SAND2012-6142P

Exceptional service in the national interest

Sandia National Laboratories

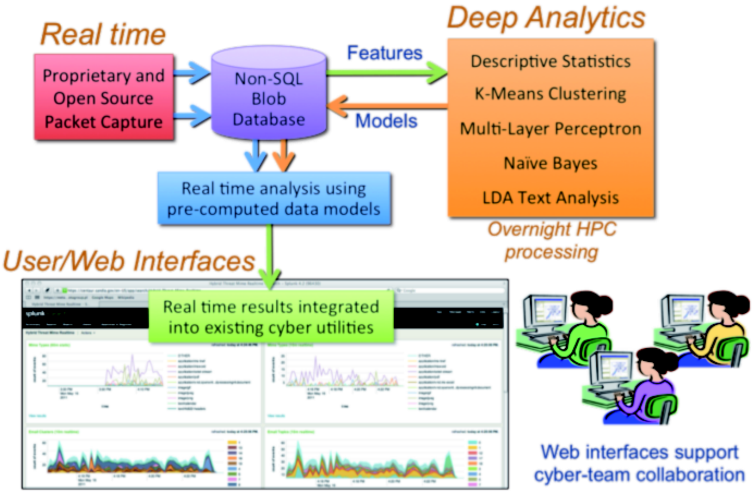
Cyber Security LDRD Enables New Level of Collaboration between Researchers and Front-Line Guardians for Sandia's Network Defense

Challenge

Despite continual efforts in this direction, novel emerging threats to Sandia's network cyber security makes it essential that defenders continue to find new protective solutions. *The Hybrid Methods for Cybersecurity Analysis* LDRD project engages an interdisciplinary team from divisions 1000, 5000, and 9000 focused on the mission-critical goal of defending Sandia against malicious network attacks. With a lab-wide interdisciplinary team including researchers, operators, and analysts, the project concentrates its efforts on defending Sandia against computer network attacks.

Research

The innovative "hybrid" approach of this project — augmenting real time analysis with offline computation — provides a fundamentally new capability to cyber analysts. Computationally expensive models that may take hours or days to compute are applied in real time to incoming network streams. The application of machine learning algorithms, text analysis, and statistics to real-time network data enables unprecedented insight into some of the toughest problems in cyber security. The team assembled for this project includes Brian Wylie (PI, 1461), Danny Dunlavy (1464), Warren Davis (1461), Eunsil Han (9317), Christopher Nebergall (9312), Roger Suppona (9317), and Justin Doak (5635). According to Suppona, "The value of teaming world-class researchers with operational cyber security professionals to solve difficult security problems is evidenced by *Hybrid Methods*. Participating in this project and watching *Hybrid Methods* develop into an essential component of our security toolkit is a terrific experience."


Real time
Proprietary and Open Source Packet Capture
Non-SQL Blob Database
Features
Models
Deep Analytics
Descriptive Statistics
K-Means Clustering
Multi-Layer Perceptron
Naive Bayes
LDA Text Analysis
Overnight HPC processing
Real time analysis using pre-computed data models
User/Web Interfaces
Real time results integrated into existing cyber utilities
Web interfaces support cyber-team collaboration

Significance

Already in prototype deployment, the technique is being applied to real-time detection of email phishing attacks by modeling both metadata and textual content of the emails. These models are computed on large segments of collected emails and known attacks. They are then applied to the real-time network stream where emails are given a composite "threat score" based on their evaluations against the analytic models. This hybrid methods approach enables the use of computationally expensive models in a real-time environment where response times can be critical. Even at this early stage of

U.S. DEPARTMENT OF ENERGY

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND 2012-xxxxP.

development, the project is successfully identifying attacks missed by the large collection of existing defenses. Although the project is continuing to refine its methods and is expected to manifest additional progress in its ability to rapidly detect attacks, this new approach greatly eases the delivery of advanced analytic methods into the hands of Sandia's front line defenders, paving the way for future collaboration between key organizations involved in cyber security.

Point of Contact: Brian Wylie bnwylie@sandia.gov

Funding: LDRD