Anomaly Detection Methods and Applications

Organizers: *Robert A. Bridges* (bridgesra@ornl.gov),
Oak Ridge National Laboratory
Blair D. Sullivan (sullivanb@ornl.gov), Oak Ridge National Laboratory

February 15, 2013

Abstract

Anomaly detection enjoys a myriad of useful applications including fraud and intrusion detection, image processing, and quality control to name a few; consequently, the methods for anomaly detection have equally wide scope with techniques combining elements from statistics, machine learning, probability, graph theory, and other subjects, often tailored to a specific application domain. This mini-symposium seeks to foster collaboration of these varied fields, application domains, and researchers through presentations on current advances in the subject.

Speakers:

- **Speaker**: *Varun Chandola, Ph.D.* (chandolav@ornl.gov), Oak Ridge National Laboratory, Computer Science and Engineering Division

  Title: *Anomaly Detection: From Data to Decisions*

  Abstract: Anomalies and scientific discovery go hand in hand. In his book, "The Structure of Scientific Revolutions", Thomas Kuhn states, "Discovery commences with the awareness of anomaly ..." The scientific argument applies to social and engineering domains, where often identifying and interpreting anomalies is the objective of the discovery. In the present data rich world, detecting anomalies from data has unprecedented application and value. Current landscape of anomaly detection is dominated by methods that are closely tied to the end application domain. In this talk I will present an overview of the existing anomaly detection research from an algorithmic and data perspective. I will also discuss the challenges within key application domains and the applicability of existing methods. Finally, I will discuss some of my experiences with anomaly detection in two important domains, cyber-security and environmental monitoring.

- **Speaker**: *William Eberle, Ph.D.* (WEberle@tntech.edu), Tennessee Technological University, Department of Computer Science

  Title: *Graph Based Anomaly Detection*

  Abstract: Traditional methods for discovering anomalies in data consist of "machine learning" approaches such as classification, clustering, nearest neighbors or other statistical techniques. However, data that represents actions or relationships between people, such as whom they know, or whom they call, can be difficult to analyze. The advantage of graph-based anomaly detection is that relationships between elements can be analyzed for structural oddities that could represent activities such as fraud, network intrusion, or suspicious associations in a social network.

- **Speaker**: *Danfeng Yao, Ph.D.* (danfeng@cs.vt.edu), Virginia Polytechnic Institute and State University, Department of Computer Science

  Title: *User-Intention Based Anomaly Detection For Cyber Security*

  Abstract: The proliferation and sophistication of malware (malicious software) activities – as well as their growing capacity to do serious harm – requires constant vigilance and upgrading. We aim to develop anomaly detection solutions that can be applied to identify suspicious network and file system activities. Specifically, we focus on identifying characteristic human-user behaviors (namely

application-level user inputs via keyboard and mouse), developing protocols for analyzing inputs and system calls, and preventing forgeries and attacks by malware. We present several projects based on this human-behavior driven malware detection approach, including drive-by-download detection, HTTP-based input-traffic analysis, and cryptographic-based traffic provenance verification.

- **Speaker**: *Polo Chau, Ph.D.* (polo@gatech.edu), Georgia Institute of Technology, School of Computational Science and Engineering

  Title: *Mining Massive Graphs: Visualization & Anomaly Detection*

  Abstract: Given a large graph, how to help people make sense of it? How to spot anomalies? I will present several works to answer these questions. Apolo, a system that guides the user to explore large graphs via a mixed-initiative approach combining visualization and machine learning. NetProbe, which detects auction fraud by identifying the collaboration networks among fraudsters. Polonium, a patented technology that uses network effects to unearth malware from 37 billion machine-file relationships.