Mitigating Inadvertent Insider Threats with Incentives

Debin Liu, XiaoFeng Wang and L. Jean Camp

School of Informatics, Indiana University

Abstract. Inadvertent insiders are trusted insiders who do not have malicious intent (as with malicious insiders) but do not responsibly managing security. The result is often enabling a malicious outsider to use the privileges of the inattentive insider to implement an insider attack. This risk is as old as conversion of a weak user password into root access, but the term inadvertent insider is recently coined to identify the link between the behavior and the vulnerability. In this paper, we propose to mitigate this threat using a novel *risk budget* mechanism that offers incentives to an insider to behave according to the risk posture set by the organization. We propose assigning an insider a risk budget, which is a specific allocation of risk points, allowing employees to take a finite number of risk-seeking choice. In this way, the employee can complete her tasks without subverting the security system, as with absolute prohibitions. In the end, the organization penalizes the insider if she fails to accomplish her task within the budget while rewards her in the presence of a surplus. Most importantly. the risk budget requires that the user make conscious visible choices to take electronic risks. We describe the theory behind the system, including specific work on the insider threats. We evaluated this approach using human-subject experiments, which demonstrate the effectiveness of our risk budget mechanism. We also present a game theoretic analysis of the mechanism.

Keywords: Insider Threat, Incentive Engineering, Human Subject, Game Theory

1 Introduction

Organizations have long been struggling with the dilemma of how to protect themselves from those parties they must trust in the ordinary course of business. These parties, called insiders, include employees, contractors, consultants and others who have access to critical aspects of the organization. An insiders privileged position gives him the opportunity to easily abuse organizational trust for personal gain. This creates a grave risk to the confidentiality, integrity and availability of critical information assets. For example, the National Association of State Chief Information Officers [1] reports 80% of publicized data breaches came from organizational threats instead of outside threats in 2006, and in 2005, more than half were attributed to insider threats. Another report from [2] estimates around half of survey participants experienced an insider incident in 2007.

Generally speaking, there are two types of insider threats based on the insiders intents. Malicious insiders are the individuals with varying degrees of malicious intent to cause harm. Inadvertent insiders are the individuals who do not have malicious intent. The E-Crime Watch survey investigates insider malicious attacks, but most IT experts

agree that most leaks of information and security breaches are not criminal but the result of accidents and human errors [3]. According to the Department of Commerce, U.S. businesses use more than 76 million PCs and laptops. The 200 million business users worldwide of Microsoft Office send over 100 million documents over email daily [4]. And this is only the information shared over email, and does not consider any other electronic means. Undeniably much of this data are work-related and require transmission. Yet for that which is not, the risk to the organization is invisible to the insider making the decision to take an electronic risk.

In this paper we focus on how to mitigate the inadvertent insider threats. Inadvertent insiders are usually defined as inattentive, complacent or untrained people who require authorization and access to an information system in order to perform their jobs. Such people may not realize the risks incumbent to having access to their system resources. For example, they are operating in a network-centric environment, which creates the possibility that a virus downloaded to one computer could infect a myriad of other computers connected to the same network. Some have jobs that are dominated by routine activities. As their tasks become more mundane, the likelihood will increase that a complacent user may not fully appreciate the potential consequences should an error lead to the leaking of sensitive information [5] [6]. Even a person with significant experience in computing may not have an appreciation for security risks. For example, an individual employee may not understand the value of updating anti-virus signatures on a regular basis. For untrained users, it is not a matter of intending to do harm, it is a matter of not having the requisite information to make informed choices about security.

Risk communication has the potential to mitigate the inadvertent insider threats [4]. A properly designed warning message could help an inadvertent insider understand the potential risk of their actions. A reliable informative alert can reduce the possibility that a complacent user makes a mistake when his activity is risky. With the detailed information, an untrained user can receive educational information from the risk communication and make an informed choice. Yet even excellent risk communication is no panacea to the inadvertent insider threats. Previous research has shown that even a well-delivered risk communication message cannot fully educate and inform most common users [7]. To many people, such risk communication messages are annoying rather than thankworthy.

Thus the problem of the inadvertent insider is two fold. First the individual does not know of the risk and may reject or avoid risk communication. Second, the incentives are incorrectly aligned for the individual insider. The insider or employee wants to keep his or her job. The insider wants to finish the tasks assigned without being interrupted to update an application; or even forced to seek entertainment at alternative sites.

Given that the insiders are usually rational and motivated by realizing their personal gains [8], we believe incentive modeling can help us understand an inadvertent insiders motives and strategies. We consider the following scenario. An inadvertent insider is about to download a football sport screensaver to his company computer. There are two websites offering free downloads of such screensavers. One of them is rated as “high risk website containing adware, spyware and viruses download” by security vendors while the other one is rated as low risk. Although some messages may pop up and warn the user to keep away from that risky website, he may still visit the risky website re-

gardless of any risk caused by his download posed to his company computer system. Currently browser-centric warnings would be identical for both websites. As an inadvertent insider, the user is only motivated by his personal gain, in this case the sport screensaver. Thus he decides to ignore the risk-warning message. In other words, the risk communication is not effective.

In our scenario, the cost of downloading from a risky website is born by the company rather than the user, and there is no incentive for the user to take risk communication seriously and worry about any potential risk caused by his actions. In this paper, we propose to shift the cost of risk from the organization to the inadvertent insider. By using incentive engineering, we designed a mechanism to encourage the users to self-manage their risks, discourage the users against their risky activities, and thus mitigate the inadvertent insider threats. Our approach gives each user a bucket of risk points called risk budget, and every move the user takes could cost him some points. If the user runs out of his budget before having his job done, he could be subject to certain penalty from his organization. On the other hand, if he behaves prudently and finishes his task before using up his points, the organization can reward him. The assignment of risk budgets is determined by the natures of individual positions. Our research shows that such a simple approach turns out to be very effective at suppressing irresponsible behaviors, according to our experimental studies. We also analyzed our approach using game theory.

The rest of the paper is organized as follows. Section 2 reviews the related work. In Section 3, we introduce our risk budget mechanism, and move on to describe human-subject experiments that evaluated our approach in Section 4 and Section 5. A game theoretic analysis is presented in Section 6 for better understanding of our mechanism. We conclude the paper and describe the future work in Section 7.

2 Related Work

The apparent irrationality of end users in choosing with whom to share information creates inadvertent insiders. The inadvertent insider can be informed by incentive mechanisms and deterred from making risky choices. The incentives have to be aligned with the interests of the users [9] [10]. For example, security incentives that prevent users from performing critical tasks will be ignored or disabled.

The core research challenge our design addresses is how to engineer incentives so either the risk behaviors incur some cost, or enable the end user to detect the security costs of a misbehaving account. Essentially the research question is how to encourage users not be risk-seeking (e.g., inadvertent insiders) by utilizing incentives.

Solutions to the problem of inadvertent insiders have included insurance that correlates with security practices [11], changing defaults so that security is difficult to avoid [12], more careful accounting of actual costs [13] and improved system usability [14] [15]. It is the core contention of the proposed research agenda that there is a clear and profound need for effective risk communication. While there have been studies of user conceptions of privacy [16] [17] and usable security [18]; these have focused on system design rather than contextual behaviors.

We assert that effective security communication is critical for handling the problem of the inadvertent insider. Changing behavior requires both communicating security information and motivating the appropriate security behaviors. The essential point is that the purpose of security communication is not conveying the perfect truth to the users, but rather to prompt them to take an appropriate action to defend their system against a certain threat [9] [10]. Mitigation of a security risks that are behavior-based does not requires that the user have the knowledge of the risk, but rather a general idea of the nature of that risk.

For the inadvertent insider considering violating security policy, the risks corresponding to the policy-forbidden actions are rarely clearly identified. In no case is there an indicator of risk-averse action that might be taken in order to reduce the risks should the user choose the particular action [19]. For example, if users choose to subvert a policy by using public email providers (e.g. Gmail) to share documents, there is no education about readily available encryption options. Yet a communication about the risks of sending documents and the option of encryption could be included should the employee go to a free email site. The efficacy of incentive technologies is to some degree a function of the assumptions of human risk behaviors in the network [20]. We will design and build our incentive mechanisms upon foundational insights that have emerged from studies on human-computer interaction and game theoretic studies of behavior.

The combination of the game theory, incentives and human interaction is what makes this work unique. In comparison, [21] proposed an access control system that used a market to distribute access tokens where the price may be set by the data owner. In this case, the response is statics and the system does not evaluate the responses in order to identify the nature of user. Nor does the system embed risk communication or risk mitigation. Horizontal Integration [22] proposes the use of risk tokens and risk calculations to manage access control. Tokens are distributed to employees in a hierarchical approach, by the organization. Again, employees trade tokens for access. Similarly the system does not use any game theoretical pricing, does not address user behavioral history, and ignores issues of risk mitigation and communication. [23] describes the mechanisms for distributing risk token to employees for access control. While these proposal use an approach that is conceptually similar to the risk budget concept, none of these approaches offers the employers opportunities for risk mitigation. Nor do the approaches engage the benign employee in risk communication in order to enable a more informed decision by the employee. But the most significant difference between the proposed research and the work described here is that we conceptualize the use of resources as a game, with different types of players. For example, in the systems above, an insider could abuse her the tokens for her personal gains. We add incentives (e.g., punishment and rewards) to regulate insiders and mitigate possible risk budget abuse. We also limit the possible damage, by tracking and responding to insider behaviors in a strategic manner (to the extent that the game theoretic model is solvable).

We will also build on the insights in the work in [24]. FuzzyMLS considers access control as an exercise in risk management. Access control decisions are a function of the risk of action or access, risk tolerance of the individual requesting access, and risk mitigation. FuzzyMLS also computes a quantified estimate of risk associated with a human subject. FuzzyMLS utilizes risk tokens in that zone of uncertainty, a fuzzy or

gray area, between permission and denial and proposes an unspecified market for risk exchanges. Such a risk exchange could prove hazardous to an organization, as an insider could build significant risk rights while remaining invisible to the organization. FuzzyMLS does not address the state of the machine requesting the access. FuzzyMLS uses the organizational level of the individual to determine the risk characteristic associated with the user. The past behaviors or choices of the user (e.g., risk seeking or risk averse) are not considered. While FuzzyMLS uses the language of economics, it is not informed in any way by the economics of security nor does it embed incentives that are understandable by the user. For example, they propose using a ROI (return on investment) model to reward users who avoid risk and market to trade risk, yet no implementation or method of calculation is proposed. In contrast, we have built a proof of concept and seek support to build a more complete prototype. While there were no user tests of FuzzyMLS, the fact that the decisions are opaque to the user indicates that the incentive structure may be ineffective in practice.

3 Risk Budget Mechanism

The problem of inadvertent insider threats is that the cost of risk is born by the organization rather than the users who initiate risky activities. In order to shift the cost back to the users themselves, we propose a risk budget mechanism. The principles of our risk budget mechanism are as follows.

- Every user is assigned a bucket of risk points for his task.
- A users risky activity will cost him some risk points.
- A user will be punished once a users risk budget gets exhausted.
- The more points remain the more rewards a user gets when he complete his task.

The requirement of consuming risk points, together with the punishment and the reward, shift the cost of risk to the users. The risk budget mechanism visualizes the cost to user and produces incentives that motivate the users to avoid risky activities.

3.1 *Risk Budget Assignment*

We denote the bucket of risk points for a user i by B_i. The size of the bucket is determined by the organization based on the user's task description, and the organizations preference. For example, if a user's job requires exploring the Internet and visiting various websites with a potential high risk, he will have a higher risk budget than someone whose main work is database maintenance. For instance, an employee who visits rating sites and social network sites to manage the companys reputation will have a large risk budget. An employee in human resources who can access the payroll database will have a very small risk budget. A user's security preference may also be considered when assigning him a risk budget. To put it simple, a risk-seeking user will be given a more limited risk budget.

3.2 *Points Payment*

As we focus on inadvertent insiders, it is reasonable to recognize that all the potentially harmful insiders are not malicious and thus they only take actions based on their privileges and access. Since the organization knows the insiders access, it knows all the possible valid actions a user can take. In addition, we assume the organization is able to associate a risk rating with each action or access right. Each action the user i has the privilege to take, a_j, is associated with a given price in terms of risk points, p_{aj}. Our current research uses web-surfing activities to study the general idea of risk budget. In this case, the point price of visiting a website can be identified from the website's ratings given by various sources [12] [25]. A further study on this direction could lead to risk-aware access control, which we plan to pursue in the follow-up research.

3.3 *Punishments*

An incentive against risk-seeking behaviors our approach offers is the punishment inflicted on the users once they empty their risk budget. Such punishment refers to some form of cost that is enforced by the organization and triggered by the risk budget exhaustion. It could be an audit or mandatory training program or a loss of access. The budget size implies a risk limit that the organization could bear for a specified task. And the punishment translates the exhausted budget into a cost that directly aligns the companys and users incentives. The risk budget connects the risk suffered by the organization and the posted cost born by the users. As a result, the risk points spent by a user can reflect his willingness to launch a risky action.

3.4 *Rewards*

The punishment caused by an exhausted risk budget brings an incentive to the user against risky action. However, such a punishment only happens only when the user empties his risk budget, which can be late. Moreover, it is desirable that the user can be encouraged to choose the least risky path for accomplishing his task, which minimizes the risk the organization is exposed to. To this end, we take a measure that rewards the user according to the surplus of his risk budget. Simply speaking, the fewer risk points consumed the more rewards the user will get. Formally, we define a reward as a function $R(p)$ of the remaining risk point p after a task is completed. In practice, the rewards can be paid in the form of welfare. For example, the unspent risk points are accumulated from day to day. Once the points reach some level, the user can then redeem his points in exchange of a vacation or a bonus or a prize. Prior research shows that a combination of penalties and rewards is more effective in employee motivation than penalties alone [26].

3.5 *An example*

Within the risk budget mechanism, users can no longer abuse their privileges without bearing any cost. As an example, consider an Internet commerce researcher whose job

demands a daily Internet surfing. Suppose the user has a daily risk budget B_i for downloading documents the Internet. He can visit a website w_j that costs him risk points p_j to perform the downloading, which costs him another p_k. Alternatively, he can visit another website w_j that requires p_j for visit and p_k for document downloading. The prices p_j, p_k, p_j and p_k are set by the organization based on its perception and evaluation of potential risks. Assuming $B_i > (p_j + p_k) > (p_j + p_k)$, we expect user i voluntarily chooses the second website, which incurs lower risks, under our risk budget mechanism.

4 Experiment Design

We conducted two human-subject experiments in order to evaluate our risk budget mechanism. The first experiment was designed for understanding users' risk behaviors, and the second one aimed at studying the change of these behaviors under our incentive mechanism. The outcomes of these experiments are elaborated in Section 5. These experiments were based upon a firefox browser extension we implemented for monitoring a user's web browsing behaviors, adjusting his risk points and enforcing penalty/reward policies.

4.1 *Recruitment*

We recruited 40 participants for the experiments and divided them randomly into two groups: 20 for the first experiment and the other 20 for the second experiment. All participants were recruited voluntarily from the undergraduates at Indiana University, Bloomington. None of the participants were majored in information security related fields. Most of them were in their freshmen year.

4.2 *Ratings*

We determined the risk rating of a website using a mechanism proposed in the prior research [25]. The mechanism rates websites as follows.

1. those that have been previously visited are trusted unless otherwise identified;
2. those that have not been previously visited are considered untrusted;
3. the ratings of an untrusted website comes from McAfee SiteAdvisor [27].

Detailed information on the reputation system itself can be found in [25]. However, this mechanism was used for convenience and in fact nearly random ratings could have been used in the experiments without loss of generality of the results. In fact, because of the nature of the reputation system, all negative ratings were a result of McAfee. Note that McAfee SiteAdvisor is a system of automated testers that continually search the Internet via browsing websites with human browsers and honey monkeys. The searchers download files, clicks on adds, and enter information on sign-up forms. The results are documented and supplemented with feedback from users, comments from website owners, and analysis from researchers. In our experiments, a participant was charged with a randomly-generated price ranging from 10 to 20 points whenever he/she was

about to visit a risky website. The reason why we ask for a random charge is that we would like to discover the risk payment distribution. Such risk payment distribution will help us determine an effective and reasonable risk price in our future study on risk-aware access control.

4.3 *Task Descriptions*

There are arguably thousands of websites offering free downloads of screensavers on the Internet. Many of them contain malicious content, yet distinguishing between the dangerous, potentially annoying, and benign websites is difficult. Downloading active or potentially active content can be high-risk activity. Thus it was this risk activity that was chosen as the basis of the experiments.

In the experiments, each participant was asked to locate five screensavers from five different websites respectively. In other words, the experiment consisted of five tasks. Each task was to locate and select a screensaver from any website. All participants were free to choose any website to surf and download the requested screensavers. They had multiple choices to complete their tasks.

Following are the detailed instructions these participants received:

1. Search for the websites offering free screen savers downloads from the web.
2. From the search results, choose five websites: website-1, website-2, website-3, website-4 and website-5.
3. From website-1, please take a screenshot of an animal screensaver.
4. From website-2, please take a screenshot of a nature screensaver.
5. From website-3, please take a screenshot of a sport screensaver.
6. From website-4, please take a screenshot of a space screensaver.
7. From website-5, please take a screenshot of a flower screensaver.
8. Thank you. You have completed the experiment.

The goal was to create a somewhat mundane set of tasks when the completion of the task resulted in immediate payment. Rather than testing the security interaction as if security were the goal, our experimental design was to create a set of tasks that are orthogonal (or even in opposition to) security.

4.4 *Experiment One*

In the first experiment, the participants were asked to pick five different websites from their previous search results as described above. All websites were rated according to the security vendors websites risk ratings [12] [25]. A website was considered high risk if it were rated as “high risk website containing adware, spyware and viruses download”. When a participant clicked on the link of a high risk website, a warning message appeared. Such warning messages communicated with participants about the potential risks of the website and asked for their confirmation. A screenshot of the warning message is shown in Figure 1.

Certainly others have documented the general tendency to swat security boxes out of the way in order to complete tasks. Determining the prevalence of this behavior and

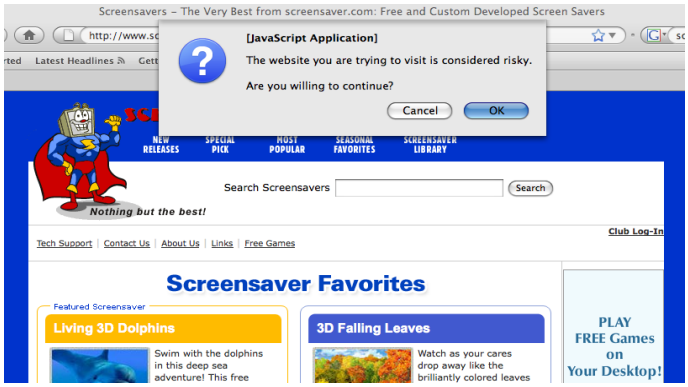

Fig. 1. The Screenshot of The Warning Message in Experiment One

ensuring consistency of the wording of the messages were critical reasons for this first experiment.

We recorded the browsing history, the participants responses to the warning messages and the time used for completing the task. The outcomes of the experiment, designated as data set R1, is presented as a baseline of local users' risk behaviors.

4.5 *Experiment Two*

In the second experiment, every participant was given an identical initial risk budget. If a website was tagged as high risk, it was then associated with a risk price from our rating mechanism. This second set of 20 participants was asked to complete the same task under the additional constraint of their risk budgets. If they successfully accomplished their tasks, they received $\$10$ plus an additional amount based on the risk budget. If any participant exhausted a risk budget, that participant forfeited their compensation. In addition, if any participant failed to complete the experiment, that participant would similarly forfeit compensation.

Participants were also rewarded with risk budget surplus with a bonus, whose amount depended on the amount of points left in his bucket. For instance, a participant who saves 20 points receives $\$10$ for completing his task and an additional $\$2$ for the saving. The formula we used to calculate the bonus is $\$10 \times (B - P_c)/B$, where B is the budget size and P_c is the points consumed in the experiment. Thus participants could make up to $\$20$ and a little as nothing. When a participant clicked on the URL of a high risk website, a warning message appeared. The warning contained not only the same text as the previous warning but also an indicator of the risk cost for the visit. A screenshot of the warning message is shown in Figure 2.

As with the first experiment we recorded the participants browsing history, their responses to the warning messages, and the total time used for completing the task. In addition we recorded prices (in risk points) paid for web activities, and the risk points remaining when the task was complete. The set of results is denoted as R2.

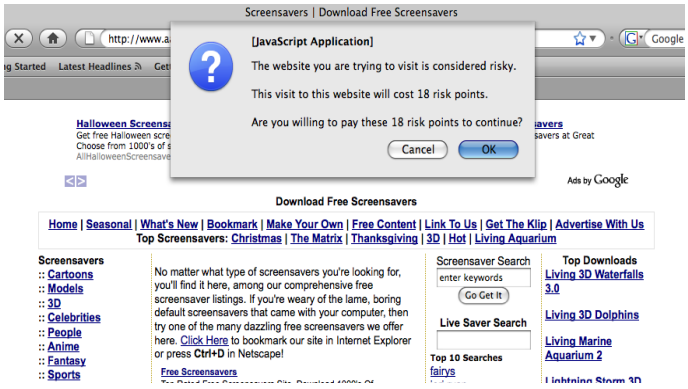

Fig. 2. The Screenshot of The Warning Message in Experiment Two

4.6 *Firefox Browser Extension*

Both experiments were based upon a Firefox browser extension, which was triggered whenever a browser was launched. The extension performed the following operations:

1. Detect a new page being loaded;
2. Check the domain name of a webpage;
3. Maintain a list of target high risk websites and their reputations according to [25];
4. Pop up a warning message when a high risk website was visited;
5. Ask for confirmation and or rejection of the visit choice from the participant; and
6. Record the experimental results,
   (In experiment two, the extension also took the following actions:)
7. Generate a price based on a website's reputation,
8. Track of participants risk budgets.

5 Data Analysis

We recorded the results of Experiment one, R1 and Experiment two,R2, as noted above. These data consists of participants browsing history, their responses to each pop-up warning message and the time they spent to accomplish their tasks. Furthermore, R2 also contains participants' payments for risky websites in terms of risk points and their risk budgets. A snippet of R1 and R2 is shown in Figure 3. At the end of each record is the time that a participant took to complete the experiment. The notation "@Y@" indicates a decision to perform a risky activity, for example, visiting a dangerous website, and "N" points to the action that avoids potential risks, for example, refraining from surfing dangerous sites. In R2, the numbers posterior to these notations is the prices a participant paid in the experiment and his remaining risk points.

```
R1
......
www.google.com,www.google.com,images.google.com,www.freesaver.com,coolscreensavers.googlepages.com,www.tnpsc.com,@Y@,www.google.com,ww
w.fabuloussavers.com,@Y@,www.fabuloussavers.com,@Y@,www.google.com,www.space-screensavers.com,@Y@,www.google.com,www.3d-screensaver-
downloads.com,www.3d-screensaver-downloads.com,,www.google.com,www.freesaver.com,Monday, October 13, 2008,7:52:13,5
......

R2
......
www.google.com,www.google.com,www.freesaver.com,www.google.com,www.webshots.com,www.google.com,www.scenicreflections.com,$N$,
13,100,www.google.com,www.google.com,www.google.com,www.google.com,findarticles.com,fbgdc.com,popularscreensavers.smileycentral.com,ww
w.widgetworld.nl,www.google.com,www.sports-logos-screensavers.com,@Y@,17,83,www.google.com,www.space-screensavers.com,$N$,
15,83,www.google.com,$N$,11,83,www.google.com,www.google.com,www.3deepspace.com,www.google.com,www.thegardenhelper.com,83,Sunday,
October 12, 2008,10:33:31,6
......
```

Fig. 3. The snippet of R1 and R2

5.1 *Risk Behaviors*

During the first experiment, the first group of participants received 104 pop-up warning messages in total. In the second experiment, there were 106 pop-ups for the other 20 participants. In other words, to complete the same task, the participants in both experiments encountered statistically similar numbers of risk warnings. However, their risk behaviors were significantly different. Among 104 warning messages, the participants in Experiment one made 81 risk-seeking decisions (i.e., continuing to visit dangerous websites) and 23 risk-averse decisions (i.e., avoiding risk websites). Under our risk budget mechanism, the participants in Experiment two responded with 11 confirmations of risk-seeking behaviors and 95 responses of risk-averse behaviors. The following figures show their risk behavior distributions.

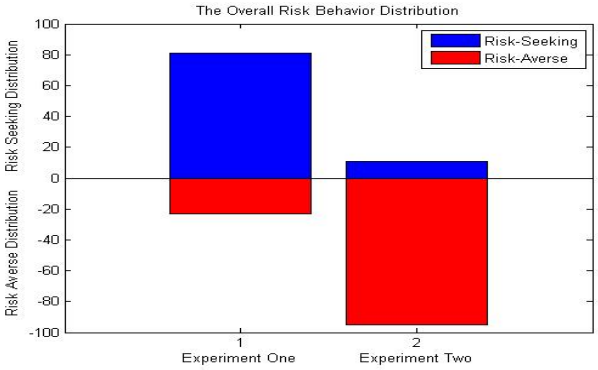

Fig. 4. Differences of Risk Behavior Distributions in two experiments

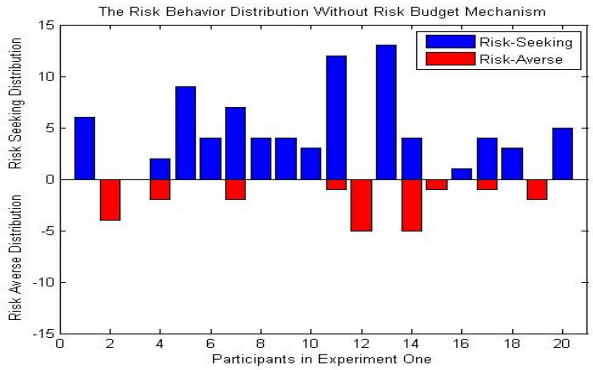

Fig. 5. Risk Behavior Distributions in Experiment One

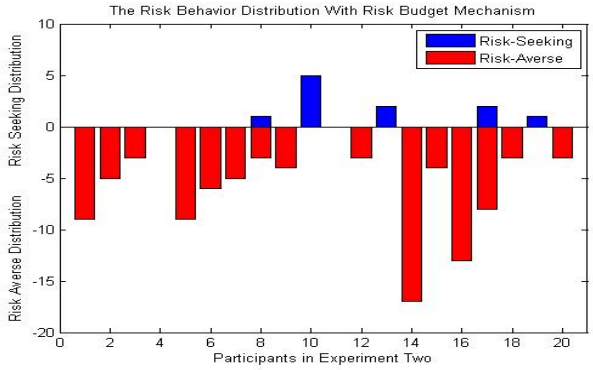

Fig. 6. Risk Behavior Distributions in Experiment Two

From these figures, we can observe the significant impact our mechanism can exert on users to suppress their risk-seeking behaviors. Through issuing proper rewards and penalties, the risk budget approach shifted the participants' risk behaviors from a strong preference of risk seeking to a strong preference of risk averse.

5.2 *Risk Boundary*

There were 11 positive responses from participants that confirmed risk-seeking behaviors in Experiment two. The average payment made by those who chose to bear risks was merely 16 points. This is in a stark contrast to what happened in Experiment one, where 20% participants each made more than 7 positive responses, which amounts to

depletion of their risk budgets if they were assigned ones with the sizes of those used in Experiment two. In Experiment two, we actually did not observe any participants failed the task and exhausted their risk budgets. These experimental results clearly indicate that penalty and rewards based upon risk budgets can effectively motivates users against abuse of their privileges. Meanwhile such an incentive helps establish a boundary for organizations and helps them to manage their risk.

5.3 *Regulation Friction*

The results of Experiment one show that without any incentive users are not willing to change their behaviors. We consider this is caused by a regulation friction that defines the efforts made by the users to adopt a risk-averse strategy instead of a risk-seeking strategy. In our experiments, we measured this regulation friction using time interval. The average time interval for completing the task in Experiment one is 5 minuets and 45 seconds. It becomes 6 minuets in Experiment two. Therefore, the regulation friction is only 15 seconds, merely 4.3% of the efforts participants made in Experiment one. Such a small friction can be easily overcome with a penalty/reward mechanism, as demonstrated in Experiment two.

6 Analysis

Our risk budget mechanism offers incentives to users to behave responsibly, shifts the cost of risk to insiders themselves, and encourages them against risk activities. The experiment results demonstrate its positive impacts to users risk behavior. In this section, we first analyze the risk budget mechanism using game theory, and then discuss how to implement our mechanism in practice.

6.1 *Risk Budget Mechanism as A Game*

Game theory studies the strategic interactions among rational players in which every player chooses its optimal move based upon her counter-speculation of others optimal moves. A solution of a game is determined by the point of equilibrium, which defines a xed point of players strategic interactions [26].

Inadvertent insiders are rational and motivated by incentives. Therefore game theory is an ideal tool to model their interactions with their organization. Application of game theory to the insider problem can predict the best move an intelligent and knowledgeable insider may take and enable organizations to prepare for that move.

A typical game consists of a set of players, their action spaces, and their payoff functions. We model the risk budget mechanism as a game played between a user and an organization administrator. Both players are rational and their objectives are to maximize their payoffs. A users payoff is calculated based on the penalty and the rewards he receives. In addition, a cost is incurred by his efforts to choose a path with minimal risk to accomplish his tasks. The administrators payoff is measured by the cost brought in by risky activities and the rewards provided to the user. In the presence of a reasonable risk budget and penalty for failing a mission, it is conceivable that depletion of

one's budget before having his job done is not an option he is going to take. Therefore, here, we only consider the situation where users choose between whether to take an optimal path to do her job, which avoids excessive risks but introduces the costs for planning, and a suboptimal one that will spends all his budget on the task. The administrators action space contains two actions: "not rewarding the user whose risk budget is not empty" and "rewarding such a user". The first action reflects the organizations strategy in Experiment one, while the other one reflects the organizations strategy in Experiment two. The users action set includes two strategies: the risk-seeking strategy and the risk-averse strategy. The game is presented in the normal form as follows.

	Risk-Seeking	Risk-Averse
No Reward	$(-P_1, 0)$	$(-P_2, -C)$
Reward	$(-P_1 - R_1, R_1)$	$(-P_2 - R_2, R_2 - C)$

Table 1. Structure of User Response Game

The notations are explained below.

- P_1 represents the cost of risk to the organization when the user adopts a risk-seeking strategy.
- P_2 represents the cost of risk to the organization when the user adopts a risk-averse strategy.
- $P_1 > P_2$
- R_1 represents the reward given to the user when a risk-seeking strategy is adopted.
- R_2 represents the reward given to the user when a risk-averse strategy is adopted.
- $R_1 << R_2$
- C represents the friction between the risk-seeking and the risk-averse strategy, namely, the cost for saving risk points while still accomplishing one's task.

The objective of a player in the game is to maximize his payoffs. An optimal strategy for a player is contingent on the strategy of the other player. When both players strategies are optimal with regards to their counterparts, their interactions are "xed" in a way that none of them has the incentive to change to another strategy. Such a strategy pair is called a Nash Equilibrium. A Nash Equilibrium offers a credible prediction of the users moves, as it gives the user the best he can get given the administrators strategy. It also identifies the system administrators best countermeasure to the users strategy.

In our game, when the administrator chooses not to reward the user whose budget is not empty then the users best response is the "risk-seeking" strategy. This explains the reasoning of the results of Experiment one. When the administrator chooses the "reward" action, the user will choose the "risk-seeking" strategy if $R_1 > R_2 - C$, otherwise he will choose the "risk-averse" strategy. As we explained in previous section, the friction C is small. Thus $R_1 < R_2 - C$ and the users optimal strategy is the "risk-averse".

Interestingly, in this game, the Nash equilibrium is *(No reward, Risk seeking)*. Such an outcome, however, is not in the organization's interest as there is a result *(reward,*

Risk averse) giving it a better payoff $-P_2 - R_2$ when $R_2 < P_1 - P_2$. This situation is similar to the classic prisoner's dilemma game [26], where the Equilibrium does not offer players desirable payoffs. This dilemma can be avoided when the game is played repeatedly, which makes *(reward, Risk averse)* part of an equilibrium strategies: this is because the organization knows if it does not reward the users this round, they will take the path of risk seeking next time.

6.2 *Application of our mechanism*

From the game theoretic analysis, we can see that in order to make the mechanism work the inequality, $R_1 < R_2 - C$, must hold. Therefore it is critical to determine the parameters of the risk budget mechanism before it can be applied to a practical scenario.

As described in previous section, the friction C can be measured in the time interval. In practice, this friction could be estimated from observed time differences between taking different paths to accomplish the same task. Another way to parameterize our mechanism is to adjust the reward functions and monitor the risks brought in by users' activities, until the distribution of risk behaviors becomes acceptable.

7 Conclusion and Future work

Inadvertent insider poses a grave security threat to the security of organizations. To mitigate this threat, we proposed in this paper a novel risk budget mechanism that encourages insiders to behave responsibly. Our mechanism assigns individual users a risk budget, which represents the amount of risks an organization can tolerate to let its employees accomplish their tasks. Each action of a user will cost him certain risk points. Once the budget is depleted and the user does not finish his work, a big penalty ensues. On the other hand, those who diligently seek the path that reduces the organization's risk, which is manifested from the surplus of their budget, will be rewarded. Our experimental study shows that our approach exerts significantly impacts to rational users' risk attitudes, and evidently shifts their behaviors from risk seeking to risk averse. In the future, we plan to study the effectiveness of our approach beyond the scenario of web browsing, and explore the possibility of combining the idea of risk budgeting with existing access control mechanisms.

References

1. NASCIO. State cios take action now! Technical report, National The Association of State Chief Information Officers, 2007.
2. CSO. The 2007 ecrime watch survey. Technical report, the U.S. Secret Service, Carnegie Mellon University Software Engineering Institute's, 2007.
3. Homeland defense journal, 2007.
4. Report to the nation on occupational fraud and abuse. Technical report, Association of Certified Fraud Examiners, Inc, 2006.
5. Richard Zeckhauser. Behavioral versus rational economics: What you see is what you conquer. *Journal of Experimental Psychology*, 59(4):435–449, October 1986.

6. David Gefen. E-commerce: the role of familiarity and trust. *The International Journal of Management Science*, 28:725–737, 2000.
7. S. Stolfo, S. Bellovin, S. Hershkop, A. Keromytis, S. Sinclair, and S. Smith, editors. *Insider Attack and Cyber Security: Beyond the Hacker*. Springer, 2008.
8. Alessandro Acquisti and Ralph Gross. Imagined communities: Awareness, information sharing and sharing on facebook. In *Privacy Enhancing Technologies*, volume 4258, pages 36–58. Springer, June 2006.
9. Marisa Reddy Randazzo, Dawn M. Cappelli, Michelle M. Keeney, Andrew P. Moore, and Eileen F. Kowalski. Insider threat study: Illicit cyber activity in the banking and finance sector. Technical report, 2004.
10. Farzaneh Asgharpour, Debin Liu, and L. Jean Camp. Mental models of security risks. In Sven Dietrich and Rachna Dhamija, editors, *Financial Cryptography*, volume 4886 of *Lecture Notes in Computer Science*, pages 367–377. Springer, 2007.
11. L. Jean Camp. Mental models of computer security. In *FC: International Conference on Financial FC: International Conference on Financial Cryptography*. LNCS, Springer-Verlag, 2007.
12. L. Jean Camp. Net trust: Signaling malicious web sites. *I/S A Journal of Law and Policy in the Information Socirty*, 3(2):211–235, 2007.
13. J. Kesan and R. Shah. Establishing software defaults: Perspectives from law, computer science, and behavioral economics. *The Notre Dame Law Review*, 82(2):583–634, 2006.
14. R. Adkins. An insurance style model for determining the appropriate investment level. In *Third Workshop on the Economics of Information Security*, Minneapolis, MN, 2004.
15. E Karofsky. Return on security investment: calculating the security investment equation. In *Secure Business Quarterly*, volume 1, 2001.
16. C. Masone and S.W. Smith. Towards usefully secure email. In *IEEE Technology and Society (Special Issue on Security and Usability)*, volume 26, pages 25–34. Springer, 2007.
17. Nathaniel Good, Jens Grossklags, David Thaw, Aaron Perzanowski, Deirdre K. Mulligan, and Joseph Konstan. User choices and regret: Understanding users' decision process about consensually acquired spyware. *I/S: A Journal of Law and Policy for the Information Society*, 2(2), January 2006.
18. Ian Goldberg, Austin Hill, and Adam Shostack. Trust, ethics and privacy. *Boston University Law Review*, 81:407–422, 2001.
19. Lorrie Faith Cranor and Simson Garfinkel. *Security and Usability*. O'Reilly, Cambridge, MA, 2005.
20. Urszula Chajewska, Daphne Koller, and Ronald Parr. Making rational decisions using adaptive utility elicitation. In *Proceedings of the 7th Conference on Artificial Intelligence (AAAI-00) and of the 12th Conference on Innovative Applications of Artificial Intelligence (IAAI-00)*, Menlo Park, CA, 2000. AAAI Press.
21. A. Yemini, D. Dailianas, Florissi, and G. Huberman. Marketnet: Market-based protection of information systems. In *The 12th Int. Symp. on Dynamic Games and Applications*, 2006.
22. MITRE Corporation. Horizontal integration: Broader access models for realizing information dominace. Technical Report JSR-04-132, JASON Defense Advisory Panel Reports, 2004.
23. I. Molloy, P. Cheng, and P. Rohatgi. Trading in risk: Using markets to improve access control. In *New Security Paradigms Workshop*, Olympic, California, September 2008. Applied Computer Security Associates.
24. Pau-Chen Cheng, Pankaj Rohatgi, Claudia Keser, Paul A. Karger, Grant M. Wagner, and Angela Schuett Reninger. Fuzzy multi-level security: An experiment on quantified risk-adaptive access control. In *IEEE Symposium on Security and Privacy*, pages 222–230, 2007.
25. Alex Tsow, Camilo Viecco, and L. Jean Camp. Privacy-aware architecture for sharing web histories. *IBM Systems Journal*, 2007.

26. Martin J. Osborne and Ariel Rubenstein. *A Course in Game Theory*. The MIT Press, Cambridge, Massachusetts, 1994.
27. Mcafee siteadvisor.