Open Access License Notice

Enhancing Cybersecurity: Applying Advanced Analytic Techniques

Gregory Laidlaw
laidlags@udmercy.edu

Charles E. Wilson

University of Detroit Mercy
Center for Cyber Security and
Intelligence Studies
Detroit, MI 48221

Abstract - This paper describes several techniques for enhancing cybersecurity and information assurance. Specifically, the paper describes the vulnerability of public and private sector enterprises to advanced and persistent cyber-attacks. To counter these attacks the paper proposes an innovative approach for enhancing cybersecurity by fusing cyber forensics, data mining of big data, and advanced analytic techniques to improve the operational cybersecurity posture of all enterprises operating in the cyber space environment. The intent of the paper is to advance the knowledge in the critical areas of cybersecurity and information assurance by suggesting the creation of an integrated cybersecurity framework to guide analysis of intelligence left behind by attackers; monitoring of networks and systems to identify persistent threats; and eventually using advanced analytics to anticipate and prevent future attacks

Keywords

Advanced Analytics, Big Data, Cyber Forensics, Cybersecurity, and Information Assurance

1. INTRODUCTION

There is a growing body of literature and evidentiary indicators that suggest cybersecurity threats and attacks are an emerging national security issue. The establishment of cybersecurity as a national security threat was formally codified in 2008 by the Comprehensive National Cybersecurity Initiative (CNCI) signed by President George W. Bush in National Security Presidential Directive 54 / Homeland Security Presidential Directive 23 (Department of Homeland Security, 2008). It was also recognized and continued as a national agenda issue in 2015 by President Barack Obama when he signed Executive Order 13694 in April1 2015. Moreover, in 2016, the President reinforced the scope and severity of cybersecurity attacks as a national security threat by declaring a national emergency due to the growing frequency of cyber-attacks against U.S. interests, critical infrastructure, other key assets (Boyd, 2016).

A cadre of key federal leaders have also acknowledged the growing danger and risk associated with cybersecurity threats and attacks. In the 2013 Worldwide Threat Assessment, James Clapper, the director of National Intelligence, for the first time noted that cyber-attacks and cyber espionage have replaced terrorism as the top security threat facing the country (Clapper, 2013). That same year, in his first appearance before a Senate committee, Federal Bureau of Investigation (FBI) Director James Comey Jr. testified that the risk of cyberattacks will exceed terrorist threats as the top national security challenge and will become the prevailing issue for law enforcement and intelligence communities in America (Comey, 2013).

The media has repeatedly published numerous headlines highlighting mega data breaches and other cyber-attacks. The media reports and industry white papers have been validated by the steady onslaught of cyber-attacks, allegedly attributed to countries like China, Russia, North Korea, and the Ukraine, which marks the immediacy of the threat to national security areas influenced by cyber operation. Even more telling is the occurrence of blatant cyber-attacks, such as the 2013 hacks of both the former President George H.W. Bush's and Secretary of State Colin Powell's email accounts allegedly by the same perpetrator.

Attacks from both national and international cyber attackers could cripple critical infrastructure assets, compromise classified information sources, access critical intelligence data, and reveal essential economic or intellectual data. In the contemporary information age, the promotion of U.S. national interests is highly dependent on technological innovation centered in the cyber domain. Cyber-attacks targeting highly sensitive data, key essential information, and highly classified intelligence related to the nation's power projection capabilities could potentially allow our adversaries to effectively counter or undermine our economic, military and technological advantages. Given that cyber threats originate from various and diverse sources, it is difficult to determine whether current cybersecurity prevention actions are effective counter measures within the traditional scope of existing cybersecurity strategies. However, one could deduce from the ever increasing frequency, severity, and scope of successful cyber-attacks that the pervasive trend of cyber-attacks is not being effectively countered by contemporary cybersecurity approaches. Therefore, cybersecurity analysts, network defense personnel, and enterprise leaders must recognize that an innovative approach is needed to effectively address current and emerging cyber threats.

As the global security landscape rapidly changes, it is no longer adequate to lock down our systems and hope for the best. New avenues of commerce and an increasing focus on cloud services open up new avenues of attack that are being exploited by evermore sophisticated attackers. In order to cope with this rapidly changing threat landscape, new approaches must be considered. This paper proposes that intelligence gathering from internal sources provided by logs and events, combined with intelligence from external sources, and framed by models developed and tested in the criminal justice field, can be useful to predict, prevent, and detect new forms of attack.

As we attempt to secure our systems against attack and yet keep them useable we face a dilemma as to where to focus our efforts. As security practitioners, once we have implemented and audited best practices where do we invest scarce security dollars to maximize the return on investment? The field of Predictive Analytics may offer us an avenue to move beyond the reactive nature of best practice based

standards and allow us to focus on emerging and targeted threats to our organizations

Predictive Analytics, using data from multiple sources to inform our actions by predicting the most likely avenues of attack and focusing our efforts to prevent or mitigate them as early as possible in the cyber kill chain As such it can be used as a supplement to traditional risk management and a way to focus our attention on our high value targets and the most likely avenues of attack.

2. DRIVING FORCES IN CYBERSPACE

The world is full of generic security advice, standards, and frameworks, but we are facing increasingly intelligent and targeted attacks in addition to common everyday threats. These frameworks are good in a generic sense, but since we need to assume intelligent and persistent attackers, with targeted attacks, we need to answer the following questions:

1) What information do we have the attackers want, or has value to them?

2) Where is the information located (stored and transmitted)?

3) In what ways can that information be accessed or attacked?

4) What is the most effective way to protect that information?

In an attempt to answer these questions, this paper proposes that by taking a more proactive approach combined with the implementation of a security architecture tailored to our organization's specific risk profile, we will have a much stronger security posture. We specifically intend to combine predictive analytics, with the concept of the cyber kill chain to drive the adoption and verify a risk based security architecture based on the Cyber Security Framework.

3. PREDICTIVE ANALYSIS

Predictive Analytics "analyzes large datasets with technologies that enable rapid and accurate analysis, correlation and reporting to identify events and patterns of interest that may indicate malicious behavior in the environment." (Shackleford,

Using Analytics to Predict Future Attacks and Breaches, 2016). While using data to quickly detect and respond to incidents more rapidly, and decreasing the time that attackers are on our systems is a primary goal, information gleaned from analytic systems can also assist us in assessing the likely goals and targets of our attackers. The source of much of the data required for analysis is already generated by the existing network. "Machine data lives in the IT infrastructure: network logs, event logs, firewall and security system data, web logs, email logs – anything and everything operating in the infrastructure. But machine-generated data can be quite problematic for aggregation, data mining and analytic" (Julie Hunt Consulting, 2015).

Network device events: Events from network devices such as firewalls, routers, and switches can generate event logs that will serve as some of the first indicators of an attack or intrusion.

Server event logs: Application, security, and system logs on a Windows-based server provide thousands of entries about the actions of the services and server as well as log on and log off information for users.

Service based event logs: DNS, DHCP, and other services provide information on the actions of workstations in the networked environment.

Application based event logs: Email, SQL, and other sources and stores of data are likely the intended target of the attacker.

Intrusion detection and prevention systems: staples of network event generation that can often detect well-known signatures of attacks or unusual patterns in network traffic.

Antivirus: Antivirus programs, while less valuable over time for prevention, do provide information on issues and anomalies on workstations and within the network.

The options appear endless, and that is one of the primary issues with building an analytic capability both in the number of sources and the large volumes of data. "The amount of data that can be generated by internal and external systems can

range from tens to hundreds of gigabytes daily for some organizations. It is not unusual for large enterprises to generate hundreds of millions of event logs on a daily basis" (Enterprise CIO Forum, 2015). The second issue with the data is the variable formats in which the data is collected. Machine data sources are quite variable, many of which are in multi-structured formats that further challenge data mining efforts. (Julie Hunt Consulting, 2015). Building an infrastructure and normalizing the data is a daunting task, even seemingly simple tasks such as synchronizing time stamps can be troublesome dealing with multiple time zones and devices that have variances in time.

A third, but not final issue to face is the lack of cybersecurity talent, both the gathering and the analysis of this data requires increasingly scarce human resources. "Most organizations are struggling mightily with finding the right skill sets to properly operate and maintain a security analytics platform for detection and response. In fact, this was overwhelmingly cited as the top impediment to discovering and following up on attacks." (Shackleford, 2015 Analytics and Intelligence Survey, 2015). In spite of the large issues to be overcome, the majority of the organizations surveyed in the 2015 SANS Analytics and Intelligence Survey saw value in the work done thus far and were continuing with their efforts. The most compelling reason to overcome these issues in an effort expand the use of security analytics is the continuing change in the threat landscape.

4. CYBER KILL CHAIN

Developed as incident response / analyst framework by Lockheed Martin, the cyber kill chain breaks down the steps of a targeted attack into its component steps as a means of tailoring preventative actions to early steps in the process. Attacks prevented early are less expensive to recover from; therefore early detection and prevention is key.

The general steps in a targeted attack are:

1) **Reconnaissance**
   Studying public information about the target, the target's environment,

software, and practices. Much of this information can be gathered from the internet and other public forums.

2) **Weaponization**
Preparing a backdoor and a penetration plan intended to deliver a successful attack.

3) **Delivery**
Launching the attack and injecting the backdoor.

4) **Exploitation**
Triggering the backdoor, usually an OS or application vulnerability.

5) **Installation**
Installing the backdoor as a bootstrap and any additional remote access tools to retain a persistent connection to the target.

6) **Command and Control**
Use of the tools to establish remote access, and expand capabilities.

7) **Actions on Objectives**
Take action on the original objective, the collection and exfiltration of information, or additional actions against the target.

Detecting cyber threats is much like software testing, where errors caught early are dramatically less expensive to correct than those discovered later. By focusing on early stages, we discourage the attack or break the chain so that it doesn't proceed further. But also as in software development, while we focus on the early stages, we also have a plan for reasonable coverage through the entire cycle.

Stage	Cost of Failure
Reconnaissance	Zero
Weaponization	Zero
Delivery	Zero

Stage	Cost of Failure
Exploitation	Proactive security
Installation	Cleanup
Command and Control	Forensic response
Actions on Objectives	Business response

Table 1: Significance of Failure (Ranum, 2014)

Stopping the attack earlier in the cycle decreases costs for the defender and simultaneously increases the costs for the attacker. The kill chain describes how to block an attack. However, it is important to realize that, once thwarted an attacker does not need to start over, merely retreat to the prior step and regroup. The farther down the chain the attacker can move, the more reusable intelligence and tools they have access to.

Stage	Preparation
Reconnaissance	None – we remain unaware
Weaponization	None – we remain unaware
Delivery	None – we remain unaware
Exploitation	▪ Firewalls ▪ Anti-virus software ▪ Desktop security
Installation	▪ File tamper monitoring ▪ Configuration management

Stage	Preparation
	▪ System monitoring
Command and Control	▪ Network monitoring ▪ Audit logging ▪ Traffic logging ▪ Network trace logging
Actions on Objectives	▪ Server-level file access monitoring ▪ Network trace analysis ▪ Event analysis

Table 2: Preventative / Defensive Measures (Ranum, 2014)

Predictive Analytics and the cyber kill chain concept provide insight into the specialized threats to our organization and provide understanding about the best places to thwart attacks. Risk management and a risk tailored security framework is how we focus our efforts and energies into areas that will strengthen our security and enhance our current practices.

5. RISK MANAGEMENT

"Intelligence-driven computer network defense is a risk management strategy that addresses the threat component of risk, incorporating analysis of adversaries, their capabilities, objectives, doctrine and limitations" (Hutchins, Cloppert, & Amin, 2015). "Risk management is a systematic and structured approach to managing the potential for loss that is related to a threat. To manage risk properly, organizations should understand the likelihood that an event will occur and the resulting impact. This understanding drives the prioritization of security initiatives

throughout the organization. In information security, a risk is the likelihood that a threat agent will exploit a vulnerability" (Ciampa, 2008).

6. ASSET MANAGEMENT

The first step in risk management is asset identification to determine the assets that need to be protected. Traditionally questions are asked about the organizations computing and data assets regarding the implications and impact of an asset being unavailable. This inward looking reflection ignores the value of assets to outsiders, both customers and attackers. In a majority of the most recent attacks, availability was not even an issue as the assets were exfiltrated and not destroyed. It was the confidentiality that was compromised, not the integrity or the availability. Incorporating the perspective of outsiders to our asset valuation is a must if we are to properly protect our assets. And while we read about insider threats to our systems, these threats when motivated by material gain can be treated as outsiders because though the method of attack is different, the motivation is the same.

7. THREAT IDENTIFICATION

Once the important assets have been inventoried and valued, the next step is to focus on potential threats to the asset. While Predictive Analytics can influence what is considered a valuable asset to the company, it provides real value in its ability to identify potential threats and keep us informed about the ever changing threat landscape. By analyzing both threats from an external perspective and data from attempted attacks, organizations can continually refine their view of organizational threats.

This area is where risk management becomes difficult and subjective, answering the question of what threats are likely and which can be ignored. Not only are there almost unlimited threats per asset, but our perspective is continuously skewed by reports on the latest breaches and the horrible consequences for the victim organization. As humans making decisions without data, it is difficult not to react even when we know that the threat posed by the latest headline attack is not as

relevant to our organization. Data provided by Predictive Analytics will at least add perspective to the list of threats faced by the organization.

8. VULNERABILITY APPRAISAL

After the major assets have been inventoried and prioritized and the threats have been estimated, the next issue to consider is the existing vulnerabilities that could enable the identified threats. By supplementing our traditional risk management processes with Predictive Analytics to assist in the identification of assets that are of high value to attackers and to provide guidance for the threat identification and vulnerability stages, we can continually refine our risk management process to be more adaptive to a continually changing threat landscape. There has never been a more compelling reason to develop, implement, and sustain a more proactive cyber strategy, complemented by an integrated process comprised of cyber threat intelligence, cyber forensics, data mining of big data, and application of advanced analytic techniques than the current cyber threat landscape. Therefore, this paper proposes that all enterprises operating in the cyber space environment adopt and utilize an innovative and integrated strategy based on several existing cyber techniques that will improve their operational and cybersecurity posture. Currently, there are several novel cyber intelligence methods that can be used to identify potential cyber threats and targets. For example, Lee (2014) suggests there are three distinct components that can aid in the identification of a potential cyber threat:

- Intent – a malicious actor's desire to target your organization
- Capability – the means to successfully execute an attack
- Opportunity – the opening or vulnerability the actor needs to attack the target

These three components are consistent elements and as such they can be profiled utilizing techniques, such as: cyber threat intelligence, cyber forensics, and advance analytic techniques to determine patterns of behavior, modus operandi of attacks, and attack surface selection processes.

9. CONCLUSION

The problem of cybersecurity attacks is rapidly growing, requiring increasing expertise in the areas of attack detection, response, and prevention. Most public and private sector enterprises use traditional perimeter defense strategies or ineffective reactionary procedures for protecting their cyber systems and networks, which can limit defensive capabilities. This paper has presented a comprehensive framework for establishing and implementing proactive cybersecurity procedures using the attributes of cyber forensics, data mining of big data, and advanced analytics. This is done with a six stage model that specifies a progression of steps focused on detecting, responding, eliminating or minimizing the effects of cyber-attacks. To demonstrate the use of the model, specific descriptions of each stage are provided. Enterprises using the model and following the six stages will able to apply a more proactive and effective cybersecurity strategy. Predictive Analytics is still a new field of study beset by implementation issues that need to be refined and resolved. The existing issues while daunting, do not appear at this point to be insurmountable. The promise and the payoff are too great to be ignored. Also as our adversaries and attackers gain skill and knowledge, we must also gain skill and knowledge to stay competitive. Standing still is not an option.

REFERENCES

[1] Boyd, A. (2016, April). *Obama: Cyberattacks continue to be national emergency*. Federal Times. Retrieved from. Retrieved from Federal Times: http://www.federaltimes.com/story/government/cybersecurity/2016/03/30/cyber-national-emergency/82423306/.

[2] Ciampa, M. (2008). *CompTIA Security+* 2008 In Depth. Boston, MA: Course Technology / Cengage Learning.

[3] Clapper, J. (2013). United States Intelligence Community worldwide threat assessment for 2013: Hearing Before the Senate Select Committee on Intelligence. (1. Congress, Interviewer)

[4] Comey, J. J. (2013). Statement Before the Senate Homeland Security and Governmental Affairs Committee. (1. C. (2103, Interviewer)

[5] Department of Homeland Security. (2008). *Fact Sheet: DHS 2008 End of Year Accomplishments*. Retrieved from DHS.gov: http://www.dhs.gov/xnews/releases/pr_1229609413187.shtm. 3

[6] Enterprise CIO Forum. (2015, November 19). *Big data analytics and an intelligence-driven security strategy*. Retrieved from Enterprise CIO Forum: http://www.enterprisecioforum.com/article/big-data-analytics-and-an-intelligence-driven-security-strategy/

[7] Hutchins, E. M., Cloppert, M. J., & Amin, R. M. (2015). *Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains*. Foxborogh, MA: Lockheed Martin. Retrieved March 10, 2016, from Lockheed Martin Corporation: http://www.lockheedmartin.com/content/dam/lockheed/data/corporate/documents/LM-White-Paper-Intel-Driven-Defense.pdf

[8] Julie Hunt Consulting. (2015, November 19). *Connecting the Cyber-Threat Dots through Big Data*. Retrieved from Highly Competitive Software Insights: http://jhcblog.juliehuntconsulting.com/2015/07/connecting-the-cyber-threat-dots-through-big-data.html

[9] Lee, R.M. (2014). Cyber threat intelligence. Tripwire online. Retrieved from HYPERLINK "http://www.tripwire.com/state-of-security/security-data-protection/cyber-threat-intelligence/" http://www.tripwire.com/state-of-security/security-data-protection/cyber-threat-intelligence/.

[10] Ranum, M. J. (2014, October 29). *Breaking Cyber Kill Chains*. Retrieved December 5, 2015, from Tenable Security: https://www.tenable.com/blog/breaking-cyber-kill-chains

[11] Shackleford, D. (2015, November). 2015 *Analytics and Intelligence Survey*. Retrieved from SANS Reading Room: https://www.sans.org/reading-room/whitepapers/analyst/2015-analytics-intelligence-survey-36432

[12] Shackleford, D. (2016). *Using Analytics to Predict Future Attacks and Breaches*. Bethesda, MD: SANS.