# A QUANTITATIVE MAN-MACHINE MODEL FOR CYBER SECURITY EFFICIENCY ANALYSIS

A Dissertation

by

SUNG-OH JUNG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2005

Major Subject: Computer Science

# A QUANTITATIVE MAN-MACHINE MODEL FOR CYBER SECURITY EFFICIENCY ANALYSIS

A Dissertation

by

SUNG-OH JUNG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

| | |
|---|---|
| Co-Chairs of Committee, | Jyh-Charn Liu |
| | Hoh Peter In |
| Committee Members, | Dick Simmons |
| | William Lively |
| | Marshall Scott Poole |
| Head of Department, | Valerie Taylor |

December 2005

Major Subject: Computer Science

# ABSTRACT

A Quantitative Man-Machine Model for Cyber Security Efficiency Analysis. (December 2005)

Sung-Oh Jung, B.S., Dankook University;

M.S., University of Southern California

Co-Chairs of Advisory Committee: Dr. Jyh-Charn Liu
Dr. Hoh Peter In

The analysis of security defense processes is of utmost importance in the management of various cyber-security attacks, which are increasing in scope and rapidity. Organizations need to optimize their resources based on a sound understanding of the level of their security defense processes' efficiency and the impact of their investment.

Modeling and characterization of the dynamics of cyber security management are essential to risk prediction, damage assessment, and resource allocations. This dissertation addresses the interactions between human factors and information systems. On the basis of the spiral life cycle model of software development processes, we develop a realistic, holistic security attack-defense model – Man-Machine Model ($M^3$), which combines human factors and information systems' (i.e., machine) states under an integrated analytical framework. $M^3$ incorporates man and machine components. The man component is comprised of several variables such as Skill & Knowledge (SKKN) and Teamwork Quality (TWQ). The machine component is composed of variables such as traffic volume and the amount of downtime. $M^3$ enables the analysis of intrusion detection and incident response process efficiency, i.e., security defense team performance.

With data analysis, we formulate and test four major research hypotheses based on the data collected during security experiments. Through hypothesis testing, we evaluate regression models to estimate the security defense team performance (i.e. efficiency) at different levels of human intelligence (e.g., skill and knowledge) and

teamwork (e.g., teamwork quality). We assess the fitness and significance of the regression models, and verify their assumptions. Based on these results, organizations can hire those who have an appropriate level of skill and knowledge when it concerns investments to increase the level of skill and knowledge of security personnel. They also can attempt to increase the level of skill and knowledge of security personnel.

# DEDICATION

To my Lord, Jesus Christ

To my wife, Dr. Hye-Kyung Cho

To my parents, Tae-Je Jung & Yong-Ok Hwang

To my father-in-law, Keun-Haeng Cho, and mother-in-law, Hwa-Ja Lee

## ACKNOWLEDGEMENTS

I would like to thank my committee co-chairs, Dr. Hoh Peter In and Dr. Jyh-Charn Liu, for their scholarly guidance and continuous support throughout the course of this dissertation. I appreciate them for their reasoned and intelligent guidance on my research and for their motivation to keep me thinking creatively – both at the academic and entrepreneurial levels.

I also would like to thank my committee members, Dr. Dick Simmons and Dr. William Lively, for their assistance and help through my preliminary examinations. They also gave me valuable feedback regarding my proposal draft.

I especially want to thank Dr. Scott Poole for his guidance and support throughout the course of this dissertation. He directed me to perform the information security experiments in class, and gave me a deep understanding of how the security and group theory fields can merge together for a new vision.

I also give my gratitude to Dr. Michael Grimaila, who made the security experiments possible by allowing me to work with students in his class throughout two semesters, and to Dr. Thomas Rodgers, who helped me think about the experiments and their connection with this dissertation.

Thanks also to my friends, colleagues, and staffs for helping me have an excellent experience at Texas A&M University through sharing good times and bad times as well as research ideas.

I would like to give my tremendous gratitude to my wonderful family – father, mother, sister, and brother – for their support and encouragement and to my beautiful wife for her patience and love.

Finally, I would like to give my best gratitude to my Lord, Jesus Christ, for His unutterable grace and wisdom, which He provided moment by moment throughout the course of this dissertation.

# TABLE OF CONTENTS

Page

## LIST OF FIGURES

Page

Page

# LIST OF TABLES

Page

Page

# 1. INTRODUCTION

## 1.1. Motivation

Protection of critical information systems and global network infrastructure is an on-going process that evolves around technologies, policies and procedures, and organizations in the view of business continuity and reputations. The Internet and network technologies allow users to enjoy a higher quality level of electronic communication and, thus, more efficient and convenient lives. However, due to the characteristics (such as openness) of the technologies enjoyed, security threats or vulnerabilities become commonplace for those who use them (especially in the case of the Internet). Any system without a connection to the Internet or any type of network is strong in terms of the protection of information; however, as soon as the system is hooked up to the Internet or other networks, it becomes vulnerable since it opens an enormous opportunity to those who want to break in the system; that is where the information assurance issues are born from.

To protect valuable information from attackers, a Security Incident Response Team (SIRT) or at minimum a group of security personnel needs to be formed; however, forming a SIRT or security group is a difficult task due to constraints on budgets and resources; however, security breaching is expected to reduce when more resources are invested. In order to decide on the allocation of resources, such as human resources for information assurance, an organization must have knowledge of how the system and human groups (i.e. SIRT) interact.

Obviously, technology is an important factor for security assurance; however, the installation of cutting-edge technology does not mean that systems are free from potential security threats or breaches. Grant Gross [28], a reporter for the IDG News Service, explains that, “Security is being more ingrained within the business and within daily operations … It's not just a technology solution any more -- you can't just throw a

---

This dissertation follows the style of *IEEE Transactions on Software Engineering.*

firewall in ... and the problem is solved. You have to address security from a people, processes, and technology standpoint in order to really have a successful security strategy in place." In essence, the technology must accompany human groups to strengthen information assurance since it is human groups (i.e. SIRT) that implement the technology; people make the technology take effect and fulfill the goals of information security defense because they are the critical decision makers in the process. The basic role of a team is to defend against any type of security attack with their combined knowledge, experience, and interactions with the other team members. Thus, the effects of the security defense team's defense efficiency (e.g. detection rates) and quality (i.e. teamwork quality or TWQ [33]) on the overall system status in emergency situations (e.g. security breaches) need to be analyzed.

IT security guidelines [30] state potential problems that impede the ability to defend against security attacks include poor maintenance of IT systems and failure to install the available security updates. System administrators often do not install security updates promptly, and much of the damage caused by viruses or worms only becomes apparent some time after the existence of the pest (e.g. a security breach) has become known.

A disastrous result of the potential *people* problem would be the Slammer worm [14], [71], which spread over the Internet (especially in many Asian countries such as Korea) on January 25, 2003. At unprecedented speed and scale, the worm paralyzed airline/train reservation systems, Internet shopping and banking, and information sharing and searching. The primary factor in this disaster is obviously the worm itself; however, human errors were evident. For instance, security patches were not installed, servers were not properly configured, and administrators could not coordinate their actions in an orderly fashion.

J. E. Canavan [8] defines this kind of people problem as *human vulnerabilities*, and characterizes them as follows:

> Human stupidity, carelessness, laziness, greed, and anger represent the greatest threats to networks and systems and will do more damage than the rest (system, physical, media, etc.) of the others combined. Moreover, human vulnerabilities and the risks associated with them are the most difficult to defend against.

Recognizing the importance of people problem, i.e. human vulnerabilities, this dissertation proposes a quantitative Man-Machine Model ($M^3$) to analyze the effects of human performance into security defense processes – intrusion detection and incident response. The proposed hybrid modeling approach can model systems or processes with hybrid natures that have both discrete and continuous characteristics. But, to the best of our knowledge, there has never been a study done on the hybrid modeling of man-machine interactions for information assurance.

## 1.2. Problem Statement

This dissertation will address the problem of measuring the efficiency of security defense processes for the protection of Information Systems (IS). It is acknowledgeable that any system contains vulnerabilities which provide good starting points for various types of security attacks performed by malicious remote users (e.g. crackers) or insiders. Humans are another major variable in disastrous situations. Without understanding the interactions between systems and humans, it is not easy to find or eliminate the potential bottlenecks which reside in information security defense processes and reduce protection levels. The rationale of this study is to help create efficient resource allocations (e.g. systems and humans) while providing a better degree of IS protection. The following is the problem statement we pose to tackle throughout this study:

> *How can we quantitatively measure the effects of key human factors on the intrusion detection and incident response process in terms of efficiency?*

## 1.3. Research Questions

Since we approach this study using a human-centric viewpoint, our primary focus is on how human performance can enhance the efficiency of the process of intrusion detection and incident response. Through our modeling and analysis work toward the measurement of human performance, we hope to present a guideline for human allocation problems in the management of security defense processes with a software engineering viewpoint. To draw this guideline, we begin with our primary research questions, which can be posed as: (1) Where can human vulnerabilities occur? (2) What approach can be effective in handling them? Why? (3) What are the goals using such an approach?

## 1.4. Dissertation Goal

The goal of this dissertation is to derive quantitative models of security defense (intrusion detection and incident response) team performance. The models will be used to analyze the efficiency of the security defense team. The goal is achieved by developing the Man-Machine Model ($M^3$) using a simulated attack (i.e., TCP SYN flooding Denial of Service (DoS) attack) and performing data analysis on the team performance using regression models.

The benefits we expect to gain through this dissertation are as follows:

- To develop a more realistic, holistic-view model of the security attack and defense process by incorporating group dynamics into the system.
- To explore not only the vulnerabilities of system and technology, but the vulnerabilities of human groups (i.e. SIRT) as well.
- To realize the relationship among key candidate variables and the conjectured security defense team performance variables.

## 1.5. Overview of the Dissertation

This dissertation presents a quantitative model to guide security managers' decision making for better intrusion detection and incident response processes. Section 2 reviews

the background of information assurance and group behavior; that section also examines the efforts exhibited by industry or academic security people toward better intrusion detection and incident response. Section 3 presents the model components and a model example. Integrating these components, we create and examine a TCP SYN flooding Denial of Service (DoS) attack to show that the model can be used to represent the process of intrusion detection and incident response. Section 4 presents research hypotheses that we will test throughout this dissertation, and describes the security experiments including data collection, experimental setting, and process. Section 5 presents the data analysis used to test the research hypotheses and analyze the efficiency of the security defense process; in that section we also present and verify necessary hypotheses assumptions. Section 6 summarizes this dissertation, and covers key contributions, discussions, and future work.

# 2. RELATED WORK

Security researchers and experts have pointed out the importance of intrusion detection and incident response for information assurance and risk management for more than 20 years. Thus, many organizations developed thousands of security tools, algorithms, policies, etc., to protect themselves from malicious software and hackers. However, their main focus and efforts were not on the measurement of human performance to improve the efficiency of the security defense process; rather, they were focused on the systems or technology only. Without the measurement of human performance, it will be extremely hard to appropriately allocate human resources for the construction of good, solid information assurance programs.

## 2.1. Information Assurance

Assuring information systems in IT organization is crucial since information assets should not be leaked out to *bad* guys (i.e. security attackers) who can possibly use the assets for their own purposes, usually in an illegal way. To secure information systems, much research has been conducted including vulnerability assessment, which is one of the most important issues that one must address helping users determine the best approaches for preventing attacks [4]. In view of human-behavior modeling complexity, we must start with certain hypotheses about the different motivations of attackers, i.e. fame, money, and privacy [25], [26]. The approach proposed in [25], [26] is to reduce risk in software life cycles by using a software security assessment instrument.

To protect the valuable information and monetary assets from attackers, a Security Incident Response Team (SIRT) or more generally an Incidence Response Team (IRT) would inevitably need to be formed; however, forming the SIRT or security group is a difficult task due to budget and resource constraints (e.g. human resources). Danny [78] addresses the issue by examining the roles of an IRT while stating that forming an IRT is a daunting task. Even if an IRT is commonly used to respond to

attacks in large software development or user organizations [90], [93], there are few descriptions of the behavioral dynamics that shape their actual behavior.

## 2.2. Group Behavior

Besides information assurance, group behavior is another field we study and apply in this dissertation research. Since good information assurance programs cannot be maintained by single-individual efforts, group behavior should be investigated. We hope that the application of group theory into this study enables us to approach a more realistic and effective program. We address the background - including importance - of these areas in this section.

The importance of group (i.e. team) behavior-related issues has been long recognized in areas of behavioral, human, social, and organization sciences, etc. For instance, the importance of group decision-making brings the need for computerized decision support systems [36]. Efforts that measure important factors and rating scales in modeling group behavior and decision making have also been investigated by many researchers. Teamwork quality and team performance variables have been constructed by Hoegl and Gemuenden [33]; a rating scale for subjective workload assessment was defined by Hart and Staveland [29]; stress measurements were described and compared in [54], [42], [89], [9], [87], [47] and a social readjustment rating scale to measure stress was described by Holmes and Rahe [34]; a confidence scoring index for speech understanding systems is presented in [63] and various confidence measurements are introduced in [27], [64], [72]; a way to measure surprise is introduced in communication theory by Shannon [74] who claims that the amount of information is a measure of surprise and is closely related to the chance of one of several messages being transmitted.

## 2.3. Security Modeling and Analysis

Based on the knowledge of two major fields, this dissertation focuses on the measurement of human performance while considering system or technology issues to be important as well. To the best of our knowledge, this study is an innovative effort

[38], the first of its kind to integrate a machine-model with a man-model for improving the efficiency of the security defense process. A survey of the literature on the measurement of human performance shows that very little is done in the measurement of human performance for better security defense process. Even so, we present other's work on security defense – intrusion detection and security defense – because their work also aims for better intrusion detection and incident responses, either partially or exclusively. We draw potential key contributions of this study from the observations of their work.

**Preventive security:** Jonsson and Olovsson [44], [45] show an effort to model preventive security based on empirical data collected in the experiment. Based on the data, they formulate and test a statistical hypothesis that the times to breach are exponentially distributed. They come up with "a typical attacking process" (learning phase – standard attack phase – innovative attack phase) in a graph of 'number of breaches' vs. 'time' (both initially inexperienced attackers and initially experienced attackers).

**Operational security:** Littlewood and Brocklehurst [51] address some quantitative aspects of operational security in an analogous manner to operational reliability. They raise several questions concerning a probability-based framework for operational security measurements in discussion.

Another similar approach towards the measurement of operational security is presented in [7]. The approach is based on the analogy between system failure and security breach. To examine the raised issues, they conduct a pilot experiment to assess the feasibility of collecting data. Attackers, not defenders, fill out the reports during the experiment related to their attacking activity and breaches; the data collected is a mixture of quantitative (efforts) and qualitative (rewards).

Ortalo and Deswarte [62] propose a theoretical model (privilege graph) that describes the system vulnerabilities in order to evaluate operational systems security. They also present a mathematical model to evaluate the mean effort for an attacker to reach the specified target, denoted as 'mean effort to security failure' (METF).

**Intrusion scenario detection:** A modeling work is presented in [18] with the objective of an intrusion scenario – an organized set of actions, which have to be executed by intruders following a certain order. Their model represents an attack, intrusion objective (final purpose of an intruder, which justifies all its actions) corresponding to security policy violation, domain rules, and intrusion scenarios. They introduce a notion of anti correlation, and claim that it is useful to recognize a sequence of correlated attacks that no longer enables intruders to achieve an intrusion objective.

**Probabilistic properties of computer audit data:** Comparative studies on probabilistic properties of computer audit data that are important to intrusion detection are performed by Ye and Li [94]. The data of intrusive activities in an information system (consisting of host machines and communication links according to their definition) are generated with the simulation of 15 different intrusion scenarios. The authors provide answers to which properties of the data are necessary for intrusion detection by applying various techniques of detection such as Hotelling's $T^2$ test.

**Cost analysis and modeling:** Wei and Frinke [85] perform a cost-benefit analysis for network intrusion detection systems with the objectives of (1) quantitative and qualitative analysis of security risks, and (2) construction of a cost model, which can be developed into an on-line system for real-time use. The cost model calculates the total costs of specific attacks. To test the model, attack data from network intrusion detection systems are gathered.

Developing a data mining framework, Lee and Fan [49] build a cost-sensitive intrusion detection model and examine cost factors associated with IDS (Intrusion Detection System). One of the main objectives of their modeling work is to reduce the total expected cost, which is summed from operational and consequential costs.

# 3. PROPOSED WORK: MAN-MACHINE MODEL ($M^3$)

## 3.1. Petri Nets (PNs): A Conceptual Modeling

To achieving the described objective, the main focus of our research employs the analysis of relationships between security defense team performance (e.g. detection rates) and human factors (e.g. TWQ), and the analysis of security defense team performance effects on the system resources. Before the analyses, we need to investigate the relationship between users (e.g. defenders), information systems, and attackers. For the purpose, we construct the Man-Machine Model ($M^3$), which entails both a discrete event part and continuous part using Petri Nets (PNs) [55].

Petri Nets (PNs) such as stochastic hybrid Petri-net models [65], [12], [20], [11] and simulation tools [16], [70], [84] have been developed in the control system area to represent, understand, and manipulate complicated states of system components. In this dissertation, our model approximates the *positive* and *negative* relationships commonly used in the man model. Our modeling approach is to develop basic components (e.g. continuous and discrete places/transitions) and their interconnections, i.e. firing rules for continuous-continuous, discrete-discrete, and continuous-discrete state transitions, so that he/she can freely develop his/her own models in his/her domains.

In a continuous model, marks are considered real quantities by subdividing whole marks into infinitesimally smaller parts (called "tokens"), whereas marks are treated as integers in a discrete model [12]. Even if the mere passing of time does not have a direct effect on the state of a discrete event model, general Petri-net models are extended to Discrete Petri Nets (DPNs) by introducing time variables in the firing vector, V, similar to the one proposed in [12]. In DPNs, state changes when a transition is fired are represented as:

$$M(t^+) = M(t) + C^d \bullet V(t) \quad (1)$$

where $M(t^+)$, the next following marking function, is driven by the current marking function, $M(t)$, according to the firing vector, $V(t)$. The firing speed is represented by $V(t)$. $M(t^+)$ and $M(t)$ are elements of a set of integer-number marking vectors **M**. The initial marking, $M_0$, is defined by $M(t)$ at time $t = 0$. Driven by an event, $V(t)$ determines an instant transition with zero duration. $C^d$ is an incident matrix of DPNs corresponding to the weights of the links (or arcs). Thus, in DPNs, the amount of marking change caused by a state change, $M(t^+)$ *minus* $M(t)$, is $C^d \bullet V(t)$.

Several approaches to defining Continuous Petri-Nets (CPNs) are presented in [65], [12], [20], [11], [13], [67], depending on their compatibility with DPNs. Instead of firing the transitions at certain instances with zero duration, our approach is a continuous firing with flow $V(M(t),t)$ that may be externally generated by an input signal and may also depend on the continuous marking vector $M(t)$ [65]. The amount of marking change caused by a state change, in CPNs, is described as a nonlinear differential equation in (2) [65].

$$\dot{M}(t) = C^c(M(t))\,V(M(t),t), \quad M(t) \geq 0 \qquad (2)$$

where $C^c$ is the incidence matrix corresponding to the continuous weights. A transition is continuously fired with flow speed, $V(M(t),t)$, if the markings of all places in this transition are greater than zero. Note that (2) *is* a *differential equation* for representing the marking change amount. In order to represent positive and negative relationships in the man model, we use $\dot{M}_{in}(t)$ and $\dot{M}_{out}(t)\,(\in \dot{M}(t))$, which change amounts in an input and output place, respectively. If $\dot{M}_{in}(t)\ \dot{M}_{out}(t)$ is greater than zero, the firing vector represents a *positive* relationship between a place and the following one; however, if $\dot{M}_{in}(t)\ \dot{M}_{out}(t)$ is less than zero, it represents a *negative* relationship. If $\dot{M}_{in}(t)\ \dot{M}_{out}(t)$ is equal to zero, it means the states of one place or both places did not change, i.e. constant value at time $t$. Note that positive relationship means that the input flow amount of the following place increases (or decreases) as the output flow amount of the previous place increases (or decreases). Likewise, a negative relationship means that the input flow

amount of the following place decreases (or increases) as the output flow amount of the previous place increases (or decreases).

Various interfaces between continuous and discrete models are shown in Figure 1. Continuous places (states) can be transformed into discrete places through state quantization techniques, or vice versa, as shown in Fig. 1 (a). For example, decision styles of *vigilance* are discrete states transformed from continuous states such as *overconfidence*, *information process efficiency*, *stress*, and *preparedness*. Different quantization techniques are used in this type of transformation. An example [65] of the simple quantization transformation is that the threshold of the intermediate arc ($\Delta\chi$) decides the amount of the corresponding jump, and is used to interpret the discrete marking as a quantization of the continuous making. The development of quantization methods is beyond the scope of our focus in this dissertation. However, more sophisticated methods can be found in the literature, including fix-rate scalar quantization [Ree38], feedback vector quantization [DG81], multi-stage vector quantization [JG82], and universal quantization [Kie93]. The interface shown in Fig. 1 (b) enables us to control the flow of continuous states by discrete states (just like "on/off" switches). Off-line or normal service modes shown in Fig. 2, for example, are discrete system states that can turn on or off continuous group factors such as *information load*. Another type of interface shown in Fig. 1 (c) is used to represent discrete states affected by continuous states.

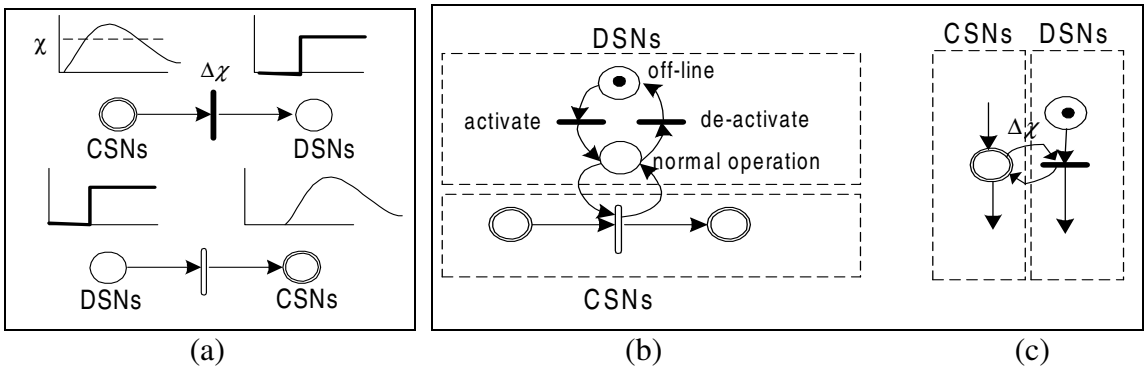


**Fig. 1. Interfaces between CPNs and DPNs [38]**

### 3.2. Machine Modeling

A key to our modeling of system (i.e. machine) behavior is achieving scalability for the simulation and computation of analytical models. For our purposes, it is neither practical nor necessary to fully capture the details of the software (e.g. intrusion detection tool) development or its operations. In its simplest form, we assume that three major types – *deployment, operation,* and *analysis* - are needed to characterize the major activities related to an information system (see Fig. 2). In the machine model, the information system can be a business application suite, a server with an operating system (e.g. Linux), or even an attack-monitoring tool. The deployment phase – which can also be considered as the preparation phase – can be further refined into multiple stages such as learning and installation. The operation phase can be formed in one of the three states: off-line, normal, or non-normal. The non-normal states are expected to lower business utility and are generalized as "degraded modes". Case by case, each state would have multiple input variables that would affect the state transitions and/or the information system's security assurance. The analysis phase is when the system and attack-related data are analyzed by the security defense team. As needed, the machine model can be scaled to add more stages.

After the state configuration for a system is defined, the next step is the creation of state transitions and their firing rules. Users use different applications to access data and/or to affect other applications or systems; intruders use similar behavior to launch attack programs. Their behavior determines the state transitions, including the creation and destruction of various states. It is fairly easy to simulate system behavior using a wide array of simulation and analytical tools – for example, stochastic/hybrid Petri Nets where state transitions and control or data flow can be represented according to distributions of significant events such as security attacks.

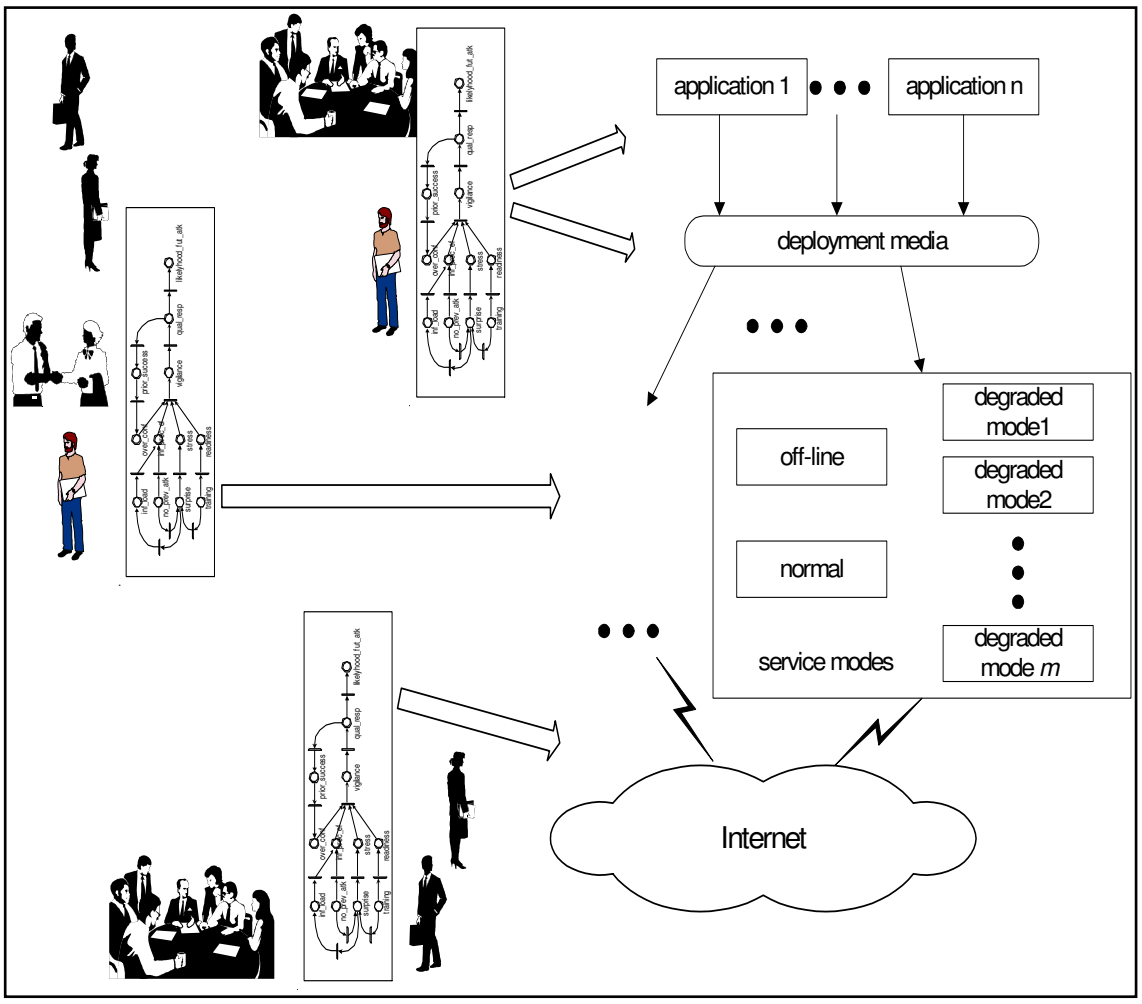


**Fig. 2. A high-level machine model connected to the group model for information assurance [38]**

### 3.3. Man Modeling

Current research on defender behavior treats a SIRT (Security Incident Response Team) as if it were a smoothly operating machine. In real-world situations, however, SIRTs vary greatly in terms of preparedness, ability to work together, and overall fitness of the team to rise to the challenge presented by the attackers. An adequate model of a defense will have to incorporate these human factors; otherwise, an important source of errors and costs will be omitted.

Two main characteristics of attack behavior considered in the Man-Machine Model are attack frequency and creativity. While there is no definitive model of intruder behavior, the work by Jonsson and Olovsson [44], [45] suggests that attack frequency follows an exponential distribution. We assume that creativity — directly correlated to attack severity — follow a normal distribution. On the basis of these two assumptions, we build a black box model of attack behavior, which is sufficient given our focus on SIRTs.

We assume that a SIRT will be formed explicitly in response to attacks. Such teams are commonly discussed in the literature on cyber incidents [90] and network forensics [93]. To determine how teams contribute to the defense in the Man-Machine Model, it is necessary to develop a model of group (team) dynamics that determine the development rate for detection (and response) and quality of the detection (and response). Our man model is presented in Fig. 3.

One of the key variables in the model is Teamwork Quality (TWQ) [33], which refers to the degree to which the team members communicate, coordinate, support, make an effort, and act cohesively together. That is, it is a measure for the quality of collaboration in teams. As the description explains, it is comprised of five factors (one factor is under investigation for adoption). *Communication* measures how frequent the team members communicate directly and personally with each other, whether or not they have mediators through whom much of the communication is conducted, etc.; *coordination* measures how closely harmonized the work done on sub-tasks within the experiment project was, etc.; *mutual support* measures the ease and speed with which conflicts were resolved when they came up, etc.; *effort* measures how high a priority the project was for each team member, etc.; *cohesion* measures how important project involvement was for each team member, etc.

A system of factors influences the TWQ enacted by the team. *Perceived mental workload* (PCMW) is the degree to which the team is under physical, mental, and temporal demands. It also includes frustration level, performance satisfaction level, and mental and physical effort level. *Preparedness* (PRPD) is the degree to which the team is

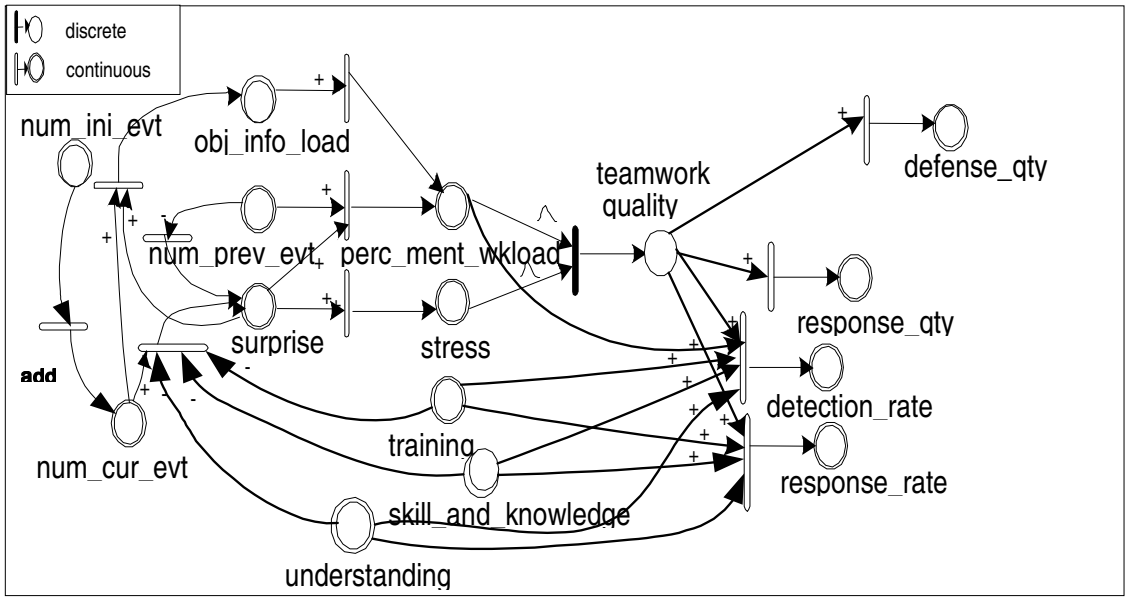


**Fig. 3. Man model**

ready to deal with attacks. *Stress* (STRS) is the mean level of negative arousal the team has. *Overconfidence* (OVCF) is the degree to which the team assumes it can handle any problem without much difficulty.

Tracing the system to the determinants of the four factors, Perceived Mental Workload (PCMW) has a positive linear relationship with *Objective Information Load (OBIN)* (the amount of cues and messages to be processed) in an inverted-U curve. The number of prior attacks is positively related to Perceived Mental Workload (PCMW), as the team members are "on edge" and ready to recognize attacks. As attacks are fewer, the team is less likely to be ready to recognize cues signaling an attack. Both Objective Information Load and Stress increase in a positive linear manner by *Surprise*. Objective Information Load increases because Surprise causes the team to scan for information, picking up many relevant and irrelevant items. Stress increases because Surprise arouses the team.

Preparedness has a positive linear relationship with *Training* (TRNG). Teams that train for attacks are more likely to be prepared than those that do not. Stress reduces Preparedness in a negative linear fashion because of Stress reactions. Surprise is a conduit for two indirect effects on Stress. The number of previous attacks reduces Surprise; other things being equal, teams that have experienced prior attacks are less likely to be surprised. Also, Preparedness reduces Surprise as well, for obvious reasons. Overconfidence is a positive function of Prior Success, which is a function of Quality of Response. Teams that have handled problems effectively in the past tend to become confident. This is not a problem except for the high range of Confidence (Overconfidence).

Providing appropriate training and education is often critical to the successful implementation of information security within an organization [5], [6]. Without appropriate training or education to every personnel (including security staff), success in information security can not be expected. In this dissertation, we consider training only. Training, as another key variable, has a goal of building knowledge and skills to facilitate the job performance [88]. Therefore, training is assumed to have a positive relationship with *Skill and Knowledge* (SKKN). The more trained the security staff, the more they can gain skills and knowledge associated with current security issues. However, allocating an appropriate amount of time and money for training is all too often overlooked as a requirement for preventing stagnation of staff expertise, but training is absolutely essential for keeping those skills current with technology issues [91]. Without keeping the skills and knowledge up-to-date with current issues, it is hard to effectively and efficiently respond against different kinds of security incidents. That is one of the benefits which training can provide.

Not only can training can have a positive impact on skill and knowledge, but it can reduce the errors made by different types of users - including system administrators. The usually forgotten but important factor of *Human Error* is emphasized by Schultz and Shumway [73], stating that, “Granted, the incident could be very serious and potentially costly, but human error costs organizations far more than security-related

incidents do." An example of human error is input errors and omissions. According to the National Computer System Security and Privacy Advisory Board, 1991 Annual Report [56], errors and omissions are top rated in the economic loss attributed to this threat. These errors are caused by many types of users such as end users, system operators, and programmers. A sound awareness and training program can help an organization reduce the number and severity of errors and omissions, which are an important threat to data and system integrity [58].

Skill and Knowledge (SKKN), as another key variable, poses a solid impact on *Detection Rate* (DTR) and *Response Rate* (RPR), for obvious reason. During a research committee meeting, two experienced network/system administrators in the Computer Science Department at Texas A&M University asserted that skill sets and knowledge are extremely important in security-related attack detection and response. To measure skill sets and knowledge, we should have SKKN that covers both skill sets and knowledge.

### 3.4. Example: TCP SYN Flooding Denial of Service (DoS) Attack

This section presents a hybrid man-machine interaction model for a TCP SYN flooding Denial of Service (DoS) attack using the proposed model. The hybrid TCP SYN flooding attack model is composed of three model components: monitoring, control, and group behavior modules. The monitoring module shown in Fig. 4 (a) models the process of monitoring the input traffic patterns of service systems from the Internet using a backward propagation feedback control algorithm, presented in [92], as an initial intrusion detection process. The algorithm detects abnormal traffic patterns, called "hot spots," inside a machine (upper dot-line box) shown in Fig. 4 (a-1) and outside the machine (bottom dot-line box) shown in Fig. 4 (a-2). Total traffic rate (total_traff_rate) is determined by these two inputs.

The control module depicted in Fig. 4 (b) is used to control the throttle of network (throttling_on/off) via traffic a control indicator (under_attack/normal) for the real system states through human verification of a TCP SYN flooding attack. The verification is performed based on total traffic rate in the monitoring module. Through the notification transition (notify) to the control module, if the TCP SYN flooding attack really happened, the systems' state is changed to 'under attack' (i.e. pass a token into under_attack). Under attack, administrators turn on the throttling to reduce network traffic (i.e. pass a token in throttling_on from throttling_off) to block suspicious packets. This is one of degraded modes under attack, as shown in Fig. 2.

The group behavior module is presented in Fig. 4 (c), which interacts with the system modules such as the control module. Typical group behavior or to-do list under a TCP SYN flooding attack is explained in [15]. The administrators verify potential victim systems when the monitoring software makes warning flags, then identify, test, install, and execute effective defense mechanisms. If a large number of systems (measured by num_cur_atk) is attacked at the same time, both *information load* (info_load) and *surprise* of administrators increases, and they finally affect the *decision style*.

Based on three styles of the decisions, quality of response (quality_response) and rate of response development (development_rate_ for_response) are determined, which are themselves key factors in how effectively the group responded with defensive mechanisms and how fast the systems can get back to normal operations by deactivating the throttling from on to off (i.e. moving a token from throttling_on to throttling_off), respectively.

A PN tool to support hybrid PNs, called Visual Object Net ++ [22], is used to develop the example. The tool supports mixed continuous and discrete event PNs.

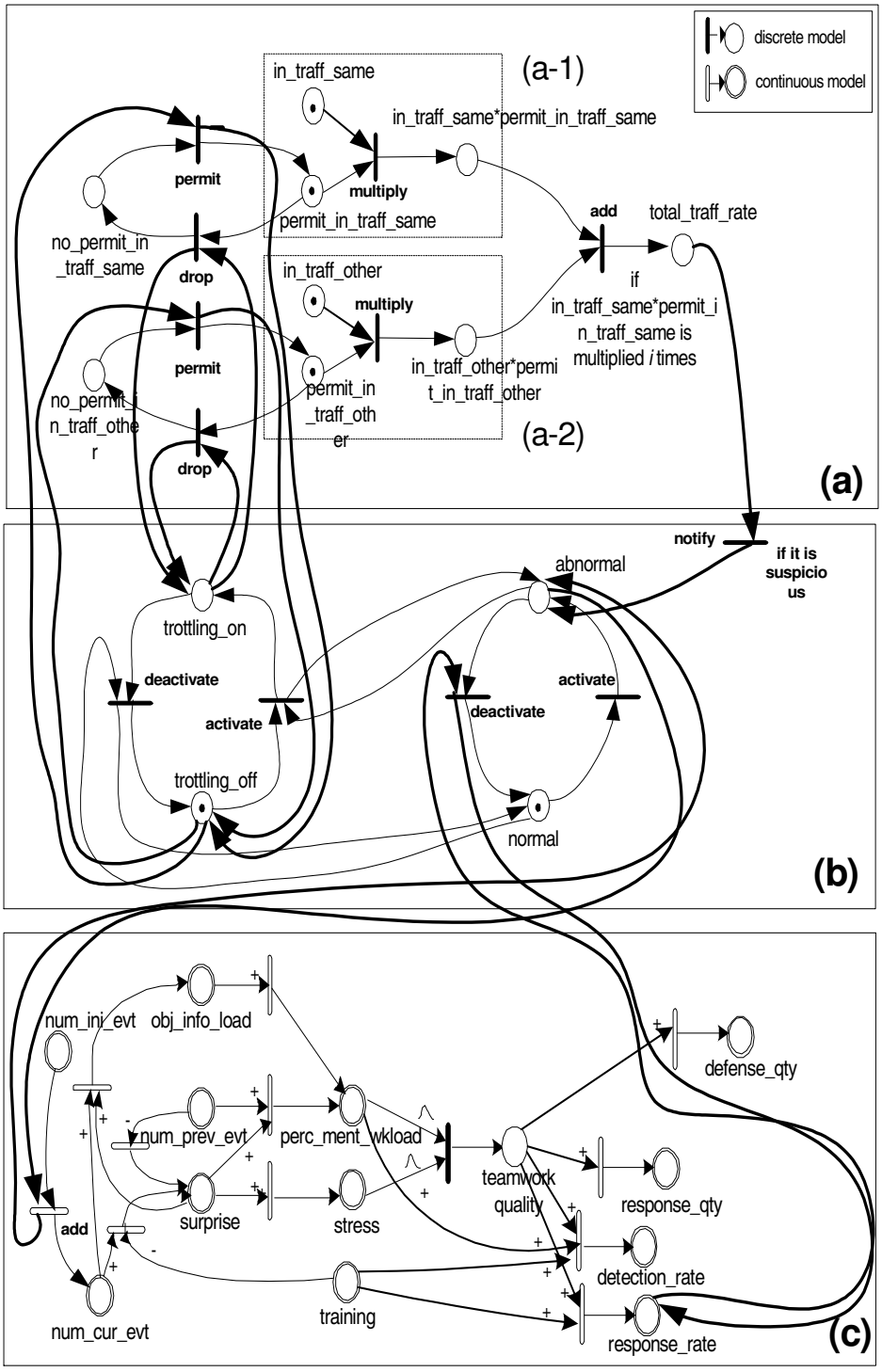


**Fig. 4. M$^3$: A PN modeling diagram**

### 3.5. Research Focus

Our modeling aims at a tradeoff analysis between resource allocations and the impact of team performance (efficiency). To make more accurate decisions, organizations should consider the most critical and impacting factors to team efficiency factors: Detection Rate (DTR) and Response Rate (RPR). Based on the current literature and experts' opinions, those critical factors include Skill and Knowledge (SKKN), Training (TRNG), and Teamwork Quality (TWQ). Thus, we shift focus to these three factors to measure team efficiency throughout this dissertation. Our research focus throughout this dissertation is summarized in Fig. 5.

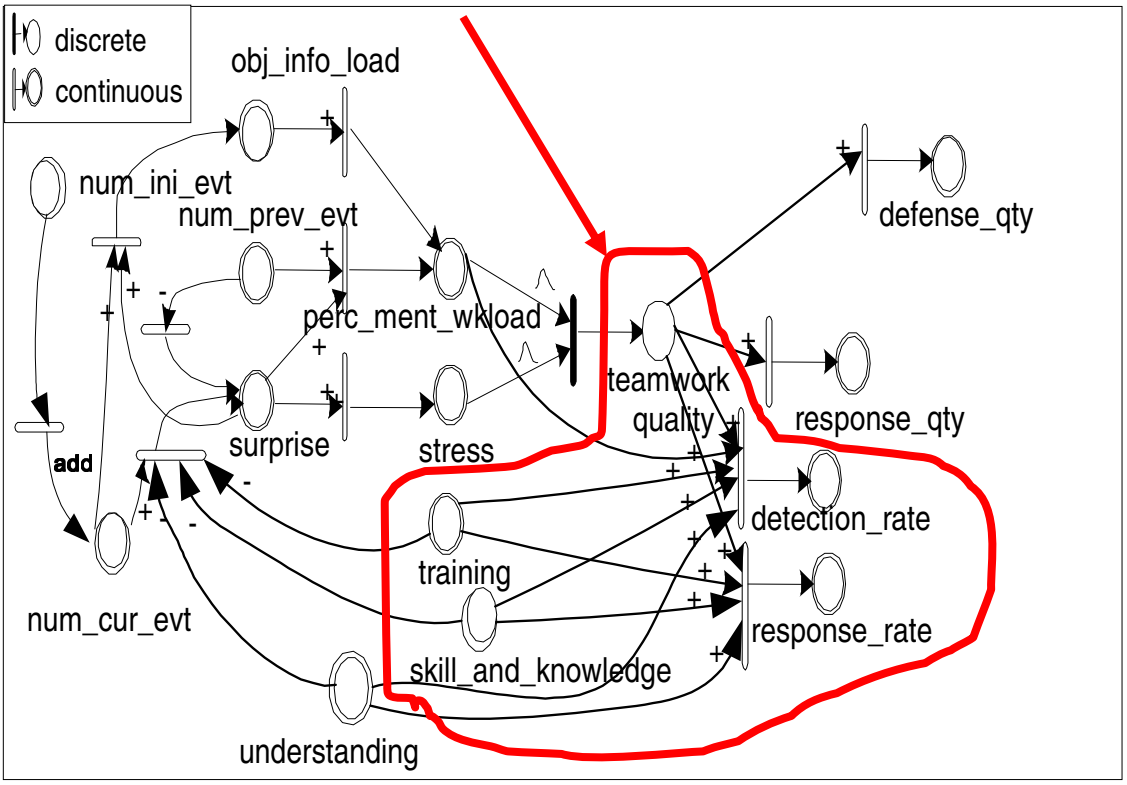


**Fig. 5. Research focus**

# 4. EXPERIMENTS

## 4.1. Hypotheses

*Hypothesis 1: Skill and Knowledge (SKKN) of security personnel is statistically significantly related to the Detection Rate (DTR)*

*Hypothesis 2: Training (TRNG) and Skill and Knowledge (SKKN) of security personnel are statistically significantly related to the Response Rate (RPR)*

*Hypothesis 3: High-SKKN level groups provide better detection (i.e. faster detection rate) than Medium-SKKN level groups*

*Hypothesis 4: High-TRNG level groups provide better response (i.e. faster response rate) than Medium-TRNG level groups*

The first hypothesis sets Skill and Knowledge (SKKN) as independent variable and Detection Rate (DTR) as dependent variable. In other words, SKKN is linearly (or nonlinearly) related to DTR. This hypothesis is derived from the experience of two network/system administrators in the Computer Science Department at Texas A&M University; they assert that both trust and skill sets (and knowledge) are among the most important factors in security-related attack detection and response.

The second hypothesis, also as important as the first, is that Training (TRNG) and Skill and Knowledge (SKKN), and Response Rate (RPR) are dependent. This hypothesis is derived from current literature and other sources, including two network/system administrators in the Computer Science Department at Texas A&M University. ISO/IEC 17799 [5], [6] claim that providing appropriate training and education is often critical to the successful implementation of information security within an organization. A NIST Handbook [58], in a similar manner, states that a sound awareness and training program can help an organization reduce the number and severity of errors and omissions, which are an important threat to data and system integrity.

This hypothesis is as important as the first one since according to the degree of effect of Training and Skill Sets and Knowledge on security personnel's performance

(i.e. detection rate) each organization may revise its decisions on spending money on (1) their own training programs, and (2) human resource allocations in the hope of improving the level of security defense in both detection and response.

### 4.2. Experimental Setting

Week-long security experiments were held twice in 2003 (April 11 through 18) and 2004 (April 16 through 23). With the help of Dr. Michael Grimaila, who taught the Business Information Security course (INFO689) at Texas A&M University, Dr. Marshall Scott Poole, Dr. Hoh Peter In, and me set up the security experiments at the Security Lab (324E), located at the Department of Information & Operations Management (INFO).

The human subjects as defenders were graduate students enrolled in INFO689. In the spring of 2003, we had six different groups of three to four students. Similarly in 2004, we had five different groups containing three to four students. There were two types of human subjects – either technical or managerial. Technical people worked mostly on setting up the experiments and were actively involved in the experiments. Meanwhile, managerial people worked mostly on planning a security budget, establishing security policies, and documenting their work. This is because these experiments were linked together with their own security project in INFO 689.

The attackers were students enrolled in the Advanced Networks and Security course (CPSC665). Attackers at the Department of Computer Science were called the Black Team; as a part of their job throughout the class, they were instructed to launch security attacks against the defense groups at the INFO Department. The organization of the Black Team was not known to the defenders.

However, both defenders and attackers had to follow certain rules, which are presented in Appendix B. The rules were experimental rules applied to every subject participating in the security experiments; however, they were general rules – and may be common to other security experiments. The rules instructed participants on how and how often to fill out the data collection forms we provided. Since these experiments were part of their original job in their class exercise, they had their own rules to follow as well as

the rules given by us. Defenders at INFO Department had read and obeyed the rules, made up by their class instructor, including goals, initial conditions, objectives, and rewards. Attackers in the CS Department had their own lab policies and procedures [66] as well.

Both defenders and attackers used intrusion detection tools and hacking tools, respectively. Each defense group might have its own preference over the tools it uses. However, generally speaking, the security tools installed (loaded) in each group's computers were Bastille Linux Firewall, TripWire, Nessus, and Nmap. Some groups installed other tools as well, including Snort, Ethereal, Rootkit Checker, and Logcheck.

While each group had different kinds of tools loaded into its system, every group was loaded with the Redhat Linux as its OS. However, due to preference differences, each group had a different version of Redhat Linux (generally, version 7.3 or 8.0). The system of each group was also loaded with various services, such as Web services. Thus, each system had installed Web, SSH, FTP, NFS, Email, DNS, News, and Database servers. Every system had Apache as its Web server and MySQL as its Database server, even if systems were running different versions.

### 4.3. Experimental Process

The protocol for these experiments, presented in Appendix A, guided the students on how to participate in the experiments and record demographic information and what they did throughout the entire week. As the experiment coordinator, I followed the protocol and helped the students to fill out the forms we distributed. Whenever they had questions, they were directed to contact me by phone or email in order to receive prompt answers. Since poor or improper responses on the forms was a concern, participants were systematically taught how to best fill out the forms (content and intervals) using examples. The defenders were directed to fill out five different forms – Background, Preparation, Activity Record, Teamwork Quality, and Post-defense forms, as displayed in Appendix D. Before the security experiments begin, they were directed to fill out Background and Preparation forms. During the experiments, they were directed to fill

out Activity Record and Teamwork Quality forms. After the experiments, they were directed to fill out Post-defense form.

After the defenders had the concepts of what they could do, the security experiments began. As soon as the experiments began, the defenders were busy in performing their job – protection of their systems and information. Their working hours toward this experimental job varied according to their other class schedules and willingness. For each activity they undertook during the period they worked, they recorded it on the Activity Record form when they detected or found a suspicious event or problem. They recorded the time they found the problem, the nature of the problem, and the way they found the problem. They recorded if they think the problem a suspicious security attack incidents, and they were asked to record why they think so. For a response activity, they recorded what actions they took as a reaction against the suspicious problem, what time they took the actions, and why they did.

Besides, they recorded any other activity not mentioned in the activity of detection and response that could be significant. Sources of information they received during that activity and type of the activity were also recorded. Examples of sources could be '/etc/initd.conf', 'snort log', etc. Example of type are email, phone calls, hearing from team members (off-line), discussion (off-line), electronic bulletin board such as Yahoo Messenger, and other types. In addition to sources and type, they recorded the number of messages and phase (e.g., detection phase). They also recorded downtime they had during the period they worked. The downtime was recorded based on sources – server, system (computer), service, network, application, etc. They recorded time duration that a downtime occurred, and the reasons.

While they were working for detection of any suspicious event or problem, they were directed to fill out the information regarding how they worked together on Teamwork Quality form, so that we can better understand how they worked as an incident response team. They recorded the degree, to which they were surprised during the experiments due to the security attack incidents they found, on a scale of 1 (not surprised at all) to 10 (extremely surprised), and the confidence level each team member

had for their team based on beliefs about the team performance, on a scale of 1 (to no extent) to 5 (to a great extent). Stress level they had during the period they worked was recorded based on fifteen judgments – demanding, pressured, hectic, calm, relaxed, many things stressful, pushed, irritating, under control, nerve-wrecking, hassled, comfortable, more pressure than I'd like, smooth running, and overwhelming. Level of vigilance, describing how their team worked together when they made decisions or solve detected problems, were recorded on a scale of 1 (strongly disagree) to 6 (strongly agree). Perceived mental workload, composed of six subcomponents (mental demand, physical demand, temporal demand, performance, effort, and frustration level) was recorded on a scale of 1 (low) to 10 (high).

Teamwork quality variables [33] which represent teamwork quality of their team during the period they worked were also recorded on Teamwork Quality Form. They include communication, coordination, cohesion, efforts, mutual support, and balance of member contributions.

While they were working in the lab by participating in their security experiments and filling out two forms – Activity Record and Teamwork Quality Forms, I checked their activity by visiting the lab in which the security experiments were held. Whenever they had questions regarding the forms, I gave them succinct and accurate answers. In case the forms were not available, I created additional copies and put them in the lab. Since the lab was guarded by a secretary and the instructor, we didn't lose any completed forms.

I collected all the data forms after each security experiments week ended. Additionally, I obtained the students' log files thanks to the instructor's help. The students' log files were expected to be useful to check out what tools each team used in the experiments, who participated, what their roles were, and most importantly the existence of missing information from their data forms.

After collecting their data forms, I found out that there were complexities in collecting considerable amounts of meaningful data from defenders. The experiments were a complex event. While defenders worked for detecting any suspicious events or

problems and responding against them whenever detecting, they had to fill out two forms simultaneously as well as they secured their e-commerce web servers so that the web servers would not be compromised from attack incidents. In other words, they performed not only filling out two data forms for our data collection purpose, but secured and managed web servers for their own class project. Even data form filling could be a daunting task for defenders because they had to answer many question items on the forms with vigilance and quickness, and in addition, it took some time for them to get a concept of all the question items. Besides, I could rarely meet them in the Lab since all the defenders had their own schedule.

As a result of complexities of security experiments and data form filling, unpleasant things occurred. For instance, sometimes, they filled out the Teamwork Quality form appropriately, but they forgot to fill out the Activity Record form, which should have been filled out simultaneously. Often, they failed to provide important data – especially timing data – while successfully filling out the rest of a form. In these cases, their log files would have helped me check out against their original data forms; however, they were usually not helpful, meaning I rarely found the missing information in their log files either.

I could collect only 22 (15 and 7 from 2003 and 2004, respectively) meaningful data out of two-time security experiments due to complexities of experiments and data form filling, and carelessness of defenders, etc. The data collected are explained and analyzed in detail in the next section; and, suggestions on much more meaningful data collection are presented in Section 6.3.

# 5. DATA ANALYSIS

## 5.1. Data Collection

Both the technical and managerial people participated in the security experiments and filled out the forms we provided. Table 1 represents those who not only participated in the experiments but filled out the given data forms with the inclusion of performance-related data (e.g. detection time), which was a must. It represents the year they participated in the security experiments, each subject's group name, his/her code number (which were given for the purpose of his/her anonymity), his/her role in the duration of the experiments, and his/her type – either technical or managerial.

Out of twenty two samples collected, only three data were the ones filled out by managerial people who might not be able to provide meaningful information. However, since these experiments are not for individual project-type experiments, it is quite possible they may have asked their team members about what happened, what actions they took, and how they were supposed to react, and so on. This reasoning became a reality when I discussed issues with both managerial and technical people after the experiments. Thus, even if not all the data came from the technical people, we accepted all completed forms provided for the data analysis. However, whenever we found something odd, especially in hypothesis testing, we could eliminate the data for better or accurate analysis.

For the hypothesis testing and regression analysis, we collected the empirical data, both original and weighted (both are displayed in Table 2 and 3, respectively). The collected data were used to construct statistical models of intrusion detection and incident analysis for the efficiency of the security defense process. SAS regression analysis was used on the data. The regression analysis was also used to test the hypotheses we presented in Section 4.

**TABLE 1**

**Human Subjects and Their Roles**

| year | group name | code number | role | tech/mngr |
|---|---|---|---|---|
| 2003 | Alpha | 1 | security programmer | T |
| | | 1 | | T |
| | | 16 | server administrator | T |
| | | 17 | security budget & policy analyst | N |
| | Bravo | 3 | networking specialist | T |
| | | 3 | | T |
| | | 11 | team leader | T/N |
| | Charlie | 6 | policy group | N |
| | | 7 | tech group | T |
| | | 7 | | T |
| | | 8 | policy group | N |
| | Echo | 4 | project manager | T/N |
| | | 9 | system administrator | T |
| | | 10 | team leader | T/N |
| | | 18 | security officer | T |
| 2004 | Charlie | 3M | technical leader | T |
| | Echo | 2A | technical/security analyst | T |
| | | 2A | technical/security analyst | T |
| | | 2K | technical administration | T |
| | Golf | 2F | technical/policy maker | T/N |
| | | 2G | technical administrator | T |
| | | 3O | technical | T |

**TABLE 2**

**Original Data**

| Skill & Knowledge | | Training | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| system experience | security experience & knowledge | security training | other training | security project understanding | number of tools | vigilance | teamwork quality | detection rate | response rate |
| 6 | 2 | 3 | 2 | 3 | 1 | 3.7 | 4.7 | 1 | |
| 6 | 2 | 3 | 2 | 3 | 1 | 3.6 | 4.7 | 0.1 | 0.1 |
| 5 | 1 | 3 | 1 | 4 | 1 | 4.2 | 4.9 | 0.1 | 0.1 |
| 3 | 1 | 3 | 2 | 2 | 1 | 4.5 | 5.3 | 0.1 | 0 |
| 6 | 1 | 3 | 1 | 5 | 1 | 4.8 | 5.7 | 1 | 0.2 |
| 6 | 1 | 3 | 1 | 5 | 1 | 5 | 5.9 | 1 | 0.2 |
| 3 | 1 | 3 | 6 | 3 | 1 | | | 0.1 | 1 |
| 2 | 1 | 3 | 1 | 4 | 3 | | | 0.2 | 0 |
| 1 | 1 | 3 | 1 | 3 | 3 | 4.8 | 5.5 | 0.1 | 0.1 |
| 1 | 1 | 3 | 1 | 3 | 3 | 4.5 | 5.2 | 0.1 | 0 |
| 1 | 1 | 3 | 1 | 4 | 3 | 4.5 | 5.1 | 0.2 | 0 |
| 3 | 1 | 3 | 1 | 1 | 1 | 4.3 | 4.9 | 0.2 | 0.1 |
| 3 | 3 | 3 | 3 | 2 | 1 | 4.2 | 5.3 | 0.1 | 0.1 |
| 3 | 1 | 3 | 3 | 3 | 1 | 4.1 | 4.3 | 0.1 | 0 |
| 2 | 1 | 3 | 1 | | 1 | 5 | 4.8 | 0 | 1 |
| 2 | 1 | 3 | 1 | 5 | 5 | 4.2 | 4.7 | 0 | 0.1 |
| 6 | 6 | 4 | 2 | 5 | 5 | 3.7 | 4.6 | 1 | 1 |
| 6 | 6 | 4 | 2 | 5 | 5 | 3.7 | 4.6 | 1 | 1 |
| 1 | 2 | 4 | 3 | 4 | 5 | 3.9 | 4.4 | 0 | 1 |
| 1 | 1 | 3 | 1 | 5 | 4 | 3.9 | 5.2 | 0 | |
| 3 | 1 | 5 | 3 | 5 | 4 | 5 | 5.9 | 0 | |
| 3 | 1 | 3 | 4 | 5 | 4 | 5 | 5.9 | 0 | |

## 5.2. Measurement

*Skill & Knowledge (SKKN)* is a compound variable of skill and knowledge. Skill is defined as the ability to do something well, especially through learning and practice; knowledge is defined as understanding one has obtained, especially through learning or experience. In our context, SKKN is the ability to perform well protection of information systems such as web servers and understanding one has obtained through various experience or learning. It was measured by having human subjects indicate the number of months they had experience with systems such as UNIX and LINUX, and the number

**TABLE 3**

**System Experience Scale**

| System Experience | 1: <= 6 mo | 2: <= 12 mo | 3: <= 24 mo | 4: <=36 mo | 5: <= 60 | 6: >60 mo |
|---|---|---|---|---|---|---|

**TABLE 4**

**Security Experience and Knowledge Scale**

| Security Experience and Knowledge | 1: <= 0.5 mo | 2: <= 2 mo | 3: <= 6 mo | 4: <=12 mo | 5: <= 24 | 6: >24 mo |
|---|---|---|---|---|---|---|

of months of security experience and knowledge they had. The number of months they had experience with systems, presented in Table 3, are rated on the scale of one to six. The number of months of security experience and knowledge they had, presented in Table 4, are rated on the scale of one to six.

Since relevance of Security experience and knowledge to SKKN could be assumed to be greater than that of System experience to SKKN, a weight of 0.3 was given to System experience, and 0.7 to Security experience and knowledge. The weighted System experience, Security experience and knowledge, and their summed value of SKKN are shown in Table 7.

*Training (TRNG)* is defined as the level of training experience which includes both off-line (e.g., class, seminars. conferences, telephone, etc.) and on-line (e.g., internet, e-mail, etc.) training experience. It was measured by having human subjects indicate the number of months they had experience with security training, and the number of months of they had experience with other training. The number of months they had experience with security training, presented in Table 5, are rated on the scale of one to six.

**TABLE 5**

**Security Training Scale**

| | | | | | | |
|---|---|---|---|---|---|---|
| Security training | 1: <= 0.5 mo | 2: <= 2 mo | 3: <= 4 mo | 4: <=6 mo | 5: <= 12 | 6: >12 mo |

**TABLE 6**

**Other Training Scale**

| | | | | | | |
|---|---|---|---|---|---|---|
| Other training | 1: <= 1 mo | 2: <= 4 mo | 3: <= 12 mo | 4: <=24 mo | 5: <= 48 | 6: >48 mo |

The class – INFO689 – they took was counted four months of security training experience because the class could be considered security training. The number of months of they had experience with other training, presented in Table 6, are rated on the scale of one to six.

Other training includes network or system training that could be useful to perform security experiments. Relevance of Security training to TRNG is greater than that of Other training to TRNG, a weight of 0.7 was given to Security training, and 0.3 to Other training. The weighted security training and other training, and their summed value of TRNG are shown in Table 7.

*Security project understanding* was measured by having human subjects indicate the degree to which they understand the project they will perform in security experiments, on a scale of 1 (little understanding) to 5 (thorough understanding). The detailed description of each scale is shown in Appendix D. Security project understating data values are shown in Table 7.

*The number of tools* was measured by having human subjects indicate how many security tools their team installed on their systems to use as tools for protecting their

systems. The tools include Bastille Linux Firewall, Tripwire, Nessus, Nmap, Ethereal, Rootkit Checker, etc. The number of tools data values are shown in Table 7.

*Vigilance* was measured by having human subjects indicate how their team works together as they decide or solve detected problems, on a scale of 1 (strongly disagree) to 6 (strongly agree). Subjects indicated their degree of agreement with ten statements as its measurements, which are shown in Appendix D. Vigilance data values are shown in Table 3.

*Teamwork Quality (TWQ)* [33] was measured by having human subjects indicate the degree their team performed their tasks together. TWQ, a complex variable, is composed of six facets (communication, coordination, cohesion, effort, mutual support, and balance of member contributions) [33] which have same scale of 1 (strongly disagree) to 6 (strongly agree), but have different measurements. Cronbach's alpha coefficient of TWQ facets were 0.91 [33], which are excellent according to a rule of thumb [24]. Due to the high reliability of measurements, we chose the variable for our research. Communication measures if there is sufficiently frequent, informal, direct, and open communication. Coordination measures if individual efforts are well structured and synchronized within the team. Cohesion measures if team members are motivated to maintain the team, and if there is team spirit. Effort measures if team members exert all their efforts to the team's tasks. Mutual effort measures if team members help and support each other in carrying out their tasks. Balance of member contributions measures if all team members are able to bring in their expertise to their full potential. Measurements of six facets are shown in Appendix D. TWQ data values are shown in Table 7.

**TABLE 7**

**Weighted Data**

| weighed system experience | weighed security experience & knowledge | skill & knowledge total value | weighted security training | weighted other training | training total value | security project understanding | number of tools | vigilance | team work quality | detection rate | response rate |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.8 | 1.4 | 3.2 | 2.1 | 0.6 | 2.7 | 3 | 1 | 3.7 | 4.7 | 1 | |
| 1.8 | 1.4 | 3.2 | 2.1 | 0.6 | 2.7 | 3 | 1 | 3.6 | 4.7 | 0.1 | 0.1 |
| 1.5 | 0.7 | 2.2 | 2.1 | 0.3 | 2.4 | 4 | 1 | 4.2 | 4.9 | 0.1 | 0.1 |
| 0.9 | 0.7 | 1.6 | 2.1 | 0.6 | 2.7 | 2 | 1 | 4.5 | 5.3 | 0.1 | 0 |
| 1.8 | 0.7 | 2.5 | 2.1 | 0.3 | 2.4 | 5 | 1 | 4.8 | 5.7 | 1 | 0.2 |
| 1.8 | 0.7 | 2.5 | 2.1 | 0.3 | 2.4 | 5 | 1 | 5 | 5.9 | 1 | 0.2 |
| 0.9 | 0.7 | 1.6 | 2.1 | 1.8 | 3.9 | 3 | 1 | | | 0.1 | 1 |
| 0.6 | 0.7 | 1.3 | 2.1 | 0.3 | 2.4 | 4 | 3 | | | 0.2 | 0 |
| 0.3 | 0.7 | 1 | 2.1 | 0.3 | 2.4 | 3 | 3 | 4.8 | 5.5 | 0.1 | 0.1 |
| 0.3 | 0.7 | 1 | 2.1 | 0.3 | 2.4 | 3 | 3 | 4.5 | 5.2 | 0.1 | 0 |
| 0.3 | 0.7 | 1 | 2.1 | 0.3 | 2.4 | 4 | 3 | 4.5 | 5.1 | 0.2 | 0 |
| 0.9 | 0.7 | 1.6 | 2.1 | 0.3 | 2.4 | 1 | 1 | 4.3 | 4.9 | 0.2 | 0.1 |
| 0.9 | 2.1 | 3 | 2.1 | 0.9 | 3 | 2 | 1 | 4.2 | 5.3 | 0.1 | 0.1 |
| 0.9 | 0.7 | 1.6 | 2.1 | 0.9 | 3 | 3 | 1 | 4.1 | 4.3 | 0.1 | 0 |
| 0.6 | 0.7 | 1.3 | 2.1 | 0.3 | 2.4 | | 1 | 5 | 4.8 | 0 | 1 |
| 0.6 | 0.7 | 1.3 | 2.1 | 0.3 | 2.4 | 5 | 5 | 4.2 | 4.7 | 0 | 0.1 |
| 1.8 | 4.2 | 6 | 2.8 | 0.6 | 3.4 | 5 | 5 | 3.7 | 4.6 | 1 | 1 |
| 1.8 | 4.2 | 6 | 2.8 | 0.6 | 3.4 | 5 | 5 | 3.7 | 4.6 | 1 | 1 |
| 0.3 | 1.4 | 1.7 | 2.8 | 0.9 | 3.7 | 4 | 5 | 3.9 | 4.4 | 0 | 1 |
| 0.3 | 0.7 | 1 | 2.1 | 0.3 | 2.4 | 5 | 4 | 3.9 | 5.2 | 0 | |
| 0.9 | 0.7 | 1.6 | 3.5 | 0.9 | 4.4 | 5 | 4 | 5 | 5.9 | 0 | |
| 0.9 | 0.7 | 1.6 | 2.1 | 1.2 | 3.3 | 5 | 4 | 5 | 5.9 | 0 | |

*Detection Rate (DTR)* was measured by calculating one over the detection time, which is equal to the time when human subjects detected any suspicious event or problem minus the time when they began working. DTR is between zero and one. DTR data values are shown in Table 7. Similarly, *Response Rate (RPR)* was measured by calculating one over the response time, which is equal to the time when human subjects reacted against any detected problem minus the time when they detected. Like DTR, Response rate (RPR) is between zero and one. RPR data values are shown in Table 7.

### 5.3. Correlation Analysis

Several potential candidates for the independent variables are presented here. To assess how strongly each independent variable statistically related to the dependent variables, we performed a correlation analysis. Results of correlation analysis between the independent and dependent variables are presented in Table 8.

**TABLE 8**

**Summary of the Correlation Analysis Results**

| | *SKKN* | *TRNG* | *UDST* | *NOTL* | *VIGL* | *TWQ* | *DTR* | *RPR* |
|---|---|---|---|---|---|---|---|---|
| *SKKN* | 1 | | | | | | | |
| *TRNG* | 0.289791 | 1 | | | | | | |
| *UDST* | 0.226678 | 0.187006 | 1 | | | | | |
| *NOTL* | 0.192665 | 0.349648 | 0.613097 | 1 | | | | |
| *VIGL* | -0.54192 | -0.08587 | 0.136105 | -0.18587 | 1 | | | |
| *TWQ* | -0.29351 | 0.002003 | 0.223032 | -0.09843 | 0.77402 | 1 | | |
| *DTR* | 0.716134 | -0.03214 | 0.349977 | 0.008807 | -0.1865 | 0.009058 | 1 | |
| *RPR* | 0.47751 | 0.729237 | 0.362548 | 0.376126 | -0.26469 | -0.39715 | 0.292485 | 1 |

We can see that the strongest correlation with DTR is SKKN with a correlation in excess of 0.72. Likewise, the strongest correlation with RPR is TRNG with a correlation in excess of 0.73. Even though SKKN does not have a strong correlation with RPR, it is moderately correlated with the coefficient value of 0.48.

However, surprisingly neither TWQ and DTR, nor TWQ and RPR are correlated high. Even TWQ and RPR are negatively correlated. The correlation coefficient between these two variables is -0.4, meaning that when TWQ increases by one unit, RPR decreases by a 0.4-unit. In other words, high TWQ does not necessarily increase RPR in incident response process; on the contrary, high TWQ cause RPR to be decreased. These surprising results could be explained by examining correlation between TWQ and SKKN, and TWQ and TRNG. Correlation coefficients between SKKN and TWQ, and TRNG and TWQ are -0.3 and 0, respectively. These results mean that the higher SKKN, the lower TWQ with a rather low (0.3) correlation coefficient; and, there is no correlation between TRNG and TWQ. To summarize, it is probable that a defender who is highly skillful and knowledgeable do not want to or need to work in teams. Thus, the defender is willing to take control of responding against detected security incidents rather than work together with his or her team members due to several possible reasons, including that he or she lacks of trust on less skillful and knowledgeable team members. However, it does not always guarantee RPR would go higher since SKKN is moderately – not highly – correlated (0.48) with RPR.

If we examine correlations between DTR and TWQ, SKKN and TWQ, and SKKN and DTR, there are some findings: 1) TWQ is not related with DTR, 2) one-unit increase in SKKN causes 0.3-unit decrease in TWQ, and 3) SKKN is highly correlated with DTR. These findings suggest that a team needs to hire highly skillful and knowledgeable defender to increase DTR, irrespective of how well the team works together in intrusion detection process.

To summarize, when dealing with complex, critical problem such as intrusion detection and incident response, interaction with other team members can be a distraction. As long as a team has a highly skillful and knowledgeable defender, it would

be better the defender take control of detecting intrusions and responding against those detected intrusions. However, in case of response, it would be better a highly trained defender take control of response rather than a highly skillful and knowledgeable defender since RPR is more highly correlated with TRNG than SKKN. One possible reason for that is when responding against detected intrusions, more practical experience obtained through security training or other related training would be needed more since problems defenders would face with are real, practical problems. In other words, a practical application of our knowledge and skill onto the problem of response is essential to solving the problem incident response teams are facing with.

## 5.4. Hypothesis Testing

### 5.4.1. Hypothesis 1

*Hypothesis 1: Skill and Knowledge (SKKN) of security personnel is statistically significantly related to Detection Rate (DTR)*

Based on the results of the correlation analysis, we set up Hypothesis 1. The linear regression model for the hypothesis is as follows:

$$DTR = \beta_0 + \beta_1 * SKKN + \text{random error}$$

The null and alternative hypothesis for testing of the above regression model is as follows:

$$H_0: \beta_1 = 0$$
$$H_a: \beta_1 \neq 0$$

Table 9 presents data with (weighted) SKKN and DTR, and the data were imported into the SAS system for data analysis. Before the regression model fitting and analysis, and testing the hypothesis, it can be helpful to see the correlations among the

variables, along with p values in the simple regression model. We can get the results using **proc corr** SAS command as shown in Fig. 6.

**TABLE 9**

**SKKN and DTR Data Imported into the SAS System**

| | SKKN | DTR |
|---|---|---|
| 1 | 3.2 | 1 |
| 2 | 3.2 | 0.067 |
| 3 | 2.2 | 0.067 |
| 4 | 1.6 | 0.05 |
| 5 | 2.5 | 1 |
| 6 | 2.5 | 1 |
| 7 | 1.6 | 0.05 |
| 8 | 1.3 | 0.2 |
| 9 | 1 | 0.067 |
| 10 | 1 | 0.05 |
| 11 | 1 | 0.2 |
| 12 | 1.6 | 0.2 |
| 13 | 3 | 0.05 |
| 14 | 1.6 | 0.1 |
| 15 | 1.3 | 0.033 |
| 16 | 1.3 | 0.033 |
| 17 | 6 | 1 |
| 18 | 6 | 1 |
| 19 | 1.7 | 0.033 |
| 20 | 1 | 0.017 |
| 21 | 1.6 | 0.033 |
| 22 | 1.6 | 0.033 |

```
proc corr data="C:\Program Files\SAS\SAS 9.1\My SaS Files\overalldata";
   var DTR SKKN;
run;
```

The CORR Procedure

2 Variables: DTR SKKN

Simple Statistics

| Variable | N | Mean | Std Dev | Sum | Minimum | Maximum | Label |
|---|---|---|---|---|---|---|---|
| DTR | 22 | 0.28559 | 0.40029 | 6.28300 | 0.01700 | 1.00000 | DTR |
| SKKN | 22 | 2.17273 | 1.41898 | 47.80000 | 1.00000 | 6.00000 | SKKN |

Pearson Correlation Coefficients, N = 22

Prob > |r| under H0: Rho=0

| | DTR | SKKN |
|---|---|---|
| DTR | 1.00000 | 0.71613 |
| DTR | | 0.0002 |
| SKKN | 0.71613 | 1.00000 |
| SKKN | 0.0002 | |

**Fig. 6. SAS command PROC CORR and the results**

To perform a regression analysis, we should select appropriate independent variables for the model of the response variable. If we look to the table above, it displays a correlation of 0.71613 between DTR and SKKN, which is significant with a p-value of 0.0002. That is, there exists a positive linear relationship between these two variables. As SKKN increases, DTR increases. Knowing that the variable (SKKN) is strongly associated with DTR, we predict that the variable (SKKN) would be a statistically significant predictor in the simple regression model. We expect that a better detection rate (DTR) performance would be associated with higher level of skill & knowledge (SKKN). In the following sections, we examine the output from the regression analysis.

*5.4.1.1. Testing*

Hypothesis testing for determining whether the linear model is useful for predicting Y (DTR) from X (SKKN), that is, testing usefulness of the model follows. At significance level alpha = 0.05, we can test the hypothesis that the Skill and Knowledge (SKKN) of security personnel contributes useful information for the prediction of the Detection Rate (DTR). In other words, we test the predictive ability of the least squares straight-line model:

$$Y \text{ hat} = \beta_0 \text{ hat} + \beta_1 \text{ hat} * X, \text{ where } Y = DTR, X = SKKN$$

Through the SAS Fit analysis, we get the $\beta_0$ of - 0.1533 and $\beta_1$ of 0.2020, as shown in Fig. 7.

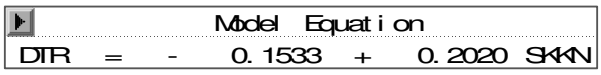


**Fig. 7. Regression model equation**

Thus, we get the predicted model equation:

$$E(Y) = - 0.1533 + 0.2020 * X, \text{ where } Y = DTR, X = SKKN$$

If we see the graph of the linear regression model using the SAS Fit Analysis, it would be best to understand what the regression model looks like. The graph is shown in Fig. 8.

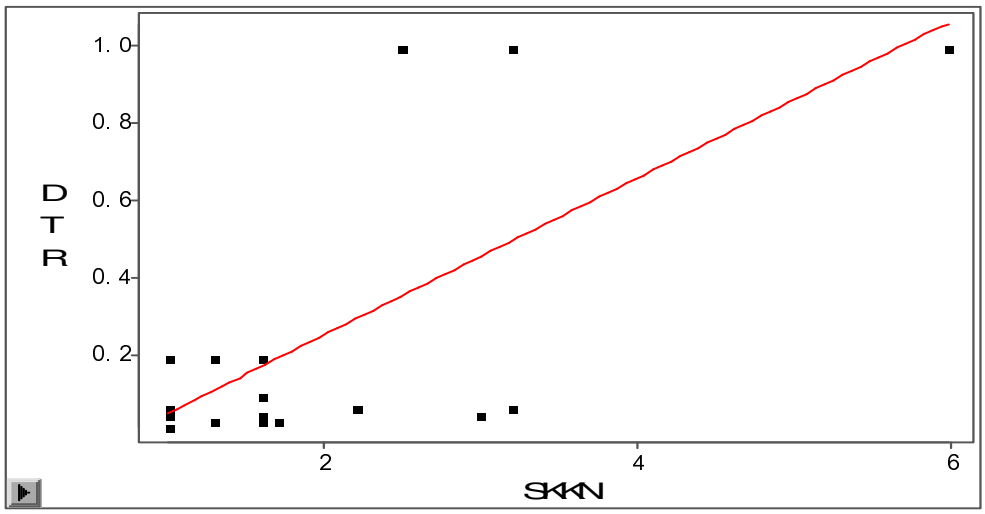


**Fig. 8. Graph of the linear regression model using the SAS Fit Analysis**

Testing the usefulness of the model requires testing the null ($H_0$) and alternative ($H_a$) hypotheses as mentioned above. To test the hypothesis for determining whether the linear model (or straight-line model) is useful for predicting Y (DTR) from X (SKKN), a test statistic T can be used. With N (total number of SKKN data points) of 22 and significance level alpha of 0.05, the critical value based on (N - 2) = (22 - 2) = 20 d.f. is as follows:

$$T_{\text{alpha/2}} = t_{0.025} = 2.086$$

Thus, we will reject $H_0$ if t < - 2. 086 or t > 2. 086. Our test statistic $t = \beta_1 \text{ hat} / (s/\sqrt{SSxx})$, where $\beta_1$ hat = SSxy/SSxx, $SSxy = \sum xy - (\sum x)(\sum y)/n$, $SSxx = \sum x^2 - (\sum x)^2/n$. The calculation summary is shown in Table 10.

$\beta_1$hat = 0.2020, $s^2$ = SSE/ (N – 2) = 1.6392 (from the Table 11) / 20 = 0.08196,

s = $\sqrt{0.08196}$ = 0.2863

$\sqrt{SSxx} = \sqrt{(\sum x^2 - (\sum x)^2/n)} = \sqrt{(146.14 - 2284.84/22)} = \sqrt{42.28} = 6.50$

$t = \beta_1 \text{ hat} / (s/\sqrt{SSxx}) = 0.2020 / (0.2863 / 6.50) = 0.2020 / 0.0440 = 4.591$

**TABLE 10**

**Summary of the Calculation of SKKN, $x^2$, DTR, $y^2$, and xy**

| SKKN (x) | $x^2$ | DTR (y) | $y^2$ | xy |
|---|---|---|---|---|
| 3.2 | 10.24 | 1 | 1 | 3.2 |
| 3.2 | 10.24 | 0.067 | 0.004489 | 0.2144 |
| 2.2 | 4.84 | 0.067 | 0.004489 | 0.1474 |
| 1.6 | 2.56 | 0.05 | 0.0025 | 0.08 |
| 2.5 | 6.25 | 1 | 1 | 2.5 |
| 2.5 | 6.25 | 1 | 1 | 2.5 |
| 1.6 | 2.56 | 0.05 | 0.0025 | 0.08 |
| 1.3 | 1.69 | 0.2 | 0.04 | 0.26 |
| 1 | 1 | 0.067 | 0.004489 | 0.067 |
| 1 | 1 | 0.05 | 0.0025 | 0.05 |
| 1 | 1 | 0.2 | 0.04 | 0.2 |
| 1.6 | 2.56 | 0.2 | 0.04 | 0.32 |
| 3 | 9 | 0.05 | 0.0025 | 0.15 |
| 1.6 | 2.56 | 0.1 | 0.01 | 0.16 |
| 1.3 | 1.69 | 0.033 | 0.001089 | 0.0429 |
| 1.3 | 1.69 | 0.033 | 0.001089 | 0.0429 |
| 6 | 36 | 1 | 1 | 6 |
| 6 | 36 | 1 | 1 | 6 |
| 1.7 | 2.89 | 0.033 | | 0.0561 |
| 1 | 1 | 0.017 | 0.000289 | 0.017 |
| 1.6 | 2.56 | 0.033 | 0.001089 | 0.0528 |
| 1.6 | 2.56 | 0.033 | 0.001089 | 0.0528 |
| $\sum x = 47.8$ | $\sum x^2 = 146.14$ | $\sum y = 6.283$ | $\sum y^2 = 5.158112$ | $\sum xy = 22.1933$ |

**TABLE 11**

**ANOVA for DTR with SKKN as the Predictor Variable**

Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Stat | Pr > F |
|---|---|---|---|---|---|
| Model | 1 | 1.7256 | 1.7256 | 21.05 | 0.0002 |
| Error | 20 | 1.6392 | 0.0820 | | |
| C Total | 21 | 3.3648 | | | |

Since t = 4.591 > 2. 086, we should reject the null hypothesis and conclude that the slope $\beta_1$ is not 0. Thus, at the alpha = 0.05 level of significance, the sample data we observed provide sufficient evidence to conclude that the Skill and Knowledge (SKKN) of security personnel contributes useful information for the prediction of the Detection Rate (DTR) using the linear model.

Testing how well the least squares line fit the data, that is, testing fitting level of the model follows. After testing the usefulness of the linear regression model, we should test how well the regression model fits the data we collected. A measure, called the coefficient of determination ($R^2$), can answer this question. It can be computed using a statistics package such as SAS. $R^2$ is the proportion of variance in the dependent variable (DTR) that can be predicted from the independent variable (SKKN). The measure is useful for assessing how many of the errors in the prediction of y (DTR) can be reduced by using the information provided by x (SKKN) [76]. $R^2$ is:

$R^2$ = (SSyy – SSE) / SSyy = 1 – SSE/SSyy,

where SSE is Residual SS and SSyy is Total SS (in Excel-like data analysis)

$SSyy = \sum y^2 - (\sum y)^2/n = 5.158 - 39.476/22 = 3.364$

$SSxy = \sum xy - (\sum x)(\sum y)/n = 22.193 - 47.8*6.283/22 = 8.542$

$SSE = \sum (y - y\ hat)^2 = SSyy - \beta_1\ hat * SSxy = 3.364 - 0.2020*8.542 = 1.639$

**TABLE 12**

**Summary of Fit for DTR with SKKN as the Predictor Variable**

| Summary of Fit | | | |
|---|---|---|---|
| Mean of Response | 0.2856 | R-Square | 0.5128 |
| Root MSE | 0.2863 | Adj R-Sq | 0.4885 |

Thus, $R^2$ = 1 – SSE/SSyy = 1 – 1.639/3.364 = 0.513. This is the value (0.5128; rounding error in the calculation) we find from Table 12. This value indicates that 51% of the variance in DTR can be predicted from the variable SKKN. In other words, 51% of the total variation, SSyy (the sum of the squared prediction errors) = $\sum$ (Actual Y – Predicted Y)$^2$ = $\sum$ (Y – Y_bar)$^2$, is explained by the model, and the remaining portion is explained by random error. $R^2$ = 0 and 1 implies a complete lack of fit of the model to the data and a perfect fit, respectively. Thus, typically, the larger the value of $R^2$, the better the model fits the set of data.

Sincich [76] warns that, however, we can use the value of $R^2$ as a measure of how useful a linear model will be for predicting Y only if the sample contains substantially more data points than the number of β parameters in the model. Thus, the more data that we can obtain from experiments, the more confident we can be.

In the 'Summary of Fit' section (see Table 12), we find another important value – Adjusted $R^2$. Adjusted $R^2$ attempts to yield a more accurate value to estimate the $R^2$ value for the population. Adjusted $R^2$ can be computed using the following formula:

$1 - ((1 - R^2)((N - 1) / (N - k - 1)))$,

where N is the number of observations and k is the number of predictors (dependent variables)

From this formula, we can see that when N is small and k is large, greater differences between $R^2$ and adjusted $R^2$ may exist since the ratio of ((N - 1) / (N - k - 1)) will be less than 1. Conversely, when N is large compared to k, $R^2$ and the adjusted $R^2$

will be much closer since the ratio of (N - 1) / (N - k - 1) will begin to approach 1. The summary of fit presents that the value of $R^2$ and adjusted $R^2$ are 0.5128 and 0.4885, respectively. The difference between the two values is 0.0243, which can be much closer to 0 when we have more data.

Testing significance of the model and coefficient follows. To see whether it is statistically significant with the predicted variable, we should look to the p-value (attained significance level) of the F-test. F value is 21.05, and the p value is shown to the right hand side of the *F*-value in the figure, i.e. 0.0002. The p value associated with this F value is small (0.0002). The p value and significance level of 0.05 are compared with each other, and the p value is smaller. Since the p-value is smaller, the model is statistically significant, and thus we may conclude that the independent variable (SKKN) reliably predicts the dependent variable (DTR). The positive coefficient (0.2020, shown below) of SKKN and its significance (p=0.0002) indicates that the higher the level of Skill and Knowledge (SKKN) that security personnel possess, the better (higher) the Detection Rate (DTR). Thus, this SAS analysis result makes sense.

The significance of the coefficient for SKKN can be assessed using the p-value of the T-test. The T-test for SKKN equals 4.59, presented in Table 13. By comparing the p value (0.0002) and alpha level (e.g. 0.05) selected, we know that p value is smaller. Thus, we can reject the null hypothesis that the coefficient for SKKN is 0. In other words, the regression coefficient of 0.2020 is significantly different from 0. Even at a lower alpha level (0.01), the coefficient for SKKN would still be significant.

**TABLE 13**

**Parameter Estimates for DTR with SKKN as the Predictor Variable**

Parameter Estimates

| Variable | DF | Estimate | Std Error | t Stat | Pr >\|t\| | Tolerance | Var Inflation |
|---|---|---|---|---|---|---|---|
| Intercept | 1 | -0.1533 | 0.1135 | -1.35 | 0.1917 | . | 0 |
| SKKN | 1 | 0.2020 | 0.0440 | 4.59 | 0.0002 | 1.0000 | 1.0000 |

*5.4.1.2. Verification of Assumptions*

We also have to check the validity of assumptions. First, we check the assumption of normal distribution. It is the residuals that need to be normally distributed to establish the validity of the t-test; note that the estimation of the regression coefficients does not require normally distributed residuals. We can exploit some graphical methods to illustrate the data: residual-by-predicted plot and normal quantile-quantile plot.

A residual-by-predicted plot is commonly used to diagnose nonlinearity or unequal variances of error. It is also used to find outliers. A residual-by-predicted plot is illustrated by the plot on the left in Fig. 9. It is a plot of residuals versus predicted responses for each observation.

A normal quantile-quantile plot of residuals is illustrated by the plot on the right in Fig. 9. The empirical quantiles are plotted against the quantiles of a standard normal distribution. If the residuals are normally distributed, the points on the residual normal quantile- quantile plot should lie approximately on a straight line with residual mean as the intercept and residual standard deviation as the slope. Even if we can see that the residuals are not perfectly normally distributed, it is hard to draw a firm conclusion since the data points are not that far away from the straight line.

As another method, we use a normal probability plot, which is often used to examine the distribution of variables. The SAS command proc capability with the ppplot statement is used for the normal probability plot. The command and its results are shown in Fig. 10. It does not seem that the plot of SKKN looks 100% normal. From the various plots we examined, we see that the variable (SKKN) does not look quite normal, but is still within a somewhat normal range.

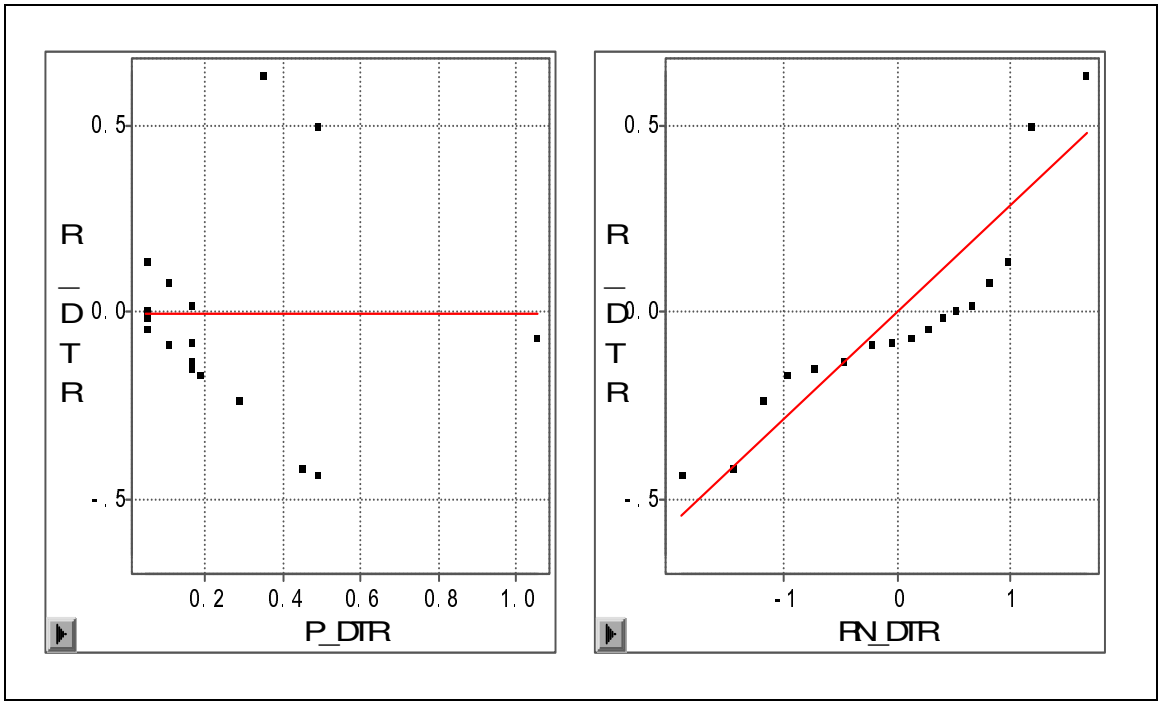


**Fig. 9. Residual-by-Predicted and Residual Normal QQ Plots**

### 5.4.2. Hypothesis 2

*Hypothesis 2: Training (TRNG) and Skill and Knowledge (SKKN) of security personnel are statistically significantly related to the Response Rate (RPR)*

The correlation between RPR and TRNG is high; however, the correlation between RPR and SKKN is moderate. We know that, from the result of the SAS command proc corr (shown in Appendix 1), the model of RPR with SKKN as the predictor variable is significant since the p-value of 0.0451 is less than the significance level of 0.05. Thus, it may be possible to build a multiple regression model of RPR with TRNG and SKKN as its predictor variables. We expect that a higher Response Rate (RPR) would be associated with higher level of both Training (TRNG) and Skill and Knowledge (SKKN).

```
proc capability data="C:\Program Files\SAS\SAS 9.1\My SAS Files\overalldata" noprint;
  ppplot SKKN;
run;
```

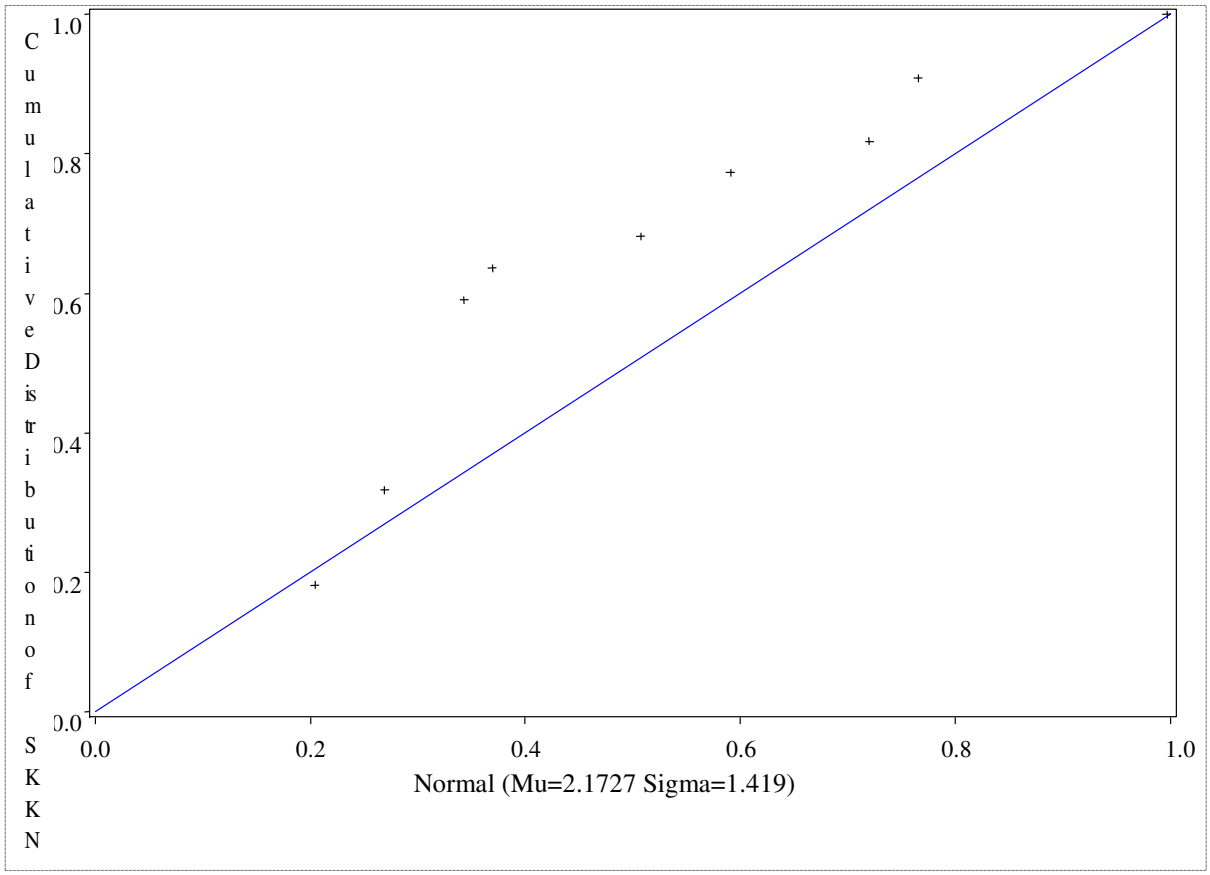


**Fig. 10. Normal probability plot with the SAS command PROC CAPABILITY with ppplot statement**

Based on the correlation analysis results, the multiple linear regression model for Hypothesis 2 can be set up as follows:

$$RPR = \beta_0 + \beta_1 * TRNG + \beta_2 * SKKN + \text{random error}$$

Table 14 presents data with (weighted) SKKN and TRNG, and RPR, and the data were imported into the SAS system for data analysis.

**TABLE 14**

**Data with (weighted) SKKN and TRNG, and RPR**

| | SKKN | TRNG | RPR |
|---|---|---|---|
| 1 | 3.2 | 2.7 | |
| 2 | 3.2 | 2.7 | 0.067 |
| 3 | 2.2 | 2.4 | 0.067 |
| 4 | 1.6 | 2.7 | 0.018 |
| 5 | 2.5 | 2.4 | 0.2 |
| 6 | 2.5 | 2.4 | 0.2 |
| 7 | 1.6 | 3.9 | 1 |
| 8 | 1.3 | 2.4 | 0.04 |
| 9 | 1 | 2.4 | 0.067 |
| 10 | 1 | 2.4 | 0.04 |
| 11 | 1 | 2.4 | 0.038 |
| 12 | 1.6 | 2.4 | 0.05 |
| 13 | 3 | 3 | 0.1 |
| 14 | 1.6 | 3 | 0.029 |
| 15 | 1.3 | 2.4 | 1 |
| 16 | 1.3 | 2.4 | 0.05 |
| 17 | 6 | 3.4 | 1 |
| 18 | 6 | 3.4 | 1 |
| 19 | 1.7 | 3.7 | 1 |
| 20 | 1 | 2.4 | |
| 21 | 1.6 | 4.4 | |
| 22 | 1.6 | 3.3 | |

*5.4.2.1.Testing*

Hypothesis testing for determining whether the overall multiple regression model is useful for predicting Y(RPR) from $X_1$ (TRNG) and $X_2$ (SKKN), that is, testing usefulness of the model follows. The hypotheses for testing whether a general linear model can be useful for predicting Y (RPR) is as follows [27]:

$H_0$: $\beta_1 = \beta_2 = 0$

$H_1$: At least one of the two $\beta$ parameters in $H_0$ is nonzero.

With the hypothesis, *F* statistic is used as a test statistic for model usefulness. The numerator degrees of freedom is k, which is the number of parameters in the model (excluding $\beta_0$), and denominator degrees of freedom is n - (k + 1), where n is the number of observations. Both values determine the value of *F*. Since n = 18 data points and k = 2, the rejection region for the test is:

$$F > F_{\alpha} = 3.68, \text{ where } \alpha = 0.05$$

The value of *F* has been computed and presented in Table 15 in the row corresponding to the model (in ANOVA section). Since the F value of 9.48 exceeds the critical value, $F_{0.05} = 3.68$, we may reject the null hypothesis and conclude that at least one of the two parameters ($\beta_1$ and $\beta_2$) is nonzero. In other words, the model appears to be useful for predicting Y, RPR.

Measuring how well the model fits the data follows. To measure how well the model fits the data collected, $R^2$ can be computed. $R^2$ is the proportion of variance in the dependent variable (RPR) that can be predicted from the independent variables (TRNG and SKKN). In 'Summary of Fit' section (see Table 15), we see that the $R^2$ is 0.5583, meaning that approximately 56% of the variability of RPR is accounted for by the two variables – TRNG and SKKN – in the regression model.

Note that the adjusted R-square is 0.4994, which indicates that approximately 50% of the variability of RPR is accounted for by the regression model, even after taking into account the number of predictor variables in the model. Note that difference between the value of $R^2$ and that of adjusted $R^2$ (0.5583 and 0.4994, respectively) is 0.0589, which can be much closer to 0 when more data is available.

Testing significance of the overall model follows. To see if the overall model is significant, we should look to the p-value (attained significance level) of the F-test. The p-value for the test is also shown to the right hand side of the *F*-value in the figure, i.e. <.0022. This means that if the model did not contribute any information for the Y

prediction, the probability of observing the *F* statistic of 9.48 would be only less than 0.0022. Because the p-value is very small (0.0022), the model is statistically significant.

Testing significance of the two predictor variables follows. We focus on whether the two predictors are statistically significant with the predicted variable, and if so, the direction of the relationship. Training (TRNG) is significant because p = 0.0044 and its coefficient is 0.5440. The coefficients for each of the predictor variables indicates the amount of change one could expect in RPR with a one-unit change in the value of each variable, given that all other variables in the model are held constant. The positive coefficient for TRNG indicates that the higher the training level the security personnel possess, the better (higher) the response rate. The SAS analysis results support this conclusion.

However, the coefficient for SKKN is not significantly different from 0 with the alpha level of 0.05 because its p-value of 0.3578 is greater than 0.05. Thus, the level of Skill & Knowledge (SKKN) seems to be unrelated to the Response Rate (RPR). This would seem to indicate that that the level of Skill & Knowledge is not an important factor in predicting the Response Rate. This result is somewhat unexpected and confusing.

*5.4.2.2. In-depth Analysis*

We should do several checks to make sure we firmly stand behind these results before drawing conclusions. First, because we are interested in residuals, we perform residual analysis against predictor variables. The results are displayed in Fig. 11. By examining residual plots, we can see one data point – located above y-value of 0.8 – is far from 0. Thus, we can suppose that the data point could be an outlier or influential value. To examine outliers or influential values, we use different statistics to catch those extreme values.

**TABLE 15**

**SAS Fit Analysis for RPR with TRNG and SKKN as Its Predictors**

RPR = TRNG SKKN

Response Distribution: Normal

Link Function: Identity

Model Equation

RPR = - 1.2900 + 0.5440 TRNG + 0.0518 SKKN

Summary of Fit

| | | | |
|---|---|---|---|
| Mean of Response | 0.3314 | R-Square | 0.5583 |
| Root MSE | 0.3039 | Adj R-Sq | 0.4994 |

Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Stat | Pr > F |
|---|---|---|---|---|---|
| Model | 2 | 1.7513 | 0.8756 | 9.48 | 0.0022 |
| Error | 15 | 1.3856 | 0.0924 | | |
| C Total | 17 | 3.1369 | | | |

Parameter Estimates

| Variable | DF | Estimate | Std Error | t Stat | Pr >\|t\| | Tolerance | Var Inflation |
|---|---|---|---|---|---|---|---|
| Intercept | 1 | -1.2900 | 0.4146 | -3.11 | 0.0071 | | 0 |
| TRNG | 1 | 0.5440 | 0.1624 | 3.35 | 0.0044 | 0.7919 | 1.2627 |
| SKKN | 1 | 0.0518 | 0.0546 | 0.95 | 0.3578 | 0.7919 | 1.2627 |

Collinearity Diagnostics

| | | | Variance Proportion | | |
|---|---|---|---|---|---|
| Number | Eigenvalue | Condition Index | Intercept | TRNG | SKKN |
| 1 | 2.7930 | 1.0000 | 0.0036 | 0.0031 | 0.0266 |
| 2 | 0.1933 | 3.8015 | 0.0333 | 0.0131 | 0.8415 |
| 3 | 0.0137 | 14.2572 | 0.9630 | 0.9839 | 0.1319 |

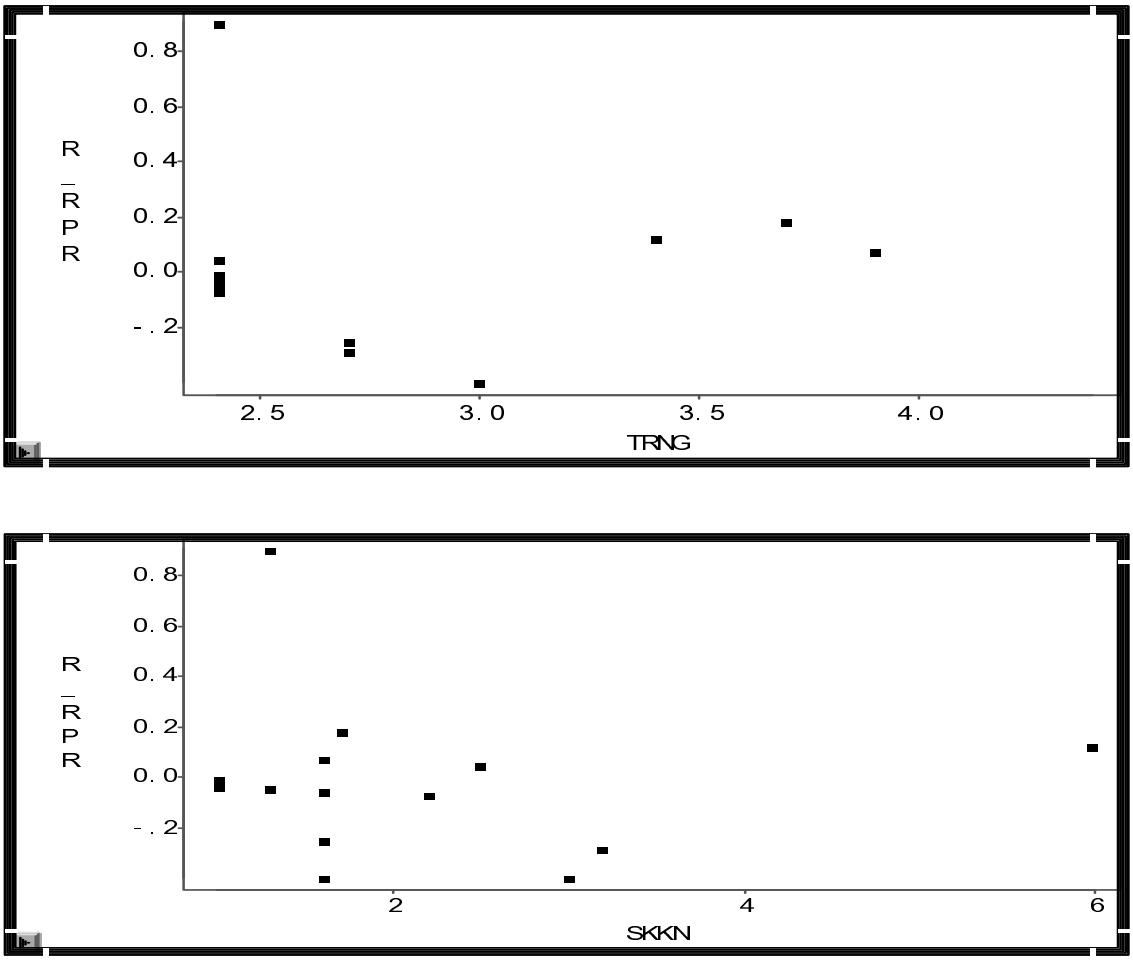


**Fig. 11. Residual plots of RPR against each of the predictor variables**

It can be helpful to look at the scatter plots of RPR against each of the predictor variables so that we may have some ideas about potential problems residing in the regression model. We can create a scatter plot matrix of these variables as shown in Fig. 12.

In each plot, we see some data points that are far removed from the rest of the data points (i.e. possible outliers). Even though the SAS commands above provide us with useful information regarding the variables, we need to exploit other statistics to identify all the potentially unusual or influential data values. These statistics include

studentized residual (named st_r), leverage (named lev), Cook's D (named ckd), and DFFITS (named dft).

```
proc insight data="C:\Program Files\SAS\SAS 9.1\My SAS Files\overalldata";
  scatter RPR TRNG SKKN*
        RPR TRNG SKKN;
run;
```

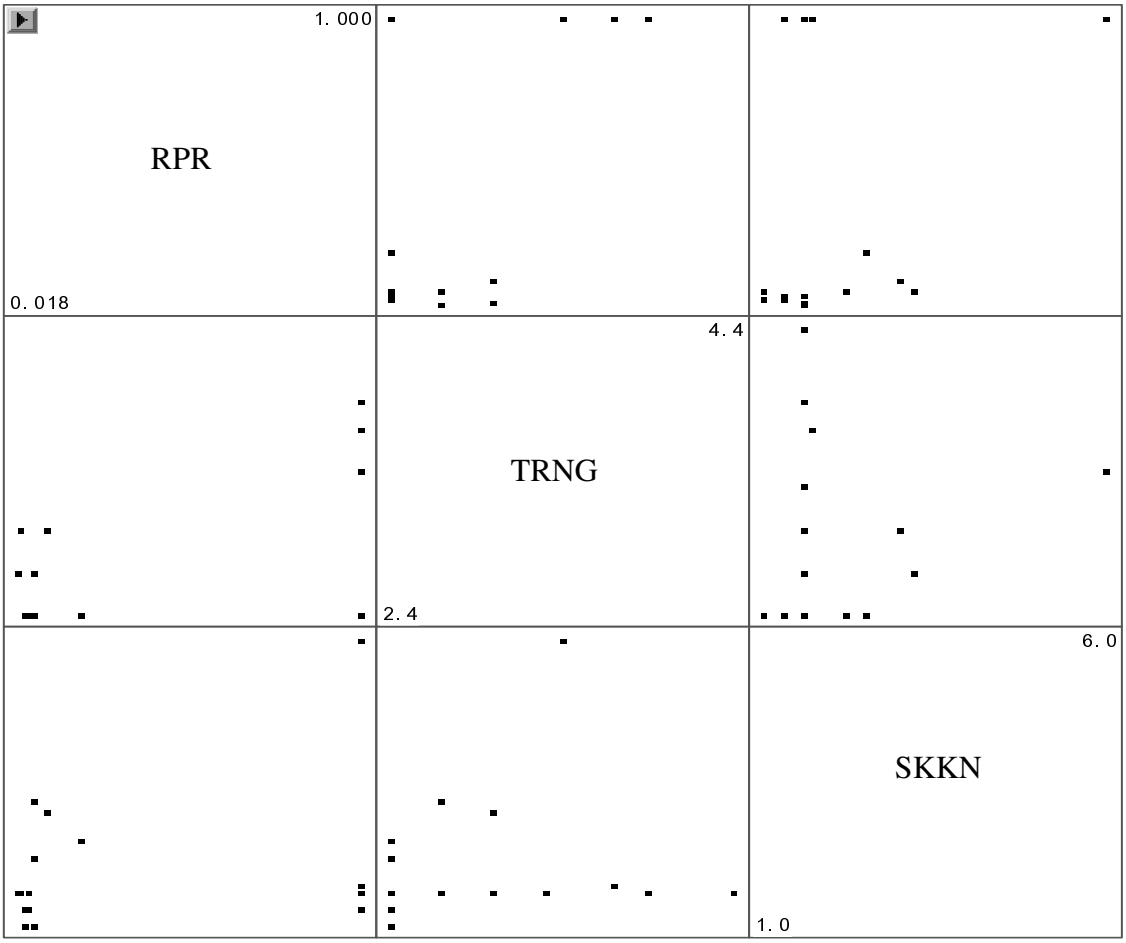


**Fig. 12. SAS PROC INSIGHT command with scatter plots of RPR against each of the predictor variables**

Before using studentized residual, we can use proc reg, of which the results are similar to Table 15 (thus, we will not show the results here). The command is shown in Fig. 13. Then, we begin a more in-depth investigation with the four statistics mentioned above.

First, we examine the studentized residuals to identify potentially unusual or influential data values such as outliers. We request the studentized residuals in the output statement and name them st_r. The command and its (partial) results are shown in Fig. 13.

From looking at the extreme observations and a stem-and-leaf display, we find a high value of 5.315190; this seems to be excessive. That value is also found in the output of all the studentized residuals and leverage against each observation of SKKN, TRNG, and RPR. The commands and its output are shown in Fig. 14.

Usually, we should be concerned about the observation where the absolute value of studentized residuals exceeds 2, and we should be even more concerned about the one where the absolute value of the residuals exceeds 3. However, in this case, the 22$^{nd}$ observation, with the studentized residual of 5.31519, is in excess of the absolute value of 5. We conjecture that even though both SKKN and TRNG have a low number each – 1.3 and 2.4, respectively – the observation has the highest RPR value.

Leverage (named lev in our analysis) is the second statistic we employ to check all the potential influence on the regression coefficient estimates. We can get the results of leverage using the SAS command proc univariate. The command and its (partial) results are displayed in Fig. 15.

```
proc reg data="C:\Program Files\SAS\SAS 9.1\My SAS Files\overalldata";
  model RPR=TRNG SKKN;
  output out=overalldata(keep= RPR TRNG SKKN st_r lev cd dffit)
                rstudent=st_r h=lev cookd=ckd dffits=dft;
run;
```

```
proc univariate data=overalldata plots plotsize=20;
  var st_r;
run;
```

Extreme Observations

| ------Lowest------ | | ------Highest----- | |
|---|---|---|---|
| Value | Obs | Value | Obs |
| -1.416973 | 14 | 0.385777 | 7 |
| -1.402450 | 13 | 0.544461 | 17 |
| -0.956025 | 2 | 0.544461 | 18 |
| -0.821109 | 4 | 0.765629 | 19 |
| -0.209424 | 3 | **5.315190** | **15** |

```
Stem Leaf              #          Boxplot
  5 3                  1          *
  4
  4
  3
  3
  2
  2
  1
  1
  0 558                3          |
  0 224                3          +--+--+
 -0 2211110            7             *-----*
 -0 8                  1          |
 -1 440                3          0
   ----+----+----+----+
```

**Fig. 13. SAS PROC commands and results of the studentized residuals**

```
proc sort data=overalldata;
  by st_r;
run;

proc print data=overalldata(obs=22);
run;
```

| Obs | SKKN | TRNG | RPR | lev | st_r |
|---|---|---|---|---|---|
| 1 | 3.2 | 2.7 | . | 0.09189 | . |
| 2 | 1.0 | 2.4 | . | 0.10399 | . |
| 3 | 1.6 | 4.4 | . | 0.92333 | . |
| 4 | 1.6 | 3.3 | . | 0.18034 | . |
| 5 | 1.6 | 3.0 | 0.029 | 0.09770 | -1.41697 |
| 6 | 3.0 | 3.0 | 0.100 | 0.07410 | -1.40245 |
| 7 | 3.2 | 2.7 | 0.067 | 0.09189 | -0.95603 |
| 8 | 1.6 | 2.7 | 0.018 | 0.06647 | -0.82111 |
| 9 | 2.2 | 2.4 | 0.067 | 0.09260 | -0.20942 |
| 10 | 1.6 | 2.4 | 0.050 | 0.08667 | -0.16170 |
| 11 | 1.3 | 2.4 | 0.040 | 0.09242 | -0.14367 |
| 12 | 1.3 | 2.4 | 0.050 | 0.09242 | -0.11025 |
| 13 | 1.0 | 2.4 | 0.038 | 0.10399 | -0.09904 |
| 14 | 1.0 | 2.4 | 0.040 | 0.10399 | -0.09232 |
| 15 | 1.0 | 2.4 | 0.067 | 0.10399 | -0.00162 |
| 16 | 2.5 | 2.4 | 0.200 | 0.10428 | 0.18426 |
| 17 | 2.5 | 2.4 | 0.200 | 0.10428 | 0.18426 |
| 18 | 1.6 | 3.9 | 1.000 | 0.49990 | 0.38578 |
| 19 | 6.0 | 3.4 | 1.000 | 0.41718 | 0.54446 |
| 20 | 6.0 | 3.4 | 1.000 | 0.41718 | 0.54446 |
| 21 | 1.7 | 3.7 | 1.000 | 0.35851 | 0.76563 |
| 22 | 1.3 | 2.4 | 1.000 | 0.09242 | 5.31519 |

**Fig. 14. SAS PROC commands and results of the studentized residuals and leverage against each observation**

```
proc univariate data=overalldata plots plotsize=20;
  var lev;
run;
```

Extreme Observations

| -------Lowest------ | | ------Highest----- | |
|---|---|---|---|
| Value | Obs | Value | Obs |
| 0.0664715 | 8 | 0.358510 | 21 |
| 0.0740961 | 6 | 0.417178 | 19 |
| 0.0866705 | 10 | 0.417178 | 20 |
| 0.0918910 | 7 | 0.499901 | 18 |
| 0.0918910 | 1 | 0.923325 | 3 |

```
Stem Leaf              #        Boxplot
   9 2                 1           *
   8
   7
   6
   5 0                 1           *
   4 22                2           0
   3 6                 1           0
   2
   1 00000008          8           +-----+
   0 779999999         9           +--+--+
     ----+----+----+----+
Multiply Stem.Leaf by 10**-1
```

**Fig. 15. SAS PROC UNIVARIATE command and results of the leverage**

Usually, we should carefully examine a single data point with leverage greater than (2K + 2) / N, where K is the number of predictor variables and N is the number of observations. In our case, K is 2 and N is 22 (missing values are included), working out to (2 * 2 + 2) / 22 = 0.272727. To check out the data points with the value greater than 0.272727, we use the SAS command proc print with the where statement, shown in Fig. 16, followed by its results.

```
proc print data=overalldata;
  var RPR TRNG SKKN;
  where lev > 0.272727;
run;
```

| Obs | RPR | TRNG | SKKN |
|---|---|---|---|
| 3 | . | 4.4 | 1.6 |
| 18 | 1 | 3.9 | 1.6 |
| 19 | 1 | 3.4 | 6.0 |
| 20 | 1 | 3.4 | 6.0 |
| 21 | 1 | 3.7 | 1.7 |

**Fig. 16. SAS PROC PRINT command and results of the data points with leverage greater than (2K + 2) / N**

Except for the missing value for RPR, we find four data points greater than 0.272727. The data point (studentized residual of 5.31519) with TRNG of 2.4 and SKKN of 1.3 was highlighted in the studentized residuals tests, but not here in the leverage tests. We have four other data points.

Third, we use Cook's D (named ckd in our analysis) to identify all the potentially unusual or influential data values. It measures the information regarding both residuals and leverage. Zero is assumed to be the lowest value in Cook's D. If the value of Cook's D is higher, the data point is assumed to be more influential. Usually, a threshold to decide Cook's D is 4 / N, where N is the number of observations. The command for Cook's D and its result are shown in Fig. 17. We see that the Cook's D for observation #22 is the highest (0.34046).

The last statistic we use is DFFITS (named dft in our analysis). A conventional threshold for DFFITS is $2 * \sqrt{(K / N)}$, where K and N are the number of predictor variables and the number of observations, respectively. The bigger the absolute value corresponding to the data point, the greater the influence of the point might be. The observation #22 is also the most influential (1.69617) observation. The command for DFFITS and its results are shown in Fig. 18.

```
proc print data=overalldata;
  where ckd > (4/22);
  var RPR TRNG SKKN ckd;
run;
```

| Obs | RPR | TRNG | SKKN | ckd |
|---|---|---|---|---|
| 22 | 1 | 2.4 | 1.3 | 0.34046 |

**Fig. 17. SAS PROC PRINT command and results of the data points with Cook's D greater than 4 / N**

```
proc print data = overalldata;
  where abs(dft) > (2 * √ (2/22));
  var RPR TRNG SKKN dft;
run;
```

| Obs | RPR | TRNG | SKKN | dft |
|---|---|---|---|---|
| 22 | 1 | 2.4 | 1.3 | 1.69617 |

**Fig. 18. SAS PROC PRINT command and results of the data points with absolute DFFITS value greater than 2 * √ (K / N)**

Through the four statistical checks, we might claim that the data point (RPR=1, TRNG=2.4, SKKN=1.3) with studentized residual of 5.31519, Cook's D of 0.34046, and DFFITS of 1.69617 is most influential. However, to be even more certain we can consider another statistical check called DFBETAS, which is a scaled measure of the change in each parameter estimate. This assesses how each coefficient can be changed by deleting the corresponding observation; large values in the DFBETAS output indicate influential observations in estimating given parameters. Usually, observations with a value bigger than the absolute value of (2 / √ N), where N is the number of observations, should cause concern. We can use ods output OutStatistics statement with proc reg

```
proc reg data="C:\Program Files\SAS\SAS 9.1\My SAS Files\overalldata";
  model RPR = TRNG SKKN / influence;
  ods output OutputStatistics=RPRdfbetas;
  id RPR TRNG SKKN;
run;
```

Output Statistics

| | | | | DFBETAS | | |
|---|---|---|---|---|---|---|
| Obs | RPR | TRNG | SKKN | Intercept | TRNG | SKKN |
| 1 | . | 2.7 | 3.2 | . | . | . |
| 2 | 0.067 | 2.7 | 3.2 | -0.1090 | 0.1143 | -0.1886 |
| 3 | 0.067 | 2.4 | 2.2 | -0.0495 | 0.0423 | -0.0179 |
| 4 | 0.018 | 2.7 | 1.6 | -0.0438 | -0.0146 | 0.0846 |
| 5 | 0.2 | 2.4 | 2.5 | 0.0459 | -0.0422 | 0.0263 |
| 6 | 0.2 | 2.4 | 2.5 | 0.0459 | -0.0422 | 0.0263 |
| 7 | 1 | 3.9 | 1.6 | -0.3040 | 0.3593 | -0.2139 |
| 8 | 0.04 | 2.4 | 1.3 | -0.0291 | 0.0179 | 0.0121 |
| 9 | 0.067 | 2.4 | 1 | -0.0003 | 0.0002 | 0.0002 |
| 10 | 0.04 | 2.4 | 1 | -0.0178 | 0.0092 | 0.0131 |
| 11 | 0.038 | 2.4 | 1 | -0.0191 | 0.0098 | 0.0140 |
| 12 | 0.05 | 2.4 | 1.6 | -0.0345 | 0.0242 | 0.0045 |
| 13 | 0.1 | 3 | 3 | 0.0738 | -0.0915 | -0.1150 |
| 14 | 0.029 | 3 | 1.6 | 0.1502 | -0.2648 | 0.2576 |
| **15** | **1** | **2.4** | **1.3** | **1.0768** | **-0.6615** | **-0.4481** |
| 16 | 0.05 | 2.4 | 1.3 | -0.0223 | 0.0137 | 0.0093 |
| 17 | 1 | 3.4 | 6 | -0.1044 | 0.0219 | 0.3712 |
| 18 | 1 | 3.4 | 6 | -0.1044 | 0.0219 | 0.3712 |
| **19** | **1** | **3.7** | **1.7** | **-0.4323** | **0.5195** | **-0.3110** |
| 20 | . | 2.4 | 1 | . | . | . |
| 21 | . | 4.4 | 1.6 | . | . | . |
| 22 | . | 3.3 | 1.6 | . | . | . |

**Fig. 19. SAS PROC REG command and results of DFBETAS**

command for DFBETAS outputs. The command and its (partial) results are shown in Fig. 19.

We see that the values bigger than the absolute value of (2 / √22) are observations #15 (-0.6615 for TRNG) and #19 (0.5195 for TRNG). The DFBETAS value for TRNG is -0.6615, which means that SKKN increases the coefficient for TRNG by 0.6615 standard errors.

By performing this DFBETAS statistic, we can have stronger confidence that the data point (#15 in the DFBETAS output; RPR=1, TRNG=2.4, SKKN=1.3) with studentized residual of 5.31519, Cook's D of 0.34046, and DFFITS of 1.69617 appears to be an outlier. Thus, we may eliminate this data point in our analysis. However,

without justification (e.g., scientific reasons), we cannot discard this data point automatically from our full data.

*5.4.2.3. Verification of Assumptions*

We check the validity of assumptions, similarly to the approach performed with the simple regression model. First, we test for normality of residuals. For that test, we examine a normal quantile graph, displayed in Fig. 20, and QQ plot, displayed in Fig. 21, to plot the quantiles of variables against the quantiles of a normal distribution.

Normal quantile graph seems to indicate that the residual distribution is not perfectly normal. The closer the residual points are to the line defining normality, the more likely the residuals are normally distributed. However, except for one data point located high above the straight line, the distribution of the residuals is somewhat near to the line.

Residual normal quantile-quantile (QQ) plot of residuals is illustrated in Fig. 21. The empirical quantiles are plotted against the quantiles of a standard normal distribution. It seems to be approximately normal because each residual point is not far from the line.

Second, we check the multicollinearity. Several methods are available to detect multicollinearity. One of the simplest methods is to use the correlation data analysis, which examines the correlation coefficient $r$ between each pair of independent variables. The result is shown in Table 16.

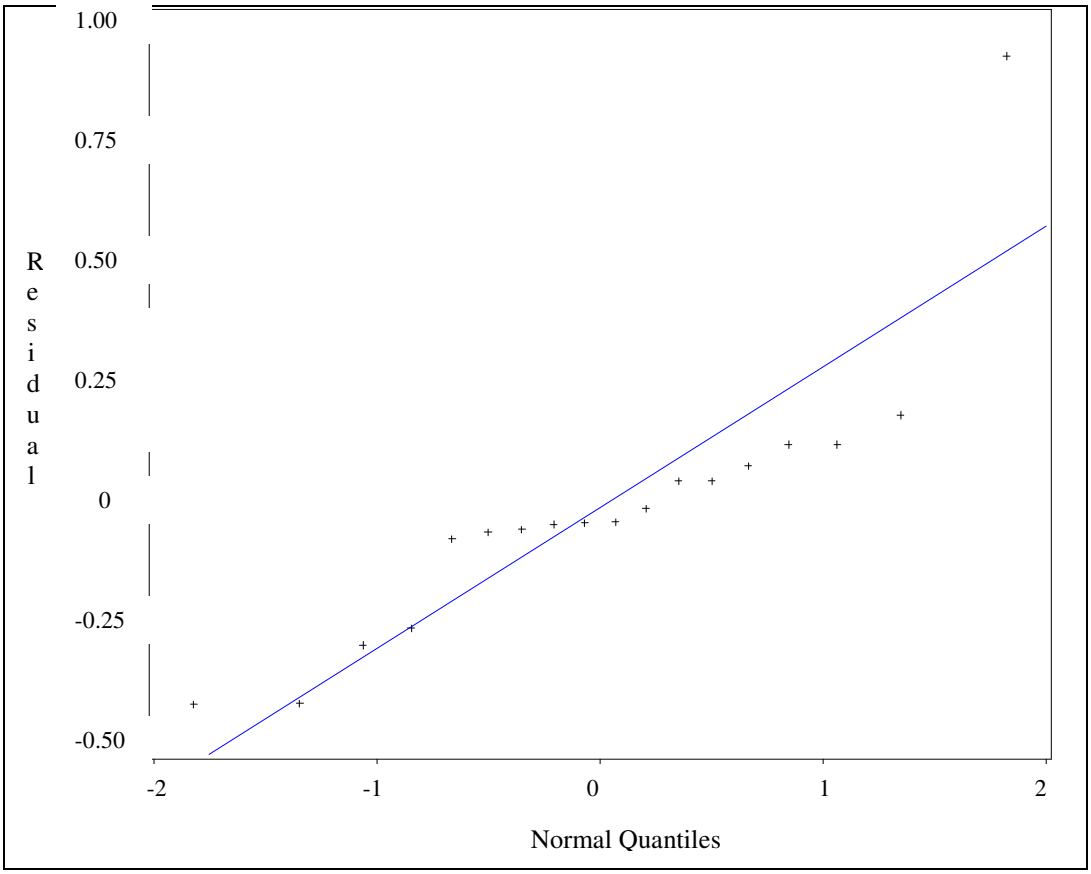


**Fig. 20. Normal quantile graph**

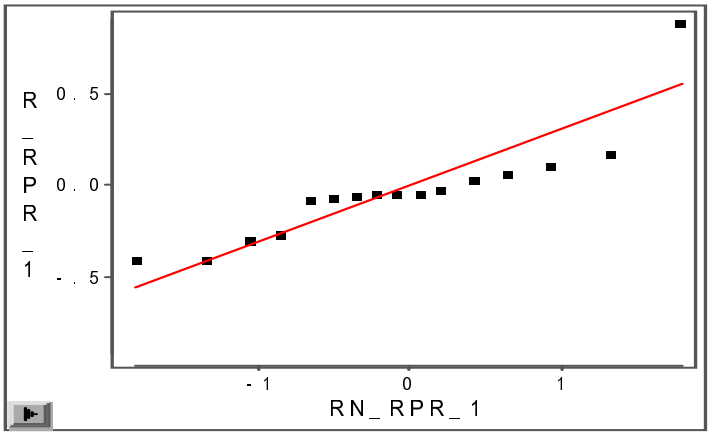


**Fig. 21. Residual normal QQ plot**

**TABLE 16**

**Correlation Data Analysis of Two Independent Variables TRNG and SKKN**

| Correlation Matrix | | |
|---|---|---|
| | TRNG | SKKN |
| TRNG | 1.0000 | 0.2898 |
| SKKN | 0.2898 | 1.0000 |

Anderson [2], in his review of linear regression model, states that there is a rule of thumb, in terms of multicollinearity, that if the correlation coefficient between two variables is greater than 0.80 then there is a problem. It seems that there is no multicollinearity between the two variables since the r value is closer to zero than it is to one. However, Miles and Shevlin [53] warn that low correlations do not indicate that there is not a problem, because it is the multiple correlations that matter, not the bivariate correlations.

Another method is to use the tolerance and variance inflation factor (VIF) [53]. Tolerance for a variable is calculated as $1 - R^2$. The variable being assessed is used as the dependent variable and other variables are used as independent variables in a regression analysis. The tolerance of value 'zero' implies that the variable being assessed is completely predictable from the other independent variables; in other words, it is a perfect multicollinearity. A tolerance of value 'one' implies that the variable being assessed is completely not predictable from the other independent variables; in other words, it is a perfect non-multicollinearity.

The variance inflation factor (VIF), which is closely related to the tolerance, is calculated as 1 / tolerance. It may explain the degree to which the standard error of a variable has increased due to multicollinearity. It is suggested that a VIF in excess of 10 is an indication that multicollinearity may be causing problems in estimation [10], [60], and the largest VIF value among all the predictor variables is often used as an indicator of the severity of multicollinearity [60]. Table 17 shows the results from the

**TABLE 17**

**Multicollinearity Analysis (VIF) for Two Independent Variables (SKKN and TRNG)**

Parameter Estimates

| Variable | DF | Estimate | Std Error | t Stat | Pr >\|t\| | Tolerance | Var Inflation |
|---|---|---|---|---|---|---|---|
| Intercept | 1 | -1.2900 | 0.4146 | -3.11 | 0.0071 | . | 0 |
| SKKN | 1 | 0.0518 | 0.0546 | 0.95 | 0.3578 | 0.7919 | 1.2627 |
| TRNG | 1 | 0.5440 | 0.1624 | 3.35 | 0.0044 | 0.7919 | 1.2627 |

multicollinearity analysis. There seems to be no violation of the VIF factor within the model since the VIF values of the two independent variables are lower than 10.

The other option regarding multicollinearity exists in an SAS command. The **collinoint** option displays the condition number – a commonly used index of the global instability of the regression coefficients. The commands and the results of collinearity diagnostics are presented in Fig. 22.

A condition number – much larger than (approximately) 30 – could be, according to many authors, a sign of harmful multicollinearity [69]. Since the numbers in the results are much lower than 30, there is no violation of multicollinearity.

### 5.4.3. Hypothesis 3

*Hypothesis 3: High-SKKN level groups provide better detection (i.e. faster detection rate) than Medium-SKKN level groups*

```
proc reg data='C:\Program Files\SAS\SAS 9.1\My SAS Files\overalldata';
 model RPR = TRNG SKKN / vif tol collinoint;
run;
```

Collinearity Diagnostics (intercept adjusted)

| | | | --Proportion of Variation- | |
|---|---|---|---|---|
| Number | Eigenvalue | Condition Index | TRNG | SKKN |
| 1 | 1.45613 | 1.00000 | 0.27194 | 0.27194 |
| 2 | 0.54387 | 1.63625 | 0.72806 | 0.72806 |

**Fig. 22. SAS PROC REG command and results of collinearity diagnostics**

*5.4.3.1.Testing*

Two different groups are compared according to their SKKN level. It is assumed that the two groups (i.e., populations) have equal variances, and the populations from which the samples are selected have approximately normal distributions. We classify their SKKN level into three regions, with rounding off SKKN level: Low (1), Medium (2 – 5), and High (6). SKKN level of a team is represented by the person who has the highest SKKN level within his/her team. Thus, if one person's SKKN level is 6 and no one in his team has more than 6, his team has the SKKN value of 6. Comparison between the two groups is presented in Table 18, and the data used are displayed in Table 19.

**TABLE 18**

**Comparison between Two Groups with Different SKKN Level**

| | Group 1 (High-SKKN) | Group 2 (Medium-SKKN) |
|---|---|---|
| Groups | Echo04 | Alpha, Bravo, Echo03 |
| Characteristics | One superstar exists | No superstar exists<br>Medium-level personnel prevails |
| SKKN level | 6 | 2 – 5 |
| Security experience & knowledge (0.7) | > 24 mo. | > 0.5 mo. – ≤ 24 mo. |
| System experience (0.3) | > 60 mo. | > 6 mo. – ≤ 60 mo. |

**TABLE 19**

**Detection Rate of Each Group with Different SKKN Level**

| Group name | SKKN | Detection rate |
|---|---|---|
| Alpha | medium | 1 |
| | | 0.067 |
| | | 0.067 |
| | | 0.05 |
| Bravo | | 1 |
| | | 1 |
| | | 0.05 |
| Echo03 | | 0.2 |
| | | 0.05 |
| | | 0.1 |
| | | 0.033 |
| Echo04 | high | 1 |
| | | 1 |
| | | 0.033 |

To test the hypothesis that High-SKKN level groups provide better detection (i.e. faster detection rate) than Medium-SKKN level groups, we perform a hypothesis test about the difference between two population mean in a small sample case. The null and alternative hypothesis is as follows:

$H_0$: $\mu_1 = \mu_2$ (i.e. no difference in detection rate)

$H_1$: $\mu_1 \neq \mu_2$ (i.e. there is difference),

where $\mu_1$: true mean DTR for High-SKKN level groups, $\mu_2$: true mean DTR for Medium-SKKN level groups

**TABLE 20**

**Comparison Results of DTR for Two Groups with Different SKKN Level**

| | DTR of Group 1 (high-SKKN) | DTR of Group 2 (medium-SKKN) |
|---|---|---|
| Mean | 0.677666667 | 0.328818182 |
| Variance | 0.311696333 | 0.187823164 |
| Observations | 3 | 11 |
| Pooled Variance | 0.208468692 | |
| Hypothesized Mean Difference | 0 | |
| df | 12 | |
| t Stat | 1.173030684 | |
| P(T<=t) one-tail | 0.13177126 | |
| t Critical one-tail | 1.782286745 | |
| P(T<=t) two-tail | 0.26354252 | |
| t Critical two-tail | 2.178812792 | |

Using the EXCEL t-test (two sample assuming equal variances), we get the comparison results, shown in Table 20. The test statistic t is 1.173030684, which is the t-value calculated from the data. Note that the calculated t-value of 1.173030684 does not exceed the critical t-value (two-tailed) of 2.178812792. Thus, using a two-tailed test at a 5% level of significance (i.e. $\alpha = 0.05$), we fail to reject the null hypothesis. In other words, on the basis of the given data, we cannot support the claim that High-SKKN level groups provide better detection (i.e. faster detection rate) than Medium-SKKN level groups.

One possible reason for this result is that the variance of each group is rather high: Medium-SKKN level groups of 0.187823164, and High-SKKN level groups of

0.311696333. A second possible reason is that potential outliers are not removed from the data. Third possible reason is that the security attack detection was not that difficult, considering the attack comes from the CS department class students – and most of them are novices for this type of experiment.

### 5.4.4. Hypothesis 4

*Hypothesis 4: High-TRNG level groups provide better response (i.e. faster response rate) than Medium-TRNG level groups*

*5.4.4.1.Testing*

Two different groups are compared according to their TRNG level. It is assumed that the two groups (i.e., populations) have equal variances, and the populations from which the samples are selected have approximately normal distributions. We classify their TRNG level into three regions, with rounding off TRNG level: Low (1), Medium (2 – 3), and High (4 – 6). The reason why these regions have different scale from the SKKN case (in hypothesis 3) is that we need to differentiate the TRNG level of each group so that we can form two different TRNG groups (i.e., High and Medium). TRNG level of a team is represented by the person who has the highest TRNG level within his/her team. Comparison between the two groups is presented in Table 21, and the data used are displayed in Table 22.

**TABLE 21**

**Comparison between Two Groups with Different TRNG Level**

| | Group 1 (High-TRNG) | Group 2 (Medium-TRNG) |
|---|---|---|
| Groups | Bravo, Echo04 | Alpha, Echo03 |
| Characteristics | One (comparably) highly-trained personnel exists | No (comparably) highly-trained personnel exists<br>Medium-level personnel prevails |
| TRNG level | 4 – 6 | 2 – 3 |
| Security training (0.7) | > 4 mo. | > 0.5 mo. – ≤ 4 mo. |
| Other training (0.3) | > 12 mo. | > 1 mo. – ≤ 12 mo. |

**TABLE 22**

**Response Rate of Each Group with Different TRNG Level**

| Group name | TRNG | Response rate |
|---|---|---|
| Alpha | medium | 0.067 |
| | | 0.067 |
| | | 0.018 |
| Echo03 | | 0.05 |
| | | 0.1 |
| | | 0.029 |
| | | 1 |
| Bravo | high | 0.2 |
| | | 0.2 |
| | | 1 |
| Echo04 | | 1 |
| | | 1 |
| | | 1 |

To test the hypothesis that High-TRNG level groups provide better response (i.e. faster response rate) than Medium-TRNG level groups, we perform a hypothesis test about the difference between two population mean in a small sample case. The null and alternative hypothesis is as follows:

$H_0$: $\mu_1 = \mu_2$ (i.e. no difference in response rate)

$H_1$: $\mu_1 \neq \mu_2$ (i.e. there is difference),

where $\mu_1$: true mean RPR for High-TRNG level groups, $\mu_2$: true mean RPR for Medium-TRNG level groups

**TABLE 23**

**Comparison Results of RPR for Two Groups with Different TRNG Level**

| | RPR of Group 1 (high-TRNG) | RPR of Group 2 (medium-TRNG) |
|---|---|---|
| Mean | 0.733333333 | 0.190142857 |
| Variance | 0.170666667 | 0.128260476 |
| Observations | 6 | 7 |
| Pooled Variance | 0.147536017 | |
| Hypothesized Mean Difference | 0 | |
| df | 11 | |
| t Stat | 2.54188611 | |
| P(T<=t) one-tail | 0.013693935 | |
| t Critical one-tail | 1.795883691 | |
| P(T<=t) two-tail | 0.027387871 | |
| t Critical two-tail | 2.200986273 | |

Using the EXCEL t-test (two sample assuming equal variances), we get the comparison results, shown in Table 23. The test statistic t is 2.54188611, which is the t-value calculated from the data. Note that the calculated t-value exceeds the critical t-value (two-tailed) of 2.200986273. The means for RPR of the two groups are significantly different at $p = 0.027387871$. Thus, using a two-tailed test at a 5% level of significance (i.e. $\alpha = 0.05$), we reject the null hypothesis. In other words, on the basis of the given data, we support the claim that High-TRNG level groups provide better responses (i.e. faster response rate) than Medium-TRNG level groups.

Thus, it is necessary to hire and allocate to the incident response team highly trained security personnel when addressing the security response process. In other words, the efficiency of the incident response team comprised of at least one security personnel

with a high level of TRNG – greater than 4 – will be better than the team which has only medium-TRNG-level personnel.

# 6. SUMMARY

To have increased security through better technology, a large number of researchers and practitioners have studied the issue in terms of attack patterns, attack incident handling policies, training for security managers and analysts, building available patches, hardening OS, and so forth. It is a well-known fact that systems are usually vulnerable to attack by remote users or insiders; however, the human groups are another source of vulnerabilities due to a variety of reasons, including human misbehavior. Within the organization's budget, the resources such as human-resources must be allocated properly in order to effectively react to emergency situations (e.g. security attacks). Thus, it will be worthwhile to investigate the problem by measuring the effects of key human factors on the intrusion detection and incident response process. In other words, it is a matter of measuring security team performance in terms of efficiency in the process of intrusion detection and incident response.

The research questions set up, as shown in Section 1.3, supports the importance of the problem. They are presented here again:

1. Where can human vulnerabilities occur?
2. What approaches can be effective in handling them? Why?
3. What are the goals of using these approaches?

To answer these questions in the hope of providing an appropriate solution to the problem statement, we proposed an interactive model – consisting not only of a machine model but also a man model – called the Man-Machine Model ($M^3$), which can entail human factors. To the best of our knowledge, this is the first innovative effort to incorporate group behavior dynamics into Man-Machine Model ($M^3$) for the improvement of security defense processes in terms of human vulnerabilities. The model enables us to assess not only potential bottlenecks from machines, but also human vulnerabilities such as misbehavior or lack of knowledge/skill. By proposing the model,

we considered potentially important factors that might impact the security team efficiency. Through correlation analysis we obtained some factors (variables) which have strong relationships with the performance variables.

Some hypotheses seem to be plausible after investigating the correlation analysis, but we had to drop some hypotheses, discussed in Section 6.2, due to weak relationships. The efforts were to set up important research hypotheses that would enable us to derive quantitative models, which might explain the efficiency of the security defense team. We created regression models through hypothesis testing, and we tested the models in a way that allows us to interpret the hypothesis testing results.

Using the analysis and hypothesis testing results, we provide the answers to the following research questions. The first question is "Where can human vulnerabilities occur from?" The answer to this question comes from many sources including literatures and modeling. As we see from the case of the Slammer worm [14], the human vulnerabilities can occur whenever security personnel lack knowledge/experience or misbehave. These problem sources can occur in any defense process, which includes intrusion detection and incident response processes. To test this claim, we conducted hypothesis testing and found out those who have lower TRNG levels perform poorer than those who have higher TRNG levels, according to the hypothesis testing results of Hypothesis 4. We can infer that if security personnel have sufficient training experience to perform their role in a security defense process, they will seldom perform misconduct or make the information systems vulnerable or risky. In other words, they will be far from having human vulnerabilities.

A systematic approach to the second question of "What approaches can be effective to handle them? Why?" is to construct a model entailing a portion that can make the measurement of human vulnerabilities possible. To measure human stupidity as an indication of human vulnerabilities, for instance, we may have SKKN (Skill and Knowledge) or TRNG (Training). Likewise, to measure human carelessness, we may have VIGL (Vigilance) or TWQ (Teamwork Quality). This is the motivation of proposing the interactive model – $M^3$, which not only contains machine-related factors

but man-related factors such as SKKN and VIGL. Thus, we claim that our modeling approach can be an effective way to handle human vulnerabilities since the model $M^3$ incorporates security personnel behavior as well as system/tool configuration or functioning.

The answer to the third question of "What are the goals of using these approaches?" is to derive quantitative performance models of a security defense process (intrusion detection and incident response), where the models can analyze the efficiency of the security defense team. If the models are applied to an optimization problem, they should be able to help in human resource allocation problems, which will be discussed in Section 6.3.

## 6.1. Key Contributions

The key contributions of dissertation research can be summarized as follows. First of all, we developed a realistic, holistic security attack-defense model – Man Machine Model ($M^3$) – to deal with human vulnerabilities in security defense process. The model is realistic in that the model can measure human vulnerabilities, seldom investigated in information assurance fields. The model is holistic in that not only system (machine) components but human (man) components are developed.

Second, we obtained and evaluated several regression models to predict the efficiency of security defense teams whose key human factors can influence the efficiency. For instance, we can use the models to predict detection rates when the level of security personnel' Skill and Knowledge (SKKN) is low (e.g. 1.6 out of 6). We can also predict response rate when the level of security personnel' Training (TRNG) is high (e.g. 4.4 out of 6). We can likewise predict the increase or decrease of the efficiency of security defense team while varying the values of key variables such as SKKN and TRNG. Thus, when it concerns investments to increase the level of skill and knowledge of security personnel, organizations can either hire personnel with an appropriate level of skill and knowledge or they can make efforts to increase the level of skill and knowledge of security personnel they are presently working with.

Third, the man model in $M^3$ can be used as a reference model to other domains. The model can be applicable or adaptable to other fields, as it is not necessarily confined to the information assurance fields. The model can be used in fields whose processes involve group behavior, allowing determinations about what group behaviors impact the processes being modeled and how best to improve those processes.

## 6.2. Discussion

### 6.2.1. Hypothesis 2 with Adjustment

After excluding the outlier we found during the examination of Hypothesis 2, we obtained data file through several regression analyses and diagnostics, as shown in Table 24. We repeated the SAS fit analysis since we already had another data file that did not include the outlier. The analysis results are shown in Table 25. With these data, we repeat the hypothesis testing that was performed in Section 5.4.2.1; however, we do this testing succinctly this time since we previously presented the testing process in depth.

Hypothesis testing for determining whether the overall multiple regression model is useful for predicting Y(RPR) from $X_1$ (TRNG) and $X_2$ (SKKN), that is, testing usefulness of the model follows. Since n = 17 and k = 2, the denominator degrees of freedom is 14. The rejection region for the test is

$$F > F_{\alpha} = 3.74, \text{ where } \alpha = 0.05.$$

Since the F value of 33.61 (shown in Table 25 in the row corresponding to Model (in ANOVA section)), exceeds the critical value, $F_{0.05} = 3.74$, we may reject the null hypothesis and conclude that at least one of the two parameters ($\beta_1$ and $\beta_2$) is nonzero. In other words, the model appears to be useful for predicting Y, RPR.

**TABLE 24**

**SKKN, TRNG, and RPR Data without the Outliers**

| | SKKN | TRNG | UDST | NOTL | VIGL | TWQ | RPR | R_RPR |
|---|---|---|---|---|---|---|---|---|
| 1 | 3.2 | 2.7 | 3 | 1 | 3.7 | 4.67 | . | . |
| 2 | 3.2 | 2.7 | 3 | 1 | 3.6 | 4.70 | 0.067 | -0.1638 |
| 3 | 2.2 | 2.4 | 4 | 1 | 4.2 | 4.91 | 0.067 | 0.0449 |
| 4 | 1.6 | 2.7 | 2 | 1 | 4.5 | 5.34 | 0.018 | -0.2128 |
| 5 | 2.5 | 2.4 | 5 | 1 | 4.8 | 5.74 | 0.200 | 0.1779 |
| 6 | 2.5 | 2.4 | 5 | 1 | 5.0 | 5.92 | 0.200 | 0.1779 |
| 7 | 1.6 | 3.9 | 3 | 1 | . | . | 1.000 | -0.0653 |
| 8 | 1.3 | 2.4 | 4 | 3 | . | . | 0.040 | 0.0179 |
| 9 | 1.0 | 2.4 | 3 | 3 | 4.8 | 5.54 | 0.067 | 0.0449 |
| 10 | 1.0 | 2.4 | 3 | 3 | 4.5 | 5.20 | 0.040 | 0.0179 |
| 11 | 1.0 | 2.4 | 4 | 3 | 4.5 | 5.13 | 0.038 | 0.0159 |
| 12 | 1.6 | 2.4 | 1 | 1 | 4.3 | 4.89 | 0.050 | 0.0279 |
| 13 | 3.0 | 3.0 | 2 | 1 | 4.2 | 5.31 | 0.100 | -0.3394 |
| 14 | 1.6 | 3.0 | 3 | 1 | 4.1 | 4.25 | 0.029 | -0.4104 |
| 15 | 1.3 | 2.4 | 5 | 5 | 4.2 | 4.67 | 0.050 | 0.0279 |
| 16 | 6.0 | 3.4 | 5 | 5 | 3.7 | 4.63 | 1.000 | 0.2824 |
| 17 | 6.0 | 3.4 | 5 | 5 | 3.7 | 4.61 | 1.000 | 0.2824 |
| 18 | 1.7 | 3.7 | 4 | 5 | 3.9 | 4.43 | 1.000 | 0.0738 |
| 19 | 1.0 | 2.4 | 5 | 4 | 3.9 | 5.22 | . | . |
| 20 | 1.6 | 4.4 | 5 | 4 | 5.0 | 5.90 | . | . |
| 21 | 1.6 | 3.3 | 5 | 4 | 5.0 | 5.90 | . | . |

Measuring how well the Model fits the data follows. $R^2$ (shown in the Summary of Fit section in Table 25) is 0.8276, meaning that approximately 83% of the variability of RPR is accounted for by the two variables – TRNG and SKKN – in the regression model. Note that the adjusted R-square is 0.8030, which indicates that approximately 80% of the variability of RPR is accounted for by the regression model, even after taking into account the number of predictor variables in the model. The difference between the value of $R^2$ and that of the adjusted $R^2$ is 0.0246, which is closer than the previous results containing the outliers.

Testing significance of the overall model follows. The p-value for the test is shown to the right hand side of the *F*-value in Table 25, i.e. <.0001. This means that if the model did not contribute any information for the Y prediction, the probability of observing *F* statistic of 33.61 would be only less than 0.0001. Since the p-value is very small (less than 0.0001), the model is statistically significant.

**TABLE 25**

**SAS Fit Analysis for RPR with TRNG and SKKN as Its Predictors (Outliers Excluded)**

RPR = TRNG SKKN

Response Distribution: Normal

Link Function: Identity

Model Equation

RPR = -1.5560 + 0.6081 TRNG + 0.0664 SKKN

Summary of Fit

| | | | |
|---|---|---|---|
| Mean of Response | 0.2921 | R-Square | 0.8276 |
| Root MSE | 0.1811 | Adj R-Sq | 0.8030 |

Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Stat | Pr > F |
|---|---|---|---|---|---|
| Model | 2 | 2.2045 | 1.1023 | 33.61 | <.0001 |
| Error | 14 | 0.4591 | 0.0328 | | |
| C Total | 16 | 2.6636 | | | |

Type III Tests

| Source | DF | Sum of Squares | Mean Square | F Stat | Pr > F |
|---|---|---|---|---|---|
| TRNG | 1 | 1.2745 | 1.2745 | 38.86 | <.0001 |
| SKKN | 1 | 0.1356 | 0.1356 | 4.13 | 0.0614 |

Parameter Estimates

| Variable | DF | Estimate | Std Error | t Stat | Pr >\|t\| | Tolerance | Var Inflation |
|---|---|---|---|---|---|---|---|
| Intercept | 1 | -1.5560 | 0.2521 | -6.17 | <.0001 | . | 0 |
| TRNG | 1 | 0.6081 | 0.0975 | 6.23 | <.0001 | 0.8058 | 1.2410 |
| SKKN | 1 | 0.0664 | 0.0327 | 2.03 | 0.0614 | 0.8058 | 1.2410 |

Collinearity Diagnostics

| | | | Variance Proportion | | |
|---|---|---|---|---|---|
| Number | Eigenvalue | Condition Index | Intercept | TRNG | SKKN |
| 1 | 2.7944 | 1.0000 | 0.0037 | 0.0031 | 0.0267 |
| 2 | 0.1915 | 3.8200 | 0.0338 | 0.0138 | 0.8549 |
| 3 | 0.0141 | 14.0956 | 0.9625 | 0.9830 | 0.1184 |

Testing significance of the two predictor variables follows. Training (TRNG) is significant because p is less than 0.0001, and its coefficient is 0.6081. The positive coefficient for TRNG indicates that the higher the training level the security personnel possess, the better (higher) the response rate. Thus, this SAS analysis results make sense.

On the other hand, the coefficient for SKKN is not significantly different from 0 with the alpha level of 0.05 because its p-value of 0.0614 is greater than 0.05.

However, it is very close to 0.05. Thus, with the alpha level of 0.05, SKKN seems to be unrelated to the response rate (RPR); however, we cannot state this with confidence. With more data collected in the future, this hypothesis could be tested again to check the significance of SKKN since it may be possible that SKKN has a small p-value of 0.05. With the alpha level of 0.1, however, SKKN seems to be related to the Response Rate (RPR). Besides, the positive coefficient (0.0664) for SKKN indicates that the higher the skill and knowledge that security personnel possess, the better (higher) the response rate.

Since it can be useful to compare the hypothesis testing results before and after getting away with the outliers, we present the comparison results in Table 26. As you see in the table, there is improvement when we exclude the outliers; actually, just one data point – a single outlier. The R-square value increased by 0.27, and the adjusted R-square by 0.30. The p-value of the F-test and p-values of the t-test for TRNG and SKKN also increased considerably. Throughout the comparison the adjustment, or exclusion of the outliers, seems to affect the testing.

**TABLE 26**

**Comparison of Important Statistics – before and after the Outliers**

| | Before (with outliers) | After (without outliers) |
|---|---|---|
| R-square (Rs) | 0.5583 | 0.8276 |
| Adjusted R-square (Ra) | 0.4994 | 0.8030 |
| Difference between Rs and Ra | 0.0589 | 0.0246 |
| p-value of F-test | 0.0022 | < 0.0001 |
| p-value of T-test (TRNG) | 0.0044 | < 0.0001 |
| p-value of T-test (SKKN) | 0.3578 | 0.0614 |

### 6.2.2. Other Hypotheses

*6.2.2.1. Hypothesis 5*

*Hypothesis 5: Skill and Knowledge (SKKN) of security personnel and Teamwork Quality (TWQ) are statistically significantly as related to Detection Rate (DTR)*

Hypothesis 5 is that the two key factors, Skill and Knowledge (SKKN) and Teamwork Quality (TWQ), and a performance (i.e., efficiency) factor of Detection Rate (DTR) are dependent. In other words, SKKN and TWQ are linearly (or nonlinearly) related to DTR. This hypothesis is derived from (1) the experience of two network/system administrators at Computer Science Department at Texas A&M University, and (2) human vulnerabilities [8]. The two network/system administrators claim that both trust and skill sets are two important factors in security-related attack detection and response. J. E. Canavan [8] defines human vulnerability as human stupidity, carelessness, laziness, greed, and anger. He stresses that human vulnerability can be considered the greatest threats to networks and systems, and can (or will) do more damage than other threats combined, i.e. system, physical, media, etc. Moreover, he mentions that human vulnerabilities and the risks associated with them are the most difficult to defend against. To measure trust, we should have TWQ, which contains two items – mutual support and cohesion – that may help measure the degree of trust among team members. Likewise, to measure skill sets, we should have SKKN that covers both skill sets and knowledge. Lack of knowledge or skills can be measured from SKKN. Human carelessness and laziness can be measured from TWQ because TWQ contains an item of effort, which can measure how careless or lazy the security personnel are.

This hypothesis is very important since the degree of effect of the key factors on security personnel's performance, i.e. detection rate, may influence both the amount of investment on overall security process and re-coordination of current security environment settings in organizations. Furthermore, it may change the mindsets of those who deal with security-related tasks because they usually focus on algorithmic or technical factors to deal with security attacks, not human-related ones.

*6.2.2.2. Hypothesis 6*

*Hypothesis 6: Training (TRNG) of security personnel is statistically significantly related to the Detection Rate (DTR)*

Hypothesis 6, also as important as Hypothesis 5, is that a key factor Training (TRNG) and Detection Rate (DTR) are dependent. This hypothesis is derived from conclusions reached by other authors encountered in the literature. For example, ISO/IEC 17799 [5], [6] claim that providing appropriate training and education is often critical to the successful implementation of information security within an organization. Also, A NIST Handbook [58] states that a sound awareness and training program can help an organization reduce the number and severity of errors and omissions, which are an important threat to data and system integrity.

This hypothesis is important because the degree of effect of training on security personnel's performance, i.e. detection rate, may affect organizations' spending on training programs in hopes of improving the level of security defense in both detection and response. Organization may conserve resources if they know how much money they should spend on their training programs for better security defense.

### 6.2.3. Reliability Analysis

To measure reliability of measuring instruments, I perform reliability analysis. Cronbach's alpha coefficient, a well-known method, is used to check the reliability. Cronbach's alpha coefficient examines if the relationship between true values and observed values is strong. Using Cronbach's alpha coefficient, reliability results with three variables that have more than two measuring instruments – Skill & Knowledge (SKKN), Training (TRNG), and Teamwork Quality (TWQ) – are presented in Table 27. Both SKKN and TRNG have two measurements; TWQ have five measurements. The analysis was performed with SPSS, well-known statistical software.

**TABLE 27**

**Reliability of Measures**

| Measure | Reliability |
|---|---|
| SKKN | 0.635 |
| TRNG | 0.273 |
| TWQ | 0.903 |

Based on a rule of thumb [24], reliability of SKKN is questionable because it is less than 0.7; reliability of TRNG may not be acceptable because it is less than 0.5; reliability of TWQ is excellent because it is greater than 0.9. The necessity to increasing the reliability of TRNG is addressed in Section 6.4.

## 6.3. Lessons Learned

One of the lessons learned through our research experiments is that unexpected things happen. Initially, we thought TWQ – one of the key variables in the group behavior model (man model) – would be a strong predictor variable in evaluating the efficiency of the security defense team. However, correlation analysis revealed that TWQ was not related to the predicted variables such as DTR (Detection Rate) and RPR (Response Rate). The reasons for this remain unclear, though it is possible to speculate.

First, it might be possible that the defenders were not eager to successfully fill out the seven-page Teamwork Quality (TWQ) form during the security experiments; not only was the form the lengthiest out of the five, they had to spend some time in thinking about their mental states, communication, efforts, etc. Second, it may be possible that the defenders filled out the TWQ form using the same rationale as the previous forms. Third, it might be possible that there were not enough potential attacks or that the defenders were already prepared to detect and react through hardening their OS and installing several good intrusion detection tools; if there were not many attacks, it is possible that the teamwork quality of the defense groups were rather good, meaning that there were

not many ups and downs in their working conditions. Alternately, if the defenders were well prepared, it might be possible that they were confident to begin with and through successful intrusion detection their confidence grew over time.

Additionally, we initially hoped to obtain considerable amounts of data from the defenders and attackers. However, we failed to obtain much data because only several defenders willingly filled out the data forms according to our instructions. The failure resulted in a bottleneck of more accurate prediction and estimation. Thus, we realized that it was difficult to have human subjects willingly participate in this data collection since they had their own jobs to perform for their class. To obtain more data, we will have to have our own experiment environments in which the systems for experiments are prepared appropriately and the human subjects recruited more willingly participate.

## 6.4. Future Work

It is our hope that future studies use our study results as a framework for extending these research hypotheses. To extend research hypotheses, further investigation of other uncovered but possibly influential factors will be needed. These factors can be applied to set up hypotheses for verifying the level of statistical relevance in the process of intrusion detection and incident response. At the same time, other processes such as preparation could be focused on as well. In that case, preparedness might be an important factor (variable) in predicting the efficiency of the security preparation process. To be more specific, preparedness – as a complex variable such as TWQ – can play an important role in combining several variables.

The quality of either the intrusion detection or incident response process is another area for future studies to measure and predict the effectiveness of security defense processes. While current research focuses more on efficiency, future studies can take effectiveness as another performance issue in security defense process. With efficiency only, no one knows how good the detection or response was; one only knows how fast the defenders' detection and response was. To address this oversight, it is worthwhile to measure the quality of security processes or security defense teams.

Our proposed model, $M^3$, and the regression models obtained through hypothesis testing can be applied by organizations towards making the most appropriate allocations in human resources with regard to improving security defense processes. We do not specifically cover the issue of optimization in this dissertation, but it is a problem of interest. The results of human resource allocations can influence security defense teams by increasing confidence, combating skill degradation, and maintaining long-term effectiveness; therefore the allocation issue will play an important role in the hiring or firing of security personnel in an organization. Through human resource allocation, an organization can arrange the tasks undertaken by specific personnel in various situations (such as security emergencies); this would impact hiring strategies by giving organizations the ability to hot-swap security team members on a per-occasion basis in order to improve effectiveness. Additionally, when hiring decisions are not immediately confident, the efficiency and effectiveness of the security defense team can be measured to investigate potential bottlenecks that might keep the team from increasing its performance.

Simulation is another important issue to be studied further. The simulation of $M^3$ was under the development of the intrusion detection and incident response processes using EXTEND [37], but it is almost ready to be used for further study such as sensitivity analysis. Through sensitivity analysis, one can identify the impact of key predictor variables (e.g. training) on the predicted variable (e.g. response rate) by increasing or decreasing the coefficient values of key predictor variables. Simulation also can be used for tradeoff analysis when comparing different levels of security teams – for instance, one group with one 'superstar' and other personnel who lack equal Skill and Knowledge, and another with two personnel with moderate levels of Skill and Knowledge and other personnel who lack similarly rated Skill and Knowledge. The simulation results can therefore help to understand tradeoffs between the two groups.

Finally, reliability problem addresses the issue of whether this instrument will produce the same results each time it is administered to the same person in the same setting [24]. High reliability may not guarantee good scientific or engineering results,

but without reliability good results cannot exist. Reliability, while not sufficient condition of the value of research results and their interpretation, is important and necessary [46]. Therefore, it is necessary that we should examine further and find appropriate methods to increase the reliability of measuring instrument of TRNG high since the variable did not have strong reliability.

# REFERENCES


[1] J. Allen, A. Christie, W. Fithen, J. McHugh, J. Pickel, and E. Stoner, "State of the Practice of Intrusion Detection Technologies," Technical Report CMU/SEI-99TR - 028, CMU/SEI, January 2000, Available at citeseer.nj.nec.com/allen00state.html.

[2] S. Anderson, "Review: Linear Regresion Model," Department of Economics, The University of British Columbia, Available at www.econ.ubc.ca/asiwan/review.pdf, Accessed May 2005.

[3] J. Banks, J. S. Carson II, B. L. Nelson, and D. M. Nicol, *Discrete-event system simulation*, 3rd Ed., Prentice-Hall, Upper Saddle River, NJ, 2000.

[4] M. Bishop, "Vulnerabilities Analysis," *Proceedings of the Recent Advances in Intrusion Detection*, pp. 125-136, September 1999.

[5] British Standards Institution, "ISO/IEC17799:2000 Information Technology - Code of Practice for Information Security Management," International Standard, 2000.

[6] British Standards Institution, "ISO/IEC17799:2005, "Information Technology - Security Techniques - Code of Practice for Information Security Management," International Standard, 2005.

[7] S. Brocklehurst, B. Littlewood, T. Olovsson, and E. Jonsson, "On Measurement of Operational Security," *COMPASS '94*, *Proc. Ninth Ann. IEEE Conf. Computer Assurance*, Gaithersburg, MD, IEEE Computer Society, pp. 257 – 266, 1994.

[8] J. E. Canavan, *Fundamentals of Network Security*, Artech House, Inc., Boston, MA, 2001.

[9] K. Chamberlain and S. Zika. "The Minor Events Approach to Stress: Support for the Use of Daily Hassles," *British Journal of Psychology*, vol. 81, pp. 469-481, 1990.

[10] S. Chatterjee, A. S. Hadi, and B. Price, *Regression Analysis by Example*, Third Edition, John Wiley & Sons, Inc., New York, NY, 2000.

[11] G. Chiola, M. Marsan, G. Balbo, and G. Conte, "Generalized Stochastic Petri Nets: A Definition at the Net Level and Its Implications," *IEEE Trans. Software Engineering*, vol. 19, no.2, pp. 89-106, Feb 1993.

[12] M. Chouikha and E. Schnieder, "Modelling of Continuous-Discrete Systems with Hybrid Petri Nets," *Proceedings of the IEEE International Conference on Computational Engineering in Systems Applications (CESA '98),* P. Borne, M. Ksouri and A. El Kamel, Eds., Nabeul-Hammamet, Tunisia, pp. 606-612, 1998.

[13] A. M. Christie and M. J. Staley, "Organizational and Social Simulation of a Software Requirements Development Process," *ProSim99*, Silver Falls, OR, pp. 103-110, June 1999.

[14] Cisco Systems, "Combating Internet Worms SQL Slammer: An Integrated Security Approach," February, 2003, Available at www.cisco.com/application/ vnd.ms-powerpoint/en/us/guest/netsol/ns128/c654/ccmigration _09186a00801 415e0.ppt.

[15] E. Cole, *Hackers Beware*, New Riders, Indianapolis, IN, 2002.

[16] Colored Petri Nets (CPN) Group, Department of Computer Science, University of Aarhus, Denmark, "Petri Nets Tools and Software website," Available at www.dai mi.au.dk/PetriNets/ tools/, Accessed April 2002.

[17] Common Criteria (CC), "Common Criteria for IT Security Evaluation", Available at http://www.commoncriteria.org, Accessed October 2003.

[18] F. Cuppens, F. Autrel, A. Miege, and S. Benferhat, "Correlation in an Intrusion Detection Process," *Securite des Communications sur Internet (SECI 2002)*, pp. 153-172, Sep. 2002.

[19] R. J. Curts and D. E. Campbell, *Building a Global Information Assurance Program*, Auerbach Publications, Boca Raton, FL, 2003.

[20] I. Demongodin and N. Koussoulas, "Differential Petri Nets: Representing Continuous Systems in a Discrete-Event World," *IEEE Trans. Automatic Control,* vol. 43, no. 4, pp. 573-578, 1998.

[21] U.S. Department of Defense (DoD), "Department of Defense Information Technology Security Certification and Accreditation Process (DITSCAP)," DoD 8510.1-M, Alexandria, VA: Defense Technical Information Center (DTIC), 31 July 2000.

[22] R. Drath, Technical University of Ilmenau, "Visual Object Net ++," Available at www.systemtechnik.tu-ilmenau.de/~drath/visual_E.htm. Accessed January 2002.

[23] P. Fites and M. P. J. Kratz, *Information Systems Security: A Practitioner's Reference*, Van Nostrand Reinhold, New York, 1993.

[24] D. George and P. Mallery, *SPSS for Windows Step by Step: A Simple Guide and Reference, 10.0 Update*, third edition, Allyn & Bacon, Boston, MA, 2001.

[25] D. Gilliam, J. Kelly, and M. Bishop, "Reducing Software Security Risk through an Integrated Approach," *Proceedings of the Ninth IEEE International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises,* pp. 141-146, June 2000.

[26] D. Gilliam, J. Kelly, J. Powell, and M. Bishop, "Development of a Software Security Assessment Instrument to Reduce Software Security Risk," *Proceedings of the Tenth IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises*, pp. 144-149, 2001.

[27] L. Goin. "A Measure of Confidence," Available at www.buyandhold.com/bh/en/education/mom/linda/2001/mom20.html, Accessed May 2004.

[28] G. Gross, IDS News Service, "Study: Information Security Field to Grow Steadily," November 2004, Available at www.nwfusion.com/news/2004/1109studyinfor.html?nl.

[29] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," *Human Mental Workload*, P. A. Hancock and N. Meshkati (Eds.), Elsevier Science Publishers B. V., North-Holland, Amsterdam, pp. 139-183, 1988.

[30] U. Helmbrecht, "IT Security Guidelines: IT Baseline Protection in Brief, Federal Office for Information Security," Available at www.bsi.bund.de/english/gshb/guidelines/guidelines.pdf, Accessed Feb. 2004.

[31] D. Herrmann, *A Practical Guide to Security Engineering and Information Assurance*, Auerbach, Boca Raton, FL, 2002.

[32] D. S. Herrmann, *Using the Common Criteria for IT Security Evaluation,* Auerbach Publications, Boca Raton, FL, 2003.

[33] M. M. Hoegl and H. G. Gemuenden, "Teamwork Quality and the Success of Innovative Projects: a Theoretical Concept and Empirical Evidence," *Organizational Science*, vol. 12, no. 4, pp. 435-449, 2001.

[34] T. H. Holmes and R. H. Rahe, "The Social Readjustment Raring Scale," *Journal of Psychosomatic Research*, vol. 2, pp. 213-218, 1967.

[35] J. D. Howard and T. A. Longstaff, *A Common Language for Computer Security Incidents*, SAND98-8667, Livermore, CA: Sandia National Laborotories, 1998.

[36] Idaho Geospatial Data Center, "Collaborative Spatial Decision Making (CSDM)," Available at geolibrary.uidaho.edu/courses/Geog427/ Lectures/8/Lecture8.doc, Accessed August 2003.

[37] Imagine That, Inc. *Extend 5 User's Guide*, San Jose, CA, 2000.

[38] H. P. In, S. Jung, S. Poole, and S. Liu, "Modeling Man and Machine Interactions for Virtual Vulnerability Defense," *The 3rd IEEE SMC Information Assurance Workshop*, West Point, NY, pp. 89-96, 2002.

[39] Information Assurance Division, School of Information Technology, Fort Gordon, Georgia, "IS Terminology," Available at ia.gordon.army.mil/iaso/lesson01.htm, Accessed January 2004.

[40] International Committee for Information Technology Standards (INCITS), "ISO/IEC 13335-1(1996-12) Information Technology – Guidelines for the Management of IT Security – Part 1: Concepts and Models for IT Security," 1996.

[41] ISO/IEC Copyright Office, International Standard, "ISO/IEC 15408-1 (1999-12-01) Information Technology – Security Techniques – Evaluation Criteria for IT Security – Part 1: Introduction and General Model," 1999

[42] J. M. Ivancevich, "Life Events and Hassles as Predictors of Health Symptoms, Job Performance, and Absenteeism," *Journal of Occupational Behavior*, vol. 7, pp. 39-51, 1986.

[43] R. Jain, *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling*, John Wiley & Sons, New York, NY, 1991.

[44] E. Jonsson and T. Olovsson, "A Quantitative Model of the Security Intrusion Process Based on Attacker Behavior," *IEEE Transactions on Software Engineering*, vol. 23, no. 4, pp. 235-245, April 1997.

[45] E. Jonsson and T. Olovsson, "Security Intrusion Process: An Empirical Model," *IEEE AES Systems Magazine*, vol. 124, pp. 7-17, April 1997

[46] F. N. Kerlinger, *Foundations of Behavioral Research: Educational and Psychological Inquiry*, Holt, Rinehart, and Winston, Inc., New York, NY,1965.

[47] P. Landreville, "A Comparison between Daily Hassles and Minor Life Events as Correlates of Well-being in Older Adults," *Canadian Journal of Aging*, vol. 11, no. 2, pp. 137-149, 1992.

[48] A. M. Law and W. D. Kelton, *Simulation Modeling and Analysis*, 3rd edition, McGraw-Hill, New York, NY, 2003.

[49] W. Lee, W. Fan, M. Miller, S. Stolfo, and E. Zadok, "Toward Cost-Sensitive Modeling for Intrusion Detection and Response," *Workshop on Intrusion Detection and Prevention*, *7th ACM Conference on Computer Security*, Athens, Greece, Available at citeseer.ist.psu.edu/cache/papers/cs/18237/http:zSzzSzwww.csc.ncsu.eduzSzfacultyzSzleezSzpaperszSzcost_modeling.pdf/lee00toward.pdf, November, 2000.

[50] R. A. Letteer, Department of Defense, "Information Operations and the DAA", Available at http://iase.disa.mil/ditscap/daav3.ppt, Accessed April 2004.

[51] B. Littlewood, S. Brocklehurst, N.E. Fenton, P. Mellor, S. Page, D. Wright, J.E. Dobson, J.A. McDermid, and D. Gollmann, "Towards Operational Measures of Computer Security," *Journal of Computer Security*, vol. 2, no. 3, pp. 211 – 229, 1993.

[52] Longman Publishing, *Longman Dictionary of Contemporary English*, New Edition, Pearson, Upper Saddle River, NJ, 1993.

[53] J. Miles and M. Shevlin, *Applying Regression and Correlation: A Guide for Students and Researchers*, Sage Publications, Thousand Oaks, CA, 2001.

[54] S. M. Monroe, "Major and Minor Life Events as Predictors of Psychological Distress: Further Issues and Findings," *Journal of Behavioral Medicine*, vol. 6, no. 2, pp. 189-205, 1983.

[55] T. Murata, "Petri Nets: Properties, Analysis and Applications," *Proceedings of the IEEE*, vol. 77, no. 4, pp. 541-580, April 1989.

[56] National Computer System Security and Privacy Advisory Board, "1991 Annual Report," Gaithersburg, MD, March 1992, Available at csrc.nist.gov/csspab/reports/91-rpt.txt.

[57] National Information Assurance Partnership (NIAP), "Guidance to Validators of IT Security Evaluations," Common Criteria Evaluation and Validation Scheme (CCEVS) for IT Security, Scheme Publication 3, National Information Assurance Partnership, February, 2002.

[58] National Institute of Standards and Technology (NIST), U.S. Department of Commerce, "An Introduction to Computer Security: The NIST Handbook," Special Publication 800-12, October 1995**,** Available at csrc.nist.gov/publications/nistpubs/800-12/handbook.pdf.

[59] National Security Telecommunications and Information System Security Committee (NSTISSI), "NSTISSI #1000: National Information Assurance Certification and Accreditation Process (NIACAP)," April, 2000.

[60] J. Neter, M. H. Kutner, C. J. Nachtsheim, and W. Wasserman, *Applied Linear Regression Models*, Third Edition, Irwin, Chicago, IL, 1996.

[61] P. Neumann, *Computer Related Risks*, Addison-Wesley, Reading, MA, 1995.

[62] R. Ortalo, Y. Deswarte, and M. Kaaniche, "Experimenting with Quantitative Evaluation Tools for Monitoring Operational Security," *IEEE Transactions on Software Engineering*, vol. 25, no. 5, pp. 633-650, September/October, 1997.

[63] C. Pao, P. Schmid, and J. Glass, "Confidence Scoring for Speech Understanding Systems," Available at www.sls.lcs.mit.edu/sls/publications/1998/icslp98-confidence.pdf, Accessed January 2003.

[64] B.L. Pellom, J.H.L. Hansen, A Duration-Based Confidence Measure for Automatic Segmentation of Noise Corrupted Speech, *ICSLP-98: Inter. Conf. on Spoken Language Processing*, vol. 6, pp. 2723-2726, Sydney, Australia, Dec. 1998.

[65] S. Pettersson and B. Lennartson, "Hybrid Modelling Focused on Hybrid Petri Nets", *The Second European Workshop on Real-time and Hybrid Systems*, Grenoble, France, pp. 303-309, June 1995.

[66] U.W. Pooch, "Lab Policies and Procedures for CPSC665," Department of Computer Science, Texas A&M University, Available at faculty.cs.tamu.edu/ pooch/course/CPSC665/Spring2002/Lab_Policies_and_Procedures_for_CPSC_665.pdf, Accessed April 2002.

[67] M. Rauterberg, M. Fjeld, and S. Schluep, "Goal Setting Mechanism in Petri Net Models of Human Decision Making," *IEEE International Conference on Systems, Man, and Cybernetics*, *Computational Cybernetics and Simulation*, vol. 3, pp. 2696-2701, 1997.

[68] D. J. Reifer (Ed.), *Software Management*, 5th Edition, IEEE Computer Society Press, Los Alamitos, CA, 1997.

[69] Resa Corporation and Licensors, "Online Economics Textbook – Regression Extensions – Multicollinearity – Detection of Multicollinearity," Available at www.xycoon.com/detection.htm, Accessed June 2004.

[70] A. E. Rizzoli, "A Collection of Modeling and Simulation Resources on the Internet," Available at www.idsia.ch/~andrea/simtools.html, Accessed April 2002.

[71] J. Roculan, S. Hittel, D. Hanson, J. Miller, B. Kostanecki, J. Gough, M. Velzen, and O. Friedrichs, "SQLExp SQL Server Worm Analysis," January 2003, Available at securityresponse.symantec.com/avcenter/Analysis-SQLExp.pdf.

[72] S. Ruben, B. Pellom, and W. Ward, "Confidence Measures for Dialogue Management in the CU Communicator System," Available at communicator.colorado.edu/papers/icassp-2000.pdf, Accessed January 2005.

[73] E. E. Schultz and R. Shumway, *Incident Response: A Strategic Guide to Handling System and Network Security Breaches*, New Riders, Indianapolis, IN, London, 2002.

[74] C. Shannon, "Communication Theory," Available at www.exploratorium.edu/turbulent/CompLexicon/Shannon.html, Accessed May 2003.

[75] G. Simmons (ed.), *Contemporary Cryptography: The Science of Information Integrity*, IEEE Press, Piscataway, NJ, 1992.

[76] T. Sincich, *Statistics by Example*, Third edition, Dellen Pub. Co., San Francisco, CA, 1987.

[77] T. Sincich, D. M. Levine, and D. Stephan, *Practical Statistics by Example Using Microsoft Excel and MINITAB*, 2nd Edition, Prentice Hall, Upper Saddle River, NJ, 2002.

[78] D. Smith, "Forming an Incidence Response Team," Available at www.auscert.org.au/render.html?it=2252&cid=1920, Accessed August 2004.

[79] W. Stallings, *Network Security Essentials: Applications and Standards*, Prentice-Hall, Upper Saddle River, NJ, 2000.

[80] G. Stoneburner, C. Hayden, and A. Feringa, "NIST Special Publication 800-27 Engineering Principles for Information Technology Security (A Baseline for Achieving Security)," U.S. Department of Commerce, pp. 800-827, National Institute of Standards (NIST), Washington, DC, June 2001.

[81] R. W. Swezey and E. Salas (Ed.), *Teams: Their Training and Performance*, Ablex Publishing, Norwood, NJ, 1992.

[82] B. D. Thomas, "Intrusion Detection Primer," Available at www.linuxsecurity.com/feature_stories/feature_story-8.html, Accessed September 2004.

[83] H. F. Tipton and M. Krause (Eds), *Information Security Management Handbook*, 4th Edition, Volume 3, Auerbach Publications, Boca Raton, FL, 2002.

[84] Vienna University of Technology, "ARGESIM Simulation Links," Available at eurosim.tuwien.ac.at/hotlinks, Accessed June 2002.

[85] H. Wei, D. Frinke, O. Carter, and C. Ritter, "Cost-benefit Analysis for Network Intrusion Detection Systems," *CSI 28th Annual Computer Security Conference*, Washington, DC, Available at www.csds.uidaho.edu/deb/costbenefit.pdf, October 29-31, 2001.

[86] Wikipedia, the Free Encyclopedia, "Information Security," Available at en.wikipedia.org/wiki/Information_assurance, Accessed May 2005.

[87] R. Williams, S. Zyzanski, and A. L. Wright, "Life Events and Daily Hassles and Uplifts as Predictors of Hospitalization and Outpatient Visitation," *Journal of Social Science and Medicine*, vol. 34, no. 7, pp. 763-768, 1992.

[88] M. Wilson (Ed.), D. E. de Zafra, S. I. Pitcher, J. D. Tressler, J. B. Ippolito, "NIST Special Publication 800-16 Information Technology Security Training Requirements: A Role- and Performance-Based Model," U.S. Department of Commerce, April 1998, Available at csrc.nist.gov/publications/nistpubs/800-16/800-16.pdf.

[89] T. M. Wolf, R. C. Elston, and G. E. Kissling, "Relationship of Hassles, Uplifts, and Life Events to Psychological Well-being of Freshman Medical Students," *Journal of Behavioral Medicine*, vol. 15, no. 1, pp. 37-45, 1989.

[90] B. J. Wood and J. F. Bouchard, "Improving Government-wide Emergency Response to Cyber Incidents," *Proceedings of the 2001 IEEE Workshop on Information Assurance and Security*, United States Military Academy, West Point, NY, pp. 195-198, June 2001.

[91] K. R. Wyk and R. Forno, *Incident Response*, O'Reilly & Associates, Inc. Sebastopol, CA, July 2001.

[92] Y. Xiong, J. Liu, and P. Sun, "On the Defense of the Distributed Denial of Service Attacks: an On-Off Feedback Control Approach," *IEEE Transactions on Systems, Man, and Cybernetics*, Part A, vol. 31, no. 4, pp. 282-293, July 2001.

[93] A. Yasinsac, "Policies to Enhance Computer and Network Forensics," *Proceedings of the 2001 IEEE Workshop on Information Assurance and Security*, United States Military Academy, West Point, NY, pp. 289-295, 5-6 June 2001.

[94] N. Ye, X. Li, Q. Chen, S. M. Emran, and M. Xu, “Probabilistic Techniques for Intrusion Detection Based on Computer Audit Data,” *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, vol. 31, no. 4, July 2001.

# APPENDIX A
# EXPERIMENTAL PROTOCOL FOR SECURITY ATTACKS AND DEFENSE

The purpose of this experiment is to have students work on a security attack and defense task using different types of attacks and defense mechanisms in order to assess vulnerability in systems and human factors. If you notice anything unusual during the conduct of this experiment period, please record your observations in the 'Comments' section of the 'Data Collection Form'.

Pre-Experiment:

**How to use the computer:**

1. Turn on the computer and login in Experiment Room 1. (You should use VPN service provided by CIS Network Group at Texas A&M University if you can't login directly.)
2. Check to make sure your computer has 'F-Secure SSH Client' and 'F-Secure SSH File Transfer' service. (You're supposed to have these services. If you don't have these services, please ask the Experiment Coordinator to install them for your work.)
3. Check to make sure you computer has 'Internet Explorer'. (If you don't have one, install one using 'Connection Wizard' on your desktop.)

Script for Experiment Coordinator

**Read:**

Hi. My name is ___(Coordinator's name here)_____ and I will be your experiment coordinator today. [If the group is not an intact group: "Please take a moment to introduce yourself to the group."] Thanks for coming today.

I appreciate your volunteering for this research. I am sure you will find this research an interesting and valuable experience for your career. The most important thing is that you take it seriously, and do your best to work to deal with every problem you may confront during the experiment. This experiment will take about seven hours.

Before we get started, I will pass out a consent form and a brief pre-experimental questionnaire. When you complete to fill out these forms, please wait quietly until everyone is finished. Please do not touch the computer to work until I give you instructions to do so. Would anyone like to have a copy of the consent form for your records?

**Do:**

1. Hand out the consent form and pre-experimental questionnaire.
2. On the data form, record students' names, their associated workstation number, coordinators' names, dates, start time, group number.
3. Collect the consent form and pre-experimental questionnaire. Be sure to offer the students a signed copy of their consent forms (if they want one, use the extras with the researcher's signature already on it and have them sign a second one.) Make sure the consent form is signed and that the pre-experimental questionnaire (background questionnaire) is filled out completely.

**Read:**

Your group work today will consist of a set of attack task and a coordinated defense task for attackers and defenders, respectively. You will need to work as a group, which means that you need to discuss and work together for better efficiency and effectiveness.

Training Guide for Attacks and Defenses will be passed out by the Experiment Coordinator to attackers and defenders, respectively. You may refer to the Training Guide to attack the computer system and defend the computer system. However, you also may use your knowledge in performing the task. In that case, you have to record about what you've done to attack the computer system or defend. You have to specify your activity on the 'Data Collection Form', which will be provided by the Experiment Coordinator soon.

**Do:**

4. Follow the instructions on the Training Guide for Attacks and Defenses for your task.

**Read:**

Now that you have received Training Guide for Attacks and Defenses, you are just about ready to begin your group work on the task. The task you'll be working on today is a group decision-making task. For each event occurred, you as a group member (defender) have to discuss and work together with your group member(s) to decide which action you'll take immediately or which procedure you'll follow to deal with the event. For each event you'll initiate to attack the computer system of the defenders, you as a group member (attacker) have to discuss and work together with your group member(s) to decide which computer system is the target and which attack methods you'll choose.

Now we'll discuss the form you'll be working on today. This form is to ask you to provide a set of security attack and defense activity information so that the researchers can better understand how each attacker and defender group works. Please answer each item as completely as possible.

**Do:**

5. Hand out the Data Collection Form to both attacker and defender groups. Make sure they understand that they are performing security attacks and defenses and filling out the form based on their activity during the experiment.

**Read:**

After a lunch break, you will go to your room where you were this morning to work as a group. You will have two hours to complete this task. I will let you know when twenty minutes are left. It is up to your group to decide how to allocate the time to your task and how to best make use the Training Guide and computer resources.

**Do:**

6. Tell the participants you will come by and collect the Data Collection Form in 10 to 15 minutes.
7. When completed the experiment, collect the form and make sure that all forms (including consent form, pre-experimental questionnaire, and Data Collection Form) have been filled out completely and correctly with the date and group number on ALL materials. If there are any missing items, have the participants fill them in. Have every participants return to Experiment Room 1.

**Read:**

To debrief you a little bit about what we hoped to gain from this study, our main purpose in this experiment was to study potential vulnerability in systems and human factors in attack-defense process, and what are most critical factors to efficiently and effectively respond to security attacks. If you are interested in further details about this study, we can provide you additional information once the entire study is completed.

<u>Post-Experiment:</u>

8. Be sure to have the participants make comments on the ‘Comments’ section of the Data Collection Form if they have some. Tell them to write down their comments before they forget.
9. Make sure every form is collected with the comments. If they have some questions, answer the questions.

# APPENDIX B
# EXPERIMENTAL RULE FOR THE ATTACKERS AND DEFENDERS

1. General Rules for the Attackers and Defenders

- You will be required to work one to two hours per day.
- You will be required to turn in the form you filled out during the experiment as soon as you finish so that the Coordinator knows the current experiment progress and take necessary actions if needed.
- You are not allowed to perform any action other than taking necessary steps to launch the attacks. For example, you should not modify any files you don't have permission on or delete any system files.
- If you are not sure the rules and have questions on the form, you should consult the coordinator.
- We suggest that you use the same computer during the whole experiment for the purpose of consistency.

2. The Rules for the Attackers

The attackers will be allowed to use three different types/levels of security attacks including ping of death (P), TCP SYN flooding attack (S), and Code Red (C). Each attack method represents software vulnerability attack [4], a protocol attack [4], and automated (autonomous) propagation attack using a blind targeting model [7], respectively. They will be provided necessary information such as tools by the coordinator since we assume that the attackers for the experiment represent the real attackers. The rules for the attackers are described as follows:

You are allowed to take necessary steps to launch security attacks against the designated system only. Any attack against other systems may cause a violation against the rules of the university's Institutional Review Board (IRB).

- You are not allowed to talk with any defender about your activity in the experiment since it may impair the integrity of the experimental result.

- You must fill out the form provided by the Coordinator once per day.

3. The Rules for the Defenders

The defenders belong to a set of groups with two students of a group. Each group has to work independently to take necessary actions to defend or mitigate the security attacks caused by the attackers. Each group will be provided necessary information such as tools by the coordinator because we assume that each group represents an Incidence Response Team of an organization. However, if they need further information to defend or mitigate, the information gathering is their responsibility. The rules for the defenders are described as follows:

- You can co-work anytime with your team member to operate effectively and correctly. You can be allowed to use books, documents, literature, information gathered from the Internet, and others.
- You are not allowed to work with other team members except your team member since it may impair the integrity of the experimental result.
- You must fill out the Data Collection Form once per day.
- You are not allowed to attack the attackers' system even though you can do so since we assume that the defenders take necessary actions only to defend/mitigate the security attacks.
- You are not allowed to ask the attackers any questions since it may impair the integrity of the experimental result.

4. Joint CS / INFO Security Exercise: Spring 2003 Semester

**GOAL:** Compromise the B2B/B2C Web Servers for team Alpha, Bravo, Charlie, Delta, and Echo located in the 10.10.50.X subnet. These web servers were setup and configured as a part of the INFO 689 Business Information Security course. No points will be awarded for attacking other systems in the 10.10.50.x network, i.e., the DNS server located at 10.10.50.2.

**INITIAL CONDITIONS:** Each team will have an account on a system "client.info689.org" (10.10.50.3) located in the 10.10.50.x subnet.

**OBJECTIVES:**

1) Steal a secret stored in /root/secret. This secret is unique to each team, so there are five possible secrets to steal.
2) Steal the credit card number authorization file. This credit card authorization file is unique to each team, so there are five possible credit card files to steal.
3) Steal the database containing the business inventory. This database is unique to each team, so there are five possible databases to steal.

**RULES:**

- No physical access to the Business Information Security Laboratory is allowed.
- No denial of service attacks will be permitted.
- All attacks MUST be conducted through the CS to INFO VPN tunnel into the 10.10.50.X network.
- **ONLY SYSTEMS in the 10.10.50.x subnet are to be attacked.** Attacking other sandboxed systems in the 10.10.x.x network will result in disqualification from the exercise as well as loss of all points.
- The period of engagement will begin 5:00PM CST on Friday, April 11, 2003 and end at 5:00PM CST Friday, April 17, 2003.

**REWARD:**

Bonus points may be awarded to teams that are successful in compromising server(s). In order to receive credit, a detailed description of the attack including tools used, date and time, and the objective information must be included.

# APPENDIX C
# DEFINITIONS

- **Attack** – 1) [1] An action conducted by an adversary, the attacker, on a potential victim

  2) [23] action(s) that prevent any part of an Automated Information System (AIS) from functioning in accordance with its intended purpose. This includes any action which causes the unauthorized destruction, modification, or delay of service; the act of aggressively trying to bypass security controls on an Automated Information System (AIS). The fact that an attack is made does not necessarily mean that it will succeed. The degree of success depends on the vulnerability of the system or activity and the effectiveness of existing countermeasures.
- **Attacker** [1] – an adversary who conducts an attack on a victim (e.g., host)
- **Incident** ([35] and [1]) – one or more related attacks that can be distinguished from other attacks because of the distinctiveness of attacker, type of attack, objectives, sites, or timing
- **Information assurance** [1] **-** The subfield of information science that focuses on the conditions necessary to assure users of information systems and services that they can expect:
  1. the information and services they use actually did originate with whom they claim and are exactly as the originator intended
  2. the information and services they use will be available when needed
  3. the information and services for which they are responsible will be made available only to those they intend and only in the manner that they intend
- **Intrusion** [1] **-** Actual illegal or undesired logical entry into an information system; The act of violating the security policy or legal protections that pertain to an information system

- **Intrusion detection** [82] – The process of preventing and detecting security breaches by monitoring user and application activity
- **Intrusion detection system (IDS)** [1] **-** A combination of hardware and software that monitors and collects system and network information and analyzes it to determine if an attack or an intrusion has occurred. Some ID systems can automatically respond to an intrusion
- **Monitoring** [1] **-** Observing a data stream for specified events to provide data for subsequent action or analysis
- **Response** [1] **-** Actions taken to protect and restore the normal operating condition of computers and the information stored in them when an attack or intrusion occurs
- **Security** [1] **-** The subfield of information science concerned with ensuring that information systems are imbued with the condition of being secure, as well as the means of establishing, testing, auditing, and otherwise maintaining that condition

  [21] (mentioned in [19]) measures and controls that ensure confidentiality, integrity, availability, and accountability of the information processed and stored by a computer
- **System** – 1) [1] one or more interconnected physical machines (hosts) operating in cooperation with one another to meet a particular mission. Systems are generally, although not necessarily, contained within one site. Hosts may participate in multiple systems. Systems may be wholly contained within one host or distributed across multiple hosts

  2) [3] a group of objects that are joined together in some regular interaction or interdependence toward the accomplishment of some purpose
- **Victim** [1] **-** That which is the target of an attack. An entity may be a victim of either a successful or unsuccessful attack
- **Vulnerability**

- 1) [1] **-** A feature or a combination of features of a system that allows an adversary to place the system in a state that is both contrary to the desires of the people responsible for the system and increases the risk (probability or consequence) of undesirable behavior in or of the system. A feature or a combination of features of a system that prevents the successful implementation of a particular security policy for that system. A program with a buffer that can be overflowed with data supplied by the invoker will usually be considered a vulnerability. A telephone procedure that provides private information about the caller without prior authentication will usually be considered to have a vulnerability
- 2) [8] - A vulnerability is an inherent weakness in the design, configuration, or implementation of a network or system that renders it susceptible to a threat. Most vulnerabilities can usually be traced back to one of three sources:
    - poor design
    - poor implementation
    - poor management

  While there are only threes sources of vulnerabilities, they can manifest themselves in many ways.
    - Physical vulnerabilities
    - Hardware and Software vulnerabilities
    - Media vulnerabilities
    - Transmission and Emanation vulnerabilities
    - Human vulnerabilities
- 3) [40], [61] (mentioned in [32]) - Vulnerability is weakness in the design, operation, or operational environment of an IT system or

product that can be exploited to violate the intended behavior of the system relative to safety, security, and/or integrity.

- **Team** ([81] and [73]) – a distinguishable set of one or more individuals who interact, dynamically, interdependently, and adaptively toward a common and valued mission/goal/objective, who have each been assigned a specific set of roles or duties – related to security (for this dissertation) - to perform, and who have a limited life-span of membership
- **Teamwork quality*** [33] – a comprehensive concept and measure of the quality of interactions, i.e., collaborations, in teams; consist of six facets: communication, coordination, cohesion, effort, mutual support, and balance of member contributions
- **Skill** [52] – special ability to do something well, especially as gained by learning and practice
- **Knowledge** [52] – the facts, information, skills, and understanding that one has gained, especially through learning or experience
- **Experience** [52] – (the gaining of) knowledge or skill which comes from practice in an activity or doing something for a long time, rather than from books
- **Training** [88] – strive to produce relevant and needed security skills and competencies
- **Objective information load** – the amount of cues (actions, events, etc., that provides a signal for something to be done or standard that can be copied [52]) and messages to be processed
- **Perceived mental workload** – the degree to which the team is under physical, mental, and temporal demand. It also includes frustration level, performance satisfaction level, and mental and physical effort level.
- **Verification** – 1) [68] - the process of evaluating a software system or component to determine if the products of a given development phase satisfy the conditions imposed at the start of that phase.

2) [43] - a step to determine whether the model implements the assumptions correctly. It is related to the correctness of the implementation of the assumptions. It is also called debugging, that is, ensuring that the model does what it is intended to do.

- **Model** [3] – a representation of a system for the purpose of studying the system
- **Simulation model** [3] - a particular type of mathematical model of a system
- **Mathematical model** [3] - a model that uses symbolic notation and mathematical equations to represent a system
- **Event** [43] - a change in the system state
- **Factors** [48] - the input parameters and structural assumptions composing a model, in experimental-design terminology
- **Preventive Security** [44] - system's ability to protect itself from external attacks
- **Security Attack** [79] - any action that compromises the security of information owned by an organization; Four general categories of attack would be the followings: 1) Interruption, 2) Interception, 3) Modification, 4) Fabrication
- **Authorization** [80] (mentioned in [19]) - The granting or denying of access rights to a user, program, or process
- **Integrity** [57] (mentioned in [32]) Prevention of unauthorized modification of information
- **Threat** – 1) [8] A threat is anything that can disrupt the operation, functioning, integrity, or availability of a network or system. The different categories of threats are the followings:
    - natural threats – occurrences such as floods, earthquakes, and storms
    - unintentional threats – the result of accidents and stupidity
    - intentional threats – the result of malicious indent

  Each type of threat can be deadly to a network.

2) [61] (mentioned in [32]) - Threat is potential danger that a vulnerability may be exploited intentionally, triggered accidentally, or otherwise exercised.

3) [40] (mentioned in [32]) - Threat is a potential cause of an unwanted incident which may result in harm to a system or organization.

4) [59] (mentioned in [32]) - Threat is any circumstance or event with the potential to harm an IT system through unauthorized access, destruction, disclosure, modification of data, and/or denial of service.

5) [23] A threat is any circumstance or event with the potential to harm to a system in the form of destruction, disclosure, modification of data, and/or denial of service. Common usage today is from the press, which uses the word to describe people who "break into" computers for various purposes.

- **Human Vulnerabilities** [8] - Human stupidity, carelessness, laziness, greed, and anger represent the greatest threats to networks and systems and will do more damage than the rest of the others combined. Moreover, human vulnerabilities and the risks associated with them are the most difficult to defend against. It is important to keep in mind that every network or system designed, configured or implemented has vulnerabilities. There is no such thing as a totally secure network or system. It does not exist!
- **Countermeasures** -
    - 1) [8] - the techniques or methods used to defend against attacks and to close or compensate for vulnerabilities in networks or systems
    - 2) [23] any action, device, procedure, technique, or other measure that reduces the vulnerability of an ADP system or activity to the realization of a threat
- **Information Security (INFOSEC) –**

- 1) [79], [75] - Information Security is about how to prevent cheating or, failing that, to detect cheating in information-based systems wherein the information itself has no meaningful physical existence.
- 2) [8] - Information security = confidentiality + integrity + availability + authentication

- **Information Assurance (IA)** –
  - 1) (U.S. DoD Directive 5-3600.1 in 1996, mentioned in [31]) - Information operations that protect and defend information and information systems by ensuring their availability, integrity, authentication, and nonrepudiation; including providing for restoration of information systems by incorporating protection, detection, and reaction capabilities.
  - 2) [31] - an engineering discipline that provides a comprehensive and systematic approach to ensuring that individual automated systems and dynamic combinations of automated systems interact and provide their specified functionality, no more and no less, safely, reliably, and securely in the intended operational environment(s).
  - 3) [39] - The protection of systems and information in storage, processing, or transit from unauthorized access or modification; denial of service to unauthorized users; or the provision of service to authorized users. It also includes those measures necessary to detect, document, and counter such threats. This regulation designates IA as the security discipline that encompasses COMSEC, INFOSEC, and control of compromising emanations (TEMPEST).
  - 4) [86] Information security deals with several different "trust" aspects of information. Another common term is information

assurance. Information security is not confined to computer systems, nor to information in an electronic or machine-readable form. It applies to all aspects of safeguarding or protecting information or data, in whatever form.

- **Assurance** [17] (mentioned in [32]) - Grounds for confidence that an entity meets its security objectives
- **Security Assurance** [41] (mentioned in [32]) - Grounds for confidence that an entity meets its security objectives
- **Security Objective** [41] (mentioned in [32]) - Statement of intent to counter identified threats and/or safety identified organization policies and assumptions
- **Information Systems (IS)** –
    - 1) [83] - Information Systems can be classified into three types:
        - Servers/mainframes: usually the most physically secure class of systems
        - Workstations: usually located in more open or accessible areas of a facility
        - Portable devices: can be an organization's security nightmare
    - 2) [50] Information System (a.k.a: Automated Information System, Information Technology System) - Any equipment or interconnected system or subsystem of equipment that is used in the automatic acquisition, storage, manipulation, management, movement, control, display, switching, interchange, transmission or reception of data and includes computer software, firmware, and hardware
- **Safeguard** [40] (mentioned in [32]) - Practice, procedure, or mechanism that reduces risk
- **Hypothesis** [76] - A statistical hypothesis is a statement about the value of a population parameter

- **Null Hypothesis** [76] - The hypothesis against which we hope to gather evidence is called the null hypothesis, and is denoted by $H_0$
- **Alternative Hypothesis** [76] - The hypothesis for which we wish to gather supporting evidence is called the alternative hypothesis, and is denoted by $H_a$
- **Significance Level** [76] - The probability, $\alpha$, of making a Type I error is called level of significance (or significance level) for a hypothesis test.
- **Outlier** [77] - An unusual observation that lies outside the range of the data values we want to describe
- **Multicollinearity** [77] - When the independent variables in a multiple regression analysis exhibit a high degree of correlation, multicollinearity exist.
- **Correlation** [53] - a measure of the extent to which two variables are linearly related
- **Tolerance** [53] - the tolerance of an independent variable is the extent to which that independent variable cannot be predicted by the other independent variables
- **Linear regression model** [10] - Regression model that all parameters enter the equation linearly, possibly after transformation of data
- **Simple linear regression model** [10] - Regression model that only one predictor (independent) variable exist
- **Multiple linear regression model** [10] - Regression model that two or more predictor (independent) variables exist

**Teamwork Quality*** [33]

| | |
|---|---|
| Communication | There was frequent communication within the team. |
| | The team members communicated often in spontaneous meetings, phone conversations, etc. |
| | The team members communicated mostly directly and personally with each other. |
| | There were mediators through whom much communication was conducted. [R] |

| | |
|---|---|
| | Project-relevant information was shared openly by all team members. |
| | Important information was kept away from other team members in certain situations. [R] |
| | In our team there were conflicts regarding the openness of the information flow. [R] |
| | The team members were happy with the timeliness in which they received information from other team members. |
| | The team members were happy with the precision of the information received from other team members. |
| | The team members were happy with the usefulness of the information received from other team members. |
| Coordination | The work done on subtasks within the project was closely harmonized. |
| | There were clear and fully comprehended goals for subtasks within our team. |
| | The goals for subtasks were accepted by all team members. |
| | There were conflicting interests in our team regarding subtasks/subgoals. [R] |
| Balance of member contributions | The team recognized the specific potentials (strengths and weaknesses) of individual team members. |
| | The team members were contributing to the achievement of the team's goals in accordance with their specific potential. |
| | Imbalance of member contributions caused conflicts in our team. [R] |
| Mutual support | The team members helped and supported each other as best they could. |
| | If conflicts came up, they were easily and quickly resolved. |
| | Discussions and controversies were conducted constructively. |

| | |
|---|---|
| | Suggestions and contributions of team members were respected. |
| | Suggestions and contributions of team members were discussed and further developed. |
| | Our team was able to reach consensus regarding important issues. |
| Effort | Every team member fully pushed the project. |
| | Every team member gave the project the highest priority. |
| | Our team put much effort into the project. |
| | There were conflicts regarding the effort that team members put into the project. [R] |
| Cohesion | It was important to the members of our team to be part of this project. |
| | The team did not see anything special in this project. [R] |
| | The team members were strongly attached to this project. |
| | The project was important to our team. |
| | All members were fully integrated in our team. |
| | There were many personal conflicts in our team. [R] |
| | There was personal attraction between the members of our team. |
| | Our team was sticking together. |
| | The members of our team felt proud to be part of the team. |
| | Every team member felt responsible for maintaining and protecting the team. |

[R] = reverse coded item

# APPENDIX D
# DATA FORMS

## 1. Background Form

Project Pre-Questionnaire Form

This questionnaire will ask you to provide some information so that we can better understand how much knowledge or capability you have as a background. Please answer each item as completely as possible. You will fill out this questionnaire only once.

Date: ____/___/___ Code Number: _____

Team Name: _____ Role: _____

Briefly describe your project (experiment):

Development (Working) Schedule: When is the starting date and ending date of your work?

a. Starting Date: ___ /___ /___
b. Ending Date: ___/___/___

1. Computer Languages: In the blanks below indicate the languages you are familiar with and your level of experience with them. When indicating your level of experience, please indicate the length of time in months that you have actively worked with the language.

| **Language** | **Level of experience, where 1 = novice (≤ 2 months); 2 = some experience (≤ 6 months); 3 = moderate experience (≤ 1 year); 4 = a good deal of experience (≤ 3 years); 5 = extensive experience (≥ 3 years)** | **Months** |
|---|---|---|
| Ada(83/95) | | |
| Smalltalk | | |
| Modula-2 | | |
| Modula-3 | | |
| Assembly | | |
| Basic | | |
| COBOL | | |

| | | |
|---|---|---|
| Fortran (77/95) | | |
| Lisp | | |
| Pascal | | |
| Prolog | | |
| C | | |
| C++ | | |
| Visual C++ | | |
| Visual Basic | | |
| VBScript | | |
| C# | | |
| CGI | | |
| Perl | | |
| Unix shell (c/korn/borne/tc/etc.) | | |
| Java | | |
| JavaScript | | |
| Java Servlet | | |
| Python | | |
| Tcl/Tk | | |
| UML | | |
| HTML | | |
| XML | | |
| ASP | | |
| JSP | | |
| PHP | | |
| (Others) | | |
| (Others) | | |
| (Others) | | |
| (Others) | | |
| (Others) | | |

2. Applications: In the blanks below indicate the applications you are familiar with and your level of experience with them. For the purposes of this questionnaire, we will define application (or application program) as a program designed to perform a specific function directly for the user or, in some cases, for another application program. Examples of applications include database programs, web browsers, development tools, and communication programs. Applications use the services of the computer's OS and other supporting applications. When indicating your level of experience, please indicate the length of time in months that you have actively worked with the application.

| | **Level of experience, where 1 = novice (≤ 2** | |
|---|---|---|

| Application | **months);<br>2 = some experience (≤ 6 months); 3 = moderate experience (≤ 1 year); 4 = a good deal of experience (≤ 3 years); 5 = extensive experience (≥ 3 years)** | Months |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

3. Systems: In the blanks below indicate the systems you have worked with and your level of experience with them. For the purposes of this study we will define system as comprised of hardware components that have been carefully chosen so that they work together and an OS that manage and provide services to other programs that can be run in the computer. Examples of systems are UNIX, LINUX, Microsoft Windows (95/98/2000/XP and NT4.0), Sun Solaris 7.0/8.0, etc, as an OS and server or router as a hardware component. When indicating your level of experience, please indicate the length of time in months that you have actively worked with the system

| System | **Level of experience, where 1 = novice (≤ 2 months);<br>2 = some experience (≤ 6 months); 3 = moderate experience (≤ 1 year); 4 = a good deal of experience (≤ 3 years); 5 = extensive experience (≥ 3 years)** | Months |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

4. Training: In the blanks below please indicate the training you have had and when you got it. Include training obtained both off-line (through class, seminars, conferences, telephone) and online (teleconference, e-mail exchange, internet). Please mark in the 'Security' field if the training you had is related to the security area.

| **Training** | **Security (S)** | **Duration (From mo/day/yr To mo/day/yr)** |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

5. Other Security Experience & Knowledge: Please describe any other experience or knowledge you have related to the security area. If you have experience in specific tools, please indicate them. You may include experience or knowledge obtained from self-study.

| **Category** | **Name** | **How many days?** |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

6. Experience working on teams: How much experience have you had working in teams?

___ I have seldom (≤ 2 months) worked in teams in the past
___ I have done some (≤ 6 months) work in teams in the past
___ I have done a good deal of (≤ 1 year) work in teams in the past
___ I have worked in teams a lot (≤ 3 years) in the past
___ I have extensive (≥ 3 years) experience working in teams

7. Previous work with your team: Have you had previous experience with one of your team members? Please indicate who they are and describe the nature of the work.

Name: 1. _____
2. _____
3. _____

Nature of the work:
_____

_____

_____

8. To what degree do you understand your project? (Choose one)

___ I have little understanding of my project [Vaguely understand project goal and importance; don't have a 'big picture'; don't know what methods (algorithms/languages/applications/systems) may be used]

___ I have a basic understanding of my project [Somewhat understand project goal and importance; have a partial 'big picture'; know roughly what methods (algorithms/languages/applications/systems] may be used)

___ I have a good understanding of my project [Roughly understand project goal and importance; roughly have a 'big picture'; know what methods (algorithms/languages/applications/systems) may be used]

___ I have considerable understanding of my project [Understand project goal and importance; have a clear 'big picture'; know what methods (algorithms/languages/applications/systems) may be used and how they are going to be used for the project; know the constraints and assumptions a little]

___ I have thorough understanding of my project [Very clearly understand project goal and importance; have a very clear 'big picture'; exactly know what methods (algorithms/languages/applications/systems) may be used and how they are going to be used for the project; know what are the constraints and assumptions]

**2. Preparation Form**

Preparation Form

Please record any activity that you took to prepare for the exercise or to prepare defenses against attacks.

1. Team Name:
2. Code Number:
3. Date:
4. Time from: (am/pm) Time to: (am/pm)
5. Activity:
6. Notes:
7. Important Points:

8. Please indicate the sources of information you received during this activity and the type (e.g., email) of activity it pertained to. If you received more than one message of the same type, please indicate the number of messages you received.

   Examples of sources might be "/etc/inetd.conf", "snort log", and so forth. For the type, please type in 'E' for email, 'P' for phone calls, 'H' for hearing from team members (off-line), 'D' for discussion (off-line), 'B' for Electronic Bulletin Board (ex. Yahoo Messenger), 'O' for other type (in case of 'O', please indicate the type).

| Sources | Type | Number of messages |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

**3. Activity Record Form**

Activity Record Form

For each activity you undertake during the exercise, please fill out a record form. Activities include cases where you have detected or discovered a problem or event, worked toward a solution for a problem or event, or any other type of action you took during the exercise that was not one of the previous two types.

1. Team Name:

2. Code Number:

3. Date:

4. Time from: (am/pm) Time to: (am/pm)

5. Problem/Event Detection

    a. At what specific time did you find the problem?

    b. What was the nature of the problem? Please describe in detail.

    c. How did you find the problem?

    d. Is the problem suspicious of any type of hacking (security attack)? (Yes/No/Don't know)

    e. If yes to 'b', please tell us the type and name of hacking (security attack) if you know:
        i. Type: _____
        ii. Name: _____

    f. If no to 'b', why do you think so?

6. Problem/Event Solving

    a. What actions did you take?

    b. When and at what time did you take the actions?
        i. Date: ____/_____/____
        ii. Time: _____:_____ (am/pm)

    c. Why did you take the actions?

7. Other activity: Please describe any other activity that was not mentioned in ‘problem/event detection’ or ‘problem/event solving’ that was significant during this period.

8. Please indicate the sources of information you received during this activity and the type (e.g., email) of activity it pertained to. If you received more than one message of the same type, please indicate the number of messages you received. Please also indicate which phase you were in.

   Examples of sources might be “/etc/inetd.conf”, “snort log”, and so forth. For the type, please type in ‘E’ for email, ‘P’ for phone calls, ‘H’ for hearing from team members (off-line), ‘D’ for discussion (off-line), ‘B’ for Electronic Bulletin Board (ex. Yahoo Messenger), ‘O’ for other type (in case of ‘O’, please indicate the type). For the phase, please type in ‘P’ for the prevention phase, ‘D’ for detection, ‘R’ or response.

| Sources | Type | Number of messages | Phase |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

9. Downtime

   Please indicate the downtime you had during the experiment. Please fill in the time duration ([hh:mm] - [hh:mm]) that the downtime occurred, and the reasons (Why?).

| *No.* | *Server* | *System (Computer)* | *Service* | *Network* | *Application* | *Others* | *Why?* |
|---|---|---|---|---|---|---|---|
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |

**4. Project Teamwork Form**

Project Teamwork Form*

This questionnaire will ask you to provide some information so that we can better understand how the incidence response teams work. Please answer each item as completely as possible. You will fill out the questionnaire several times as requested throughout the exercise.

Date: ____/___/___ Code Number: _____

Team Name: _____ Role: _____

Start Time [hh:mm]: _____ End Time [hh:mm]: _____

1. How surprised were you by the attacks you have received since you last filled out this questionnaire?

Not surprised at all: ___: ___: ___: ___: ___: ___: ___: ___: ___: ___: Extremely surprised

2. Please indicate your beliefs about the following statements that could be applied to your team for the period since you last filled out this questionnaire.

| | *To no extent* | *To a limited extent* | *To some extent* | *To a considerable extent* | *To a great extent* |
|---|---|---|---|---|---|
| This group has confidence in itself | | | | | |
| This group believes it can become unusually good at producing high-quality work. | | | | | |
| This group expects to be known as a high-performing team. | | | | | |
| This group believes it can solve any problem it encounters. | | | | | |
| This group believes it can be very productive. | | | | | |
| This group can get a lot done when it works hard. | | | | | |
| No task is too tough for this group. | | | | | |
| This group expects to have a lot of influence around here. | | | | | |

3. Please place a check in the box of the response ("Yes", "No", or "?") that indicates whether each of the following words or phrases describes your feeling about your work during the last work period (since you last filled out these scales.)

| | *Yes* | *No* | *? (Don't know)* |
|---|---|---|---|
| Demanding | | | |
| Pressured | | | |
| Hectic | | | |
| Calm | | | |
| Relaxed | | | |
| Many things stressful | | | |
| Pushed | | | |
| Irritating | | | |
| Under control | | | |
| Nerve-wracking | | | |
| Hassled | | | |
| Comfortable | | | |
| More stressful than I'd like | | | |
| Smooth running | | | |
| Overwhelming | | | |

4. The following questions ask you to describe how your work group works together as it makes decisions or solves problems. Please indicate your degree of agreement with each of the following statements as they apply to your team since the last time you filled out this questionnaire.

| | ***Strongly Disagree*** | ***Disagree*** | ***Somewhat Disagree*** | ***Somewhat Agree*** | ***Agree*** | ***Strongly Agree*** |
|---|---|---|---|---|---|---|
| This group weighs all the potential effects of all possible options or solutions carefully. | | | | | | |
| This group carefully considers possible negative consequences of options or solutions. | | | | | | |
| This group does not capitalize on the wisdom and experience of all members when making decisions or solving problems. | | | | | | |
| This group thoroughly diagnoses the problems it faces. | | | | | | |
| In group decisions, key issues are neglected or not fully considered. | | | | | | |
| This group carefully considers questions and issues when they run counter to the general consensus. | | | | | | |
| This group conducts a broad search for information about the problem or decision. | | | | | | |
| After making a decision, this group often stops to reexamine it one more time to make sure it is making the right choice. | | | | | | |
| This group makes careful plans for implementing its decisions or problem solutions. | | | | | | |
| If new, relevant information comes up, this group considers it carefully, even it already has closure on the decision, problem, or solution. | | | | | | |

5. Please indicate your degree of agreement with these statements about communication within your team since the last time you filled out this questionnaire.

| | ***Strongly Disagree*** | ***Disagree*** | ***Somewhat Disagree*** | ***Somewhat Agree*** | ***Agree*** | ***Strongly Agree*** |
|---|---|---|---|---|---|---|
| There was frequent communication within the team. | | | | | | |
| There was intensive communication within our team | | | | | | |
| Important information was kept away from other team members in certain situations. | | | | | | |
| The team members were happy with the timeliness in which they received information from other team members. | | | | | | |
| The team members were happy with the accuracy of the information received from other team members. | | | | | | |
| The team members were happy with the usefulness of the information received from other team members. | | | | | | |

6. Please indicate your degree of agreement with each of the following statements about how your team worked since the last time you filled out this questionnaire.

| | ***Strongly Disagree*** | ***Disagree*** | ***Somewhat Disagree*** | ***Somewhat Agree*** | ***Agree*** | ***Strongly Agree*** |
|---|---|---|---|---|---|---|
| The work done on sub-tasks within the project was closely harmonized. | | | | | | |
| There were clear and fully comprehended goals for sub-tasks within our team. | | | | | | |
| Our team avoided duplication of effort | | | | | | |
| There were conflicting interests in our team regarding sub-tasks/sub-goals. | | | | | | |

7. Please indicate your degree of agreement with each of the following statements about your team since the last time you filled out this questionnaire.

| | ***Strongly Disagree*** | ***Disagree*** | ***Somewhat Disagree*** | ***Somewhat Agree*** | ***Agree*** | ***Strongly Agree*** |
|---|---|---|---|---|---|---|
| It was important to the members of our team to be part of this project. | | | | | | |
| The team did not see anything special in this project. | | | | | | |
| The team members were strongly attached to this project. | | | | | | |
| The project was important to our team. | | | | | | |
| All members were fully integrated in our team. | | | | | | |
| There were many personal conflicts in our team. | | | | | | |
| There was personal attraction between the members of our team. | | | | | | |
| Our team was sticking together. | | | | | | |
| All team members were equally engaged to achieve common goals | | | | | | |
| All members were fully contributing to our team | | | | | | |

8. Please indicate your degree of agreement with the following statements about your team's work since the last time you filled out this questionnaire.

| | ***Strongly Disagree*** | ***Disagree*** | ***Somewhat Disagree*** | ***Somewhat Agree*** | ***Agree*** | ***Strongly Agree*** |
|---|---|---|---|---|---|---|
| Every team member fully pushed the project. | | | | | | |
| Every team member gave the project the highest priority. | | | | | | |
| Every team member felt fully responsible for team goals | | | | | | |
| There were conflicts regarding the effort that team members put into the project. | | | | | | |

9. Please indicate your degree of agreement with each of the following statements about your team since the last time you filled out this questionnaire.

| | ***Strongly Disagree*** | ***Disagree*** | ***Somewhat Disagree*** | ***Somewhat Agree*** | ***Agree*** | ***Strongly Agree*** |
|---|---|---|---|---|---|---|
| There was a cooperative work atmosphere in our team | | | | | | |
| Discussions and controversies were conducted constructively. | | | | | | |
| Suggestions and contributions of team members were respected. | | | | | | |
| Suggestions and contributions of team members were discussed and further developed. | | | | | | |
| Our team was able to reach consensus regarding important issues. | | | | | | |

10. Please indicate your degree of agreement with each of the following statements about your team since the last time you filled out this questionnaire.

| | *Strongly Disagree* | *Disagree* | *Somewhat Disagree* | *Somewhat Agree* | *Agree* | *Strongly Agree* |
|---|---|---|---|---|---|---|
| Our team recognized the specific potentials (strengths and weaknesses) of individual team members. | | | | | | |
| Every team member was contributing to the achievement of the team's goals in accordance to their specific potentials. | | | | | | |
| Imbalance of member contributions caused conflicts in our team. | | | | | | |

11. Place a mark on each scale that represents the magnitude of each factor during the last work period (since you last filled out these scales).

Mental demand: How many mental and perceptual activities were required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the work easy or demanding, simple or complex, exacting or forgiving?

Low: ___: ___: ___: ___: ___: ___: ___: ___: ___: ___: High

Physical demand: How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the work easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Low: ___: ___: ___: ___: ___: ___: ___: ___: ___: ___: High

Temporal demand: How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

Low: ___: ___: ___: ___: ___: ___: ___: ___: ___: ___: High

Performance: How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

Poor: ___: ___: ___: ___: ___: ___: ___: ___: ___: ___: Excellent

Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance?

Low: ___: ___: ___: ___: ___: ___: ___: ___: ___: ___: High

Frustration level: How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the work?

Low: ___: ___: ___: ___: ___: ___: ___: ___: ___: ___: High

*: 2004: #10 was omitted.

**5. Post-Defense (Analysis) Form**

Post-Defense (Analysis) Form

Please indicate anything you learned through this exercise including final thoughts after defenses against attacks.

1. Team Name:
2. Code Number:
3. Date:
4. Time from: (am/pm) Time to: (am/pm)
5. Final Thoughts & Lessons Learned:

6. Please indicate the sources of information you received during this activity and the type (e.g., email) of activity it pertained to. If you received more than one message of the same type, please indicate the number of messages you received.

   Examples of sources might be "/etc/inetd.conf", "snort log", and so forth. For the type, please type in 'E' for email, 'P' for phone calls, 'H' for hearing from team members (off-line), 'D' for discussion (off-line), 'B' for Electronic Bulletin Board (ex. Yahoo Messenger), 'O' for other type (in case of 'O', please indicate the type).

| Sources | Type | Number of messages |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

# VITA


| | |
|---|---|
| Name: | Sung-Oh Jung |
| Address: | Dae-Rim Apartment 1204-302<br>Jeong-Wang Dong<br>Siheung, Kyung-Ki Do<br>429-762<br>South Korea |
| Email: | jungs@cs.tamu.edu |


Education:

B.S., Dankook University, Computer Science & Statistics, 1996

M.S., University of Southern California, Computer Science, 1999

Ph.D., Texas A&M University, Computer Science, 2005