Graph Analytics and Visualization for Criminal Network Identification

LEAD PI: DIRK REINERS

Criminal Network Analysis

SUMMARY

Investigators have an enormous amount of data and information sources to explore and integrate in order to uncover and understand complex criminal networks, and they need a visual representation of the connections between people and events to perform their work. This project will develop new methods for graph display and layouts targeted at the kind of data typically encountered by investigators and analysts, allowing them to focus on identifying criminal activity and networks rather than data processing.

PROBLEM STATEMENT

Identification of criminal networks is a highly challenging task. A wide variety of evidence need to be collected and correlated in order to form and prove hypotheses that stand up under judicial scrutiny. Current systems have an abundance of data, so the challenge has evolved from not having the information needed to finding it inside the vast array of data that is available. A wide variety of analytical tools are being developed in order to automatically find valuable knowledge, but in some cases automated extraction may not be sufficient and human ingenuity and intuition will be required to find the obscure connection that breaks a case. This work addresses the challenge of finding efficient ways to feed as much information as possible to the human analysts in order to trigger their natural and unique abilities.

APPROACH

The goal of this research is to design and develop new methods for graph display and graph layout that are targeted at the kind of data typically encountered in DHS and LE situations. The first phase of this work is focused on developing a new layout algorithm for a 3-layer (events/people/attributes) dataset such as currently used by investigators. Such datasets are often shallow but wide, and as an added complication a large number of other events are typically linked, leading to a much larger number and more complex distribution of connecting edges.

Finding a layout that at the same time shows the data and connections clearly without a large number of overlapping and crossing edges is a non-trivial problem, and this work is focused on finding effective solutions for it. Additional work is addressing annotation challenges and presentation to varied audiences, including judges and juries.

ANTICIPATED IMPACT FOR DHS

This project will provide algorithms and code to enable effective graphical data representations appropriate for DHS and LE investigations, and designed to maximize the unique cognitive abilities of human analysts. Investigators will be able to more quickly identify criminal network structures and related data, facilitating more effective, efficient, and complete investigations.