSN)*, *8*(4), 57-69. https://doi.org/ 10.4018/IJVCSN.2016100104

[12] Singh, S. S., Singh, K., Kumar, A., Shakya, H. K., & Biswas, B. (2019). A survey on information diffusion models in social networks. In *Advanced Informatics for Computing Research: Second International Conference, ICAICR 2018, Shimla, India, July 14–15, 2018, Revised Selected Papers, Part II 2* (pp. 426-439). Springer Singapore. https://doi.org/10.1007/978-981-13-3140-4

[13] Arrami, S., Oueslati, W., & Akaichi, J. (2018). Detection of opinion leaders in social networks: A survey. In *Intelligent Interactive Multimedia Systems and Services 2017 10* (pp. 362-370). Springer International Publishing. https://doi.org/10.1007/978-3-319-59480-4_36

[14] Ribeiro, A. C., Azevedo, B., Oliveira e Sá, J., & Baptista, A. A. (2020). How to measure influence in social networks? In *International Conference on Research Challenges in Information Science* (pp. 38-57). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-50316-1_3

[15] Singh, R. R. (2022). Centrality measures: a tool to identify key actors in social networks. *Principles of Social Networking: The New Horizon and Emerging Challenges*, 1-27. https://doi.org/10.1007/978-981-16-3398-0_1

[16] Yujie, Y. (2020). A survey on information diffusion in online social networks. In *Proceedings of the 2020 European Symposium on Software Engineering* (pp. 181-186). https://doi.org/10.1145/3393822.343232

[17] Li, M., Wang, X., Gao, K., & Zhang, S. (2017). A Survey on Information Diffusion in Online Social Networks: Models and Methods. *Information*, *8*(4), 118. https://doi.org/10.3390/info8040118

[18] Khatri, I., Choudhry, A., Rao, A., Tyagi, A., Vishwakarma, D. K., & Prasad, M. (2023). Influence Maximization in social networks using discretized Harris' Hawks Optimization algorithm. *Applied Soft Computing*, *149*, 111037. https://doi.org/10.1016/j.asoc.2023.111037

[19] Dhingra, S., Dodwad, P. S., & Madan, M. (2016). Finding strongly connected components in a social network graph. *International Journal of Computer Applications*, *136*(7), 1-5. https://doi.org/10.5120/ijca2016908481

[20] Bonifazi, G., Buratti, C., Corradini, E., Marchetti, M., Parlapiano, F., Ursino, D., & Virgili, L. (2025). Defining, Detecting, and Characterizing Power Users in Threads. *Big Data and Cognitive Computing*, *9*(3), 69. https://doi.org/10.3390/bdcc9030069

[21] Puigbò, J. Y., Sánchez-Hernández, G., Casabayó, M., & Agell, N. (2014). Influencer detection approaches in social networks: A current state-of-the-art. In *Artificial intelligence research and development* (pp. 261-264). https://doi.org/10.3233/978-1-61499-452-7-261

[22] Guzman, J. D., Deckro, R. F., Robbins, M. J., Morris, J. F., & Ballester, N. A. (2014). An analytical comparison of social network measures. *IEEE Transactions on Computational Social Systems*, *1*(1), 35-45. https://doi.org/10.1109/TCSS.2014.2307451

[23] Zafarani, R. (2014). *Social Media Mining: An Introduction*. Cambridge University Press. https://doi.org/10.1017/CBO9781139088510

[24] Wang, Y., Zhang, L., Yang, J., Yan, M., & Li, H. (2024). Multi-factor information matrix: A directed weighted method to identify influential nodes in social networks. *Chaos, Solitons & Fractals*, *180*, 114485. https://doi.org/10.1016/j.chaos.2024.114485

[25] Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry* https://doi.org/10.2307/3033543

[26] Pereira, F. S., Gama, J., de Amo, S., & Oliveira, G. M. (2018). On analyzing user preference dynamics with temporal social networks. *Machine Learning*, *107*, 1745-1773. https://doi.org/10.1007/s10994-018-5740-2

[27] Rashid, Y., & Bhat, J. I. (2024). Topological to deep learning era for identifying influencers in online social networks: a systematic review. *Multimedia Tools and Applications*, *83*(5), 14671-14714. https://doi.org/10.1007/s11042-023-16002-8

[28] Borgatti, S. P., & Everett, M. G. (2006). A graph-theoretic perspective on centrality. *Social networks*, *28*(4), 466-484. https://doi.org/10.1016/j.socnet.2005.11.005

[29] Duda, R. O., & Hart, P. E. (2006). *Pattern classification*. John Wiley & Sons

[30] Lü, L., Chen, D., Ren, X. L., Zhang, Q. M., Zhang, Y. C., & Zhou, T. (2016). Vital nodes identification in complex networks. *Physics reports*, *650*, 1-63. https://doi.org/10.1016/j.physrep.2016.06.007

[31] Liao, H., Mariani, M. S., Medo, M., Zhang, Y. C., & Zhou, M. Y. (2017). Ranking in evolving complex networks. *Physics Reports*, *689*, 1-54. https://doi.org/10.1016/j.physrep.2017.05.001

[32] Wenji, C., Bo, Y., Ruigang, Z., Qiong, W., Binsha, Z., Zengbo, L., & Kai, N. (2022). Research on Key Node Identification Method of Transmission Network based on Improved PageRank Algorithm. In *2022 41st Chinese Control Conference (CCC)* (pp. 5056-5061). IEEE. https://doi.org/10.23919/CCC55666.2022.9902364

[33] Zhao, L., Sun, P., Zhang, J., Peng, M., Zhong, Y., & Liang, W. (2023). A complex network important node identification based on the KPDN method. *Applied Sciences*, *13*(14), 8303. https://doi.org/10.3390/app13148303

[34] 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* (pp. 137-146). https://doi.org/10.1145/956750.956769

[35] Pastor-Satorras, R., Castellano, C., Van Mieghem, P., & Vespignani, A. (2015). Epidemic processes in complex networks. *Reviews of Modern Physics*, *87*(3), 925. https://doi.org/10.1103/RevModPhys.87.925

[36] Badr, N., Abdel-Kader, H., & Ali, A. H. (2021). Diffusion Models for Social Analysis, Influence and Learning. *IJCI. International Journal of Computers and Information*, *8*(2), 162-169. https://doi.org/ 10.21608/ijci.2021.207863

[37] Tian, S., Mo, S., Wang, L., & Peng, Z. (2020). Deep reinforcement learning-based approach to tackle topic- aware influence maximization. *Data Science and Engineering*, *5*, 1-11. https://doi.org/10.1007/s41019-020-00117-1

[38] Khalife, S., Read, J., & Vazirgiannis, M. (2021). Structure and influence in a global capital–ownership network. *Applied Network Science*, *6*, 1-21. https://doi.org/10.1007/s41109-021-00359-6

[39] Ji, F., & Jin, J. (2025). A map-reduce algorithm to find strongly connected components of directed graphs. *Soft Computing*, 1-20. https://doi.org/10.1007/s00500-025-10451-z

[40] Zhong, Z., Lin, L., Jiang, Z., Yuan, X., Ngai, E., Lam, J., & Kwok, K. W. (2025). Connectivity Determination Algorithm for Complex Directed Networks. *IEEE Transactions on Network Science and Engineering*. https://doi.org/10.1109/TNSE.2025.3549777

[41] Tarjan, R. (1972). Depth-first search and linear graph algorithms. *SIAM journal on computing*, *1*(2), 146-160. https://doi.org/10.1137/0201010

[42] Tarjan, R. E., & Zwick, U. (2024). Finding strong components using depth-first search. *European Journal of Combinatorics*, *119*, 103815. https://doi.org/10.1016/j.ejc.2023.103815

[43] Sobhi, N. A. M. *Detection of terminal strongly connected components using a novel DFS-like algorithm* (Doctoral dissertation, Universiteit van Amsterdam).

[44] Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012). The role of social networks in information diffusion. In *Proceedings of the 21st international conference on World Wide Web* (pp. 519-528). https://doi.org/10.1145/2187836.2187907

[45] Yanchenko, E., Murata, T., & Holme, P. (2024). Influence maximization on temporal networks: a review. *Applied Network Science*, *9*(1), 16. https://doi.org/10.1007/s41109-024-00625-3

[46] Yang, W., Shi, Q., Yan, J., Wang, C., Song, M., & Wu, M. (2024). Complementary influence maximization under comparative linear threshold model. *Expert Systems with Applications*, *238*, 121826. https://doi.org/10.1016/j.eswa.2023.121826

[47] Jaouadi, M., & Romdhane, L. B. (2024). A survey on influence maximization models. *Expert Systems with Applications*, *248*, 123429. https://doi.org/10.1016/j.eswa.2024.123429

[48] Monserrat, T. J. K. P., Pabico, J. P., & Albacea, E. A. (2015). A Hybrid Graph-drawing Algorithm for Large, Naturally-clustered, Disconnected Graphs. *arXiv preprint arXiv:1507.02766*. https://doi.org/10.48550/arXiv.1507.02766

[49] Veerman, J. J. P., & Kummel, E. (2019). Diffusion and consensus on weakly connected directed graphs. *Linear Algebra and its Applications*, *578*, 184-206. https://doi.org/10.1016/j.laa.2019.05.014

[50] Wang, B., Zhang, J., Dai, J., & Sheng, J. (2022). Influential nodes identification using network local structural properties. *Scientific Reports*, *12*(1), 1833. https://doi.org/10.1038/s41598-022-05564-6

[51] Ullah, A., Wang, B., Sheng, J., Long, J., Khan, N., & Sun, Z. (2021). Identification of nodes influence based on global structure model in complex networks. *Scientific Reports*, *11*(1), 6173. https://doi.org/10.1038/s41598-021-84684-x

[52] Wang, G., Alias, S. B., Sun, Z., Wang, F., Fan, A., & Hu, H. (2023). Influential nodes identification method based on adaptive adjustment of voting ability. *Heliyon*, *9*(5). https://doi.org/10.1016/j.heliyon.2023.e16112.

[53] Wang, Y., Zhang, L., Yang, J., Yan, M., & Li, H. (2024). Multi-factor information matrix: A directed weighted method to identify influential nodes in social networks. *Chaos, Solitons & Fractals*, *180*, 114485. https://doi.org/10.1016/j.chaos.2024.114485

[54] Csókás, E., & Vinkó, T. (2025). A Heuristic for Influence Maximization Under Deterministic Linear Threshold Model. *Informatica*, *48*(4). https://doi.org/10.31449/inf.v48i4.4805

[55] Tang, L., Zhou, Y., Wang, L., Purkayastha, S., Zhang, L., He, J., ... & Song, P. X. K. (2020). A review of multi-compartment infectious disease models. *International Statistical Review*, *88*(2), 462-513. https://doi.org/10.1111/insr.12402

[56] Bhat, N., Aggarwal, N., & Kumar, S. (2020). Identification of influential spreaders in social networks using improved hybrid rank method. *Procedia Computer Science*, *171*, 662-671. https://doi.org/10.1016/j.procs.2020.04.072

[57] Zhao, X., Liu, F. A., Wang, J., & Li, T. (2017). Evaluating influential nodes in social networks by local centrality with a coefficient. *ISPRS International Journal of Geo-Information*, *6*(2), 35. https://doi.org/10.3390/ijgi6020035

[58] Talukder, A., Alam, M. G. R., Tran, N. H., Niyato, D., Park, G. H., & Hong, C. S. (2019). Threshold estimation models for linear threshold-based influential user mining in social networks. *IEEE Access*,*7*,105441-105461. https://doi.org/10.1109/ACCESS.2019.2931925

[59] Wang, F., Zhu, Z., Liu, P., & Wang, P. (2019). Influence maximization in social network considering memory effect and social reinforcement effect. *Future internet*, *11*(4), 95. https://doi.org/10.3390/fi11040095

[60] Liang, J. C., Gong, Y. J., Wu, X. K., & Li, Y. (2024). Customized influence maximization in attributed social networks: heuristic and meta-heuristic algorithms. *Complex & Intelligent Systems*, *10*(1), 1409-1424. https://doi.org/10.1007/s40747-023-01220-2

[61] Ishfaq, U., Khan, H. U., & Iqbal, S. (2022). Identifying the influential nodes in complex social networks using centrality-based approach. *Journal of King Saud University-Computer and Information Sciences*, *34*(10), 9376-9392. https://doi.org/10.1016/j.jksuci.2022.09.016

[62] Şimşek, A. (2022). Lexical sorting centrality to distinguish spreading abilities of nodes in complex networks under the Susceptible-Infectious-Recovered (SIR) model. *Journal of King Saud University-Computer and Information Sciences*, *34*(8), 4810-4820.

[63] Waniek, M., Michalak, T. P., Wooldridge, M. J., & Rahwan, T. (2018). Hiding individuals and communities in a social network. *Nature Human Behaviour*, *2*(2), 139-147. https://doi.org/10.1038/s41562-017-0290-3

[64] Ghalmane, Z., El Hassouni, M., & Cherifi, H. (2018, October). Betweenness centrality for networks with non-overlapping community structure. In *2018 IEEE workshop on complexity in engineering (COMPENG)* (pp.1-5).IEEE. https://doi.org/10.1109/CompEng.2018.8536229

[65] Doostmohammadian, M., Rabiee, H. R., & Khan, U. A. (2020). Centrality-based epidemic control in complex social networks. *Social Network Analysis and Mining*, *10*, 1-11. https://doi.org/10.1007/s13278-020-00638-7

[66] More, J. S., & Lingam, C. (2019). A gradient-based methodology for optimizing time for influence diffusion in social networks. *Social Network Analysis and Mining*, *9*(1), 5. https://doi.org/10.1007/s13278-018-0548-4

[67] Shelke, S., & Attar, V. (2019). Source detection of rumor in social network–a review. *Online Social Networks and Media*, *9*, 30-42. https://doi.org/10.1016/j.osnem.2018.12.001

[68] Dey, P., Chaterjee, A., & Roy, S. (2019). Influence maximization in online social network using different centrality measures as seed node of information propagation. *Sādhanā*, *44*(9), 205. https://doi.org/10.1007/s12046-019-1189-7

[69] Sengupta, A., Middya, A. I., & Roy, S. (2024). Centrality measures on segmented entropy networks to identify influencers and influencees for financial market scenario. *International Journal of Data Science and Analytics*, 1-22. https://doi.org/10.1007/s41060-024-00608-8