Journal of Information Hiding and Multimedia Signal Processing ©2019 ISSN 2073-4212
Ubiquitous International Volume **10**, Number **1**, January **2019**

A Robust Community Detection in Temporal Networks Based on Random Walks

Xiao Chen and Zhe-Ming Lu*

School of Aeronautics and Astronautics
Zhejiang University
Hangzhou, 310027, P. R. China
*Corresponding author: zheminglu@zju.edu.cn

Received March 2018; revised July 2018

Abstract. *Time-varying networks are common in many domains ranging from social media analytics to biology. An important task in these domains concerns the detection of community structures. Many existing algorithms for community detection in temporal networks attempt to detect communities in the whole evolution life of a network. In this paper, we propose a method to enhance the accuracy and robustness of the detection result of current network snapshot with the aid of previous and next snapshots. Our algorithm named STRW (Space-Time Random Walk) is based on a random walk on the multilayer network consisting of snapshots of a temporal network, and the transition probabilities defining the random walk are adapted by the local topological and temporal similarity between successive snapshots. Based on this space-time random walk, nodes are clustered in a hierarchical fashion. Experimental results show that the performance of STRW is superior to related community detection algorithms.*
Keywords: Community detection, Temporal network, Multilayer network, Random walk.

---

1. **Introduction.** Many real world systems, such as social and biological groups, are often represented as complex networks that capture the interaction among individuals [1, 2]. The individuals are interpreted as the nodes of the network and the relations among them are modelled by the links of the network. An important aspect of network analysis is the discovery of community structures that are defined as groups of nodes that are more densely connected to each other than they are to the rest of the network. The detection of such modules has proven important in the investigation of the underlying principles governing complex systems and has been a very active area of research over the past decade [3, 4, 5].

However, most real-world networks are evolving constantly and the increasing availability of time-varying network data such as on-line social networks has brought more attention to study time-varying networks [6, 7]. Although it is possible to create a static network by aggregating over the temporal evolution of the system, such temporally aggregated representations may ignore essential features of the system or may confound structures that can only be distinguished by retaining the time-varying information of the data. For example, a group of nodes may belong to different communities at different time, so aggregate the network over time will merge those communities and create a cluster that does not represent the reality of the system at any point in the time axis. Moreover, groups of nodes may share similar activity patterns over time, so an aggregated view on the network will only retain the topology of the interactions and lose the activity patterns and temporal correlations [4].

In this paper, we investigate the task of detecting communities that utilizes the topological correlation between two adjacent snapshots in a temporal network. We offer a novel approach of encoding the structure information of adjacent layers into the transition probabilities of the target layer. The walker is encouraged to jump between nodes in different snapshot layers that share similar topology and discouraged to jump across layers that do not share local topology. Consequently, the structure information of the adjacent snapshots is implicitly facilitated to explore communities of the target layer. Our algorithm STRW (Space-Time Random Walk) introduces a random walker that is able to jump not only in spatial dimension but in temporal dimension. The walker spends longer time moving between nodes with similar connections across layers. Then a distance measure between nodes based on these random walk procedures is introduced and finally an agglomerative clustering procedure is used to detect communities. The resulting algorithm can also be thought of as a temporal extension of the well-known Walk-Trap algorithm [8]. The remainder of this paper is organized as follows. In Section 2, we provide a concise literature review of time-varying cluster detection methods. Section 3 introduces our STRW algorithm. In Section 4, we provide an illustrative example to show the ability of STRW to discover more evolution patterns of a temporal network and compare the performance of STRW with other community detection methods based on a real world dataset. Section 5 provides conclusion and directions for future work.

2. **Proposed Evolving Network Model.** In this section, for a compact representation, we first introduce the adjacent tensor of a temporal network and discuss how the weights at each node between two snapshots can be defined. Taking the topology of the adjacent snapshots into account, we adapt the transition probabilities of a random walk and briefly discuss their properties.

2.1. **The Adjacency Tensor.** The snapshots of a time-varying network can be easily modelled as a multilayer network where each layer represents a snapshot. A multilayer network $M(V_M; E_M; V; L)$ is a graph $G_M(V_M; E_M)$ with more than one layer where V is a set of nodes, L is a set of layers, $V_M \subseteq V \times L$ is a set of nodes in different layers, and $E_M \subseteq V_M \times V_M$ is a set of edges. We use $i\alpha$ to denote node $i \in V$ in layer $\alpha \in L$ and $(i\alpha, j\beta) \in E_M$ to denote a directed edge from node $i\alpha$ to node $j\beta$. The structure of a multilayer network including both intra-layer and inter-layer edges can then be encoded in an forth order adjacency tensor A with elements

$$A_{i\alpha}^{j\beta} = \begin{cases} \omega_{i\alpha}^{j\beta}, & (i\alpha, j\beta) \in E_M \\ 0, & otherwise \end{cases} \tag{1}$$

where $\omega_{i\alpha}^{j\beta}$ is the weight of the edge between nodes $i\alpha$ and $j\beta$. In the case of temporal networks, we set $\omega_{i\alpha}^{j\beta} = 0$, if $i \neq j$ and $\alpha \neq \beta$, which means we do not set any connection between different nodes in any pair of snapshots. The inter-layer weights $\omega_{i\alpha}^{j\beta}$ of the same node, i.e. $i = j$ and $\alpha \neq \beta$, is usually interpreted as coupling strength. On the one hand, values of the weights close to 1 encourage the random walker transfer to another layer, which results in a similar community assignment of a node in two different layers. On the other hand, values of the weights close to 0 does not consider the effect of structures from different layers. Here we define the inter-layer weight reflecting the similarity in local topology between $i\alpha$ and $i\beta$ as the Jaccard similarity coefficient:

$$\omega_{i\alpha}^{i\beta} = \frac{|N_{i\alpha} \cap N_{i\beta}|}{|N_{i\alpha} \cup N_{i\beta}|} \tag{2}$$

where $N_{i\alpha}$ is the set of neighbours of node i in layer α and $N_{i\beta}$ is the set of neighbours of node i in layer β. It follows that $\omega_{i\alpha}^{i\beta} \in [0,1]$.

2.2. **Transition Probabilities.** When a discrete random walk process is running on a single layer network, at each time step a walker is on a node and moves to a node chosen randomly and uniformly among its neighbours. The sequence of visited nodes is a Markov chain, whose states are the nodes of the graph. As for the multi-layer network, the random walker should be allowed to move within and across layers. The structure of multi-layer network allows four possible moves that a random walker can make from node $i\alpha$. The corresponding transition probabilities associated to these four possible moves are defined as:

$$P_{j\beta}^{i\alpha} = \frac{A_{j\beta}^{i\alpha}}{d_{i\alpha}} \tag{3}$$

where $d_{i\alpha}$ is the generalized degree of node $i\alpha$ in A defined as $d_{i\alpha} = \sum_{j,\beta} A_{j\beta}^{i\alpha}$. As can be seen above, the transition probabilities not only depend on the intra-layer structure but also the inter-layer topological similarity. The idea behind these definitions consist in the assumption that a node in two adjacent snapshots tends to have several common neighbours if it retains the same community. The more similar they are, the more possibly the walker will jump between them. From Eq. (3) we can see that the maximal transition probability between layers $P_{j\beta}^{i\alpha}$ is equal to $\frac{1}{d_{i\alpha}}$. This means if the node $i\alpha$ already has large number of neighbours in layer α, the transition probability from $i\alpha$ to other layers will be very small. In other word, when the node $i\alpha$ already has many neighbours in layer α, there is enough information to identify its community and we hardly need any extra structure information from adjacent snapshots to revise the community in layer α. It is also notable that there exist inter-layer connections only between corresponding nodes. Therefore the probability to move from a node $i\alpha$ to other node $j\beta$ for $i \neq j$ and $\alpha \neq \beta$ is zero since there cannot exist a direct move where there is no connection.

The t-step transition probability is then defined by following multiple tensor multiplications:

$$(P^t)_{j\beta}^{i\alpha} = P_{j_1\beta_1}^{i\alpha} P_{j_2\beta_2}^{j_1\beta_1} \cdots P_{j\beta_1}^{j_{t-1}\beta_{t-1}} \tag{4}$$

To reduce the notational complexity in the tensor equations the Einstein summation convention is adopted as follows:

$$P_{j\beta}^{i\alpha} P_{i\gamma}^{j\beta} = \sum_{i=1}^{N} \sum_{\beta=1}^{L} P_{j\beta}^{i\alpha} P_{i\gamma}^{j\beta} \tag{5}$$

2.3. **Agglomerative Clustering.** In order to group the nodes into communities, we follow the definition of the distance between the nodes in [6]. Here we only consider the distance between two nodes $i\alpha$ and $j\alpha$ in the same layer. The distance is small if they are in the same community and it can be directly computed from the transition probability tensor:

$$r(t)_{j\alpha}^{i\alpha} = \sqrt{\sum_{k=1}^{N} \frac{[(P^t)_{k\alpha}^{i\alpha} - (P^t)_{k\alpha}^{j\alpha}]^2}{d_{k\alpha}}} \tag{6}$$

The distance is small when two nodes from the same layer and in the same community because the probabilities to reach any other node in the layer starting from the two nodes will be approximately equal. The time step length t of the random walks must be sufficiently long to get enough information about the topology of the graph, but t must not be too long (see [6] for detail). The information of the adjacent layers has been implicitly encoded in the above definition and it plays the key role in making community detection robust against the noise in the current layer. Finally, we use agglomerative clustering to merge nodes in communities by the same method introduced in [6,7].

3. **Experimental Results.** In this section we study the effectiveness of our approach first on illustrative networks for which the partitioning in communities is known. In such a case, we show that our algorithm can not only successfully detect the network communities from noise but also detect the hidden structures that can hardly be discovered by traditional detection methods designed for single layer networks. We then use a high resolution dataset that describes the social interactions of children in a primary school[8]. It is shown that our approach achieves a competitive result to the state-of-art methods.

3.1. **Illustrative Networks.** The network consists of 35 nodes divided into four communities: three fixed communities with 10 nodes and a transitional community with 5 nodes. The node in the three fixed communities is fully connected and each pair of the communities is connected by a link. The nodes of the transitional community jump among the fixed communities. Each time the nodes join a fixed community, they first fully connect themselves and then set up connection to all the nodes in the fixed community. As a result, the transitional community successfully disguises?themselves to the fixed community. It is clear to see that few community detection algorithms are able to tell the transitional community from the fixed community if only one snapshot of the temporal network is given. Fig. 1 shows the temporal network with three snapshots.

In the rest of this section, we consider 2 scenarios representing different patterns that the network may evolve:

Scenario S1: We consider three consecutive snapshots where the transitional community jump to the three fixed communities. In this case, we represent snapshots of the network by an ordered set: layer1, layer2, layer3. This representation means that the network evolves from layer1 to layer2 and finally to layer3 where layer2 is the layer to be investigated. Our purpose is to see whether our method can discover the transitional communities hidden in the fixed community in layer2 with the help of the layer1 and layer3. In the rest of the paper, we adopt the same representation. With the prior information that there exist four communities, we run the agglomerative clustering process until it gets four clusters. The result in Fig. 2 shows the capability of our method correctly detect the transitional community and the other fixed communities.

Scenarios S2: We consider the case when the nodes of the traditional community jump to a fixed community and become fixed members of that community. This means once the node jumps to a community, it does not jump to another community for a relative long period. The evolution of the network is represented by layer1, layer2, layer2. Therefore the number of communities in the network is in fact reduced to three. We set the number of clusters to 3 and 4 respectively when running the agglomerative clustering process. Fig. 3 shows that the transitional community cannot be detected when the number of clusters is set to 4 and the detected forth community containing 2 nodes does not turn out to be meaningful. When the number of clusters is set to 3, the three community can be correctly detected, the same as scenarios S1 in Fig.2.

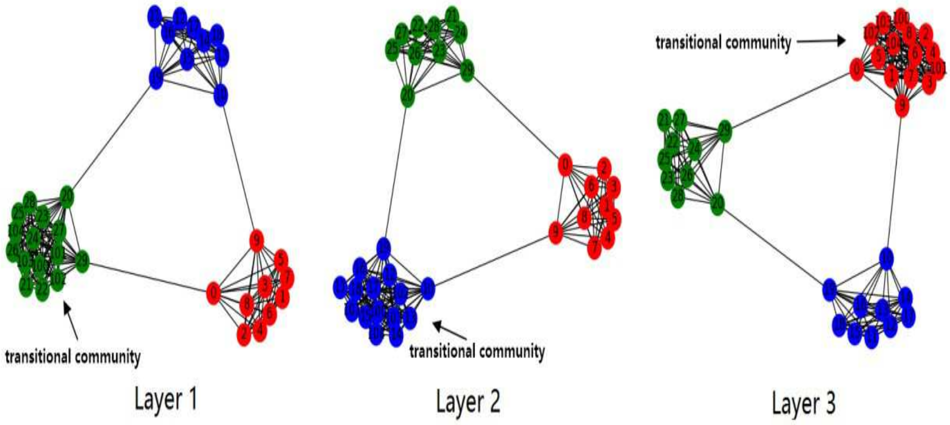

FIGURE 1. Three snapshots of a temporal network, where three communities in each snapshot are distinguished by their color.

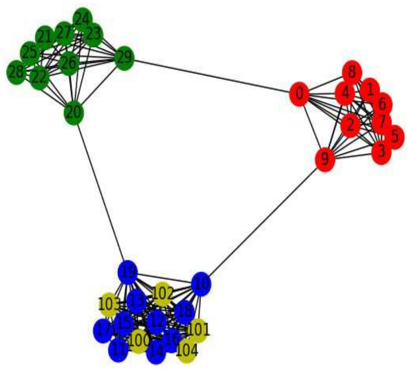

FIGURE 2. All four communities are correctly detected by STRW (The transitional community is indicated by the yellow nodes.

3.2. **Real-Life Dataset.** Here we use a high-resolution dataset that records the close-range social interactions of children in a primary school. The data were collected by the SocioPatterns collaboration (http://www.sociopatterns.org) using wearable sensors that sense the face-to-face proximity relations of individuals wearing them. The population of the school consisted of 231 children organized in 10 classes and 10 teachers. The data were collected over two consecutive days.

We first retrieve the temporal network data within an hour and evenly divide it into 3 parts: the first 20 minutes, the middle 20 minutes and the last 20 minutes and aggregate the temporal network for each 20-minutes interval. The state of a network during one interval is represented by an adjacency matrix where the binary-valued entry indicates the presence of a certain link. The temporal network can thus be represented as 3 successive adjacency matrices combined into a 3-order tensor. We then sample the dataset every half hour and finally obtain 33 3-order tensors. It is notable that aggregate temporal networks are usually comprised of many small disconnected components and most detection algorithms, including ours, only work in the connected network, so we only run the agglomerative clustering process in the largest components of the middle layer.

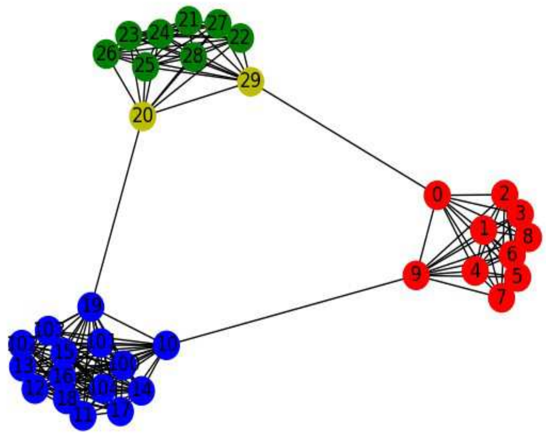

Figure 3. The four communities detected by STRW(The transitional community cannot be detected and the forth community indicated by the yellow nodes is not meaningful).

Table 1. The number of pure cluster detected by 3 method

Algorithms	**STRW**	**RANDOM-WALK**	**LOUVAIN**
`Number of Detected Pure Clusters`	52	41	43

Here we introduce the method to measure the detection performance. We say that a cluster is pure if the classes of the students in this cluster do not appear in other clusters. For example, if all the students in a cluster $C1$ are from class $1A$ and class $1B$ and no student in other clusters is from $1A$ and $1B$, then the cluster $C1$ is pure because it finds all the members of the two classes $1A$ and $1B$ in the largest component. The number of the pure clusters indicate the ability of the detection algorithm to discover the ground truth, so it is a reasonable quantity to measure the performance of community detection. We choose the two state-of-arts algorithms, the Louvain algorithm and the random walk algorithm, as the baseline algorithms. The Louvain Method for community detection is a method to extract communities from large networks. The method is a greedy optimization method for optimization of Modularity. Modularity is a scale value between -1 and 1 that measures the density of edges inside communities to edges outside communities. Optimizing this value results in the best possible grouping of the nodes of a given network. The Louvain algorithm first finds small communities by optimizing modularity locally on all nodes, then each small community is grouped into one node and the first step is repeated. The random walk method considers a discrete random walk process on a graph. At each time step a walker is on a node and moves to a node chosen randomly and uniformly among its neighbours. The sequence of visited nodes is a Markov chain which defines the transition matrix of random walk processes. The authors further utilize the transition probability to get the distance between nodes. A hierarchical clustering process (merging iteratively the nodes into communities) is followed to discover the communities.

Table 1 and Fig.4 show the number of pure clusters detected by the three algorithms. As can be seen in Table 1, STRW is obviously superior to the two baseline algorithms. The original random walk and Louvain algorithms obtain similar performance, detecting 41

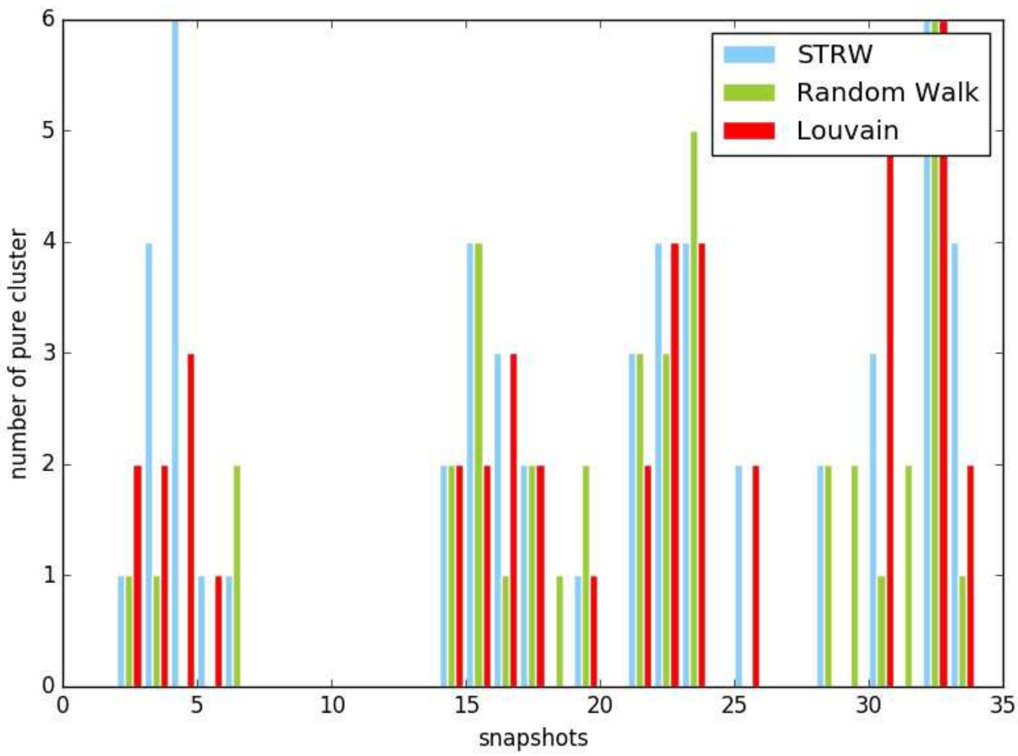

FIGURE 4. Comparison of the number of pure clusters detected by the 3 methods

and 43 pure clusters respectively while STRW discovers 53 pure clusters, about 25% more than the two baseline methods. It is also notable that STRW is specially well-performed in the first 6 snapshots. We find that the first 6 snapshots correspond to the morning class during which little extra-class activity exists. This reduces variation among snapshots. We think the reasons why STRW outperforms the baseline methods are: 1)STRW utilizes not only the current network but also its neighbour networks, which smooths the noise in real world dataset; 2)many real world networks evolve quite slowly so the neighbour network topology could acts as redundant information to correct the community detection error.

4. **Conclusions.** This paper proposes an algorithm named STRW to enhance the accuracy of community detection with the aid of adjacent snapshots in a temporal network. The algorithm takes advantage of the temporal correlation among a temporal network, and adapts the transition probabilities of the random walk to depend on the local topological similarity between adjacent snapshots. Experimental results show that by utilizing the correlation among a temporal network, our algorithm STRW is superior to the state-of-art methods.

Acknowledgements. This work was supported by the Zhejiang Provincial Natural Science Foundation of China under grant No.LY17F030008.

REFERENCES

[1] Z. Kuncheva, and G Montana, Community detection in multiplex networks using Locally Adaptive Random walks, *International Conference on Advances in Social Networks Analysis and Mining*, pp.1308-1315, 2015.

[2] S. H. Strogatz, Exploring complex networks, *Journal of Nature*, vol.410, no.6835, pp.268-276, Mar. 2001.

[3] S. Fortunato, Community detection in graphs, *Journal of Physics Reports,* vol. 486, no.3, pp.75-174, Feb. 2010.

[4] L. Lovasz, Random walks on graphs: A survey, *Journal of Combinatorics*, vol. 2, pp. 1-46, 1993.

[5] M. Hmimida and R. Kanawati, Community detection in multiplex networks: A seed-centric approach, Networks and Heterogeneous Media, vol. 10, no. 1, pp. 71-85, 2015.

[6] H. T. Nguyen, T. N. Dinh, T. Vu, Community detection in multiplex social networks, C*omputer Communications Workshops IEEE*, PP. 654-659, 2015.

[7] P. Holme, and J. Sarama, Temporal networks, *Journal of Physics Reports*, vol.519, no.3, pp.97-125, 2012.

[8] P. Pons, and M. Latapy, Computing communities in large networks using random walks, Computer and Information Sciences, vol. 10, no. 2, pp. 191-218, 2006.

[9] T. Y. Berger-Wolf, C. Tantipathananandh, and D. Kempe, Dynamic community identification, Link Mining: *Models, Algorithms, and Applications, Springer*, pp. 307-3036, 2010.

[10] J. Stehle, N. Voirin, A. Barrat, C. Cattuto, L. Isella, et al., High-resolution measurements of face-to-face contact patterns in a primary school, *PloS one, e23176,* 2011.