ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/385834125

Visual Analysis of Complex Temporal Networks Supported by Analytic Provenance

Conference Paper · October 2024

DOI: 10.1109/VASTChallenge64683.2024.00016

CITATIONS
0

READS
33

6 authors, including:

Yuhan Guo
Peking University

7 PUBLICATIONS **36** CITATIONS

SEE PROFILE

All content following this page was uploaded by Yuhan Guo on 25 January 2025.

The user has requested enhancement of the downloaded file.

Visual Analysis of Complex Temporal Networks Supported by Analytic Provenance

Yuhan Guo* — Peking University
Yuchu Luo† — Peking University
Xinyue Chen‡ — Peking University
Hanning Shao§ — Peking University
Xiaoru Yuan¶ — Peking University
Kai Xu‖ — University of Nottingham

Figure 1: The system we developed for MC3. The main part is a graph view (c) for exploring the structure of the network and a temporal view (d) for exploring the events with temporal information at different levels of details. The filter widgets (a) and the list view (b) enable flexible filtering and selection. At the bottom is a provenance map (e) that captures and visualizes the history of user interactions and system states.

Abstract

We present an interactive visual analysis tool to explore large dynamic graphs. Our system provides users with multiple perspectives to analyze the network. The graph view presents the node-link structure and offers various layout options. To complement, a temporal view shows both the overall temporal distribution and detailed event timelines. The system also supports flexible filtering to reduce the graph size and identify interesting entities. One bonus feature of our system is the provenance map, which visualizes the automatically captured user interactions and allows users to record their findings. The provenance map is helpful for organizing the exploration process and synthesizing analysis results.

1 Introduction

The 2024 VAST Challenge Mini-Challenge 3 requires temporal analysis of a large knowledge graph extracted from business activities. In the graph, the nodes are entities, i.e., people and organizations, and the edges are relationships between the entities, mostly ownership relationships with start dates and possibly end dates.

The complexity in the MC3 data, i.e., the dynamics of the network structure, along with the large scale (over 60k nodes and 70k edges), makes it difficult to understand the events and patterns in the graph. To tackle the challenge, we designed a visual analysis system with the following features: (1) flexible filtering to narrow down the exploration, (2) two coordinated views that emphasize complementary perspectives, namely graph structure and temporal information, to support analysis of specific behaviors, and (3) provenance capturing support that enhances the analytic process.

The analysis of complex graph often starts with exploration of sub-graphs or local neighborhoods, from which the analysts examine cases and formulate hypotheses on data patterns. Such an exploratory process might involve a lot of analysis options, such as different filters to apply and different entities to inspect. To facilitate the exploration, our system leverages analytic provenance, specifically the automatically captured user interactions and manually recorded notes (findings or unexplored options).

The key insights we found when addressing the challenge are that, firstly it is useful to offer flexible filtering and layout options when analyzing complex graphs; secondly analytic provenance is helpful for structuring exploration, synthesizing findings, and communicating results.

An online demo of our system is available at https://vis4sense.github.io/vast-2024-mc3/.

*e-mail: yuhan.guo@pku.edu.cn
†e-mail: luoyuchu@pku.edu.cn
‡e-mail: cxyapril@stu.pku.edu.cn
§e-mail: hanning.shao@pku.edu.cn
¶e-mail: xiaoru.yuan@pku.edu.cn
‖e-mail: kai.xu@nottingham.ac.uk

2 Design of the Main Components

Figure 1 shows the interface of our system. The graph is decomposed into connected components, which are shown at the top of the list view (Figure 1b). There is also a list of component thumbnails at the bottom of the graph view (Figure 1c) to help user find interested components. The connected components are frequently used for filtering nodes and usually the analysis is focused on a selected component. While the size of the components is much smaller compared to the whole graph, it is still challenging to explore the biggest component (with over 20k nodes). A set of scented widgets (Figure 1a) further support filtering nodes and edges by various attributes (e.g., type, date, and country). After narrowing down the focus to a small set of nodes, the user can use the graph view (Figure 1c) and the temporal view (Figure 1d) to select entities of interest and analyze their relationships in detail. The provenance map (Figure 1e) helps users orient themselves in the exploration and capture analytic thoughts. The following introduces the design of the main views.

2.1 Graph View

The graph view uses node-link representation to show the structure of the network. The nodes are encoded as icons that indicate the type of the entity and colored by the country. The color of links represent the type of the relationship. The most common relationships are shareholdership (colored red) and beneficial ownership (colored blue). A dash line is used if the relationship has an end date. Not all the nodes in the network are shown initially if the network is too large. By default, the system selects nodes according to the current filtering. The user can add more nodes by clicking the “add” button at the bottom right of the node to expand its neighborhood.

We offer several layout options so that users can select the one that best presents the selected sub-graph. *Force-directed* layout is suitable for most cases as there are a lot of chains in the network. *DAG* layout works for complex components with hierarchical ownership structures. We also provide options for reducing and aggregating nodes. For example, the *company only* mode shows only company nodes and derived links between closely connected companies. The *aggregated* mode groups nodes that have the same neighbors. The *hide leaf* mode hides the leaf person nodes which are less interesting. The *focus* mode hides the nodes that are irrelevant to the current selection.

2.2 Temporal View

The temporal view shows the temporal information at different levels of details. At the top is a chart showing the overall distribution of the events (Figure 1-d1). Users can choose to view all events or events in the selected connected component. At the bottom is the entity-level activity distribution (Figure 1-d3). In the activity matrix each row represents an entity, and the events are represented as pixels. The activity matrix provides a summary of the entity’s behaviors and is helpful for detecting anomalies.

The timeline (Figure 1-d2) adds another level of detail by displaying links between entities at the time of the event. Only selected nodes are shown in the timeline. A curved arrow is shown when both the source and target of the event are selected. Otherwise, the event is represented as a square (organization) or circle (person) with a small arrow above it, which can be added to selection on demand. The encoding is consistent with the graph view.

2.3 Provenance Map

The provenance map is an application of HistoryMap [1], which was initially designed for tracking browser history and collecting relevant information, to visual analytics. We model the history of system states as a hierarchy, i.e., a node is the child of its previous state. A branch is created when the user goes back to a previous state and explores a different path.

In our system the user interactions and resulting system states are automatically captured and visualized as nodes in the provenance map. When there is a finding, the user can manually attach screenshots and notes to the state that leads to it.

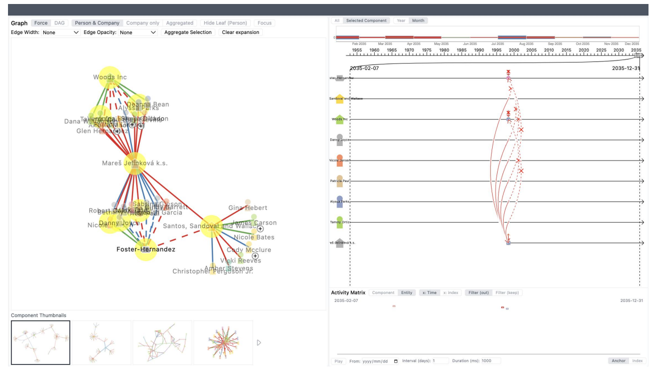

Figure 2: *Foster-Hernandez* (a fishing company) and *Woods Inc* were merged into *Mareš Jelínková k.s.* (a seafood provider).

3 Case: Discovering Suspicious Transactions

In this section, we demonstrate the use of our system by presenting a discovery process to identify suspicious transactions. Inspired by the changes associated with the illegal behavior of SouthSeafood Express Corp (Figure 1), i.e., reorganization of company ownership structures, we assume that illegal transactions might involve endings of relationships. We started by filtering events around the time of the illegal behavior (May 2035), events with ending dates, and fishing companies. With the filters, several entities were highlighted on the activity matrix. The state was marked as a branching point and the options were selecting different entities of interest. For example, Figure 2 shows the result of selecting the first entity (the fishing company *Foster-Hernandez*) and relevant nodes. The spindle-shaped node-link diagram indicated a merger of *Foster-Hernandez* and *Woods Inc* to *Mareš Jelínková k.s.*, which is a seafood provider. We assume this might be an attempt to escape punishment for illegal fishing.

4 Reflection and Discussion

The lessons we learned from the design of the system include:

- Breaking down the graph into smaller components is an effective way to start exploration of large networks. This can be achieved by decomposing the graph into connected components, filtering by various attributes, etc.
- Providing rich graph layout options helps scale the analysis to address different sub-graph structures and user tasks.
- The collection of analytic provenance can reduce cognitive load in planning the exploration and reasoning.

While the provenance map facilitates the generation of hypotheses, it is less helpful for finding evidence for validation. Also, it can be tedious to manually explore the analysis options at each branching point. We believe that the system can be further improved by integrating automated methods and learning from the collected provenance to provide more active support for analysts.

References

[1] P. H. Nguyen, K. Xu, A. Bardill, B. Salman, K. Herd, and B. W. Wong. Sensemap: Supporting browser-based online sensemaking through analytic provenance. In *IEEE Conference on Visual Analytics Science and Technology (VAST)*, 2016. doi: 10.1109/vast.2016.7883515 2

View publication stats