Discover Causality of Battlefield Sequential Events Based on THPM Algorithm

Chaoyang Li[1](✉), Yang Li[2], Zhimin Zhuo[2], and Yongjian Zhang[2]

[1] Beijing Institute of Technology, Beijing 100081, China
1402753041@qq.com
[2] Beijing Institute of Electronic System Engineering, Beijing, China

Abstract. Granger causality is a causality based on prediction, which exists among various events in complex systems. It is an important and challenging task to predict the causality of various combat events on the battlefield. In this paper, aiming at the combat events generated by each agent on the battlefield, the causality network between sequential combat events is constructed. With causality network, we can locate the fundamental events and conduct powerful countermeasures. In order to solve the problem caused by topological relations between agents in discovering event causality, we use Topology Hawkes Process Method (THPM) algorithm to construct causality network. We use simulate annealing algorithm to improve the performance of THPM. Compared with the traditional PCMCI algorithm, experiments show that the THPM algorithm based on simulated annealing algorithm has better ability to search the optimal causal graph and higher prediction accuracy.

Keywords: Granger causality · Battlefield events sequences · Topology hawkes process

1 Introduction

Learning Granger causality from the sequence of events is an important task in real life. Sociologists are concerned with causality between social events [1], network maintenance specialists focus on the causality among network alarms to find the root alarm, economists are interested in causality between economic activities for proposing policies conducive to economic development. In an adversarial situation, commanders analyze causality among combat events in order to make decisions that determine the situation.

Learning Granger causality is an important task of multi-dimensional point process [2]. From the perspective of graph model, learning Granger causality is to infer a directed acyclic graph from historical event data. In the directed acyclic graph, it can be inferred that event e_j is caused by event e_i when node e_i has one edge pointing to node e_j. Various methods have been proposed by researchers in the study of Granger causality of event sequence. These methods can be classified into two categories. One line is to construct model based on constraint, which mainly studies the independence between various

Z. Ren et al. (Eds.): *Proceedings of 2021 5th Chinese Conference on Swarm Intelligence and Cooperative Control*, LNEE 934, pp. 773–780, 2023.
https://doi.org/10.1007/978-981-19-3998-3_74

events. The typical method is PCICM which consists of two stages [3]: PC condition selection (PC) and the momentary conditional independence (MCI). Another line is the model-based Hawkes point process [4], which focuses on the process of generating event. The typical method is Proximal Graphical Event Models (PGEMs) whose learning process is entirely data driven and without the need for additional inputs.

The above methods assume that the event sequence is independent and identically distributed, and the occurrence of the event sequence is not affected by the network topology in the system. In the real complex system, especially in the battlefield situation, the generation of an agent's combat event is not only related to the historical events of it, but also affected by the topology network between agent group it belongs to. Topological Hawkes Processes Method [5] (THPM) is an algorithm that takes topology relations of nodes in the network as knowledge to learn Granger causality. The graph convolution in time domain is applied to deal with topology relations in network. And Expectation-Maximum (EM) is used to optimize the parameters of the maximum likelihood function to obtain the correct causality graph.

2 Preliminary and Problem Formalization

In this section, we mainly model and analyze the causality of combat events occurred in agents in the battlefield situation, and illustrate the relevant contents involved in the model.

2.1 Hawkes Process

Multivariate point process is a random process which can be expressed by $\varepsilon = \{v_i, t_i\}_{i=1}^{m}$, where $v_i \in V$ indicates the type of event and $t_i \in T$ indicates occurrence moment of the i-th event, m is the number of events in an event sequence [6]. This point process can also be represented equivalently as $C = \{C_v(t) | t \in T, v \in V\}$, where $C_v(t) \in R$ denotes the occurrence number of event whose type is v. Hawkes process is a typical multivariate point process, which mainly describes the influence of historical events on future events in a complex event sequence. A special intensity is proposed to measure the excitation or inhabitation degree that past events transferred to present event. The intensity function is shown in the formula (1).

$$\lambda_v(t) = \mu_v + \sum_{v' \in V} \int_{t' \in T_{t-}} \phi_{v',v}\left(t - t'\right) dC_{v'}\left(t'\right) \tag{1}$$

In formula (1), $T_{t-} = \{t' \in T,\ t' < t\}$ denotes the set of times less than t. This intensity function consists of two parts. The first part μ_v is the basic intensity that the event of type v occurs in time t. The second part is the excitation intensity collected from past events where $\phi_{v',v}(t)$ is an impact function characterizing the time-decay of the casual influence.

2.2 Graph Convolution

There are topological relations among combat agents on the battlefield. An undirected graph $g_N = (N, E_N)$ can be used for this topology, where N is the set of nodes, E_N is the set of edges between nodes in the graph [7]. Normalized Laplacian matrix of graph g_N is represented as $L = I - D^{-\frac{1}{2}} AD^{-\frac{1}{2}}$, where I is the identity matrix whose dimension is number of nodes in the graph and D is the degree matrix of graph g_N. The convolution operation on the graph g_N is generally represented as formula (2).

$$H^{l+1} = \sigma\left(LH^l W^l\right) \tag{2}$$

In formula (2), $\sigma(*)$ denotes activate function, W^l is weight matrix of the l-th layer and $H^0 = X$. X represents the matrix consisted of values of nodes in topology. The difference between ordinary convolution on the image and graph convolution operation is shown as Fig. 1.

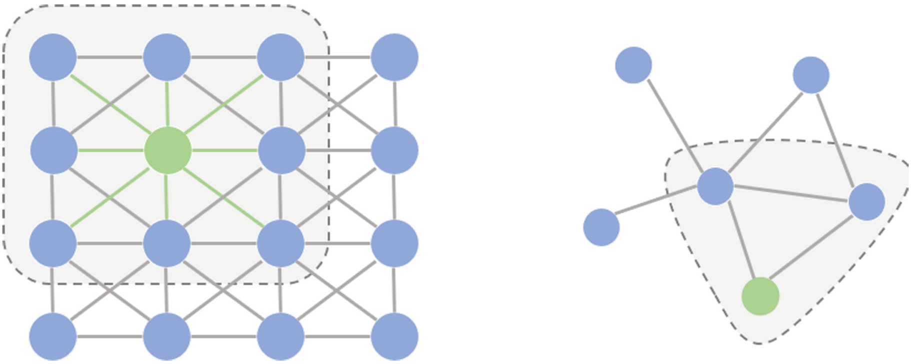

Fig. 1. Ordinary convolution and graph convolution

2.3 Problem Formalization

In order to accurately express the Granger causality between combat events, we formalize this problem as a topological Hawkes process. The undirected graph $g_N = (N, E_N)$ represents the topological relations among agents on the battlefield, while the directed acyclic graph $g_V = (V, E_V)$ represents the causality structure among various combat events. Adding topological relations to the Hawkes process, the traditional representation of a sequence of events $\varepsilon = \{v_i, t_i\}_{i=1}^m$ transform to $\{n_i, v_i, t_i\}_{i=1}^m$ where n_i means that the event whose type is v_i occurs in the agent whose id is n_i. The problem of learning causality of combat events can be described as: to find the causality graph of events in the event set V, given a set of observed sequence of combat events $\varepsilon = \{n_i, v_i, t_i\}_{i=1}^m, n_i \in N, v_i \in V, t_i \in T$ and topology graph g_N which abstracted from agents.

The intensity function in Hawkes process is essentially a convolution operation in time domain. In order to tackle topology relations, graph convolution operation is introduced into the intensity function. To simplify the problem, the continuous time is discretized as $T = \{0, \Delta t, 2\Delta t, \ldots, T\}$. The differential part in the intensity function can be discretized as a set of observed data $X = \{X_{n,v,t} | n \in N, v \in V, t \in T\}$. By convolution operation and time discretization processing, we can estimate the intensity function which includes various kinds of parameters. The intensity function including various kinds of parameters is represented as $\lambda_v(n, t, \Theta)$. In order to better fit the data set, we adopt the *log* maximum likelihood function as the objective function of optimization. The parameters to be optimized are g_V and the set of parameters Θ of the intensity function. The log maximum likelihood function is shown as formula (3).

$$L(g_V, \Theta; X, g_N) = \sum_{v \in V} \sum_{t \in T} \sum_{n \in N} \left[-\lambda(n, t)\Delta t + X_{n,v,t} log(\lambda_v(n, t)) \right] \tag{3}$$

Using maximum likelihood function as objective function will generate redundant edges in causal graph g_V. In order to suppress over-fitting, Bayesian Information Criterion penalty is added to objective function, which plays a similar role to l_0 regularization, whose purpose is to keep the sparsity of the learned parameters.

3 Experiments

3.1 Datasets

The dataset used in the experiments is sampled from the simulation environment, in which V combat events with causality relation are defined, a group composed of N agents with a certain topological relation is initialized, and a set of rules for generating combat events is defined. At a fixed time, the operational agent group continuously generates combat events according to established rules. When a combat event is detected, to record the type of the current event, the occurrence moment of the event and the agent that generated the event. The format of the generated dataset is shown in Table 1.

To generate several datasets, multiple simulation experiments were conducted. In every simulated experiment, the number of agent was conducted as N, the number of combat event was conducted as N.

Table 1. Format of dataset

Event_id	Agent_id	Start_time
9	28	33
6	37	41
4	17	43
4	25	48
…	…	…

3.2 Flowchart of Prediction Experiment

The flowchart of the prediction experiment is shown as Fig. 2.

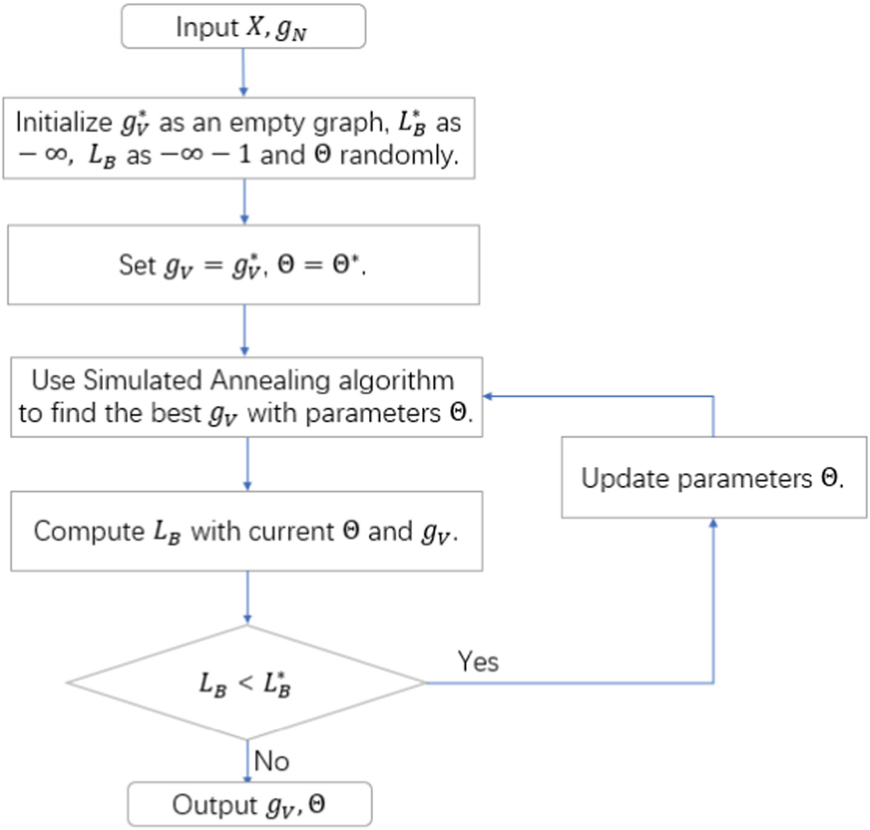

Fig. 2. Flowchart of learning causality

In this part, the simulated annealing algorithm is applied to the iterative optimization of THPM algorithm. In the process of iteration, poor g_V is accepted with a certain probability. Compared to traditional THPM, improved method has better results.

3.3 Experiment Results

In this part, we mainly compare with the traditional PCMCI algorithm. The main comparison indexes include *precision*, *recall* and *F1*. The confusion matrix is as shown in Table 2.

The formula of precision is as show in formula (4).

$$precision = \frac{TP}{TP + FP} \tag{4}$$

The formula of recall is as show in formula (5).

$$recall = \frac{TP}{TP + FN} \tag{5}$$

Table 2. Confusion matrix

True	Prediction	
	Positive	Negative
Positive	TP (True positive)	FN (False negative)
Negative	FP (False positive)	TN (True negative)

The formula of F1 is as show in formula (5).

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \tag{6}$$

On the same data set, the results of the two algorithms are shown as Table 3. and Figs. 3, 4, 5 and 6.

Table 3. Comparison of results predicted by two algorithms

Datasets	Algorithms	*Recall*	*Precision*	*F1*
$N = 38$ $V = 10$ $m = 21768$	PCMCI	1.0	0.8462	0.9167
	THPM	1.0	**1.0**	1.0
$N = 39$ $V = 11$ $n = 92845$	PCMCI	1.0	0.9	0.9474
	THPM	1.0	**1.0**	1.0
$N = 28$ $V = 12$ $n = 2297$	PCMCI	0.8	0.8	0.8
	THPM	1.0	**0.833**	0.9091
$N = 38$ $V = 13$ $m = 93329$	PCMCI	1.0	**1.0**	1.0
	THPM	1.0	0.9524	0.9756

In the comparison diagram, the element bits with black markings indicate a causal relationship between two combat events. It can be seen that THPM algorithm can find the causal relationship between events more accurately.

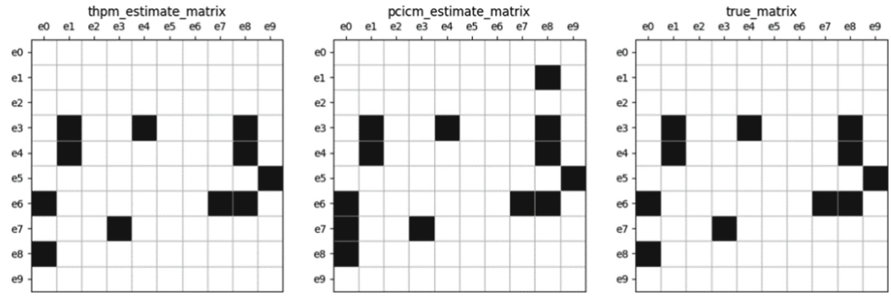

Fig. 3. Dataset with $N = 38$, $V = 10$, $m = 21768$

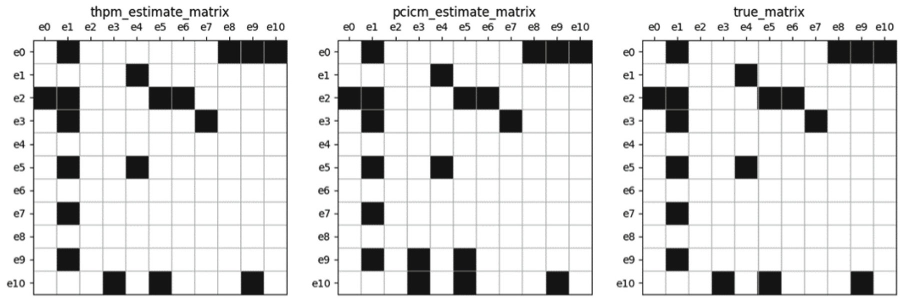

Fig. 4. Dataset with $N = 39$, $V = 11$, $m = 92845$

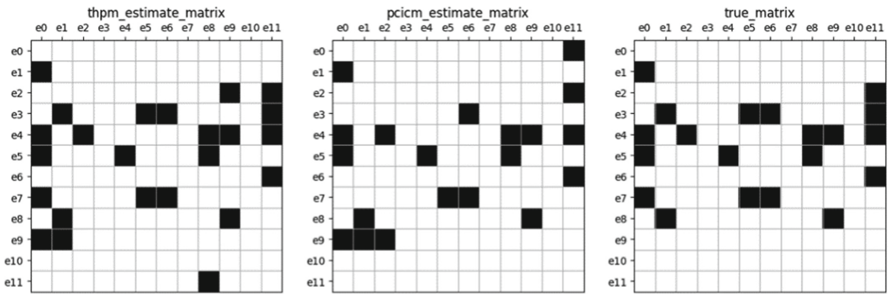

Fig. 5. Dataset with $N = 28$, $V = 12$, $m = 2297$

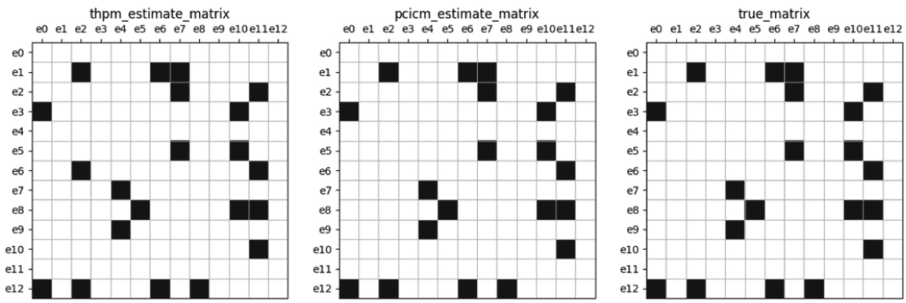

Fig. 6. Dataset with $N = 38$, $V = 13$, $m = 93329$

4 Conclusion

In this paper, we use simulated annealing to improve the topological Hawkes method to solve the problem of causality discovery of battlefield event sequences. When tested on several different data sets, the improved method achieves high accuracy. In future studies, we plan to extend our work to the problem of causality in the sequence of operational events in which clear topological relationships of operational units cannot be obtained.

References

1. Zhou, K., Zha, H., Song, L.: Learning social infectivity in sparse low-rank networks using multi-dimensional hawkes processes. In: Proceedings of the Artificial Intelligence and Statistics, pp. 641–649 (2013)
2. McGraw, M.C., Barnes, E.A.: Memory matters: A case for Granger causality in climate variability studies. J. Clim. **31**(8), 3289–3300 (2018)
3. Runge, J., Nowack, P., Kretschmer, M.: Detecting and quantifying casual associations in large nonlinear time series datasets. Sci. Adv. **5**(11), 4996 (2019)
4. Bhattacharjya, D., Subramanian, D., Gao, T.: Proximal graphical event models. Adv. Neural. Inf. Process. Syst. **31**, 8136–8145 (2018)
5. Cai, R., Wu, S., Qiao, J., Hao, Z., Zhang, K., Zhang, X.: THP: Topological hawkes processes for learning granger causality on event sequences. arXiv preprint arXiv:2105, 10884 (2021)
6. Salehi, F., Trouleau, W., Grossglauser, M., Thiran, P.: Learning hawkes processes from a handful of events. In: Proceedings of the Advances in Neural Information Processing Systems, pp. 12694–12704 (2019)
7. Defferrard, M., Bresson, X., Vandergheynst, P.: Convolutional neural networks on graphs with fast localized spectral filtering. In: Proceedings of the Advances in Neural Information Processing Systems, vol. 61(7), pp. 1644–1656 (2013)
8. Xu, H., Farajtabar, M., Zha, H.: Learning granger casuality for hawkes processes. In: Proceedings of the International Conference on Machine Learning, pp. 1717–1726 (2016)
9. Bonnet, A., Herrers, M., Sangnier, M.: Maximum likelihood estimation for Hawkes processes with self-excitation or inhibition. Statist. Probab. Lett. **179**, 214–221 (2021)