Calhoun
Institutional Archive of the Naval Postgraduate School

Calhoun: The NPS Institutional Archive

DSpace Repository

Acquisition Research Program | Acquisition Research Symposium

2017-03

Decision Support for Cybersecurity Risk Assessment

Hibshi, Hanan; Breaux, Travis D.

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/58940

This publication is a work of the U.S. Government as defined in Title 17, United States Code, Section 101. Copyright protection is not available for this work in the United States.

Downloaded from NPS Archive: Calhoun

DUDLEY KNOX LIBRARY

Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

http://www.nps.edu/library

SYM-AM-17-064

Proceedings

of the

Fourteenth Annual Acquisition Research Symposium

Wednesday Sessions Volume I

Acquisition Research: Creating Synergy for Informed Change

April 26–27, 2017

Published March 31, 2017

Approved for public release; distribution is unlimited.

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

Acquisition Research Program
Graduate School of Business & Public Policy
Naval Postgraduate School

Decision Support for Cybersecurity Risk Assessment

Hanan Hibshi—is a PhD candidate and a Research Assistant in the Societal Computing program at Carnegie Mellon University. Hibshi's research area includes usable security, security requirements, and expert's decision-making. Hibshi's research involves using grounded theory and mixed-methods user experiments to extract rules for use in intelligent systems. Hibshi received an MS in Information Security Technology and Management from the Information Networking Institute at Carnegie Mellon University and a BS in Computer Science from King Abdul-Aziz University in Jeddah, Saudi Arabia. [hhibshi@cs.cmu.edu]

Travis D. Breaux—is an Associate Professor of Computer Science, appointed in the Institute for Software Research of the School of Computer Science at Carnegie Mellon University. Dr. Breaux's research program searches for new methods and tools for developing correct software specifications and ensuring that software systems conform to those specifications in a transparent, reliable, and trustworthy manner. This includes demonstrating compliance with U.S. and international accessibility, privacy and security laws, policies, and standards. Dr. Breaux is the Director of the Requirements Engineering Laboratory at Carnegie Mellon University. Dr. Breaux has several publications in ACM and IEEE-sponsored journals and conference proceedings. Dr. Breaux is a member of the ACM SIGSOFT, IEEE Computer Society, and USACM Public Policy Committee. [breaux@cs.cmu.edu]

Abstract

The U.S. DoD transition to a multi-tier, risk management framework aims to streamline information assurance assessments by promoting alignment with NIST information assurance control sets. While these control sets are broadly applicable and comprehensive, those responsible for accreditation will continue to struggle with assessing security risk in dynamically reconfigurable systems. Security analysts rely largely on background knowledge and experience to make security-related decisions. With increasingly dynamic software, analysts need to resolve dependencies among components and understand how those dependencies affect security requirements. Analysts need new decision-support tools based on models that predict how analysts reason about security in distributed systems. We present an approach that formalizes security expert assessments of security requirements nested in scenarios into threat mitigation rules. The assessments are collected empirically using factorial vignettes. The vignette results are statistically analyzed to yield membership functions for a type-2 fuzzy logic system. The corresponding type-2 fuzzy sets encode the interpersonal and intrapersonal uncertainties among security analysts in their decision-making. This work establishes an early foundation for a digital cyber-security decision-support service where an IT professional with any level of security background can benefit from efficiently receiving security assessments and recommendations.

Introduction

The U.S. DoD acquisition process goes through well defined and documented security guidelines. Security guidelines and checklists are widely available and well-documented, but organizations and government agencies like the DoD are still relying on human security analysts to evaluate the security of their systems and reason over these guidelines. The DoD transition to a multi-tier, risk management framework aims to streamline information assurance assessments by promoting alignment with National Institute Standards and Technology's (NIST) information assurance control sets (Marzigliano, 2014; Swenson, 2009). The DoD considers NIST security controls to be the minimum and requires an additional set of controls that vendors need to meet before they can work on classified networks (Swenson, 2009). NIST controls such as the 800-53 ("NIST/ITL Special Publication (800)," 2015) would need to go through a process of implementation to create system design and development requirements. Each control represents a class of technology aimed at mitigating a security threat.

Review of the controls is done by human security analysts who are supposed to have sufficient expertise to reason over potentially millions of scenarios that account for various permutations of controls. Under NIST SP 800-53, the analyst decides if a specific system is high, medium, or low impact and then the analyst satisfies the impact rating by selecting security controls (e.g., audit events, lock sessions, etc.).

Human security assessment can be impacted by context, where security requirements apply; *priorities* that some requirements have over other requirements; *uncertainty* due to human's experts' memory constraints; and the *stove-piped knowledge* among security experts who come from a variety of backgrounds, such as systems, networks, databases, and web applications. We try to use an approach that helps address these challenges. Figure 1 summarizes the steps of our overall approach. In each step, we use methods and techniques that would help understand and address the challenges mentioned above. We have completed two phases of this project, where in each phase we sun a series of studies. In the first phase, we ran all the steps shown in Figure 1. In the second phase, we completed steps 1 through 4 and we are still conducting more studies to refine our results. In the upcoming sections, we will explain the studies conducted, research methodology, and technical challenges of each phase.

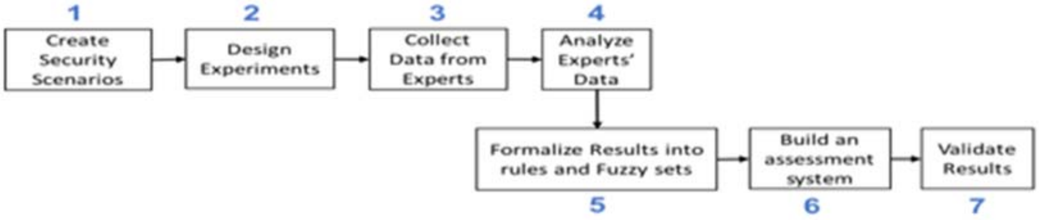
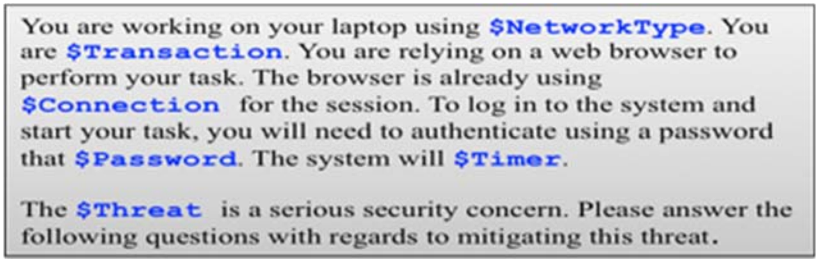

1
Create Security Scenarios
2
Design Experiments
3
Collect Data from Experts
4
Analyze Experts' Data
Formalize Results into rules and Fuzzy sets
5
Build an assessment system
6
Validate Results
7

Figure 1. The Overall Process to Build Security Assessment Digital Solutions

Phase 1: Secure Web Transaction Scenarios

Step 1: Creating Security Scenarios

The main theme in this phase is *secure web browsing*. We conducted a user study using scenarios where participants are asked to rate the security of performing an online transaction (e.g., reading email or credit card purchasing) (Hibshi, Breaux, & Broomell, 2015). Participants were presented with a variety of settings (security requirements) that were manipulated throughout the study to measure which different requirements compositions contribute to the overall security of the scenario and to understand the priorities that exist among requirements. Figure 2 shows the template used to generate security scenarios or vignettes.



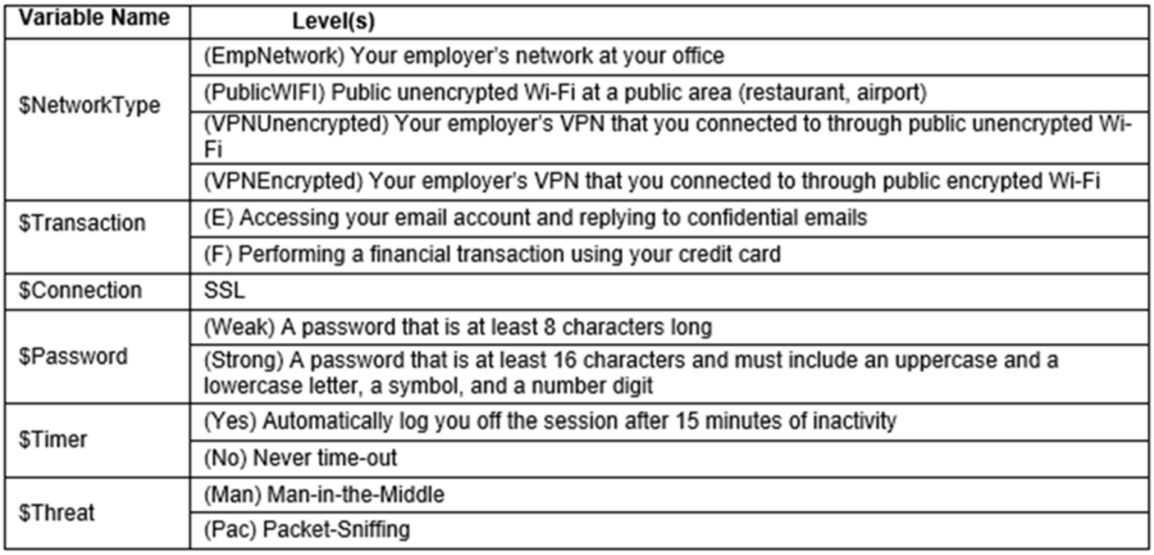
You are working on your laptop using $NetworkType. You are $Transaction. You are relying on a web browser to perform your task. The browser is already using $Connection for the session. To log in to the system and start your task, you will need to authenticate using a password that $Password. The system will $Timer.

The $Threat is a serious security concern. Please answer the following questions with regards to mitigating this threat.

Figure 2. Vignettes Template Used in Phase 1
(Hibshi et al., 2015)

The template shows variables (starting with $) that are each replaced by a security requirement. We call values that replace the variable *levels* of the requirements variable. We generate different vignettes by using different combinations of requirements levels. In Table 1, we show the variables used in phase 1 vignettes and their levels.

Table 1. Variables Used in Phase 1 and Their Levels



Variable Name	Level(s)
$NetworkType	(EmpNetwork) Your employer's network at your office
	(PublicWIFI) Public unencrypted Wi-Fi at a public area (restaurant, airport)
	(VPNUnencrypted) Your employer's VPN that you connected to through public unencrypted Wi-Fi
	(VPNEncrypted) Your employer's VPN that you connected to through public encrypted Wi-Fi
$Transaction	(E) Accessing your email account and replying to confidential emails
	(F) Performing a financial transaction using your credit card
$Connection	SSL
$Password	(Weak) A password that is at least 8 characters long
	(Strong) A password that is at least 16 characters and must include an uppercase and a lowercase letter, a symbol, and a number digit
$Timer	(Yes) Automatically log you off the session after 15 minutes of inactivity
	(No) Never time-out
$Threat	(Man) Man-in-the-Middle
	(Pac) Packet-Sniffing

This technique of using scenarios with discrete factors that get manipulated to study human judgment is called *factorial vignettes* (Rossi & Nock, 1982; Wallander, 2009). We chose this empirical method because it was shown to be more reliable to evaluate and collect human judgment as compared to direct questioning (Rossi & Nock, 1982; Wallander, 2009). Our purpose is to measure the effect of security requirements' composition on the analysts' risk perception and therefore their overall ratings; to identify priorities among requirements, and to understand the effect of ambiguity on analysts' security decisions (Hibshi et al., 2015).

Step 2: Designing the Experiments

We use a mixed-effect design (a combination of within-subject and between-subject factors) for the user study. We ran two separate experiments. First, we invited participants to evaluate scenarios for *Man-in-the-Middle* threat, then we re-invited them after two weeks for the second experiment where the participants evaluate the *packet-sniffing* threat. In each experiment, each participant is assigned a condition where they see four scenarios. They see all the four different values for the `$NetworkType` variable (within-subject effect), but they only see one value for all the other variables (between-subject effect; see Figure 2 and Table 1; Hibshi et al., 2015).

For each scenario, the participant is asked to rate the overall security of the scenario choosing one of the three following ratings:

- **Excessive** security measures that exceed the requirements to mitigate the threat
- **Adequate** security measures that are enough to mitigate the threat
- **Inadequate** security measures that are not enough to mitigate the threat

After rating the overall security, we ask participants to rate each individual requirement shown in the scenario. The ratings 5-point scale, where point 1 is labeled “inadequate mitigation,” point 3 is labeled “adequate mitigation,” and point 5 is labeled “excessive mitigation.” Participants are also given the opportunity to list additional security requirements that they believe contribute to increasing the security level to adequate (Hibshi et al., 2015).

Participants in this study need to have sufficient security expertise. Therefore, we ask participants to answer a list of security knowledge questions that would help assess their level of security understanding, followed by background demographic questions about their years of expertise, number of courses in security, and their job roles. We also collect general demographics such as age, gender, and highest level of education (Hibshi et al., 2015).

Step 3: Collecting Data From Experts

In this study, we sent invitations to security mailing lists at Carnegie Mellon University and North Carolina State University, and we offered participants a $10 Amazon Gift card as a compensation. A total of 174 participants responded to the Man-in-the-Middle threat survey, of which, 116 returned to respond to the Packet-Sniffing survey. The sample has 101 graduate students, 42 undergraduate students, and 2 university professors (Hibshi et al., 2015).

Step 4: Analyzing Experts’ Data

In this step, we use two methods of analysis:

- Multi-level modeling of the user security assessments. This method is suitable to analyze data from our study that instruments a mixed effect design.
- Grounded analysis (Corbin & Strauss, 2007) of additional requirements. We code the statements provided by participants, and then we categorize the codes into one of six categories: server, client, encryption, network, encryption, attack detection/prevention, and integrity and authentication.

Our study results show that security requirements’ composition affect the experts’ risk perception and security assessment. For example, in scenarios where the password level is strong, participants rated the overall security of the scenario to be less than adequate if the `$Networktype` is public Wifi. Participants view the network to be have higher priority than the three requirements: password, timer, and SSL. Once the `$Networktype` is raised to an adequate level, then other requirements will start impacting the risk assessments (Hibshi et al., 2015).

Step 5: Formalizing Results Into Rules and Fuzzy Sets

We formalize the results of the empirical study to derive if-then rules that we use in a security assessment system based on rule-based interval type-2 fuzzy logic (Hibshi, Breaux, & Wagner, 2016). The following is an example rule that we derived from the results:

$$R^1\text{: } IF\ \boldsymbol{NetworkType}\ is\ \boldsymbol{Inadequate}\ THEN\ \boldsymbol{OverallRating}\ is\ \boldsymbol{Inadequate}$$

Any fuzzy logic system needs fuzzy sets that can be constructed using experts’ input.Type-2 fuzzy sets allow us to model the uncertainty in the data by providing *a footprint of uncertainty* (FOU; Mendel, 2001). It is important to point out that uncertainty in our data is always present because it relies on experts’ input. Experts’ data include interpersonal uncertainty, which is the uncertainty between different experts, and intrapersonal uncertainty, which is the uncertainty that the same expert may experience on two different occasions due to the nature of humans’ memory (Hibshi et al., 2016).

To build fuzzy sets in our security assessment system, we conducted another empirical study on security experts where we asked participants to provide an interval on a range from 1 to 10 to represent linguistic labels for adequacy (Hibshi & Breaux, 2016; Hibshi et al., 2016). Figure 3 shows the results of the collected intervals from 38 security experts. Then, we use the data collected to construct type-2 fuzzy sets and their membership functions. Figure 4 shows the membership functions for the three fuzzy sets: inadequate, adequate, and excessive.

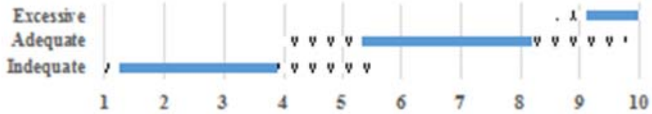


Figure 3. The Fuzzy Sets With the Start and End Means and Standard Deviation (Hibshi et al., 2016)

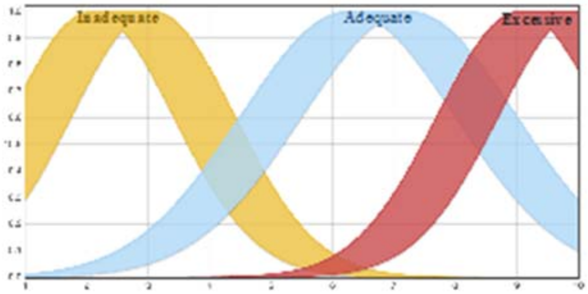


Figure 4. . Membership Function for Inadequate, Adequate, and Excessive (Hibshi et al., 2016)

Step6: Building a Security Assessment System

We will explain how we build and evaluate a security assessment system that would help security analysts evaluate their security decisions. Figure 5 shows the overall architecture of our assessment system.

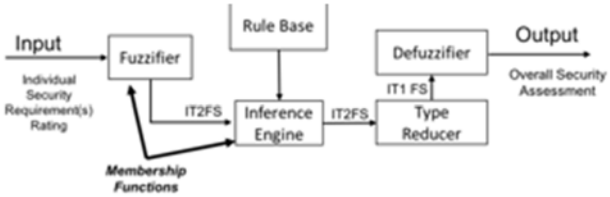


Figure 5. Mamdani IT2FLS for Security Assessment (Hibshi et al., 2016)

We use the type-2 fuzzy sets and rules formalized in step 5 of the process to build our assessment system.

Step 7: Validate Results of the Assessment System

Our goal here is to evaluate the system and to measure how well it mimics human experts' reasoning. Hence, we designed a survey similar to the survey used in the data collection step (see steps 1 and above), and we sent the survey to 13 security experts who rated 52 scenarios (four scenarios per expert). Then, we used the individual security requirements ratings as inputs to our assessment system, which will produce an overall security assessment output. Later, we interviewed the experts and asked them to provide reasoning behind their ratings. Finally, we showed them the system's rating and asked them to compare it to their overall ratings. The results show that the security analysts found the assessment system to provide reliable security ratings, generating more conservative assessments in 19% of the test scenarios compared to the experts' ratings (Hibshi et al., 2016).

Phase 2: Security Administration Scenarios

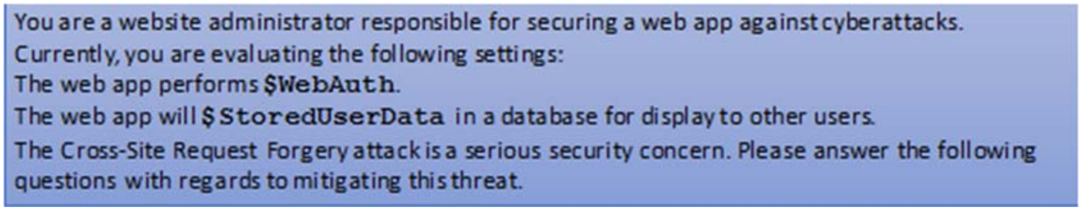
The studies conducted in phase 1 have the following limitations:

- We used a single security scenario that puts the participant in the role of a user.
- We recruited mostly graduate security and privacy students for the user study.

To address these limitations:

- We increased scenario coverage by selecting scenarios from four security domains: networking, operating systems, databases security, and web applications security.
- We used a language in the scenarios that puts the study participant in an expert role analyzing the requirements shown in the scenarios.
- We recruited security professionals from industry and government.

Figure 6 shows the template used to generate vignettes for the web applications security study.



You are a website administrator responsible for securing a web app against cyberattacks.
Currently, you are evaluating the following settings:
The web app performs **$WebAuth**.
The web app will **$StoredUserData** in a database for display to other users.
The Cross-Site Request Forgery attack is a serious security concern. Please answer the following questions with regards to mitigating this threat.

Figure 6. Scenario Template Used in the Web Applications Security Study

The `$WebAuth` variable represents the type of authentication used in the web application and it can take on one of many values. To illustrate, we consider two extremely different values: "basic authentication," which is a weak form of web-based authentication, or "form-based authentication using encrypted credentials stored in a database," which is stronger. Similarly, the `$StoredUserData` variable represents how the user input is being collected and could take the values: "collect user-supplied content from GET request" or "require CSRF tokens and escape and validate user supplied content from POST requests before storing." Again, the latter value is stronger than the former.

In a similar fashion, we constructed scenario templates to generate vignettes for the remaining security domains: networking, operating systems, and database security. Currently, we are still collecting and analyzing data for this phase to help design and build our next security decision-support system.

Summary and Future Work

In this paper, we summarized our research that consists of a series of empirical studies where we study how security experts make their decisions. We used the data collected from experts to formalize and model the human reasoning to build decision-support tools. One of the major challenges in security decision-making is the amount of interpersonal and intrapersonal uncertainties present in the data. Hence, we choose to model experts' data using interval type-2 fuzzy logic, which can handle these uncertainties. We continue to create scenarios, design experiments, and collect data from experts so we can build decision support tools that would better assist the security experts as they make their decisions and evaluate requirements. These smart tools would help security analysts with their acquisition process, as it is a building step towards semi-automating the currently manual process of reviewing systems and evaluating them against security requirements.

References

Corbin, J., & Strauss, A. (2007). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage.

Hibshi, H., & Breaux, T. D. (2016). *Evaluation of linguistic labels used in applications* [Technical report]. Carnegie Mellon University.

Hibshi, H., Breaux, T., & Broomell, S. B. (2015). Assessment of risk perception in security requirements composition. *2015 IEEE 23rd International Requirements Engineering Conference (RE)* (pp. 146–155).

Hibshi, H., Breaux, T. D., & Wagner, C. (2016). Improving security requirements adequacy: An interval type 2 fuzzy logic security assessment system. In *2016 IEEE Symposium Series on Computational Intelligence* (pp. 1–8). IEEE. Retrieved from http://ieeexplore.ieee.org/abstract/document/7849906/

Marzigliano, L. T. (2014, March 14). Defense Department adopts NIST security standards. Retrieved from http://www.informationweek.com/government/cybersecurity/defense-department-adopts-nist-security-standards/d/d-id/1127706

Mendel, J. M. (2001). *Uncertain rule-based fuzzy logic systems: Introduction and new directions*. Prentice Hall PTR.

NIST/ITL Special Publication (800). (2015, January 2). Retrieved from http://www.itl.nist.gov/lab/specpubs/sp800.htm

Rossi, P. H., & Nock, S. L. (1982). *Measuring social judgments: The factorial survey approach*. SAGE.

Swenson, G. (2009, June 11). NIST, DOD, intelligence agencies join forces to secure U.S. cyber infrastructure. Retrieved from https://www.nist.gov/news-events/news/2009/06/nist-dod-intelligence-agencies-join-forces-secure-us-cyber-infrastructure

Wallander, L. (2009). 25 years of factorial surveys in sociology: A review. *Social Science Research, 38*(3), 505–520. https://doi.org/10.1016/j.ssresearch.2009.03.004

Acquisition Research Program
Graduate School of Business & Public Policy
Naval Postgraduate School
555 Dyer Road, Ingersoll Hall
Monterey, CA 93943

www.acquisitionresearch.net