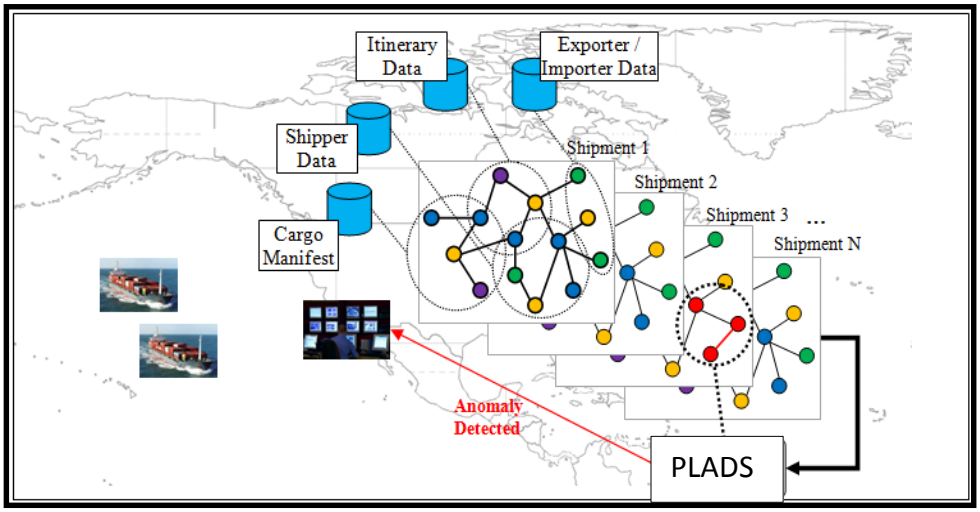
Pattern Learning and Anomaly Detection in Streams (PLADS)



Need

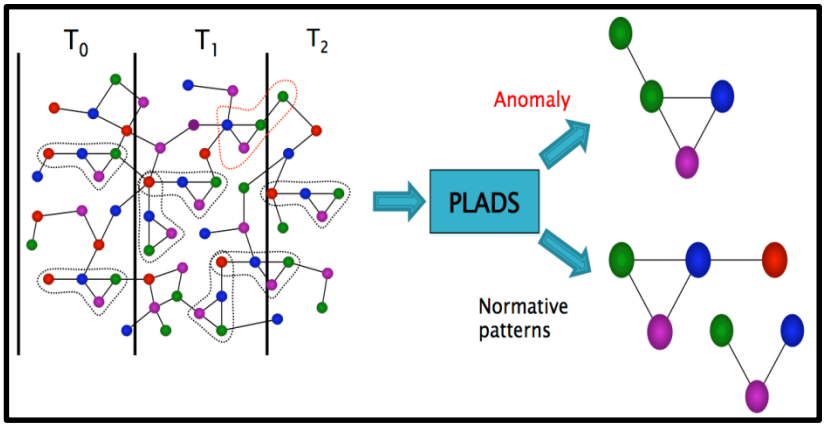
Protecting our nation's cyber infrastructure and securing sensitive information are critical challenges for both industry and government. Due to the plethora of information that is transmitted daily, attacks can pose potentially serious consequences to individuals, corporations, governments, and society as a whole. In order to address this issue, one must provide methods of monitoring and rapidly detecting security compromises. However, due to the *complex* and *diverse* nature of the environments which are vulnerable to cyber-crime or cyber-terrorism activity, one must not only be able to deal with attacks that are *dynamic*, or constantly changing, but also take into account the *structural* aspects of the targeted networks and the relationships among communication events.

Approach

Pattern Learning and Anomaly Detection in Streams (or PLADS) is a novel approach that focuses on *streaming data* that represents the **relationships** and **transactions** between entities (people, organizations, etc.) by discovering and analyzing a **graph structure** of their social and behavioral network. Through the analysis of the structure surrounding actions, such as can be found in communications or data access, our approach first searches for the normative patterns of activity, then discovers those small, unusual deviations and presents the patterns and anomalies to an analyst. This approach uses novel algorithms that can be applied to the *dynamic* discovery of anomalies in a variety of areas, including the detection of insider threats, terrorist activity, fraud, and cybercrime.

Benefits

PLADS is an unsupervised approach (no training needed) that can provide an automated, behind-the-scenes analysis of existing data that can be streaming in real-time. By understanding the benefits of structure for detecting possible threats, early detection can be accomplished that will minimize the damage, potentially protecting valuable IP, reclaiming lost revenue and resources, and maintaining trust. Organizations need better, automated tools and methods to improve detection ability that will combat serious threats such as the leaking of sensitive information.



Existing Approaches

Most existing approaches deal with the numerical and statistical attributes of their data sources. Behavioral profiles are generated based on simple attributes where training data is needed. In addition, existing graph-based approaches are unable to handle large, heterogeneous data sets. Our approach proposes the handling of "big data" through a streaming graph approach that can discover interesting **structural** patterns and anomalies.

Key Personnel

Dr. Larry Holder (holder@wsu.edu) and Dr. Bill Eberle (weberle@tntech.edu) have worked together over the last twelve years developing new techniques for analyzing graph-based representations of data in order to detect structural anomalies. Working on multiple DoD, DHS, and NSF projects, they have analyzed various domains, such as TCP/IP networks, business processes, social networks, and cargo shipments, for both normal and unusual patterns.