Anomaly analysis and visualization for dynamic networks through spatiotemporal graph segmentations

Qi Liao, Ting Li, Benjamin A. Blakely

PII: S1084-8045(18)30303-5

DOI: 10.1016/j.jnca.2018.09.016

Reference: YJNCA 2216

To appear in: *Journal of Network and Computer Applications*

Received Date: 23 March 2018

Revised Date: 17 August 2018

Accepted Date: 27 September 2018

Please cite this article as: Liao, Q., Li, T., Blakely, B.A., Anomaly analysis and visualization for dynamic networks through spatiotemporal graph segmentations, *Journal of Network and Computer Applications* (2018), doi: https://doi.org/10.1016/j.jnca.2018.09.016.

Anomaly Analysis and Visualization for Dynamic Networks through Spatiotemporal Graph Segmentations

Qi Liao, Ting Li, and Benjamin A. Blakely

***Abstract*—With recent technology advance in Internet of Things (IoT) that involves human, sensors, and mobile devices, networks are not only growing much larger but more complex and dynamic in nature. The spatiotemporal dynamics of networks are represented by both topological changes and temporal shifts of attribute information associated with network components. Understanding the pattern and trend of dynamic networks is increasingly important. While data mining approaches are generally useful in analyzing the statistical properties of networks, there has been recent trend to consider bringing human into the loop, and to examine how feedback from visualizations of large-scale dynamic networks can further improve data mining and machine learning. Traditional visualization methods based on animation and sequences of snapshot graphs are also limited by human cognitive capability. We present a dynamic network analysis and visualization (DNAV) tool which explores the spatiotemporal dimensions of graph components. In particular, nodes and edges are augmented with spatiotemporal segmentation based on both topological and attribute dynamics (e.g., time and locations of connectivity). To further facilitate analysis of large dynamic networks, DNAV includes statistical dynamic overviews alongside graph views, as well as data filtering modules for scalable analysis. Using case studies on public datasets, we demonstrate the effectiveness of DNAV in understanding and analyzing anomalies in dynamic networks such as computer communication networks.**

***Index Terms*—Dynamic networks, anomaly detection, visualization, graph segmentation, spatiotemporal analysis, analytic tool and applications**

I. Introduction

Communication networks formed by recent technological advance in Internet of Things (IoT) are not only becoming much larger in size but more complex and dynamic in nature [1]. Consequently, the large amount of network traffic log generated by the movement of people and mobile devices are difficult to analyze. Dynamic network analysis [2] is an emergent scientific field to address the challenges and has profound implications for a number of important tasks. For example, a network operator may need to detect anomalies in information technology network for potential faults and security breaches. A military commander may need to gain situational awareness, and understand the relationships in terrorists networks. Advertisers may need to target high-impact users based upon user spatiotemporal actions in social networks. Or, cancer researchers may need to understand interactions in protein networks and detect anomalous tumor cells.

Dynamic network analysis is highly challenging due to spatiotemporal network dynamics in terms of both topological structure and attribute evolution. Dynamic networks are known for their temporally changing topology as a result of on/off patterns of network connectivity. For communication networks, for example, network connections may be established or removed at any moment, making static analysis of a moving target more difficult. Often, dynamic networks are also encoded in higher dimensional space. For example, edges may contain properties such as link quality metrics (e.g., loss rate, latency, or bandwidth) and the locations where the connections are made. Similarly, nodes may contain attributes such as user profile or computing status (e.g., CPU, memory, or disk utilizations). Determining the most effective manner to analyze the many complex relationships between these attributes, especially in a temporally dynamic context, presents significant challenges.

While classic statistics and graph theories are useful in analyzing static networks, such methodologies may become insufficient with large dynamic networks. There has been research [3] to understand and detect anomalies in dynamic networks using methods such as clustering [4] and link prediction [5], [6]. However, with rapid growth of dynamic network size, volume, dimension and complexity, data mining approaches *alone* become inadequate. This is due to inherent computational infeasibility when considering all possible prediction outcomes. Alternatively, visualization techniques [7] have been proposed to explore dynamic networks for patterns and potential abnormal behaviors. Among them, small multiples and animation [8], [9] are two typical visualization methods for dynamic network analysis. However, it is still difficult to track changes over time due to limitations of human cognitive ability. Other dynamic network visualizations that explore three-dimensional (3D) space [10], [11] or utilize superimposition and layer juxtaposition [12]–[14] have been

Q. Liao (corresponding author) is with the Department of Computer Science, Central Michigan University, USA. Email: liao1q@cmich.edu.

T. Li was with the Department of Computer Science, Central Michigan University, USA. Email: tingli0919@gmail.com.

B. Blakely is with the Strategic Security Sciences Division, Argonne National Laboratory, USA. Email: bblakely@anl.gov. The work presented in this paper was partially supported by the U.S. Department of Energy, Office of Science under DOE contract number DE-AC02-06CH11357.