Synthetic Cyber Environments for Training and Exercising Cyberspace Operations

Stephanie D. Harwell & Christopher M. Gore
Camber Corporation
O'Fallon, IL 62269
sharwell@camber.com, chgore@camber.com

ABSTRACT

To combat the cyberspace threat facing the nation, an integrated combination of technology, education, training, and exercising is needed. The Air Force cyber simulator journey began in 2001 with a small exercise. Today synthetic-live environments (cyber simulators) are in use for training and exercises, mission rehearsal, and tool development for cyberspace operations. The Air Force has 78 simulators at 3 locations in Illinois, Mississippi, and Florida. Solutions similar to the Air Force are also in use by the Navy (Navy Cyber Operations Range (NCOR) in Norfolk, US Strategic Command (STRATCOM), STRATCOM Cyber Operations Range (SCOR) in Nebraska, and the National Guard, Army Guard Enterprise Network Training Simulator (ARGENTS) in Arkansas and seven other States. In all, there are over 100 active simulators in the US. Evolving over time, the requirements of the cyber simulator have grown from just replicating the operational day-to-day environment of the blue force to modeling the environment of the red threat. The environment now encompasses a world-wide routable gray space and is interoperable with other synthetic environments.

Cyber simulators expose operators to various network situations and threats and advance their technical skills. They are used in validating solutions and the development of innovative approaches enhancing operational competencies. The risk-free environment of a cyber-simulator and scenario based stimuli allow crews to experience and conduct aggressive activities to: disrupt, obstruct, and destroy the integrity of the network; infiltrate a simulated computer network for intelligence collection; and train on procedures and tactics to defend and protect the network. Fidelity and realism throughout the physical and virtualized platform, appliances, and applications is paramount and must also be present in traffic generation, data, and the synthetic Internet. While these key factors are critical to an immersive experience, the simulator must be constructed within a rapidly reconstitutable environment with the capability to start, stop, and re-roll scenarios from a requisite state.

ABOUT THE AUTHORS

Stephanie D. Harwell, Vice President, Cyber Solutions of Camber Corporation, has been involved in the evolution of cyber simulators since Air Force's first implementation in 2001. Her experience includes designing, developing, and maintaining integrated synthetic-live cyber environments. Ms. Harwell is a former Air Force Communications Chief Master Sergeant, retiring in 2004, where she pioneered the concept of using simulators for training and exercising the Cyber Operations crew force. She manages the systems design, software development, and systems integration for Camber's Cyberoperations Enhanced Network and Training Simulator (CENTS®) product line. She holds a Master of Science in Information Technology and a Bachelor of Science in Information Systems.

Christopher M. Gore is a Research Scientist with Camber Corporation's Cyber Development and Integration team. He received his Master of Science in Computer Science from Missouri University of Science and Technology and his Bachelor of Science in Mathematics and Computer Science from Eastern Illinois University. His graduate research focused on evolutionary algorithms and their application to difficult problem domains in the financial arena. He has worked professionally on embedded avionics systems and cyber-operations systems for several years. His future plans involve applying methods from evolutionary computation towards problems in the cyber-operations arena.

Synthetic Cyber Environments for Training and Exercising Cyberspace Operations

Stephanie D. Harwell, Christopher M. Gore
Camber Corporation
O'Fallon, IL 62269
sharwell@camber.com, chgore@camber.com

INTRODUCTION

So how do you model or simulate cyberspace? Is the realm of cyber a venue for modeling and simulation? When taken to its root form, it is using a network of computers to model a network of computers. Adding to that, it is using virtualization and compression to simulate an environment that is already virtualized and compressed. For the cyber arena, the purpose of the model or simulator drives the composition.

For the Air Force (AF), the purpose of the cyber simulator is to:

- Assess and train defensive and offensive forces to decisively operate in cyberspace.
- Develop, validate and train rigorous, relevant and standardized cyber tactics and Command and Control (C2) procedures.
- Evaluate and refine information dissemination, Indicators and Warnings (I&W), and synchronization of U.S. computer network operations.
- Determine effectiveness and priority areas to refine cyber readiness and mitigate the full spectrum of rapidly-evolving threats and vulnerabilities.
- Provide simulator-based education, training, crew certification, mission rehearsal and exercise capabilities at the individual, crew position, unit and AF levels to ultimately increase AF cyber operations effectiveness.

The AF Cyber modeling and simulation objectives are:

- Provide realistic threat emulation
- Be interoperable with Modeling and Simulation (M&S) live-virtual-constructive environments
- Create a simulated environment to exercise fighting through a cyber attack
- Adapt to current threats (0-day)

The overall goal is *to provide the best training to the Cyber Network Ops community*.

The term cyberspace conjures up a vast virtual electronic universe that is increasingly becoming the center of our ability to exist in a modern world. The term denotes the Internet – an interwoven world of computer technology, networks, sensors, infrastructure, control mechanisms, processing end units and users. Cyberspace contains the information and the networks over which information is transmitted and on which digitized information is stored.

As critical as cyberspace is, cyber security is its potential Achilles Heel. Computer networks that are not properly protected with adequate security software, hardware and trained personnel are vulnerable to aggressive and malicious activities that can, at the very least, disrupt information flow. Establishing robust communications, computer networks, information assurance, and cyber security is more important now than ever before if we are to protect the vital networks that play such a critical role in achieving national security, economic independence, and secure and organized daily lives. Effective cyber operations must be employed and managed by professionals who are well versed in protecting their networks and have a firm understanding of security policy and procedures and the tactics and tools of the cyberspace adversary.

Commercial certifications and vendor product courses will never be able to teach the integrated solution of: people/communication, processes/tactics, and the Service technology set. Acquiring and honing this type of skill can only be done when the cyber operator is immersed in a training environment that provides:

- The cyber weapons in the operational arsenal
- Realism and stressors of the watch floor
- Exposure to repeatable events with realistic effects

Conducting training and exercises in a risk-free environment is paramount. A risk-free environment ensures each individual has the freedom to explore without fear of catastrophic system failure or a security breach to operational systems.

ORIGINS

In 2002 the AF conducted a "first of its kind" computer network defense exercise called Black Demon. The AF wanted to develop tactics for responding to a large-scale computer network attack and provide the network defender their first 10 cyber warfare combat "sorties". The focus of the initial exercise was on developing tactics, techniques, and procedures for reconnaissance, insider threat, web defacement, viruses, intrusion detection and other malicious threats. Ancillary objectives included improving network operator situational awareness, response to multiple threats, and network defense reconfiguration.

The exercise was conducted on a first-generation (simple) network simulator (referred to as the range) designed to emulate the operational AF network. Components were borrowed from wherever they could be found (bench stock, test networks, programs) and software was acquired from the program office or trial licenses were used. It provided a fairly realistic training environment for network defenders and gave them the ability to interact with other participants. However, there were many shortcomings:

- No configuration control between the ranges so each of the four solutions was slightly different
- Network traffic to mask the activities of the red team (attackers) was nominal
- Resetting the simulator took hours
- Exercise inter-connectivity was constrained to a 56K (serial) VPN connection at the player locations. (This was the approved solution to preclude network saturation and "spillage" of attack events onto the operational network)

Despite the range environment shortcomings, the After Action Report (AAR) from Exercise Black Demon 2002 praised the exercise and recommended that the AF develop a permanent environment. The range would provide a risk free environment where network operators can continuously exercise and practice their skills and develop additional tactics to defend against cyber threats. This recommendation generated the original requirements for what is now the AF Simulator Training and Exercises (SIMTEX) program.

In 2003, the AF followed up on their Black Demon successes and developed the SIMTEX network which was first used for quarterly training exercises. The training events generally focused on providing operational training on new or specific network operations or defense tools used throughout the AF. For the 2004 Black Demon event, the AF unveiled a standardized simulator suite –SIMTEX– to be used for exercises modeled after its network core, the Combat Information Transport System (CITS) (Figure 1). The SIMTEX network has been used to support training exercises, operational exercises, and Joint network exercises over the past nine years. Thousands of cyber operators have participated and been trained on the latest cyber defense tools, tactics, techniques, and procedures (TTPs), cyber C2, and current threat signatures utilizing SIMTEX. Using the SIMTEX network and AF Combat Training Exercises, AF cyber operators receive practical experience on the cyber battlefield – their "first 10 combat sorties" in network defense; exposure to real-world threats, and training on cyber C2 processes.

The lack of professionally trained cyber operators led the AF to recognize the need to increase the available avenues for simulator training. SIMTEX provided the solution through the use of scenario based training within the synthetic environment and increased the numbers of those trained while greatly improving retention of material taught. The AF has placed variants of its SIMTEX simulators in formal school houses at Keesler AFB, Mississippi and Hurlburt Field, Florida, for Communications/Cyber Operations, Undergraduate Cyber Training, and Defensive Counter Cyber (Intermediate Network Warfare Training) courses.

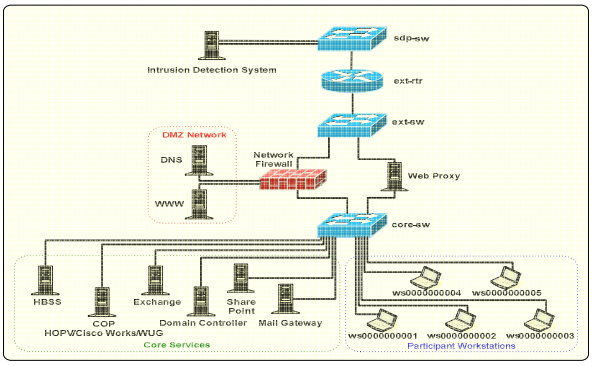

Figure 1. Notional Simulator Architecture

Even with SIMTEX as the standard solution, its routine use in exercises identified shortcoming and new requirements:

- Realistic network traffic to obfuscate attacks was lacking
- Network traffic produced by the traffic generators did not have payloads that triggered alerts from the network security devices
- An automated means for executing events across the entire interconnected range was needed; use of Information Warfare Squadron personnel to execute the attacks was very expensive
- A status window showing scenario execution status was needed

- Rapid simulator reconstitution capability needed to be developed so scenarios could be quickly re-rolled
- The serial connection between the ranges limited training realism; an approved Wide-Area-Network (WAN) Virtual Private Network (VPN) for world-wide interconnections was needed

Over the past nine years, through technology advancements and lessons learned, SIMTEX has evolved into an interoperable network environment based on an open-systems architecture that includes physical, virtual, and simulated network components. SIMTEX models the architecture of the AF enterprise network and has expanded to include wide-area-network connectivity through the Joint Cyber Operations Range (JCOR) VPN for Joint and Inter-Service exercises and training. Through the JCOR VPN, SIMTEX connects to other Service and Combatant Command (COCOM) cyber simulators and ranges.

SIMTEX's synthetic environment is provided through the commercial application SLAM-R® (Sentinel-Legion-AutoBuild-Myrmidon-Reconstitution.) The SLAM-R© application provides:

- IEEE Request For Comment (RFC) compliant real-world network traffic
- Over 3000 simulated users
- An attack manager
- Simulated network events (attacks) within the simulated real-world traffic
- Social media services (comparable to Facebook®/Twitter®)
- A simulated Internet
- Reconstitution capability

The integration of commercial applications, appliances, and infrastructure (either virtualized or physical) and SLAM-R© provide a synthetic-live simulator/trainer. The integration of the commercial products with SLAM-R© results in true-life system response either from user actions or from the attacks/events. As in real-world operations, user actions can impact the simulator's network (for example, a self-inflicted denial of service). AF cyber operators and decision makers (as well as other Services, Joint, and 5-Eyes) utilize the SIMTEX risk-free environment for classroom training, small and large-scale exercises, team competitions, tool development, and mission rehearsal.

LESSONS LEARNED AND INNOVATION LEAD TO EVOLUTION

Governments, corporations, small businesses, and individuals spend millions of dollars annually for commercial training programs and vendor courses. These programs and courses, while teaching industry best practices, accepted standards, and tips and tricks of vendor products, are not enough. To truly be considered an expert in the cyber defense community, cyber operators not only need to know how to use specific products according to vendor guidelines, they must know how their capabilities, when intertwined within their network, provide integrated situational awareness.

The Network Environment–Physical or Virtualized

Many a Masters thesis has been written extolling the virtues of virtualization and how it can be used to create a cyber simulator/trainer with a small footprint at a low cost. In theory this is true, if the goal is to create a "generic" cyber environment to only teach basic principles.

The AF cyber simulators provide network professionals opportunities to practice classroom learning in a realistic environment that does not impact any operational network. The simulator provides the participants with the same "look and touch" of the computer network environment they manage and defend day-to-day.

At the start of the AF program in 2001, virtualization was not as evolved as it is today. Each core service application, infrastructure device, or security appliance was a physical server or device in the simulator. The typical "base" solution filled a 42Unit (U) rack. Today, with virtualization, that same simulator is still approximately 9U even though the servers for the solution only take up 2-3U. This is because not all of the infrastructure and security devices in the AF network come in a virtual appliance. The core services that are virtualized are rapidly reconstitutable. An entire simulator's system baseline can be restored in less than 10 minutes.

For the sake of a smaller footprint and being able to snapshot all components of the simulator, advocates of new cyber simulators entering this arena are encouraging adopting a simulator that is completely virtualized. Air Force lessons learned point toward a different solution – a hybrid of virtualized machines and hardware-in-the-loop. For training/exercising operational forces, replacing brand-name physical devices (loaded with proprietary operating systems) that cannot be virtualized with available open-source virtual devices falls short of meeting the requirements

the cyber simulator program evolved from – *train like you fight*.

Attack Engine

During the early stages of cyber exercises, attacks were executed by members of information warfare squadrons (the red cell). In late 2003 the AF decided to put a SIMTEX type capability in the Communications school house. Having live players (red aggressors) physically execute each attack wasn't cost effective or feasible. This meant that an alternative solution to the way attacks were delivered to participants had to be developed – an automated method. The attacks/events students would be exposed to had to execute the exact same way for each student for each class. This generated the initial need for simulated attackers – the attack engine.

The attack engine used in SIMTEX (the Myrmidon module) generates network attacks within the simulator's network environment. The core includes a module configured for creating one or more attack events against the network devices (physical or virtualized). Individual attack events are grouped into scenarios. The attack events include exploitations of published vulnerabilities (e.g. SANS Top 10) and failures of hardware and software within the simulator. Scenarios are also created to replicate actual occurrences that affected AF operations.

As a result of the comments received from the '07 Bulwark Defender, a graphical user interface (GUI) was developed. As the scenarios got more complex, controllers required insight into the execution status of each attack/event in a scenario. Further, controllers needed a quick view of the details of each attack/event (description of the attack, objective of the attack, system indications and warnings for the event, and attacker and target information.)

The GUI provides the interface for controlling and monitoring the creation and execution of the attack events (Figure 2). The GUI includes an attack event editor configured for writing the attack events into a standard XML file, and wherein the control module is configured for automatically generating unique attributes within each attack event. Attributes include: the source of the attack (both Internet Protocol (IP) and Media Access Control (MAC) address), the attack target, how long into the scenario should the attack start, and how long the attack should run.

The early attack engine required the controller or instructor to be at the keyboard/mouse to execute an event. Exercise participants and students keyed in on

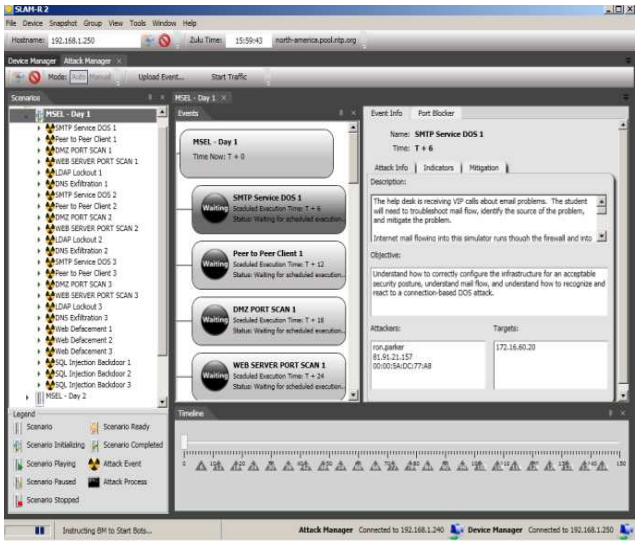

Figure 2. GUI Displaying the Attack Engine

the controller/instructor's position at the keyboard in relation to when events would occur. As a result, the capability for the events in a scenario to auto-execute on a timeline was added. Controllers/instructors can start, stop, pause, re-roll and adjust the event kickoff time on the timeline within a scenario.

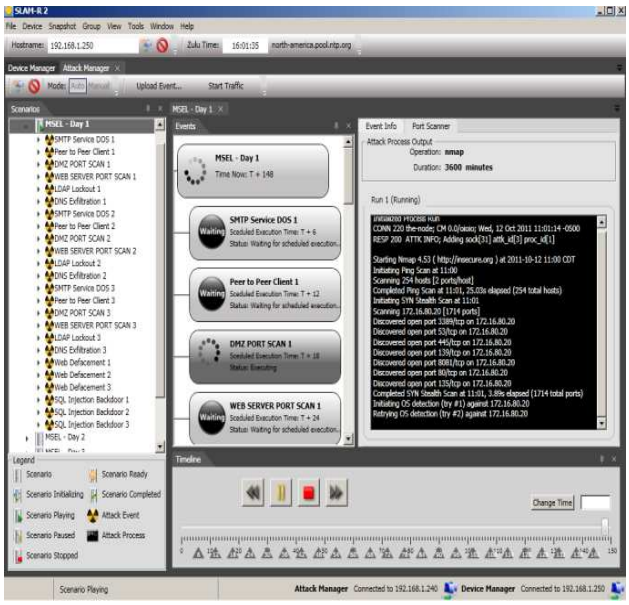

Figure 3. GUI Displaying Attack Status

To show the status of attack/event, an attack scenario execution manager tab populates when an event starts. At execution, a bot server module is generated within a bot of the simulator utilizing at least one of the created attack events. The execution module is configured for monitoring the creation and transmission of the attack events including the success of the attack event within the simulator and attributes of the attack event (Figure

3). This information is relayed back to controller/ instructor via the bot to the control window.

Network Traffic

When the standardized simulator suite was being designed one of the requirements was for traffic generation. The purpose of the traffic was to mask the activities of the adversaries within a "normalized" traffic flow representative of an installation's day-to-day traffic pattern. Over the course of five years, the Air Force integrated two different commercial traffic generators in an effort to populate the simulator with realistic cyber operations traffic. The available solutions were not satisfying the "realistic" requirements. The selected solutions were fashioned for performance testing and did not generate RFC compliant packets to the degree that the network devices (firewall, intrusion detection system, proxy server) were able to inspect the packets (deep packet inspection). This shortfall meant the security devices and applications did not throw the correct indicators and warnings. Network traffic, representative of day-to-day activity that was RFC compliant and attributable to the simulators domain (source and/or destination IP) was illusive. At the time a cost effective solution providing traffic that met the requirements for the cyber simulator wasn't found. This led to the development of a traffic generation capability focused on producing cyber effects.

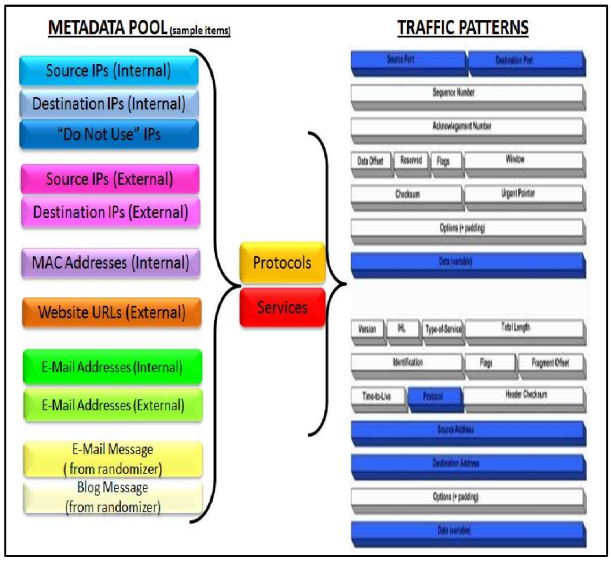

Figure 4. Sample Traffic Metadata Types

The traffic generator in SIMTEX (the Legion module) creates network traffic patterns within the simulator replicating actual network traffic patterns within the AF's operational network environment. The created patterns generate network traffic between a plurality of network devices within the simulator (router to router, router to server, server to server, server to workstation, workstation to server). The module is configured for creating a network traffic agent utilizing one or more of the patterns. The traffic agent is a group of one or more patterns. (Figure 4).

The module is further configured for creating a traffic scenario that includes a group of created agents and traffic scenario virtual machines (VM). The VMs act as senders and receivers of packets (patterns) defining a relationship between one or more of the agents in the scenario (Figure 5). An interface is configured for:

- receiving pattern metadata and adding the received metadata to the associated patterns
- adding the patterns to the traffic profile
- generating the scenario VMs and adding the VMs to the traffic scenario

The incorporated network traffic patterns include one or more network traffic protocols selected from the group. (e.g. Domain Name Service (DNS) requests and DNS responses, Hyper Text Transfer Protocol (HTTP) and HTTPS secure (HTTPS) requests and responses, Simple Mail Transfer Protocol (SMTP) send and SMTP receive, Internet Control Message Protocol (ICMP), Transmission Control Protocol (TCP), and various Remote Procedures Calls (RPC)). The result is network traffic that has source/destination IP addresses, valid e-mail addresses, is RFC compliant, has valid data payloads, and has IP addresses or URLs that resolve with the simulator's DNS structure. Feedback from participants operating the AF's Information Operations Platform (IOP) during the May '12 Global Lightening exercise was that Legion's traffic is the most realistic they had ever seen.

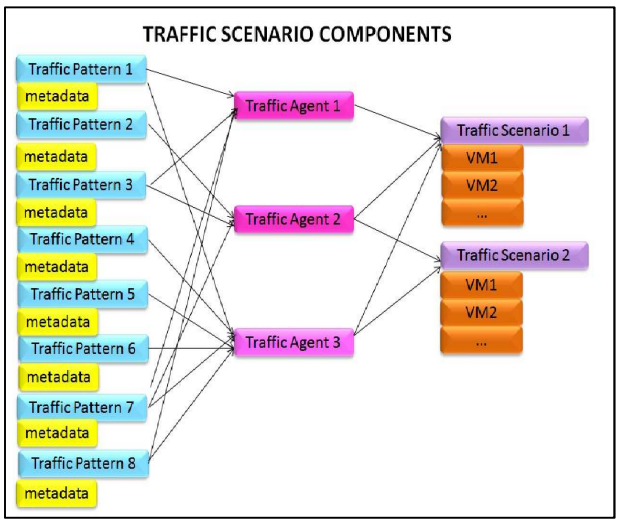

Figure 5. Traffic Scenario Elements

An example algorithm for generating these patterns is: A user enters the names of 100 different web sites. The user then selects an integer which can be used as input for the level of variance between the basic traffic patterns. A mathematical algorithm is then applied, producing a sequence of number pairs such that they represent the web sites surfed to and the length of time in seconds until the next pair is to be read. Take [23,30:99,13:40] means the 23rd web site is surfed to immediately, then the 99th web site is surfed to 30 seconds later, and then the 40th web site is surfed to 13 seconds later. In this example, the list of 100 different web sites and the integer for variance provides the specified criteria. Based on the requirements of the attack/event scenarios in the exercise or training, the individual traffic scenarios are created by applying the algorithm.

The Internet and Beyond

Bulwark Defender '08 saw the inclusion of robust HTTP traffic added to the simulator and web defacement exploits added to the available attacks/events. A lesson learned from this event generated the requirement for root (tier 1) DNS services. Although the HTTP traffic was realistic, the URLs being outside of the local simulator structure were generating errors because there was nowhere to get the A records. A requirement for a simulated Internet of websites followed. The simulated web-site Internet is "surfable" by all participants, all web-site URLs resolved in DNS, and generated HTTP traffic (both inbound and outbound) has actual source and destination points (Figure 6).

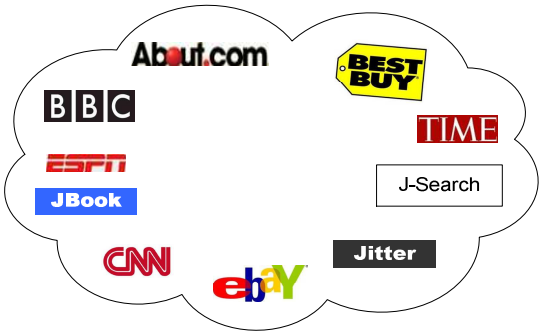

Figure 6. Simulated Internet

In 2009, the requirements for the European Command (EUCOM) exercise Austere Challenge, dictated more realistic adversaries, targets, and launching platforms. With botnets compromising "innocent victim" machines around the world, a world-wide botnet attack scenario was detailed in the exercise requirements. To complement the current SIMTEX simulators/JCOR, a simulated Range Global Internet (RGI), a Synthetic Non-Kinetic Bombing Range of sorts, was designed and implemented (Figure 7.)

The RGI provides a look and feel comparable to the actual internet. It provides for controlled and secure training scenarios outside of the public realm. The RGI is completely virtualized, using open source utilities where possible, and utilizes real IP addresses found in the global Internet structure.

The RGI is made up of 30+ backbone routers, with more than 150 class C subnets, supporting 150+ domestic and international web-sites and 35 fully functional e-mail servers along with global DNS and Network Time Protocol (NTP) services. J-Services provides social media services ranging from domestic to foreign personal blogs, and Facebook® and Twitter® like services. The RGI also includes RFC compliant Internet traffic-generation providing routine traffic activities between Internet routers, DNS queries to actual servers, website "GET" request, e-mail generation, along with other miscellaneous random traffic (e.g. ICMP). The RGI is comprised of four (4) interconnected networks spread across six (6) continents. Multiple location types are represented around the globe: hospitals, banks, universities, cyber cafés, commercial business, churches, government entities, and the military. The locations have full domain services and defense in depth construction.

With the true-IP global routing infrastructure in place and various location types populating the subnets around the world, exercise engineers built out the gray-space locations for Austere Challenge. During the exercise, a trace of an attack showed gray-space machines (and victim machines of the botnet) virtually located around the world.

FUTURE WORK

Modeling and simulation work in the area of cyber is still in its infancy. The is a large need for additional capabilities and interconnections for synthetic-live Cyber environments

SCADA

It can be assumed that any major engagement in the near-term with a capable opponent will involve a major component in the cyber arena. It can also be assumed that one of the main targets within the cyber arena for any such opponent will be our critical infrastructure and industrial control systems. Therefore, it is vital that we train a new type of cyber-defender specializing in their defense. Integrating this capability into SIMTEX/JCOR and the RGI is a logical next step.

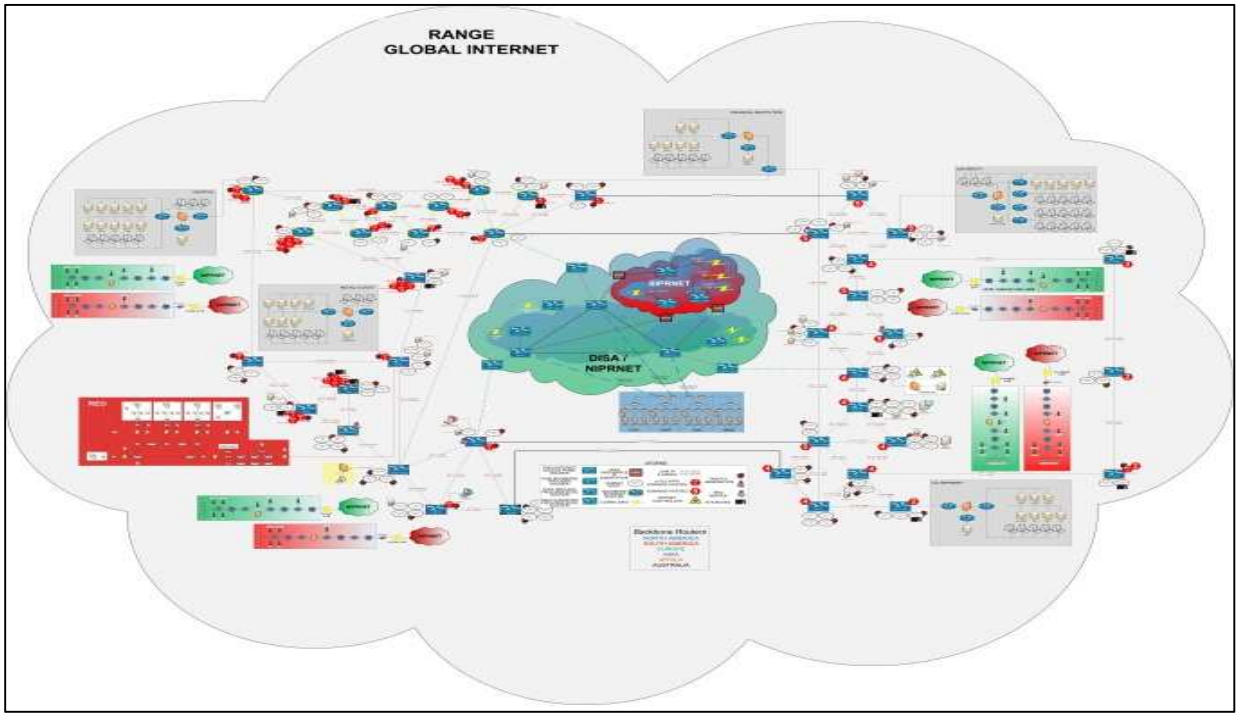

Figure 7. Range Global Internet

Supervisory Control and Data Acquisition, SCADA, refers to industrial control systems (ICSs) that monitor and control industrial, infrastructure, and facility-based processes. SCADA systems are used to monitor and control a plant or equipment in industries such as telecommunications, water and waste control, energy, oil and gas refining and transportation. These systems encompass the transfer of data between a SCADA central host computer and a number of Remote Terminal Units (RTUs) and/or Programmable Logic Controllers (PLCs), and the central host and the operator terminals.

The current SCADA master station architecture is an open system architecture rather than a vendor controlled, proprietary environment. The architecture consists of multiple networked systems sharing master station functions. While there are still RTUs utilizing protocols that are vendor-proprietary, it opens the system architecture, utilizing open standards and protocols and making it possible to distribute SCADA functionality across a WAN and not just a LAN. With critical infrastructure control systems existing on WANs, connected through the public Internet, the threat of remote disruption by hostile agents moves out of the arena of science fiction and into reality.

The first publically-acknowledged real-world example of a government-sponsored attack against critical infrastructure was Stuxnet, a computer worm first discovered in June 2010. Stuxnet targeted Siemens industrial software and equipment, reprogramming PLCs and disrupting the Iranian uranium enrichment infrastructure.

While such sophisticated systems generally require specialized knowledge and are difficult to produce without state support, we can fully expect such systems to be developed and used against our critical infrastructure systems, and therefore we need to be equipped to defend against them.

We cannot expect defenders to be adequately trained to defend modern critical infrastructure from attacks if they are not exposed to realistic training systems. In much the same way that aircraft pilots are trained first in on-the-ground aircraft simulators and then in real aircraft, the need for extremely realistic advanced training systems cannot be overstated. While simulation is useful in the earlier stages of training, these defenders need to be trained on the real-world SCADA systems controlling simulated infrastructure instead of actual real-world infrastructure. Fully integrating SCADA and simulated infrastructure into cyber simulator environments will be critical for our defense in the near future.

Evolutionary Process for Automatic Scenario Generation

One of the key problems with training is that exercises cannot be easily repeated by the same student, since they will have previous knowledge of the problem space from their previous run. It is unrealistic to manually generate a new environment for them on multiple occasions, but if we can generate their problem environment computationally then the student could be continually exposed to similar situations repeatedly and therefore develop a deeper understanding of the methods for defending their systems.

A typical educational program aimed at young children learning arithmetic is a good example of this idea. Such a system might ask the child to determine the result of 12/55 one time and 7/43 the next. It does not have a large store of predetermined division problems but instead randomly generates those problems for the student to answer. In a similar, albeit vastly more complicated manner, it is feasible using concepts from evolutionary algorithms (EA) to computationally generate useful training environments for cyber operations.

In artificial intelligence (AI), EA's are a style of generic population-based meta-heuristic optimization algorithms whose processes are inspired by those of natural biological evolution (Figure 8).

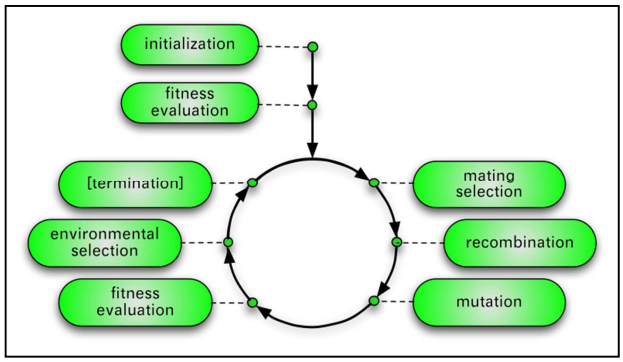

Figure 8. Evolutionary Process

The primary mechanisms employed in EA's to evolve a population of possible solutions towards an optimal one are:

- parent selection based on fitness
- recombination
- mutation
- survivor selection based on fitness

Evolution serves as a powerful metaphor and demonstrates great creativity in both the natural world and in the world of computer science.

A learning classifier system (LCS) is an EA that operates on a population comprised of rules referred to as the rule set: this rule set is used to attempt to classify a situation. The first LCS was created shortly after genetic algorithms (GA) were created and is considered one of the classical types of evolutionary algorithms. Since then there have been several improvements to the field.

Genetic programming (GP) is an EA-based methodology to find computer programs that perform a user-defined task. GP is a specialization of genetic algorithms (GA) where each individual is a computer program, either partial or complete. It is a machine learning technique used to develop and optimize a population of computer programs according to a fitness landscape determined by a program's ability to perform a given computational task. Techniques derived from GP could be applied to the domain of generating training environments. Many seemingly different problems in artificial intelligence, symbolic processing, and machine learning can be viewed as requiring discovery of a computer program that produces some desired output for particular inputs. When viewed in this way, the process of solving these problems becomes equivalent to searching a space of possible computer programs for the "best fit" individual computer program.

There exists the potential to add computational opponents to the training simulation by employing GP. In a game, there are two or more independently-acting players who make choices (moves) and receive a payoff based on the choices they make. A "strategy" for a given player in a game is a way of specifying what choice (move) the player is to make at a particular point in the game from all the allowable moves at that time and given all the information about the state of the game that is available to the player at that time. Strategies for games may be expressed in several different ways, even in terms of the state of the game or in terms of various features abstracted from the state of the game. By abstracting the simulation state space, we could then use that abstracted representation as a basis for evolving computational opponents. These opponents might be defensive, defending their systems from the human student. The opponents may also be offensive, attacking a network that the human student is trying to protect. Both of these scenarios would be useful for training.

Using LCS and GP to computationally generate an attack scenario based on previous responses of the cyber operator would allow for cyber-operations training and simulation as a game that could compliment the human-driven environment. It is well-known that serious games provide some of the best training methods available. Game-based learning

(GBL) is a branch of serious games that deals with applications that have defined learning outcomes. GBL has the potential of improving training activities and initiatives by virtue of its engagement, motivation, role playing, and repeatability. Integrating serious gaming via GP into a cyber simulator would potentially be of great value allowing for automatic scenario generation based on the skill/progress of the players.

CONCLUSION

The AF, either in AF only exercises/ events/competitions or in Joint activities has had tremendous success with SIMTEX and JCOR. The consistent feedback from Blue Force operators, students, and Senior Leaders is that the cyber range is a must have commodity. Being on the simulator/range where they are challenged to fight through the attack with the toolset they have available is invaluable.

Depth and breadth of knowledge is required by cyber crews to understand the technically complicated infrastructure and network "system of systems" selected by program offices. Managing and defending it against an ever-increasing number of highly motivated adversaries only comes from using a hands-on training environment comprised of the components used in daily operations so theory can be put into practice.

Much of what cyber operators do is intuitive. Constant exercise of those thought processes provides the continued skill level improvements and innovative approaches needed to stay ahead of the technical problems and hostile activities. Doing this in an environment that does anything other than truly replicate the cyber operator's environment (or the adversaries) falls short of satisfying the goal: achieving and maintaining a cyber security posture for our critical national computer network infrastructure. Training and exercising with a synthetic-live cyber environment provides a foundation for ensuring our critical infrastructure is adequately protected from any and all deliberate attacks and provides the information and mission assurance expected and needed by all levels of leadership.

REFERENCES

Air Force Network Integration Center (2011). *Cyber Simulator Requirements*.

Andel, T., Stewart, K., Humphries, J. (2010). Using Virtualization for Cyber Security Education and Experimentation. *14th Colloquium for Information System Security Education (CISSE).*

Boyd, Marcus, Colonel, USAF (2012). Air Force Cyberspace Modeling & Simulation (M&S). *AFCEA-Orlando: Air Force LVC Operational Training for the Cyber Warrior*

Corrin, Amber (2012). Seismic Shift Occurs in Air Force Cyber Planning. *Defense Systems*, Vol 6, Number 2, page14.

Eiben, A. E., Smith, J.E. (2003). Introduction to *Evolutionary Computing.* Springer-Verlag.

Falliere, N. (2010). Exploring Stuxnets PLC Infection Process. Retrieved April 12, 2012, from http://www.symantec.com/connect/blogs/exploring-stuxnet-s-plc-infection-process

Gilmore, J. Michael, Director, DOT&E (2011). Information Assurance (IA) and Interoperability (IOP). *Director, Operational Test and Evaluation FY 2011 Annual Report*, pages 285-291.

Hansen, Andrew P., Major, USAF (2008). Cyber Flag A Realistic Cyberspace Training Construct. *Masters Thesis.* Wright-Patterson AFB, Ohio: Air Force Institute of Technology.

Institut für Neuroinformatik, Ruhr-Universität Bochum (2012). *Shark – EALib Documentation.* Retrieved April 21, 2012, from http://shark-project.sourceforge.net/2.1.2/doc/EALib/index.html

Koza, J. R. (1990). Genetic programming: A Paradigm for Genetically Breeding Populations of Computer Programs to Solve Problems. *Technical report*. Stanford, CA: *Computer Science Department, Stanford University*.

Liljenstam, M., Liu, J., Nicol, D., Yougu, Y., Uan, G., Grier, C., (2006). RINSE: the Real-time Immersive Network Simulation Environment for Network Security Exercises. *Simulation*. Volume 82 Issue 1. San Diego, CA: Society for Computer Simulation International.

McBride, Aaron. (2007). Air Force Cyber Warfare Training. *Defense Standardization Program Journal*, pages 9-13, April/June 2007.

National Communications System (2004). NCS TIB 04-1. *National Communications System Technical Information Bulletin 04-1: Supervisory Control and Data Acquisition (SCADA) Systems*. Arlington, VA: Office of the Manager, National Communications System

Network Centric Operations Industry Consortium (NCOIC) (2010). *Net-Centric Cyber Simulator Capability Pattern*.

Sanger, David E. (2012). Obama Order Sped Up Wave of Cyberattacks Against Iran. *The New York Times*, June 1, 2012, retrieved from http://www.nytimes.com/2012/06/01/world/middleeast/obama-ordered-wave-of-cyberattacks-against-iran.html?hp&pagewanted=all

Susi, T., Johannesson, M., and Backlund, P. (2007). Serious games – an overview. *Technical Report.* School of Humanities and Informatics, University of Skovde, Sweden.