



# Cyber Security Division Technology Guide

Volume 1



**Homeland  
Security**

Science and Technology

This page left blank intentionally.

# Introduction

Thank you for your interest in the U.S. Department of Homeland Security (DHS) Science and Technology Directorate's (S&T) Cyber Security Division (CSD) Research and Development (R&D) portfolio. This Technology Guide is the culmination of extensive efforts to identify and develop cybersecurity technologies for homeland security application within industry, academia, and our national lab partners. Many of these technologies have been funded through Broad Agency Announcements (BAA) and Small Business Innovative Research (SBIR) programs. We're excited to share these promising cybersecurity technologies with you.

Our Technology Guide, which is updated and published annually, features innovative R&D Cyber Security Forensics, Cybersecurity Competitions, Cybersecurity Incident Response Teams, Insider Threat, Internet Measurement and Attack Modeling, Mobile Device Security, Moving Target Defense, Security of Cloud-Based Systems, Software Assurance Marketplace, and Software Quality Assurance.

Through partnerships and commercialization CSD is identifying innovative, federally-funded research that addresses cybersecurity needs and is helping transition our developed tools and technologies into the Homeland Security Enterprise. All of the technologies included here are mature and ready to be piloted in an operational environment, or transitioned into a commercially available product. If you're interested in becoming an S&T partner, piloting, licensing, or commercializing one of our technologies please be sure to [connect with us](#).

In addition to these technologies, we are interested in future research areas that solve cybersecurity capability gaps in your own organizations. We encourage you to share your thoughts with us and your input will help us identify real-world solutions and inform future research efforts. Again, it's our pleasure to introduce you to our CSD Technology Guide and these potentially groundbreaking cybersecurity tools from the federal government R&D community.

Sincerely,



**Douglas Maughan**  
DHS S&T Cyber Security Division  
Director





# CONTENTS

<b>CYBER SECURITY DIVISION</b>	<b>2</b>
<b>CYBER SECURITY FORENSICS</b>	<b>4</b>
Autopsy: Enabling Law Enforcement with Open Source Software	5
<b>CYBER SECURITY COMPETITIONS</b>	<b>7</b>
Web-based Interactive Stories for [Cyber] Education (CyberWISE)	8
<b>CYBER SECURITY INCIDENT RESPONSE TEAMS</b>	<b>9</b>
Improving CSIRT Skills, Dynamics and Effectiveness	10
<b>INSIDER THREAT</b>	<b>11</b>
Monitoring Database Management System (DBMS) Activity for Detecting Data Exfiltration by Insiders	12
<b>INTERNET MEASUREMENT AND ATTACK MODELING</b>	<b>14</b>
Alembic: Machine-Based Detection of Domain Ownership Change	15
Cartographic Capabilities for Critical Cyber Infrastructure	17
Clique: Safeguarding Cyber Systems with Visualization	19
Enhancing Watchdog System for Internet Routing (WIT)	21
Netalyzr: Measuring the Network from the Edge	22
Stucco: A Cyber Intelligence Platform	24
Symbiote® Defense	26
<b>MOBILE DEVICE SECURITY</b>	<b>28</b>
MobileRoT – Software-only Roots-of-Trust for Mobile Devices	29
<b>MOVING TARGET DEFENSE</b>	<b>31</b>
Self-shielding Dynamic Network Architecture (SDNA)	32
<b>SECURITY OF CLOUD-BASED SYSTEMS</b>	<b>33</b>
Silverline: Assessment System for Secure Cloud Computing	34
<b>SOFTWARE ASSURANCE MARKETPLACE</b>	<b>36</b>
The Software Assurance Marketplace (SWAMP)	37
<b>SOFTWARE QUALITY ASSURANCE</b>	<b>39</b>
Code Ray: Better software vulnerability management through hybrid application security testing	40
CodeHawk Automated Malware Analyzer	42
Hybrid Analysis Mapping: Software Assurance Enhancement Technology	44
Tunable Information Flow (TIF): Policy-Driven Software Analysis & Assurance Toolset	46



# Department of Homeland Security (DHS) Science and Technology Directorate (S&T) Cyber Security Division (CSD)

## The Cyber Security Division (CSD) is a Key Component in the President's National Strategy

Threats on the Internet change fast and cybersecurity is one of the most challenging areas in which the Federal government must keep pace. Next-generation cybersecurity technologies are needed to enhance the security and resilience of the nation's current and future critical infrastructure and the Internet.

In the Department of Homeland Security (DHS) Science & Technology Directorate (S&T), the CSD enables and supports research, development, testing, evaluation, and transition for advanced technologies in cybersecurity and information assurance. This full lifecycle of activities evolved in response to the President's National Strategy to Secure Cyberspace and the Comprehensive National Cybersecurity Initiative (CNCI).



The CNCI establishes a multi-pronged approach the Federal government will take in identifying current and emerging cyber threats, shoring up current and future vulnerabilities in telecommunications and cyberspace, and responding to or proactively stopping entities that wish to steal or manipulate protected data on secure Federal systems.

The S&T Cyber Security Division addresses these objectives by:

- Discovering new solutions for emerging cybersecurity threats to the nation's critical infrastructure;
- Driving security improvements to close critical weaknesses in today's technologies and emerging systems; and
- Delivering new, tested technologies to defend against cybersecurity threats and making them available to all sectors through technology transfer and other methods.

## CSD Focuses on Critical Vulnerabilities in the Cyber Security Landscape

**Internet Infrastructure Security**—Developing security protocols for the existing Internet infrastructure (browsers and routers, essential to daily Internet operation) so that users are not redirected to unsafe websites or pathways by malicious actors.

**Critical Infrastructure/Key Resources**—Securing the information systems that control the country's energy infrastructure including the electrical grid, oil and gas refineries, and pipelines, to reduce vulnerabilities as legacy, standalone systems are networked and brought online.

**National Research Infrastructure**—Providing the infrastructure that enables development and testing of technologies to address cybersecurity issues including botnets, worm propagation and defense, and denial-of-service defenses that protect Internet websites against attack; providing a data repository to support the cybersecurity research community.

**Leap-Ahead Technologies**—Develop "leap-ahead" technologies that will achieve orders-of-magnitude improvements in cybersecurity. One of CNCI's goals is to achieve a reliable, resilient, and trustworthy digital infrastructure.

**Cyber Security Education**—Helping to foster adequate training and education programs critical to the nation's cybersecurity needs by providing opportunities for high

*Our vision is a cyberspace that supports a secure and resilient infrastructure, that enables innovation and prosperity, and that protects privacy and other civil liberties by design. It is one in which we can use cyberspace with confidence to advance our economic interests and maintain national security under all conditions.*

— Quadrennial Homeland Security Review, 2010

school and college students to develop their skills and by giving them access to advanced education and exercises through team competitions.

**Identity Management**—Evaluating and developing proof-of-concept solutions, and conducting pilot experiments of identity and access control architectures and technologies, as well as data privacy protection technologies for the homeland security community.

**Cyber Forensics**—Developing new cyber forensic analysis tools and investigative techniques to help law enforcement officers and forensic examiners address cyber-related crimes.

**Software Assurance**—Developing tools, techniques, and environments to analyze software, address the presence of internal flaws and vulnerabilities in software, and improve software security associated with critical infrastructure (energy, transportation, telecommunications, banking and finance, and other sectors).

## **S&T: Preparing for Next-Generation Cyber Threats**

In the coming years, several cybersecurity challenges must be addressed. The most critical of these include Enterprise-Level Metrics, Combating Insider Threats, Combating Malware and Botnets, Digital Provenance, Situational Understanding and Attack Attribution, and Usable Security.



# CYBER SECURITY FORENSICS:

© Autopsy: Enabling Law Enforcement with Open Source Software





# Autopsy: Enabling Law Enforcement with Open Source Software

**Brian Carrier**

**Megan Mahle  
Cyber Forensics Program Manager**

## Overview

Autopsy is open source digital forensics software that can be used by law enforcement and corporate investigators to determine what a digital device was used for. It has thousands of users around the world and can be used in a variety of investigation types, from fraud to terrorism to child exploitation.

The Department of Homeland Security (DHS) Science and Technology (S&T) funded development that focused on features that law enforcement needed to more quickly perform investigations. The results have been released into the public as features in the Autopsy open source program.

## Customer Need

As digital devices have become an essential part of everyone's lives, they have also become a critical part of nearly every criminal investigation at the local, state, and federal levels. And as the storage capacity of the devices increase, so do the challenges of investigating them and finding the evidence in the midst of the other data on the device.

Meanwhile, not all state and local law enforcement organizations have the budget to purchase commercial tools or have contacts to obtain government-only software. This effort focused on providing analytic features that did not exist in commercial software and providing them as part of free open source software. This enables the investigation community to have access to better tools and have the features maintained by the community.

## Our Approach

Our general approach was to first survey state, local, and federal law enforcement officials to identify their biggest challenges and where they spent the bulk of their time. We identified several areas that we felt could be improved.

We then worked with users to better understand their workflow and behaviors in order to automate that process. These features were then incrementally released into the Autopsy software.

Autopsy was first released in 2000 and had a major rewrite in 2011 to enable it to focus on ease of use and being a platform with many frameworks. Out of the box, it comes with the standard features that a typical investigator would need, but its modularity allows for workflows to be automated so that examinations can be more efficient.

A recent release of Autopsy had nearly 40,000 downloads over the course of five months. This allows the many thousands of investigators to have the benefit of the S&T's efforts. In addition, those users can provide feedback on how they can be improved.

## Benefits

Over the course of this effort, we have focused on three areas: analyzing large numbers of images, timeline analysis, and looking for indicators of compromise. Each has different benefits.

The Image Gallery module allows investigators to review large numbers of images by grouping them by folder, or by some other criteria. Autopsy prioritizes which group of files to display to the user based on how many files in the group are "known bad and on MD5 hash values. The user can categorize the entire group as good or bad without needing to categorize each one.

The Timeline module allows investigators to answer questions about how many and what events occurred in a given time range. It combines events from files and application-level data into a single interface. It clusters events together to help reduce data overload for the investigator.

### Competitive Advantage

Many solutions for investigating cases involving large number of images that do not take disk images as input. You must export the images from a forensics tool and then import them into the picture analysis tool. You may also need to wait for all images to be hashed and analyzed before any can be displayed. Other tools support disk images and allow immediate access, but the interface is simply a very long window of thumbnails and it is easy to get lost in the process.

Many other timeline solutions have minimal methods for dealing with an overload of data. The clustering approach of Autopsy is unique to help the user ignore some events and focus on others. It also incorporates events from many data sources and not simply file system times.

### Next Steps

We are continuing to build these features and more by engaging with a multitude of users to understand their needs and building an initial version to receive additional user feedback. By releasing this version into open-source, we will be able to receive feedback from users beyond those who were our original group of survey participants.



## **CYBER SECURITY COMPETITIONS:**

© **Web-based Interactive Stories for [Cyber] Education (CyberWISE)**



# Web-based Interactive Stories for [Cyber] Education (CyberWISE)

Laurin Buchanan

Anita D'Amico

Edward Rhyne, Cybersecurity Competitions Program Manager

## Overview

CyberWISE is a tool for educators, students, employers, subject matter experts and non-subject matter experts to teach or evaluate knowledge using interactive, graphic branching stories. These branching stories, or “choose your own adventure” (CYOA) comics, let readers make choices that determine a characters actions and the story’s outcome. Readers can make decisions on topics of cyber security and explore the consequences in the safe environment of a comic, and no artists or programmers are needed to develop the branching interactive stories.

## Customer Need

Raising safe computing awareness and changing risky behavior is a known difficult problem. The target audience must understand how and why the risk applies to them, that the risk brings real consequences, and that they can do something to reduce the risk. Explaining the causes and effects of cyber events can be especially difficult, as they do not occur in a context that is easily visible to individuals. What is needed is a way to help people of all ages and backgrounds explore both risky and safe cyber behaviors, and see the consequences of choices made in a safe environment.

## Our Approach

In order to address this need, CyberWISE uses visual storytelling to help people comprehend the interaction of cause and effect of cyber events. Learners read the story and then make a choice that affects the storyline. In order to simplify and accelerate the creation of delivery of these interactive educational materials, the CyberWISE tool provides a unique system that enables non-programmers and non-artists to easily develop branching storylines using advanced automation technologies and pre-rendered art assets.

## Benefits

Developing graphic interactive stories the traditional way is costly and time consuming, and requires specialized skills which present barriers to creation and dissemination. CyberWISE automates the most technically and artistically intensive aspects of production, from initial concept generation to the creation of graphical multi-path storyboards.



## Competitive Advantage

There are currently no other known integrated solutions for the creation of interactive storylines that guide users through the process of creating branching, graphic stories that educate learners. There are also not many web-based comic creation tools that allow the creator to have complete control over comic assets.

## Availability

CyberWISE is currently available for piloting, testing and evaluation.



# CYBER SECURITY INCIDENT RESPONSE TEAMS:

© Improving CSIRT Skills, Dynamics and Effectiveness

# Improving CSIRT Skills, Dynamics and Effectiveness

Heather Drinan

Scott Tousley,  
Incident Response Program Manager

## Overview

This applied research project provides recommendations to improve the skills, dynamics and effectiveness of Cyber Security Incident Response Teams (CSIRT). This research will determine and validate the principles of creating, running and sustaining an effective CSIRT. The output will include descriptions of needed knowledge, skills and abilities for key CSIRT roles, viewed from individual, team and multi-team system perspectives, as well as simulation-derived recommendations for optimal CSIRT performance.

## Customer Need

CSIRT teams are often both dynamically formed and temporary in nature, assembled in response to specific incidents. In cyber incident response, the teams are often responding to problems or incidents that have not been seen before. There is no social science anchored set of guiding principles and best practices that CSIRTs can apply to their organization, training and execution. Guiding principles are essential because the growth of the interconnected world is driving a need for better organized and dynamic CSIRT teams that can respond effectively to an array of incidents, the precise nature of which may not be known until it occurs.

## Our Approach

The CSIRT research team applies a multi-disciplinary, multi-step approach to determine and validate the principles of creating, running and sustaining an effective CSIRT. The team, led by Dartmouth College, includes Organizational psychologists from George Mason University looking at knowledge, skills and abilities (KSAs); teams; multi-team systems (MTS); and interactions between and across CSIRTs. Hewlett-Packard, as the contractor for the Navy-Marine Corps Intranet (NMCI), provides access to CSIRT teams and performs process modeling. Specific approach steps include:



1. Define effectiveness of CSIRTs, specifically identify team characteristics and outcome measures,
2. Define response triggers: What starts CSIRT actions, determines their size, and escalation criteria?,
3. Describe CSIRT processes for the types of teams involved, developed through interviews and observations,
4. Developing a common taxonomy of CSIRT performance dimensions at individual, team and MTS levels, and finally,
5. Encouraging change, through participation in workshops, developing and transitioning a handbook of best practices and actionable recommendations for CSIRT managers.



The background features a network of white padlock icons connected by dashed white lines. The padlocks are arranged in a grid-like pattern, with some appearing to be highlighted or glowing. The overall color scheme is dark blue with white highlights.

## INSIDER THREAT:

- © **Monitoring Database Management System (DBMS) Activity for Detecting Data Exfiltration by Insiders**

# Monitoring Database Management System (DBMS) Activity for Detecting Data Exfiltration by Insiders

Northrop  
Grumman

Donald Steiner, Ph.D.

Jennifer Smith Miller

Megan Mahle, Insider Threat  
Program Manager

## Overview

Within the Monitoring database management system (DBMS) Activity for Detecting Data Exfiltration by Insiders (MDBMS) project, Northrop Grumman and their subcontractor Purdue University are developing an anomaly detection system, DBSAFE, to protect relational databases from data exfiltration attempts. DBSAFE uses machine learning techniques to develop models of normal Structured Query Language (SQL) statements used by users within roles as determined by standard database Role-Based Access Control (RBAC) procedures. These models are then used to identify anomalous queries made by individual users that may be indicative of data exfiltration, manipulation, or sabotage.

## Customer Need

Data represents one of the most important assets of an organization. The undesired release (exfiltration) or manipulation of sensitive or proprietary data is one of the most severe threats of insider cyber-attacks. A malicious insider who has the proper credentials and authorization to access organizational databases may, over time, send data outside the organization's network through a variety of channels. Likewise, the authorized insider may manipulate data with undesired and damaging consequences.

While data exists throughout the organization, major harm can be done by exfiltrating large quantities of sensitive data that reside in an organization's relational database management system (RDBMS). By studying the patterns of interaction between users and an RDBMS, it is possible to detect anomalous activity that may indicate early signs of exfiltration. Customers need an anomaly and misuse detection system that operates at the data source (i.e., the RDBMS) in order to prevent data from leaving the source before it enters the organizational network where it is hard to track.

## Our Approach

The MDBMS project is researching and developing techniques for detecting and countering efforts by insiders to extract, exfiltrate, and/or manipulate sensitive data. The approach comprises:

1. Profiling normal database interactions,
2. Detecting anomalous queries against the data store; and
3. Deploying countermeasures.

The project's goals and objectives are to:

- Perform research in techniques to detect data exfiltration attempts (Purdue University),
- Implement the techniques in successively comprehensive prototype software systems,
- Install a pilot system in an operational environment,
- Evaluate the prototype and pilot systems for their efficacy and accuracy, and
- Report on the results of the evaluation.

## Benefits

Benefits to the customer include:

- Dynamic and automated generation of behavioral profiles based on roles rather than individuals,
- Near-real time alerts of anomalous database activity that may be indicative of insider threat,
- Automated response, such as blocking query results, according to pre-defined policies, and
- History of anomalous queries, associated metadata, and explanation for forensic purposes.

## Competitive Advantage

Most existing cyber security defense tools focus on protecting an organization from external attacks and are

# Monitoring Database Management System Activity for Detecting Data Exfiltration by Insiders

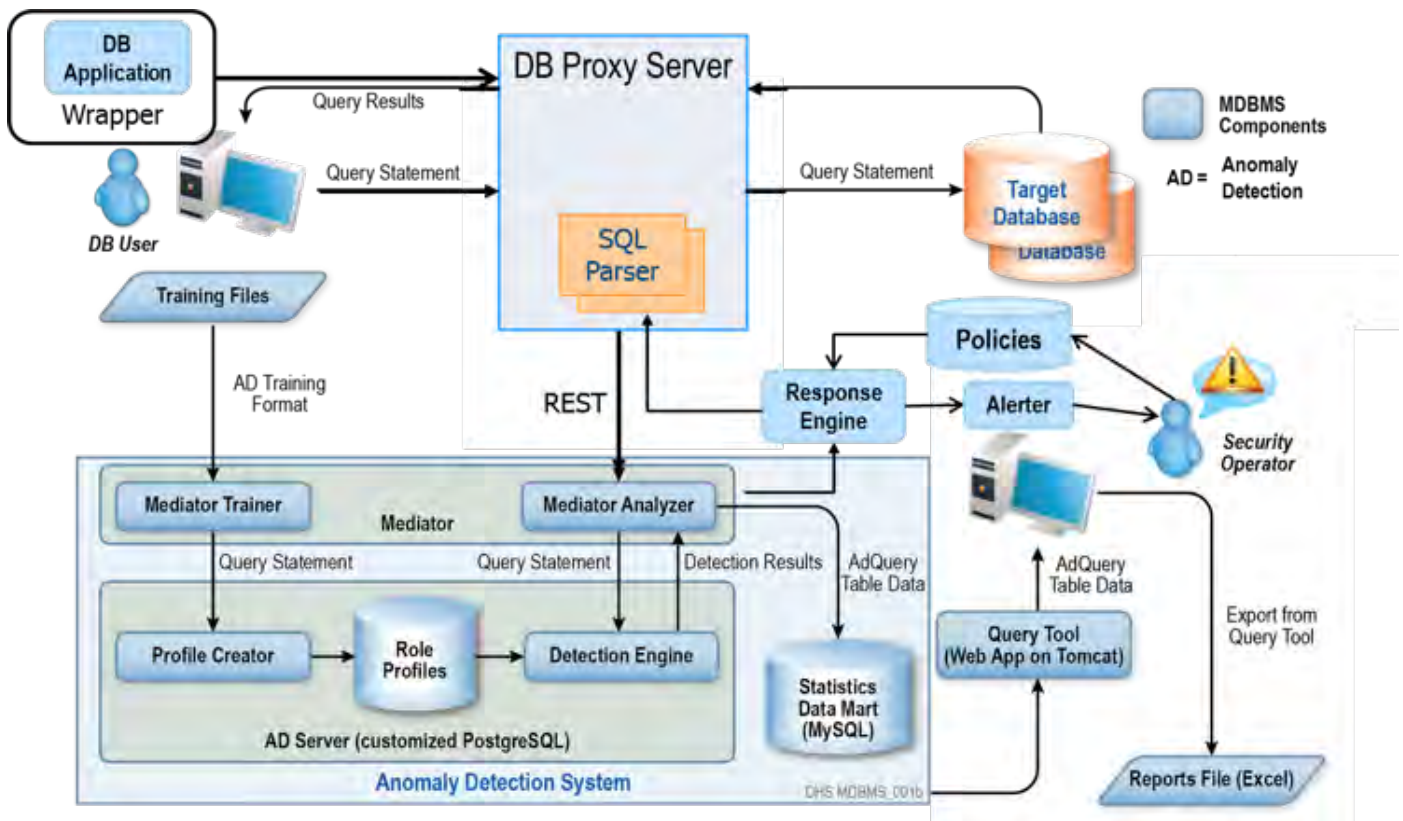
ineffective in the case of insider exfiltration attempts. Many insider threats detection tools focus on collecting behavioral patterns of individual users as opposed to protecting the data at the source; the overhead is cumbersome and threat identification is susceptible to delay. Those that do monitor data sources use hard-coded rules that may be complex and hard to manage. Currently deployed analytic approaches use static logs causing a delay in response. Profiling the users may catch only the “obvious” cases. Two-factor access mechanisms are cumbersome to deploy.

DBSAFE aims to overcome these hurdles by providing an

automated and readily adaptable mechanism to detect potential insider threats in near-real time.

## Next Steps

The DBSAFE tool is currently available for pilot deployment to evaluate the efficacy of the machine learning approaches in real-life enterprise environments.





# INTERNET MEASUREMENT AND ATTACK MODELING:

- ◎ **Alembic: Machine-Based Detection of Domain Ownership Change**
- ◎ **Cartographic Capabilities for Critical Cyber Infrastructure**
- ◎ **Clique: Safeguarding Cyber Systems with Visualization**
- ◎ **Enhancing Watchdog System for Internet Routing (WIT): Coupling Physical Infrastructure with Logical Infrastructure, Part II (WIT-II)**
- ◎ **Netalyzr: Measuring the Network from the Edge**
- ◎ **Stucco: A Cyber Intelligence Platform**
- ◎ **Symbiote® Defense**

# Alembic: Machine-Based Detection of Domain Ownership Change

Georgia Tech  
Research Corporation

David Dagan

Ann Cox, Internet Measurement and  
Attack Modeling Program Manager

## Overview

This document describes a new machine learning system that helps identify potential changes in domain ownership. Over 80% of all expired domains that end up being re-registered are used for malicious purposes, and many advanced persistent threat (APT) attacks use expired domains, since they often have a benign historic reputation, to evade detection, WHOIS lookups can be used to detect the re-registration of a domain, but doing this at scale, for hundreds of millions of domains, is difficult due to rate limiting of WHOIS queries by registrars. The Alembic system uses machine learning and features computed from passive DNS to permit light-weight, fast monitoring of expired domains. Domains that change ownership can then have new reputation scores computed, effectively reexamining the residual trust placed in these domains.

## Customer Need

Network operators and security researchers must curate and maintain their domain reputation systems so that expired domains no longer carry residual trust. Testing whether a single domain has changed ownership is easy; one merely has to compare WHOIS output by checking the registration date fields—perhaps the only fields in WHOIS output that cannot be manipulated. But doing this at scale, for hundreds of millions of domains, is quite difficult. Most registries rate limit WHOIS lookups and some even permit only three WHOIS lookups per day per IP. Even commercial sources of WHOIS data rate limit queries or make it expensive and difficult to check hundreds of millions of WHOIS records.

## Our Approach

To help automate light weight domain ownership detection, we built Alembic, a machine learning system that uses only DNS properties to infer a change in ownership. When domains expire, they enter into a

“redemption grace period” created by Internet Corporation for Assigned Names and Numbers (ICANN), so that accidental expirations can be caught and remedied. Similarly, former name servers treat expired domains differently (often by refusing to resolve them or by dropping out of the delegation path). Our Alembic system uses these and other signals to infer domain ownership changes.

Users can then identify domains that have changed ownership in order to evaluate whether they deserve any existing reputation or residual trust. For example, formerly whitelisted domains can be moved to a neutral status or provided a new reputation score. Since most re-registered domains end up being used for malicious purposes, forensic investigators might wish to prioritize their focus on “old/re-registered” resource record (RR) data during investigations.

## Conclusion

Domain reputation systems have proved effective, and attackers have taken notice. Many malicious campaigns now use expired whitelisted domains to evade detection. Detecting domain ownership changes is resource-intensive and difficult to scale. Our Alembic system therefore uses light-weight DNS properties to infer a domain ownership change. This can be used to discard previous domain reputation scores, curate white lists, and otherwise improve network security and forensic investigations.

## Alembic: Machine-Based Detection of Domain Ownership Change

---

### Next Steps

We have identified several security companies that use domain reputation in their product offerings, and Alembic presents an interesting “high signal/high value” capability. To enable wider use of this technology, the Georgia Tech Research Corporation plans the following:

- We will create an “Alembic domain blacklist” or rbdnsd zone of domains that have changed ownership based on the output of our system. This will help users understand how domain ownership changes can be tracked and incorporated into their existing technologies.
- We will work with several interested companies to transition the technology away from academia and into commercial use. This will include a refresh and packaging of the Alembic machine learning toolkit; in addition, it will necessitate updating the tool to support various output formats (e.g., zone editing, domain intelligence data storage, security information and event management (SIEM) integration, etc.)



# Cartographic Capabilities for Critical Cyber Infrastructure

Center for Applied Internet Data Analysis (CAIDA)

KC Claffy

Ann Cox, Internet Measurement and Attack Modeling Program Manager

## Overview

Researchers at the Center for Applied Internet Data Analysis (CAIDA) are maintain active Internet measurement infrastructure, and developing new techniques to collect, analyze, and process resulting measurement data. As part of this effort CAIDA makes two tools available for pilot deployment:

- scamper - an open-source packet-prober for active measurement of the Internet supporting IPv4, IPv6, ping, and several traceroute variants (Originally developed by Matthew Luckie at University of Waikato; enhanced at CAIDA.)
- Monotonic ID-based Alias Resolution (MIDAR) – a powerful scalable tool for mapping collected IP topology data to router-level granularity (alias resolution).



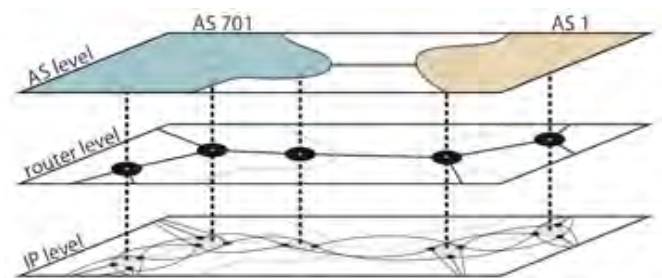
## Customer Need

The “cyber threat” is one of the most serious challenges we face as a nation. America’s economic prosperity in the 21st century crucially depends upon a secure and trustworthy communications fabric. Yet, we lack a thorough understanding of the vulnerabilities of the global Internet. Versatile, secure measurement infrastructures; reliable, representative, high-quality Internet data sets; and advanced measurement and analysis tools are rarely available to researchers and developers. CAIDA researchers bridge this gap, by offering tools and data sets that advance situational awareness of the Internet

by supporting crucial studies of the structure, dynamics, performance, and vulnerabilities of global Internet topology.

## Our Approach

We use both scamper and MIDAR for regular data collections on the Archipelago (Ark) measurement platform designed, deployed, and maintained by CAIDA. Distributed all over the world, Ark monitors (135 and growing, see map) are tailored to support active network measurements. The monitors use scamper for sending probes and collecting traceroute-like data to all routed /24 networks in the IPv4 address space and to all announced IPv6 prefixes (/48 or shorter) every 2-3 days. This data is used to derive maps of the Internet at various granularity levels: IP, router, and Autonomous Systems. The Ark monitors run the MIDAR tool every 4-6 months for alias resolution measurements to obtain the input data necessary for constructing router-level Internet topologies. These topologies become one of the components of the Internet Topology Data Kits (ITDK), regularly produced and distributed to network and security researchers and analysts.



## Benefits

The integration of strategic measurement and analysis capabilities has enabled us to provide comprehensive annotated Internet topology maps, as well as provides a platform capable of critical Internet infrastructure security assessments.

## Cartographic Capabilities for Critical Cyber Infrastructure

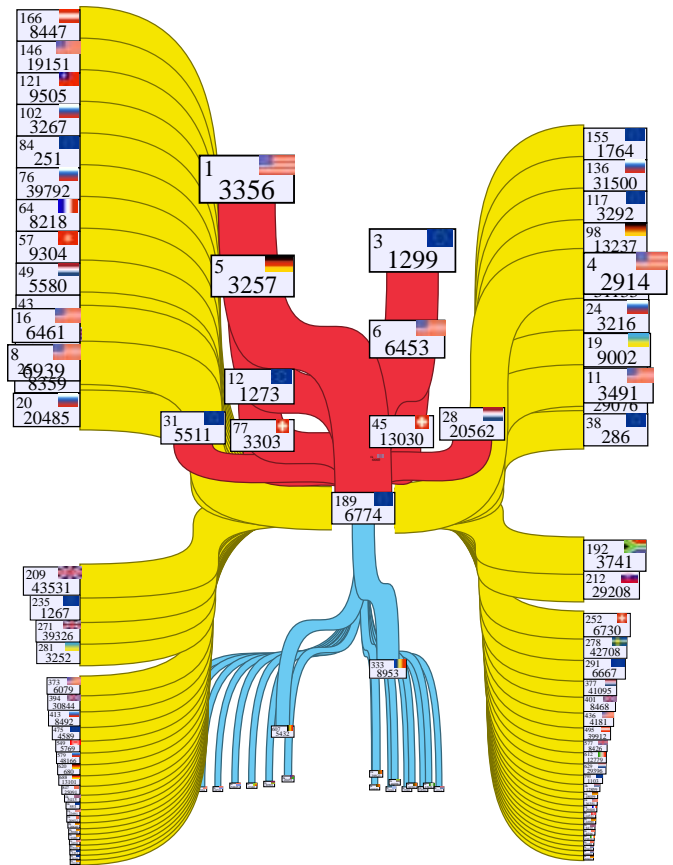
ITDK, CAIDA's flagship product, contains richly annotated topology maps of the observable Internet at multiple granularity levels, providing a more detailed and validated topological view than ever achieved before. These data sets enable crucial empirical research in network and security fields, advance our ability to identify, monitor, and model critical cyber infrastructure, and deepen our insight into the structure, behavior, and evolution of the global Internet.

CAIDA researchers continue to increase the number of Ark vantage points, refine measurement methods, develop new tools, and improve analysis and inference algorithms. The DHS S&T supported Ark active measurement platform supports cybersecurity-related situational awareness through macroscopic active measurements, including projects that integrate interdomain routing data to understand the Autonomous System (AS)-level structure of the ecosystem (depicted for one AS in the diagram on the right), and darknet traffic data to infer macroscopic disruptions. Cybersecurity-related uses of the platform include measurement of security best practice compliance, IPv4 and IPv6 stability, TCP security vulnerabilities, middlebox behavior, and detection of large-scale outages and BGP hijacks.

### Competitive Advantage

scamper is a flexible, powerful parallelized packet-probing tool that makes it easy to conduct large-scale measurements and archive collected data in a well-defined format. Its design allows researchers to focus on their measurement idea rather than infrastructure and logistical details of safe execution of global Internet measurement.

CAIDA's MIDAR tool uses pre-existing topology data as well as real-time measurements to infer aliases (IP addresses on the same router) based on similarities in IP ID time series produced by different IP addresses. We developed a precise IP ID comparison test based on monotonicity rather than proximity. To achieve greater scalability, MIDAR utilizes multiple probing methods & vantage points, and a novel sliding-window probe-scheduling algorithm. This approach minimizes the false positive rate sufficiently to achieve a high positive predictive value at Internet scale, with greater precision and sensitivity than previously achieved.



### Next Steps

scamper and MIDAR are available for download and pilot development. scamper compiles on most UNIX-like operating systems, Windows, and DragonFly. MIDAR requires POSIX or a UNIX-like operating system. Users have a choice between three different front-ends: the small-scale resolution tool (< 200) IP addresses from a single monitor host; the medium-scale MIDAR resolution system, capable of testing a medium-size (< 40000) set of IP addresses from a single monitor host; and the large-scale system, capable of testing an Internet-scale (at least 2 million) set of IP addresses from multiple monitor hosts. The MIDAR package includes software support for replicating the infrastructure in a new deployment. For other uses of the Ark platform, [contact Caida](#) or see [Caida tools](#) and [Archipelago \(Ark\) Measurement Infrastructure](#).

# Clique: Safeguarding Cyber Systems with Visualization

Pacific Northwest  
National Laboratory

Daniel Best

Ann Cox, Internet Measurement and  
Attack Modeling Program Manager

## Overview

Researchers at Pacific Northwest National Laboratory (PNNL) developed the Correlation Layers for Information Query and Exploration (Clique) capability as scalable cyber visualization tool that offers insight into massive data sets, and helps users discover early indicators of potentially malicious activity.

## Customer Need

Protecting communications networks against attacks that aim to steal information, disrupt order, or harm critical infrastructure requires the collection and analysis of staggering amounts of data. The ability to detect and respond to threats quickly is a paramount concern that spans government, utilities, financial and private sectors. These organizations share a common burden of identification of threats buried within billions of network transactions each day. To better equip analysts, state-of-the-art data intensive visual analytics tools are needed to address the unique challenges within cyber security.

## Our Approach

In response to this need, PNNL developed Clique, a capability with multiple views, Trace and Cadence, to visualize network data.

The Cadence view within Clique displays high-level overviews of network traffic using a behavioral model-based anomaly detection technique. This technique builds models for learning and classifying expected behavior of individual hosts on a network and compares these modeled behaviors to current data. The result is a deviation score that provides indicators of “non-normal” network activity.

The effectiveness of Cadence is enhanced by a graphical user interface that enables the highlights anomalous activity using color saturation and progressive disclosure,

empowering users with the ability to identify deviations from expected activity. Users can navigate through their data temporally, viewing time periods as short as a few minutes or as long as many hours. Cadence models help analysts to see departures from normal behavior at any time scale. The user interface enables drill down capability so that analysts can view detailed displays of network activity to determine the machines, buildings, sites, or other sources of traffic behaving anomalously.

The Trace view within Clique provides analysts with a flexible and scalable two-dimensional scatter plot. This enables identification of patterns that exist in large volumes of network data. This visualization approach was specifically requested by defenders on the front lines to display raw network traffic using multiple attribute-based views and supports millions of communications events in a single view. Trace enables users to use human cognition to identify patterns contained within large volumes of data that is tremendously difficult using other analytic techniques. This enables analysts to view typical communication patterns contained in their data and provides a mechanism to highlight patterns of interest.

Trace empowers analysts with the ability to interact and explore their data at scale. The view provides a mechanism for analysts to color encode the data in meaningful ways to highlight features of interest. Once a feature is identified, analysts can view summary statistics about a selection or display the underlying data in a common tabular view where they can export for reporting and use in other tools.

## Benefits

The multiple views provided in Clique enable users to readily move from high-level views of millions of transactions in Cadence, down to detailed representations in Trace. The result is significantly improved situational awareness of network activity, which provides more efficient investigation to support prevention, response, and mitigation of harmful attacks.



# Clique: Safeguarding Cyber Systems with Visualization

## Competitive Advantage

Clique provides visibility and command of network data in ways not previously possible. In addition, the tools are designed to work together to support an investigative workflow.

## Next Steps

These tools have already been tested and demonstrated at several government organizations. Now, further pilot partners are being sought to validate the technology's operational readiness.



# Enhancing Watchdog System for Internet Routing (WIT):

## Coupling Physical Infrastructure with Logical Infrastructure, Part II (WIT-II)

Christos Papadopoulos

Ann Cox, Internet Measurement and Attack Modeling Program Manager

### Overview

Watchdog System for Internet Routing (WIT) builds a Critical Prefix Monitoring System for the Australian government. WIT-II has now been deployed in CERT Australia.

### Customer Need

A nation relies on Critical Network Infrastructure to provide services to its citizens. Critical infrastructure exists at both the private and public sectors. Monitoring such infrastructure enables a nation to detect and respond quickly to potentially catastrophic threats.

### Our Approach

The WIT-II project monitors the global Internet routing infrastructure looking for threats such as Border Gateway Protocol (BGP) route hijacks, loss of prefix visibility and anomalous paths. WIT-II relies on the BGPmon monitoring platform with over 400 monitoring points distributed worldwide. WIT-II also deploys private instances of the platform to monitor internal, private networks. WIT-II has access to a nearly 20-year routing archive to provide historical context of the Internet control plane.

### Benefits

Unlike other centralized platforms, the ability of WIT-II to be deployed privately enables secure monitoring of critical infrastructure without disclosing information to other parties. WIT-II provides a queryable database based on a modern distributed, scalable NOSQL database with a high degree of resilience and redundancy. Users have several options: (a) deploy a completely private instance, (b) a read-only instance where global routing information is received from the global BGPmon system or (c) a public deployment that shares data with other users. All monitoring points in the BGPmon system have been accurately geolocated to enable location-based characterization of prefix activity. The project can also characterize routing detours, paths between endpoints within the same country that briefly traverse a foreign country.

### Competitive Advantage

WIT-II offers the option of a private deployment and the use of proven cloud-based software that scales to levels far beyond current needs. WIT-II integrates with current geolocation services to characterize observed paths in terms of transient and long-lasting anomalies.

### Next Steps

WIT-II is currently undergoing fine-tuning of its database parameters. It is available for deployment and use in any network that wishes to monitor routing prefixes. Plans are in progress to develop a dataplane probing mechanism triggered by events at the control plane to verify anomalies and collect fine-grain information.



# Netalyzr: Measuring the Network from the Edge

International Computer  
Science Institute

Nicholas Weaver

Christian Kreibich

Ann Cox, Internet Measurement and  
Attack Modeling Program Manager

## Overview

A user's Internet Experience is often not defined by the overall Internet but is often constrained by the last mile and even the last meter - with many issues occurring due to local equipment or the Internet Service Providers (ISP). Netalyzr provides an easy to use network measurement tool on both Android phones and Java-enabled web browsers that comprehensively measures the edge network, diagnoses problems, and provides a detailed report.

## Customer Need

"The net is broken" is a common frustration, but often the cause may be hard to determine. Likewise, there is a long history of ISPs deliberately manipulating traffic for their own purposes, ranging from intercepting sites (to replace web searches with advertising or adding affiliate tags to shopping) to intercepting all HTTP requests to inject tracking "permacookies".

## Our Approach

Netalyzr measures from the edge, providing an automatic free detection and diagnosis service. By using standard networking functions, either as an Android app or a Java applet, connecting to a custom suite of servers, Netalyzr probes the network connection searching for a wide variety of possible misconfigurations or errors, ranging from almost ubiquitous performance defects to serious security vulnerabilities.

## Benefits

Netalyzr provides a free, comprehensive network measurement service available to anyone with an Android phone or a Java-capable web browser. Not only are the tests thorough, but the report instantly highlights possible

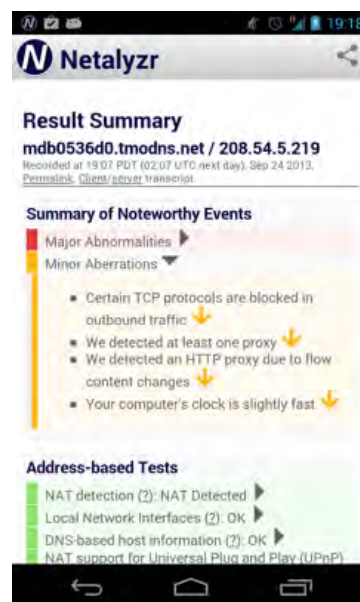
anomalies, including several security vulnerabilities. At the same time, Netalyzr has enabled a comprehensive picture of the edge of the network, including detecting limitations on DNS, problems with the TLS certificate stored on Androids, the prevalence of "bufferbloat" and other issues.

## Competitive Advantage

Although there are numerous network performance testers available both as browser applications and phone apps, none of these systems provides the comprehensive diagnostic and detection suite in Netalyzr.

## Next Steps

Netalyzr is now operating in a sustained mode: although the service is no longer being significantly developed, we have optimized for low cost of operation and intend to maintain the free Netalyzr service for years to come.





**This page left blank intentionally**

**John Goodall**

**Ann Cox, Internet Measurement and Attack Modeling Program Manager**

### Overview

The Lawrence Livermore National Laboratory Network Mapping System (NeMS) is a software-based network characterization and discovery tool. NeMS produces a comprehensive representation of Internet Protocol (IP)-based computer network environments constructing visual representations of the targeted network based on observed behavior. The tool provides an iterative analysis platform from which network security managers and information technology (IT) personnel can explore the findings of each mapping operation.

### Customer Need

Security event data, such as intrusion detection system alerts, provide a starting point for analysis, but are information impoverished. To provide context, analysts must manually gather and synthesize relevant data from a myriad of sources within and external to their enterprise. Analysts search system logs, network flows, and firewall data, along with IP blacklists and reputation lists, software vulnerability information, malware and threat data, operating system (OS) and application vendor blogs, and news sites. All of these sources are manually searched for data relevant to the event being investigated. Relevant results must then be brought together and synthesized to put the event in context and make decisions about its importance and impact. This is a manual, tedious process, but the results of this process are required to know how to react to events.

### Our Approach

Stucco is a cyber intelligence platform to help automate this process and provide relevant information to analysts quickly and easily. Stucco collects data not typically integrated into security systems, extracts domain concepts and relationships, and integrates that information into a cyber security knowledge graph to accelerate decision making.

Stucco addresses a fundamental problem in cyber security: quickly putting security events in context

Stucco's approach includes the following:

- Continuous collection and processing of documents from endogenous and exogenous sources
- Domain Specification Language for parsing and extracting domain concepts and relationships from structured data
- Natural language processing for extracting domain concepts and relationships from unstructured text documents
- Alignment methods for instantiating the knowledge graph
- API for programmatically accessing the knowledge graph
- Visualizations for exploring the knowledge graph to derive context

Some of the data sets Stucco can currently ingest includes:

- Observables
  - Network flows (Argus)
  - Process à Port (Hone)
  - User logins (auth.log)
  - Package list (deb)
  - GeoIP (Maxmind)
  - AS mapping (CAIDA)
  - Software (CPE)
- Exploit targets
  - Vulnerabilities (NVD, Bugtraq, Metasploit)
- TTPs
  - Exploits (Metasploit)
  - Malware (F-Secure, Emerging Threats, 1d4, Zeus Tracker)
  - Virus (CleanMX, Sophos)

# Stucco: A Cyber Intelligence Platform

- Indicators
  - Attacker IPs (Malware domain list, Sophos)
- Course of Action
  - Remediation (Bugtraq)

## Benefits

By organizing data into a knowledge graph, security analysts will be able to rapidly search for domain concepts, speeding up access to the information needed for decision-making. The information returned will only be that which is pertinent to their search. Our approach enables analysts to more quickly identify events that can be discarded as false positives and to perform more thorough analysis with the relevant context to make decisions.

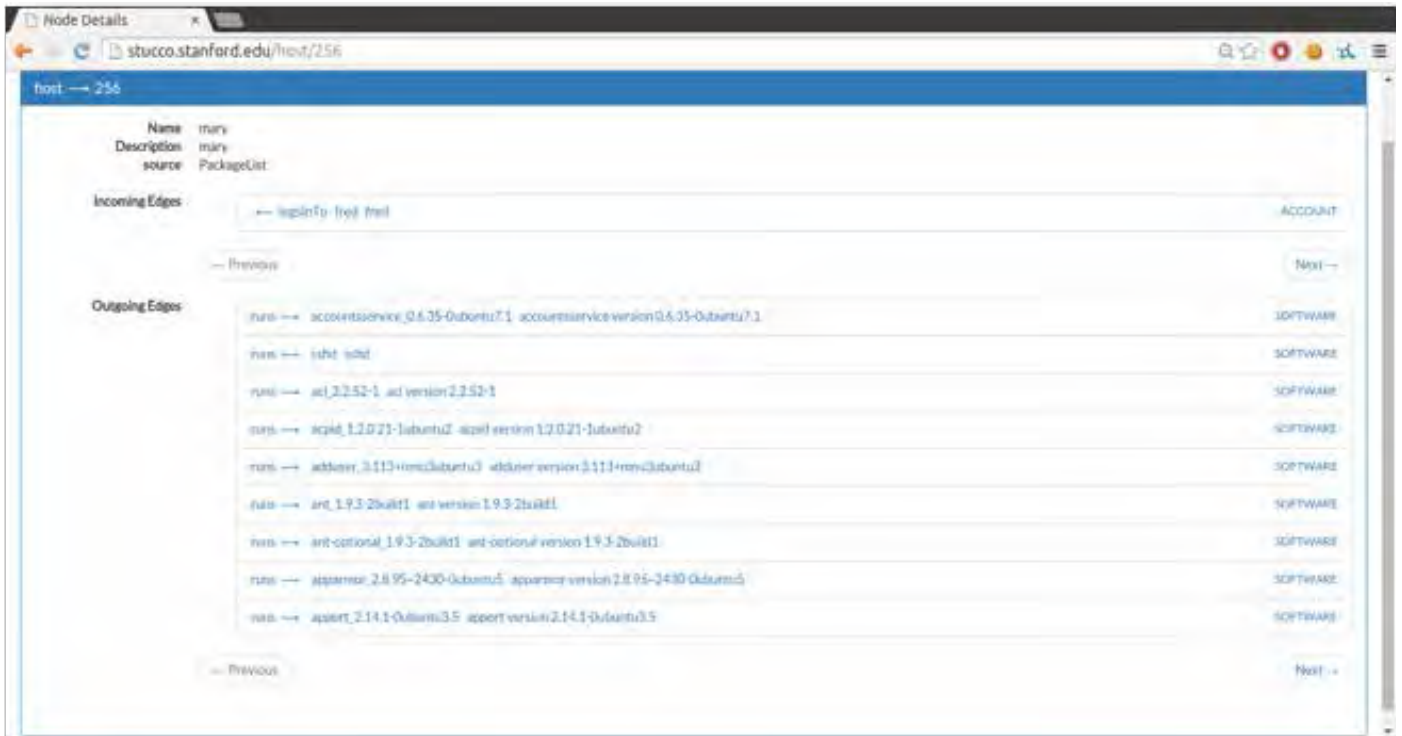
## Competitive Advantage

Current tools tend to focus entirely on local, enterprise data (e.g. log aggregators or security information and event management (SIEM)s), whereas Stucco integrates these enterprise data sources with external data, such as exploit and vulnerability databases, as well as malware or spam blacklists. Current tools also typically only ingest structured data sets, whereas Stucco includes information extraction capabilities to pull out relevant information from unstructured data sources, such as security blogs or vendor posts.

## Next Steps

Stucco is currently available for pilot deployment to enhance situation awareness and augment analysis.

Stucco is open-source (Massachusetts Institute of Technology (MIT) license) and freely available. For more information, with links to a demonstration site, source code, instructions for building a development and demonstration environment, [visit the project page](#).





**Dr. Ang Cui**

**Calvin Chu**

**Ann Cox, Internet Measurement and Attack Modeling Program Manager**

## Overview

Red Balloon Security is the leading embedded device security company, delivering deep host-based defense for all devices. Our technology was invented at Columbia University in the City of New York under funding sponsorship with Department of Homeland Security Science and Technology (S&T) and Defense Advanced Research Project Agency (DARPA).

Symbiote® Defense enables true host-based intrusion defense and situational awareness, capable of defending against nation-state level advanced zero-day exploitation of embedded devices. A robust defense-in-depth security posture requires host-based protection in the event the preceding layers of defense are compromised. Attacks resulting in persistent modification of device firmware and data may result in an advanced persistent threat. These threats can direct devices to be used as staging points to access and compromise other high value assets within an organization.

## Key Benefits

- Continuous firmware integrity attestation
- Real-time forensic information extraction
- Zero false-positives, zero signatures
- Meets hard real-time constraints
- OS agnostic, universally compatible
- Deployable via firmware update
- Enterprise scalable, centrally managed
- Integrates with Splunk and Arcsight

Enterprise embedded assets are not currently reporting to a security operations center in a consistent way, and typically not at all. Symbiote® Defense conforms entire enterprise fleets of embedded assets onto common security information and event management (SIEM) software. Specific security policies can be dialed in to allow for advanced tactical response.

## Accurate Detection, Versatile Response

Upon detection of an attack, Symbiote® can respond according to enterprise security policy. For example, it may automatically thwart the attack and capture the malware for analysis. It can also be directed to allow the attack so as not to yield the field before forensic information about the adversary can be recorded. Symbiote®'s versatile attack mitigation capabilities allows the managed device population to be protected rapidly, reliably and intelligently. It provides unmatched forensics collection capabilities that take embedded incidence response to the next level.

## Universal Compatibility

Symbiote® is compatible with all embedded devices running ARM, MIPS and PowerPC. Symbiote® is currently deployed in the following devices:



## Symbiote® Defense

- Cisco Routers, Switches and Firewalls
- Cisco & Avaya IP Phones
- HP Enterprise LaserJet printers
- Programmable logic controllers
- Avionic platforms
- Robotic platforms
- Many other device platforms

### Other Important Capabilities

- Strong anti-tampering and self-monitoring
- Preserves host-performance

Symbiote® Execution Manager (SEM) uses smart context aware scheduling algorithms to interleave execution processes without affecting real time device characteristics. SEM also self-monitors itself against attack.

## Symbiote® Defense Specifications

<b>ISA Supported</b>	ARM, MIPS, PowerPC and more
<b>SOC Integration</b>	McAfee ePo, HP ArcSight, Splunk ES
<b>Scope of Protection</b>	Protect against APT, Rootkits, ROP attacks and threats that cause a persistent change
<b>Detection Latency</b>	Typical threat detection in the millisecond regime

## Symbiote® Defense Capabilities

- Symbiote® Binary Structure Randomization (BSR)**
  - Provides strong anti-propagation by breaking firmware monoculture
  - Ensures binary diversity across populations of similar devices.
- Symbiote® Continuous Integrity Attestation (CIA)**
  - Active enforcement of firmware region and static data integrity against unauthorized changes
- Symbiote® Secure Telemetry**
  - Upstreams key system information to an aggregation and analytics appliance for situational awareness integration with SIEM
  - Uses a covert encrypted channel to transfer data and ensure against man in the middle (MiM)
- Symbiote® Implant Apprehension**
  - Upon detection of a threat, forensics information, including the actual malcode is captured and placed into audit log
  - The malcode is disassembled in real time in the AESOP appliance and deep forensics is sent to SIEM
- Symbiote® Context-Aware Response**
  - An automated incidence response security policy can be dialed in to respond to threats in real time.



# MOBILE DEVICE SECURITY:

© MobileRoT - Software-only Roots-of-Trust for Mobile Devices



# MobileRoT – Software-only Roots-of-Trust for Mobile Devices

Kristopher Carver

Dr. Andras Moritz

Jeffrey Gummeson

Vincent Sritapan, Mobile Device Security Program Manager

## Overview

BlueRISC's MobileRoT measures and verifies a mobile device's static and runtime state to enable trust and overall device security. It can be utilized to detect malicious system change or activity and to ensure that access to critical information and software can only be performed in a trusted state. MobileRoT requires no modifications to the underlying operating system kernel, nor any manufacturer or service provider support for insertion, which greatly reduces hurdles to adoption.

## Customer Need

The mobile device market has grown tremendously. Individuals, businesses, and governments rely on mobile devices to access critical infrastructure and share vital information (e.g. banking, medical, intellectual property, etc.). This growth in adoption has also brought about a parallel surge in attacks. Malware, ransomware and spyware are targeting mobile platforms to steal sensitive data, access private networks, track users and do other nefarious activities. Particularly for governments using mobile technology, mobile attacks can disrupt life-saving operations, endanger personnel and expose government systems to exploitation.

Roots-of-Trust (RoTs), which are highly trustworthy tamper-evident components, can provide a foundation to build security and trust. RoTs are usually provided as a specialized hardware chip (e.g., Trusted Platform Module) on desktop or laptop systems. However, mobile devices lack dedicated hardware mechanisms for providing RoTs. This leaves a single solution, namely to provide RoTs via software. Unfortunately, this is challenging to realize given the sophistication of current threats and the ease in which a mobile device's state and information can be extracted and altered. Moreover security specifications such as Trusted Computing Group's Mobile Trusted Module do not address how to support mobile RoTs in software nor do they address dynamic verification of device and software behavior while applications are running.

## Our Approach

To overcome the array of surface attacks designed against software-based systems, MobileRoT utilizes a new architecture for enabling transitive trust based on a Core Root of Trust for Measurement (CRTM). The CRTM is hardened code that acts as the root-of-trust for reliable integrity measurements and is the foundation for additional trusted services. The MobileRoT architecture includes a layer of encrypted CRTM code that is tied to a cryptographic key that is generated at boot-time. With the CRTM established, the resulting system does not require any sensitive information to be stored persistently in an unprotected state, closely mimicking the level of security achievable via a dedicated hardware. A secure cryptographic sealing and unsealing procedure tied to the boot-time and runtime measurements performed by the solution enables application and data protection. Since all protected data and applications are sealed, they remain protected even in the case of an attacker's attempt to alter or bypass the MobileRoT technology.

Traditional solutions focus primarily on boot-time validation, establishing the validity of each component prior to a complete boot, while providing only minimal support for runtime activities. Unfortunately, it is widely known that sophisticated attacks can target applications that are already running and devices are rarely rebooted these days. To address the shortcomings of one-time static verification, MobileRoT provides dynamic verification and attestation by also performing runtime measurements of the system state of the device. These runtime agents harden themselves from attack and modification by creating a self-validating network, which can instantly respond to a threat to the system or the protection technology itself.

While cybercrime targeting mobile devices is becoming pervasive, MobileRoT can preserve and confirm the integrity of the device while at rest or in use. BlueRISC's MobileRoT technology has overcome barriers to bring RoT

## MobileRoT – Software-only Roots-of-Trust for Mobile Devices

to a mobile platform, providing a foundation of security features to accelerate development of secure mobile devices.

### Benefits

The value proposition for BlueRISC’s MobileRoT product is the establishment of software-based static and dynamic RoTs that can be leveraged for providing application and data protection and trusted MDM policy enforcement. The automated installation methodology and the lack of any modifications to the underlying operating system kernel drastically reduce barrier-to-entry.

MobileRoT reliably allows all levels of software, including user applications, to have access to its trusted services through an open application programming interface (API). This enables the creation of secure off-the-shelf third-party and proprietary applications and data, and strengthens key management and policy enforcement technology, such as Mobile Device Management (MDM). MobileRoT also provides fine-grained protections integrated directly into an application. For example, BlueRISC has taken a standard Android Calendar application and modified to support the concept of a “Secure Event”. This secure event is established in cooperation with the MobileRoT and persistently protected. To view a secure event, proper authorization and authentication is required and the system state must be verified.

### Competitive Advantage

In the mobile device protection, there are two main types of solutions: those provided by the device manufacturers and those designed to operate on top of the OS to provide some user-land security services. Out of these two types of products, the former represent the main competition. Table 1 provides a more detailed competitive analysis between mobile protection solutions.

BlueRISC’s solution complements the user-land security solutions (such as MDM), which could take advantage of the RoTs provided by MobileRoT to harden their system/ approach via the open trusted services API. The provided features are valuable to traditional anti-virus/MDM companies because recent trends in security suggest that they are losing their value proposition as the attacks are

becoming more sophisticated. Lastly, one of the goals of MobileRoT is to provide a U.S.-made alternative to vendor-specific technologies such as Samsung’s Knox that is also open to third-party developers. This is also expanding upon the protections and trusted services while enabling flexibility.

**Table 1: Competitive analysis**

Features	BlueRISC	Samsung	Arxan	McAfee
Software-only Roots of Trust	1	0	1	1
Chain-of-trust: Boot through Runtime	1	1	0	0
Dynamic System Attestation	1	0	1	1
MTM Compatible*	1	0	0	0
Open API	1	1	0	0
Automated Technology Insertion	1	0	1	0
Provisioning for FIPS Certification	1	1	0	1
Supports Government Credentials	1	1	0	0
Owned & Operated in USA**	1	0	1	1
<b>Total:</b>	<b>9</b>	<b>4</b>	<b>4</b>	<b>4</b>

\* Enables 3rd party MTM compatible software to run

\*\*Critical for US Government & Defense Use Cases

### Next Steps

We are currently finalizing implementation of the Trusted Services API to provide beta versions to our existing partners for use-case development and security evaluation. We are always interested in exploring additional use-cases with new partners.

## MOVING TARGET DEFENSE:

© Self-shielding Dynamic Network Architecture (SDNA)



# Self-shielding Dynamic Network Architecture (SDNA)

Intelligent  
Automation Inc.

Nick Evancich

Sapna George

Edward Rhyne, Security of Cloud-Based  
Systems Program Manager

## Overview

The static nature of today's networks provides ample opportunity for attackers to gather intelligence, perform planning, and execute attacks at will. To address this problem, Intelligent Automation Inc. (IAI) has developed Self-shielding Dynamic Network Architecture (SDNA), a dynamic defense that alters network architecture and behavior to stop and contain cyber attacks while remaining transparent to the legitimate user.

## Customer Need

Today's networks are "sitting ducks" waiting for attackers to exploit them. To a determined adversary, there are many ways to get inside your network, bypass any current protection technologies, and attack intended targets. None of the current protection technologies stop the now common practice of attacking your network from within using zero-day exploits, stolen credentials, and other sophisticated tactics. An innovation in cyber security technology is needed that goes beyond what the current state of the art has to offer.

## Our Approach

SDNA prevents an attacker from targeting, entering, or spreading through your network by adding dynamics that present a constantly changing view of the network over space and time. The dynamics are IPv6-based and cryptographically strong. SDNA prevents malicious packets from even reaching the hosts. SDNA increases the attacker's effort, risk of detection, and time required to successfully conduct an attack. If an attacker gains a foothold inside your network, for example via a malicious insider or host compromised by a phishing attack, SDNA limits the attacker's ability to spread by constraining each host to an abstract, modified view of the network.

## Benefits

In order to protect against compromise, SDNA mutates the network's "DNA" through packet manipulation, policies, and rules to manage the competing goals of securing the network while providing legitimate users transparent access to needed services. Because of this, attackers (even with unlimited resources) cannot send traffic directly into an SDNA-protected enclave.



## Competitive Advantage

Current defenses check against signatures, behaviors, and artifacts of known attacks, but do not protect against unknown attacks. Firewalls are good at stopping attacks from entering the network but there is no protection once the attacker gets past them. Basic randomization can improve resilience, but does not prevent against misuse of credentials.

## Availability

SDNA is currently available for piloting, testing and evaluation.



# SECURITY OF CLOUD-BASED SYSTEMS:

© Silverline: Assessment System for Secure Cloud Computing



# Silverline: Assessment System for Secure Cloud Computing

Architecture Technology Corporation

Rob Joyce

Gene Proctor

Edward Rhyne, Security of Cloud-Based Systems Program Manager

## Overview

Cloud computing brings new security risks to enterprises and homeland security organizations. Architecture Technology Corporation's Silverline is an innovative software tool for automatically detecting, evaluating, and mitigating security risks in cloud-based software applications. Silverline saves analysts time, reduces costs, and promotes the secure deployment of real-world cloud applications.

## Customer Need

Cloud computing offers a flexible and cost-effective means for computation and enjoys a rapidly growing user base. Despite investment by cloud infrastructure vendors in various security solutions, in practice it is very difficult for their customers to evaluate their exposure to security risks such as data loss, theft of service, and exfiltration of user passwords. Programs such as FedRAMP are a great start but are generic baselines and cannot address a given organization's specific cloud applications.

## Our Approach

Silverline helps cloud service customers and software developers discover actual technical vulnerabilities in their systems, evaluate the risks those vulnerabilities pose in their specific applications, mitigate those risks, compare alternatives, and monitor for continuing compliance.

Silverline models risks to an application using attack trees that describe the high level goals of a potential attacker, such as exfiltrating sensitive data or compromising a service's availability. Goals are then divided into sub-goals, down to the level of individual items that can be tested and verified.

Using the attack tree model, Silverline helps automate testing and analysis. Individual sub-goal tests for configuration items, either on a cloud instance or using

a cloud provider's API, are candidates for automatic execution. These can be implemented as remote executables, scripts, or NIST Security Content Automation Protocol (SCAP) data streams. Silverline can also automatically perform database lookups, using the NIST National Vulnerability Database (NVD) or other on-line data sources, to assess whether a particular sub-goal's requirements have been met. Sub-goals may also depend on policy compliance, training, or other non-technical factors that can be recorded alongside the automated tests.

Silverline runs all the automatable tests, gathers any manually-entered information, and "bubbles up" the results and metrics for a high-level view of the threats to a cloud application. This enables an analyst to assess the overall impact (e.g., to data confidentiality) of remediating individual vulnerable configuration items, as well as to pinpoint the fixes that have the greatest impact for the lowest cost.

## Benefits

Silverline increases visibility into the security impact of individual design and configuration decisions, helping analysts and developers make targeted, efficient improvements with the highest impact first—optimizing already-stretched security budgets.

Improved security removes a barrier to the widespread adoption of cloud computing and mitigates risks posed by migration to the cloud. As utilities, customs and border agencies, transit authorities, travel monitoring organizations, financial and healthcare entities, first responder organizations, and the DoD move computing operations into the cloud, cloud security is becoming critical for homeland security.

Silverline provides a systematic, standards-compliant approach to evaluating security risks to applications running on cloud platforms—especially Infrastructure



# Silverline: Assessment System for Secure Cloud Computing

as a Service (IaaS) and Platform as a Service (PaaS) clouds. Silverline's approach also applies in the areas of requirements analysis, cost/benefit/risk analysis, impact analysis, software and distributed systems testing, dependency tracking, forensic analysis, and vulnerability tracking in complex systems.

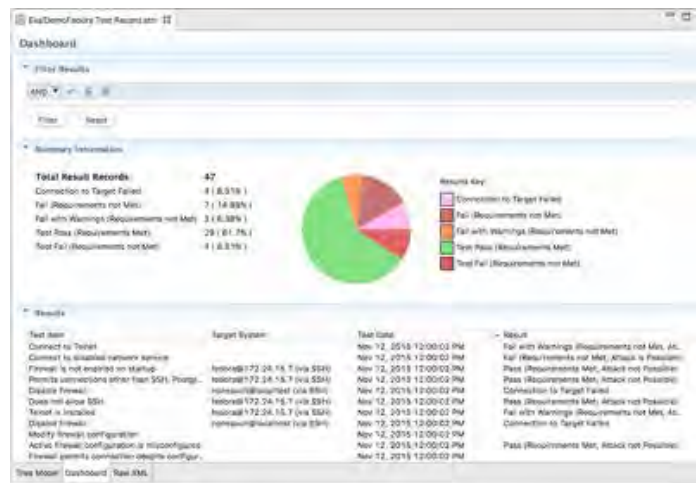
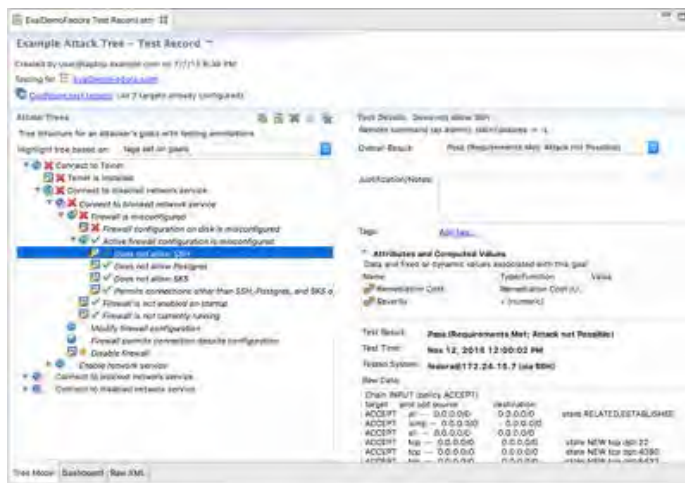
## Competitive Advantage

Current attack tree modeling tools do not scale well to large trees and few handle the complex metadata and attributes that Silverline provides. Competing security modeling tools do not offer automated testing or metric computation. Meanwhile, enterprise-scale automated testing and compliance tools provide only rudimentary modeling capabilities, leaving security personnel with little insight into the broader impact of potential problems and no roadmap for fixing them.

## Availability

Silverline is currently available for evaluation.

Public release is scheduled for Spring 2016.



Silverline's cloud attack model provides drill-down capabilities (left) and high-level summaries (right) of automated testing results—and their impact on overall security goals.



# SOFTWARE ASSURANCE MARKETPLACE:

© The Software Assurance Marketplace (SWAMP)



# The Software Assurance Marketplace (SWAMP)

Advancing continuous software assurance, open-source innovation, & cybersecurity

Irene Landrum

Miron Livny

# The Software Assurance Marketplace (SWAMP)

Kevin Green, SWAMP Program Manager

## Overview

The Software Assurance Marketplace (SWAMP) is an open facility that is designed, built, and operated by four research institutions. Launched in February 2014, the SWAMP offers a no-cost software assurance testing platform that combines an array of open-source and commercial software assurance tools with advanced high throughput computing. The SWAMP also includes a growing library of open-source applications with known vulnerabilities to help developers improve the effectiveness of their static and dynamic analysis tools. SWAMP users' activities and results are private to the user and those with whom they choose to share. The SWAMP offers an open marketplace of software packages and analysis tools, with the ability to control how packages, tools, and expertise are shared with the entire software community. With the computing capacity required to support continuous assurance, the SWAMP provides the automation to continuously run multiple analysis tools against software packages. Software and tool developers can use an integrated results viewer to display weakness reports with integrated CWEs (common weakness enumerations) from multiple tools.

## Customer Need

Software is integrated into nearly every aspect of our lives and heavily utilized by devices at all sizes. Security breaches are regular news headlines. Software applications need to be built securely at the code level and tested regularly to ensure security and protect privacy. The SWAMP promotes continuous assurance technologies and practices through an open and collaborative framework that protects confidential data and facilitates sharing, thus, making it easier for software and tool developers to adopt continuous assurance practices.

## Our Approach

The Software Assurance Marketplace is a collaboration of four research institutions, each providing expertise to enhance the security and robustness of the SWAMP.

- The Morgridge Institute for Research, a private, non-profit research institute located on the University of Wisconsin-Madison campus, leads the initiative and hosts the infrastructure, core development, and software testing teams.
- The University of Wisconsin-Madison's Middleware Security and Testing team supplies the framework and research on software analysis tools.
- Indiana University's Center for Applied Cybersecurity Research and High Throughput Computing group manages cybersecurity and 24/7 end user support.
- The University of Illinois Urbana-Champaign's National Center for Supercomputing Applications maintains identity access management.
- Together, the team promotes the practice of continuous software assurance, evaluating software and remediating weaknesses throughout the software development lifecycle.

## Benefits

The SWAMP is a no-cost resource available to the global software community, providing a powerful, flexible, and secure facility for organizations and open-source developers to institute software assurance practices. The supported platforms, tools, and packages are maintained by the SWAMP team, lowering the obstacles to performing software security assessments. SWAMP encourages software developers, software assurance researchers, infrastructure operators, educators, students, and individuals from open-source, government, and commercial groups to assess their software, both developed and acquired, to promote a more stable and secure software ecosystem.



## The Software Assurance Marketplace (SWAMP)

### Competitive Advantage

The SWAMP facility is unique in offering vendor-neutral access to multiple software assurance tools, the automation and high throughput computing capacity needed to support continuous assurance, and an integrated viewer to display assessment results from multiple tools. Unlike similar offerings of no-cost software assessment services by commercial entities, the SWAMP is designed, built, operated, and supported by a partnership of four not-for-profit research institutions that have a long, demonstrated commitment to open-source, cybersecurity, privacy, and software assessment, and are driven by an underpinning vision of an open continuous software assurance framework that facilitates easy adoption of new software analysis technologies.

### Next Steps

- Start using the [SWAMP](#).
- [Learn more about continuous assurance.](#)
- [Subscribe to the SWAMP mailing list.](#)
- Integrate your software packages or analysis tools into the SWAMP by [contacting us](#).

The screenshot shows the SWAMP website homepage. At the top, there is a navigation bar with links for 'About', 'Contact', 'Resources', 'Policies', and 'Help', along with a 'Sign In' button. The main header features the SWAMP logo, which includes a gear icon with 'CONTINUOUS ASSURANCE' and the text 'SWAMP SOFTWARE ASSURANCE MARKETPLACE'. Below the logo is the tagline 'Do It Early. Do It Often.' To the right of the logo, there is a 'Welcome to the SWAMP' section with a brief description of the service and two 'Sign Up' buttons: 'Sign Up!' and 'Sign Up with GitHub!'. Below this, a section titled 'Get results in just three steps:' outlines the process: 1) Upload your package, 2) Run your assessment, and 3) View your results. Each step is accompanied by a small screenshot of the corresponding interface.

## SOFTWARE QUALITY ASSURANCE:

- ◎ **The Software Assurance Marketplace (SWAMP) Advancing continuous software assurance, open-source innovation, & cybersecurity**
- ◎ **Code Ray: Better software vulnerability management through hybrid application security testing**
- ◎ **CodeHawk Automated Malware Analyzer**
- ◎ **Hybrid Analysis Mapping: Software Assurance Enhancement Technology**
- ◎ **Tunable Information Flow (TIF): Policy-Driven Software Analysis & Assurance Toolset**



# Code Ray: Better software vulnerability management through hybrid application security testing

**Kenneth Prole**

**Kevin Green, Software Assurance Program Manager**

## Overview

Code Ray is a technology that combines the results of both static and dynamic application security testing. It highlights those software vulnerabilities that are both present in the source code and is exploitable by an external attacker who does not have access to the source code. Code Ray also maps the software vulnerabilities to industry standards to make it easier for users to find and fix the highest priority vulnerabilities first. Code Ray will be transitioned into the Code Dx software vulnerability discovery and management system as well as into the Department of Homeland Security (DHS) Software Assurance Marketplace (SWAMP). It will also be used in an educational version of Code Dx available for free to institutions teaching secure coding practices.

## Customer Need

Up to 90% of computer security incidents are traceable to vulnerabilities in software that were exploited by an attacker. To avoid this, software developers, testers and security analysts must run application security tests to discover the vulnerabilities before the attackers do. To find most of the vulnerabilities in an application, to achieve the the greatest “vulnerability coverage”, users must run several static source code analyzers as well as dynamic penetration testing tools and manual code analyses. Unfortunately, each tool presents its results in a different format and on different severity scales. It is very difficult and time consuming for users to create a consolidated set of results that show all the vulnerabilities in the source code, including which vulnerabilities are visible to an external attacker. It is also time consuming to prioritize the thousands of vulnerabilities that are typically found so that the most critical vulnerabilities are identified first and then corrected.

## Our Approach

Code Ray engages in Hybrid Application Security Testing (HAST) and first correlates and normalizes the output of dynamic application security testing (DAST) and static application security testing (SAST) tools, using runtime instrumentation and the Common Weakness Enumeration (CWE). The output is a consolidated set of results from all the tools, with duplicate results removed identifying which source code vulnerabilities are accessible by the end-users of the application software. The consolidated results are then mapped to industry standards such as the Open Web Application Security Project (OWASP) Top 10 and SysAdmin, Audit, Network, and Security (SANS) Top 25 to show which vulnerabilities are deemed most severe by these standards. Compliance auditors may perform additional analyses to identify those vulnerabilities relevant to regulations such as Health Insurance Portability and Accountability Act (HIPAA) or Payment Card Industry (PCI).

The specific workflow of Code Ray is as follows:

1. Code Ray monitors a running application during DAST using runtime instrumentation to store the precise execution path that produced DAST test results.
2. It uses those runtime instrumentation traces to map SAST result source locations to the observed execution paths.
3. The correlation is enhanced by normalizing the DAST and SAST results using the CWE as a common frame of reference.
4. Using the DAST-to-SAST merged results, Code Ray maps the correlated findings to selected industry standards and widely recognized compliance standards.
5. Finally, the results are displayed in a simplified risk management framework, with the ability for users to drill down for details.



# Code Ray: Better software vulnerability management through hybrid application security testing



## Benefits

Code Ray provides software developers, testers, security analysts and auditors with:

### **Better vulnerability coverage**

Detects more vulnerabilities by combining several analysis techniques.

### **Improved accuracy:**

Reduces false positive by confirming exploitability.

### **Easier prioritization**

Highlights vulnerabilities that are considered the most severe based on industry standards, or are associated with Health Insurance Portability and Accountability Act (HIPAA) or Payment Card Industry (PCI) compliance requirements.

### **Remediation guidance**

Pinpoints where to fix the code to remediate vulnerabilities.

### **Improved communication**

Visual interface fosters collaboration between development and security teams. Provides ability to filter findings into a summary of results for management.

## Competitive Advantage

Code Ray's consolidated results are easier and quicker to interpret than the current state of the art, which requires the sequential reviewing and mental correlations of disparate results from multiple application security testing techniques.

The transition path into Code Dx, which is an easy-to-use and affordable software vulnerability management system, jumps the hurdle that many small and medium-sized businesses (SMBs) face: time to learn and cost.

## Next Steps

Secure Decisions is transitioning the Code Ray technology into a software vulnerability management system called Code Dx, which is a commercially available and also accessible in the DHS SWAMP. After incorporation into the Code Dx product, a free educational version will be available to institutions engaged in teaching application security testing.

# CodeHawk Automated Malware Analyzer

**Henny Sipma**

**Kevin Green, Software Assurance Program Manager**

## Overview

The Kestrel Technology (KT) CodeHawk Automated Malware Analyzer (CHAMA) is a tool for semantic static analysis of malware using KT's abstract interpretation technology. The tool is easy to use. Invoked on any x86 PE executable, it performs a fully automatic deep semantic analysis of the executable and outputs a report that identifies significant data flows and type information, providing a detailed view of the behavior of the executable that can be used to extract IOC's (Indicators of Compromise) as well as semantic features for machine learning.

## Customer Need

CHAMA is targeted at several use cases:

- Public and Government entities responsible for identification and analysis of the latest malware used in critical attacks
- Commercial threat intelligence companies needing to augment their "sandbox" dynamic analysis with static analysis that identifies and analyses advanced threats and can analyze malware that doesn't execute in dynamic analysis environments
- Commercial companies needing advanced "semantic level" feature extraction to up-date machine learning/big data systems to identify advanced malware
- Service companies that analyze new advanced malware threats that require semantic level analysis of the malware
- Human analysts needing to automate malware analysis when using disassemblers like Interactive DisAssembler (IDA) Pro

## Our Approach

Using the CodeHawk abstract interpretation analysis engine, CHAMA uses intra and inter-procedural data propagation to collect information on:

- What data is retrieved from the system.
- What data is received from the network.
- What data is sent to the network.
- What actions are performed on the computer / peripherals.

The analysis results provide

- Host-based indicators (file names, registry keys, environment variables, etc.)
- Network-based indicators (IP addresses, domain names, etc.)
- Input indicators (likely input strings)
- Output indicators (format strings)

CHAMA assists in detecting suspicious activity by associating predefined and user-defined predicates to library functions that identify the call itself as suspicious or identify unusual combinations of arguments or flags.

## Benefits

CHAMA complements current malware identification and analysis tools by:

- Automating static analysis of malware that today requires skilled human effort: equivalent semantic information is automatically produced in minutes by CHAMA that may require hours to obtain when using tools like IDA Pro

## CodeHawk Automated Malware Analyzer

- Enabling machine learning based on semantic features to allow detection of complex sophisticated malware not detectable by existing malware machine learning tools based only on syntactic-level feature extraction.

CHAMA complements current malware identification and analysis tools by:

- Providing insight into sophisticated malware functionality using static analysis even if the malware will not execute in a “sand-box” environment or must be triggered by a specific event
- Providing detailed CISO level reporting of what data was accessed by malware and where it was sent to understand the damage done by the malware

### Next Steps

The CodeHawk Automated Malware Analyzer is now available as an easy to install/use automated tool. Parties interested in evaluating the tool should contact Kestral Technologies about using the tool under an evaluation license.

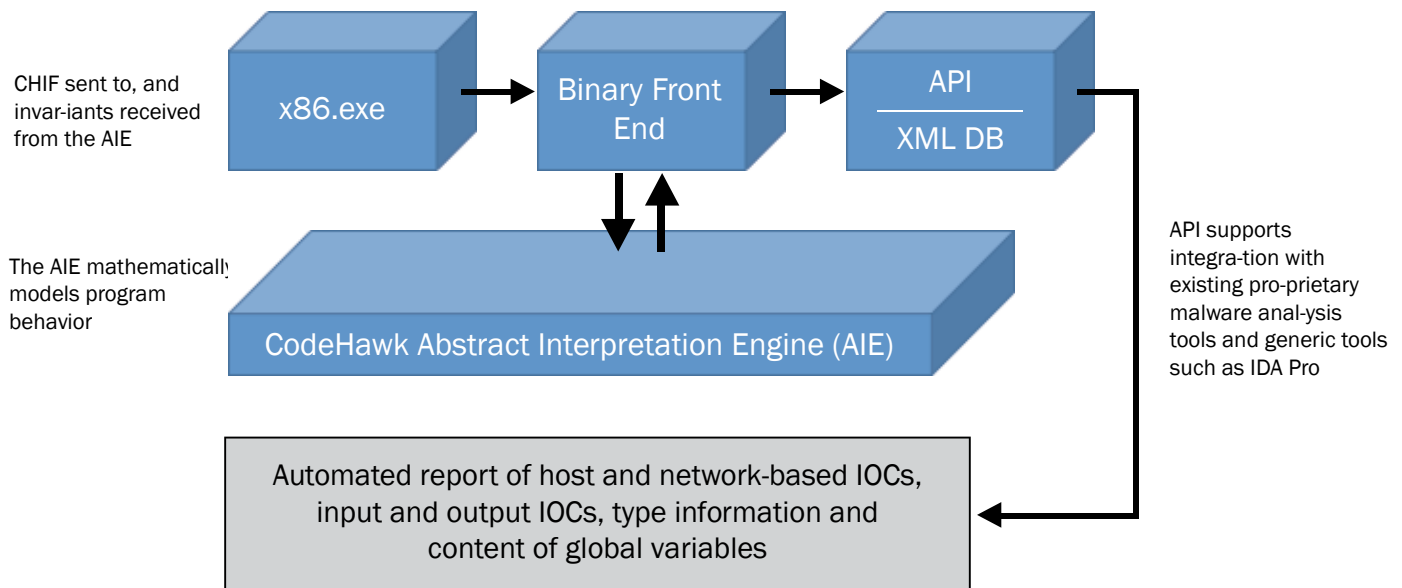
### Competitive Advantages

Existing tools for static analysis of malware provide syntactic analysis with no behavioral analysis to understand malware functionality. To overcome this

limitation, symbolic execution has been added to some malware static analysis tools but this technique has two critical limitations not present in abstract interpretation:

- Code coverage is often low, maybe 20% vs the 80% to 100% coverage achieved by KT’s tool, consequently symbolic execution may miss much of the malware functionality.
- Path explosion occurs when symbolically executing programs with large, complex functions, which are prevalent in modern malware.

Existing tools for dynamic analysis of malware cannot analyze malware that detects that it is run in a debugger or virtual machine and alters its behavior. Further, malware that is triggered by a specific event cannot be analyzed by dynamic analysis tools. CHAMA analysis does not have these limitations.





# Hybrid Analysis Mapping: Software Assurance Enhancement Technology

**Dan Cornell**

**Kevin Green, Software Assurance Program Manager**

## Overview

Organizations are at risk because of vulnerabilities that exist in the software systems that they develop and deploy. Static Analysis Security Testing (SAST) and Dynamic Analysis Security Testing (DAST) technologies can help software assurance teams identify potential weaknesses and vulnerabilities in software, but individual tools cannot provide sufficient coverage, requiring the use of multiple technologies. Denim Group's Hybrid Analysis Mapping (HAM) technology provides software assurance teams the ability to efficiently correlate the results of different security testing technologies to provide them with better insight into the security state of software systems.

## Customer Need

Organizations testing the security of software systems must use multiple technologies in order to obtain sufficient test coverage. Unfortunately, the results produced by these tools are provided in incompatible formats, leaving organizations with volumes of data requiring extensive manual analysis and correlation. Hybrid Analysis Mapping (HAM) allows security analysts to quickly and efficiently correlate and analyze this data to make decisions about vulnerability remediation. In addition, HAM can be used to increase the quality of security testing, resulting in superior test coverage and deeper analysis.

## Our Approach

When provided with application source code, Denim Group's Hybrid Analysis Mapping (HAM) technology works by detecting the language and application framework used by the application. Based on this, HAM creates an attack surface model of the application – identifying all URLs that the running application will respond to as well as all inputs that can change the behavior of the application. For each of these attack surface points, HAM tracks the location in

the application source responsible. Based on this attack surface model, HAM allows for the coordination of the attack surface of DAST scanning results with the source code location of SAST scanning results.

## Benefits

Both DAST and SAST scanning tools can provide software assurance teams with large volumes of data that are expensive and time consuming to properly triage. HAM allows these teams to get much more comprehensive testing results based on running multiple tools and to much more efficiently consume these results and turn them into actionable remediation recommendations.

In addition to correlation, this attack surface model can be used for additional operations that aid software assurance teams to increase the fidelity of security testing as well as to accelerate the remediation of identified vulnerabilities. Capabilities include:

- Pre-seeding DAST scans with comprehensive attack surface information, reducing false negatives in testing
- Mapping DAST scanning results to specific lines of source code, even in the absence of SAST results. These findings can then be mapped to development tools to aid developers in more efficiently identifying problem areas of code.

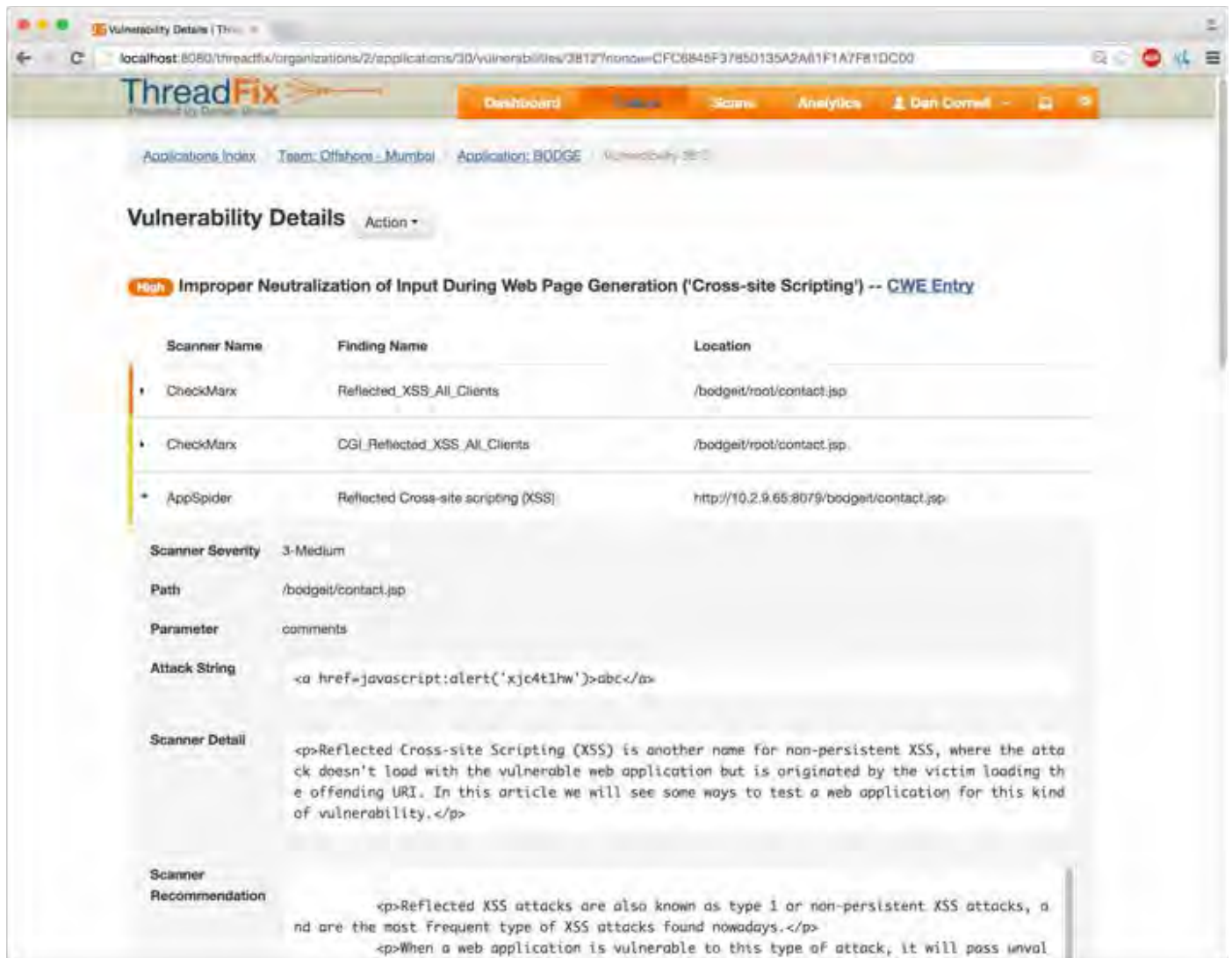
## Competitive Advantage

Denim Group includes its Hybrid Analysis Mapping (HAM) technology in its award-winning [ThreadFix](#) application vulnerability management platform. This allows software assurance teams to utilize HAM across their entire portfolio of applications. In addition, making aspects of the HAM technology available under an open source license has allowed the Denim Group to create a community of users and contributors that accelerates the adoption and evolution of the technology.

# Hybrid Analysis Mapping: Software Assurance Enhancement Technology

## Next Steps

HAM is currently available for pilot usage via the open source ThreadFix Community Edition and commercial ThreadFix Enterprise Edition software platforms.



# Tunable Information Flow (TIF): Policy-Driven Software Analysis & Assurance Toolset

Aleksey Nogin

Kevin Green, Software Assurance  
Program Manager

## Overview

HRL Laboratories, LLC's Tunable Information Flow (TIF) toolset is a flexible software supply chain mechanism for assuring information flow security. The TIF toolset allows enforcing information flow security policies in software through a flexible combination of high-precision static source code analysis and low-overhead inlined run-time monitoring.

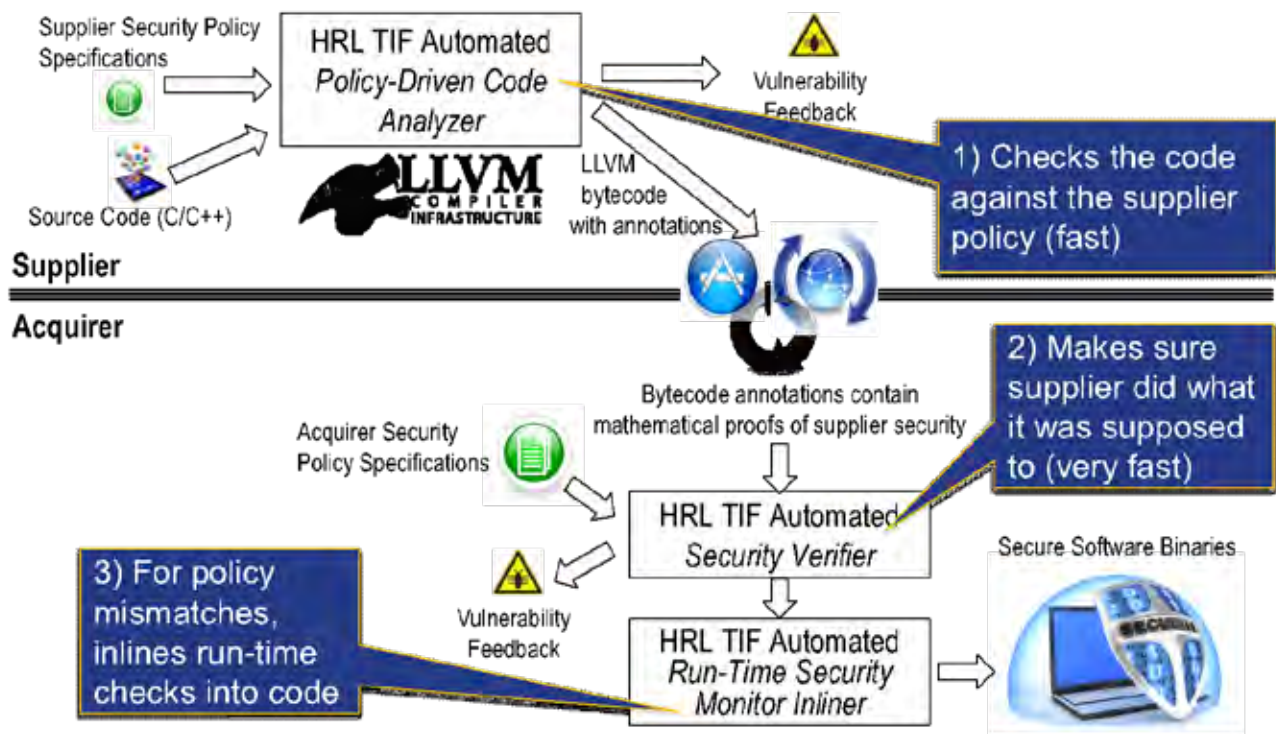
Developers of a software module would initially make assumptions on the information security properties of the module's environment. At a later time, the module is then integrated, the software is deployed, and the actual security properties of the module's environment could end up being fairly different. The resulting mismatch between the software and the security requirements would often result in security vulnerabilities.

## Customer Need

Ensuring a software implementation satisfies the intended security properties can be especially challenging in cases of long software supply process – whether a supply chain is stretching across organizations, within a single large organization, or simply extended in time.

## Our Approach

The TIF toolset is focusing on enforcing information security properties of software, which account for about 60% of the top Common Weakness Enumeration (CWE)/ Open Web Application Security Project (OWASP) vulnerability types. As one of its inputs, the TIF toolset





---

takes an information flow security policy file that specifies information integrity and confidentiality assumptions and requirements for the interface between the software being analyzed and its environment.

The TIF toolset is implemented as a set of extensions to the (Low Level Virtual Machine) LLVM compiler tool chain. It consists of three components:

1. TIF Analyzer is a static software analysis tool that can check whether the software source code satisfies the information flow policy. It leverages an existing Data Structure Analysis available for LLVM, and computes a range of possible security labels for each memory pool. The TIF Analyzer can run when software is compiled by the TIF-instrumented LLVM compiler, analyzing a single source code file. The TIF analyzer can also run during link-time optimization (LTO) pass of the compiler, for full-program analysis. TIF Analyzer can be configured to abort the compilation on detecting any policy violations, or to issue a warning, and defer handling of the error to TIF Inliner.
2. TIF Checker can quickly re-run the analysis to re-validate LLVM bytecode files previously created by the TIF Analyzer.
3. TIF Inliner inserts run-time checks for policy violations detected by the TIF Analyzer. It uses the ranges computed by the TIF Analyzer to only insert run-time checks where TIF Analyzer detected a potential violation, and only require additional run-time monitoring to those portions of the code where the information flow security consequences could not be fully determined statically.

## Benefits

To software acquirers: TIF policy-driven mechanisms allow acquirers to get software that satisfies their custom security needs – sometimes even when suppliers were not aware of those needs. TIF Checker also allows acquirers to quickly re-validate the software they acquire.

For software developers: TIF helps software developers rapidly adapt software to evolving security requirements. TIF policies can serve as a common language for specifying security requirements. TIF compiler integration helps ensure that software that is analyzed is the same software that is executed.

For tool developers: TIF LLVM integration allows us to take advantage of existing LLVM analyses, and provide a common environment for integration with other tools/analyses.

## Competitive Advantage

Current information flow tools tend to focus narrowly on taint analysis, typically with a single implicit taint policy. In contrast, the TIF toolset supports flexible policies and explicitly handles the evolution of policies over time. Most static analysis tools capable of capturing information flow requirements only capture explicit flow (through data), whereas TIF tools also track implicit flow (through control flow) and allow policies to state whether implicit-only violations should be allowed. Most static analysis tools suffer from poor precision, and most run-time monitoring approach suffer from high overhead; by choosing the combined hybrid approach, we are able to avoid both issues.

## Next Steps

We are currently evaluating and benchmarking the performance of the TIF toolset by performing an experimental security analysis of a suite of open source programs. We are also in the process of documenting our work and preparing several papers describing both the theoretical underpinning of TIF work and TIF implementation for publications.



Website



Email