ACSC/NEACE, AY15

AIR COMMAND AND STAFF COLLEGE

AIR UNIVERSITY

Measuring Cyber Operations Effectiveness

by

Denney L. Neace, Major, USAFR

A Research Report Submitted to the Faculty

In Partial Fulfillment of the Graduation Requirements for the Degree of

MASTER OF OPERATIONAL ARTS AND SCIENCES

Advisor: Major Marc Flores

Maxwell Air Force Base, Alabama

Nov 2014

Disclaimer

The views expressed in this academic research paper are those of the author(s) and do not reflect the official policy or position of the US government or the Department of Defense. In accordance with Air Force Instruction 51-303, it is not copyrighted, but is the property of the United States government.

Abstract

This paper outlines the various methods currently used to show the effectiveness of Network Operations, Cyber Defensive, and Offensive operations. Network operations effectiveness depends on the view of the person measuring it. It is a delicate balance of usability versus security with the mission of the network providing a guide. Network Defense can be measured using many automated tools, included in the defense hardware and software itself. These devices include hardware and software firewalls, Network Intrusion Detection and Prevention Systems. Security Information Event Management software allow the ingestion of all of the above system data into a single cohesive picture better able to detect advanced threats. The effectiveness of these devices and software can be measured through auditing using network scanners and frameworks like the SANS Critical Security Controls. Finally, a network vulnerability penetration test can be performed to test the systems and controls put in place to ensure they are working as designed. Penetration tests should use a framework like the Penetration Test Execution Standard in order to provide standardized and reproducible results. Measuring success in Offensive and Computer Network Exploitation depends on the goal. If the goal is to exfiltrate information and it was obtained without being detected the operation was a success. Attacks to degrade or harm systems can be measured if the end goal of the attacker can be determined.

Measuring Cyber Operations Effectiveness

It is human nature to wonder how effectively you are doing your job, playing the game, completing your work. Depending on what you are trying to measure it can be visually easy to distinguish success from failure. Military operations can be easy to visually determine success due to ISR identifying the target area and battle damage or front lines. Planners can see the battle damage assessment and where their lines are in relation to the enemy and know what they need to do next. Cyber operations do not have as many visual queues of success or failure. Effectiveness must be determined through complex programs or interdependencies between defensive measures. Determining the right combination to measure effectiveness in network, defensive, and offensive operations is a difficult task. Measures may be static, automated, some combination of both or simply whether each objective has been met by performing penetration testing.

Modern governments and militaries are increasingly inter-connected through cyberspace, making them both more efficient in operating across distance and time, to organize and manage their operations. Efficient operations and decreasing amounts of time between seeing the results of an action drives military observe, orient, decide and act (OODA) cycles to shorten. That is the information is received by the sensor, delivered to the analyst and sent to the leaders or warfighters to guide their next action. It also makes their infrastructure a target that must be effectively managed and maintained to decrease the threat. In 2012 Maj Gen Suzanne Vautrinot stated "DoD networks are probed millions of times per day…the Air Force blocks roughly two billion threats and denies two million emails each week". Network operations are the day-to-day upkeep and management of the infrastructure and software to maintain their viability.

The US Air Force network contains the 2rd largest Microsoft Active Directory forest in the world behind the country of Germany; its workstations are spread out at over 120 geographically separated locations throughout the world. In order to manage this infrastructure and software to include patching over 1 million computers, automated means are used to push patches, provide reports, detect trouble and maintain anti-virus software. An operator of this designated weapons system relies on automation to make management possible but is it enough to print out a report and call the job complete? Or view a dashboard that shows all green lights instead of red or yellow indicating a problem? It seems that every week there are reports of new attacks on business and government networks that were supposedly secure and up to date. Network security is only as good as your least secure system, network segment, or system user. Balance between security and usability must be maintained depending on the level of security desired.

Measuring effectiveness in network operations then must be a balance between security and the ability of users to perform their mission aided by technology, without being unduly hindered by its operation. This can run the gamut from very loose security practices that allow access to nearly everything, to the opposite end of the spectrum, a stand alone system that is secure but provides little utility if information sharing is needed. Commercial companies reliant on their Internet connectivity often strive toward five nines reliability, meaning an uptime of 99.999%. Breaking that number down into minutes give them a mere 5.26 minutes of downtime per year. Redundant and distributed web servers make it possible to keep websites and services like email and Microsoft SharePoint collaboration sites up while still accomplishing needed updates.

Military organizations often don't have the resources or people to make 5 nines a reality for 100% of their missions or users. It becomes more important for them to identify critical systems or missions that require that level of resourcing. Once identified, the AF uses stoplight charts in measuring whether a system is operationally effective. They are easy to understand for commanders who need quick feedback on the status of a system (US Air Force, 2011, p. 2). At the operator level stoplight charts do not give the level of detail needed and automated dashboards are used. When looking at these measures of effectiveness administrators must also take into account two viewpoints, that of the user and of the operator/maintainer.

System administrators will point to the fact that the server or computer is up and running and say that it is working. While in fact it is on and the operating system is running it may still be unable to perform its mission. Keeping that in mind, we need to look at the system from the view of the user. If the system is on and the user needs a specific program or network connectivity for email in order to perform their job and it is unavailable, then the system is down in their eyes. Using the AF example above, if a user in Colorado Springs is expecting an email from Washington D.C. and they never receive it the system is effectively down. So then, whose point of view do we take in measuring the effectiveness of network operations? The answer is both. The system exists for them to complete their mission and without their need the system would not exist. However, systems must be managed and operated to maintain efficiency to enable the users mission to be completed quickly and effectively.

Defensive Measures of Effectiveness

Measuring the effectiveness of cyber defense may be the easiest of the three to quantify because there are specific devices whose configuration can be displayed and audited or validated against a baseline. They may also be malicious sites or domains that are blocked, an easier

measurement than users feelings about the system or how long a patch took to load and reboot. Effectiveness relies on a solid understanding of the risk being mitigated and the countermeasure used to control that risk (Wright, 2006, p. 1). After all, defense is all about minimizing the risk to the system or operation being performed. Network defense uses passive, semi active, and active defenses to both restrict potential attack vectors and actively search and eliminate malicious programs or users. First, we will explore the types of passive defenses and then how to measure their effectiveness as a system in blocking malicious programs or actors from gaining access to a workstation or network.

Passive defenses such as hardware and software firewalls and gateway router configurations shape the environment being defended. Shaping the environment by limiting the ports and protocols able traverse the first lines of defense limits the capabilities attackers may use against the defended network. These systems are then monitored and audited for data on how effective they are at decreasing the attack surface. Monitoring and auditing systems must be clear in their ability to provide defined, measurable and contrastable data points to the organization (Wright, 2006, p. 6). Further consideration must also be given to where the device is in your infrastructure and whether it is hardware or software when defining measurements. Placement defines just how much or little a given tool will be able to tell the security team.

Custom hardware devices include robust monitoring and configuration tools to enable their use and configuration; they are also able to monitor the physical hardware for potential problems and alert the security team. Software tools have many of the same capabilities but vary in the amount of data they can give based on where they are located in the infrastructure. If installed on end user workstations they will only be able to see the system and not the surrounding environment. Most hardware-based devices are located around critical points in the

network to give them a view into traffic transiting those nodes and the ability to further restrict transit if desired. Managing a large enterprise makes it necessary to use management tools to remotely configure and audit many devices at once from a central console. These systems have their own built in audit tools to verify configurations and ensure functionality. Auditing can be a combination of automatic and manual reviews by the operators to ensure the system has not been modified by unauthorized users and is configured appropriately to be most effective.

The Critical Security Controls advocated by SANS identifies using port scanners as a way to verify the setup has been performed correctly and only needed ports and protocols are allowed through the devices (CyberSecurity, 2014, p. 59). Additionally, effectiveness is measured by the amount of time taken to identify and generate alerts of unauthorized ports being enabled. Using separate programs to audit the configuration is critical in the face of advanced adversaries whose attacks may include altering the output of the management screens. The parties responsible for Stuxnet demonstrated why it is important to use vulnerability scanners separate from the operational system when the management system of Iranian centrifuges showed normal operation. Physical inspection showed what was truly happening and was hidden from view on the control room dashboards (Langner, 2013, p. 12).

Email Gateways scan email traffic for known signatures and domain names in order to block potentially malicious or spam traffic from being received by the end user. Enterprise solutions hold and release messages to users based on proprietary algorithms, which determine if it could possibly be a threat or is just unwanted spam email. Dropping unwanted messages can improve the responsiveness of the internal network by decreasing the bandwidth needed for email traffic. These systems must be tuned to the organizations needs and specific domain names to be effective. Without administrator tuning, valid messages can be delayed or dropped

entirely by the baseline software algorithms. Management software is used to tune all devices in the enterprise to ensure the same settings are used throughout, providing the most reliable experience to the user. The data provided to system administrators is very similar to that of firewall and router management interfaces. Email gateways are more automated than static firewalls and provide a robust security layer when properly configured.

Firewalls and routers are the first line of defense in network infrastructure and provide static defense by allowing only certain ports, protocols and IP addresses to pass traffic in and out of the protected network. They are static devices that must be configured by system administrators to provide some level of protection from threats. As stated above enterprise applications of firewalls and routers have robust management tools, which provide the system administrator data showing their effectiveness. Data such as how much traffic was blocked by what rule, the processor utilization and bandwidth passing through the device are all given. Each of the data points is a measure of how effectively it has been configured and the level of security provided. Traditional static firewalls are cost effective and provide a good measure of security if configured properly to allow only the data needed through the network. Next generation firewalls are able actively process what they are seeing and take defensive measures autonomously if allowed.

Network Intrusion Detection and/or Prevention Systems (NIDS/NIPS) are advanced firewalls capable of taking limited action to block malicious traffic or hacking attempts. Their capabilities vary widely and must be carefully configured in order to not disrupt legitimate traffic. They work by comparing traffic to known attack methods and either taking action to counter the attack or alerting the security team of the event as it is happening. They differ from more traditional firewalls in this capability and are more effective at countering modern threats.

A “learning mode” is used to characterize traffic and provide a baseline to later compare against while in operation. If learning was not used the devices would need to be constantly monitored and manually learn what “normal” traffic looks like in order to block possible malicious traffic. As with the above devices NIDS or NIPS come with their own management software to configure and maximize their effectiveness in protecting the network.

In the overview of specific defensive devices and software above it is clear they all have built in software to manage and tune their capabilities. Cyber defense operators cannot focus on each device in a vacuum in order to measure our overall effectiveness in combatting adversaries; they must take a system wide approach. Ignoring the big picture leaves gaps and seams for adversaries to find and exploit. Operators must be versed in the data displays for each of the security layers and how to fuse what they are seeing in real time to detect and anticipate problem areas or attacks in progress. Advanced systems to aggregate the data known as Security Information Event Management (SIEM) software provide a cohesive picture are available as long as all of the devices support open standards or are supported by the data integration software chosen.

SIEM software is overlaid on top of the previously described management software in order to provide real-time aggregation of security data. Adding the context missing from stove piped systems, provide enhanced abilities for incident investigations and detect advanced or unknown threats that would fail to be detected by single source systems. In order to install such a system the organization needs to perform a detailed analysis of their needs and their current environment. Requirement validation should include what is needed from the system and overall project cost to include additional servers, Storage Area Networks, and training to install and operate the SIEM.

After the defensive layers and systems are in place the only way to truly know if they are working as engineered is to perform penetration testing. Comprehensive pen testing will both measure the effectiveness of the current systems or it will uncover additional holes to be closed within the security architecture. The Penetration Testing Execution Standard (PTES) is a recent effort to better codify how a pen test should be coordinated and executed within industry. It is meant to give a baseline for the activities needed by organizations to gain the full use out of a pen test. The sections include pre-engagement interactions, intel gathering, threat modeling, vulnerability analysis, exploitation, post-exploitation, and reporting results.

Before a pen test is started it is important to have a meeting between the testing team and organization being testing to go over the scope of the test i.e. what is on or off limits, rules of engagement, evidence handling, timeline, is the test being conducted with the security team's knowledge, and locations being tested. The testers also need to know if it being conducted for a specific legal compliance requirement. Some examples are HIPPA, PCI-DSS that require specific tests to be performed and documented for compliance records. Finally, goals should be identified and documented before starting.

Intel gathering and threat modeling can vary based on the scope of the test. Web and network tests are less involved than a full red team test of the entire organizations security posture. The framework identifies three levels of intel gathering ranging from automated tools and open source web lookups to advanced red teaming that is manpower and time intensive. The requirements and scope of the test drive the detail needed from this step. Once information is gathered threat modeling using the criticality of the assets being defended and the attackers capabilities. Business Asset Analysis and Processes are key to threat modeling in order to understand how the business operates and what assets are high priority for potential threats.

Vulnerability Analysis is used to discover flaws or holes in applications and systems being tested. Automated tools like network vulnerability scanners are used to sweep the environment and find any unpatched and misconfigured systems in addition to what ports are open into and out of the target. There are also web and application scanners, which detect flaws in the actual applications or web pages running on the targeted network. Once data has been gathered it is often validated and correlated against vulnerability databases and category of threat or compliance framework i.e. HIPPA. After the vulnerabilities have been categorized testers can move into the exploitation phase and try to "hack" the system by bypassing the security restrictions put in place to access the high priority targets found in intel gathering and threat modeling.

After the targeted system has been compromised by exploitation the goal is to determine what is on the machine and the value of the information or machine to the organization. In this step rules of engagement are very critical to ensure that the host system and any sensitive data are protected. Extensive documentation of what was changed or found should be kept to ensure the configuration can be secured once the test is complete. Ultimately the key to a successful pen test is the report generated to give to the client. It should give an overall effectiveness grade and the level of risk found in the organization with recommendations for improvement. Part of the recommendation section is the roadmap prioritizing items requiring remediation and the level of risk they present to the organization. Pen testing uses many of the same steps as offensive operations and will give the organization an accurate representation of what actual attackers could do.

Offensive Measures of Effectiveness

Offensive cyber operations by non-nation state actors and nation-states offer more traditional military examples of visual queues on their effectiveness. Non-state actors are in news reports nearly every week as the result of active attacks to gain intellectual property, personally identifiable information, and credit cards. Other attacks on organizations include Distributed Denial of Service (DDoS) attacks used to harm the target organization or the customers who do business with them. Nation-state offensive operations are not as prevalent in the news and are by design hard to attribute back to their originators, lest they draw retribution in another domain by the target. This is in stark contrast to non-state actors who wish their exploits to be known and attributed to them in order to draw attention to their cause and skills.

Non-state actors can be a single person or a group of people aligned by common ideals in their interests. An example of a person who showed his skills but did not intentionally cause harm by his actions is Kevin Mitnick, who was pursued for years by the Federal Bureau of Investigation for his exploits. The group Anonymous gained fame for DDoS attacks against banks, Government websites, and e-commerce sites resulting loss of confidence, revenue or ability to convey information by the targets. Cyber crime is a global concern and costs at least $375 billion annually according to a McAfee report (Intel Security, 2014, p. 1). Not all attacks are for monetary gain but organizations affected by them report fiscal losses for their inability to continue business or the loss of intellectual property (IP). In contrast nation-state actors often perform cyber network exploitation (CNE) in order to gain intelligence on adversaries.

Nation state goals in CNE are to obtain the information they seek from the targeted system without being detected. Cyber espionage is a cat and mouse game not unlike that played out on the streets of nations the world over by intelligence services. Through planning of the mission is required in order to gain the required information while remaining undetected or

untraceable to the actual home country. Non-attribution is important because CNE starts with gaining access to the system in order to exploit it for the information it contains. US government policy states that certain cyber attacks may constitute a "use of force" and invoke a nation's right to self-defense through conventional or cyber means (Department of Defense, 2011).

Russia or actors in support of Russian goals used Distributed Denial of Service attacks against both Estonia and Georgia in order to cripple their use of the Internet for command and control as well as distribution of news. Banking and other services were also affected, causing greater affects against the population than a purely military objective. Soon after the attacks in 2007, NATO stood up the Cooperative Cyber Defence Centre of Excellence in Estonia. NATO reaffirmed its commitment to cyber defense through its Enhanced Cyber Defence Policy in the Wales Summit Declaration in Sept 2014 (NATO, 2014). Governments continue to struggle with attribution and policy toward cyber crime and attacks within their borders.

CNE operations are conducted against targets determined to be of use to the nation or entity performing the mission. Before the information can be accessed offensive operations are conducted using exploits of vulnerabilities to a host operating system or device connected to a network needed to gain access to the targeted information. There are publically available tools such as Kali Linux, an operating system specifically designed with penetration testing toolkits as its core functionality. These tools contain thousands of known exploits for Internet connected devices and computers. Advanced hackers and nation-states can use such an operating system as a base and include their own exploits. "Zero-day" (previously unknown) exploits are found and created for use against high priority well-defended targets. Effectiveness of zero-days is extremely high due to system vulnerability being unknown before the attack being launched. Their use must be against the right target operating system and patch level to be successful.

The choice to use previously unknown exploits is not taken lightly as it will most likely end up being patched or mitigated by the vendor once the attack has been discovered. Therefore, attackers must use a gain/loss analysis to determine if the use of the exploit is worth the gain of the information or ability to control the system being targeted. A high level example of complex malware taking advantage of zero-day exploits is Duqu. It is assumed that a nation-state wrote the Duqu malware to gather intelligence on industrial control systems and targeted computers. Using the windows kernel vulnerability zero-day outweighed the loss of disclosing it to the responsible party. Malware designed for espionage is harder to determine its effectiveness; it must be assumed that finding it on a system means that all files have been compromised. Attacks meant to cause harm or change the characteristics of a target are easier to determine success or failure as the initiator.

Stuxnet is widely considered to be the first use of a cyber "weapon" against a physical target. Its creators set out to cause the failure of uranium enrichment centrifuges in Iran by manipulating their industrial control system. Reports after its discovery linked the infection with the Iranians swapping out centrifuges at an increased rate and therefore slowing down their enrichment activities. It also caused the engineers to spend valuable time trying to figure out why their equipment was malfunctioning and second-guessing themselves (Langner, 2013, p. 16). It did not cause widespread destruction but had the capability to if the responsible party had wanted to tip their hand that they were in the control system.

Offensive operations can by easier to assess damage or ability to manipulate the target by having visual queues. It is still difficult to judge how effective they were in accomplishing their intended effects on the target from the outside looking in. The ultimate goals of the attacker may not be revealed or may be a third order effect against another target entirely. CNE effectiveness

often hinges on whether the targeted data or system could be exploited without discovery. In the future cyber weapons like Stuxnet that target specific systems to cause destruction in preparation for more traditional kinetic attacks are likely to be used.

Conclusions

Cyber operations are broken down into network operations, network defense, and network attack/network exploitation. Network operations primary job is to keep the network running and systems operational and patched for users to complete their missions. Effectiveness can be determined by tradeoffs in security and the ability of the end user to perform their mission without technology inhibiting completion.

Dashboards and SIEM tools aggregating data to find and track threats or changes in the network during operations enhance defense effectiveness. Ultimately the best tool to judge whether the network security architecture is effective is the pen test. These tests must be pre-coordinated and ROE's put in place to ensure critical information is secure and both sides know what to do in the case of unintended consequences. Frameworks like PTES are useful in standardizing procedures and the expected documentation received at the conclusion of the test.

Offensive operations measures of effectiveness often depend on the goal of the attack. It may be easy to tell via news media or whether the information sought was gained in the case of CNE. In the future battle damage assessment of cyber weapons is a field that must be more fully developed but could be determined in real time as the attack is taking place depending on the mission.

References

CyberSecurity, C. o. (2014). *Critical Security Controls.* Retrieved September 27, 2014, from SANS: http://www.sans.org/critical-security-controls/

Department of Defense. (2011). *Department of Defense Cyberspace Policy Report.*

Federal Bureau of Investigation. (2013). *2013 Internet Crime Report.*

Intel Security. (2014). *Net Losses: Estimating the Global Cost of Cybercrime* . Center for Strategic and International Studies.

Langner, R. (2013). *To Kill a Centrifuge A Technical Analysis of What Stuxnet's Creators Tried to Achieve.* Arlington: The Langner Group.

NATO. (2014, Sep 05). *Wales Summit Declaration.* Retrieved Oct 18, 2014, from North Atlantic Treaty Organization: http://www.nato.int/cps/en/natohq/official_texts_112964.htm#cyber

Penetration Testing Execution Standard. (2014, August 16). *Penetration Testing Execution Standard*. Retrieved September 27, 2014, from Penetration Testing Execution Standard: http://www.pentest-standard.org/index.php/Main_Page

US Air Force. (2011, November 30). Annex 3-12 Cyberspace Operations.

Williams, B. T. (2014). The Joint Force Commander's Guide to Cyberspace Operations. *Joint Forces Quarterly , 73*, 19.

Wright, S. (2006). Measuring the Effectiveness of Security using ISO 27001.

National Security Archive,

Suite 701, Gelman Library, The George Washington University,

2130 H Street, NW, Washington, D.C., 20037,

Phone: 202/994-7000, Fax: 202/994-7005, nsarchiv@gwu.edu