Computing and Informatics, Vol. 39, 2020, 156–173, doi: 10.31577/cai_2020_1-2_156

REVERSE INTERVENTION FOR DEALING WITH MALICIOUS INFORMATION IN ONLINE SOCIAL NETWORKS

Deyu YUAN, Haichun SUN*

College of Police Information Engineering and Cyber Security
People's Public Security University of China
Beijing 102623, China
&
Key Laboratory of Safety Precautions and Risk Assessment
Beijing 102623, China
e-mail: {yuandeyu, sunhaichun}@ppsuc.edu.cn

Abstract. Malicious information is often hidden in the massive data flow of online social networks. In "We Media" era, if the system is closed without intervention, malicious information may spread to the entire network quickly, which would cause severe economic and political losses. This paper adopts a reverse intervention strategy from the perspective of topology control, so that the spread of malicious information could be suppressed at a minimum cost. Noting that as the information spreads, social networks often present a community structure and multiple malicious information promoters may appear. Therefore, this paper adopts a divide and conquer strategy and proposes an intervention algorithm based on subgraph partitioning, in which we search for some influential nodes to block or release clarification. The main algorithm consists of two main phases. Firstly, a subgraph partitioning method based on community structure is given to quickly extract the community structure of the information dissemination network. Secondly, a node blocking and clarification publishing algorithm based on the Jordan Center is proposed in the obtained subgraphs. Experiments show that the proposed algorithm can effectively suppress the spread of malicious information with a low time complexity compared with the benchmark algorithms.

Keywords: Malicious information, social network, reverse intervention

* Corresponding author

1 INTRODUCTION

In recent years, social media has become an important platform for online users to participate in the Internet, such as Facebook, Twitter, Sina Weibo, WeChat, QQ, etc. Users on social media have formed OSN (Online Social Networks). The expansion of the social category from physical space to virtual space is a process from quantitative change to qualitative change. On the one hand, the deep integration of social media and politics, economy and culture releases positive energy, and the highly connected OSN provides infrastructure for the realization of "Internet +". On the other hand, malicious information such as rumors and fake news often hide in the massive social data, which brings unprecedented challenges for national security and social stability, making the high-speed diffusion of information in OSN a double-edged sword. The "information security" in online social networks has attracted more and more attention.

The control of malicious information in OSN is mostly studied from two aspects: credibility evaluation and information dissemination dynamics. In the perspective of credibility evaluation, classification or sorting methods are often used, and the text content of social media, supplemented by user information and communication characteristics could be analyzed. Kwon et al. [1] used the timing characteristics, combined with the structure and semantics of the message to identify rumors. Song et al. [2] analyzed the statistical characteristics of text content (such as number of repeated microblogs in the last 20 Weibo contents, number of external links, number of @ symbols, number of topic tags) and characteristics of user relationships (number of followers, reputation of each user) to identify malicious information on Twitter. In the perspective of information dynamics, Fang [3] used life cycle theory to divide the information fermentation process into four stages: gestation, diffusion, transformation and attenuation. Lan et al. [4] established a differential equation model based on the forming process and influencing factors to study the information evolution in the network, and the authors proposed three characteristic time points and four periods for public opinion diffusion.

The above studies provide the basis for reverse control (i.e. manual intervention) of malicious information, even though none of them mentioned the intervention of malicious information. In the analysis of the propagation of malicious information, OSN is regarded as a closed system. The attacker can choose a reasonable publishing strategy to make information spread quickly and achieve his purpose. However, in reality, the system is open. From a theoretical point of view, the multi-layered information dissemination process could be interfered by adding disturbance variables. From a practical point of view, it is possible to issue clarification or block rumor accounts, making malicious information and the clarification disturb each other, which could hinder the rapid spread of malicious information. Therefore, to timely and effectively disturb the evolution of malicious information is a challenging and important issue.

The current literatures of reverse control mostly compromise on effectiveness and efficiency. In this paper, we go further on reverse control in OSN and try

our best to suppress the spread of malicious information at a minimum cost. The main contributions of this paper can be summarized as follows. Firstly, we propose a novel community partitioning algorithm to reduce the complexity in large-scale networks. Secondly, we introduce a mechanism which incorporates both blocking and clarification publishing methods to impede the spread of malicious information. Thirdly, we utilize the Jordan Center to select key nodes for publishing clarifications. Finally, we verify the effectiveness of the model through the experiments.

The remainder of this paper is organized as follows. We review the related works in Section 2. In Section 3, we present the problem formulation. The reverse intervention algorithm based on subgraph partitioning is proposed in Section 4. Experimental results are presented in Section 5. Finally, we summarize this paper in Section 6.

2 RELATED WORK

The research of reverse control in OSN originated from the invulnerability of complex networks, in which different measurement and control indicators have been proposed and analyzed. For example, the authors of [5] studied the invulnerability of ad hoc network, in which "k-connectivity" and power control were used to protect the network against random failure. In [6], the critical removal ratio was used as the measurement for the networks with incomplete information, and the invulnerability of the network was analyzed based on characteristic spectrum. Albert et al. [7] used generating function to analyze the critical removal ratio under the random failure conditions. Cohen et al. [8] extended the problem to the generalized random graph. Callaway et al. [9] studied the percolation problem on graphs with completely general degree distribution and proposed some specific solutions for a variety of cases, including site percolation, bond percolation, and models in which occupation probabilities rely on vertex degree. In [10], highly optimized tolerance (HOT) theory and node preference attachment mechanism were used to build the invulnerable dynamic evolution model for the studied network.

The controllability and information diffusion were also analyzed in opportunistic social networks [11] and location-based social networks (LBSNs). For example, weight distribution between nodes and communities reconstitution were established in [12] to solve the problem of message delivery for social opportunistic networks. In [13], the authors proposed a routing algorithm called sensor communication area node extend (SCANE) to select relevance nodes and to recombine communication areas. In [14], a method for recommending points of interest (POIs) was proposed based on a collaborative tensor factorization (CTF) technique. Luan et al. [15] proposed a maximal-marginal-relevance-based personalized trip recommendation method that considers both relevance and diversity of trips in a trip planning. These literatures are inspiring and instructive to analyze the propagation of rumors.

In order to restrain the propagation of rumors, scientists have proposed many methods. The literatures can be roughly categorized as controlling influential users (links), and clarifying the rumors by spreading the truths under different diffusion models. For blocking strategies, the evaluation of important nodes and links plays an important role in the blocking strategies. The centrality indicators such as degree centrality, clustering coefficient, betweenness centrality, closeness centrality, k-shell decomposition [16], HITS algorithm [17], PageRank algorithm [18], network efficiency [19], Laplace centrality [20] , structural hole [21], minimum spanning tree index [22], mutual information method [23], and node contraction method [24] could be drew on in the proposed strategies. Fan et al. [25] explored the Least Cost Rumor Blocking (LCRB) problem to prevent rumors from spreading. The authors tried to minimize the number of people infected from the originate community to other communities by identifying a minimal bridge end set which diffuse the positive (protector) cascade. However, the authors assumed that the cascade of rumor and protector start at the same time, which was not in line with the real situation that the positive cascade was usually released after the rumor has been noticed. For clarifying the rumors, Wan et al. [26] proposed a novel model of competitive coupling to describe the complex process of information diffusion in online social networks and introduced the constrained intervention strategies. The analysis of coupling diffusion among different information is very inspiring when we introduce the clarifications. Wen et al. [27] numerically evaluated the two types of strategies used for restraining rumors in OSNs, including blocking rumors at important users and clarifying rumors by spreading truths, thus introduced a mathematical model to present the spread of rumors and truths. The authors found that the truth clarification method could eliminate more rumors in the long run while the blocking method based on degree could provide better performance in the early stage of the rumor spread.

Credibility analysis of posts and users was also used to control the rumor spreading. Bao et al. [28] proposed a novel immunization strategy called MST based on trust network. The authors established a weighted trust network based on the trust relationship between users, and determined the most important information diffusion paths to cut down. However, the trust weight of the links was hard to determine and the proposed algorithm was time consuming. Bao et al. also proposed a SPNR model in [29], in which the authors split the infected states into two separate states according to whether the user support or oppose the information. That is, the paper assumed the users in OSN could spontaneously oppose the rumor. However, only parameters' influence was analyzed and effective rumor control strategies need further discussion. Bhattacharya et al. [30] proposed a belief surveillance approach for specific propositions, which is inspired by studies on disease surveillance. The authors demonstrated that although factual statements garner a high degree of belief, some are still being questioned, and some fictional statements also garner a high degree of belief, which was instructive for the control of malicious information.

Inspired by the above literatures, we incorporate both blocking and clarification publishing methods to control the diffusion of malicious information. In this paper,

we break the assumption of closed systems and implement reverse interventions to impede the spread of malicious information.

3 PROBLEM FORMULATION

In this section, we give the definition of the Jordan Center and present information diffusion models and symbols used in this paper.

3.1 Diffusion Model

We use directed graph $G = (\mathbb{V}, \mathbb{E})$ to represent OSN, where $\mathbb{V}$ is the set of nodes, and $\mathbb{E}$ is the set of edges in the network. Two nodes connected by edges are called neighbors (e.g., there is a relationship of following). At some certain moment, an attacker in the network issues a malicious message m, and other nodes in the network will receive message m and forward it to its neighbor nodes.

Next, we describe the propagation model used in this paper. Existing information dissemination models can be roughly divided into two categories: epidemiological infection models such as SI (Susceptible-Infected), SIR (Susceptible-Infected-Recovery) and SIS (Susceptible-Infected-Susceptible) model and influence diffusion models such as IC (Independence Cascade) and LT (Linear Threshold) model. This paper focuses on the influence diffusion model, namely LT and IC model. These two models have received extensive attention since they were first proposed in the pioneering work of Kempe et al. [31].

IC model: An infected node v has only one chance to infect its susceptible neighbors, and each neighbor node $w \in N(v)$ ($N(v)$ represents the neighbor set of node v) can be infected with an independent probability $p_{v,w}$.

LT model: Each node in the network independently selects a threshold $\theta_v \in [0, 1]$ at the initial stage. Whether a susceptible node w adopts the information depends on the sum of all its neighbors' weights $p_{v,w}$, where $v \in N(w)$. When the sum of the weights for susceptible node w satisfies $\sum_{v \in N(w)} b_{v,w} \geq \theta_v$, the node w will be infected.

3.2 Problem Formulation

The reverse intervention of malicious information is closely related to the influence diffusion model of OSN. Malicious information spread together with other information in the network, and information holding the opposite opinion will compete with each other. In the real world, users who receive clarification usually should no longer accept the malicious information (rumors). In order to prevent people from being misled by malicious information, a natural way is to introduce clarifications to uninfected users as soon as possible, at least earlier than the arrival of malicious information. Once malicious information is detected, the network administrator (e.g.

police department) can generate a competitive positive cascading (clarification) to minimize the number of infected (propagating) users. In this paper, we assume that the clarification has higher priority than malicious information to activate nodes. Therefore, according to the IC and LT model discussed above, the problem that needs to be solved in this paper is described as the following optimization problem.

$$\min |\mathbb{S}| \tag{1}$$

s.t.

$$\mathbb{S} \subset \mathbb{V}, \tag{2}$$

$$\frac{|\mathbb{I}(G)|}{|\mathbb{V}|} \leq \beta. \tag{3}$$

That is, according to the propagation situation of malicious information m, we try to select a minimum set of nodes to block or publish clarification to control the spread of malicious information, so that the infection rate of the whole network after a time window T is less than β ($\mathbb{I}(G)$ in Equation 3 represents the set of infected nodes in the network).

3.3 Jordan Center

In this subsection, we give the definition of the Jordan Center according to the previous work [32, 33].

Definition 1 (Jordan Center)**.** Let $d(s, u)$ represent the distance between nodes s and u in graph G (i.e. length of the shortest path). $\mathbb{A}$ is a collection of randomly selected nodes in G, and $\bar{d}(s, \mathbb{A})$ is defined as the eccentricity of node s, i.e., the maximum distance between s and any selected node of $\mathbb{A}$, yielding:

$$\bar{d}(s, \mathbb{A}) = \max_{u \in \mathbb{A}} d(s, u). \tag{4}$$

Jordan Center of $\mathbb{A}$ is defined as the node with the smallest eccentricity in G.

4 REVERSE INTERVENTION ALGORITHM BASED ON SUBGRAPH PARTITIONING

In online social networks, algorithms based on community partitioning have been proven to be effective [34, 35]. In the actual network, we can usually observe the fragments of the propagation, take SIR model as an example, some nodes will change from infected state to recovery state. In this paper, we ignore the problem of incomplete observation. Considering the huge advantages of community partitioning, we propose a community-based heuristic method according to the network topology to solve the problem of reverse intervention for malicious information. Specifically, our approach consists of two main phases:

1. subgraph partitioning based on community structure to quickly reveal the community structure of the network;
2. node selection based on the Jordan Center to effectively control the spread of malicious information by means of high-influence nodes.

4.1 Subgraph Partitioning Algorithm Based on Community Structure

In networks with distinct community structures, information is more likely to spread within the community and then spread to other areas of the network. As shown in Figure 1, the network often presents a community structure. As the malicious information spreads, users will hold different opinions on the current event, thus malicious information and external disturbances will form a competition process. The red arrow in Figure 1 represents the opponent flow of malicious information, the blue arrow represents the supporter flow, and the two will form a hedge. By observing the current information dissemination, this paper uses the community structure of the network to distinguish the spread of malicious information, and then suppress the spread of malicious information by publishing clarification in each community.

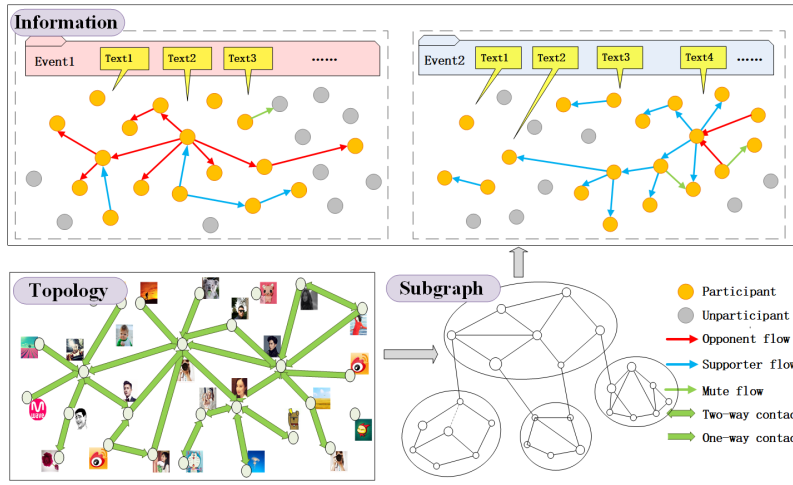


Figure 1. Reverse intervention of malicious information

Lots of measures of the strength of division of a network into communities have been proposed by experts in this field, such as conductance, normalized cut, cut ratio, triangle participation ratio (TPR), etc. [36, 37]. In this paper, we adopt the classic index of modularity proposed by Mark Newman [38] to measure the quality of the community partitioning algorithm, which compares the connection density between the original network and the reference network in the same community. The reference network is defined as a random network having the same degree sequence

as the original network. Suppose $\mathbb{A}$ is the adjacency matrix of a network, where $k_v(k_w)$ is the degree of node $v(w)$ and the total number of edges in the network is N. Then, the modularity is defined as follows:

$$Q = \frac{1}{2N} \sum_{v,w} \left[\mathbb{A}_{vw} - \frac{k_v k_w}{2N} \right] \delta(\mathbb{C}_v, \mathbb{C}_w) \tag{5}$$

where $\mathbb{C}_v$ is the community to which node v belongs. If node v and node w belong to the same community, i.e., $\mathbb{C}_v = \mathbb{C}_w$, then $\delta(\mathbb{C}_v, \mathbb{C}_w) = 1$; otherwise, $\delta(\mathbb{C}_v, \mathbb{C}_w) = 0$.

The higher the modularity, the better the community partitioning algorithm. As an important indicator to measure the quality of community division, modularity has been widely used [39].

In this paper, in order to evaluate the community characteristics of subgraphs, we use the definition of subgraph fitness function [40]:

$$Q(\mathbb{C}) = \frac{\sum in}{2N} - \left(\frac{\sum tot}{2N} \right)^2 \tag{6}$$

where $\sum in$ denotes the number of inner edges of a subgraph $\mathbb{C}$, and $\sum tot$ denotes the total number of edges connected to the nodes inside subgraph, including the edges inside the subgraph and the edges outside the subgraph. The subgraph fitness function measures the degree of "cohesion" of the edges in the subgraph. Obviously, if the community structure of $\mathbb{C}$ is more obvious, the value of Q is larger, and vice versa.

For a particular node, when its edges are mostly inside a subgraph, it is more likely to belong to the subgraph. When most of the edges point to the external nodes of the subgraph, it is unlikely to belong to the subgraph. Therefore, we define the evaluation function f of node n as follow:

$$f(n) = k_n^{\mathbb{C}} / k_n^G \tag{7}$$

where $k_n^{\mathbb{C}}$ is the degree of node n inside subgraph $\mathbb{C}$. k_n^G is the degree of node n in the entire network G.

Accordingly, we propose the subgraph partitioning algorithm based on community structure in this subsection. The basic idea of the algorithm is to randomly select nodes in the network, and then gradually expand the subgraph until the local subgraphs satisfying the given conditions are constructed. That is, the existing structure of the network is divided according to the local subgraph, and the specific process is shown in Algorithm 1.

The subgraph partitioning algorithm starts from a randomly selected set of nodes $\{V_1, V_2, \ldots, V_k\}$ and extends the subgraph along the edges. In order to ensure the community structure of the obtained subgraph, the nodes are first screened in the process of expansion. The algorithm selects the node with the highest evaluation function (most likely belongs to the subgraph) (Step 6), and judges whether adding the node to the current subgraph can increase the subgraph fitness function

Input: Online social network $G = (\mathbb{V}, \mathbb{E})$. The number of nodes initially selected (initial number of subgraphs) k, the maximum number of nodes m in each subgraph.
Output: Subgraphs $\mathbb{C}_i = \{\mathbb{C}_i, i = 1, 2, \dots\} \subset \mathbb{V}$.
Initialize: Randomly select k nodes V_i from the node set, let $\mathbb{C}_i = \{V_i\}, i = 1, 2, \dots, k$
1. **repeat**
2. **for each** $\mathbb{C}_i$ **do**
3. **if** $size(\mathbb{C}_i) < M$ **then**
4. $N_{\mathbb{C}_i} = neighbor(\mathbb{C}_i)$,$increase.\mathbb{C}_i = false$
5. **repeat**
6. $m = \arg\max_{m \in N_{\mathbb{C}_i}} f(m)$
7. **if** $Q(\mathbb{C}_i \cup \{m\}) > Q(\mathbb{C}_i)$ **then**
8. $\mathbb{C}_i = \mathbb{C}_i \cup m, increase.\mathbb{C}_i = true$
9. **end if**
10. $N_{\mathbb{C}_i} = N_{\mathbb{C}_i} - \{m\}$
11. **until** $size(N_{\mathbb{C}_i}) = 0$
12. **end if**
13. **end for**
14. **if** $\mathbb{C}_i \cap \mathbb{C}_j \neq \phi$ **and** $Q(\mathbb{C}_i \cup \mathbb{C}_j) > \max(Q(\mathbb{C}_i), Q(\mathbb{C}_j))$
15. **then** $\mathbb{C}_i = \mathbb{C}_i \cup \mathbb{C}_j, \mathbb{C}_j = \phi$
16. **end if**
17. **until** $size(\mathbb{C}_i) > M$ or $increase.\mathbb{C}_i = false$

Return: Subgraphs $\mathbb{C}_i, i = 1, 2, \dots$

Algorithm 1: Subgraph Partitioning Algorithm

(Steps 7–9). If yes, the node is added to the subgraph, otherwise the node is abandoned. Repeat the above steps until the subgraph reaches the specified scale m or the subgraph fitness stops growing (Step 17).

Due to the randomness of the initial node selection, subgraph initialized from different nodes may overlap. For overlapped subgraphs, the algorithm chooses to merge them according to whether the combined fitness function Q increases (Steps 14–16). Therefore, when selecting k, the subgraph merge situation that may occur should be considered. In order to suppress all possible sources of malicious information, this paper chooses k to be larger than the estimated number of sources in the network.

It is not necessary to divide the network into complete community structures to achieve perfect reverse intervention for malicious information. Therefore, in order to reduce the complexity of the algorithm, by setting the value of a reasonable subgraph size m, the algorithm stops when the subgraph has been extended to the expected size.

4.2 Reverse Intervention Algorithm Based on Jordan Center

Once the first phase is completed, we get a subgraph structure $\mathbb{C}_i = \{\mathbb{C}_i, i = 1, 2, \dots\} \subset \mathbb{V}$, where $\mathbb{C}_i, i = 1, 2, \dots$ is a disjoint subset, now we need to select nodes from these subgraphs to block or post clarification. For simplicity, we assume that the subgraphs $\mathbb{C}_i = \{\mathbb{C}_i, i = 1, 2, \dots\}$ are sorted in a non-incremental order with number of nodes (i.e., $|\mathbb{C}_1| \geq |\mathbb{C}_2| \geq |\mathbb{C}_3| \dots$). Since users exchange information more frequently with users in the same community, and nodes from different subgraphs typically have a small chance to spread malicious information (or clarification) to nodes in other subgraphs. Therefore, our problem is equivalent to finding nodes in each subgraph to control the infection rate of malicious information, so that the infection rate of the whole network can be lower than β.

This paper uses the Jordan Center to find the key node in each subgraph to control the propagation of malicious information. The specific process is summarized in Algorithm 2. The algorithm selects the most influential node in each subgraph according to the definition of Jordan center (Step 4), and determines if it is an infected node. If yes, we delete the node (block the account), otherwise we select it as the clarification publishing node (Steps 5–9). Repeat the above steps until the infection rate of the whole network is lower than β.

Input: Online social network $G = (\mathbb{V}, \mathbb{E})$, number of subgraph p, the infection rate of malicious information β
Output: Set $\mathbb{S} \subset \mathbb{V}$ makes $\frac{|\mathbb{I}(G)|}{|\mathbb{V}|} \leq \beta$
1. **Let** $\mathbb{S} = \phi$
2. **for** i **from** 1 to p **do** $\mathbb{S}_i = \phi$
3. **while** $\frac{|\mathbb{I}(\mathbb{C}_i)|}{|\mathbb{V}|} \leq \beta$ **do**
4. $v = \min_{s \in \mathbb{C}_i} \overline{d}(s, \mathbb{C}_i)$
5. **if** $v \in I(G)$ **then**
6. **block** and **delete** v;
7. **else**
8. $\mathbb{S}_i = \mathbb{S}_i \cup \{v\}$
9. **break**
10. **end if**
11. **end while**
12. **if** $\frac{|\mathbb{I}(G)|}{|\mathbb{V}|} \leq \beta$ **then**
13. **break**
14. **end if**
15. **end for**
Return $\mathbb{S}$

Algorithm 2: Reverse Intervention Algorithm

5 EXPERIMENT RESULTS

In this section, we used real large-scale networks to experimentally evaluate the performance of our proposed method in this paper. The datasets we used were downloaded from Stanford dataset collection (`http://snap.stanford.edu/data`). The first dataset is ego-Facebook, which contains 88 234 edges and 4 039 nodes, and the average clustering coefficient is 0.6055. The second dataset is cit-HepPh, which is a paper citation network, containing 34 546 nodes and 421 578 edges, and the average clustering coefficient is 0.2848. The experimental environment in which the algorithm ran is: processor Intel® Core™ i7-7500M @ 2.70 GHz, memory 8 GB, operating system Windows 10, programming language is Python.

We chose the following four benchmark methods to compare with our proposed algorithm:

1. Random: Randomly selected nodes in the network to block (or publishing clarification) until the infection rate met the requirements.
2. High-degree: The degree based heuristic algorithm, which selected nodes with the highest degree in the network to block (or publishing clarification) until the infection rate met the requirements.
3. Topcgo: A method proposed by Eftekhar et al. [41], which selected nodes with the greatest margin of information spread until the stopping criterion was met.
4. Greedy: The basic greedy algorithm proposed by Kempe et al. [31], which calculated the information dissemination range of each node under the IC model.

In all experiments, Monte Carlo simulation was implemented to estimate the effectiveness of the algorithms. That is, the results were averaged over 1 000 runs for consistency. We chose $p_{v,w}$ in the IC model as 0.25 for any node. And parameter β changed from 0.1 to 0.5. For each β, our proposed algorithm and the benchmark algorithms were independently implemented to get the number of required nodes to achieve the inhibitory effect.

We first consider the IC model using the two datasets. As depicted in Figures 2 and 3, the number of required nodes in our proposed method was highly competitive in comparison with those of others, especially in case that large number of nodes need to be immunized with the malicious information. In particular, when β was small ($\beta \in [0 \ldots 0.09]$), our proposed method did not performed as good as other methods. However, it became much better than other methods except Greedy algorithm as β gets larger. This is because the benchmark methods chose the candidate nodes within the whole network and our proposed method chose the candidate nodes based on community structure. When a small number of nodes were required, the Random, High-degree and Topcgo algorithm could easily select the influential nodes while our proposed algorithm must select nodes in each subgraph. In fact, influential nodes were often distributed in different subgraphs. As the number of required nodes increased, our proposed algorithm could effectively pick up the key nodes in each subgraph, which had influence to other nodes within the subgraph.

However, the benchmark methods had to choose these influential nodes in the whole network.

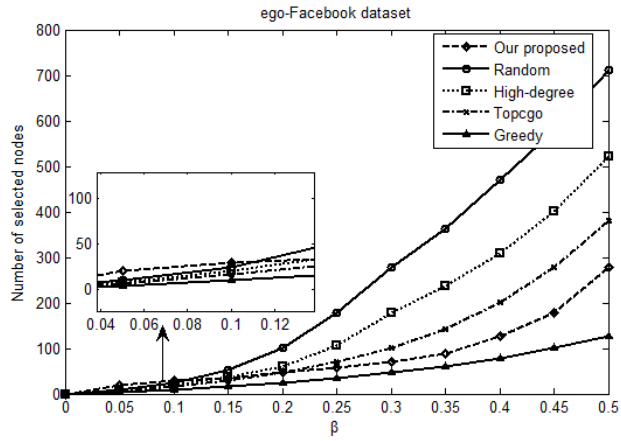


Figure 2. Nodes selected in different algorithms for ego-Facebook dataset

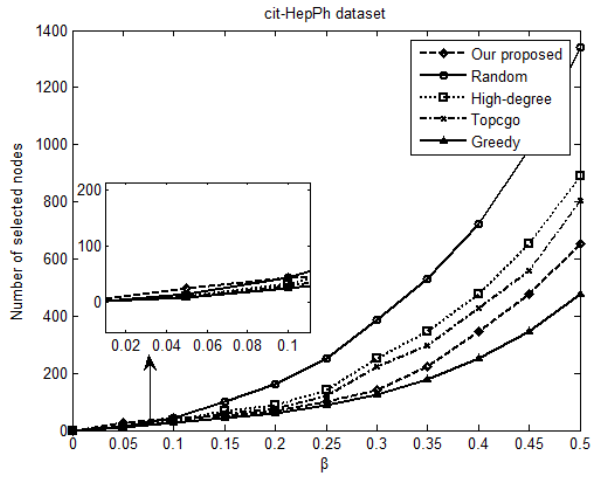


Figure 3. Nodes selected in different algorithms for cit-HepPh dataset

We next illustrate the difference when our proposed method is used under IC model and LT model. As shown in Figure 4, under the LT model, the proposed method could select a slightly fewer nodes to block or release clarification to achieve the desired effect than the IC model. This is because in Steps 5–9 of Algorithm 2, once the Jordan Center in the community was an infected node, we blocked it

and selected the node with the second largest influence (i.e., the second smallest eccentricity) to release clarification if it was not infected, and so on. In the IC model, the infected node had only one chance to affect its neighbor nodes. Whether to block this node or not had no effect to depress the propagation of the malicious information. But the infected nodes still had an impact under the LT model. It is worth noting that the time when to intervene was very important. This is beyond the scope of this paper and will be discussed in our future work.

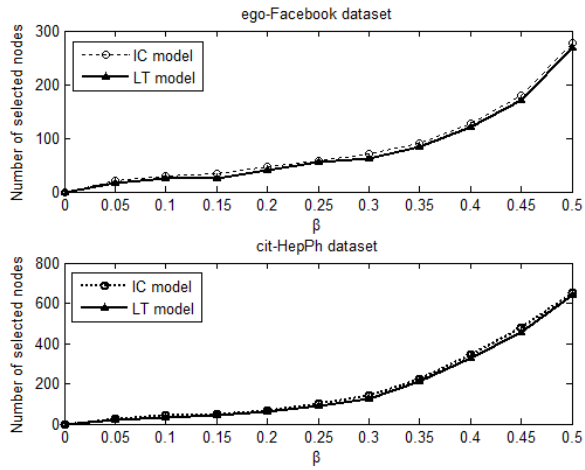


Figure 4. Comparison of the proposed algorithm under different propagation models

We finally evaluated the running time of our proposed algorithm and benchmark methods in Figure 5. Since the time consumption of the Random algorithm was very small, it is not depicted in this figure. As shown in the figure, although Greedy algorithm had the best performance (fewest nodes required to suppress the spread of malicious information), its time complexity was too high, especially on cit-HepPh dataset where it took more than 7500 seconds to meet the condition. Compared with other benchmark algorithms, our proposed algorithm had not only the advantage in intervention performance, but also had the advantage in time complexity. This is because we first divided the entire network into community structures, which could reduce much processing time during the influential node selection period. Therefore, our proposed algorithm could effectively impress the propagation of malicious information in a timely manner.

6 CONCLUSIONS

In this paper, we propose a reverse intervention algorithm based on subgraph partitioning, which impede the spread of malicious information from the perspective

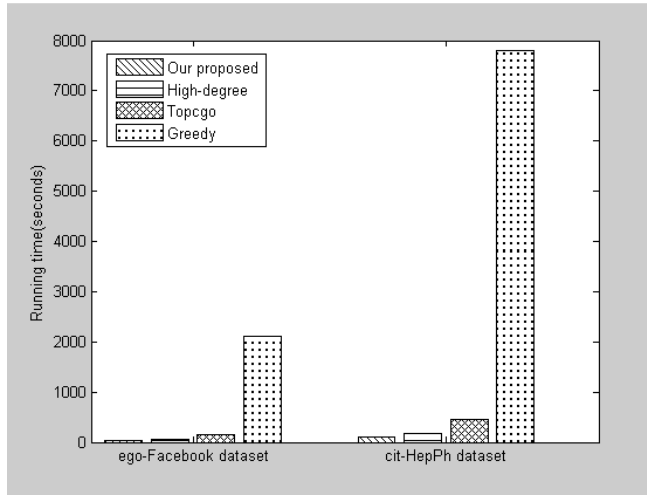


Figure 5. Time consumption of different algorithms

of network topology. Firstly, a subgraph partitioning method based on community structure is given. Secondly, a node blocking and clarification publishing algorithm based on the Jordan Center is proposed in the obtained subgraphs. Experiments on real-world networks including ego-Facebook and cit-HepPh show that the proposed algorithm can effectively suppress the spread of malicious information under a low time complexity.

Acknowlegement

The authors contributed equally to this study and share the first authorship. This work was supported by the National Key R & D Program of China (Grant No. 2017YFC0803700), the Beijing Natural Science Foundation Program (Grant No. 4184099), the National Natural Science Foundation of China (Grant No. 61771072), and the National Social Science Fund of China (Grant No. 17CXW014).

REFERENCES

[1] Kwon, S.—Cha, M.—Jung, K.—Chen, W.—Wang, Y.: Prominent Features of Rumor Propagation in Online Social Media. 2013 IEEE 13th International Conference on Data Mining (ICDM), IEEE, 2013, pp. 1103–1108, doi: 10.1109/ICDM.2013.61.

[2] Song, J.—Lee, S.—Kim, J.: Spam Filtering in Twitter Using Sender-Receiver Relationship. In: Sommer, R., Balzarotti, D., Maier, G. (Eds.): Recent Advances in Intrusion Detection (RAID 2011). Springer, Berlin, Heidelberg, Lecture Notes in Computer Science, Vol. 6961, 2011, pp. 301–317, doi: 10.1007/978-3-642-23644-0_16.

[3] Fang, F.: Study on the Evolution of Public Opinion on Network of Unexpected Event. Ph.D. Dissertation, Huazhong University of Science and Technology, China, 2011.

[4] Lan, Y.—Deng, X.—Ma, M.: Construction of Public Opinion Security Evaluation Index System for Group Events. Information Exploration, Vol. 10, 2011, pp. 37–39.

[5] Hu, X.—Zhang, X.—Wu, J.—Deng, H.: Research for Invulnerability of Ad Hoc Network Topologies. Computer Technology and Development, Vol. 20, 2010, No. 1, pp. 185–188 (in Chinese).

[6] Tan, Y.—Wu, J.—Deng, H.: Progress in Invulnerability of Complex Networks. Journal of University of Shanghai for Science and Technology, Vol. 33, 2012, No. 6, pp. 653–668.

[7] Albert, R.—Jeong, H.—Barabási, A.-L.: Error and Attack Tolerance of Complex Networks. Nature, Vol. 406, 2000, No. 6794, pp. 378–382, doi: 10.1038/35019019.

[8] Cohen, R.—Erez, K.—ben-Avraham, D.—Havlin, S.: Resilience of the Internet to Random Breakdowns. Physical Review Letters, Vol. 85, 2000, No. 21, pp. 4626–4628, doi: 10.1103/PhysRevLett.85.4626.

[9] Callaway, D. S.—Newman, M. E. J.—Strogatz, S. H.—Watts, D. J.: Network Robustness and Fragility: Percolation on Random Graphs. Physical Review Letters, Vol. 85, 2000, No. 25, pp. 5468–5471, doi: 10.1103/PhysRevLett.85.5468.

[10] Liu, Y.: Network Invulnerable Dynamic Evolution Model Based on HOT Theory. Computer Engineering, Vol. 39, 2013, No. 1, pp. 97–101, doi: 10.3969/j.issn.1000-3428.2013.01.021 (in Chinese).

[11] Wu, J.—Chen, Z.—Zhao, M.: Information Cache Management and Data Transmission Algorithm in Opportunistic Social Networks. Wireless Networks, Vol. 25, 2019, No. 6, pp. 2977–2988, doi: 10.1007/s11276-018-1691-6.

[12] Wu, J.—Chen, Z.—Zhao, M.: Weight Distribution and Community Reconstitution Based on Communities Communications in Social Opportunistic Networks. Peer-to-Peer Networking and Applications, Vol. 12, 2019, No. 1, pp. 158–166, doi: 10.1007/s12083-018-0649-x.

[13] Wu, J.—Chen, Z.: Sensor Communication Area and Node Extend Routing Algorithm in Opportunistic Networks. Peer-to-Peer Networking and Applications, Vol. 11, 2018, No. 1, pp. 90–100, doi: 10.1007/s12083-016-0526-4.

[14] Luan, W.—Liu, G.—Jiang, C.—Qi, L.: Partition-Based Collaborative Tensor Factorization for POI Recommendation. IEEE/CAA Journal of Automatica Sinica, Vol. 4, 2017, No. 3, pp. 437–446, doi: 10.1109/JAS.2017.7510538.

[15] Luan, W.—Liu, G.—Jiang, C.—Zhou, M.: MPTR: A Maximal-Marginal-Relevance-Based Personalized Trip Recommendation Method. IEEE Transactions on Intelligent Transportation Systems, Vol. 19, 2018, No. 11, pp. 3461–3474, doi: 10.1109/TITS.2017.2781138.

[16] Carmi, S.—Havlin, S.—Kirkpatrick, S.—Shavitt, Y.—Shir, E.: A Model of Internet Topology Using K-Shell Decomposition. Proceedings of the National Academy of Sciences of the United States of America (PNAS), Vol. 104, 2007, No. 27, pp. 11150–11154, doi: 10.1073/pnas.0701175104.

[17] Kleinberg, J. M.: Authoritative Sources in a Hyperlinked Environment. Journal of the ACM (JACM), Vol. 46, 1999, No. 5, pp. 604–632, doi: 10.1145/324133.324140.

[18] Page, L.—Brin, S.—Motwani, R.—Winograd, T.: The Pagerank Citation Ranking: Bringing Order to the Web. Technical Report, Stanford InfoLab, 1999. Available at: `http://ilpubs.stanford.edu:8090/422/`.

[19] Hu, J.—Wang, B.—Lee, D.: Evaluating Node Importance with Multi-Criteria. Proceedings of the 2010 IEEE/ACM International Conference on Green Computing and Communications and International Conference on Cyber, Physical and Social Computing, IEEE Computer Society, 2010, pp. 792–797, doi: 10.1109/GreenCom-CPSCom.2010.26.

[20] Qi, X.—Duval, R. D.—Christensen, K.—Fuller, E.—Spahiu, A.—Wu, Q.—Wu, Y.—Tang, W.—Zhang, C.: Terrorist Networks, Network Energy and Node Removal: A New Measure of Centrality Based on Laplacian Energy. Social Networking, Vol. 2, 2013, No. 1, pp. 19–31, doi: 10.4236/sn.2013.21003.

[21] Burt, R. S.: Structural Holes: The Social Structure of Competition. Harvard University Press, 2010, pp. 150–188.

[22] Chen, Y.—Hu, A.—Hu, X.: Evaluation Method for Node Importance in Communication Networks. Journal of China Institute of Communications, Vol. 25, 2004, No. 8, pp. 129–134 (in Chinese).

[23] Zhang, Y.—Liu, Y.—Xu, K. et al.: Evaluation Method for Node Importance Based on Mutual Information in Complex Networks. Computer Science, Vol. 38, 2011, No. 6, pp. 88–89 (in Chinese).

[24] Tan, Y.—Wu, J.—Deng, H.: Evaluation Method for Node Importance Based on Node Contraction in Complex Networks. System Engineering – Theory and Practice, No. 11, 2006, pp. 79–84 (in Chinese).

[25] Fan, L.—Lu, Z.—Wu, W.—Thuraisingham, B.—Ma, H.—Bi, Y.: Least Cost Rumor Blocking in Social Networks. 2013 IEEE 33rd International Conference on Distributed Computing Systems (ICDCS), IEEE, 2013, pp. 540–549, doi: 10.1109/ICDCS.2013.34.

[26] Wan, P.—Wang, X.—Wang, X.—Wang, L.—Lin, Y.—Zhao, W.: Intervening Coupling Diffusion of Competitive Information in Online Social Networks. IEEE Transactions on Knowledge and Data Engineering, 2019, doi: 10.1109/TKDE.2019.2954901.

[27] Wen, S.—Jiang, J.—Xiang, Y.—Yu, S.—Zhou, W.—Jia, W.: To Shut Them Up or to Clarify: Restraining the Spread of Rumors in Online Social Networks. IEEE Transactions on Parallel and Distributed Systems, Vol. 25, 2014, No. 12, pp. 3306–3316, doi: 10.1109/TPDS.2013.2297115.

[28] Bao, Y.—Niu, Y.—Yi, C.—Xue, Y.: Effective Immunization Strategy for Rumor Propagation Based on Maximum Spanning Tree. 2014 International Conference on Computing, Networking and Communications (ICNC), IEEE, 2014, pp. 11–15, doi: 10.1109/ICCNC.2014.6785296.

[29] Bao, Y.—Yi, C.—Xue, Y.—Dong, Y.: A New Rumor Propagation Model and Control Strategy on Social Networks. Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013), 2013, pp. 1472–1473, doi: 10.1109/ASONAM.2013.6785909.

[30] Bhattacharya, S.—Tran, H.—Srinivasan, P.—Suls, J.: Belief Surveillance with Twitter. Proceedings of the 4th Annual ACM Web Science Conference (WebSci '12), ACM, 2012, pp. 43–46, doi: 10.1145/2380718.2380724.

[31] Kempe, D.—Kleinberg, J.—Tardos, É.: Maximizing the Spread of Influence Through a Social Network. Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '03), 2003, pp. 137–146, doi: 10.1145/956750.956769.

[32] Hage, P.—Harary, F.: Eccentricity and Centrality in Networks. Social Networks, Vol. 17, 1995, No. 1, pp. 57–63, doi: 10.1016/0378-8733(94)00248-9.

[33] Dekker, A. H.: Centrality in Social Networks: Theoretical and Simulation Approaches. Proceedings of the Simulation Technology and Training Conference (SimTecT) 2008, Melbourne, Australia, 2008, 6 pp.

[34] Nguyen, N. P.—Dinh, T. N.—Xuan, Y.—Thai, M. T.: Adaptive Algorithms for Detecting Community Structure in Dynamic Social Networks. 2011 Proceedings IEEE INFOCOM, Shanghai, China, IEEE, 2011, pp. 2282–2290, doi: 10.1109/INFCOM.2011.5935045.

[35] Nguyen, N. P.—Dinh, T. N.—Tokala, S.—Thai, M. T.: Overlapping Communities in Dynamic Networks: Their Detection and Mobile Applications. Proceedings of the 17th Annual International Conference on Mobile Computing and Networking (MobiCom 2011), Las Vegas, USA, 2011, pp. 85–96, doi: 10.1145/2030613.2030624.

[36] Yang, J.—Leskovec, J.: Defining and Evaluating Network Communities Based on Ground-Truth. Knowledge and Information Systems, Vol. 42, 2015, No. 1, pp. 181–213, doi: 10.1007/s10115-013-0693-z.

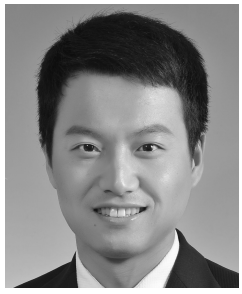
[37] Fortunato, S.: Community Detection in Graphs. Physics Reports, Vol. 486, 2010, No. 3–5, pp. 75–174, doi: 10.1016/j.physrep.2009.11.002.

[38] Newman, M. E. J.: Fast Algorithm for Detecting Community Structure in Networks. Physical Review E, Vol. 69, 2004, No. 6, Art. No. 066133, 5 pp., doi: 10.1103/PhysRevE.69.066133.

[39] Blondel, V. D.—Guillaume, J.-L.—Lambiotte, R.—Lefebvre, E.: Fast Unfolding of Communities in Large Networks. Journal of Statistical Mechanics: Theory and Experiment, Vol. 2008, 2008, Art. No. P10008, 12 pp., doi: 10.1088/1742-5468/2008/10/P10008.

[40] Zhang, X.—Zhang, Y.—Lv, T.—Fu, S.—Zhang, B.: A Multi-Diffusion Source-Localization Method for Online Social Networks Based on Sub-Graph Extraction. Scientia Sinica Informationis, Vol. 46, 2016, No. 4, pp. 496–510, doi: 10.1360/N112015-00190.

[41] Eftekhar, M.—Ganjali, Y.—Koudas, N.: Information Cascade at Group Scale. Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '13), ACM, 2013, pp. 401–409, doi: 10.1145/2487575.2487683.



Deyu Yuan received his Ph.D. degree from the School of Electronic Engineering, Beijing University of Posts and Telecommunications (BUPT), China in 2015. He is now working as Lecturer at the College of Police Information Engineering and Cyber Security, People's Public Security University of China (PPSUC). His research interests include cyber security and complex networks. He has published over 20 papers in journals and conferences such as China Communications, HCC 2014.

Haichun Sun received her Ph.D. degree in computer software and theory from the Tongji University, Shanghai, China, in 2015. She is currently Assistant Professor at the Police Information Engineering and Cyber Security, People's Public Security University of China, Beijing, China. She is a member of Professional Committee of Internet Information Service of the Chinese Association of Automation. Her current research interests include information service, Petri nets, and service-oriented computing. She has published over 10 papers in journals and conferences such as IEEE Transactions on Systems, Man, and Cybernetics, WISE 2014.