HiSPO - A Novel Threat Analysis and Risk Mitigation Approach to Prevent Cyber Intrusions

Shuangbao (Paul) Wang, Ph.D.
Professor and Program Director for Cybersecurity
University of Maryland, University College

William Kelly
Metonymy Corporation

Xiaoming Wang
Karlsruhe Institute of Technology

Abstract - In this paper, we study recent data breaches from both technical and business operation perspectives and propose an approach that calculates threat factors of information systems based on various features in hardware, software, policies and business operations. The assessment process takes more than 200 features into account. The data are then imported into an algorithm that calculates the threat factor and normalizes the value to [0-1]. A higher threat factor means the information systems would be hacked at higher risk. Mitigation strategies are provided to reduce risks of information systems from being hacked into and to protect data from being misused, stolen or identifiable. Experiments show that the threat factor reduced from 0.71 to 0.38 in one month for the company we worked with. It was further reduced to 0.18 after finishing a four-month assessment and mitigation period.

This comprehensive approach can reduce data breaches caused by cyber intrusions to corporations such as Anthem, Sony, JP Morgan, Home Depot and Target. It can also deal with privacy concerns in this big data arena. Government agencies and private sectors can reduce risks of cyber intrusions by adopting this innovative threat analysis and risk mitigation strategy.

***Keywords**: threat analysis, risk mitigation, threat modeling, data protection, privacy*

1. INTRODUCTION

Though most government agencies and companies have adopted technologies to protect information systems, incidents of cyber attacks are still on the rise. For example, 4 million government employee's data were stolen from Office of Personnel Management (OPM); Anthem data breach exposes 80 million records; Sony was hit by hackers resulting in a companywide computer shutdown and leak of corporate information, including Social Security Numbers (SSNs) of many employees; 76 million people and 7 million small business data were stolen from Chase; some 250,000 users' personal information was stolen from Twitter; the "backdoor" leak of the world's largest social network site Facebook; the third party vendor breach at Home Depot and Target affected 56 million and 70 million people respectively.

National Institute of Standards and Technologies (NIST) published Guide to Cyber Threat Information Sharing [1]. The aim is to assist organization in establishing incident response capabilities and sharing threat intelligence. Microsoft proposed a STRIDE (Spoofing identity, Tampering with data, Repudiation, Information disclosure, Denial of service, and Elevation of privilege) model for charactering known threats [9]. It also puts forward a risk assessment model - DREAD (Damage, Reproducibility, Exploitability, Affected users and Discoverability) [10]. Trike is a security auditing framework for risk management [11]. NRAT (Network risk Assessment Tool) is a network risk assessment framework implemented at Department of Defense (DoD) [12]. NRAT can be used to guide expert analysis of operational risk from the exploitation of a supporting information system. PASTA (Process for Attack Simulation and Threat Analysis) is a process for attack simulation and threat analysis [13]. It is a seven-step process that is applicable to most application development methodologies. CVSS (Common Vulnerability Scoring System) is a vulnerability assessment system. It provides base score as well as temporal (current) and environmental (asset-based) scores [16].

Undo computing [14] is a project at MIT to help computer users restore system integrity after an intrusion. The approach is to record a system-wide dependency

graph to undo the attack that was made by attackers. When an intrusion is detected, the user uses the graph to track down the attack to its source and recursively re-execute legitimate computations, such as processes or system calls. This innovative approach is able recover from attacks that had been recorded in the dependency graph. The only issue is if an incident happens, attackers may steal data from the compromised computer and recovering from the data loss becomes impossible or meaningless.

Though corporations and government have developed a lot of innovative technologies and approaches, due to the complexity of the problems, threats cannot be determined by technology alone. Good policies and effective business operations have to be put in place to make information systems free from data breaches.

2. DATA BREACHES AND TECHNIQUES USED BY HACKERS

Techniques used by hackers range from low tech ones such as phishing and social engineering to more advanced techniques such as malware, backdoors, third party supply chains attacks, zero-day attacks , etc. From our research, we discovered that many of the known attacks could have been prevented. For easy analysis, we categorize the techniques that hackers use to launch attacks into four areas:

- System architecture, firewalls, software patches
- Malware, security policies and human factors
- Supply chains and insider threat
- Database schemas and encryption technologies

Next we look into these areas and analyze the techniques aiming to find comprehensive solutions to counter those attacks.

2.1 Firewalls, Patches and Architecture

Sony has an external intrusion on PlayStation Network (PSN) in April 2011. An unauthorized person has obtained names, addresses, emails, dates of birth, PSN usernames and passwords, credit card numbers, billing addresses and password

security questions of 101.6 million users. Twelve million credit card numbers were unencrypted and were stolen and could easily be read. In July 2014, Sony paid $15 million settlement to the victims.

Not only did Sony fail to use firewalls to protect its networks, it was using outdated versions of the Apache Web server with no patches applied on the PlayStation Network. These problems were flagged on security forums two or three months prior to the April data breach, which were monitored by Sony employees.

2.2 Malware, Policies and Human Factors

On November 24, 2014, the corporate network of Sony Pictures had been hacked. The attackers took terabytes of private data, deleted the originals from Sony computers, and left messages threatening to release the information if Sony didn't comply with the attackers' demands.

In July 2014, George Mason University (GMU) had security incident involving a malware intrusion into the university's network. This is after an earlier incident in 2005.

Experts at Norse say based on both forensic and other evidence that the attack on Sony may not be orchestrated or initiated from outside the company. They suggest an ex-employee was to blame. If this is true, it raises another question: what policies and procedures were followed after that IT person left the company?

2.3 Supply Chains and Insider Threat

Following Target data breach which exposed 70 million customer data in 2013, The Home Depot appears to be another victim of a data breach of their Point of Sale (POS) systems from supply chains, reportedly by the same Russian hacking group that hit Target, Michaels, Neiman Marcus and P.F. Chang's. As much as 56 million customer data were stolen.

All breaches mentioned above were related to corporations. Banks are more secure, thanks for the independent networking and secure electronic data exchange

service. However this is no longer true. In August 2014, 76 million customer data was stolen from JP Morgan Chase due to data were partially encrypted.

Government and corporations especially third party vendors are vulnerable to cyber intruders. Universities, as an open freedom of information platform also suffer hard from the attacks.

2.4 Partial Data Encryptions and Weak Encryptions

Four million data were stolen from OPM. The compromised data contains government security clearances and federal employee records including SSNs and other Personal Identifiable Information (PII). The story published in June 2015. However the breach was first detected in April but it appears to have begun at least late 2014. The intrusion came before OPM implement new security procedures that restrict remote access.

The University of Maryland, College Park (UMD) had one of their records databases hacked [6]. This particular database holds information dating back to 1998 and includes names, SSNs, dates of birth and university identification numbers for 309,079 people. UMD suffered a second cyberattack on the heels of the first data theft just one month later [7].

As a result, UMD moved a number of its websites offline, asked people to change passwords, and purge sensitive data records that are no longer needed. Those actions are apparently not adequate to address the data breach problems.

In January 2005, GMU was hacked. Names, photos, and SSNs of 32,000 students and staff were compromised. It took a week for GMU IT staff to identify the attack. Sensitive data were stolen and used. Partial data encryption was to blame.

For one incident, there may be many areas to look into. For example, networks and systems at GMU had weak encryption, malware and operations issues. Imagine it took a week to discover the intrusion. A hacker could steal 300MB data in less than a second. OPM had data encryption and policy issues. Sony had malware, weak encryption, firewall, operational and insider threat problems all combined.

The threat factor was so high that systems could easily get hit and once hacked, the consequences would be significant.

So far there have been many ways to harden firewalls, monitor patches installations, detect malwares and apply encryptions. As a matter of fact, government agencies and corporations have implemented a lot of such technologies to protect their networks and systems, such as threat intelligence at FireEye, vulnerability score system at Forum of Incident Response and Security Teams (FIRST), alerts and bulletins at the United States Cyber Emergency and Readiness Team (US-CERT) and National Vulnerability Database (NVD) at NIST. On the other hand, incidents of intrusions are still on the rise. Not only to information systems, threats to critical infrastructure such as power grids become a concern. A new comprehensive approach combing technologies with policies and business operations is needed to analyze threats and mitigate the risks [5].

3. THREAT ANALYSIS AND RISK MITIGATIONS

We have proposed a new approach to analyze threats to information systems by gathering more than 200 features from system architecture, networks, operating systems, database schemas, encryption techniques, security policies, business operations, corporate data, threat intelligence, vulnerability scores and threat bulletin boards. This Hardware, intelligence, Software, Policies and Operation (HiSPO) approach [8] uses an algorithm we developed to calculate threat factors based on those features. The threat factor gives us how robust an information system is facing the cyber threats.

3.1 Threat Intelligence

Threat intelligence is to gather and share global threat information, alerts, actors, malware and provide analysis to the government and industries. More advanced analysis include trends, news and profiles so that trusted partners can detect and defer adversaries more effectively. The key benefits of good threat intelligence include:

- Detect unknown attacks

- Increase security analyst efficiency
- Accelerate incident response
- Reduce risk
- Improve Return On Investment (ROI) and
- Effective countermeasures

Continuous monitoring and intelligence sharing make it very useful in threat analysis.

Identifying threats can be done by classifying threats into several board categories, such as spoofing, tampering, session hijacking, denial of service and elevation of privilege, or putting together a threats list with categories. Here are areas being considered in the HiSPO approach:

- Identify network threats
- Identify host threats
- Identify application threats
- Inspect security policies
- Inspect operational security (including insider threats)
- Analyze attack trees and attack patterns

Identifying network threats is to analyze the network topology and the flow of packets, and inspect routers, firewalls and switch configurations.

Identifying host threats can be done by examining the security setting of system servers, application server, patches, open ports, services, access control, authentication, password cracking, viruses, Trojan horses, worms etc.

Identifying application threats is to check authentication, authorization, code vulnerability, input validation, session hijacking, password policy setting, data encryption, sql injection, exception handling, auditing and logging, etc.

Inspecting security policies includes server, router and switch policy, remote access policy, wireless and Bluetooth policy, database credential policy, technology equipment disposal policy, logging policy, lab security policy, software installation policy, workstation security, privacy protection policy, web application security policy and compliances.

Inspecting operational security is to analyzing systems without updated virus definitions, insider threats, security policy enforcement, account managements, authorized connections on firewall, restricted/banned site access attempts, etc.

In addition, the HiSPO approach also integrates data from public and commercial threat analysis and threat intelligence systems including PASTA, CVSS, NRAT, WASC (Web Application Security consortium), and FireEye as a service to get more up-to-date threat data.

3.2 Threat Modeling

Threat modeling process starts with gathering information in network and system architecture, operating systems and updates, components and configurations of applications, data and data storage, database schemas, services and roles, encryptions and external dependencies. Then an assessment team examines the business objectives, security policies, procedures and compliance with interviews from executives ranging from CISO and IT managers. After this step, the team looks at the business operations of the company and interviews top executives including CIO, COO, and CEO.

Next, the team conducts a series of vulnerability assessments. Based on the data collected, the team starts modeling threats [2] to the organization. It uses a Data Flow Diagram (DFD) to represent the system, networks and processes graphically. Figure 1 shows the initial threat modeling diagram from the case study we conducted.

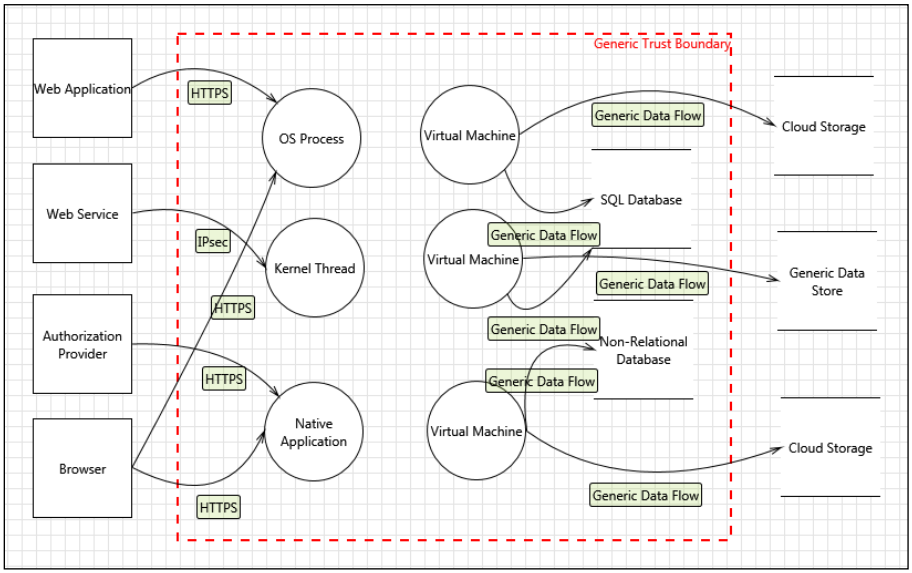

Figure 1: Threat modeling diagram

The diagram here only shows partial network configurations. For the assessment work we conducted, the full diagram contains more than 200 nodes and 800 links.

3.3 Risk Assessment

Based on the threat modeling, the system generates a list of threats and associated risks [3]. Threats are divided into different categories: spoofing, DoS, elevation of privilege, and tempering etc. Human intervention is required at this step to determine whether the threat is at "high risk", "medium risk" or "low risk". Actions can be taken either by marking a threat "need investigation", "mitigated", "not applicable" or an action has "not started". A snapshot of risk analysis view is shown on Figure 2.

The approach also use data from

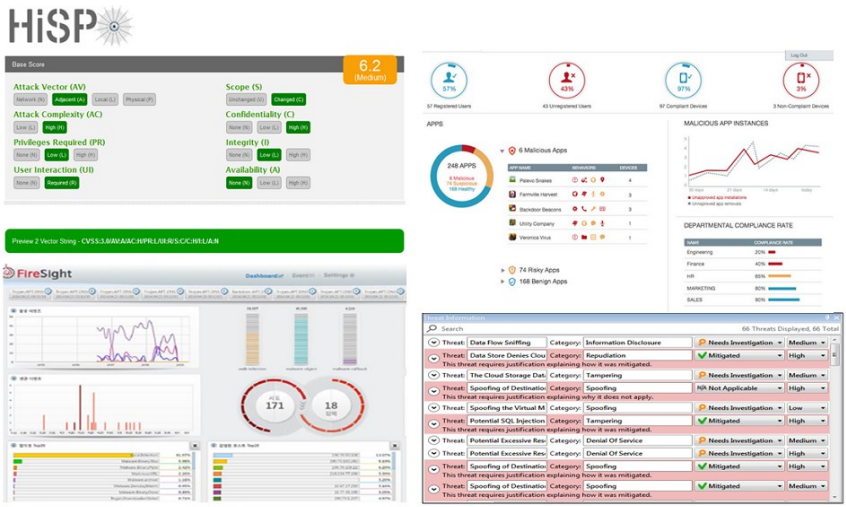

Figure 2: Threat information and risk assessment

This step requires experienced professionals to make judgments. The dynamic threat library that comes with the HiSPO algorithm provides tremendous help.

3.4 Threat Factors and HiSPO Algorithm

To measure threat, we use a threat factor that calculated based on more than 200 features gathered from the previous steps. Threat library contains all threats and updated from time to time. Each threat is assigned with a weight. The value of threat factor is calculated using the formula:

$$t = 0.5 * \sum_{i=1}^{n} w_i * (t_i + \delta) + 0.01 * (c_B + c_T + c_E) + 0.02 * f_{TI}$$

where

t_i - *value of threat i*
w_i - *weight of threat i*
t - *overall threat factor*
δ_i - *weight adjustment for threat i*

C_B, C_T, C_E - base, temporal and environmental scores in CVSS
f_{TI} - threat intelligence value

For example exploitability score can be calculated using this formula [16]:

$$t_{exp} = 8.22 * Attack\ Vector * Attack\ Complexity * Privilege\ Required * User\ Interaction$$

Threat value is to measure the risks associated with the threat. Unlike some other systems such as CVSS that has only three levels (high, low, none). HiSPO assigns threats with continuous value from 0-1. Group threats mean threats measurement imported directly from other threat intelligence systems or vulnerability assessment platforms. The weight for group threats is calculated by the group threat average value multiplies the number of threats included in that approach.

Figure 3 is WASC threat classification in a tag cloud format. HiSPO approach uses the same principle to calculate the weight for each threat [18].

Figure 3. Tag cloud of the web application security consortium threat classification

Table 1 shows the random selection of 10 threat activities of "Insider Threat" along with their threat values and weight factor. The values were assigned based on other score systems and our own experiments.

No	Activities	Threat Value	Weight
1	Attempts to obtain classified information by an individual not authorized to receive such information.	.73	1.0
2	Persons attempting to obtain access to information inconsistent with their duty requirements.	.58	1.0
3	Discovery of suspected listening or surveillance devices in classified or secure areas.	.67	1.0
4	Transmitting or transporting classified information by unsecured or unauthorized means.	.45	1.0
5	Unauthorized copying, printing, faxing, e-mailing, or transmitting classified material.	.71	1.0
6	Adverse changes to financial status.	.56	1.0
7	Any hospitalization for a mental health condition.	.47	1.0
8	Unexplained storage of encrypted data.	.38	1.0
9	Unexplained user accounts.	.66	1.0
10	Hacking or cracking activities.	.82	1.0
	Group Threat Score	0.603	1.0

Table 1. Insider threats measurement

The group average score can be used to calculate the overall threat factor using the formula listed in this section.

The HiSPO algorithm considers threats and risks of most common attack surfaces including hardware, software, policies, business operations and other threat

intelligence data. So the threat factor provides an overview of security of information systems. Reducing the threat factor will in return enhance the security and reduce the risks of data breaches to information systems.

3.5 Threat Report

Threat report contains threat modeling executive summary, model name, owners, reviewers, contributors, description, and a model diagram. It also lists a detailed description about names and nature of the threat, actions that have been taken and a data flow diagram that corresponding to the threat surface. Figure 4 is an example of three threat descriptions.

31. Potential SQL Injection Vulnerability for SQL Database [State: Not Started] [Priority: High]

Category: Tampering is the act of altering the bits. Tampering with a process involves changing bits in the running process. Similarly, Tampering with a data flow involves changing bits on the wire or between two running processes.

Description: SQL injection is an attack in which malicious code is inserted into strings that are later passed to an instance of SQL Server for parsing and execution. Any procedure that constructs SQL statements should be reviewed for injection vulnerabilities because SQL Server will execute all syntactically valid queries that it receives. Even parameterized data can be manipulated by a skilled and determined attacker.

Justification: <no mitigation provided>

59. OS Process May be Subject to Elevation of Privilege Using Remote Code Execution [State: Not Started] [Priority: High]

Category: A user subject gains increased capability or privilege by taking advantage of an implementation bug.

Description: Web Application may be able to remotely execute code for OS Process.

Justification: <no mitigation provided>

64. Elevation Using Impersonation [State: Not Started] [Priority: High]

Category: A user subject gains increased capability or privilege by taking advantage of an implementation bug.

Description: Kernel Thread may be able to impersonate the context of Web Service in order to gain additional privilege.

Justification: <no mitigation provided>

Figure 4: Threat modeling report

The report also contains vulnerability assessment results with data discovered during the process. Figure 5 shows the data stored in the database that were retrieved by our blue-hat team from the website of the company we worked with.

```
[19:46:53] [INFO] analyzing table dump for possible password hashes
Database: db337433205
Table: Hotel_Reservation
[98 entries]
+-----+--------------+---------+------+------+--------------------+-------------------
--+--------+--------+--------+--------------------+---------------+-------------------
--------+-----------------------+
| ID  | Tel          | Code    | Star | Area | Hotel              | Ck_In
  | Room_S | Nights | Room_D | Ck_Out             | Tour_Code     | Hotel Infomation
        | Confirmation Number   |
+-----+--------------+---------+------+------+--------------------+-------------------
--+--------+--------+--------+--------------------+---------------+-------------------
--------+-----------------------+
| 41  | 703-818-1234 | MBRIDGE | 4    | DC   | Hyatt              | Apr  1 2009 12:00A
M | 7      | 1      | 0      | Apr  2 2009 12:00AM | XICHENGQU     | 1 suite + 6 single
        | Confirmed             |
| 42  | 703-273-6161 | HYDE    | 4    | DC   | CY                 | Apr  2 2009 12:00A
M | 6      | 1      | 2      | Apr  3 2009 12:00AM | LL26-MAR      | <blank>
        | Confirmed             |
| 46  | 703-591-5900 | FCC     | 3    | DC   | CF                 | Apr  3 2009 12:00A
```

Figure 5: Database data revealed by blue-hat team

The report also contains threat factors that were calculated before mitigation and after the assessment and mitigation period. For the company we worked with, the first month of assessment and mitigation leads to the threat factor dropping down from originally 0.71 to 0.38.

After the first round, many areas of the information systems are secured. However the blue-hat team was still able to reveal data from the system. The second round of assessment and mitigations took additional three months. When it was done, the threat factor was further reduced to 0.18. At this point, our blue-hat team was no longer able to find any vulnerable data from the system. Spear phishing tests also yielded no negative results. With system admin accounts, our white-hat team extracted data from the databases but was unable to identify any PII content due to the encryption we applied.

3.6 Cybersecurity Curriculum Development

As one of the national Centers for Academic Excellence (CAE) in cybersecurity education, we have integrated this threat analysis and risk assessment research into the curriculum. More than 500 students participated the research ranging from vulnerability testing, threat modeling technique study and risk assessment. Students are very much interested in the real world scenario. Sometime they work independently, other times they work in a team.

On the other hand, one semester seems short for some students getting ready to work. It would be better to have doctoral students and students from other institutions work together. We are looking for programs in the federal government to assist in this effort.

4. CONCLUSION

Threat analysis and risk mitigation are important for corporations and government agencies. In the past, people focus more on installing firewalls and patches, but less on configuring and monitoring firewalls, encryptions, access control and business operations. Even with huge money invested, intrusions still could not be prevented or mitigated.

The new HiSPO approach takes more than 200 features from various areas into consideration. The approach looks at information system architecture, firewalls and malware protection programs. It also looks at database schemas, data encryption technologies, security policies, and corporation operations. Threat intelligence data are also included to keep the system up to date. The vulnerability assessment stage is an iterated process with several threat analysis life cycles. Based on the data collected and imported, the HiSPO algorithm calculates threat factor and normalizes it. The approach also uses defense in depth and threat mitigations strategies, and provides recommendations.

We will further study threat intelligence, threat modeling and risk mitigation technologies, and improve the threat library and value calculation to shorten the threat assessment life cycle. We will further study the relationship between security

and privacy and techniques to protect PII data [19].The adaptation of this innovative approach can minimum data breaches caused by intrusions to government agencies and private sectors and reduce the threats and risks to information system in this cyber in-security space.

5. ACKNOWLEDGEMENT

This research is funded in part by a grant from US National Science Foundation (NSF) [EAGER: 1419055].

REFERENCES

[1] NIST 800-150. (2014). Guide to Cyber Threat Information Sharing. *National Institute of Standards and Technology*.

[2] Rulz, J. et.al. (2012). A Methodology for the Analysis and Modeling of Security Threats and Attacks for Systems of Embedded Components. *20th Euromicro International Conference on Parallel, Distributed & Network-based Processing*. p261-268.

[3] Xu, D. et. al. (2012). Automated Security Test Generation with Formal Threat Models. *IEEE Transactions on Dependable & Secure Computing*. Vol. 9 Issue 4, p526-540.

[4] National Initiative for Cybersecurity Education. (2014). *National Institute of Standards and Technology*.

[5] Mousavian, S., Valenzuela, J. & Wang, J. (2015). A Probabilistic Risk Mitigation Model for Cyber-Attacks to PMU Networks. *IEEE Transactions on Power Systems*. Vol. 30 Issue 1, p156-165.

[6] Privacy Clearinghouse. (2014). *University of Maryland, College Park*. Retrieved from https://www.privacyrights.org/data-breach-asc?title=maryland

[7] UMCP reports another cybersecurity breach. (2014). Retrieved from http://www.baltimoresun.com/news/maryland/education/blog/bs-md-umd-another-cyberattack-20140320,0,798878.story#ixzz3EAtmElmM

[8] Wang, S. P. & Ledley, R.S. (2013). *Computer Architecture and Security*. Wiley, ISBN 978-1-1181-6881-3.

[9] Hernan, S. et. al. (2006). *Uncover Security Design Flaws Using The STRIDE Approach*. Retrieved from https://msdn.microsoft.com/en-us/magazine/cc163519.aspx

[10] Threat Modeling. (2015). *Microsoft Corporation*. Retrieved from https://msdn.microsoft.com/en-us/library/ff648644.aspx

[11] Saitta, P., Larcom, B, & Eddington, M. (2005). *Trike v.1 Methodology Document*. Retrieved from http://octotrike.org/papers/Trike_v1_Methodology_Document-draft.pdf

[12] Whiteman, B. (2008). Network Risk Assessment Tool (NRAT). *IA newsletter, Vol 1*. Retrieved from http://iac.dtic.mil/iatac/download/Vol11_No1.pdf

[13] PASTA. (2015). *Process for Attack Simulation and Threat Analysis Risk-Centric Threat Modeling*. OWASP. Retrieved from https://www.owasp.org/images/a/aa/AppSecEU2012_PASTA.pdf

[14] Hutchins, E. M., Clopperty, M. J. & Amin, R. M. (2010). *Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains*. Retrieved from http://papers.rohanamin.com/?p=15

[15] WASC Threat Classification v2.0. (2015). *Web Application Consortium*. Retrieved from http://projects.webappsec.org/w/page/13246978/Threat%20Classification

[16] FIRST. (2015). *Common Vulnerability Score System v3.0*. Retrieved from https://www.first.org/cvss/cvss-guide

[17] OWASP. (2015). Application Threat Modeling. *The Open Web Application Security Project*. Retrieved from https://www.owasp.org/index.php/Application_Threat_Modeling

[18] Wang. S., Chen. J. & Behrmann. M. (2003). Visualizing Search Engine Result of Data-Driven Web Content. *World Wide Web Consortium*. Retrieved from http://www.w3.org/WAI/RD/2003/12/Visualization/VisSearch/VisualSearchEngine.htm

[19] NIST 800-122 (2010). Guide to Protecting the Confidentiality of Personally Identifiable Information (PII), *National Institute of Standards and Technology*.