Topology Sensing in Wireless Networks by Leveraging Symmetrical Connectivity

Zitong Liu[1], Jiachen Sun[2], Feng Shen[1], Guoru Ding[2](✉), and Qihui Wu[1]

[1] Key Laboratory of Dynamic Cognitive System of Electromagnetic Spectrum Space, Ministry of Industry and Information Technology, College of Electronics and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China
zt_liu@126.com, sfjx_nuaa@163.com, qihuiwu2014@sina.com

[2] College of Communications Engineering, Army Engineering University, Nanjing 210007, China
sun_jiachen@outlook.com, dr.guoru.ding@ieee.org

Abstract. With the popularization of wireless networks, the role of machine intelligence is becoming more and more important, where the core is that the network needs to make its own decisions through learning. Topology sensing is a fundamental issue in the field of network intellectualization, but most of the related existing studies have focused on wired networks, while the characteristics of wireless networks are relatively few investigated. In this paper, a wireless channel-oriented topology sensing method based on Hawkes process modeling is proposed for the wireless network with symmetrical connectivity. Simulation are carried out to demonstrate that how to combine wireless channel with Hawkes process and how to further process the results to improve performance.

Keywords: Topology sensing · Symmetrical connectivity · Hawkes process · Wireless channel

1 Introduction

Research on intelligent analysis of network behavior, dynamic analysis of network topology and real-time analysis of key nodes/links can realize automatic identification of network nodes, dynamic analysis of network connectivity and automatic generation of network topology. Traditional network topology sensing usually requires a lot of prior information, such as operating mode and network protocol. However, considering reconnaissance on the wireless network of adversary Unmanned Aerial Vehicle (UAV), we are more concerned about the information inside the network, which is difficult to obtain for an observer outside the network. Therefore, topology sensing with limited information is what we need to do next.

In [1], it introduces the background, methods and applications of graph topology sensing from a macro perspective and points out that learning-based approach is the

X. B. Zhai et al. (Eds.): MLICOM 2019, LNICST 294, pp. 49–57, 2019.
https://doi.org/10.1007/978-3-030-32388-2_5

development trend of topology sensing. Reference [2] aims at identifying network connections by modelling the information transmission process as Granger Causality (GC) and puts forward the method of time window to solve the problem of data fusion. Dynamic Causality Models (DCMs) is proposed in [3] to model interactions among neuronal populations, which can be extended to the field of communication. Hawkes process has only recently been applied to topology sensing, [4, 5] use it to recognition connection and present a method named Low Cost Paths for Acyclic Graphs (LCPAG) to discover event chains. However, the above works have their limitations: First, a clear physical model considering wireless channel has not been formulated. Second, the application conditions are so harsh that they unable to adapt to complex real scenarios.

Thus, improvement and extension on the Hawkes are made in response to the above problems. The main contributions of this paper are summarized as follows:

– We formulated a physical model for external topology sensing considering wireless channels.
– Using threshold setting and symmetry, we proposed a wireless channel-oriented topology sensing method based on Hawkes process.
– We not only presented topology sensing simulation of virtual data, but also proved performance by actual database.

The rest of this paper is organized as follows. Section 2 introduces the system model. In Sect. 3, a wireless channel-oriented topology sensing method based on Hawkes process is proposed by using threshold and symmetry. Then, in Sect. 4, we give the simulation results to prove the effectiveness of the algorithm. In addition, the conclusions and future prospects are mentioned in Sect. 5.

2 System Model

2.1 Network Physical Model

As shown in the Fig. 1, we formulate the physical model of the whole process as follows:

Target. As an external observer of the network, how to restore the network topology through effective algorithm under the condition of limited data, so as to get the key node information of the network.

Assumptions for Simplifying the Problem. There are only two kinds of information available to external observer: the node that transmits the data and the time when the transmission takes place. The information transmission channel within the network is an ideal channel, i.e. channel fading is not considered. The information transmission channel outside the network is AWGN channel without considering the influence of distance. For a long enough period of time, all links are used several times.

Process Description. The network S is an unknown wireless network, and the sensor O is an external observer. It monitors the transmission of information in the network S from $t = 0$. Once the information transmission is monitored, the corresponding event n_i and time t_i are automatically recorded. In the end, we can get two one-to-one corresponding sequences. The sequence of the originator is called the set of events, and the sequence of the recording time is called the set of times. Through the corresponding calculation of these initial data, we can get the interconnection of the network nodes.

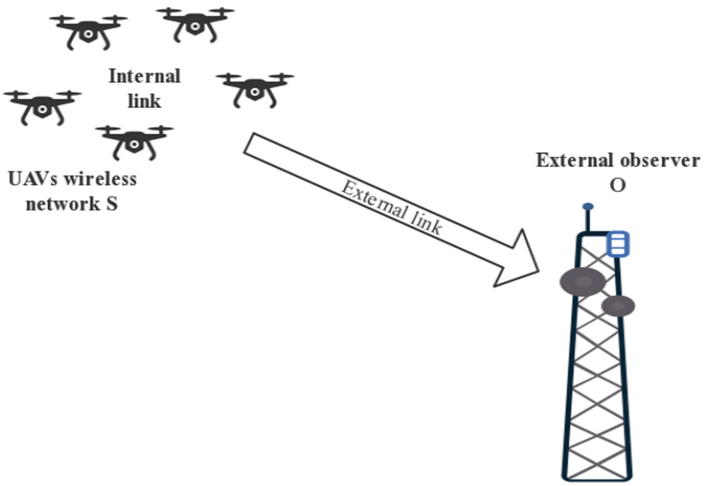

Fig. 1. Network physical model

2.2 Wireless Channel Model

There are two wireless channels in this problem, as shown in Fig. 1: One is the transmission channel of information within the network, and the other is the channel through which information is monitored by external observers outside of the network. For the first channel, because the target network we want to know is mostly used in military, such as UAV group communication network, we have reason to believe the high reliability of the network link. In order to simplify the problem, we do not consider the fading of this channel, as we said in the hypothesis. For the second channel, since the external observer listens in private without the permission of the internal network, it is likely to be subject to various external interference, so the fading of this channel cannot be ignored. As an exploration, we complete the simulation of AWGN channel without considering the distance.

2.3 Network Mathematical Model

Hawkes process is a point process which depends on autoregressive past events. Its main idea is that the transmission rate of an event at any time is a function of the recent events in the process [5].

For an event, in a case which has given t_k, the transmission rate of the event is

$$\lambda(t) = \mu(t) + A\sum_{k=1}^{K} \gamma(t - t_k), \tag{1}$$

which can also be called Conditional Intensity Function (CIF) [5]. The parameter $\mu(t) \geq 0$ is the basic rate of the event, or the rate of innovation; the parameter $A \geq 0$ indicates the degree of self-motivation of the event, that is, how much influence the event at time t_k has on the occurrence of the event at time t, so we can call it the self-motivation matrix; the kernel function $\gamma(t)$ represents the time relationship between the events, which is known, causal, non-negative and integrable. Here we take it as $\gamma(t) = e^{-t}$. This model can easily be extended to processes containing multiple subprocesses, where the transmission rate of one subprocess is affected not only by its own behavior, but also by the behavior of other subprocesses. We can use the multi-dimensional Hawkes process to model the communication process in wireless networks, and regard every information transmission in the network as a subprocess. For process with N subprocesses, the CIF of the ith subprocess can be obtained by formula (1) as

$$\lambda_i(t) = \mu_i + \sum\nolimits_{j=1}^{N} A_{ij} \sum\nolimits_{k \in K_j} \gamma(t - t_k), \tag{2}$$

in which K_j represents the event set in jth subprocess, μ_i represents the basic rate of the ith subprocess, which we think is not changing with time, A_{ij} quantifies how much the ith subprocess reacts to jth subprocess, $A_{ij} = 0$ represents the occurrence of the jth subprocess has no effect on the ith subprocess, and $A_{ij} > 0$ represents that the occurrence of the jth subprocess will lead to the temporary increase of the occurrence probability of the ith subprocess. We call it the influence matrix [5].

3 Topology Sensing Methods in Wireless Channels

3.1 Determination of Hawkes Process Parameters

In many cases, parameters μ_i and A_{ij} cannot be known beforehand, so we need to use mathematical tools to reasonably infer the parameters according to the observed values in a certain period of time to find the most realistic parameters. Here we use the maximum likelihood estimation method to determine μ_i and A_{ij} in $t \in [0, T]$, the negative log-likelihood function of the ith subprocess [5] is

$$L_i(\mu, A) = \int_0^T \lambda_i(t)\mathrm{d}t - \sum\nolimits_{K \in K_i} log\lambda_i(t_k). \tag{3}$$

Maximum likelihood function requires us to minimize this convex function. There are many ways to choose. Through practice, we find that the quasi-Newton method is simple and effective, so we finally use quasi-Newton method in the program.

3.2 Threshold and Symmetry

Using the influence matrix $\boldsymbol{A}$ obtained by Hawkes process and maximum log-likelihood function, redundant links can easily be generated due to a little interference. At this time, we can set a threshold to reduce redundant links, which can achieve self-adaptation by analyzing the preliminary simulation results. However, we must admit that it is contradictory to find real links and reduce redundant links. We need to find a balance between the two things according to different application scenarios and requirements.

Considering the cooperative communication in the scenario of UAVs, links usually have no direction. Therefore, the influence matrix should be symmetrical, using which we can simply optimize the results. When $A_{ij}\ \&\ A_{ji} > 0$, we think there is a connection, which we call the “and” rule; when $A_{ij} \,|\, A_{ji} > 0$, we think there is a connection, which we call the “or” rule. Selecting the “and” rule will cause some real links not to be found, while the “or” rule will cause link redundancy.

4 Performance Evaluation

4.1 Basic Simulation Setup

In the simulation, we first set up a simple tree network topology. The network has 25 nodes, including 5 root nodes and 20 leaf nodes. We do not consider the communication between the nodes and themselves. The true physical topology is shown in Fig. 2.

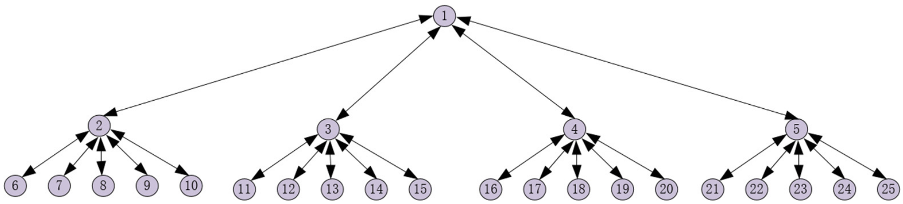

Fig. 2. True physical topology

Because the simulation tool we use is MATLAB, we cannot really restore the transmission of network node information. Firstly, we use the true CIF to weigh the sub-processes, so as to generate the ideal set of times and the ideal set of events. In order to explore the influence of AWGN channels on the performance of Hawkes process, we first simulate the channel and calculate its correct detection probability P_d and false alarm probability P_f. Then we filter the original data by P_d, and supplement the original data by P_d. Finally, we can get the initial data under the channel.

4.2 Simulation Result

As shown in Fig. 3, the simulation results are given when SNR = 10, $\lambda = 20$ (decision threshold of AWGN channel), $P_d = 0.9862$, $P_f = 0.4561$. It can be seen that the performance of the system is satisfying when SNR is relatively large. Figure 3(a) shows the true influence matrix, in which yellow represents connected, blue represents disconnected. Comparing the results of Fig. 3(b) and (c), we can find that all redundant links can be removed by setting threshold and utilizing symmetry. Finally, Fig. 3(d) can be got by binarizing the Fig. 3(c). For a binary classification problem, there are four kinds of discriminant situations. If a sample is true and the prediction is true, we call it true positive; if a sample is false but the prediction is true, we call it false positive. The Receiver Operating Characteristic (ROC) curve can well reflect the performance of the algorithm, which represents the relationship between the true positive weight and the false positive weight. The ROC curve at this point is shown in Fig. 4, in which the red dotted line represents the ROC when nothing is known, and the blue solid line represents the simulation results.

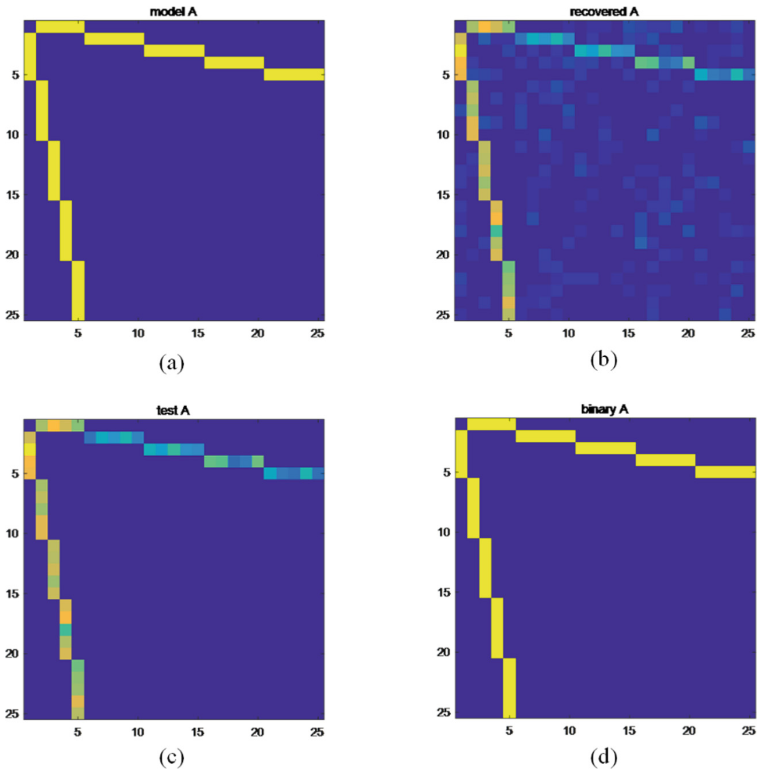

Fig. 3. Comparison of simulation results (a) True influence matrix (b) Reductive influence matrix (c) Reductive influence matrix with threshold (d) Influence matrix after binarization (Color figure online)

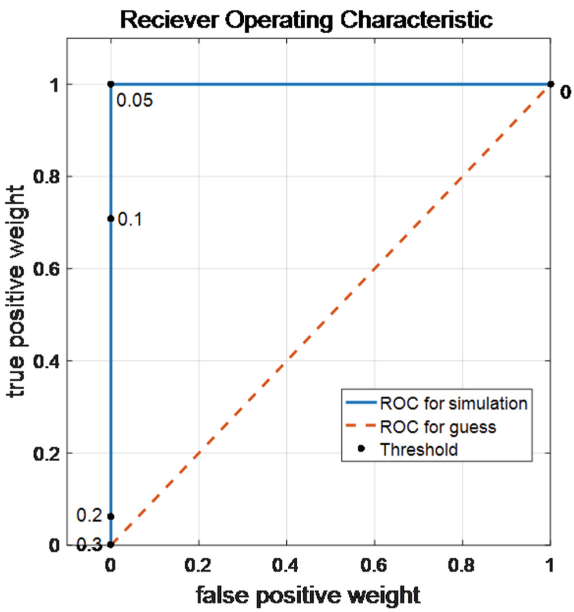

Fig. 4. Receiver operating characteristic curve (Color figure online)

4.3 Real Data Simulation

Table 1 is the part of the initial data in actual database, from left to right is the Index, Source ID, Destination ID, Start Time, Receive Time, Destination Abscissa, Destination Ordinate respectively. Our specific approach is to extract the first 10,000 data from the database as raw data, assuming that we only know srcID and starttime. And input the two sequences into the Hawkes process, we will eventually get the influence matrix, then compare with the real influence matrix, so that the conclusion will be drawn.

Table 1. Partial initial data

idx	srcID	dstID	starttime	receivetime	DstxPos	DstyPos
1	93	24	7.234305	17.234356	0.067886	−0.04369
2	24	93	17.23461	17.234668	0.05631	−0.04242
3	93	24	17.234822	17.23488	0.067886	−0.04369
4	93	24	17.234999	17.235057	0.067886	−0.04369
5	93	24	17.235491	17.235549	0.067886	−0.04369
6	93	24	17.235768	17.235826	0.067886	−0.04369
7	86	24	17.235945	17.236003	0.067886	−0.04369

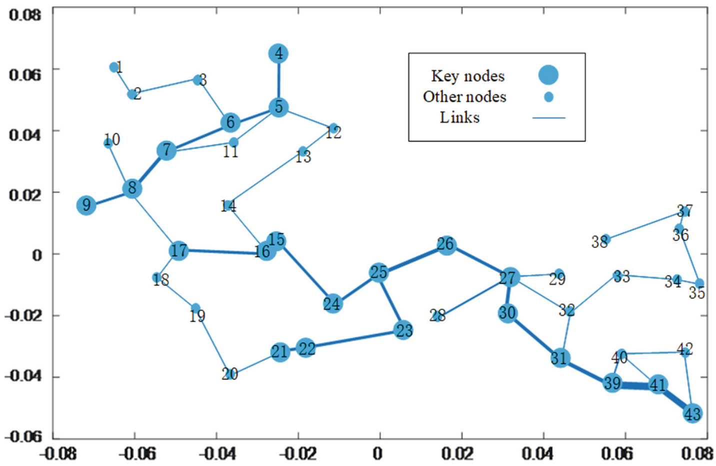

Fig. 5. True physical topology

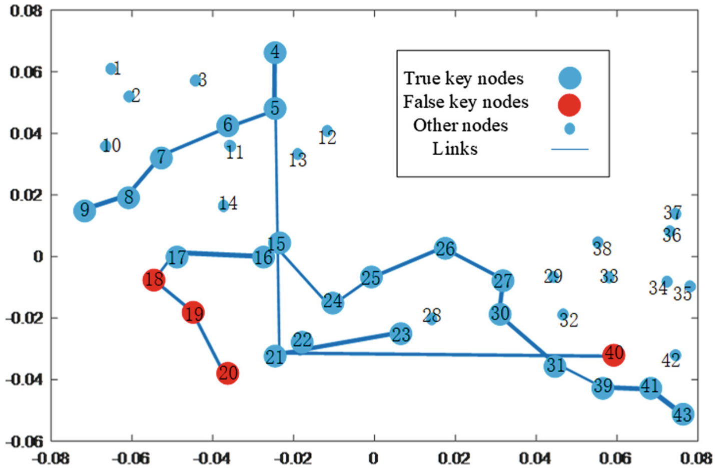

Fig. 6. Strong topological links under "or" rule with threshold of 0.3

Figure 5 is a true physical topology, in which the horizontal and vertical coordinates represent the physical location of nodes, and the thickness between nodes represents the strength of links. For convenience, we renumber the nodes, and if two nodes communicate more than 200 times, we consider these two nodes as key nodes. Since we want to find all key nodes, we choose "or" rule to symmetry the influence matrix. At this time, a large number of redundant links will be generated. In order to remove the redundant links, we set corresponding thresholds and get the final results, as shown in Fig. 6.

From the results, we can see that after symmetry and threshold setting, most redundant links have been eliminated, and all key nodes and strong links have been found. But what we have to see is that there is a case of error detection, i.e. redundant links that cannot be eliminated even if a higher threshold is set. Therefore, there is room for further performance improvement.

5 Conclusion and Outlook

Topology sensing is an important research direction. This paper has carried out exploratory research from three aspects: system model, method improvement and data simulation. Firstly, we combine topology sensing with wireless channel characteristics for the first time and formulate a clear physical model of topology sensing over wireless channel. Then, we propose a wireless channel-oriented topology sensing method based on Hawkes process using threshold and symmetry. Finally, we verify the reliability of the algorithm by real communication database. The expected research results have been achieved and the predetermined technical requirements have been fulfilled. Through the simulation results, we can see that the Hawkes process can still maintain good performance in AWGN channel, which is related to the threshold in the channel. How to set the threshold is also a very worthwhile problem. However, in real data simulation, the performance of Hawkes process is not ideal, but after symmetry and threshold processing, it can identify key nodes in the topology. We still have a long way to go to adapt the Hawkes process to the real situation.

At present, the work of topology sensing using Hawkes process is still in the immature stage. There are still many improvements in the future. This paper has also stimulated many new and interesting research directions, which is worth further research and exploration. such as the impact of distance on the channel, data preprocessing and so on. In addition to the Hawkes process, the use of machine learning for topology sensing is also a very challenging task, which we will leave to future work.

References

1. Giannakis, G.B., Shen, Y., Karanikolas, G.V.: Topology identification and learning over graphs: accounting for nonlinearities and dynamics. Proc. IEEE **106**(5), 787–807 (2018)
2. Tilghman, P., Rosenbluth, D.: Inferring wireless communications links and network topology from externals using granger causality. In: 2013 IEEE Military Communications Conference, MILCOM 2013, San Diego, CA, pp. 1284–1289 (2013)
3. Friston, K.J., Harrison, L., Penny, W.: Dynamic causal modelling. Neuroimage **19**(4), 1273–1302 (2003)
4. Moore, M.G., Davenport, M.A.: Analysis of wireless networks using Hawkes processes. In: 17th International Workshop on Signal Processing Advances in Wireless Communications, pp. 1–5. IEEE, Edinburgh (2016)
5. Moore, M.G., Davenport, M.A.: A Hawkes' eye view of network information flow. In: IEEE Statistical Signal Processing Workshop (SSP), Palma de Mallorca, pp. 1–5 (2016)