Received: 12 April 2026 | Revised: 23 June 2026 | Accepted: 3 July 2026 | Published online: 6 August 2026

Journal of Data Science and Intelligent Systems
2026, Vol. 00(00) 1–12
DOI: 10.47852/bonviewJDSIS62029981

RESEARCH ARTICLE

A Hybrid Approach for Prominent Node Detection in Social Media Networks Using Modified Cluster Walktrap with Analytical Hierarchy Process

Faz Mohammad[1,*] **and Shweta Vikram**[1]

[1] *Department of Computer Science and Engineering, Maharishi University of Information Technology, India*

Abstract: The rapid growth of social media platforms has significantly increased the complexity of online interaction networks, creating new challenges in suspicious community detection and influential node identification. Conventional community detection approaches often face limitations in handling heterogeneous social network structures and identifying structurally significant nodes using multiple decision parameters simultaneously. To address these limitations, this study proposes a hybrid framework integrating modified cluster Walktrap (MCW) with the Analytical Hierarchy Process (AHP) for suspicious community detection and prominent node analysis in social media networks. In the first phase, the proposed MCW approach utilizes probability-driven random-walk analysis and modularity optimization to identify structurally similar communities in bipartite graph structures. In the second phase, AHP is applied to evaluate influential nodes using multiple graph-theoretic parameters, including mean distance, degree centrality, and betweenness centrality. The proposed framework was experimentally evaluated using the publicly available KONECT Facebook social network dataset. Experimental results demonstrate that the proposed MCW + AHP framework achieved improved clustering effectiveness, reduced overlapping community formation, and more structurally consistent community detection compared with conventional Cluster Edge Betweenness and Walktrap approaches. The framework also enabled balanced multi-criteria influential node evaluation, where Node 101 was identified as the most structurally significant node within the analyzed network. Overall, the proposed framework provides a computationally feasible and effective approach for structural community analysis and influential node identification in complex social media environments.

Keywords: modified cluster Walktrap, social network analysis, community detection, influential node identification, Analytical Hierarchy Process

1. Introduction

Social media platforms have emerged as significant channels of communication in the modern digital age as they provide for communication, information sharing, media exchange, and virtual community creation despite geographical barriers. Sites such as Facebook, Instagram, Telegram, X, and LinkedIn have revolutionized the mode of communication, collaboration, marketing, and dissemination of information on the internet [1]. Ease of access and convenience have made many social media sites attract billions of users all over the world, leading to the rapid development of the interconnected social network systems. However, while there are many advantages associated with social media sites, they can also become vulnerable to cybercrime owing to the sharing of personal information, anonymity, and quick information dissemination [2].

Misuse of social networking sites has led to various forms of cyberattacks such as phishing, identity theft, fraud, cyberbullying, unauthorized access, distribution of false information, and financial scams [3]. Cybercriminals often use virtual social networks as well as communication systems to perform illegal operations without revealing their identities in large social networks [4]. It is seen that victims often fail to report cyberattacks because of embarrassment, among other issues, such as unawareness, fear of publicity, or lack of technical knowledge [5]. The process of detecting malicious groups or people playing influential roles in social networks, therefore, is considered an important research problem.

The social network analysis (SNA) constitutes an efficient framework for analyzing relations, interactions, and behaviors among the users in the online environment [6]. Graph-based modeling of social networks treats users as nodes, whereas relationships or interactions among the users are treated as edges [7]. With the help of graphs, one can efficiently analyze the topologies of networks, information diffusion, behavior similarity, and community formations. Several parameters from graph theory, for instance, degree centrality, betweenness centrality,

***Corresponding author:** Faz Mohammad, Department of Computer Science and Engineering, Maharishi University of Information Technology, India. Email: faiz.conatctid@gmail.com

closeness centrality, eigenvector centrality, and average distance, have been used for discovering the structurally important nodes in the network [8]. High connectivity nodes and nodes with higher intermediary influence play a vital role in spreading information or coordinating cyber-actions.

The identification of hidden communities, influencers, and interaction patterns that are involved in criminal activity is one of the current research areas in SNA known as the Criminal Network Analysis (CNA) [9]. There exist many challenges in the analysis and detection of criminal activities, for example, incomplete data, network evolution over time, the overlapped nature of communities, and the scale-free nature of the network [10]. It has been observed that users and interaction features in criminal networks constitute two different categories of nodes to model heterogeneity in relationship behavior [11]. Traditional techniques for community detection face many challenges in detecting dense and suspiciously structured communities in such networks.

Methods of community detection such as Louvain, Girvan–Newman, Cluster Edge Betweenness (CEB), and the Walktrap algorithm have already found application in the field of SNA [12]. The Walktrap algorithm is considered among the best techniques due to its capability of detecting densely connected subgraphs through the utilization of random-walk traversal behavior. Conventional methods of Walktrap, however, have been developed to deal with traditional graphs and fail to adequately capture structural similarities as well as dependencies of interactions within bipartite social networks [13]. Also, while many previous works focus mainly on the detection of suspicious communities, there is less emphasis placed on the identification of the influential node in them.

Multi-criteria decision-making (MCDM) approaches have recently received considerable attention in terms of the analysis of influential nodes within suspicious communities. There are many MCDM methods, which include Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), Elimination et Choix Traduisant la REalité (ELECTRE), decision trees, and the Analytical Hierarchy Process (AHP). All these techniques enable evaluation of several structural parameters at once [14, 15]. AHP is one of these approaches, which enables the evaluation of alternatives based on pairwise comparison and weighing of several criteria. Therefore, combining the graph technique of community detection with that of MCDM will improve the accuracy of finding an influential node [8, 16].

Inspired by the gaps outlined in the literature, this paper proposes a combined approach called modified cluster Walktrap with Analytical Hierarchy Process (MCW + AHP) for the detection of suspicious communities and prominent nodes in social media networks. To achieve this goal, firstly, a probability-enhanced modified cluster Walktrap (MCW) algorithm would be applied for the detection of structurally similar suspicious communities in bipartite social media networks using transition-probability analysis and modularity optimization [16]. Then, second, the AHP will be used for ranking influential nodes based on several graph theory parameters: mean distance, degree centrality, and betweenness centrality.

The proposed solution is primarily innovative in its probability-enhanced MCW framework for the detection of structurally similar suspicious communities in bipartite social media networks. This framework will use the fusion of a transition-probability-based structural similarity evaluation technique along with modularity optimization in order to ensure higher stability and to avoid the overlapping community problem. Moreover, AHP is coupled with such graph-theoretic parameters as mean distance, degree centrality, and betweenness centrality to perform multi-criteria analysis of structurally influential nodes. The empirical evaluation of the MCW + AHP framework will be conducted using the publicly available KONECT dataset.

2. Literature Review

2.1. Community detection in criminal and social networks

The fast-growing online social media networking platforms have brought increased complexity in the network of digital interactions, thus inspiring research on sophisticated approaches of community detection to detect any suspicious behavior and hidden user groups. The process of community detection in social networks involves detecting communities in which the nodes interact in a similar way in terms of interactions, communication, or common interest areas [13, 17]. During cybercrime investigations, these communities are normally composed of suspicious groups who are involved in different malicious actions, including fraud, phishing, cyber harassment, misinformation spreading, and illegal data sharing [1].

One of the approaches used by many researchers to model and analyze social networks is graph-based analysis [7]. In the process of graph-based analysis, social interactions are considered as edges connecting nodes, while users constitute the nodes themselves. Some of the algorithms widely used to analyze social networks include Girvan–Newman, Louvain, CEB, and Walktrap [12, 13]. These algorithms involve the grouping of a social network into clusters. The most effective among these algorithms is Walktrap because of its random-walk approach for identifying nodes with structural similarity.

The recent research has focused on the improvement of the conventional approaches in community detection in complex and dynamical social networks. Ullah et al. [17] developed neural embedding-based approaches to find communities that could discover hidden dependencies in the structure of complex large networks. On the other hand, Li et al. [18] introduced an improved version of the Walktrap method by introducing K-means clustering to enhance the process of clustering and modularity maximization. Li et al. [19] presented Walktrap-Structural Perturbation Method (SPM) to detect overlapping communities from the complex graph. Though these methods improve clustering efficiency and accuracy, numerous existing community detection methods can be used only for homogeneous graphs, where they may experience certain issues related to applying them to heterogeneous social bipartite social networks with complex interaction structures.

Many researchers have recently pointed out the use of bipartite networks and heterogeneous networks to represent complex social interactions using a more accurate depiction of real-world conditions than homogeneous networks do. In bipartite networks, nodes fall into either of two categories (e.g., people vs the items they have purchased) and have interaction-based connections between them. By representing structural heterogeneity (i.e., by keeping much of a relationship intact), the amount of information lost by projecting the data to one mode is reduced through the use of bipartite graphs. For this reason, the analysis of bipartite graphs is becoming increasingly popular for the analysis of social media, the investigation of criminal networks, and the modeling of user–activity interactions—since the relationships that different types of entities have with each other substantially affect the behaviors and communities created by the network [11, 17].

The CNA is one of the subfields of SNA and focuses on discovering hidden criminal organizations and suspicious user communities [9]. The use of graph-based clustering to detect suspicious communities was presented by Mao et al. [20] and emphasized the importance of structural similarity in investigating cybercrimes. However, criminal networks tend to include incomplete data, evolving interaction structures, overlapping communities, and covert interactions, which makes it harder to detect such communities [9, 10].

2.2. Influential node detection and centrality analysis

In addition to community detection, identifying influential or prominent nodes within a social network is another major research challenge [16]. Influential nodes are users who possess significant structural importance and can strongly affect information propagation, behavioral influence, or coordinated suspicious activity within a network. Graph-theoretic centrality measures are widely used to evaluate the importance of nodes in social networks [9, 12].

Degree centrality is one of the simplest and most commonly used measures, representing the number of direct connections associated with a node. Nodes with higher degree values are generally considered highly active or influential users within the network. Betweenness centrality measures the extent to which a node acts as an intermediary between other nodes and therefore plays an important role in information flow and communication control. This method is also used to figure out how easy it is to get to certain points in a graph [8].

Several current studies have concentrated on influential node discovery using various hybrid and optimization methods. Kazemzadeh et al. [21] have put forward a community structure-based approach to influence maximization as a means for discovering influential nodes in social networks. Mukhtar et al. [22] have considered fuzzy centrality as a measure of uncertainty in SNA. Hafiene et al. [23] have developed a fuzzy logic-based influence maximization method to facilitate the ranking of influential nodes in large social networks. However, despite these recent developments, many existing models still utilize only single-parameter analyses, neglecting the complex nature of nodes' influence in social interaction networks [16].

Recent studies have demonstrated that the use of several graph parameters is highly beneficial in terms of increasing the effectiveness of the node discovery process [16, 24]. However, the problem of choosing the optimal influential node out of several candidates remains a complex MCDM problem. It has become one of the reasons for applying decision support methods in graph-based SNA problems.

Researchers have also recently demonstrated that employing a multi-criteria evaluation process and integrating community-level structural information into influential node identification can considerably enhance the effectiveness of a variety of hybrid node ranking and decision support methodologies. Instead of solely relying on individual centrality measures, hybrid methodologies take into consideration the clustering information of nodes, as well as multiple different structural attributes, to provide nodes with a more thorough ranking and to aid decision-making. Hybrid node-rank management methodologies are especially beneficial when analyzing suspicious communities since the behaviors associated with influential nodes can vary when utilizing a combination of different characteristics of the network [16, 24].

2.3. Multi-criteria decision-making approaches in social network analysis

The MCDM methodologies provide a systematic way of considering decisions on several decision criteria at once [14, 15]. With regard to SNA, MCDM methodologies are used increasingly often for the problem-solving of influential node selection, priority setting, risk management, and behavioral analysis. A number of MCDM methodologies, such as TOPSIS, ELECTRE, scoring methodology, and the decision tree approach, have already been applied to network-based decision-making problems. However, the AHP methodology suggested by Thomas Saaty still belongs to the most popular methodologies because of its pairwise comparison and hierarchical structure [15, 25].

The AHP breaks down an entire problem into a hierarchy consisting of objectives, criteria, sub-criteria, and alternatives. Pairwise comparisons are used to determine the weightage assigned to each of these criteria and alternatives. The process allows both qualitative and quantitative assessment of the problem, making it suitable for complicated decision-making scenarios where several structural parameters affect the end result. Niemcewicz [26] successfully used it for MCDM in the context of cloud computing, while recent research has also applied the AHP method for environmental monitoring, network optimization, and intelligent systems.

Although there are many applications of AHP in different areas, the use of AHP in conjunction with the modified graph clustering technique for suspicious community analysis is relatively unexplored in the existing literature. Earlier studies mainly focused on community detection in isolation or on the use of individual centrality measures for ranking nodes. Relatively fewer studies exist that integrate community detection techniques with multi-criteria evaluation of influential nodes for suspicious communities.

In recent years, intelligent decision support systems have addressed clustering techniques and multi-criteria decision analysis together to analyze complex networks. Most studies have either focused on cluster formation or on alternate ranking independently, and very few have proposed a unified framework that combines probability-based community detection with AHP-based evaluations of influential nodes. This is a critical limitation for decision makers working on social media and cybercrime investigations who need to analyze community structure and the influence of nodes at the same time when making decisions.

2.4. Research gap

Despite recent advancements in community detection techniques, influential node ranking techniques, and MCDM processes related to SNA, many research challenges remain unaddressed. There are proven algorithms used for community detection, such as Louvain, Girvan–Newman, CEB, and traditional Walktrap, which have been developed to identify groups of highly connected users; however, these algorithms are primarily based on a homogeneous network model. When applied to a heterogeneous or bipartite model, the performance of these algorithms tends to decline. Conventional methods for Walktrap rely heavily on structural connections as well as the random-walk behavior of nodes to form communities, causing interference with their clustering stability due to overlapping communities to create clusters in sparse/non-uniformly structured networks.

In exactly the same manner as other researchers have done, it has been done in the past for the identification of an influential node using individual graph-theoretic measures such as degree centrality, closeness centrality, and betweenness centrality. Although these centrality measures provide meaningful and relevant information about the structure of a complex network, one centrality measure does not necessarily indicate that the influence of a node is unique. With regard to nodes with high influence, depending on the method used to determine centrality, other node features may not reflect an accurate representation of their degree of influence.

Recent research has looked at applying MCDM techniques (such as AHP) to assist in the decision-making process in both network analysis and intelligent systems. The majority of AHP-based approaches published today have placed focus mainly on ranking alternatives based on the features extracted from the data and not on integrating the community detection process into a single analytical framework along with the identification of the influential nodes. The existing body of research on the combination of probability-based random-walk community detection with multi-criteria influential node ranking to analyze suspicious communities in social media environments is limited.

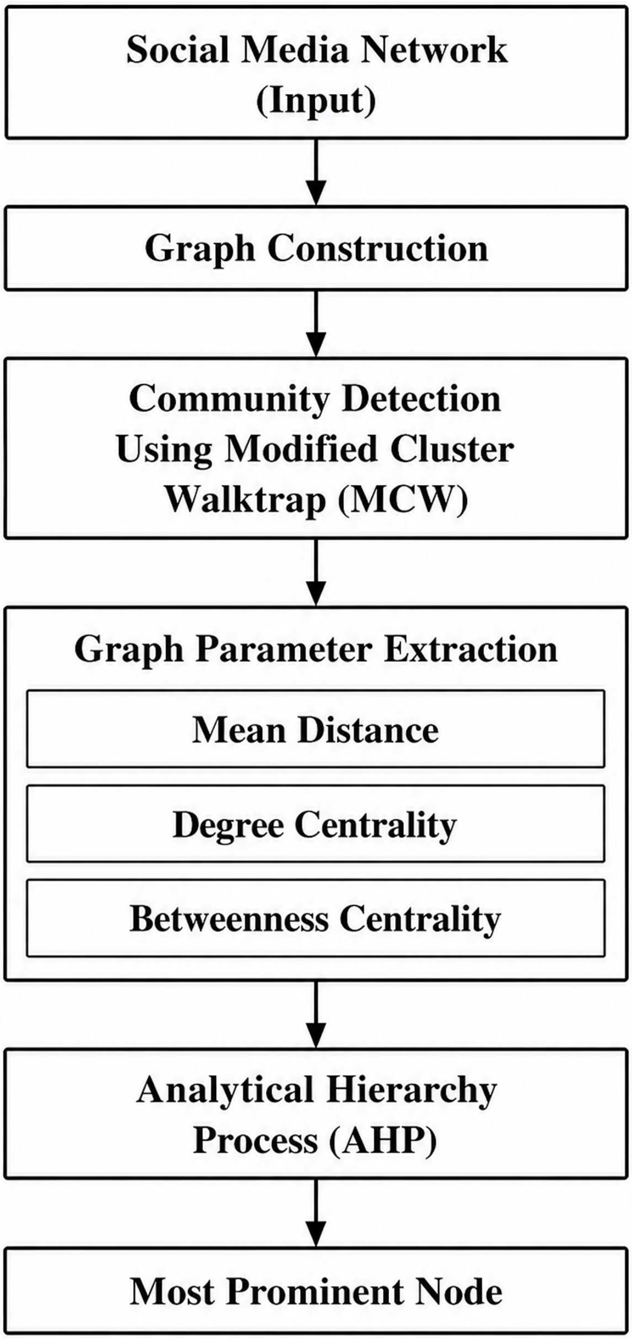
To address the limitations outlined above as it relates to community stability and colliding community structures, the study develops a hybrid approach by combining probability-based random-walk community detection through the use of MCW and multi-criteria influential node rankings through the use of AHP within an analytical framework (MCW + AHP). The development of a new hybrid approach to community detection and multi-criteria influential node ranking will enhance the identification of overlapping community structures within a stable community model while reducing the impact of transient nodes on community integrity.

3. Proposed Hybrid Framework for Community Detection and Prominent Node Identification

The growing complexity of online social networks makes it harder to detect suspicious communities and influential nodes. Traditional approaches to community detection mainly focus on finding dense subgraphs, whereas methods to find influential nodes generally use standalone centrality measures. However, the fact that real-world social interaction networks tend to exhibit heterogeneity, overlapping communities, and hidden communication makes traditional approaches less effective [11, 16]. To address this issue, this research proposes an integrated approach that utilizes the combination of the MCW and AHP to detect suspicious communities and influential nodes in bipartite social media networks.

The proposed integrated approach consists of two phases. The first phase involves the utilization of the MCW to detect similar structural communities based on random walks guided by transition probabilities in bipartite graph structures [13, 27]. The second phase applies AHP to determine the most influential node among detected communities using different graph parameters. The overall process of the proposed integrated approach is illustrated in Figure 1.

Figure 1
Overall architecture of the proposed MCW + AHP framework illustrating bipartite graph construction, community detection using modified cluster Walktrap, graph parameter extraction, and influential node identification through the Analytical Hierarchy Process



3.1. Network representation and graph construction

The social media network can be represented by a bipartite graph that contains

$$BG = (U, C, E) \tag{1}$$

Where:

$U = \{u_1, u_2, u_3, ..., u_n\}$ represents users,

$C = \{c_1, c_2, c_3, ..., c_m\}$ denotes interaction attributes or communication types,

E comprises edges linking users to interaction entities.

For this proposed graph model, nodes represent either users or interaction entities, whereas edges represent communication relations, common behaviors, or interactions between the nodes. Graph theory helps analyze node connectivity, interaction similarities, and community formation in the social network [6, 7]. Unlike the conventional one-mode graph approach, bipartite graph modeling retains heterogeneity in interrelation patterns, which helps prevent loss of information inherent in the projection process. Consequently, such a graph model is better suited for examining suspicious behavior involving different types of interaction.

3.2. Modified cluster Walktrap (MCW) for community detection

The traditional Walktrap algorithm detects communities based on conducting short random walks on a graph, which is done on the basis of the assumption that random walks occur within densely connected regions of a graph [13, 27]. Even though the traditional Walktrap algorithm works well on homogeneous graphs, it faces some difficulties when used for detecting communities in heterogeneous bipartite social networks due to the lack of consideration of heterogeneous interaction and probability-based similarity among nodes [17, 19]. In order to overcome this problem, the MCW algorithm applies a probability-based distance evaluation approach to assess the structural similarity of nodes in bipartite networks [18, 19]. In particular, this algorithm improves the decision-making about community merging when conducting a random walk.

First, each node is considered a separate community:

$$P_0 = \{\{v_1\}, \{v_2\}, \{v_3\}, \dots, \{v_n\}\} \tag{2}$$

Then, the transition-probability matrix is calculated based on the connection behaviors of nodes,

$$P_{ij} = \frac{A_{ij}}{d\,(i)} \tag{3}$$

Where:

A_{ij} is the adjacency relation between node i and node j,
$d\,(i)$ is the degree of node i,
P_{ij} is the transition probability from node i to node j.

The probability-based distance between any two nodes is calculated as the distance between their probability transition matrices.

$$PD(i, j) = \frac{|\, P_i - P_j \,|}{\sum_{k=1}^{n} |\, P_i - P_k \,|} \tag{4}$$

where P_i and P_j represent the transition-probability distributions associated with nodes i and j, respectively.

In contrast to traditional algorithms that use purely structural analysis, the modified Walktrap algorithm uses the normalized probability distribution of transitions to evaluate the structural similarity between nodes. As a result, this approach makes it possible to detect structurally related nodes in heterogeneous and sparse social networks. The merging of communities with minimum probability-based distance continues until reaching the modularity convergence point. The modularity of the resulting communities is calculated using the following formula:

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{d(i)d(j)}{2m} \right] \delta(C_i, C_j) \tag{5}$$

Where:

m is the total number of edges,
d(i) and $d\,(j)$ are the degrees of the nodes,
$\delta\,(C_i, C_j)$ equals 1 if the nodes belong to the same community and 0 otherwise.

3.3. Algorithmic steps of MCW

The procedural process of the suggested MCW approach is outlined as follows:

Algorithm 1: Modified Cluster Walktrap (MCW)

Inputs: Bipartite graph $BG = (U, C, E)$
Output: Communities found $Com = \{C_1, C_2, \dots, C_k\}$
1. Creation of the bipartite graph using the social media data.
2. Each node is assumed to be a single community.
3. Calculation of the adjacency and transition-probability matrices.
4. Calculation of probability-based distance between neighboring nodes.
5. Determination of pairs of nodes having minimal distance values.
6. Community merging based on structural similarity.
7. Calculation of modularity after each community merge.
8. Repeating steps 4–7 until the modularity converges.
9. Output of the detected communities.

The proposed modification of the Walktrap algorithm incorporates random-walk-based probability evaluation of structural similarity, which improves clustering stability and community formation.

3.4. Time complexity analysis

The computational complexity of the suggested MCW method mostly relies on the calculation of transition probabilities, distance calculations, and the iterative procedure of modularity maximization and community merging. For a network containing n nodes and m links:

The calculation of transition probabilities involves:

$$O(m) \tag{6}$$

The calculation of probability-distance between adjacent nodes involves the following:

$$O(n^2) \tag{7}$$

The iterative process of modularity maximization and communities merging involves approximately:

$$O(n^2 \log n) \tag{8}$$

Therefore, the computational complexity of the entire MCW algorithm can be estimated as:

$$O(n^2 \log n) \tag{9}$$

and is still computationally tractable for medium-sized social networks [17].

3.5. Analytical Hierarchy Process (AHP) for influential node identification

After detecting suspicious communities using MCW, the next step involves determining the most influential node among those detected communities. Considering the influence of many factors in the structure of the social network, using one centrality measure to determine the most influential node can be inaccurate or incomplete [8, 16, 22]. As such, this research uses the AHP for MCDM [14, 15, 28].

Some of the graph-based criteria evaluated in the AHP approach include:

1) Mean distance
2) Degree centrality
3) Betweenness centrality

The steps involved in the AHP method include the following [28]:

1) Generation of a pairwise comparison matrix of the selected criteria.
2) Calculation of priority weights of criteria.
3) Ranking of nodes according to selected criteria.
4) Calculation of overall weighted score for node ranking.

The calculation of the weighted score of each node is based on the formula:

$$\mathrm{Score(i)} = \sum_{k=1}^{n} W_k \times PV_{ik} \tag{10}$$

Where:

W_k is the weight of the criterion k,

PV_{ik} is the priority vector value of the node i for the criterion k.

The node having the highest weighted score will be considered the most prominent node among the detected community.

3.6. Consistency verification in AHP

For evaluating the consistency of the pairwise comparison method, the Consistency Ratio (CR) is calculated in compliance with the Saaty criteria of consistency verification [28]. The Consistency Index (CI) is found using the formula:

$$CI = \frac{\lambda_{max} - n}{n - 1} \tag{11}$$

The CR is defined as:

$$CR = \frac{CI}{RI} \tag{12}$$

Where:

λ_{max} is the largest eigenvalue of the matrix of comparisons,

n is the number of criteria,

RI is the Random Index.

If $CR < 0.1$, the consistency of the pairwise comparison is considered acceptable. This validation improves the reliability and objectivity of the proposed MCDM process [15, 28].

4. Implementation and Experimental Setup

4.1. Dataset description

Evaluation of the proposed model was done using the public KONECT Facebook Social Network Dataset [29]. The data represents social interaction between users through graph structures where nodes represent users, while edges represent social ties between them. This choice of data is based on its massive structure along with its ability to be studied concerning community formation and influential nodes in social media environments [7, 11, 30].

In the current study, the abovementioned data was analyzed using a structural approach toward graph network analysis in order to detect suspicious communities and influential nodes depending on their topology and interaction properties [8, 16, 22]. In order to enable probability-based community detection along with multiple criterion evaluation of the nodes, the dataset was represented in the form of a bipartite graph [11, 31].

As a pre-step to constructing a graph structure, the data was pre-processed in order to improve network consistency and delete redundant links. The preprocessing procedure included the following actions:

1) Deletion of duplicate edges,
2) Isolation of single nodes,
3) Normalization of the network,
4) Extraction of adjacency properties.

Finally, the processed data was used as input for the proposed MCW model.

4.2. Graph construction and network representation

The processed social network data was transformed into the graph format illustrated below:

$$BG = (U, C, E) \tag{13}$$

Where:

U is the set of users,

C is interaction or activity data,

E is the connection between nodes.

The network graph was constructed using tools of graph visualization and analysis. In the obtained network:

The size of nodes denotes the degree value of the node.

The intensity of colors of nodes denotes betweenness centrality.

The connection between edges is based on interactions between users.

The connection between edges is based on interactions between interconnected nodes.

Such a visual representation allows for revealing highly interconnected clusters and structurally important nodes [8].

Figure 2 illustrates the graph generated from the KONECT Facebook dataset after preprocessing and network construction.

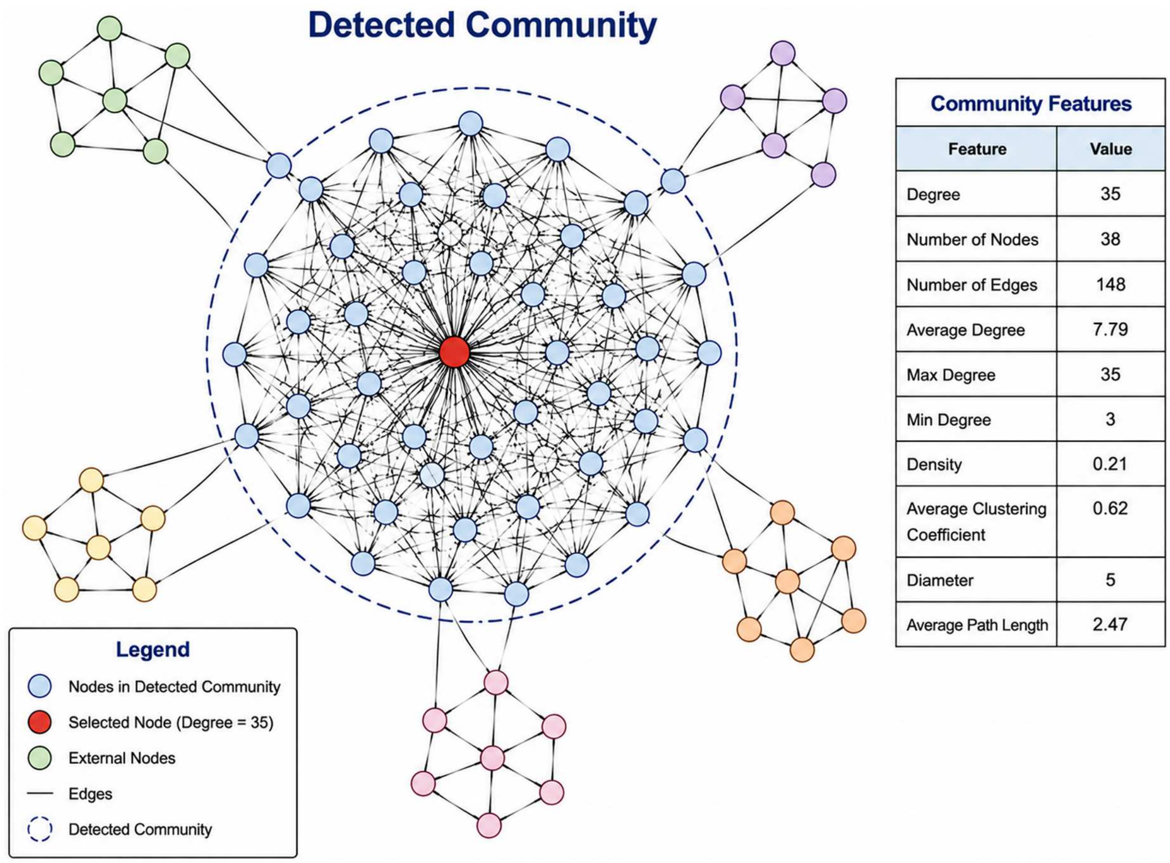
The generated graph demonstrates the existence of highly connected node groups and structurally dominant regions within the network. Such regions are potential indicators of suspicious communities and influential interaction behavior [20, 32].

4.3. Community detection using modified cluster Walktrap (MCW)

The first step involved in the new framework entails the detection of suspicious communities using the MCW algorithm. The MCW algorithm framework involves the integration of analysis through random walk with probability and distance measure for bipartite networks [13, 31]. Initially, each individual node in the graph was considered a community by itself. Transition probabilities between neighboring nodes were calculated based on the connectivity and node degree distribution of the network. Afterward, probability-distance measures were used to find structurally similar communities [18, 19]. Communities having lower probability-distance were continuously merged, with modularity values being taken note of at every point in the process. The merging was continued till the point when no further changes could be seen in the value of modularity [12, 31].

Results from the implementation process show that the proposed MCW framework successfully detected denser communities within the social network.

Figure 2
Graph representation generated from the KONECT Facebook dataset



Note: Schematic illustration of the detected community structure. The illustration was manually prepared using Draw.io and arranged with the assistance of ChatGPT for presentation purposes. All scientific content, network structure, and parameter values were verified by the authors.

4.4. Extraction of graph-theoretic parameters

After discovering the communities, a series of graph parameters were calculated using the discovered communities to discover influential nodes in the network. These included:

1) Mean distance
2) Degree centrality
3) Betweenness centrality

Figure 3 represents the values of the selected parameters for Node ID 101 using the developed graph-analyzing methodology.

As shown in the figure below, Node ID 101 has relatively high degrees and betweenness compared to neighboring nodes, indicating the node's structural influence within the community [16, 24, 33]. Below is the table that contains the parameter values for the important nodes whose node numbers are 49, 101, 105, and 314.

The parameter values extracted from prominent nodes are summarized in Table 1.

4.5. Influential node identification using AHP

The AHP method was applied in the second stage of the proposed procedure to find the influential node among the identified communities. AHP has been used here because the influential behavior depends upon the analysis of several structural properties at once. In contrast to the single criterion-based analysis, AHP allows us to perform weighted and multi-criteria evaluation of network parameters through pairwise comparisons [15, 25].

The following tasks have been executed in the application of the AHP procedure: pairwise parameter comparison, calculation of the priorities, evaluation of the nodes in terms of parameter ratings, weighted scores calculation, and selection of the node with the highest value [25]. The specific stages include:

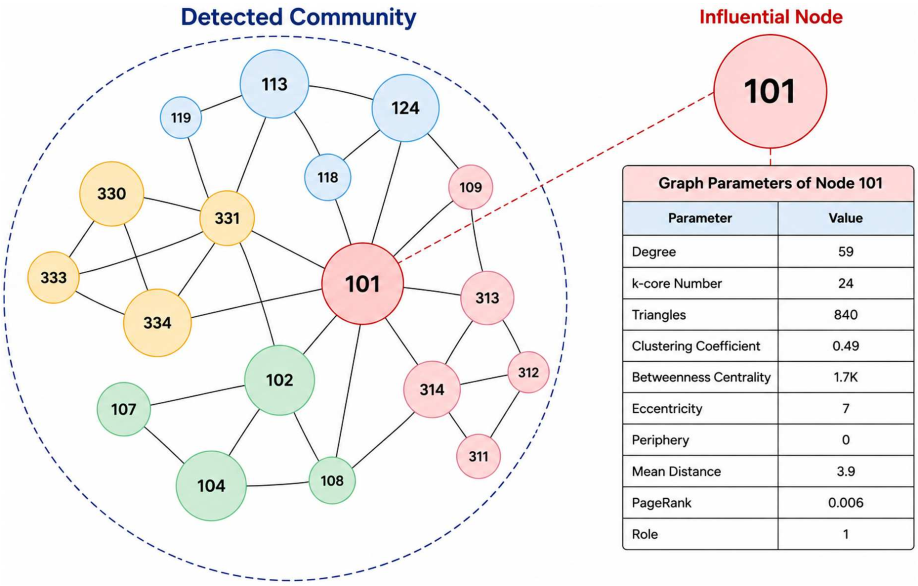
1) Creation of the pairwise comparison matrix for the graph's parameters.
2) Priorities calculation for the selected parameters.
3) Calculations of the parameter ratings for each node.
4) Weighted score calculation for all possible candidates.
5) Identification of the node with the maximum score.

The entire process of AHP application for influential node identification has been visualized in Figure 4.

The pairwise comparison matrix was formed using the scale suggested by Saaty to describe the comparative importance of the parameters [15]. The validation for consistency was performed through CR in order to ensure the correctness of pairwise comparison values. Table 2 illustrates the pairwise comparison matrix, while Table 3 shows calculations of priority vectors.

The vector for all the nodes generated using all the parameters is calculated. The parameters' rating per node using mean

Figure 3
Extraction and evaluation of graph-theoretic parameters, including degree centrality, betweenness centrality, and mean distance, for influential node analysis within detected communities



Note: Schematic illustration of influential node identification. The illustration was manually prepared using Draw.io and arranged with the assistance of ChatGPT for presentation purposes. All scientific content and node parameters were verified by the authors.

Table 1
Value of mean distance, degree, and betweenness for all prominent nodes with Node ID 49, 101, 105, and 314

Parameter (row)/node ID (column)	Mean distance	Degree	Betweenness
49	8.000	15.000	8.000
101	4.000	59.000	17.000
105	5.000	63.000	15.000
314	6.000	37.000	18.000

Figure 4
AHP-based influential node ranking process illustrating criteria weighting, pairwise comparison, and final prioritization of candidate nodes within the identified community

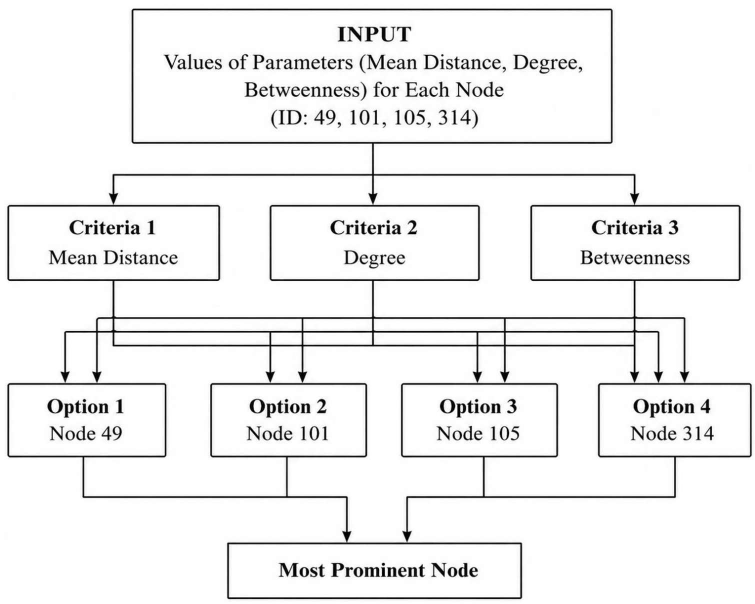


Table 2
Pairwise comparison matrix for parameters

Parameters	Mean distance	Degree	Betweenness
Mean distance	1.000	3.000	7.000
Degree	5.000	1.000	5.000
Betweenness	3.000	2.000	1.000

Table 3
Calculation of priority vector

	Mean distance	Degree	Betweenness	3rd root of product	PV
Mean distance	1.000	3.000	7.000	2.759	0.368
Degree	5.000	1.000	5.000	2.924	0.390
Betweenness	3.000	2.000	1.000	1.817	0.242
Sum				8.149	1.000

distance, degree, and betweenness is generated as shown in Tables 4, 5, and 6, respectively.

The final weighted score for every node was calculated using the following formula:

$$\mathrm{Score(i)} = \sum_{k=1}^{n} W_k \times PV_{ik} \tag{14}$$

Where:

W_k denotes the weight of the parameter,

PV_{ik} denotes the priority vector value of node i.

The node with the highest score at the end is considered to be the most influential node inside the discovered community. From the experiments conducted, it became evident that Node 101 had scored the highest among all nodes being tested, hence indicating high influence and importance within the social network being studied, as shown in Table 7.

4.6. Experimental environment

The MCW-AHP framework implementation was carried out using Python 3.11 on Windows 11. NetworkX, Pandas, NumPy, Matplotlib, and igraph libraries were used for the graph processing and community analysis. The experiments were conducted on a computer with an Intel Core i7 processor and 16GB of RAM, which provided sufficient computational power for graph construction, community detection, and extraction of the parameters needed for AHP-based ranking of the influential nodes.

4.7. Evaluation metrics

Performance of the proposed integration of the MCW and AHP algorithm was evaluated based on four parameters including the number of communities generated, iterations required, effectiveness, and CPU time. Effectiveness is considered to be directly proportional to the quality of the community detection process and the decrease in overlapping communities [17, 31]. The performance of the proposed method was compared to the CEB

Table 4
Priority vector of each node on the basis of mean distance parameter

Node ID	49	101	105	314	4th root of product	PV
49	1.000	0.500	0.630	0.750	0.697	0.169
101	2.000	1.000	1.250	1.500	1.391	0.337
105	1.600	0.800	1.000	1.200	1.114	0.270
314	1.330	0.670	0.830	1.000	0.928	0.224
SUM					4.130	1.000

Table 5
Priority vector of each node on the basis of degree parameter

Node ID	49	101	105	314	4th root of product	Priority vector
49	1.000	0.250	0.260	0.410	0.404	0.089
101	3.930	1.000	1.040	1.590	1.598	0.351
105	3.800	0.970	1.000	1.540	1.544	0.339
314	2.470	0.630	0.650	1.000	1.003	0.221
SUM					4.549	1.000

Table 6
Priority vector of each node on the basis of betweenness parameter

Node ID	49	101	105	314	4th root of product	Priority vector
49	1.000	0.470	0.530	2.000	0.840	0.180
101	2.130	1.000	1.130	4.250	1.790	0.390
105	1.880	0.880	1.000	3.750	1.580	0.340
314	0.500	0.240	0.270	1.000	0.420	0.090
SUM					4.630	1.000

Table 7
Final score for each node

Node ID/criteria	Mean distance	Degree	Betweenness	Score
49	0.174	0.086	0.138	0.137
101	0.348	0.337	0.293	0.335
105	0.278	0.360	0.259	0.297
314	0.232	0.217	0.310	0.246

and Conventional Cluster Walktrap (CWT) methods. Results showed that the proposed method performed better than others in terms of effectiveness and stability in generating communities [18, 29, 34].

4.8. Scalability and practical considerations

Even though the presented approach shows better performance in detecting communities, the computational complexity related to calculating the similarity of nodes based on probability becomes larger when the size of the network grows [17, 19, 34]. Compared to traditional edge-based clustering algorithms, MCW involves extra calculations to analyze the probability and distance between each node in the network. However, these extra calculations improve the quality of clustering results by increasing their stability and reducing the likelihood of finding structurally overlapped communities. Thus, the presented approach is better suited for analyzing suspicious interactions in cases when precision is of higher importance than fast execution of the algorithm. The approach was evaluated based on a static set of social media data. The framework may be used in social networks with dynamically changing structure, where nodes and edges keep evolving.

5. Experimental Results and Analysis

The experimental test for the proposed framework was carried out using the Facebook social network data collected from the KONECT database in order to determine its effectiveness in the context of network analysis. It was pitted against two other methods of graph clustering: the CEB and CWT. These comparison tests are displayed in Table 8.

The proposed MCW-AHP framework achieved the highest effectiveness score (10.43), outperforming the Cluster Edge Betweenness (CEB) method (9.69) and the Conventional Cluster Walktrap (CWT) method (7.22) approaches evaluated CEB 9.69 and CWT 7.22, while generating significantly more meaningful communities with good computational efficiency than both other approaches. Even though MCW-AHP took a little longer to execute than CEB, the improvement in the quality and structural consistency of the communities created from this framework justified the extra computational cost. Thus, the results of this research indicate there is an effective balance between the clustering performance and influential node analysis in social media networks using MCW-AHP.

6. Conclusion

In this study, the MCW algorithm and the AHP were integrated into a hybrid framework for detecting suspicious communities and identifying influential nodes in social networks. The framework employs a probability-based approach to detect a community and a multi-criteria approach to assess individual nodes to overcome the issues associated with traditional clustering algorithms and with single centrality-based ranking.

The evaluation performed on data extracted from the KONECT Facebook social network indicated that the integration of the MCW + AHP methodologies produced the highest level of effectiveness (10.43) compared to the CEB method (9.69) and the Conventional Walktrap method (7.22). The MCW + AHP framework successfully identified and generated structurally coherent communities while maintaining the same computational efficiency as a traditional clustering method. The integration of the mean distance, degree centrality, and betweenness centrality methodologies within the AHP framework also allowed for a balanced assessment of influential nodes, resulting in Node 101 being identified as the most significant node in the network that was analyzed.

In a practical sense, the proposed MCW + AHP methodology can support investigations into cybercrime, monitoring of suspicious communities, social media intelligence operations, and network security analysis through the identification of hidden communities and users with structural influence.

7. Future Work

It is possible to extend the suggested model further and incorporate features that facilitate an iterative process of influential node identification within changing social networks. Potential improvements may include the inclusion of multiple stages of community detection and graph updates, as well as repeated

Table 8
Comparative performance analysis of community detection approaches on the KONECT Facebook dataset

Algorithm	Number of clusters	Non-overlapping communities	Overlapping communities	Effectiveness score	CPU execution time (s)
Proposed MCW + AHP	146	13	14	10.430	3.810
Cluster Edge Betweenness (CEB)	126	20	13	9.690	3.250
Conventional Cluster Walktrap (CWT)	130	10	18	7.220	4.860

identification of influential nodes to develop a list of the most influential ones. Moreover, future directions for this research will include dynamic community detection in social media networks that evolve over time, adding more graph-theoretic features, testing on larger heterogeneous datasets, and evaluating other MCDM techniques for ranking of the influential nodes such as TOPSIS, ELECTRE, and fuzzy-AHP.

Acknowledgement

The authors would like to thank Maharishi University of Information Technology, Lucknow, for providing academic support and research facilities for conducting this study.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available in KONECT at https://networkrepository.com/socfb.php.

Author Contribution Statement

Faz Mohammad: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration. **Shweta Vikram:** Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration.

Generative AI Disclosure

Figures 2 and 3 were manually prepared using Draw.io (https://app.diagrams.net) based on the original network analysis results obtained in this study. The individual graphical components were subsequently arranged into publication-quality figures with the assistance of ChatGPT (GPT-5.5) to improve visual presentation and readability. No AI tool was used to generate, modify, or interpret the scientific data, network topology, experimental results, or conclusions. All scientific content, node relationships, parameter values, captions, and final figures were carefully reviewed, verified, and approved by the authors, who take full responsibility for their accuracy and integrity.

All graphical elements included in Figures 2 and 3 were created specifically for this manuscript and do not contain copyrighted or third-party licensed materials.

References

[1] Shakya, S. R., & Ceh-Varela, E. (2025). Social media analytics for investigations: A survey of recent trends, challenges and future research directions. *Journal of Social Media Research*, *2*(4), 297–318. https://doi.org/10.29329/jsomer.57

[2] Abid, A., Roy, S. K., Lees-Marshment, J., Dey, B. L., Muhammad, S. S., & Kumar, S. (2025). Political social media marketing: A systematic literature review and agenda for future research. *Electronic Commerce Research*, *25*(2), 741–776. https://doi.org/10.1007/s10660-022-09636-7

[3] Tomassi, A., Falegnami, A., & Romano, E. (2024). Mapping automatic social media information disorder. The role of bots and AI in spreading misleading information in society. *PLOS One*, *19*(5), e0303183. https://doi.org/10.1371/journal.pone.0303183

[4] He, Z., Davila, D., Bi, S., Wang, T., & Hou, T. (2025). Machine learning for cybersecurity: A survey of applications, adversarial challenges, and future research directions. *Electronics*, *14*(23), 4563. https://doi.org/10.3390/electronics14234563

[5] Marques, P., Rhode, M., & Gashi, I. (2021). Waste not: Using diverse neural networks from hyperparameter search for improved malware detection. *Computers & Security*, *108*, 102339. https://doi.org/10.1016/j.cose.2021.102339

[6] Newman, M. E. J. (2018). *Networks*. UK: Oxford University Press.

[7] Ju, W., Fang, Z., Gu, Y., Liu, Z., Long, Q., Qiao, Z., . . . , & Zhang, M. (2024). A comprehensive survey on deep graph representation learning. *Neural Networks*, *173*, 106207. https://doi.org/10.1016/j.neunet.2024.106207

[8] Ait Rai, K., Machkour, M., & Antari, J. (2023). Influential nodes identification in complex networks: A comprehensive literature review. *Beni-Suef University Journal of Basic and Applied Sciences*, *12*(1), 18. https://doi.org/10.1186/s43088-023-00357-w

[9] Ekle, O. A., & Eberle, W. (2024). Anomaly detection in dynamic graphs: A comprehensive survey. *ACM Transactions on Knowledge Discovery from Data*, *18*(8), 1–44. https://doi.org/10.1145/3669906

[10] Qing, H. (2024). *Community detection in multi-layer bipartite networks*. arXiv. https://doi.org/10.48550/arXiv.2405.04711

[11] Li, J., Lai, S., Shuai, Z., Tan, Y., Jia, Y., Yu, M., . . . , & Lu, Y. (2024). A comprehensive review of community detection in graphs. *Neurocomputing*, *600*, 128169. https://doi.org/10.1016/j.neucom.2024.128169

[12] Diboune, A., Slimani, H., Nacer, H., & Beghdad Bey, K. (2024). A comprehensive survey on community detection methods and applications in complex information networks. *Social Network Analysis and Mining*, *14*(1), 93. https://doi.org/10.1007/s13278-024-01246-5

[13] Jin, D., Yu, Z., Jiao, P., Pan, S., He, D., Wu, J., . . . , & Zhang, W. (2021). A survey of community detection approaches: From statistical modeling to deep learning. *IEEE Transactions on Knowledge and Data Engineering*, *35*(2), 1149–1170. https://doi.org/10.1109/TKDE.2021.3104155

[14] Alvarez, P. A., Ishizaka, A., & Martinez, L. (2021). Multiple-criteria decision-making sorting methods: A survey. *Expert Systems with Applications*, *183*, 115368. https://doi.org/10.1016/j.eswa.2021.115368

[15] Stofkova, J., Krejnus, M., Stofkova, K. R., Malega, P., & Binasova, V. (2022). Use of the analytic hierarchy process and selected methods in the managerial decision-making process in the context of sustainable development. *Sustainability*, *14*(18), 11546. https://doi.org/10.3390/su141811546

[16] Meccawy, M., Bayazed, A. A., Al-Abdullah, B., & Algamdi, H. (2023). Automatic essay scoring for Arabic short answer questions using text mining techniques. *International Journal of*

Advanced Computer Science and Applications, *14*(6), 768–775. https://doi.org/10.14569/IJACSA.2023.0140682

[17] Ullah, A., Khan, N., Khawaja, F. R., Sun, Z., & Mendes, J. F. F. (2026). Network community detection through enhanced effective distance information: A scalable and parameter-free mechanism. *Chaos, Solitons & Fractals*, *202*, 117537. https://doi.org/10.1016/j.chaos.2025.117537

[18] Li, Y., Jia, C., & Yu, J. (2015). A parameter-free community detection method based on centrality and dispersion of nodes in complex networks. *Physica A: Statistical Mechanics and its Applications*, *438*, 321–334. https://doi.org/10.1016/j.physa.2015.06.043

[19] Li, X., Zhen, X., Qi, X., Han, H., Zhang, L., & Han, Z. (2023). Dynamic community detection based on graph convolutional networks and contrastive learning. *Chaos, Solitons & Fractals*, *176*, 114157. https://doi.org/10.1016/j.chaos.2023.114157

[20] Mao, W., Zhu, H., Wu, H., Lu, Y., & Wang, H. (2023). Forecasting and trading credit default swap indices using a deep learning model integrating Merton and LSTMs. *Expert Systems with Applications*, *213*, 119012. https://doi.org/10.1016/j.eswa.2022.119012

[21] Kazemzadeh, F., Safaei, A. A., & Mirzarezaee, M. (2022). Influence maximization in social networks using effective community detection. *Physica A: Statistical Mechanics and its Applications*, *598*. 127314. https://doi.org/10.1016/j.physa.2022.127314

[22] Mukhtar, M. F., Abas, Z. A., Rasib, A. H. A., Anuar, S. H. H., Zaki, N. H. M., Rahman, A. F. N. A., . . . , & Shibghatullah, A. S. (2023). Hybrid global structure model for unraveling influential nodes in complex networks. *International Journal of Advanced Computer Science and Applications*, *14*(6), 724–730. https://doi.org/10.14569/IJACSA.2023.0140677

[23] Hafiene, N., Karoui, W., & Ben Romdhane, L. (2019). Influential nodes detection in dynamic social networks. In *Business Information Systems: 22nd International Conference*, 62–73. https://doi.org/10.1007/978-3-030-20482-2_6

[24] Sheikhahmadi, A., Veisi, F., Sheikhahmadi, A., & Mohammadimajd, S. (2022). A multi-attribute method for ranking influential nodes in complex networks. *PLOS one*, *17*(11), e0278129. https://doi.org/10.1371/journal.pone.0278129

[25] Abdel-Basset, M., Mohamed, R., & Chang, V. (2025). A multi-criteria decision-making framework to evaluate the impact of industry 5.0 technologies: Case study, lessons learned, challenges and future directions. *Information Systems Frontiers*, *27*(2), 791–821. https://doi.org/10.1007/s10796-024-10472-3

[26] Niemcewicz, P. (2021). The use of the multi-criteria AHP method to select a cloud computing provider. *Procedia Computer Science*, *192*, 2558–2567. https://doi.org/10.1016/j.procs.2021.09.025

[27] Zehavi, M. (2022). Parameterized analysis and crossing minimization problems. *Computer Science Review*, *45*, 100490. https://doi.org/10.1016/j.cosrev.2022.100490

[28] Madzík, P., & Falát, L. (2022). State-of-the-art on analytic hierarchy process in the last 40 years: Literature review based on Latent Dirichlet Allocation topic modelling. *PLOS One*, *17*(5), e0268777. https://doi.org/10.1371/journal.pone.0268777

[29] Lenharth, A., Nguyen, D., & Pingali, K. (2016). Parallel graph analytics. *Communications of the ACM*, *59*(5), 78–87. https://doi.org/10.1145/2901919

[30] Rozemberczki, B., Allen, C., & Sarkar, R. (2021). Multi-scale attributed node embedding. *Journal of Complex Networks*, *9*(2), cnab014. https://doi.org/10.1093/comnet/cnab014

[31] Javed, M. A., Younis, M. S., Latif, S., Qadir, J., & Baig, A. (2018). Community detection in networks: A multidisciplinary review. *Journal of Network and Computer Applications*, *108*, 87–111. https://doi.org/10.1016/j.jnca.2018.02.011

[32] Alharbi, N., Alkalifah, B., Alqarawi, G., & Rassam, M. A. (2024). Countering social media cybercrime using deep learning: Instagram fake accounts detection. *Future Internet*, *16*(10), 367. https://doi.org/10.3390/fi16100367

[33] Mahdiraji, H. A., Razghandi, E., & Hatami-Marbini, A. (2021). Overlapping coalition formation in game theory: A state-of-the-art review. *Expert Systems with Applications*, *174*, 114752. https://doi.org/10.1016/j.eswa.2021.114752

[34] Rafieizadeh, H., Zare, H., Ghassemi Parsa, M., & Cherifi, H. (2025). Dual contrastive learning-based reconstruction for anomaly detection in attributed networks. *PLOS One*, *20*(11), e0335135. https://doi.org/10.1371/journal.pone.0335135

How to Cite: Mohammad, F., & Vikram, S. (2026). A Hybrid Approach for Prominent Node Detection in Social Media Networks Using Modified Cluster Walktrap with Analytical Hierarchy Process. *Journal of Data Science and Intelligent Systems*. https://doi.org/10.47852/bonviewJDSIS62029981