

# Traffic Generation and Simulation in Cyber Security Research

Realistic traffic generation is an important input to cyber security research, providing grist for research and testing, without putting actual systems at risk, or inappropriately disclosing sensitive systems.

Traffic can be simulated or emulated, or real traffic captured and replayed.

Skaion has supported research on information assurance technologies for Intelligence Community and Defense Department R&D efforts, providing network traffic for testing.

Cyber security test beds now being established as synthetic network laboratories for attack and defense analysis will also need a means to create realistic traffic, if modeling is to address large-scale (even “Internet-scale”) simulations.

# The Need for Realistic Traffic

Artificial generation of realistic system activity and Internet traffic is essential to tools research, since security and privacy concerns restrict access to actual systems and networks.

The “ground truth” behind artificially-generated traffic is known, providing an answer key to evaluate systems under test.

Federal agencies addressing cyber security are constructing simulation infrastructure to support research and development. The DETER research testbed and Evaluation Methods for Internet Security Technology (EMIST) program (funded by the NSF and DHS) provide a means for simulated environments within which both attacks and defenses against them are studied without risk to the “open Internet.”

Supporting systems such as the TGS are needed to render these environments as realistic as possible.

# Emulation vs. Synthesis

*“The use of ‘complete traffic’ is necessary for realistic information assurance system testing. In general the issue of traffic generation is one of the most difficult ones to tackle. Synthetic traffic does not represent the realities of an actual network.” - Athanasiades, et al., “Intrusion Detection Testing and Benchmarking Methodologies,” 2003, First IEEE International Information Assurance Workshop*

Skaion’s approach to traffic generation is to create real traffic using synthetic users. While the network topology and population is virtual, every benign connection, probe, and attack is real.

Emulation should minimize transition time between test bed development, and real world deployment. Generating realistic traffic on testbeds may alert developers to potential problems long before deployment.

# Traffic Generation System (TGS)

(see large poster at right)

White team controls the simulation

Red hosts are malicious – they can be:

- background scanners and attackers
- live red team
- scripted attack hosts
- some combination

The Blue team defends a protected network, the large poster shows two sample Blue enclaves. Blue team hosts and networks report data sources to the blue team defender. (This reporting can be on-line or offline data sources collected onto DVD).

# Traffic Generation

## Simulated Users

- User-centric traffic generation
- Configured to match real traffic statistics
- Real traffic against real servers
- Real client software or scalable Skaion emulations

## Traffic Generation System Implementations

- Single host, Multi host  
Multi-host systems scale to hundreds or thousands of virtual systems
- Either Linux or Windows

## Can capture sensor information from a variety of logfiles:

- application logs
- host logfiles
- host instrumentation
- tcpdump files
- netflow records
- IDS records

# TGS WebBot

The **WebBot** is the most recent addition to the TGS architecture, and simulates human use of the World Wide Web.

Both HTTP and HTTPS protocols are supported. Each WebBot works from a “hot list,” and makes requests from it, or from outbound site links, to simulate human Web browsing behavior.

The WebBot pauses for random intervals when “viewing” Web content, and recovers from errors (e.g., “404 – page not found”) as appropriate, backing up and pursuing alternate links or “hot list” pages.

An ideal Web client simulation would show as much diversity in types of users and modes of use as is seen in the “real world.” Research programs such as the “Glass Box” might provide valuable insights:

*P. Cowley, L. Nowell, J. Scholtz, “Glass Box: An Instrumented Infrastructure for Supporting Human Interaction with Information,” 2005, HICSS*

# AFRL Hackfest

AFRL's Advanced Course in Engineering (ACE) summer Cyber Security Boot Camp culminates in the "Hackfest," a Red/Blue team exercise against a simulated network environment:

**<http://www.afrlhorizons.com/Briefs/Dec04/IF0408.html>**

The Skaion TGS was used to create simulated hosts to fill in the network topology, instead of only having just the four victim hosts populating the network. The TGS "cruft" bots also mimicked the background maliciousness from the Internet, so that there were scans and other IDS alarms not associated with the Red Team.

The TGS provided both simulated servers and clients, which influenced both Red Team and Blue Team activity.

The TGS is scheduled to be used again in support of the summer 2005 Hackfest exercise.

# Support to ARDA BAA “Information Assurance for the Intelligence Community”

Skaion is an awardee under **ARDA BAA BAA 03-03-FH** – Skaion’s role is to support other ARDA-funded principal investigators, delivering realistic data sets to describe notional IC environments for R&D and testing.

Skaion created a notional model of an Intelligence Community network with external access, given the IAIC focus on external threats. The network was patterned on the Open Source Information System (OSIS) interagency network. An overview of IC networks and threats for researchers is also available from Skaion by request, to qualified recipients (document is FOUO).

Skaion has thus far produced three data sets, escalating in complexity, incorporating both non-malicious background activity and attacks.

Producing sanitized descriptions of the nature and traffic of Intelligence Community systems and networks for less cleared/uncleared outside researchers is a challenge, and might be an area for fruitful research.

# ARDA Data Release Statistics

	<u>Data Set 1</u>	<u>Data Set 2</u>	<u>Data Set 3</u>
<b>Release Date</b>	June 2004	December 2004	Feb/Mar 2005
<b>Compressed Data Size</b>	1.1 gigabytes	12 gigabytes	11 gigabytes
<b>Background Runs</b>	2	6	1
<b>Distinct Multi-Stage Scenarios</b>	none	2	2
<b>Attack Runs</b>	4	15	6 total
<b>Attack Tools</b>	Metasploit	Metasploit, CORE Impact	Metasploit, CORE Impact
<b>Background Attacks</b>	6	12	13
<b>New Variants</b>	varying # of sources, fuzzers attack speeds	two anomaly	both labeled and unlabeled data
<b>Lessons Learned</b>	add complexity	improve quality control	detailed ground truth



# Small Business Innovation Research (SBIR)

Skaion has received an award under USAF SBIR 05.1, “Training Simulations for Decision Effectiveness in Computer Network Defensive Operations”

Information security training is sorely needed, yet difficult to provide because of scarce cybersecurity trainers. Automated training can fill this gap, but only if it can provide a realistic and flexible training environment.

Skaion’s proposal is to create a vendor-independent information assurance training environment on top of the TGS. By feeding realistic network traffic to an IDS, we can create training courses that could later be used for different COTS or GOTS systems.

Using the Skaion TGS as a framework, we will research the steps necessary to expand our system and network traffic simulation capabilities into an interactive cybersecurity training environment, initially for single-user blue-team training, but anticipating a comprehensive, multi-party blue- and red-team exercise.

# Presenter Information

## **Sam Gorton**

Worked on the 1998 DARPA IDS evaluation at MIT Lincoln Laboratory. Co-founder of the security tools vendor Sandstorm Enterprises. At Skaion, he simulates malicious activity on the testbed, both background activity and scenarios under test.

## **Skaion Principals:**

### **Terrence Champion, President**

Research scientist for ten years at the Air Force Research Laboratory's Information Security Office. His ideas for the test and evaluation of DARPA research into intrusion detection led to the development of a unique real-time testing capability.

### **Robert Durst, R&D Engineer**

Served as a U.S. Air Force officer at AFRL (Hanscom AFB), developing computer security product test techniques.

# Skaion Corporation

Skaion's expertise is in simulating real-world threats to information systems, modeling attacks, and evaluating defenses. Skaion's technologies synthesize realistic user and network behavior to test working systems, design new defenses, and conduct innovative research.

- Veteran-owned small business, established in 1999.
- Personnel and facilities cleared to SECRET.
- Providing support to federally-sponsored information security R&D efforts, for AFRL, ARDA and DARPA

Contact: Sam Gorton, [sam@skaion.com](mailto:sam@skaion.com)

Skaion Corporation  
51 Middlesex Street, Suite #114  
North Chelmsford, Massachusetts 01863  
Tel: (978) 251-3963  
Fax: (978) 418-9175

