Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Locating multiple information sources in social networks based on the naming game

Xue Yang\textsuperscript{a,b}, Zhiliang Zhu\textsuperscript{a,*}, Hai Yu\textsuperscript{a}, Yuli Zhao\textsuperscript{a}, Ying Wang\textsuperscript{a}

\textsuperscript{a} Software College, Northeastern University, Shenyang 110819, China
\textsuperscript{b} College of Mathematics and Information Science, Anshan Normal University, Anshan 114005, China

\textbf{A R T I C L E  I N F O}

Article history:
Received 1 August 2020
Received in revised form 14 September 2020
Accepted 20 September 2020
Available online 25 September 2020
Communicated by M. Perc

Keywords:
Information propagation
Multiple sources localization
Social networks
Naming game theory

\textbf{A B S T R A C T}

Identifying the source of information in a network plays a key role in controlling the impact of information. Herein, we study the problem of multiple source localization in the context of information propagation in social networks. We use the theory of the naming game to conduct observations. Moreover, we divide the observations into different sets based on the information provided by them and then estimate the source of each set. Finally, we combine the source of each observation set to obtain all the estimated information sources. The proposed method can locate sources without knowing the number of information sources. Simulations on four real data sets are provided to verify the performance of our method.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

In January 2020, we witnessed the outbreak of the Coronavirus Disease 2019 (COVID-19) [1]. The outbreak and propagation of the virus threatened health worldwide and changed people’s lives. However, there are other kinds of information spreading, such as viruses on social networking sites, self-Media and WeChat Moments. There are “goodwill reminders”, “good wishes” and sensational “revelations” in this information. Some of the false information has bad social influences and even worse consequences. In essence, this situation can be modeled as information spreading through social networks. Thus, effectively locating the sources of information propagation in a network is significant for forecasting the range of information propagation, controlling the propagation process, etc.

Generally, information propagation in a social network originates from a few information sources and then spreads quickly through the underlying social network [2–4]. According to the number of information sources, existing studies can be divided into single source location and multiple source location. Zhu et al. [5] provide a greedy optimization algorithm for selecting the minimum number of observers in cyber-physical systems and combine the propagation delay and back-diffusion methods to locate the source node of information diffusion. Louni and Subbalakshmi [6] propose a two-stage source localization algorithm, and the experimental results show that the algorithm decreases the percentage of sensors significantly. Spinelli et al. [7] propose the first online approach for source localization. This approach allows for source localization with an extremely small number of sensors.

The works listed above accurately identify the propagation source when there is only one source. In many applications, there may be multiple information sources in the network, such as rumors that often originate from more than one person to propagate quickly and widely [8]. In social networks, there is some uncertainty in the propagation paths of the sources of information propagation, and the influence ranges of multiple sources are mixed with each other, which makes it difficult to locate multiple sources compared to a single source. Some researchers use infection timestamps to investigate this problem. Such methods use the observations of the infection times of a subset of nodes in the network [9–11]. Fu et al. [12] propose a backward diffusion-based source localization method and find that multiple sources can be located with high accuracy even when the fraction of observers is small and the time delay along the links is not known exactly. Tang et al. [13] introduce a new heuristic involving an optimization over a parametrized family of Gromov matrices to develop an estimation algorithm for both a single source and multiple sources. However, the above studies assume that the propagation delays along each edge and the number of sources are known. Such assumptions can be restrictive for practical applications of multiple source localization.
Many of the works on information source localization are based on the knowledge of the network topology, the infection status and the infection time of a portion of nodes in the network. However, information propagation usually occurs in a large scale social network, and it is challenging to obtain the infection status and the infection time of a large number of nodes. Thus, Yang et al. [14] introduce a novel deployment scheme based on the naming game theory [15], which does not need to obtain the infection states or infection times of the nodes. The experimental results show that the method can achieve better localization accuracy with a small number of observations for single source information propagation. To improve the speed and scope of information propagation, more than one source is selected at the beginning of information propagation. However, the multiple source localization problem is much more complicated than the single source localization problem.

In this paper, our goal is to propose a strategy to solve the problem of multiple information source localization when the number of information sources is unknown a priori. We hope to achieve a high location accuracy with as few observations as possible. We first need to know the topology of the network. Then, we select some nodes as observations. Next, we need to know the local sources of each observation. Here, ‘local sources’ are defined as the nodes that first spread the information to the observation, as in [14]. On these bases, this paper proposes a method to locate multiple information sources in social networks based on naming game theory. According to the information provided by the observations, we modify the network topology by removing edges that are not needed for information propagation. Moreover, label propagation is used to divide the observations into different sets [16]. Then, a maximum likelihood estimation function is used to calculate the estimated sources of each set of observations. Finally, all the expected sources of the information propagation are obtained by filtering the results. In this method, we place the selection of observations in the information source localization process, and the number of observations is set on demand rather than predetermined. Moreover, the number of sources is not required beforehand. The simulation results demonstrate that our approach is able to infer multiple sources from a portion of the observations.

The rest of this paper is organized as follows. Section 2 introduces the model and methods for multiple information source localization in social networks using the naming game. Section 3 describes the simulation results with analyses and comparisons. Section 4 is the conclusion section.

2. Model and methods

In this section, we first describe our social network model, the fundamentals of the Naming Game, and the assumptions.

Generally, in a social network, a node is an individual and an edge exists between any two individuals if they have any kind of social relationship, such as friendship or followership. In most social networks, information propagation is directed. In particular, if the connection of edges in a network is undirected, we treat each undirected edge as two directed edges. Thus, the information propagation network is modeled by a directed graph $G = (V, E)$, where $V$ and $E$ are the node set and edge set of $G$, respectively. An edge from node $u$ to $v$ is noted by $e_{uv}$. We then define $N = |V|$ and $M = |E|$ to represent the number of nodes and the number of edges in the network. When information propagates in a social network, there must be a small number of nodes that first spread the information to their neighbors, and they are called information propagation sources. Subsequently, the nodes that receive the information spread it to their neighboring nodes. This repetition of the process allows information to spread quickly across the network. Many propagation models can well represent the propagation process, such as the classical susceptible-infected (SI) [17], susceptible-infected-susceptible (SIS) [18], and susceptible-infected-removed (SIR) [19] models. Many information source localization methods are designed for a certain type of propagation model. However, it is impossible to know the type of propagation and its parameters in the actual localization of information propagation sources. Thus, our work does not focus on modeling the information propagation process. For a better practical application, we should estimate the propagation sources of the information without knowing anything about the propagation process, including the propagation type, the parameters and the number of sources.

The naming game studies the emergence of shared lexicons in a population of agents about some objects they observed, such as naming a new object [20]. The rules of a naming game model [21–23] are as follows. Initially, every agent has a memory. In each iteration, randomly pick a node as the speaker and a node as the hearer that are neighbors to each other in the network. Then, the speaker says a word to the hearer. The word is randomly picked from the speaker’s memory or, if it has an empty memory, from external lexicons. If, by coincidence, the hearer has the same word in its memory, which means that the speaker and hearer reach a consensus, then both of them will clear their memories except for the common word; otherwise, the hearer will add the new word to its own memory. This process iterates until the whole population reaches the global consensus for which all the agents have one and only one common word in their memories.

2.1. Methods for selecting observations

It is essential to obtain some specific information in the propagation process to accurately locate the information sources. Many works identify the most likely information sources based on the knowledge of the infection status of all the nodes in the network. However, it is impractical to track the infection status of each node in a large scale social network. Thus, we obtain information from a specific set of a limited number of nodes in a network. We denote a node in this specific set as an observation. Different from the diffusion of an epidemic, information propagation has specialized characteristics, such as the decision to spread information is determined by the personal will of individuals, and each individual clearly knows which one is its local source (the node that first spreads information to the individual). Inspired by the naming game, we adopt a similar method developed in [14] for selecting observations. Specially, we add an extra node $n^*$ that is connected to all the nodes in the underlying network and represents an individual or organization that detects the information sources. Thus, $n^*$ could obtain the information provided by each node in $G$. We perform the edge removal operation on the network according to the local source information provided by the observations. Then, we iteratively determine the observations until a certain number of edges have been removed. In each iteration $t$, we randomly chose an observation $o_t$ as the speaker and let $n^*$ be the hearer. The speaker $o_t$ tells its local sources to the hearer $n^*$. We denote $LS_{o_t} = \{S_{o_t,1}, \ldots, S_{o_t,k}\} \subset V$, which contains the local sources of observation $o_t$. We assume that all the speakers are honest. This method has good practicability because it makes it possible to obtain the local sources of each observation. Table 1 clarifies the main symbols used in the model developed in this paper and their specific meanings.

Through the information provided by each observation, we can obtain a propagation path graph $G_2$ and it can be simplified by $G$. Initially, $G_2 = G$. For each $o_t$, we remove some of the edges in $G_2$ according to the contents of set $LS_{o_t}$. We let $F_{o_t} = \{f_{o_t,1}, \ldots, f_{o_t,k}\} \subset V$ denote the set that contains the neighbors of observation $o_t$. In addition, we denote $NL_{o_t} = F_{o_t} - LS_{o_t}$, which contains the neighbors of observation $o_t$ that are not re-
Table 1
Model notations and definitions.

| Notation | Definition |
|----------|------------|
| n^*      | The individual or organization that detects the information sources |
| a_t      | The observation in iteration t |
| LS_0     | The set that contains the local sources of observation a_t |
| S        | The set that contains the actual information sources |
| G_2      | The propagation path graph |
| F_0      | The set that contains the neighbors of observation a_t |
| C_F_0    | The set that contains the common neighbors of the observation node a_t and its local sources |
| O        | The set of observations |
| ALS      | The set of local sources for all observations |
| N_Lot    | The set that contains the neighbor of the observation a_t that is not a local source |
| ANL      | The set that contains the nonlocal source neighbors of all observations |
| \(\rho_k^{(i)}\) | The probability vector \(\rho_k^{(i)} = (\rho_k^{(i)}_1, ..., \rho_k^{(i)}_N)\) where \(\rho_k^{(i)}_j\) represents the probability that node \(i\) is the propagation source of the \(k\)-th observation set when considering the information provided by the first \(t\) observations |
| d(u, v)  | The length of the shortest path between \(u\) and \(v\) in \(G\) |
| \(\Omega\) | The set of the estimated information sources |
| \(|d|\)  | The network diameter |
| \(|\beta|\) | The average degree |
| \(|c|\)  | The average clustering coefficient |

garded as local sources. We let \(C_{F_0} = \{c_{F_0, 1}, ..., c_{F_0, k}\} \subset V\) denote the set that contains the common neighbors of observation node \(a_t\) and its local sources. It is obvious that there are some edges in the network that information does not propagate through. Thus, we identify and remove these edges in \(G_2\) as accurately as possible. If \(LS_0 = \emptyset\), all the neighbors of \(a_t\) do not spread the information. Thus, \(G_2\) must not contain the edges starting from the neighbor of \(a_t\), and they should be removed. Similarly, if \(LS_0 \neq \emptyset\), there are three types of edges that can be removed according to the information transmission order: (1) The edges whose start node is in the \(N_L_0\) set and whose target node is in the \(LS_0\) should be removed. (2) The edges whose start node is in \(N_L_0\) and whose target node is \(a_t\) should be removed. (3) The edges whose start and target nodes are all in \(LS_0\) should be removed. Our method does not determine all the observations beforehand, and it only confirms the existence of an observation at each iteration. We denote \(NE\) as the percentage of the removed edges. We stop selecting the observations when \(NE\) reaches a certain value. Then, we obtain the simplified network graph \(G_2\), the set of observations \(O = \{a_1, ..., a_t\} \subset V\), and the set of local sources for all the observations \(ALS = LS_0 \cup ... \cup LS_0\). We also denote \(ANL\) as the set that contains the nonlocal source neighbors of all the observations. It is clear that the nodes in set \(ANL\) must not be the information source.

2.2. Method for multiple source localization based on the pre-set observations

We denote \(S = \{s_1, ..., s_k\} \subset V\), which contains the actual information sources. Our goal is to identify set \(S\) from the network topology and the local sources provided by the observations. In the previous Section, we introduced a method for selecting observation nodes. In this Section, we divide the observations into sets according to the simplified network \(G_2\). Then, we find the source of each observation set according to the information provided by observations.

The localization of multiple information sources is more complicated than identifying a single source. This is because different observations may receive information from different sources in the case of multiple sources. In this section, we discuss how to divide observations into several sets. Compared with the observations in different set, the observations in the same set have higher probability of receiving information from the same source. We overcome this challenge through label propagation. We propose the observation division algorithm in Algorithm 1 to find some suitable observation sets. First, we take all the observations as the sources to carry out reverse label propagation and find the nodes that can spread the information to each observation in a certain number of steps. In a network with diameter \(|d|\), the longest distance between any node and its source is \(|d|\). Without any loss of generality, we choose \(|d|/2\) steps to find the nodes that can spread the information to each observation. Thus, if a node receives labels from multiple observations, these observations are considered to be propagated by the same source, that is, they belong to the same set. We denote an \(N \times N\) matrix \(OD\) to record the observation labels received by each node. Initially, if \(j \in O\) and \(i = j\), \(OD(i, j) = 1\), else \(OD(i, j) = 0\), where \(1 < i, j < N\). It means that label propagation starts at each observation. Subsequently, the value of \(OD(i, j)\) changes from 0 to 1, indicating that node \(i\) propagates the label of observation \(j\). After \(|d|/2\) rounds of tag propagation, we obtain the matrix \(OD\) for the result of the reverse label propagation, and the \(i\)-th row of this matrix contains the observations that can be spread by node \(i\) in \(|d|/2\) steps. It is straightforward that if a node can spread the information to an observation, then it also can spread the information to the nodes that the observation can reach. Thus, we reorganize the result matrix in an iterative way. Then, we eliminate all-zero rows and the rows that can be contained by others. Finally, we obtain the matrix \(OD\) in which the non-zero items in each row indicate the observations getting information from the same source node. A parameter \(top\) is defined as the number of rows of matrix \(OD\), and it also represents the number of observation sets.

Algorithm 1. Observation Division

```plaintext
01 for m starts from 1 to \(|d|/2\) do 02 \(OD_1 = OD\) 03 for each node \(i \in V\) do 04 for each node \(j \in V\) do 05 if \(e_{i,j}\) is an edge of graph \(G_2\) 06 for each node \(a_t \in O\) do 07 if \(OD_1(j, a_t) = 1\) and \(OD_1(i, a_t) = 0\) then 08 \(OD_1(i, j) = 1\) 09 end if 10 end for 11 end if 12 end for 13 end for 14 end for 15 flag=1 16 while flag=1 17 \(OD_2 = OD\) 18 for each node \(u \in V\) do 19 for each node \(a_t \in O\) do 20 if \(OD_2(u, a_t) = 1\) then 21 if \(u \in O\) and \(u \in ALS\) then 22 \(OD(u, :) = OD(u, :) \cup OD(\cdot, a_t)\) 23 end if 24 end if 25 end for 26 end for 27 if \(OD = OD_2\) then 28 flag=0 29 break the while loop 30 end if 31 end while 32 delete the rows of all zeros in OD. 33 delete the rows in OD that are subsets of others.```
After dividing the set of observations, another problem to be solved is to find the source of each set. It is obvious that the observations receive information later than their local source. However, it is impossible to set all the nodes as observations. In general, the longer the propagation path from the source to the node, later the node receives the information. Hence, in each observation set, for each observation \(o_i\), we update the probability that a node is the source of the observation set with the difference between the distance of the node to \(o_i\) and the distance of the node to its local source set. When all the observations are considered, the node with the highest probability is the source of the underlying set. In detail, in set \(k(1 < k < r_{OD})\), we denote \(p_k^{(i)} = (p_{k,1}^{(i)}, ..., p_{k,N}^{(i)})\) as a probability vector, where \(p_{k,i}^{(i)}\) represents the probability that node \(i\) is the source of the \(k\)-th observation set when considering the information provided by the first \(k\) observations. Initially, any node in set \(V\) is equally likely to be the source, i.e., for \(\forall i \in 1, 2, ..., N\), \(p_{k,i}^{(0)} = 1/N\). Then, for each observation \(o_i\), we quantify the effect of the information provided by \(o_i\) on the value of \(p_{k,i}^{(i)}\). If \(OD(k, o_i) = 0\), it means that \(o_i\) does not belong to the \(k\)-th observation set. Thus, we have \(p_{k,i}^{(i)} = p_{k,i}^{(i-1)}\). If \(OD(k, o_i) = 1\), it means that \(o_i\) belongs to the \(k\)-th set. It is obvious that the node in set \(NLS_{k,i}\) must not be the source. We denote \([u, v]\) as the shortest path between nodes \(u\) and \(v\). Let \((u, v)\) be the length of the shortest path between \(u\) and \(v\) in \(G_2\). When \((v, o_i) = \text{inf}\), node \(i\) cannot propagate the information to the observation \(o_i\). Thus, node \(i\) must also not be the propagation source. We denote \(Q\) as the set of all nodes that satisfy \(d(i, o_i) = \text{inf}\) \((i \in V)\). For the sake of calculation, we denote \(T = LS_{k,i} \cup Q\) and update the value of each \(p_{k,i}^{(i)}\) according to (1).

\[
p_{k,i}^{(i)} = \begin{cases} 
p_{k,i}^{(i-1)} + \frac{\sum_{j \in Q} p_{k,j}^{(i-1)}}{n_2}, & i \notin T \\
0, & \text{otherwise}
\end{cases}
\]

where \(n_2\) represents the number of nodes which are not in \(T\).

Moreover, we denote \(d(i, LS_{k,i}) = \max(d(i, LS_{k,i}, r))\) \(\text{ls}_{i tolerate} \in LS_{k,i}\), \(d(i, LS_{k,i}, r) = \text{inf}\) as the distance from node \(i\) to set \(LS_{k,i}\). Then, we denote \(\eta = d(i, o_i) - d(i, LS_{k,i})\) as the deduction between \(d(i, o_i)\) and \(d(i, LS_{k,i})\), and \(\eta\) is larger than zero. The node in \(LS_{k,i}\) would receive the information earlier than node \(o_i\). Thus, the larger the value of \(\eta\), the greater the probability that node \(i\) is the propagation source of the \(k\)-th observation set. In order to avoid negative distance difference, we define \(\epsilon = \min \eta\). Based on this, we further update the probability that node \(i\) is the source of the \(k\)-th observation set.

\[
p_{k,i}^{(i)} = \begin{cases} 
p_{k,i}^{(i-1)} + \frac{\alpha}{\alpha+1} + \frac{\eta - \epsilon - 1}{\sum_{j \in Q} (\eta - \epsilon - 1) \times (\alpha+1)}, & p_{k,i}^{(i-1)} \neq 0
\\p_{k,i}^{(i-1)} = 0, & \text{otherwise}
\end{cases}
\]

where \(\alpha = \sum_{r=1}^{r_{OD}} OD(k, r)\), it means that \(o_i\) is the \(\alpha\)-th observation in the \(k\)-th set.

When \(t\) equals \(0\), the information provided by each observation is evaluated in the estimation process of the source node, that is, the probability update process stopped. The node with the biggest probability of (3) is considered as the source of the \(k\)-th observation set.

\[
\beta_k = \arg \max_{i \in V} p_{k,i}^{(i)}, k = 1, 2, ..., r_{OD}
\]

Finally, a set of the estimated multiple sources information is obtained by merging the source of each observation set, which we denote as \(\Omega = \{\beta_1, ..., \beta_{r_{OD}}\} \subset V\).
From Fig. 1 and Fig. 2, it can be observed that the range of the error quantity $\Delta$ is larger and the range of the error distance $\phi$ is smaller in the network that has a lower network diameter. It is because the smaller the network diameter is, the smaller the distance between the nodes in the observation set is. This results in a large number of observation sets. More the number of sets is, the more estimated sources are and the smaller the distance between the estimated sources and the actual sources is.

3.2. Efficiency of the proposed method

In our method, the more observations there are, the more iterations and time required. Thus, the number of observations should be used to evaluate the effectiveness of our information source location method. We expect to locate the sources accurately by selecting as few observations as possible. The smaller the number of observations, the better the efficiency that can be obtained. Corresponding to the accuracy in Fig. 1 and Fig. 2, Fig. 3 shows the average percentage of observations versus different NE in the four social networks. As NE increases, the accuracy of the algorithm improves, and we can see that the average number of observations increases. Furthermore, when NE=50%, the average percentage of observations we need is less than 49.53% in the Enron network, 43.5% in the Simmons81 network, 42.78% in the Hamilton46 network, and 40.13% in Wake73 network.

3.3. Comparison with other source location methods

The experimental results in Sections 3.1 and 3.2 show that our proposed method can achieve high-localization accuracies. Fig. 4 shows a comparison of the average error distance between our method and the method proposed in [13]. Our proposed method outperforms the method proposed in [13] in terms of accuracy when setting up a similar number of observations.

4. Conclusions and future work

In this study, we develop a method to estimate multiple information sources using a subset of observations when the number of sources is unknown. Our method uses the theory of the naming game and label propagation flexibly. The method we proposed can estimate the number and location of actual sources accurately without knowing the type and parameters of the propagation model. Moreover, the number of observations is determined in the localization process rather than artificially determined at the beginning of localization. Experimental evaluations with real-world data reveal that our method can locate the information sources within a small number of hops from the actual sources. Although our experiment is only based on SIS propagation model, our method focuses on the sequence of time when the neighbors first propagate messages, rather than the state transition time of each node. Therefore, different propagation models seldom affect the results. Because of this reason, our method can not only be used in other propagation models, but also do not need to know the content of the propagation model in the work of source location. However, generally, with the same fraction of observations, the multiple source estimation performance is not as good as that of single source estimation. In the future, we will focus on designing a better observations selection algorithm to improve the localization accuracy of multiple information source.
Fig. 2. Distribution of the error distance (φ) in different datasets. (a) Enron, (b) Simmons81, (c) Hamilton46, and (d) Wake73.

Fig. 3. Average percentage of observations versus different NE in the four social networks.

Fig. 4. Comparison of the average error distance between our method and the method in [13].

CRediT authorship contribution statement

Xue Yang: Conceptualization, Data curation, Methodology, Software, Writing - original draft, Writing - review & editing. Zhiliang Zhu: Conceptualization, Methodology, Supervision. Hai Yu: Conceptualization, Investigation, Methodology, Visualization. Yuli Zhao: Conceptualization, Methodology, Writing - review & editing. Ying Wang: Conceptualization, Methodology, Software, Validation.

Declaration of competing interest

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled.
Acknowledgement

This research was supported by the National Natural Science Foundation of China (Grant Nos. 61977014, 61902056, 61603082), the Fundamental Research Funds for the Central Universities (Grant Nos. N2017016, N2017011).

References

[1] V.J. Munster, M. Koopmans, N. van Doremalen, et al., A novel coronavirus emerging in China - key questions for impact assessment, N. Engl. J. Med. 7 (7) (2008) 2817–2827.
[2] S.H. Lim, S.W. Kim, S. Park, et al., Determining content power users in a blog network: an approach and its applications, IEEE Trans. Syst. Man Cybern., Part A, Syst. Hum. 41 (5) (2011) 853–862.
[3] F. Yang, X.W. Li, Y.Q. Xu, et al., Ranking the spreading influence of nodes in complex networks: an extended weighted degree centrality based on a remaining minimum degree decomposition, Phys. Lett. A 382 (34) (2018) 2361–2371.
[4] M.L. Bertotti, J. Brunner, G. Modanese, Innovation diffusion equations on correlated scale-free networks, Phys. Lett. A 380 (33) (2018) 2475–2479.
[5] K. Zhu, Z. Chen, L. Ying, Locating the contagion source in networks with partial timestamps, Data Min. Knowl. Discov. 30 (5) (2015) 1217–1248.
[6] A. Louni, K.P. Subbalakshmi, A two-stage algorithm to estimate the source of information diffusion in social media networks, in: Proceedings of the 33ed IEEE Annual Conference on Computer Communications (INFOCOM WKSHP5), 2014, pp. 329–333.
[7] B. Spinelli, L.E. Celis, P. Thiran, Back to the source: an online approach for sensor placement and source localization, in: Proceedings of the 26th International Conference on World Wide Web (WWW), 2017, pp. 1151–1160.
[8] W.Q. Luo, W.P. Tay, M. Leng, Identifying infection sources and regions in large networks, IEEE Trans. Signal Process. 61 (11) (2013) 2850–2865.
[9] S. Zejnilovic, J. Gomez, B. Sinopoli, Network observability and localization of the source of diffusion based on a subset of nodes, in: Proceedings of the 51st IEEE Annual Allerton Conference on Communication, Control, Computing, 2013, pp. 847–852.
[10] S. Zejnilovic, J. Xavier, J. Gomez, et al., Selecting observers for source localization via error exponents, IEEE Int. Symp. Inf. Theory (2015) 2914–2918.
[11] Z.L. Hu, Z.S. Shen, C.B. Tang, et al., Localization of diffusion sources in complex networks with sparse observations, Phys. Lett. A 382 (14) (2018) 931–937.
[12] L. Fu, Z.S. Shen, W.X. Wang, et al., Multi-source localization on complex networks with limited observers, Europhys. Lett. 113 (1) (2016) 18006.
[13] W.C. Tang, F. Ji, W.P. Tay, Estimating infection sources in networks using partial timestamps, IEEE Trans. Inf. Forensics Secur. 11 (12) (2018) 3035–3049.
[14] X. Yang, Z.L. Zhu, H. Yu, et al., A naming game-based method for the location of information source in social networks, Complexity 2020 (2020) 6975250.
[15] Q.M. Lu, G. Korniss, B.K. Szymanski, The naming game in social networks: community formation and consensus engineering, J. Econ. Interact. Coord. 4 (2009) 221–235.
[16] T. Wu, Y. Guo, L.T. Chen, et al., Integrated structure investigation in complex networks by label propagation, Physica A 448 (2016) 68–80.
[17] M. Barthélemy, A. Barrat, R. Pastor-Satorras, et al., Dynamical patterns of epidemic outbreaks in complex heterogeneous networks, J. Theor. Biol. 235 (2) (2005) 275–288.
[18] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, Phys. Rev. Lett. 86 (14) (2001) 3200–3203.
[19] Y. Moreno, R. Pastor-Satorras, A. Vespignani, Epidemic outbreaks in complex heterogeneous networks, Eur. Phys. J. B, Condens. Matter 26 (4) (2002) 521–529.
[20] L. Steels, A self-organizing spatial vocabulary, Artif. Life 2 (3) (1995) 319–332.
[21] B. Li, G.R. Chen, T.W.S. Chow, Naming game with multiple bearers, Commun. Nonlinear Sci. Numer. Simul. 18 (5) (2013) 1214–1228.
[22] S.K. Maitly, A. Mukherjee, F. Tria, et al., Emergence of fast agreement in an overhearing population: the case of the naming game, Europhys. Lett. 101 (6) (2013) 68004.
[23] Y. Lou, G. Chen, Analysis of the “naming game” with learning errors in communications, Sci. Rep. 5 (2015) 12191.