This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



# Exploring Influence Maximization in Criminal Social Networks to Identify Key Individuals and Patterns

Sunil Kumar Meena, Shashank Sheshar Singh, and Kuldeep Singh

***Abstract*—Criminal networks are the social networks that show the individual and their links involved in illegal activities. The influence maximization (IM) problem aims to find the top influential nodes in social networks. Existing studies of criminal networks do not utilize the IM problem to understand its dynamics. The proposed work introduces IM to investigate its role in criminal social networks and proposes an algorithm IM-C that finds the key influential nodes in criminal social networks. This work examines key influencers and their influence on criminal networks. Further, by introducing IM, this study inspects the role of IM in geographic impact, resilience and vulnerability in criminal networks and the influence of centralities-based key users. The experiments are conducted on six datasets, and the experimental results show that the proposed algorithm outperformed the benchmark algorithms. The findings of key users, their influence, geographic impact, resilience and vulnerability provide insights into disrupting criminal networks and designing targeted intervention strategies for crime reduction. Thus, through experiment results, this work demonstrates that IM in criminal networks provides an impactful novel approach to uncover insights such as identifying key influencers, understanding their impact, and assessing network resilience and vulnerability.**

***Index Terms*—Criminal social network, cybercrime, influence maximization, social network analysis.**

## I. Introduction

SOCIAL network analysis (SNA) [1] is a method of analyzing the structure of social interactions by examining relationships between individuals and groups. SNA is used to study user behavior and analyze patterns in social networks. This approach has been applied to fields such as criminology [2], industry [3], education [4], and policymaking [5]. Influence maximization (IM) [6] is a fundamental problem in SNA that focuses on identifying a small subset of influential nodes to maximize the spread of information, ideas, or behaviors. IM problem is computationally challenging due to its combinatorial nature and NP-hard complexity. Rooted in viral marketing [7] and epidemic modeling [4], [6], IM has many applications. It includes marketing [8], public health campaigns [4], [6], political mobilization [5], misinformation containment [8], etc.

In recent years, crime networks have been extensively studied using SNA [2]. However, existing approaches [9] often overlook the influence that key individuals exert on the network. Traditional methods [2], [9], [10] focus on identifying central figures based on degree or community structures, but they fail to capture the cascading effects of influence spread, resilience, and vulnerability within criminal networks. These insights yield significant patterns in criminal networks. However, existing studies [2], [9], [10] failed to utilize IM to analyze criminal networks. The proposed work introduces IM as a novel approach to analyzing crime networks. This allows for a deeper understanding of key influencers, the spread of key nodes, and network resilience and vulnerability in crime networks.

To address the existing research gap, this work proposes an IM-based novel algorithm that identifies the key influential nodes in the criminal network. By utilizing the IM approach, this work studies the impact of key nodes in the network. Additionally, the proposed work conducts key analysis from the perspective of network resilience, geographical impact, etc. The contributions of the proposed work are as follows:

1) *First application of IM in crime networks*: This work applies IM to crime analysis for the first time to study the key users and patterns in criminal social networks.
2) This work proposes a novel IM-C algorithm to find the key influential nodes in the criminal social network.
3) This work conducts five key analyses in criminal networks, i.e.
   a) Identifying influential nodes in criminal networks.
   b) Analysis of influence spread of criminal behavior.
   c) Examining their geographic impact in the network.
   d) Evaluating network resilience and vulnerability.
   e) Analysis and comparison of seed nodes based on centralities.
4) This work conducted experiments on six criminal networks and compared the results with criminal SNA-based and IM-based algorithms. This analysis provides actionable information for law enforcement by effectively targeting key nodes.

Received 23 December 2024; revised 10 April 2025; accepted 12 May 2025. *(Corresponding author: Kuldeep Singh.)*

Sunil Kumar Meena and Kuldeep Singh are with the Department of Computer Science, University of Delhi 110007, New Delhi, India (e-mail: skmeena@cs.du.ac.in; ksingh@cs.du.ac.in).

Shashank Sheshar Singh is with the Computer Science and Engineering Department, Thapar Institute of Engineering and Technology, Patiala, Punjab 147004, India (e-mail: shashank.sheshar@gmail.com; shashank.singh@thapar.edu).

Digital Object Identifier 10.1109/TCSS.2025.3571963

2329-924X © 2025 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: PSG COLLEGE OF TECHNOLOGY. Downloaded on July 16,2025 at 06:07:49 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



## II. Related Work

Influence maximization (IM) [6] has been extensively studied since its introduction by Domingos and Richardson [7] for viral marketing. Further, it is formalized as a discrete optimization problem by Kempe et al. [6]. The IM problem focuses on identifying a subset of influential nodes in a social network to maximize influence spread. To estimate the influence, it uses diffusion models like the independent cascade (IC) [6] and linear threshold (LT) [6] models. The authors [6] also proved that the IM problem is an NP-hard problem under these models. Early research primarily concentrated on Greedy algorithms [6] that provide near-optimal solutions with approximation guarantees. However, due to the expensive computation of simulations, the Greedy algorithm is computationally expensive. To address scalability challenges, researchers proposed CELF [11] that utilizes the submodular property to reduce computational overhead while maintaining solution quality. With advancements in computational resources, evolutionary algorithms [12], metaheuristics [13], and swarm-based optimization [14] have been explored to enhance performance. Dynamic networks have also attracted attention, leading to the development of algorithms [12], [15] that adapt to temporal changes. Recent trends have shifted towards incorporating diversity [15], fairness [16], and user preferences [17] in IM. To address diversity, the work [18] proposed the G-HIST algorithm for the composite community-aware diversified IM (CC-DIM) problem. It utilizes a two-stage approach, first selecting a small sentinel set to reduce the size of random generalized reverse reachable (G-RR) sets, and then applying an unbiased estimator to improve efficiency. This method yields high-quality approximate solutions. Quantum computing is emerging as a promising approach for solving IM [19], leveraging quantum diffusion models and optimization techniques for better efficiency in complex networks. Further, to handle the large-scale networks, the authors [20] proposed the ToupleGDD framework and integrated deep reinforcement learning (DRL) with graph-based diffusion dynamics to address IM. The paper [21] provides an extensive survey of the machine learning-based IM algorithms.

### A. Application-Based Influence Maximization Algorithms

Influence maximization (IM) has diverse applications across domains. In marketing, to maximize return on investment, IM [22] aids companies in selecting influential individuals to promote products or services. The domain of e-commerce benefits through personalized promotions and dynamic pricing strategies [17]. In public health [4], [6], IM helps identify key individuals for effective dissemination of health information, vaccination campaigns, and disease awareness. It plays a crucial role in political campaigns [5] for targeting influential voters or spreading awareness about social issues. Social media platforms utilize IM for content recommendation [23] and user engagement strategies [24]. Additionally, IM is used in rumor control [8], [25], ensuring misinformation is curtailed by targeting specific nodes for truthful information spread. A variant of rumor control, the work [8] protects the targeted users from being influenced by rumors. This work focuses on identifying a set of nodes for protection to minimize the spread of harmful influence. IM [26] also finds utility in crisis management, where the rapid dissemination of critical information during natural disasters or emergencies can save lives.

### B. Criminal Network-Based Social Network Analysis

Existing studies [2], [10], [27], [28], [29] conducted the scientific study on organized crime and networks. These studies have applied social network analysis tools to study the social dynamics and structure of criminal networks. Few studies [10], [27] aim to find the key nodes in the criminal networks. The study [10] discusses the use of node removal as a strategy to destabilize criminal networks. It assesses the impact of removing nodes in criminal networks. The work [29] discusses the application of social network analysis to identify individuals associated with organized criminal networks. The criminal network focuses on terrorist organizations. The authors suggest a strategy to find key nodes using identifying codes. It [29] emphasizes the identification and monitoring of key individuals involved in criminal activities.

The above literature also mentioned that few studies aim to find the key nodes in criminal networks. However, to the best of our knowledge, in existing studies, IM has not been utilized to find key influencers and other patterns in criminal social networks. Therefore, these methods do not capture the cascading influence of individuals or the resilience of the network when specific nodes are removed. Our work extends IM techniques to crime analysis to provide an IM-based understanding of influence spread rather than just structural prominence.

## III. Proposed Work

This section discusses the problem formulation, description of the proposed algorithm, details methodology used to study criminal social network analysis, and explanation of analyses using an example.

### A. Problem Formulation

This work aims to understand the dynamics of criminal networks using influence maximization (IM). Although IM has been widely studied [4], [6], [30] in social networks, its application to criminal networks remains unexplored. Traditional research fails to address unique aspects of criminal networks using IM, such as key influential users (seed nodes or seed set), geographic proximity and network resilience. Bridging this gap can improve crime prevention by understanding how influence spreads in these contexts. To bridge the gap, the proposed work introduces IM to the study of criminal social networks and proposes an algorithm IM-C to identify key influencers. Further, this work analyses the influence, assesses geographic impacts, evaluates network resilience and vulnerability, and analyses the influence of key users selected using different centralities, i.e., degree centrality [31], betweenness centrality [31], and closeness centrality [31]. To do so, the subsequent sections provide detailed descriptions of these analyses. The IM problem is defined as follows:

Authorized licensed use limited to: PSG COLLEGE OF TECHNOLOGY. Downloaded on July 16,2025 at 06:07:49 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



*Definition III.1 (Influence Maximization (IM)):* Influence maximization (IM) identifies a subset $S \subseteq V$ of $k$ seed nodes in a social network $G = (V, E)$ such that the expected spread of influence $\sigma(S)$ (using the given diffusion model) is maximized. Mathematically, it is expressed as

$$S^* = \underset{S \subseteq V, |S|=k}{\arg\max} |\sigma(S)|$$

where $\sigma(S)$ is the set of influence nodes achieved through a given diffusion model: This work utilizes the independent cascade (IC) model [6] as the diffusion model in experiments.

*Definition III.2 (Independent Cascade (IC) $\sigma(.)$ - Diffusion Model [6])*: The independent cascade (IC) model [6] represents influence propagation in a social network $G = (V, E)$, where $V$ is the set of nodes, and $E$ is the set of edges. Each edge $(u, v) \in E$ is associated with a influence probability/weight $p_{uv} \in [0, 1]$, which denotes the likelihood that node $u$ influences node $v$. At any time step $t$, let $A_t$ denote the set of nodes influenced during that step. When a node $u$ becomes influence at time $t$, it gets a single chance to influence each of its uninfluenced neighbors $v$ by performing the Bernoulli trial with probability $p_{uv}$ and probability of success $p_s$, independent of other neighbors. The process continues iteratively, with newly influenced nodes attempting to activate their neighbors in subsequent time steps until no further activations occur. The influence spread, $\sigma(S)$, starting from an initial seed set $S \subseteq V$, is defined as

$$\sigma(S) = \left\{ \bigcup_{t=0}^{\infty} A_t \,\middle|\, A_0 = S \right\}$$

where $\bigcup_{t=0}^{\infty} A_t$ represents the total set of nodes activated over time. Thus, IC model $\{\bigcup_{t=0}^{\infty} A_t\}$ is a set of the influenced nodes that have influenced using an initial seed set $S$.

### B. Proposed Algorithm of IM for Criminal Networks: IM-C

The goal of the IM for criminal networks (IM-C) algorithm is to select a seed set $S \subseteq V$ of size $k$ from a graph $G = (V, E)$, where $V$ represents the set of nodes and $E$ the set of edges. The objective is to maximize the combined score $f(S)$. Thus the key influencer $v^*$ is added to set of key influencer (seed set $S$) as

$$v^* = \arg \max_{v \in V \setminus S, |S| <= k} f(v, S)$$

where $f(v, S)$ is defined as follows:

$$f(v, S) = \lambda \cdot \Delta\sigma(v \mid S) + (1 - \lambda) \cdot D(v, S).$$

Here, $S$ represents the current seed set. $v \in V \setminus S$ be a candidate node not yet in the seed set. $\Delta\sigma(v \mid S)$ represents the marginal influence gain of the node $v$. $D(v, S)$ represent the diversity score of node $v$ with respect to $S$. $\lambda \in [0, 1]$ be a trade-off parameter balancing influence and diversity. Marginal influence gain ($\Delta\sigma(v \mid S)$) is defined as follows:

$$\Delta\sigma(v \mid S) = \frac{|\sigma(S \cup \{v\})| - |\sigma(S)|}{|\sigma(S \cup \{v\})|}.$$

**Algorithm 1:** Influence Maximization Algorithm for Criminal Networks: IM-C.

1: **Input:** Graph $G = (V, E)$, seed set size $k$, trade-off parameter $\lambda$
2: **Output:** Seed set $S$
3: $S \leftarrow \emptyset$ {Start with an empty seed set}
4: $V \leftarrow G.Nodes()$ {Take all nodes in $V$}
5: **for** each iteration $i = 1$ to $k$ **do**
6:   best_node $\leftarrow$ None
7:   best_score $\leftarrow -\infty$
8:   **for** each $v \in V$ **do**
9:     **if** $v \in S$ **then**
10:       **continue** {Skip nodes already in the seed set}
11:     **end if**
12:     Compute $\Delta\sigma(v \mid S) = \frac{\sigma(S \cup \{v\}) - \sigma(S)}{\sigma(S \cup \{v\})}$
13:     Compute $D(v, S) = \frac{1}{1 + \frac{1}{|S|} \sum_{s \in S} d(v, s)}$
14:     Score$(v) = \lambda \cdot \Delta\sigma(v \mid S) + (1 - \lambda) \cdot D(v, S)$
15:     **if** Score$(v) >$ best_score **then**
16:       best_node $\leftarrow v$
17:       best_score $\leftarrow$ Score$(v)$
18:     **end if**
19:   **end for**
20:   **if** best_node $\neq$ None **then**
21:     $S \leftarrow S \cup \{$best_node$\}$
22:     $V \leftarrow V \setminus \{$best_node$\}$
23:   **end if**
24: **end for**
25: **Return** $S$

This normalized gain measures the relative increase in influence spread when $v$ is added to $S$. Diversity score ($D(v, S)$) is defined as

$$D(v, S) = \frac{1}{1 + \frac{1}{|S|} \sum_{s \in S} d(v, s)}$$

where $d(v, s)$ is the shortest path distance between $v$ and $s$. This score encourages selecting nodes that are farther apart from the existing seeds. This component is used to select the key nodes from different places in the network. By doing this, the overlapping influence [32] is also reduced, enhancing overall influence.

Thus, $f(v, S)$ select the node $v$ which has higher marginal gain as well as far from the already selected nodes. This formulation gives us a way to select node that has maximum influence as well as targeting the different locations in the network.

*1) Algorithm Description:* This section briefly explains the proposed algorithm IM-C as depicted in Algorithm 1 using pseudocode. Lines 1 and 2 show the input and output of the algorithm. Lines 3 and 4 initialize the seed set $S$ and all nodes in $V$. The loop in lines 5–24 selects the key influential nodes iteratively. Lines 6 and 7 initialize the $best\_node$ and $best\_score$ to keep the record of the most influential node. The loop in lines 8 – 19 iterates over all the nodes and gives the key nodes to add to the seed set $S$. Lines 9–11 filter if the current node is already

Authorized licensed use limited to: PSG COLLEGE OF TECHNOLOGY. Downloaded on July 16,2025 at 06:07:49 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



selected in the seed set. Line 12 computes the marginal gain of node $v$. Line 13 computes the diversity score of node $v$. Line 14 computes the final score of node $v$. The detailed explanation of lines 12–14 is given in Section III-B. Lines 15–18 select the node that has the highest score. Lines 20–23 add the node that has the highest score in the seed set. Line 25 returns the seed set containing the key influential nodes in the network.

The complexity of the proposed algorithm is computed as follows. For every value of $k$, it finds the marginal gain and the shortest path between node $v_i$ and $v_j$. The complexity of the algorithm is computed as $O(k(V(E+RE)))$. Further, it can be written as $O(kVRE)$. Here, $k$ is the seed set size, $V$ is the number of nodes, $E$ is the number of edges, and $R$ is the number of Monte-Carlo simulations.

### C. Methodologies for Criminal Network Analysis

This section elaborates the suggested methodologies for the analysis to uncover the pattern of the criminal network.

*1) Identifying Key Influencers (Seed Nodes) in Criminal Networks:* In criminal networks, all individuals do not hold equal significance; a small subset of nodes (individuals) often plays a disproportionately large role in coordinating and propagating criminal activities. Identifying these key influencers is essential for law enforcement agencies to efficiently allocate resources to dismantle criminal organizations. This work applied the following IM methods to find the key influential nodes: the proposed IM-C (Algorithm 1), CELF [11], CSO [12], FIMMOGA [16], RIS [33], D-ALO [13], ToupleGDD [20] and SNA-based PageRank [34], and degree [10], [31] algorithms to identify key influencers in the criminal network.

*2) Influence Spread of Key Nodes (Seed Nodes) in Criminal Networks:* The existing studies in criminal networks Section II find the key nodes in criminal networks. This work applies the IM to find key nodes in the network. After finding the key nodes in the previous section, this analysis investigates the impact of the chosen key influential nodes. To estimate their impact, we utilize the IC diffusion model [6] to find the nodes that are influenced by these key nodes. This analysis enables us to model these propagation patterns, providing insights into the mechanisms of influence spread and the potential for disrupting it. This analysis bridges a significant research gap by quantitatively examining the diffusion of influence within criminal networks.

To estimate the influence, this work utilizes the IC model $\sigma(.)$ [6], as given in work [6]. The IC model takes a social graph $G$ and a seed set (key nodes) as input and returns a set of influence nodes as output. It influences the nodes in the following manner:

a) At $t=0$, the seed set $S$ is activated.

b) At each subsequent time step $t>0$, an activated node $u$ has a single chance to activate its neighbor $v$ with probability $p(u,v)$.

c) The process continues until no further activations occur.

Therefore, to compare the results, we show the number of influenced nodes $|\sigma(S)|$ as "Activated nodes" in the experimental results.

*3) Geographic Impact on Crime Networks:* The geographic impact of crime spread in criminal networks provides critical insights into how influence propagates spatially and socially. Influence nodes in these networks often influence not only their immediate connections but also extend their impact to secondary neighbors. This cascading effect can lead to the emergence of crime hotspots and geographically vulnerable regions. The 2-hop neighborhood is a well-established metric in social network analysis to assess the local influence and reach [35], [36]. It captures both direct and indirect interactions. The 1-hop neighborhood includes only immediate connections. However, it may be too restrictive to assess broader influence. The 2-hop neighborhood includes friends-of-friends, which provides a better approximation of influence spread within local clusters. Modeling and quantifying this propagation up to the 2-hop neighborhood of influential key nodes as geographic impact is useful for identifying the patterns in criminal networks.

Let $G=(V,E)$ represent the criminal network. Let $S\subseteq V$ denote the set of seed nodes or key influential nodes identified through criminal network analysis or IM algorithms. The geographical impact is defined as the number of nodes up to the 2-hop neighborhood of the seed nodes, and it is expressed mathematically as $\textit{Geographic impact} = \sum_{u\in S}|N_1(u)\cup N_2(u)|$, where

a) $N_1(u)$ is the set of direct neighbors (1-hop) of node $u$, i.e., $N_1(u)=\{v\in V:(u,v)\in E \text{ AND } u\in S\}$.

b) $N_2(u)$ is the set of 2-hop neighbors of node $u$, i.e., $N_2(u)=\{w\in V:\exists v\in N_1(u) \text{ and } (v,w)\in E\}$.

By restricting the analysis to 1-hop and 2-hop neighbors, the geographic impact captures the local propagation dynamics. Geographical terms are indicated in this analysis in the sense of nearby nodes up to 2-hop neighbor nodes. This formulation reveals how originating from seed nodes diffuses through immediate and secondary connections. Policymakers use these findings to identify geographic regions prone to crime escalation and to design targeted interventions.

*4) Resilience and Vulnerability of the Network:* To evaluate the resilience and vulnerability of the criminal network after the removal of influenced nodes, we define resilience as the fraction of the network that remains functional after the removal of these nodes. Vulnerability is considered the extent to which the network is disrupted or fragmented due to the influence spread and subsequent node removal. It is the complement of resilience, meaning that as resilience decreases (indicating a large influence spread), the network becomes more vulnerable to disruption. The resilience $R$ can be mathematically modeled as: $R=1-|\sigma(S)|/N_{\text{total}}$, where $N_{\text{total}}$ is the total number of nodes in the original graph $G$ and $\sigma(S)$ represents the set of influenced nodes (nodes activated by the seed set under the diffusion model). This formula quantifies the network's ability to sustain its structure and connectivity after the removal of influenced nodes. Thus, based on the formula, if the number of influence nodes is higher, then $|\sigma(S)|$ is higher; then $|\sigma(S)|/N_{\text{total}}$ will also be higher and $R$ becomes low. Thus, it has high resilience and low vulnerability. This indicates we have identified the higher number of influenced nodes in the criminal network and can target them for further actions. If $|\sigma(S)|$ is low, then $|\sigma(S)|/N_{\text{total}}$ will also be low and $R$ becomes high. This indicates that we have not identified the most influential key

Authorized licensed use limited to: PSG COLLEGE OF TECHNOLOGY. Downloaded on July 16,2025 at 06:07:49 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



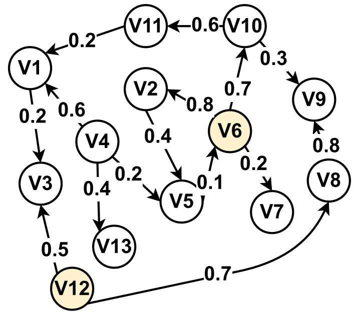


Fig. 1. Given a social graph for running example.

nodes; therefore, their influence is low, and influenced nodes in criminal networks can not be targeted at higher frequency. Hence, it has low resilience and high vulnerability. Based on the above explanations, this work aims to minimize $R$.

*5) Key Influential Nodes Based on Centralities:* In SNA, to find the central nodes, most of the work utilizes the various centralities. These centralities are defined by various factors such as degree, paths, etc. These centralities have their own importance depending on the applications. This work analyses the influence impact of the key user using centrality measures. It employs three centrality measures to find the key influential nodes i.e., degree, betweenness, and closeness centrality.

After identifying seed nodes using various centrality measures, their influence is quantified using the IC model. By simulating the influence propagation for seed nodes selected based on degree, betweenness, and closeness centralities, this analysis evaluates their effectiveness in spreading influence. For example, high-degree nodes are likely to activate their immediate neighbors, while high-betweenness nodes may bridge different network communities, influencing disconnected subgraphs. Closeness centrality nodes maximize overall efficiency in influence dissemination. It helps in analyzing the disrupting network connectivity by removing high-betweenness nodes or containing rapid influence spread by targeting high-closeness nodes.

### D. Example

This section explains all the analysis using an example. Consider the graph as shown in Fig. 1. The graph has 13 nodes and 15 edges. We utilize the IC model [6] to find the influence spread from the nodes. The IC model's parameter activation threshold is taken as 0.5 [37]. The influence probability/weight is shown on the edges in Fig. 1 and assigned randomly. By utilizing the diffusion IC model, the node and their influenced nodes are {{**V1:** V1}, {**V2:** V2}, {**V3:** V3}, {**V4:** V1, V4}, {**V5:** V5}, {**V6:** V2, V6, V10, V11}, {**V7:** V7}, {**V8:** V8, V9}, {**V9:** V9}, {**V10:** V10, V11}, {**V11:** V11}, {**V12:** V3, V8, V9, V12}, {**V13:** V13}}. Let's assume a ranking-based algorithm selects the seed nodes $v \in V$ based on the number of influenced nodes ($|\sigma(v)|$). The above analyses for the given example are done as follows:

1) *Identifying key influencers in criminal network*: This analysis finds the key influencers in the network. The ranking-based method selects nodes $V6$ and $V12$ as the key influential nodes or seed nodes in the criminal network, as shown in Fig. 1.
2) *Spread of criminal behavior*: This analysis demonstrates the influence nodes from the selected key influential nodes in the criminal network. For the given example, nodes $V6$ and $V12$ are selected as seed nodes. These seed nodes $S = \{V6, V12\}$ influences $\{V2, V6, V10, V11, V3, V8, V9, V12\}$ nodes, i.e., 8 nodes are influenced using IC model [6].
3) *Geographic impact on crime spread*: This analysis indicates the geographical impact by counting the nodes up to 2-hop neighbors of seed nodes. In the running example, the nodes $V6$ and $V12$ are the key nodes (seed nodes) and up to 2-hop neighbors nodes are $\{V3, V8, V9, V7, V10, V2, V5, V11\}$. It means that the geographical impact of the key nodes is 8.
4) *Resilience and vulnerability of the network*: This analysis indicates if the key influential nodes are identified, then how vulnerable the network remains (Section III-C4). The key nodes $V6$ and $V12$ influence 8 nodes. Thus, resilience and vulnerability score $R$ becomes $1 - |\sigma(S)|/V$, i.e. $1 - 8/13 = 0.39$. As mentioned above, if $R$ is low, then it has high resilience and low vulnerability. Thus, in this example, $R$ is comparatively low. Thus, we can say 8 out of 13 nodes have been identified, and respective actions can be taken. Thus, it has high resilience and low vulnerability.
5) *Key nodes based on centralities*: This analysis demonstrates the importance of centrality measures with influence diffusion. It shows the influence effect of the seed nodes chosen using centrality measures. In the running example, based on degree centrality, nodes $\{V4, V6\}$ are chosen as seed nodes which influences 6 nodes. Based on betweenness centrality, nodes $\{V6, V10\}$ are chosen as seed nodes, which influences 4 nodes. Based on closeness centrality, nodes $\{V3, V9\}$ are chosen as seed nodes which influences 2 nodes. Thus, degree centrality has a higher influence in this example.

## IV. Experimental Results

This section presents the analysis of experimental results for the respective analysis discussed above. We have taken the seed set size as $\{2, 5, 10, 15, 20, 25\}$, which indicates the number of key influence nodes to be extracted from the criminal networks. We utilize the IC model as the diffusion model [6] with an activation threshold of 0.75 [38]. The experimental analysis compared the proposed algorithm IM-C[1] with benchmark algorithms CELF [11], CSO [12], FIMMOGA [16], PageRank [34], Degree [31], RIS [33], D-ALO [13], and ToupleGDD [20].

### A. Criminal Social Networks

This work models the six criminal datasets as criminal social networks. To the best of our knowledge, the information on the criminal datasets is not given in the form of nodes and

[1]Code-https://github.com/sunil120meena/IM/tree/main/IM_C

Authorized licensed use limited to: PSG COLLEGE OF TECHNOLOGY. Downloaded on July 16,2025 at 06:07:49 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



TABLE I
KEY NODES SELECTED FROM THE CRIMINAL NETWORKS

| Datasets | IM-C | CELF | CSO | FIMMOGA | PageRank | DEGREE |
|---|---|---|---|---|---|---|
| CyberCrime | ‘Fraud’, ‘Causing_Disrepute’, ‘Developing_own_business’, ‘Extortion’, ‘Sexual_Exploitation’ | ‘Causing_Disrepute’, ‘Sexual_Exploitation’, ‘Disrupt_Public_Service’, ‘Extortion’, ‘Personal_Revenge’ | ‘Anger’, ‘Personal_Revenge’, ‘Extortion’, ‘Fraud’, ‘Prank’ | ‘Anger’, ‘Causing_Disrepute’, ‘Personal_Revenge’, ‘Spreading_Piracy’, ‘Abetment_to_Suicide’ | ‘Others’, ‘Abetment_to_Suicide’, ‘Psycho_or_Pervert’, ‘Spreading_Piracy’, ‘Steal_Information’ | ‘Others’, ‘Personal_Revenge’, ‘Anger’, ‘Abetment_to_Suicide’, ‘Fraud’ |
| London | ‘Camden’, ‘Bromley’, ‘Richmond_upon_Thames’, ‘Ealing’, ‘Haringey’ | ‘Bromley’, ‘Greenwich’, ‘Richmond_upon_Thames’, ‘Redbridge’, ‘Ealing’ | ‘Redbridge’, ‘Croydon’, ‘Greenwich’, ‘Richmond_upon_Thames’, ‘Bromley’ | ‘Westminster’, ‘Islington’, ‘Ealing’, ‘Kensington_and_Chelsea’, ‘Waltham_Forest’ | ‘City_of_London’, ‘Hammersmith_and_Fulham’, ‘Merton’, ‘Tower_Hamlets’, ‘Islington’ | ‘City_of_London’, ‘Croydon’, ‘Greenwich’, ‘Merton’, ‘Bromley’ |
| CAC | ‘Gujarat’, ‘Jharkhand’, ‘Chhattisgarh’, ‘Goa’, ‘Arunachal_Pradesh’ | ‘Chhattisgarh’, ‘Jharkhand’, ‘Lakshadweep’, ‘West_Bengal’, ‘Assam’ | ‘Nagaland’, ‘Rajasthan’, ‘Odisha’, ‘Maharashtra’, ‘Lakshadweep’ | ‘Meghalaya’, ‘Tamil_Nadu’, ‘Bihar’, ‘Madhya_Pradesh’, ‘Haryana’ | ‘Gujarat’, ‘Haryana’, ‘A&N_Islands’, ‘Chandigarh’, ‘Lakshadweep’ | ‘Gujarat’, ‘Bihar’, ‘Bihar’, ‘Telangana’, ‘Karnataka’ |
| Mafia Meeting | ‘N91’, ‘N18’, ‘N51’, ‘N12’, ‘N6’ | ‘N18’, ‘N12’, ‘N3’, ‘N34’, ‘N51’ | N34’, ‘N18’, ‘N36’, ‘N22’, ‘N12’ | ‘N4’, ‘N98’, ‘N97’, ‘N96’, ‘N18’ | ‘N68’, ‘N25’, ‘N18’, ‘N48’, ‘N22’ | ‘N18’, ‘N68’, ‘N47’, ‘N22’, ‘N27’ |
| Global Terrorism | ‘Anti-Capitalist_Brigades’, ‘Unknown’, ‘New_People‘s-_Army_(NPA)’, ‘Palestinians’, ‘National_Liberation-Army_of-Colombia_(ELN)’ | ‘New_People‘s-_Army_(NPA)’, ‘Unknown’, ‘Palestinians’, ‘Basque_Fatherland_-and_Freedom-_(ETA)’, ‘Palestine_Liberation-_Organization_(PLO)’ | ‘Free_and_Democratic_-Society_of_Eastern_-Kurdistan_(KODAR)’, ‘German_-Resistance_-Movement’, ‘Kim_Pyongnil_Supporters’, ‘Eva_Peron_Organization’, ‘Major_Muhammad_-Zuhnyn_Group’ | ‘Antonia_Martinez-_Student_-Commandos_(AMSC)’, ‘Orakzai_-Freedom_-Movement’, ‘Sons_-of_-Liberty’, ‘Right-Wing_National_-Youth_Front’, ‘Hekla_-Reception_-Committee-Initiative_-for_-More_-Social_-Eruptions’ | ‘Jamaat_-Nusrat_al-Islam_wal_-Muslimin_(JNIM)’, ‘Gorkha_-Janmukti_-Morcha_(GJM)’, ‘Gorkha_-Liberation_-Army_(GLA)’, "Hay‘at_-Tahrir_-al-Sham", ‘Arakan_-Rohingya_-Salvation_-Army_(ARSA)’ | ‘Unknown’, ‘New_People‘s-_Army_(NPA)’, ‘Palestinians’, ‘Basque_-Fatherland_-and_Freedom-_(ETA)’, ‘Irish_-Republican_-Army_(IRA)’ |
| Boston | ‘STANTON_ST’, ‘WASHINGTON_ST’, ‘BLUE_HILL_AVE’, ‘BOYLSTON_ST’, ‘HARRISON_AVE’ | ‘WASHINGTON_ST’, ‘BLUE_HILL_AVE’, ‘BOYLSTON_ST’, ‘HARRISON_AVE’, ‘HYDE_PARK_AVE’ | ‘TREMONT_ST’, ‘HARRISON_AVE’, ‘WASHINGTON_ST’, ‘HYDE_PARK_AVE’, ‘MASSACHUSETTS_AVE’ | ‘FARRINGTON_AVE’, ‘BOYLSTON_ST_-&_CHARLES_ST-_BOSTON_MA_02108_-UNITED_S’, ‘WASHINGTON_ST_-&_HERON_ST_-WEST_ROXBURY_MA_-02132_UN’, ‘WILLIAMS_ST’, ‘INDEPENDENCE_DR’ | ‘GORDON_AVE’, ‘DANA_AVE’, ‘RANDOLPH_ST’, ‘FRUIT_ST’, ‘CHISWICK_RD’ | ‘WASHINGTON_ST’, ‘BLUE_HILL_AVE’, ‘HARRISON_AVE’, ‘BOYLSTON_ST’, ‘HYDE_PARK_AVE’ |

TABLE II
DATASETS

| **Dataset** | **#Nodes** | **#Edges** |
|---|---|---|
| CyberCrime | 28 | 105 |
| London Crime | 64 | 528 |
| Crime against children (CAC) | 72 | 1260 |
| Mafia Meeting | 154 | 286 |
| Global Terrorism | 7062 | 2 475 523 |
| Boston | 10 668 | 94 233 |

edges. Therefore, we have modeled the networks based on the information given in their datasets. The number of nodes and edges are shown in Table II and their description is as follows:

1) *Cybercrime dataset*[2]: This dataset models nodes as crime motives (e.g., financial gain, data theft), with edges representing relationships between motives.
2) *London crime dataset*[3]: London-specific data captures various types of crime and their interrelations. In this dataset, nodes represent specific locations in London, with edges connecting locations where crimes occurred in the same year.
3) *Crime against children (CAC) dataset*[4]: This dataset includes data on crimes against children, categorized by Indian states and union territories. Here, states and union territories are represented as nodes. It is a complete graph where each edge is weighted by the influence factor of crime rates against children in each state, allowing for a state level.
4) *Sicilian Mafia meeting network*[5]: This network captures individuals connected through shared meetings. Each node represents a person who participated in at least one criminal meeting, and an edge connects two nodes if they attended the same meeting, indicating collaboration or association.
5) *Global terrorism dataset*[6]: Represents global terrorist connections and interactions. Here, nodes represent terrorist groups, and edges connect groups that share similar target relationships.
6) *Boston crime dataset*[7]: This dataset includes the crime happened in Boston in the year 2022. The nodes in the network represents the location of the crimes, and edges between them represents the crime occur on the same day.

### B. Result Analysis

As mentioned above, the nodes in the datasets represent states, locations, people, crime motives, etc. The experimental results with respect to analyses are as follows:

1) *Key influencers in criminal network*: As mentioned above, the analysis defined in Section III-C1 aims to find the top key nodes (seed nodes) in criminal networks. Experiments were conducted on different sizes of key nodes. However, we showed the top key influential node

[2]https://www.kaggle.com/datasets/seanangelonathanael/dataset-cybercrime-in-india

[3]https://www.kaggle.com/datasets/jboysen/london-crime

[4]https://www.kaggle.com/datasets/dekomorisanae09/criminal-activities-in-india

[5]https://www.kaggle.com/datasets/mingshanjia/sicilian-mafia-criminal-network

[6]https://www.kaggle.com/datasets/START-UMD/gtd

[7]https://www.kaggle.com/datasets/shivamnegi1993/boston-crime-dataset-2022

Authorized licensed use limited to: PSG COLLEGE OF TECHNOLOGY. Downloaded on July 16,2025 at 06:07:49 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



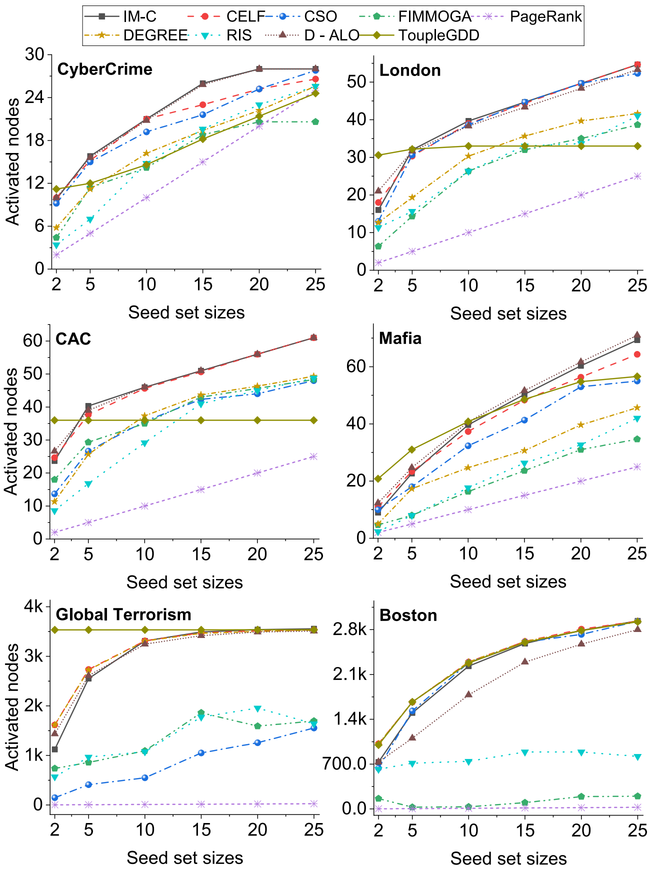


Fig. 2. Activated/influenced nodes at different sizes of key nodes.

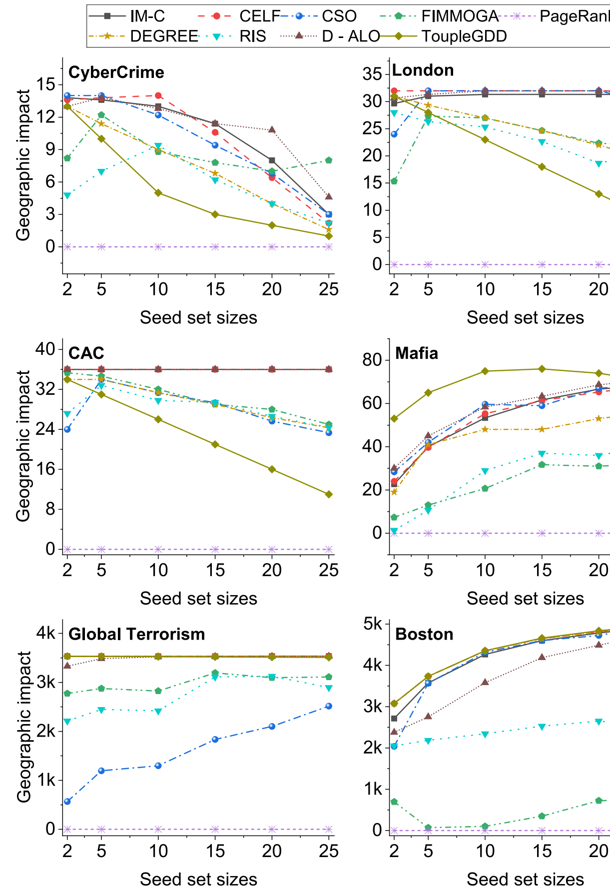


Fig. 3. Geographic impact of the key influential nodes.

of size 5 in this work. Considering size $k = 5$, the top key influential nodes are shown in Table I. Based on the experimental results, from these key influential nodes, we observed that "Fraud" and "Extortion" are the key motives in the CyberCrime dataset. In the London dataset, "Bromley" and "Greenwich" are the key cities where crime has happened. In the CAC datasets, states like "Jharkhand" and "Chhattisgarh" are commonly occurring in states where crimes against children occur at a higher rate. In the Mafia networks, key persons who frequently coordinate meetings are "N12," "N18," and "N51." Likewise, this analysis finds the key influential nodes in the criminal networks as shown in Table I. The influence of these key nodes is discussed in the subsequent section.

2) *Influence spread:* This section analyses the influence of the key nodes (seed set) of criminal networks extracted by the algorithms. The influence spread at different seed sizes of key nodes is shown in Fig. 2 as "Activated nodes." From the experimental results, as shown in Fig. 2, it is observed that the proposed algorithm IM-C has shown the maximum or competitive influence with the top-performing benchmark algorithm. Along with the proposed IM-C algorithm, the benchmark CELF, D-ALO, and ToupleGDD algorithms have given a competitive influence. The influence of most of the algorithms has increased as the seed set sizes increase (except ToupleGDD). The PageRank algorithm remains the worst-performing algorithm. The most commonly used algorithm in criminal social network analysis, degree, has given adequate influence. The proposed algorithm IM-C has outperformed the benchmark algorithm because it selects the node that has a high marginal gain and is far from the existing seed set. This leads to reducing the overlap influence. Therefore, these two components together give the nodes that have higher influence with less overlap. Thus, it has given a higher influence. The proposed IM-C and CELF utilize marginal gain; therefore, they have similar influencing behaviors.

3) *Geographic impact on crime network:* As mentioned in Section III-C3, to show the graphical impact, this analysis shows the number of nodes up to a 2-hop neighbor of the key influential nodes in the criminal network. The graphical representation of this analysis is shown in Fig. 3. The geographic impact of proposed IM-C, CELF, and CSO remains the highest on all the datasets. Fig. 3 shows that the benchmark algorithm ToupleGDD has the highest geographical impact on the Mafia, Global Terrorism, and Boston datasets. We have also observed that as the size of key users increases, it is not necessary that their geographic impact always increases. Interestingly, it is also

Authorized licensed use limited to: PSG COLLEGE OF TECHNOLOGY. Downloaded on July 16,2025 at 06:07:49 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



TABLE III
IMPACT OF DC (DEGREE CENTRALITY), BC (BETWEENNESS CENTRALITY), AND CC (CLOSENESS CENTRALITY) CENTRALITIES ON INFLUENCE IN CRIMINAL SOCIAL NETWORKS

| | CyberCrime | | | London | | | CAC | | | Mafia | | | Global Terrorism | | | Boston | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Seed set sizes | DC | BC | CC | DC | BC | CC | DC | BC | CC | DC | BC | CC | DC | BC | CC | DC | BC | CC |
| 2 | 5.8 | 5 | 2 | 12.67 | 7 | 2 | 11.34 | 11.34 | 2 | 5 | 4 | 2 | 1615.69 | 242.15 | 2 | 1018.39 | 24.65 | 2 |
| 5 | 11.2 | 9.2 | 5 | 19.34 | 14.67 | 5 | 25.67 | 25.67 | 5 | 17.34 | 12.34 | 5 | 2725.35 | 1081.39 | 5 | 1668.58 | 443.15 | 5 |
| 10 | 16.2 | 15.4 | 10 | 30.34 | 26.67 | 10 | 37.34 | 37.34 | 10 | 24.67 | 20.67 | 10 | 3309.79 | 1228.24 | 10 | 2295.39 | 700.76 | 10 |
| 15 | 19.4 | 18.6 | 15 | 35.67 | 32 | 15 | 43.67 | 43.67 | 15 | 30.67 | 27.34 | 15 | 3465.67 | 1673.13 | 15 | 2616.64 | 1158.99 | 15 |
| 20 | 22.2 | 22 | 20 | 39.67 | 38.67 | 20 | 46.34 | 46.34 | 20 | 39.67 | 35.34 | 20 | 3502.14 | 2223.85 | 20 | 2788.59 | 1285.25 | 20 |
| 25 | 25.6 | 26 | 25 | 41.67 | 42 | 25 | 49.34 | 49.34 | 25 | 45.67 | 41.34 | 25 | 3530.20 | 2393.03 | 25 | 2933.34 | 1358.04 | 25 |

observed that the geographic impact of the degree algorithm is not the highest in four datasets (out of six). The PageRank algorithm has given the minimum geographic impact. The proposed algorithm has outperformed because the nodes are selected from different locations due to the diversity component. Thus, it has a higher geographic impact.

4) *Resilience and vulnerability of the network*: To show resilience and vulnerability, this work defined $R$ in Section III-C4. If $R$ is low, the key nodes have high resilience and low vulnerability, and if $R$ is high, then key nodes have low resilience and high vulnerability in the network. Thus, this work aims to minimize $R$. The graphical analysis of $R$ is shown in Fig. 4. From this figure, it is observed that the proposed algorithm IM-C has outperformed the four datasets in the CyberCrime, London, CAC, and Mafia datasets. The proposed IM-C algorithm is competitive with the best-performing benchmark algorithm in two datasets on global terrorism and Boston datasets. On global terrorism and Boston datasets, the ToupleGDD has outperformed. The performance difference between the CELF and the proposed algorithm remains the minimum on all the datasets. The proposed algorithm has outperformed because it has achieved higher influence, and higher influence leads to giving a lesser $R$ value. Conclusively, the proposed algorithm has outperformed the benchmark algorithms in achieving high resilience and low vulnerability.

5) *Key influential nodes based on centralities*: As mentioned in Section III-C5, this work investigates the role of centralities in criminal networks. The influence of these centralities is shown in Table III. From this table, it is observed that in the CyberCrime, London, and CAC datasets, the influence of degree and betweenness centrality shows a similar influence pattern. But in the mafia meeting, global terrorism and Boston datasets, the degree centrality has given a higher influence than the betweenness centrality. Closeness centrality has not performed well because closeness centrality in directed graphs is calculated based on incoming edges, focusing on how easily a node can be accessed by others. So if closeness centrality measures incoming paths, it identifies nodes that are easily reachable but not necessarily good spreaders of influence. Therefore, closeness centrality did not perform well.

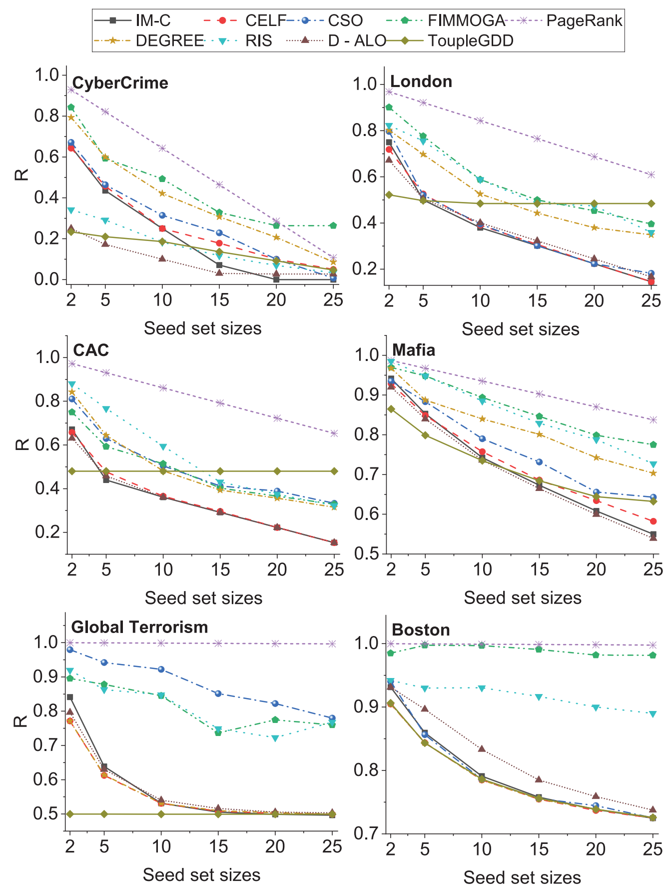


Fig. 4. Resilience and vulnerability in the criminal networks.

## C. Discussion

The proposed work uncovers several insights by applying IM strategies to criminal networks. These insights are useful regarding key influential nodes, their influence, geographic factors, resilience, etc. The experimental results showed the most influential key nodes are identified, and some key influential nodes are common in the IM-based algorithms. The experimental results have shown that the proposed IM-based algorithm IM-C has outperformed the benchmark algorithms in giving a high influence. However, in some instances, the benchmark algorithm CELF, D-ALO, and ToupleGDD are competitive with the proposed algorithm. The geographic impact of the proposed

Authorized licensed use limited to: PSG COLLEGE OF TECHNOLOGY. Downloaded on July 16,2025 at 06:07:49 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



algorithm also remains high. It is also observed that the geographical impact of degree-based key influential nodes is not the highest. The experimental results concluded that the key influence nodes identified by the proposed algorithm give high resilience and low vulnerability in criminal networks. For the key user based on the centralities measure, the experimental results concluded that the Degree algorithm has given the highest influence, but on some datasets, it gives equivalent influence as compared to betweenness centrality. However, based on the experiments, it is concluded that closeness centrality is not suitable for finding key influential users.

## V. Conclusion

Influence maximization (IM) finds the key influential nodes in social networks. The proposed work introduced IM to study and uncover patterns in criminal social networks. This work proposed an algorithm IM-C that finds the top key nodes in the criminal social network. The IM-C algorithm finds the top $k$ nodes based on the marginal gain and path-based diversity. By introducing IM to the criminal social network, this work analyses the identification of key influential nodes, their influence in the criminal network, their geographic impact, resilience and vulnerability analysis, and the influence of centrality-based key influential nodes. The experiments have been conducted on six datasets. The experimental results have shown that the proposed algorithm is the top-performing algorithm with benchmark algorithms. The results showed that the IM-based algorithms have the maximum influence in criminal networks. The IM-based algorithm also achieved the highest geographic impact, high resilience and low vulnerability. Thus, based on the experimental results, this work concluded that the IM uncovers various insights into criminal social networks. Future work may focus on tag-based IM, context-based IM, or location-based IM algorithms for customized influential key nodes in criminal social networks.

## References


[1] J. Golbeck, *Analyzing the Social Web.* San Mateo, CA, USA: Morgan Kaufmann, 2013.

[2] F. Calderoni, "Morselli, Carlo: Inside criminal networks," *Eur. J. Criminal Policy Res.*, vol. 16, pp. 131–145, Mar 2010.

[3] S. Aral and D. Walker, "Creating social contagion through viral product design: A randomized trial of peer influence in networks," *Manage. Sci.*, vol. 57, no. 9, pp. 1623–1639, 2011. [Online]. Available: https://doi.org/10.1287/mnsc.1110.1421

[4] S. S. Singh, D. Srivastva, M. Verma, and J. Singh, "Influence maximization frameworks, performance, challenges and directions on social network: A theoretical study," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 9, pp. 7570–7603, 2022.

[5] R. M. Bond et al., "A 61-million-person experiment in social influence and political mobilization," *Nature*, vol. 489, no. 7415, pp. 295–298, 2012. [Online]. Available: https://doi.org/10.1038/nature11421

[6] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in *Proc. 9th ACM SIGKDD Int. Conf. Knowledge Discov. Data Mining (KDD).* New York, NY, USA: ACM, 2003, pp. 137–146.

[7] P. Domingos and M. Richardson, "Mining the network value of customers," in *Proc. 7th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining (KDD).* New York, NY, USA: ACM, 2001, pp. 57–66.

[8] J. Guo, Y. Li, and W. Wu, "Targeted protection maximization in social networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 3, pp. 1645–1655, 2020.

[9] E. K. Anastasiadis and I. Antoniou, "Directed criminal networks: Temporal analysis and disruption," *Information*, vol. 15, no. 2, 2024.

[10] R. R. Petersen, C. J. Rhodes, and U. K. Wiil, "Node removal in criminal networks," in *Proc. Eur. Intell. Secur. Inform. Conf.*, 2011, pp. 360–365.

[11] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance, "Cost-effective outbreak detection in networks," in *Proc. 13th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining (KDD).* New York, NY, USA: ACM, 2007, pp. 420–429.

[12] S. Kumar Meena, S. Sheshar Singh, and K. Singh, "Cuckoo search optimization-based influence maximization in dynamic social networks," *ACM Trans. Web*, vol. 18, no. 4, Oct. 2024.

[13] S. K. Meena, S. Sheshar Singh, and K. Singh, "Pruning-enabled dynamic influence maximization using antlion optimization," *Knowl. Based Syst.*, vol. 317, 2025, Art. no. 113406. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950705125004538

[14] M. Mani, O. Bozorg-Haddad, and X. Chu, *Ant Lion Optimizer (ALO) Algorithm.* Singapore: Springer Singapore, 2018, pp. 105–116. [Online]. Available: https://doi.org/10.1007/978-981-10-5221-7_11

[15] S. K. Meena, S. S. Singh, and K. Singh, "Dcdimb: Dynamic community-based diversified influence maximization using bridge nodes," *ACM Trans. Web.*, vol. 18, no. 4, Oct. 2024. [Online]. Available: https://doi.org/10.1145/3664618

[16] H. Gong and C. Guo, "Influence maximization considering fairness: A multi-objective optimization approach with prior knowledge," *Expert Syst. Appl.*, vol. 214, 2023, Art. no. 119138. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S095741742202156X

[17] S. Chen, X. Yin, Q. Cao, Q. Li, and H. Long, "Targeted influence maximization based on cloud computing over big data in social networks," *IEEE Access*, vol. 8, pp. 45512–45522, 2020.

[18] J. Guo, Q. Ni, W. Wu, and D.-Z. Du, "Composite community-aware diversified influence maximization with efficient approximation," *IEEE/ACM Trans. Netw.*, vol. 32, no. 2, pp. 1584–1599, 2024.

[19] T. Dinh, A. Nguyen, U. Nguyen, and G. Nguyen, "Quantum social computing approaches for influence maximization," in *Proc. GLOBECOM 2022 - 2022 IEEE Global Commun. Conf.*, 2022, pp. 5832–5837.

[20] T. Chen, S. Yan, J. Guo, and W. Wu, "ToupleGDD: A fine-designed solution of influence maximization by deep reinforcement learning," *IEEE Trans. Comput. Social Syst.*, vol. 11, no. 2, pp. 2210–2221, 2024.

[21] Y. Li, H. Gao, Y. Gao, J. Guo, and W. Wu, "A survey on influence maximization: From an ml-based combinatorial optimization," *ACM Trans. Knowl. Discov. Data*, vol. 17, no. 9, Jul. 2023. [Online]. Available: https://doi.org/10.1145/3604559

[22] M. Richardson and P. Domingos, "Mining knowledge-sharing sites for viral marketing," in *Proc. 8th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining (KDD).* New York, NY, USA: ACM, 2002, pp. 61–70.

[23] X. Wang, W. Lu, M. Ester, C. Wang, and C. Chen, "Social recommendation with strong and weak ties," in *Proc. 25th ACM Int. Conf. Inform. Knowl. Manage. (CIKM).* New York, NY, USA: ACM, 2016, pp. 5–14. [Online]. Available: https://doi.org/10.1145/2983323.2983701

[24] A. Caliò and A. Tagarelli, "Attribute based diversification of seeds for targeted influence maximization," *Inf. Sci.*, vol. 546, pp. 1273–1305, 2021.

[25] W. Chen, Y. Yuan, and L. Zhang, "Scalable influence maximization in social networks under the linear threshold model," in *Proc. IEEE Int. Conf. Data Mining*, 2010, pp. 88–97.

[26] S. Li, B. Wang, S. Qian, Y. Sun, X. Yun, and Y. Zhou, "Influence maximization for emergency information diffusion in social internet of vehicles," *IEEE Trans. Veh. Technol.*, vol. 71, no. 8, pp. 8768–8782, 2022.

[27] X. Liu, E. Patacchini, Y. Zenou, and L.-F. Lee, "Criminal Networks: Who is the Key Player?" Fondazione Eni Enrico Mattei (FEEM), Climate Change and Sustainable Development 128707, May 2012.

[28] D. Firmani, G. F. Italiano, and L. Laura, "The (not so) critical nodes of criminal networks," in *Social Informatics*, L. M. Aiello and D. McFarland, Eds. Cham: Springer Int. Publishing, 2015, pp. 87–96.

[29] K. Basu and A. Sen, "Identifying individuals associated with organized criminal networks: A social network analysis," *Social Netw.*, vol. 64, pp. 42–54, 2021.

[30] K. R. Purba, D. Asirvatham, and R. K. Murugesan, "Influence maximization algorithm: Review on current approaches and limitations," *Eng. Appl. Sci. Res.*, vol. 48, no. 2, pp. 221–229, Mar. 2021.


Authorized licensed use limited to: PSG COLLEGE OF TECHNOLOGY. Downloaded on July 16,2025 at 06:07:49 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



[31] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social Netw.*, vol. 1, no. 3, pp. 215–239, 1978.

[32] F. Kazemzadeh, A. A. Safaei, and M. Mirzarezaee, "Optimal selection of seed nodes by reducing the influence of common nodes in the influence maximization problem," in *Proc. 13th Int. Conf. Inf. Knowl. Technol. (IKT),* 2022, pp. 1–7.

[33] C. Borgs, M. Brautbar, J. Chayes, and B. Lucier, "Maximizing social influence in nearly optimal time," in *Proc. 25th Annu. ACM-SIAM Symp. Discrete Algorithms (SODA)*. USA: Society for Industrial and Applied Mathematics, 2014, pp. 946–957.

[34] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web," Stanford InfoLab, Technical Report 1999-66, Nov. 1999.

[35] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proc. Nat. Acad. Sci.*, vol. 99, no. 12, pp. 7821–7826, 2002.

[36] Q. Jiang, G. Song, C. Gao, Y. Wang, W. Si, and K. Xie, "Simulated annealing based influence maximization in social networks," *Proc. AAAI Conf. Artif. Intell.*, vol. 25, no. 1, pp. 127–132, 2011, [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/7838

[37] S. Bhagat, A. Goyal, and L. V. Lakshmanan, "Maximizing product adoption in social networks," in *Proc. 5th ACM Int. Conf. Web Search Data Mining (WSDM)*. New York, NY, USA: ACM, 2012, pp. 603–612.

[38] M. Gong, J. Yan, B. Shen, L. Ma, and Q. Cai, "Influence maximization in social networks based on discrete particle swarm optimization," *Inf. Sci.*, vol. 367-368, pp. 600–614, 2016.

Authorized licensed use limited to: PSG COLLEGE OF TECHNOLOGY. Downloaded on July 16,2025 at 06:07:49 UTC from IEEE Xplore. Restrictions apply.