Accounting for the Human User in Predictive Security Models

Mohammad A. Noureddine*, Andrew Marturano*,Ken Keefe*, Masooda Bashir†, and William H. Sanders ‡
* *Information Trust Institute*
†*School of Information Sciences*
‡*Department of Electrical and Computer Engineering*
University of Illinois at Urbana-Champaign
Urbana, Illinois, USA
{*nouredd2, martura2, kjkeefe, mnb, whs*}*@illinois.edu*

***Abstract*—Given the growing sophistication of cyber attacks, designing a perfectly secure system is not generally possible. Quantitative security metrics are thus needed to measure and compare the relative security of proposed security designs and policies. Since the investigation of security breaches has shown a strong impact of human errors, ignoring the human user in computing these metrics can lead to misleading results. Despite this, and although security researchers have long observed the impact of human behavior on system security, few improvements have been made in designing systems that are resilient to the uncertainties in how humans interact with a cyber system. In this work, we develop an approach for including models of user behavior, emanating from the fields of social sciences and psychology, in the modeling of systems intended to be secure. We then illustrate how one of these models, namely general deterrence theory, can be used to study the effectiveness of the password security requirements policy and the frequency of security audits in a typical organization. Finally, we discuss the many challenges that arise when adopting such a modeling approach, and then present our recommendations for future work.**

***Index Terms*—Human factors, computer simulation, modeling, computer security, computer crime.**

1. Introduction

Users are often assumed to be the "weakest link" in the information security landscape. Security designers and policy makers often incorrectly assume that users have perfect compliance. The recent IBM Security Services 2014 Cyber Security Intelligence Index [1] revealed that 95% of the investigated incidents largely involved human error. Users make mistakes, choose weak passwords, write down their passwords, and more importantly, they intentionally circumvent security policies. Documents leaked from the recent Sony hack showed that employees frequently used plain text documents to save lists of user-names and passwords for various personal and company-owned accounts [2]. Additionally, system administrators often misconfigure security settings, choose to skip installing a patch, or leave default database passwords unchanged.

Meanwhile, the widespread use of computing devices has continued to increase at an incredible pace. The introduction of the Internet of Things poses many security and privacy challenges [3]. However, usable security software and effective security policy compliance continues to be a failure. Users had, and still have, wrong perceptions about security and privacy, and still underestimate risks related to information security.

It has been long realized that security mechanisms should not make the completion of a task by a human user harder than it would be if the mechanisms were not present, a concept referred to as *psychological acceptability* [4]. However, today's security solutions are still far away from achieving usability. They are either too complex to be used by inexperienced users or reduce a system's availability. Eventually, such solutions will force users to circumvent designed controls, thus increasing the security vulnerabilities. In their study of password usage in corporate environments, Beautement et al. identified several usability issues with the currently employed password policies [5]. Users are usually demotivated by the increasing number of passwords that they must keep track off. They therefore circumvent in order to avoid strict password requirements and come up with passwords that are easy to remember. For example, a common circumvention is to add a numerical character at the end of a previously used password when generating a new one. Additionally, users are often not well informed about information sensitivity and security issues. Thus, They form their own perceptions which are mostly inconsistent with reality. This behavior will lead them to making unintentionally bad security decisions.

Reasoning under risk and uncertainty is an utterly difficult task for human beings. Studies in the fields of social sciences and economics suggest that humans are "less than optimal" decision makers when reasoning under risk and uncertainty [6]. When it comes to cyber security, users do not often consider themselves at risk, and perceive the losses incurred by performing a *pro-security* action to be much larger than the benefits gained. A pro-security behavior provides some safety benefit, i.e. a reduced chance of something bad happening. Such a benefit is often neither tangible nor immediately observable. Add to that the fact that users

978-1-5090-5652-1/17/$31.00

generally have multiple other tasks that they must perform, tasks that require their attention and effort. Therefore they are more likely to prefer actions that save time and effort while accepting a certain level of risk.

Complementary to usable security studies, addressing the human compliance problem requires the integration of human decision making considerations into the design and testing processes of security systems. Bishop [4] argues that in order to produce psychologically acceptable security mechanisms, designers and policy makers need to test their mechanisms in the environments in which they will deployed. Since testing in real world scenarios with actual human participants is neither practical nor scalable, simulation would be the next resort.

In Section 2, we provide an examination of past work and argue for the need to correctly model human behavior. In Section 3, we review some of the most prominent behavioral theories that researchers and scientists have used to explain the human behavior in cyber security. We draw these theories from the fields of psychology and social sciences, and envision that they can be used for building accurate models of human users in cyber security. In Section 4, we then illustrate how the general deterrence theory [7] can be used in a stochastic simulation model in order to obtain quantitative security metrics. These metrics related to passwords policy and security audits frequency in a typical organization. We follow an approach similar to that presented in [8] and [9] in defining utility functions for the employees and the security of the organization, and modeling the decision making process as a willingness probability that we vary across our simulation. In Section 5, we conclude our work with insights obtained from our case study, and we highlight the tasks we deem challenging when building models of human users and incorporating them into modeling and simulation tools. Finally, in Section 6, we present our ideas for interesting future work.

2. Motivation

System and policy designers have acknowledged the fact that the misbehavior of human users and administrators can have devastating effects on the security of a system. Humans, although not perfectly rational, are smart, devise shortcuts, use heuristics, and can always find ways to circumvent designed controls seeking quick short-term outcomes [10]. Additionally, when a person's perception of the security state is not aligned with the real state in which a system can be, she can have unrealistic expectations and can thus exaggerate or downgrade certain security risks [11].

The study in [12] highlights the user's general lack of the necessary computer security knowledge for proper usage of computer systems. Such knowledge is deemed important for pro-security measures such as the creation and maintenance of passwords across multiple user accounts. The authors identify the number of passwords to maintain, the complexity of the password policy, the perceived compatibility of the policy with work practices, as well as the personal perception of security and information sensitivity, all as important factors affecting a user's effective creation and maintenance of passwords. The authors also note that the level of communication between the security departments and the users in organizations is inadequate. System and policy designers need to treat users as partners, as opposed to enemies, in the development of effective security measures. This highlights the need for measures and metrics that allow us to study the usability and security of system and policy designs, while taking the user's behavior into consideration.

The work in [4] affirms the findings of [12] by revisiting the concept of psychological acceptability. Adhering to this concept requires the system or policy designer to take into account the knowledge, ability, and security mental models of the system or policy users. The author draws the conclusion that developing psychologically acceptable security mechanisms depends upon the context in which the mechanisms are to be used. This gives rise to a need for testing security mechanisms by placing them in the environments in which they will be used, and analyzing the different ways in which different users practice them. Since testing in real environments with actual users is not efficient nor scalable, designers and security managers must resort to simulation. Obtaining quantitative metrics can help policy and system designers compare different policies and designs, and make informed decisions about which ones to adopt [13] . This highlights the need for the development of models of human users to be incorporated into the design and modeling tools.

Furthermore, complementary to the usability problem, the work in [6] highlights the importance of a user's risk assessment. People usually find it very difficult to estimate concepts such as risk and uncertainty. Users do not typically consider themselves at risk, and wrongly assume that firms handling sensitive data must have done enough to protect that data, regardless of their actions. Additionally, security, much like safety, is an abstract concept. It is concerned with ensuring that nothing bad will happen, which makes it harder for a user to appreciate that benefit when faced with a more immediate reward. The author finally recommends that rewarding pro-security behavior and sanctioning risky actions is essential to ensure user compliance with security policies and procedures.

These realizations motivated work towards building predictive security models that include system, adversary, as well as human user models. The work in [8] draws an analogy from the economics models in the central bank problem [14] in order to provide quantitative evidence of the usability and confidentiality trade-offs in cyber security systems. Based on the analogy, the authors propose numerical formulas that relate availability to confidentiality to level of security investment in a given firm. The security investment is budgeted between training, IT support, and monitoring. They then perform an empirical simulation of a USB usage policy in a given firm, in which human users compute local utility functions and choose accordingly whether to comply to the policy and encrypt the USB data, or to ignore the policy and favor usability.

In [15], the authors propose an information security ontology that security analysts can use to understand the implications of the human behavior, and consider alteration to the security mechanisms based on such implications. The ontology includes best practices guidelines for security, usually defined in global standards, and the human-behavioral implications associated with the application of these guidelines. In the ontology, human behavior effects are always considered as vulnerabilities to the system. Such vulnerabilities offer a chance for attackers to exploit them and threaten the security of the system.

The authors in [16] study the trade-offs between security and unavailability in enterprise information security technologies. They review several centralized information security mechanisms currently employed in organizations, such as USB access control management, disk encryption solutions and access control methods. They then use simulation to show the effects that such mechanisms can have on the availability of the systems, and thus on the productivity on the workers. They show that with current information security solutions, as the number of the employees in an organization increases, the productivity loss incurred from such security mechanisms can reach high levels; the employees are spending their time performing security mechanisms and waiting for IT support rather than spending it on daily productive tasks.

Although these techniques are attempting to provide quantitative security metrics using predictive models of human users, they often lack a strong background in psychology and social sciences. Theories developed in these fields provide descriptive models of the ways in which human users perceive computer security and make security related decisions. It is important to include such theories in the development of predictive human models as they can provide theoretical basis upon which mathematical models can be built. Factors such as risk assessment, security culture, organizational culture, workload, experience, and training must take part of any accurate predictive model.

In this paper, we present an overview of the most prominent behavioral theories that researchers and scientists have used to understand the human behavior in the field of computer security. We then illustrate how the general deterrence theory can be used to develop a predictive model of the employees in a typical organization, and present the security metrics that we can obtain from it. For our review of psychological theories of human behavior in computer security, we chose papers that have a prominent basis in both psychology and computer science, and include empirical studies to validate their findings and speculations.

3. Theories of Human Behavior

In this section, we summarize the most commonly used behavioral theories that researchers and scientists have used to describe human behavior in cyber security. We draw these theories from the fields of psychology and social sciences, and focus on the ones that admit a strong empirical evidence in terms of user studies and interviews.

3.1. The General Deterrence Theory

The general deterrence theory focuses on the effects of sanctions against committing a criminal act, where sanctions are formal punishments for failing to follow established security policy [7]. The theorys main aspects are the certainty of sanctions, or the probability of being punished for violating a security policy, and the severity of sanctions or the degree of punishment associated with the act.

Commonly associated with this theory is the security action cycle, which provides a model for handling computer abuse [17]. This model is composed of four stages. The first is *deterrence*, where most of the potential offenders are deterred through mechanisms such as policies, guidelines, and awareness programs. *Prevention* is employed when deterrence fails, mainly through physical or procedural controls. Third is *detection* where mechanisms are designed to address the realization of computer abuse, aiming to make such abuses known. Finally, *remedies* are employed when a computer abuse act is detected, where specific actions are taken against offenders, in accordance with the organization's policies and rules.

The work in [7] attempts to provide an empirical validation for this theory. The authors observe that employees were less likely to misuse information security policy if their awareness of the punishment was high. However, the study in [18] found that the severity of sanctions had a negative impact on employee security compliance, instead of a positive one. One interpretation suggested for this finding was that employees are not as affected by severity of sanctions if they have not been sanctioned before. For example, the punishment for information security misuse may be severe, but an employee will perceive that it is more likely to happen to his/her peers than to him/her. This is supported from a psychological standpoint, following established theories of risk perception [19].

3.2. The Theory of Planned Behavior

The theory of planned behavior suggests that intentions are an important factor for predicting behavior. This is most affected by the person's attitude toward the behavior, social factors, and control factors [20]. A person's attitude towards behavior is the degree to which she does or does not favor a certain behavior. If she perceives the result of a behavior as positive, she will shape a positive attitude towards it. Social Factors, or subjective norms, are normative beliefs concerning a behavior. In order for a person to adopt a behavior, she must first be motivated to comply with its social demands. Control Factors, or perceived behavioral control, means that a person shapes their intentions in regard to a particular behavior based on her personal beliefs.

The work in [21] provide empirical evidence validating that if employees are aware of any benefits (intrinsic or extrinsic) to following information security policy, they are more likely to comply. This supports the attitude towards behavior factor of the theory. Additionally, the authors found

that security awareness and compliance were positively related to intrinsic costs such as guilt or embarrassment. Such factors are examples of social demands, giving validation to the second part of the theory, social factors. Furthermore, the work in [22] found that intrinsic factors such as perceived legitimacy of policy, as well as perceived value congruence were positively related to employee security compliance. This reflects the impact of the control factors section of the theory.

3.3. The Protection Motivation Theory

The theory of protection motivation adds another factor to the aforementioned theory of planned behavior, and asserts that motivation emanates from not only the threat appraisal (certainty or severity of sanction), but from the coping appraisal as well. Coping appraisal is defined as an individuals assessment of her ability to cope with and avert the potential loss or damage arising from a threat [23]. This breaks down into three factors: self-efficacy, response efficacy, and response cost. Self-efficacy is an individuals ability or judgment regarding his or her capabilities to cope with and avert the potential loss or damage arising from the threat. Response efficacy is the compliance with the information security policy as being an effective mechanism for detecting threats. Response cost emphasizes the perceived opportunity costs in terms of monetary, time, or effort extended in adopting the recommended behavior.

The work in [20] found that self-efficacy, response cost, and response efficacy were positively related to compliance behaviors. This finding not only validates the theory but also shows that employee attitudes and perceptions, as well as their level of competence are important to encouraging security compliance. The findings in [24] echo the validation of the theory, and further emphasize the need for organizational education programs for security, in order to teach employees the benefits of pro-security behavior.

3.4. The Social Bond Theory

The social bond theory suggests that social bonds can deter a person from committing a crime, regardless of their inclination toward it. This theory is common in the field of criminology, and seeks to explain social behavior that does not conform to social rules. The work in [25] shows that strong social bonds negatively affect intentions toward computer abuse. The theory shows that bonds of involvement can strengthen beliefs that computer abuse is socially unacceptable and can deter individuals from abusing security policy. However, there is also a converse effect: when social bonds are weak, it may lead to increased inclination to break established computer security policy.

3.5. The Social Learning Theory

Similar to the social bond theory, the social learning theory suggests that an individuals perception towards committing a crime can be affected by the actions and beliefs

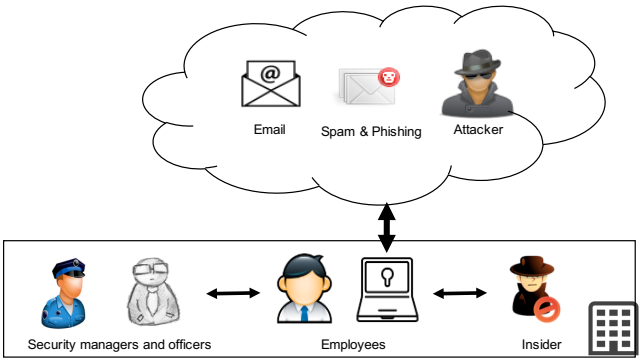


Figure 1: Top-level diagram of our case study

of those they associate with, even if the risk and punishment of the crime is high. If an individual associates with delinquent peers, her own sense of delinquency may increase [18]. The theory identifies four constructs that can influence the delinquency of an individual. (1) *differential association* is the process during which a person is exposed to normative definitions that favor or are against criminal behavior. (2) *differential reinforcement/punishment* is the balance of expected or realized reward and punishment resulting from criminal behavior. (3) *definition of behavior* is the rationalization of a behavior as good or bad, right or wrong, desirable or not, justified or not. (4) *imitation* is the engagement in behavior after observing similar behavior from others.

Studies on the impact of the social learning theory in organizations have found that computer abuse by peers, as well as those in upper management have a positive relation with individual motivation to engage in computer abuse. This effect is especially strong when individuals have negative perceptions of the organization, regardless of the level of existing security measures and policy [18].

4. Case Study

In this section, we illustrate the importance of introducing human user models on the predictive analysis of cybersecurity designs. We present a case study simulating the behavior of a customer service representative in a typical customer-based firm. We model the interactions between the employee and the firm's security policy using *Stochastic Activity Networks* (SANs) [26] and use the Möbius [27] modeling tool to obtain results. Our models are drawn from our discussions on the general deterrence theory and employ a similar approach to that in [9] to model the user's decision making process. We vary the employee's compliance probability and study its effects on the well-being of the employee as well as on the overall security of the firm.

4.1. Model Description

Figure 1 shows the top-level diagram of our case study. Employees use password protected accounts to access the

firm's workstations in order to receive and process customer e-mails. The firm employs security managers to design and maintain its security policy, and makes use of security officers to perform regular security audits. We assume the presence of two types of attackers. Outsiders use a variety of techniques to compromise the firm's security, while insiders attempt to exploit the employees' mistakes in order to obtain unauthorized access to the firm's workstations.

4.1.1. Threat and Attacker Models. We assume that the firm is trying to protect its information assets against two classes of attackers: insiders and outsiders. We assume the attackers are rational decision makers with sufficient intelligence and computational resources to pose a serious threat to the firm's information infrastructure.

Insiders attempt to acquire authentication measures to the firm's internal network by the means of stealing unprotected, written authentication information, basically user names and passwords. An employee writing their passwords on sticky notes next to their desk poses a security vulnerability that the insider attackers can make use of to infiltrate the firm's internal network. We model insider threats by the probability of a successful theft of authentication information, namely $\mathcal{P}_i$. This probability is proportional to the employee's willingness to write down their passwords; the more inclined employees are to write their passwords, the easier it is for insiders to obtain that information and thus threaten the firm's security assets.

Outsiders on the other hand, attempt to gain unauthorized access to the firm's network in order to achieve their malicious goals. We assume for simplicity that outsiders attempt two classes of attacks: (1) password cracking and (2) phishing. An attacker's ability to brute force or intelligently crack a password depends on the complexity of the firm's password requirements. The more complex the requirements are, the harder it gets for an attacker to break a password. We assume that attackers are well-resourced enough to be able to break any password (regardless of its complexity) with a given success probability $\mathcal{P}_{ap}$. The looser the password requirements are, the greater $\mathcal{P}_{ap}$ becomes.

Attackers attempt to steal the employees' credentials and fool them into downloading malicious software on their workstations using e-mail phishing techniques. A fraction of the employees' daily e-mails contains either phishing links or malicious attachments. The success of such attacks depends on the willingness of employees to provide their credentials to suspicious web forms, or download and open untrusted attachments on their workstations. We assume that $\mathcal{P}_{phish}$ and $\mathcal{P}_{mal}$ are the probabilities that an e-mail contains phishing links and malicious attachments, respectively.

4.1.2. User Models. For simplicity, we assume that customer service representatives are the only users in our case study. Additionally, we assume that the main means of communication for the firm is e-mail, and thus the representatives' tasks are to answer customer e-mails. In this section, we use the terms employee, user, and customer service representative interchangeably.

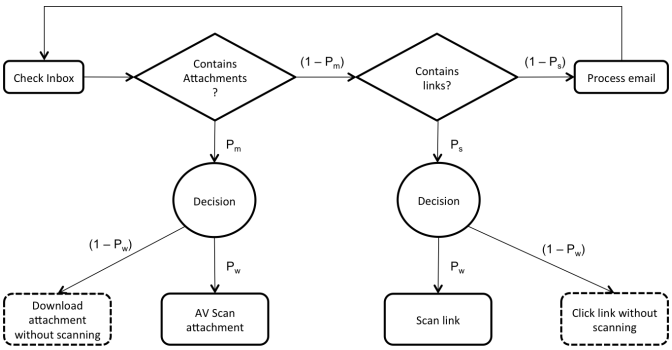


Figure 2: User's E-mail Tasks

Figure 2 shows the sub-tasks involved in the users' daily routine work. As per our assumption, the customer representative is responsible for answering customer questions via e-mail. We assume that e-mails arrive at the firm's address at an exponential rate with parameter λ_a. Once the employee receives the e-mail, she has to first check for the presence of attachments. The firm requires its employees to perform a virus scan on all e-mail attachments before opening them. However, the employee can either choose to adhere to the firm's policy or to ignore it and open the attachments directly. We denote by $\mathcal{P}_w$ as the *willingness probability*, or the compliance probability of any employee. We mark the state in which the employee has opened an attachment without scanning it as a *vulnerable state*, shown with dashed lines in Figure 2.

If the received e-mail is attachment free, the employee must then check for any links in the body of the message. Due to the fact that employees interact with both personal and work related e-mail, attackers can use phishing links to collect personal and work related information (such as authentication information). If such links are present, the firm requires its employees to run any links they receive via e-mail by a reputation scanner, and refrain from clicking on ones that do not meet the minimum safety (or reputation) guidelines. The employee in this case can again conform the firm's policy and run the check, or ignore the policy and directly open links. We mark the state in which the employee has clicked on a received link without scanning it to be a vulnerable state.

After the employee has performed (or ignored) the firm's procedure with regards to attachments and links (if any), she can process the e-mail and write a reply. We assume that the time needed for a typical employee to perform such a task is exponentially distributed with parameter λ_e.

4.1.3. Security Policy. As we previously discussed, the firm requires its employees to perform virus scans on all of the attachments received in e-mail. Similarly, it requires them to run links present in e-mail bodies through a reputation checker before going through them. In order to enforce these security requirements, the firm perform regular security audits and checks for any violations. If a violation is detected, sanctions are imposed on the violating employee.

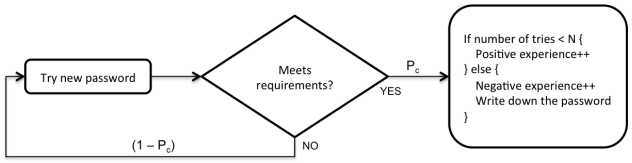


Figure 3: User's password change tasks

Additionally, the firm may choose to reward employees that have no violations on their record. We denote by $\mathcal{T}_a$ the frequency (in days) with which the firm decides to perform its security audits.

Furthermore, the firm's security managers can control the complexity requirements of the employees' passwords, which has a direct impact on the probability $\mathcal{P}_{ap}$. Additionally, the managers can control the frequency with which the employees are requested to change their passwords. We refer by $\mathcal{T}_p$ to that frequency (in days).

Every $\mathcal{T}_p$ days, the employees are requested to change their passwords. Figure 3 shows the tasks involved in an employee changing her password. We assume a very simplistic model in which the employee attempts a new password and checks whether it conforms with the password requirements, and keeps trying new ones until she succeeds. We denote by $\mathcal{P}_p$ the probability with which the employee is successful in generating a password that meets the minimum security requirements. The firm's security managers control $\mathcal{P}_p$ by setting the minimum requirements. The harder to requirements are, the smaller $\mathcal{P}_p$ is. We assume that if the employee can generate a successful password in $\mathcal{N}_p$ tries, she considers it to be a positive security experience. Otherwise, if it takes her more than $\mathcal{N}_p$ tries, she is inclined to write it down next to her office space, and considers its experience with the security policy to be negative. We refer to $\mathcal{N}_p$ by *the password write-down threshold*. We recognize that work such as [28] and [29] have proposed more accurate password usage and strength models, however, adopting such models into our simulation is beyond the scope of this paper and is left for future work.

4.2. Simulation Approach

In order to account for human decision making, we adopted an approach similar to the one presented in [9] where we simulate a human decision between two options (namely, to comply with the security policy, or not to comply) by the $\mathcal{P}_w$ probability. Varying this probability allows us to study the impact of the human decisions on the employee's happiness and the overall security of the firm.

We capture the firm's security utility (equation 1) as the fraction of insider and outsider attacks that are deemed unsuccessful. A more accurate representation of the security utility would account for the different goals of the firm's security policy, such as the data confidentiality, integrity and availability. We discuss this and other approaches in our discussion and future work sections.

$$\text{Security Utility} = \frac{\text{Failed attacks}}{\text{Total attacks}} \quad (1)$$

In order to measure the employee's level of satisfaction or "happiness," we use a utility function that is a linear combination of the employee's workload, her positive experiences with the security policy, her received sanctions and rewards, and her level of cognitive effort.

$$\text{Employee Utility} = -\alpha W + \beta E - \gamma S + \eta R - \epsilon C \quad (2)$$

Equation 2 shows the employee's utility function. We model *workload* (W) as the fraction of unanswered emails in the employee's mailbox. E is the fraction of positive security experiences that the employee encounters when doing security related tasks. S (R) represents the amount of sanctions (rewards) that the firm has imposed (rewarded) on the employee when performing their regular audits. Finally, C represents the employee's mental effort that she has to put in through her work day. This variable is increased every time the employee is required to try a new password, and increases as the work day progresses. We assume that at the end of a working day, W is reset to an initial value to model the employee resting after the end of the work day. Although our model of the employee's cognitive load is simple, our study of the literature indicates that the mental load is an important factor that affects human behavior and decision making [30]. We will investigate more accurate models of the cognitive and mental load in our future work.

$\alpha, \beta, \gamma, \eta$, and ϵ are scaling parameters that highlight the impact of each of the aforementioned factors has on the employee's utility. We assume that these parameters are inputs to the model that are to be determined by the security managers after conducting interviews and surveys of their employees. For our model, we use $\alpha = -0.5$, $\beta = 0.1$, $\gamma = -0.2$, $\eta = 0.1$, and $\epsilon = -0.1$. We adopted linear utility functions similar to the ones presented in [9] and [8]. This provides a simple yet meaningful way to capture the level of security and the employee's welfare. It also provides an convenient way to study how different factors affect the employee's utility as her willingness varies.

Our choice of the parameters in equation 2 is motivated by the general deterrence theory. For effective deterrence, the severity of the punishment must outweigh the benefits of the crime. We therefore chose to set $|\gamma| = 2|\eta|$ to increase the negative impact that sanctions have on the employee's utility, thus introducing a motivation for compliance and a deterrence from policy violation.

4.3. Results and Discussions

Table 1 shows the assignment for the different parameters that we used in our simulation. We validated our choice for parameters $\mathcal{P}_i$, $\mathcal{P}_{ap}$, $\mathcal{P}_{phish}$, and $\mathcal{P}_{mal}$ from recent security incident reports such as [1], [31], and [32]. The term *Variable* in Table 1 refers to a parameter that we vary during the simulation.

Figures 4a and 4b show the employee and the security utilities as a function of the employee's willingness proba-

TABLE 1: Möbius simulation parameters

Parameter	Exaplanation	Value
λ_a	Inter-arrival rate of e-mails	15.0 (e-mails/hour)
λ_e	Email processing rate	10.0 (e-mails/hour)
$\mathcal{P}_i$	Probability of an insider theft of authentication credentials	if (employee writes down password) 0.7 else 0.05
$\mathcal{P}_w$	Human compliance willingness probability	Variable
$\mathcal{P}_{ap}$	Probability of a successful password breaking attack	$\mathcal{P}_{ap} - 0.1$
$\mathcal{P}_{phish}$	Probability that an e-mail contains phishing links	0.01
$\mathcal{P}_{mal}$	Probability that an e-mail contains malicious attachments	0.005
$\mathcal{T}_a$	Frequency of security audits	Variable
$\mathcal{T}_p$	Frequency of password reset notices	30 (days)
$\mathcal{P}_c$	Probability to devise a valid password	Variable
$\mathcal{N}_p$	Password write-down threshold	Variable

bility $\mathcal{P}_w$ for different frequencies of security audits, respectively. We consider 4 cases: (1) no audits, (2) weekly audits, (3) bi-weekly audits, and (4) monthly audits. The results show, in accordance with the general deterrence theory, that as the frequency of the security audits increases, employee's are motivated towards compliance. This is shown from the fact that the employee's utility increases as the willingness to comply increases, achieving its highest value when full compliance is achieved (i.e. $\mathcal{P}_w \simeq 1$).

We first note that when no audits are employed, the employee's utility is at its highest since they can always choose not to comply with the security policy with no fear of retribution. Additionally, with no audits, the employees achieve maximum utility when $\mathcal{P}_w = 0$, i.e., when they ignore any security recommendations.

As the firm is considering more frequent security audits, the employees are more likely to comply in order to achieve higher utility values. We note that when the audits are the most frequent (weekly audits), the employee's utility increases at a faster rate compared to less frequent audits. According to the general deterrence theory, this reflects the benefits of having more frequent audits, and thus more frequent sanctions. The employees start to consider the negative effects of non-compliance and thus are more inclined to comply. When the audits frequency is relaxed (i.e. monthly audits), the employees are still compelled to comply but at a slower rate and with less increments in received benefit. As audits become more far apart, the employees can perform non-compliant behavior with lower risk of getting caught, thus making the pro-productivity choices more appealing.

Figure 4b shows that, as expected, the security utility increases with the increase in the willingness probability $\mathcal{P}_w$, i.e. when employees are more inclined towards complying with the security policies. The security utility also increases with the increase in the frequency of the security audits. The more the audits are frequent, the more violations of the policies are caught and dealt with, the more attacks are thwarted. We finally note that when no audits are considered, the benefit obtained from an increase in compliance is much lower than in all of the cases where audits are performed. This highlights the need to find a balance between the security utility and the employee's utility.

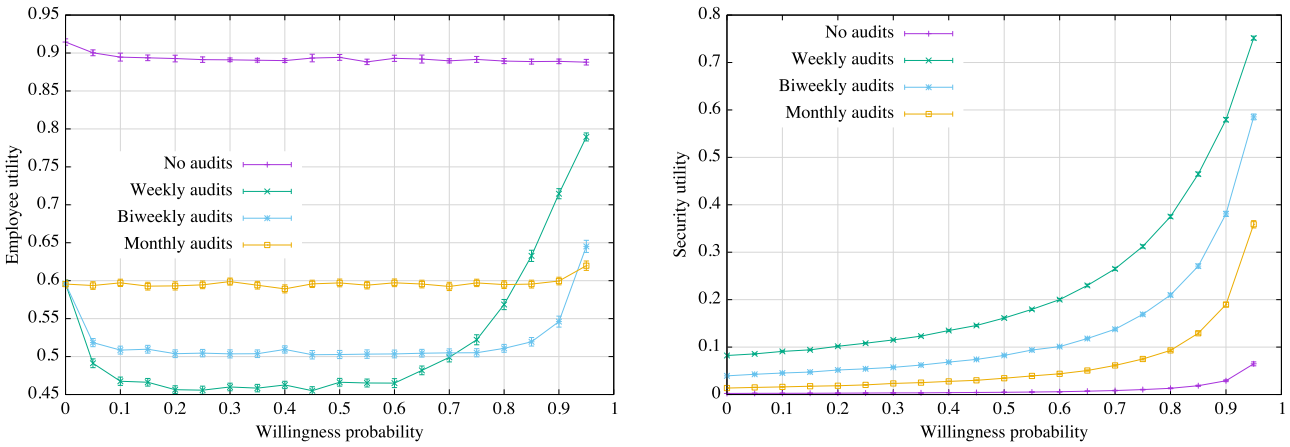
In Figures 5a and 5b, we fixed $\mathcal{P}_w = 0.5$ and $\mathcal{T}_a = 14$ (bi-weekly audits), and varied the password complexity $\mathcal{P}_c$ and the password write-down threshold $\mathcal{N}_p$. We used the same parameters for the employee utility function in order to stay within the framework of the general deterrence theory. We varied $\mathcal{P}_c$ over its entire range $(0, 1]$ and chose $\mathcal{N}_p \in \{1, 3, 5, 7\}$.

The results show that, contrary to popular belief, choosing a very complex password policy does yield the highest values for the security utility. The benefits of a very strong password policy are often outweighed by the risks of employees writing down their passwords. For the employees, due to the severity of the sanctions, they are more inclined to adopt of a high password write-down threshold since it will allow them to avoid the negative impacts of sanctions. This explains the increase in employee utility when $\mathcal{N}_p$ is increased. Intuitively, the employee's utility increases when the password policy is less complex (i.e., $\mathcal{P}_c$ increases) since it becomes easier for the employee to devise new passwords, thus increasing their positive experiences, reducing their workload and the probability of them being sanctioned.

Figure 4b shows that when $\mathcal{N}_p$ increases, the maximum security utility shifts towards more complex passwords (towards smaller $\mathcal{P}_c$ values). This is a desirable result since it reflects that the management is making use of the security benefits of a more complex security policy, as opposed to loosing these benefits due to employees writing down their passwords. When the threshold is low ($\mathcal{N}_p = 1$), there is practically no benefit from adopting a complex password policy since the risk from employees writing down their passwords become too large. This is shown by the fact that the maximum of the security utility occurs when $\mathcal{P}_p = 0.5$.

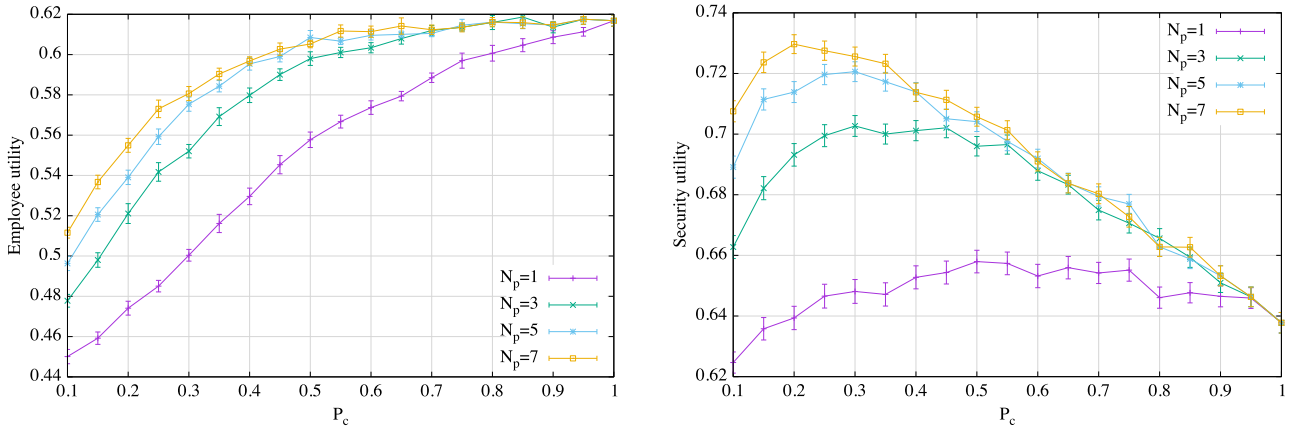
The results also show that, in accordance with the general deterrence theory, employing more severe sanctions pushes employees into more compliance, both in terms of security procedures and in putting in more effort towards creating complex passwords. This result is also in accordance with the conclusion made in [6] which states that making the outcomes of both pro and anti-security behaviors immediate and tangible would motivate compliant user behavior, thus increasing the overall security utility.

Finally, our results illustrate how the general deterrence theory can be employed as a basis for predictive models of human behavior in cyber security. Using such models, we were able to study the impact of the frequency of security audits and the complexity of the password policy on the employees' welfare and the system security in a typical firm.



(a) Employee Utility (b) Utility

Figure 4: Employee and Security Utilities for various audit frequencies



(a) Employee Utility (b) Security Utility

Figure 5: Employee and Security Utilities for various password thresholds

5. Challenges and Conclusions

Quantitative security metrics are gaining much attention within the cyber security research community. Since it is impossible to design perfectly secure systems, quantitatively assessing the security of certain designs becomes essential [33]. It is important to measure security in order to be able to improve it [34]. However, unlike hardware faults, security incidents often incorporate a human aspect. Thus, ignoring the models of the human element in system designs might lead to misleading results.

In the previous sections, we have motivated the need for involving human considerations in the design of systems and security policies. We reviewed several psychology and social sciences theories that provide descriptions of the human behavior in cyber security. We then illustrated how the general deterrence theory can be used as a basis for a model of customer representatives in an everyday firm. While this is a good step toward reconciling security modeling with human behavior models, much work is left to be done.

Security designers have devised models of cyber attackers with the aim of guiding the design process with quantitative metrics such as, among many, the probability of compromise, the mean time to compromise, and the attackers preferred attack paths [35], [36], [37]. Compared to human users and administrators, cyber attackers can be assumed to be rational decision makers aiming to maximize their profit while minimizing their relative cost. Users on the other hand, deviate from pure rational behavior and thus require, as we have discussed, more challenging models. We highlight the challenges related to building such models in the following.

The first challenge is concerned with designing accurate models of human decision making, ones that can fit within the context of stochastic simulation tools such as Möbius. Similar to attacker models, designing user and administrator models at an appropriate level of granularity is critical for quantitative security metrics. The level of detail required for such models is dependent on the scope of the experiment under design, and the type of metrics that the designer wants

to obtain. For example, a security manager looking to study the effectiveness of a proposed password policy needs to have an accurate model of how users design, memorize, and use their passwords. On the other hand, a system designer interested in evaluating the impact of security measures on the availability of a distributed file server in the presence of multiple users across multiple locations, needs models of the users at a different level of granularity.

Additionally, the nature and the uncertainty of human behavior cannot be easily captured by mathematical models. The rationality assumption that makes the reasoning about attacker models suitable for simulation tools no longer applies. The theories described in Section 3 are either (1) based on empirical studies that examine the behavior of humans, and then present explanations based on mere observations, or (2) propose a descriptive theory and conduct empirical studies to prove their applicability. Such theories are referred to as descriptive theories, i.e. theories that attempt to describe the human behavior as it actually is. Opposed to normative theories that describe what the ideal behavior should be, descriptive theories are hard to quantify.

Furthermore, human behavior is often influenced by individual aspects such as personality and culture. For example, some users are risk takers and tend to favor risky behavior if it provides them with immediate benefits (such as increased productivity), others however, prefer to always play by the rules. Therefore user models must be able to account for such different aspects. One possible solution is to create user profiles, similar to attacker profiles in [37], where users with different profiles behave differently.

Once models have been developed, the second challenge lies in obtaining valid parameter values that characterize the models. It is intuitive to represent organizational and attacker utilities in terms of monetary values, it is however more challenging to devise similar quantitative values for abstract concepts such as safety and culture. Consider for example the variable C in Equation 2, quantifying the cognitive effort that an employee puts in creating a new password is more challenging that characterizing the amount of monetary sanctions S or rewards R that the employee receives. What is needed is an extensive set of interviews with employees and system administrators that generates a large data set that allows the use of approaches such as the one in [38] to obtain model parameters.

A third and important issue that emanates from the first two challenges is the validation of the model and its results. Validation revolves around determining the degree to which a simulation model, its parameters, and its results provide an accurate representation of the real world, from the perspective of the intended use of the model [39]. The abstract nature of the theories in Section 3 and its corresponding challenges make the process of validation very arduous. However, aligned with the view presented in [13], this hardship is tempered by the fact that security metrics are highly useful in providing insights and guiding the decision making process, rather than providing exact monetary values. Currently, we do not envision that such models are to be used for finding absolute quantitative values for metrics. We do however find them very useful in providing relative values that can help compare different designs and policies, taking into consideration the presence of human users and administrators.

In summary, we envision that models of human users and administrators will make their way into the design and evaluation of security systems and policies. We imagine that adopting such an approach will provide security designers and managers with a more comprehensive view of the environment in which their designs or policies will be deployed. In this work, we have reviewed the theories of human behavior in cyber security from the fields of social sciences and psychology. In order to provide meaningful insights, any model of human users and administrators must have a valid basis from their fields. We then illustrated how the general deterrence theory can be used in a simple stochastic simulation model and highlighted the type of metrics and insights that we can obtain from it. We finally summarized the main challenges that researchers will face when taking on work in this area.

6. Future Work

In the future, we plan on investigating more accurate models than the one we used in Section 4. We plan on experimenting with different models of human behavior and comparing the metrics and insights that we can obtain from each. Additionally, we are looking at agent based modeling [40] as an alternative approach. In agent based modeling, a system is modeled as a collection of autonomous decision making entities called agents, where each agent assesses its current situation and makes decisions accordingly. This provides a way to model a system from the perspective of its users or employees, and allows designers to capture the interactions between different agents in the system. We envision that the theories we discussed in Section 3 provide a theoretical basis upon which agents depicting human users and administrators can be built.

Acknowledgments

The work depicted here was performed, in part, with funding from the the Maryland Procurement Office under Contract No. H98230-14-C-0141. We would also like to thank Jenny Applequist for her editorial comments.

References

[1] IBM Global Technology Services, "IBM security services 2014 cyber security intelligence index," Tech. Rep., 2014.

[2] C. Warzel and M. Zeitlin, "It gets worse: The newest Sony data breach exposes thousands of passwords," http://www.buzzfeed.com, 2014.

[3] H. Suo, J. Wan, C. Zou, and J. Liu, "Security in the Internet of Things: A review," in *Proc. Computer Science and Electronics Engineering (ICCSEE), 2012 International Conference on*, vol. 3, March 2012, pp. 648–651.

[4] M. Bishop, "Psychological acceptability revisited," in *Security and Usability: Designing secure systems that people can use*. Sebastopol, CA: OReilly Media, 2005, pp. 1–12.

[5] A. Beautement, M. A. Sasse, and M. Wonham, "The compliance budget: Managing security behaviour in organisations," in *Proceedings of the 2008 Workshop on New Security Paradigms (NSPW)*. New York, NY, USA: ACM, 2008, pp. 47–58.

[6] R. West, "The psychology of security," *Commun. ACM*, vol. 51, no. 4, pp. 34–40, Apr. 2008.

[7] J. D'Arcy, A. Hovav, and D. Galletta, "User awareness of security countermeasures and its impact on information systems misuse: A deterrence approach," *Information Systems Research*, vol. 20, no. 1, pp. 79–98, 2009.

[8] A. Beautement, R. Coles, J. Griffin, C. Ioannidis, B. Monahan, D. Pym, A. Sasse, and M. Wonham, "Modelling the human and technological costs and benefits of USB memory stick security," in *Managing Information Risk and the Economics of Security*, M. Johnson, Ed. Springer US, 2009, pp. 141–163.

[9] D. Eskins and W. Sanders, "The multiple-asymmetric-utility system model: A framework for modeling cyber-human systems," in *Proc. Quantitative Evaluation of Systems (QEST), 2011 Eighth International Conference on*, Sept. 2011, pp. 233–242.

[10] C. Colwill, "Human factors in information security: The insider threat - who can you trust these days?" *Information Security Technical Report*, vol. 14, no. 4, pp. 186–196, 2009.

[11] B. Schneier, "The psychology of security," in *Proc. Progress in Cryptology: First International Conference on Cryptology in Africa (AFRICACRYPT)*, S. Vaudenay, Ed., vol. 5023. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 50–79.

[12] A. Adams and M. A. Sasse, "Users are not the enemy," *Commun. ACM*, vol. 42, no. 12, pp. 40–46, Dec. 1999.

[13] W. Sanders, "Quantitative security metrics: Unattainable holy grail or a vital breakthrough within our reach?" *IEEE Security and Privacy*, vol. 12, no. 2, pp. 67–69, Mar. 2014.

[14] J. B. Taylor, "Discretion versus policy rules in practice," *Carnegie-Rochester Conference Series on Public Policy*, vol. 39, pp. 195–214, 1993.

[15] S. E. Parkin, A. van Moorsel, and R. Coles, "An information security ontology incorporating human-behavioural implications," in *Proceedings of the International Conference on Security of Information and Networks (SIN)*. New York, NY, USA: ACM, 2009, pp. 46–55.

[16] S. E. Parkin, R. Yassin Kassab, and A. van Moorsel, "The impact of unavailability on the effectiveness of enterprise information security technologies," in *Proc. Service Availability: 5th International Service Availability Symposium (ISAS)*, T. Nanya, F. Maruyama, A. Pataricza, and M. Malek, Eds., vol. 5017. Springer Berlin Heidelberg, 2008, pp. 43–58.

[17] M. Theoharidou, S. Kokolakis, M. Karyda, and E. Kiountouzis, "The insider threat to information systems and the effectiveness of ISO17799," *Computers & Security*, vol. 24, no. 6, pp. 472–484, 2005.

[18] T. Herath and H. Rao, "Encouraging information security behaviors in organizations: Role of penalties, pressures and perceived effectiveness," *Decision Support Systems*, vol. 47, no. 2, pp. 154–165, 2009.

[19] J. B. Hardee, R. West, and C. B. Mayhorn, "To download or not to download: An examination of computer security decision making," *ACM interactions*, vol. 13, no. 3, pp. 32–37, May 2006.

[20] P. Ifinedo, "Understanding information systems security policy compliance: An integration of the theory of planned behavior and the protection motivation theory," *Computers & Security*, vol. 31, no. 1, pp. 83–95, 2012.

[21] B. Bulgurcu, H. Cavusoglu, and I. Benbasat, "Information security policy compliance: An empirical study of rationality-based beliefs and information security awareness," *MIS Quarterly*, vol. 34, no. 3, pp. 523–548, 2010.

[22] J.-Y. Son, "Out of fear or desire? toward a better understanding of employees' motivation to follow IS security policies," *Information & Management*, vol. 48, no. 7, pp. 296–302, 2011.

[23] A. C. Johnston and M. Warkentin, "Fear appeals and information security behaviors: an empirical study," *MIS quarterly*, vol. 34, no. 3, pp. 549–566, 2010.

[24] B.-Y. Ng, A. Kankanhalli, and Y. C. Xu, "Studying users' computer security behavior: A health belief perspective," *Decision Support Systems*, vol. 46, no. 4, pp. 815–825, 2009.

[25] J. Lee and Y. Lee, "A holistic model of computer abuse within organizations," *Information Management & Computer Security*, vol. 10, no. 2, pp. 57–63, 2002.

[26] W. H. Sanders and J. F. Meyer, "Stochastic activity networks: Formal definitions and concepts," in *Lectures on Formal Methods and Performance Analysis*, ser. Lecture Notes in Computer Science, E. Brinksma, H. Hermanns, and J.-P. Katoen, Eds., vol. 2090. Springer Berlin Heidelberg, 2001, pp. 315–343.

[27] G. Clark, T. Courtney, D. Daly, D. Deavours, S. Derisavi, J. M. Doyle, W. H. Sanders, and P. Webster, "The Möbius modeling tool," in *Proc. Petri Nets and Performance Models, 2001. 9th International Workshop on*, 2001, pp. 241–250.

[28] B. Ur, S. M. Segreti, L. Bauer, N. Christin, L. F. Cranor, S. Komanduri, D. Kurilova, M. L. Mazurek, W. Melicher, and R. Shay, "Measuring real-world accuracies and biases in modeling password guessability," in *Proc. 24th USENIX Security Symposium (USENIX Security 15)*. Washington, D.C.: USENIX Association, Aug. 2015, pp. 463–481.

[29] C. Kuo, S. Romanosky, and L. F. Cranor, "Human selection of mnemonic phrase-based passwords," in *Proceedings of the Second Symposium on Usable Privacy and Security (SOUPS)*. New York, NY, USA: ACM, 2006, pp. 67–78.

[30] F. Greitzer, J. Strozer, S. Cohen, J. Bergey, J. Cowley, A. Moore, and D. Mundie, "Unintentional insider threat: Contributing factors, observables, and mitigation strategies," in *Proc. System Sciences (HICSS), 2014 47th Hawaii International Conference on*, Jan. 2014, pp. 2025–2034.

[31] Verizon, "2014 data breach investigations report," Tech. Rep., 2014.

[32] Symantec, "Internet security threat report," Tech. Rep., 2015.

[33] D. M. Nicol, W. H. Sanders, and K. S. Trivedi, "Model-based evaluation: from dependability to security," *Dependable and Secure Computing, IEEE Transactions on*, vol. 1, no. 1, pp. 48–65, Jan. 2004.

[34] S. M. Bellovin, "On the brittleness of software and the infeasibility of security metrics," *IEEE Security & Privacy*, vol. 4, no. 4, p. 96, 2006.

[35] D. Leversage and E. James, "Estimating a system's mean time-to-compromise," *IEEE Security & Privacy*, vol. 6, no. 1, pp. 52–60, Jan. 2008.

[36] B. Schneier, "Attack trees," *Dr. Dobbs Journal*, vol. 24, no. 12, pp. 21–29, 1999.

[37] E. LeMay, M. D. Ford, K. Keefe, W. H. Sanders, and C. Muehrcke, "Model-based security metrics using ADversary VIew Security Evaluation (ADVISE)," in *Quantitative Evaluation of Systems (QEST), 2011 Eighth International Conference on*, Sept. 2011, pp. 191–200.

[38] R. Cain and A. van Moorsel, "Proc. optimization of data collection strategies for model-based evaluation and decision-making," in *Dependable Systems and Networks (DSN), 2012 42nd Annual IEEE/IFIP International Conference on*, June 2012, pp. 1–10.

[39] P. Sanders, "DoD modeling and simulation (M&S) verification, validation, and accreditation (VV&A)," DTIC Document, Tech. Rep., 1996.

[40] E. Bonabeau, "Agent-based modeling: Methods and techniques for simulating human systems," *Proceedings of the National Academy of Sciences*, vol. 99, no. suppl 3, pp. 7280–7287, 2002.