ABSTRACT

In previous simulation studies attackers were shown to respond to changes in reward with an S shaped curve and to changes in security with a declining S shaped curve. This paper reports further experimental work that investigates the whether the response to reward is relative or absolute and influenced the framing of the rewards.

KEYWORDS: Infosec, Security, Experimental Economics

INTRODUCTION

The primary purpose of this paper is to determine whether the response to reward in an information system under attack is absolute or is it influenced by the framing of the rewards received by the attackers as described by Tversky and Kahneman (1981).

LITERATURE REVIEW

Much of the research in computer security focused on the technological systems interactions. Early research surveyed by Browne (1976) found that there were thousands of papers on computer security and risk management, but most were too narrow in scope and too fixated on technological fixes to be of much value.

Udo’s (2001) survey of privacy and security concerns as related to e-commerce focused on how users perceived threats and concluded that while many IT users felt that security is a critical issue, they didn’t believe that the government or any technological fix is capable of securing their privacy. In those areas where the human interaction factor was described, the literature tended to focus on the technology side of the interaction rather than the human side. Cranor and Garfinkel (2005) made the point that overly complex passwords can hurt the overall effectiveness of password security. While Besar and Arief (2004) and Duggan, Johnson and Gravemeyer (2012) discussed the impairment of security by legitimate users, their description of the faults focused on the technical.

Sasse et.al. (2001) made the point that the human portion of the security problem is the area of highest leverage. Adams and Sasse (1999) stated that rather than avoiding investigating the human factors, we
need to embrace them. Mitnick and Simon (2005) noted that the importance of human factor is critical because it is the basis of many threats. Saltzer and Schroeder (1975) who recognized that humans play a role, focused mainly on the technological issues of security rather than the interactions of the system with its users and attackers.

Becker (1968) explored motivation and behavior relating to crime and economics, describing criminals as rational, but nor moral. Schram (2000) explored the motivations and behaviors of various actors in economic decision making.

Kagel and Roth (1995) and Plott and Smith (2008) reviewed the current state of the art in experimental economics and their work has significantly influenced the design of these experiments. Lum and Yang (2005) showed that there is general support for using experimental research to examine criminal behavior. Jonsson and Olovsson (1997) conducted an experiment using students as proxies for both attackers and users and makes a good case concerning the issue of students used as a good proxy for stakeholders in an information system.

Christin et. al. (2011) conducted experiments that incentivized users to run potential malware and concluded that most users will run potentially harmful software if the incentive is great enough.

EXPERIMENTAL DESIGN OVERVIEW

Experimental Methodology

This experiment was described in detail in Rounds, Pendegraft, and Alves-Foss (2013). That work supported the hypothesis that attacker’s response to changes in value and security is an “S” shaped curve. These experiments used students as proxies for systems attackers. Their “attack” was to guess a password. The difficulty of this task was varied by expanding the length of the password and the number of letters in the alphabet. This experiment is based on the work of Jonsson and Olovsson (1997) and Carbone and Geus (2004).

Each student was given the security task of guessing a password from a restricted alphabet. The length of the alphabet and the length of the password were known to the subject. The numbers of attempts and successes and failures were logged at each level.

Three different experiments were run to test different aspects of this theory. The first experiment kept the reward constant and increased the complexity by increasing the number of possible passwords. The analysis in this paper does not include these results because they could not be normalized to make a meaningful comparison.

The second experiment kept the task the same with 64 possible passwords but varied the reward starting at $1.00 and reducing with each success by $0.05. This group is referred to later in the paper as the standard reward experiment.

Methodology for the third experiment is exactly the same as the second except that instead of starting the payments at a dollar ($1.00) payments were started at $1.50. The results for these experiments have been published Rounds, Pendegraft, and Alves-Foss (2013) and will be referred to later in this paper as the high reward experiment.
Hypothesizes

The hypotheses for the analysis performed in this paper are described below. They attempt to ascertain whether or not framing is occurring when dealing with the perceptions of the attacker’s appreciation of the value of the task they are performing.

First Hypothesis

$H_0 = \text{The average number of attempts for the standard reward experiment equals the average number of attempts in higher reward experiment.}$

$H_1 = \text{The average number of attempts for standard reward experiment is less than the average number of attempts for the higher reward experiment.}$

Second Hypothesis

$H_0 = \text{The average reward at termination for the standard reward experiment equals the average number of attempts in the higher reward experiment.}$

$H_1 = \text{The average reward at termination for the standard reward experiment is less than the average number of attempts for the higher reward experiment.}$

RESULTS AND DISCUSSION

All Hypotheses that follow were tested using the Student t distribution and a confidence value of 95%. The critical value is determined using the degrees of freedom which are determined using the Welch-Satterthwaite Equation (Satterthwaite, 1948) and the level of confidence required which in most research is 95%. This is a one-tailed test as only the figures greater than the mean are relevant.

The data was normalized for comparisons between the higher reward experiment and the standard reward experiment. Only the attempts that are registered for rewards from $1.00 and lower we counted as that is the range for standard reward experiment.

Table 1 showed the results of the experiments as they relate to reward levels.

<table>
<thead>
<tr>
<th>Source</th>
<th>Average Number of Attempts</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Reward experiment</td>
<td>411.4</td>
<td>247.9</td>
</tr>
<tr>
<td>Higher Reward experiment</td>
<td>193.3</td>
<td>232.2</td>
</tr>
</tbody>
</table>
Table 2 shows the statistical reduction when the higher reward experiment was compared to the standard reward experiment. These figures showed a statistically significant difference for the higher reward experiment when compared to the standard reward experiment with a 95% confidence level.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Welch's student-t test</th>
<th>Critical Value</th>
<th>Df</th>
<th>df approx.</th>
<th>Reject the Null</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher reward experiment to Standard reward experiment</td>
<td>-4.586</td>
<td>1.664</td>
<td>95.85</td>
<td>80</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Further comparison showed that not only did higher rewards reduce the number of attempts people were willing to do when the reward was below $1.05 as shown in Table 1. Starting the rewards at a higher point also showed a significant increase in the point where participants decided that they reward is not sufficient to continue as shown in Table 3.

<table>
<thead>
<tr>
<th>Source</th>
<th>Average reward at refusal</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard reward experiment</td>
<td>0.4667</td>
<td>0.3358</td>
</tr>
<tr>
<td>Higher reward experiment</td>
<td>0.7766</td>
<td>0.4102</td>
</tr>
</tbody>
</table>

In the standard reward experiment, the stopping point was approximately 46.67 cents while the high reward experiment averaged 77.66 cents which was significant at the .05 level as shown in Table 4.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Welch's student-t test</th>
<th>Critical Value</th>
<th>Df</th>
<th>df approx.</th>
<th>Reject the Null</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher reward experiment to Standard reward experiment</td>
<td>4.144</td>
<td>1.664</td>
<td>91.02</td>
<td>80</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**CONCLUSIONS**

It appears that the perception of reward is relative. This is in line with the concept of framing as described by Tversky and Kahneman (1981) and shows that understanding the target selection process that systems attackers use is more complex than previously realized. CIO’s and other security personnel deal with
determining how to deploy scarce security resources. Practical applications for CIO’s are relevant when deciding how to present the value of their systems to investors, customers, and by extension to the attacker population.

This is significant in that these factors can be used to model the behavior of attackers as they try exploit a given computer resource. These findings seem to validate the major assumptions discussed earlier and in our previously cited models.

These results can be used to determine some of the characteristics of attackers and selectly harden likely targets or re-educate users to make more efficient use of a scarce security budget.

Modeling of attackers also holds potential for more accurate penetration testing and more realistic security testing in general.

**Future Work**

This methodology, which could help in characterizing system attackers as criminals, is relatively new to computer security but it does open up many new techniques and insights into this area. An experiment could be constructed to examine the effects of enforcement by extracting a monetary penalty if a participant were to be discovered while attempting one of these simulated attacks.

A variation on the experiment above would create an environment where the perceived difficulty was different from the actual difficulty or reward and measure the same data. The goal here would be to determine if it is perception of difficulty or actual difficulty that forms the basis of attackers’ decisions as to which sites are targeted.

A further line of study would be to develop a taxonomy of targets and determine if there is a reason that a certain class of hacker would attack any specific group. This might be facilitated with Honey Pots or a laboratory experiment similar to this one.

**References**


Welch B., (1938), The significance of the difference between two means when the population variances are unequal. Biometrika 29:350–62.