Economic Analysis of the Market for Software Vulnerability Disclosure

Karthik Kannan[1]

Assistant Professor,

Purdue University, West Lafayette, IN.

kkarthik@cmu.edu

Rahul Telang

Assistant Professor,

Carnegie Mellon University, Pittsburgh, PA.

rtelang@andrew.cmu.edu

[1]Contact Author

Economic Analysis of the Market for Software Vulnerability Disclosure

Abstract

Software vulnerability disclosure has been a critical area of concern for policy makers. Traditionally, Computer Emergency Response Team (CERT) has been acting as an *infomediary* between *benign identifiers* who report vulnerability information and users of the software. After verifying a reported vulnerability, the infomediary – CERT – sends out a public "advisory" to inform software users about it. In this traditional mechanism, reporting vulnerabilities is voluntary with no explicit monetary gains to benign identifiers. Of late, firms such as iDefense have been proposing a different market-based mechanism. In this *market-based mechanism*, the infomediary rewards identifiers for each vulnerability disclosed to it. The infomediary then shares this information with its clients who are users of this software. Using this information, clients can protect themselves against attacks that exploit those specific vulnerabilities.

The key question addressed in our paper is whether movement towards such a market-based mechanism for vulnerability disclosure leads to a better social outcome. Our paper employs game-theoretic models to provide insights into this important problem. We also extend the model to compare the welfare-effects of other disclosure mechanisms. Based on our analysis, we find that under certain conditions, the market-based infomediary generates a lower industry loss than a CERT-type one and vice-versa. Also, we observe that a Federally-Funded Social Planner always performs at least as well as other mechanisms. Intuitions provided for these results aid a market-planner in gaining a better understanding of the implications of these different mechanism designs.

Acknowledgements: We wish to thank Sunil James, Manager, Vulnerability Contributor Program, iDefense.com for his support. Also, we thank Xu Hao for his contribution.

1 Introduction

Software vulnerability disclosure has been a critical area of concern for policy makers. Traditionally, Computer Emergency Response Team (CERT) has been acting as an *infomediary* between *benign identifiers* who report vulnerability information and users of the software. After verifying a reported vulnerability, the infomediary – CERT – sends out a public "advisory" to inform software users about it. In order to ensure that such public notifications are not exploited by *hackers*[2] to attack software users, CERT follows a series of steps. The steps include contacting the vendor for the appropriate patch, publicly disclosing the vulnerability after waiting an optimal time, etc. In this traditional mechanism, reporting vulnerabilities is voluntary with no explicit monetary gains to benign identifiers. Of late, firms such as iDefense[3] have been proposing a different market-based mechanism. In this *market-based mechanism*, the infomediary rewards identifiers for each vulnerability disclosed to it. The infomediary then shares this information with its clients who are users of this software. Using this information, clients can protect themselves against attacks that exploit those specific vulnerabilities.

The key question addressed in our paper is whether movement towards such a market-based mechanism for vulnerability disclosure leads to a better social outcome. The answer to this question is not obvious. On one hand, incentives to discover and disclose vulnerabilities may lead to benign identifiers investing more effort and time in finding them i.e., leading to better security. But on the other hand, the same incentives may also lead to a *race* for vulnerability discovery[4] between

[2]The information security space considers both malign and benign intentioned vulnerability identifiers as hackers. We however, employ a narrower definition in this paper and refer to "hackers" only as people who exploit vulnerabilities to attack.

[3]www.iDefense.com

[4]Similar behavior is observed in R&D competitions where firms race to be an innovator (See Dasgupta & Stiglitz (1980) and Reinganum (1982)).

benign identifiers and hackers. If this happens and the number of vulnerabilities discovered by the hacker increases, it may lead to a lesser desirable social outcome.

We investigate this question using a game-theoretic model. We extend the model to analyze and compare the welfare-effects of other mechanisms as well (the mechanisms are described below). In each mechanism, the infomediary decides on the following two variables – the reward it pays for each vulnerability disclosed to it and the one-time subscription fee it charges its clients. The only difference among these mechanisms is the objective function that the infomediary optimizes:

- Market-Based Mechanism (e.g. iDefense): In this mechanism, the infomediary maximizes its expected profits by choosing the optimal reward price and the optimal one-time subscription fee.

- CERT-Type Mechanism: This corresponds to the traditional mechanism followed for vulnerability disclosure. In this mechanism, no monetary benefits are provided to the identifier and no subscription fee is charged to software users.

- Consortium Mechanism: This is one of two mechanisms we propose in this paper and it does not have any real-world equivalents. In this mechanism, we assume the infomediary to be a zero-profit one which uses the payment from its clients to pay for vulnerabilities disclosed to it. The infomediary sets the optimal reward price and the optimal one-time subscription fee in a manner that minimizes the loss suffered by its clients.

- Federally Funded Social Planner: This is the second model we propose in this paper. We intend to use this mechanism as a reference against which we compare all other mechanisms. The federally-funded infomediary is assumed to set the reward for vulnerability disclosure at the level that maximizes social outcome. However no subscription fee is charged to users.

Notifications about vulnerabilities are sent out to all users similar to the CERT-type mechanism.

The above-mentioned mechanisms are compared on the following two metrics which we commonly refer to as *welfare-parameters*: the total loss suffered by the users and the total loss suffered by the industry. Based on our analysis, we find that under certain conditions, the market-based mechanism generates a lower industry loss than a CERT-type one and vice-versa. Also, we observe that the Federally-Funded mechanism always performs at least as well as other mechanisms on both these welfare parameters. Intuitions provided for these results can aid a market-planner in gaining a better understanding of the implications of these different mechanisms.

The paper is organized as follows. In section 2, we review the literature most relevant to this topic. Following that in section 3, we model the four mechanisms and compare their welfare-parameters. Finally in section 4, we conclude.

2 Literature Review

Most prior work in the software vulnerability and information security area has focused on the technical aspects of the problem. But in this section, we restrict our attention to papers that address "non"-technical issues. We begin with the paper by Krsul *et al.* (1998) which analyzes five common vulnerabilities. For each vulnerability, Krsul *et al.* (1998) identify the characteristics and describe the policies that get violated when that vulnerability is exploited. Their analysis contributes to the understanding of the steps needed to eradicate these vulnerabilities. Du & Mathur (1998a) take this classification one step further. They categorize and analyze software errors that led to security breaches. Based on their classification schemes, they also develop testing techniques in Du &

Mathur (1998b) that can identify security errors.

When such techniques are incorporated in the software development processes, Krishnan *et al.* (2000) and Banker *et al.* (1998) argue that software quality improves and the software has lesser number of vulnerabilities. Although these methods and processes are useful in improving software quality, it is widely believed that vulnerabilities and therefore, attacks exploiting these vulnerabilities cannot be completely eliminated.

Given this, a few papers have analyzed related problems in the information security space. Gordon & Loeb (2002) develop an economic model for information security investment decisions. They claim that the optimal information security spending does not always increase with the expected loss from attacks and that the optimal security spending has to be far less than the expected loss from attacks. They provide intuitions for this counter-intuitive result and validate their claims using empirical data.

Similarly, Arora *et al.* (2003) develop an economic model to study a vendor's decision of when to introduce its software and whether or not to patch vulnerabilities in its software. They compare the decision process of a social-welfare maximizing monopolistic vendor to that of a profit-maximizing monopolistic vendor. Interestingly, they observe that the profit-maximizing vendor delivers a product that has lesser vulnerabilities than a social-welfare maximizing vendor's. However, the profit-maximizing vendor is less willing to patch its software than its social-welfare maximizing counterpart. Arora *et al.* (2003) provide intuitive explanation for these results.

To our knowledge, no prior work has addressed issues related to the current imbroglio. Practitioners in different capacities have been proposing different legal/economic frameworks for software vulnerability disclosure (Security-Focus (2003), e Week (2003)), ZD-Net (2003)). A few researchers have suggested other mechanisms as well. For example in a New York Time article,

Varian (2000) suggests that information security can be improved by first assigning legal liability. Along with a legal framework, he argues that an insurance framework can provide the correct market-based incentive structure. In this current scenario, policy-makers are left with little guidance in understanding the implications of these proposed frameworks. Before policy-makers explore these proposed frameworks, they need a better understanding of the implications of existing frameworks and this is the contribution of our paper. Our paper employs game-theoretic models to provide insights into the welfare-effects of existing disclosure frameworks.

3 Model

There are four types of participants in our model – the infomediary, a benign identifier, a hacker and software users. We are interested in comparing the welfare-effects when the mechanism adopted by the infomediary changes. We model each of the four mechanisms as a two-period game. In the first period, the infomediary sets its pricing policy to maximize its objective function and in the second period, all other players – software users, the benign identifier and the hacker – react.

Let p_h be the reward that the infomediary is willing to pay for each vulnerability reported. Let p_s be the one-time subscription fee that the infomediary charges each of its clients – software users. These prices determine the number of subscribers (and hence the market share) to this service, number of vulnerabilities reported and the probability of attacks. Thus, they are a determinant of the welfare-parameters – the total industry loss and the total user loss. To analyze this, we begin with the second period game in section 3.1. Specifically, we model the reactions of software users, the benign identifier and the hacker to any (p_h, p_s) pair set by the infomediary. In the same subsection, we also express our welfare-parameters as functions of p_h and p_s. In each of the

following subsections, we model the first-period game for each mechanism and then compute the optimal (p_h, p_s) pair, taking into account the reactions of other players. The optimal pair for each mechanism will drive the welfare-parameters which are compared in the last subsection.

3.1 Modeling Software Users, The Benign Identifier and The Hacker

In this section, we model the reactions of software users, the benign identifier and the hacker to any (p_h, p_s) pair set by the infomediary. We begin by labeling the parameters of interest to us. Let N_c represent the number of users subscribed to the infomediary's service. This is dependent on the one-time subscription fee p_s charged by the infomediary. Recall that the infomediary acquires the vulnerability information by paying a reward of p_h for each vulnerability reported. This reward incentivizes a benign identifier to exert effort, discover vulnerabilities and report them. Without loss of generality, we deal with probability values of discovering one vulnerability instead of dealing with numbers of vulnerabilities discovered.

Let K_r be the probability that the benign identifier discovers and reports the vulnerability to the infomediary. After obtaining the vulnerability information, the infomediary notifies[5] its clients so that they can protect their systems against potential future attacks. Let the probability that the attack is prevented be K_p. This K_p corresponds to the probability that the vulnerability reported by the benign identifier is discovered later by the hacker. In such a case, the hacker can exploit the vulnerability to attack those users that are not subscribed to the infomediary's service. Sometimes the hacker may discover the vulnerability first. Let K_h be the probability that the vulnerability is first discovered by the hacker. In this case, the hacker exploits the vulnerability to attack all users. All these probabilities – K_r, K_p, and K_h – are dependent on the effort level exerted by the

[5]The infomediary may also provide value added services such delivering the patch for the vulnerability, filters to protect against attacks that exploit the vulnerability etc.

benign identifier and the hacker which are in turn, driven by p_s and p_h set by the infomediary. Our objective in this section, is to express these probabilities and N_c as functions of p_s and p_h. We intend to use these expressions to compute the optimal p_h and p_s under each mechanism.

3.1.1 Software Users

We assume that software users are heterogeneous in terms of the loss they incur when a vulnerability is exploited. Let the user "loss"-type, θ, be distributed uniformly between $[0, \overline{\theta}]$. Any software user i of type θ_i is assumed to incur a loss of θ_i^2 when the vulnerability is exploited. The non-linear choice of the loss function reflects the empirical observations quite well – many users suffer smaller losses while only a few users suffer huge losses. The software users have an option of preventing attacks on their systems by subscribing to the infomediary's service. Any user i, whose expected profit from subscribing

$$\Pi_{\text{user}} = \theta_i^2 \; K_p - p_s > 0 \tag{1}$$

subscribes to the service. In this expression, the first term corresponds to the loss prevented by subscribing to the service. The second term corresponds to the payment made to the infomediary. Rearranging the terms, we can state that only those software users whose θ_i satisfies the following condition subscribe to the service:

$$\theta_i \; > \sqrt{\frac{p_s}{K_p}}$$

Since θ is assumed to be uniformly distributed between $[0, \overline{\theta}]$, the number of clients subscribed to the infomediary's service is:

$$N_c = \overline{\theta} - \sqrt{\frac{p_s}{K_p}} \tag{2}$$

Consider the CERT-type mechanism or the Federally Funded mechanism where software users are not charged any price at all i.e., $p_s = 0$. In such a case, $N_c = \overline{\theta}$ which implies all users are provided with vulnerability information.

3.1.2 Competition Between the Benign Identifier and the Hacker

In this section, we are interested in obtaining the probabilities – K_r, K_p and K_h – as functions of p_h and p_s. As a first step, we express these probability values in terms of effort levels of the benign identifier and the hacker. Then in the second step, we compute the optimal effort level exerted by them based on their respective profits.

Let α be the effort exerted by the benign identifier and β be the effort exerted by the hacker. Then, any functional form for K_r, K_p and K_h which satisfies the following properties is sufficient for our analysis:

- $\frac{\partial K_r}{\partial \alpha} > 0$: The probability that the vulnerability is reported increases with the benign identifier's effort.
- $\frac{\partial K_r}{\partial \beta} < 0$: The probability that the vulnerability is reported decreases with the hacker's effort.
- $K_p + K_h \leq 1$ is the probability the vulnerability is discovered by the hacker.
- $\frac{\partial K_h}{\partial \alpha} < 0$: The probability that the hacker discovers the vulnerability first and exploits it decreases with the benign identifier's effort.

- $\frac{\partial K_h}{\partial \beta} > 0$: The probability that the hacker discovers the vulnerability first and exploits it to attack increases with the hacker's effort.

- $\frac{\partial K_p}{\partial \alpha} > 0$: The probability that the attack due to the vulnerability is prevented increases with the benign identifier's effort.

- $\frac{\partial K_p}{\partial \beta} \geq 0$: The probability that the attack due to the vulnerability is prevented which is the same as the probability that the hacker discovers the vulnerability after the benign identifier, never decreases with increase in the hacker's effort.

In the next few paragraphs, we demonstrate one way of expressing these probabilities which we use for comparing the different mechanisms.

The competition between the benign identifier and the hacker is modeled by considering the timeframe for the software's life cycle as T. Within this period, we assume that the probability that a player – the benign identifier or the hacker – discovers a vulnerability is distributed uniformly. Note that we are simply characterizing the probability of discovering the vulnerability and not the probability conditioned on a player discovering it first. Let γ, an exogenous parameter, correspond to the random independent probability for each player to discover the vulnerability before the end of the time period without exerting any effort. Given our distributional assumption, $\frac{\gamma}{T}$ is the probability that the vulnerability is discovered by a player at each instant without exerting any effort. Players can alter this probability value by exerting effort.

Let the benign identifier and the hacker exert efforts that increase their probabilities of finding the vulnerability before the time period, T, to $\alpha+\gamma$ and $\beta+\gamma$ respectively. α and β are determined by the Nash-equilibrium of the competition between them. These parameters – α and β – are assumed to be set for the entire duration, T, and cannot be modified during the game. Because of

their efforts, the different probability values are computed as follows:

- The probability that the vulnerability is reported – K_r – corresponds to the probability that the vulnerability is first discovered by the benign identifier and reported to the infomediary.

$$K_r = \int_0^T \text{Probability}(benign = t)\ \text{Probability}(\overline{hacker} < t)\ dt$$

Probability$(benign = t)$ is the probability that the vulnerability is identified by the benign identifier at time t and Probability$(\overline{hacker} < t)$ is the probability that the vulnerability is *not* identified by the hacker any time before t. Therefore,

$$\begin{aligned} K_r &= \int_0^T \frac{\alpha+\gamma}{T}\left(1 - \frac{(\beta+\gamma)t}{T}\right)\ dt \\ &= (\alpha+\gamma)\left(1 - \frac{(\beta+\gamma)}{2}\right) \end{aligned} \tag{3}$$

- K_p is the probability that an attack exploiting the vulnerability is prevented by subscribing to the infomediary's service. Recall that K_p also corresponds to the probability that the hacker discovers the vulnerability after the benign identifier:

$$\begin{aligned} K_p &= \int_0^T \text{Probability}(hacker = t)\ \text{Probability}(benign < t)\ dt \\ &= (\alpha+\gamma)\frac{(\beta+\gamma)}{2} \end{aligned} \tag{4}$$

- Finally, the probability that the vulnerability is first discovered by the hacker – K_h – is

$$K_h = \int_0^T \text{Probability}(hacker = t)\ \text{Probability}(\overline{benign} < t)\ dt$$

$$= \quad (\beta + \gamma)\left(1 - \frac{(\alpha + \gamma)}{2}\right) \tag{5}$$

Having defined these probabilities, we characterize the optimal efforts exerted by the benign identifier and the hacker. For this, we consider their respective expected profit functions. Recall that the effort exerted by the benign identifier increases its probability of finding the vulnerability to $\alpha + \gamma$. This effort pays p_h, if the benign identifier discovers the vulnerability before the hacker. Since K_r is the probability that the benign identifier discovers the vulnerability first, the expected revenue for the benign identifier is given by $p_h\ K_r$. Corresponding to its effort, the benign identifier's cost is $C(\alpha)$, a function of α. Mathematically, the expected profit for the benign identifier is:

$$\Pi_b = K_r\ p_h - C(\alpha)$$

For obtaining an interior optimal solution, we require that Π_b be concave with respect to α. Since the revenue increases linearly with α, any convex cost function will suffice. We assume that the cost, $C(\beta) = M\alpha^2$ where M is an exogenous constant parameter. Note that quadratic functions are commonly used in literature for such scenarios. Substituting for $C(\alpha)$ and K_r,

$$\Pi_b = (\alpha + \gamma)\left(1 - \frac{(\beta + \gamma)}{2}\right) p_h - M\ \alpha^2 \tag{6}$$

Next, let us consider the hacker's expected profit. The hacker benefits by attacking all users if he discovers the vulnerability first. But if he discovers the vulnerability after the benign identifier,

he obtains the profit only from attacking users not part of the infomediary's clientele[6]. We assume that if the hacker is successful in attacking a user of type θ_i, he gains a profit of θ_i. Note that the functional form of the hacker's profit function is intentionally made to be different from the loss for the user – θ_i^2. The hacker's cost is $C(\beta)$. Therefore,

$$\Pi_h = K_h \left(\int_0^{\overline{\theta}} \theta \, d\theta \right) + K_p \left(\int_0^{(\overline{\theta} - N_c)} \theta \, d\theta \right) - C(\beta) \tag{7}$$

In the first term, K_h corresponds to the probability that the hacker discovers the vulnerability before the benign identifier and attacks all the users. The term inside the integral is the expected profit that the hacker obtains from attacking all the users. Similarly in the second term, K_p corresponds to the probability that the hacker discovers the vulnerability after the benign identifier. The integral in the second term corresponds to the expected profit that the hacker can secure by attacking users that are not part of the infomediary's clientele. The last term corresponds to the cost of exerting effort. Substituting for K_h and K_p, and simplifying the expression, we have

$$\Pi_h = (\beta + \gamma) \left(1 - \frac{(\alpha + \gamma)}{2} \right) \frac{\overline{\theta}^2}{2} + \frac{(\beta + \gamma)(\alpha + \gamma)}{2} \frac{(\overline{\theta} - N_c)^2}{2} - M \, \beta^2 \tag{8}$$

Substituting for N_c from equation 2, we have the expected profit for the hacker as

$$\Pi_h = (\beta + \gamma) \left(1 - \frac{(\alpha + \gamma)}{2} \right) \frac{\overline{\theta}^2}{2} + \frac{(\beta + \gamma)(\alpha + \gamma)}{2} \frac{p_s}{K_p \, 2} - M \, \beta^2 \tag{9}$$

[6]We assume that the hacker never finds it optimal to sell the vulnerability.

We can simplify this expected profit function further, by substituting for K_p. Therefore,

$$\Pi_h = (\beta + \gamma)\left(1 - \frac{(\alpha + \gamma)}{2}\right)\frac{\overline{\theta}^2}{2} + \frac{p_s}{2} - M\,\beta^2 \tag{10}$$

These expected profit expressions are used to determine the optimal values for α and β. To obtain the optimal level of effort for the benign identifier – α – we differentiate the benign identifier's expected profit expression i.e., equation 6, with respect to α and equate it to zero. Thus,

$$\alpha^* = \left(1 - \frac{\beta + \gamma}{2}\right)\frac{p_h}{2\,M} \tag{11}$$

Similarly to obtain the optimal effort level for the hacker – β – we differentiate the hacker's expected profit expression i.e., equation 10, with respect to β and set it to zero.

$$\beta^* = \left(1 - \frac{\alpha + \gamma}{2}\right)\frac{\overline{\theta}}{4\,M} \tag{12}$$

Solving the simultaneous equations – equation 11 and equation 12, we get:

$$\alpha^* = \frac{(8\,M - \overline{\theta}^2)\,p_h\,(2 - \gamma)}{32\,M^2 - p_h\,\overline{\theta}^2} \tag{13}$$

$$\beta^* = \frac{(2 - \gamma)(4\,M - p_h)\overline{\theta}^2}{32\,M^2 - p_h\,\overline{\theta}^2} \tag{14}$$

These expressions are valid only for $M > M_{th} = \frac{\overline{\theta}^2}{4(2-\gamma-\sqrt{2-(2-\gamma)\gamma}}$. For $\frac{\overline{\theta}^2}{8(1-\gamma)} < M < M_{th}$, α from this equation may be greater than $1 - \gamma$. Since $\alpha + \gamma > 1$ does not make sense, we set the corresponding $\alpha^* = 1 - \gamma$ and $\beta^* = \frac{1}{8\,r}$. Similarly $M > \frac{\overline{\theta}^2}{8(1-\gamma)}$, both α and β may be greater than $1-\gamma$ in which case, we set $\alpha^* = \beta^* = 1 - \gamma$. The only interesting case to study is when $M > M_{th}$

which is the focus of the rest of this paper.

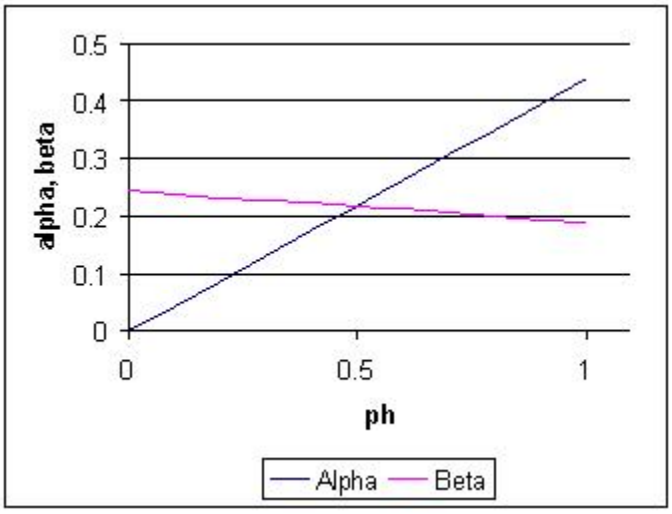

Figure 1: Shows the variation of the difference α and β with p_h.

For $M > M_{th}$, we observe the following properties in these equations: a) Both these expressions are independent of p_s, the one-time fee charged by the infomediary. b) As p_h increases, α increases but β decreases. Figure 1 captures the variation of α and β with p_h for $\gamma = 0.05$, $M = 1$ and $\overline{\theta} = 1$. Intuitively, the effort exerted by the benign identifier increases with p_h. This imposes a negative externality on the hacker's incentive. c) For a given p_h, both the benign identifier and the hacker have an incentive to increase their efforts as γ decreases. d) Finally, as M increases, i.e., the cost of exerting effort increases, the optimal effort levels – α^* and β^* – decrease as expected.

Using α^* and β^*, we can also compute the following probabilities:

$$K_p = \frac{2(16\gamma M^2 + 8Mp_h - 4\gamma Mp_h - p_h\overline{\theta}^2)(16\gamma M^2 + 4M\overline{\theta}^2 - 2\gamma M\overline{\theta}^2 - p_h\overline{\theta}^2)}{(32M^2 - p_h\overline{\theta}^2)^2} \quad (15)$$

$$K_r = \frac{4(2-\gamma)M(8M - \overline{\theta}^2)(16\gamma M^2 + 8Mp_h - 4\gamma Mp_h - p_h\overline{\theta}^2)}{(32M^2 - p_h\overline{\theta}^2)^2} \quad (16)$$

$$K_h = \frac{8(2-\gamma)M(4M - p_h)(16\gamma M^2 + 4M\overline{\theta}^2 - 2\gamma M\overline{\theta}^2 - p_h\overline{\theta}^2)}{(32M^2 - p_h\overline{\theta}^2)^2} \quad (17)$$

From these equations, we observe the following properties to be valid:

- As the incentive to disclose the vulnerability – p_h – increases, the probability that the vulnerability is reported increases i.e., $\frac{\partial K_r}{\partial p_h} > 0$, the probability that the attack due to the vulnerability is prevented increases i.e.,$\frac{\partial K_p}{\partial p_h} > 0$ and the probability that the hacker discovers the vulnerability first decreases p_h i.e., $\frac{\partial K_h}{\partial p_h} < 0$.

- As the random probability that the bug is discovered – γ – increases, all three probabilities increase i.e., $\frac{\partial K_r}{\partial \gamma} > 0$, $\frac{\partial K_p}{\partial \gamma} > 0$, $\frac{\partial K_h}{\partial \gamma} > 0$.

- As the cost of effort – M – increases, all three probabilities decrease i.e.,$\frac{\partial K_r}{\partial M} < 0$, $\frac{\partial K_p}{\partial M} < 0$, $\frac{\partial K_h}{\partial M} < 0$.

3.1.3 Welfare Parameters

Given these behaviors, we can compute the following welfare parameters – Total User Loss and Total Industry Loss. The total user loss is:

$$UL = K_h \left(\int_0^{\overline{\theta}} \theta^2 \, d\theta \right) + K_p \left(\int_0^{(\overline{\theta} - N_c)} \theta^2 \, d\theta \right) + N_c \, p_s \tag{18}$$

The first expression corresponds to the loss incurred when the hacker discovers the vulnerability first. The second expression corresponds to the loss incurred when the hacker discovers the vulnerability after the benign identifier. Since the benign identifier reports the vulnerability to the infomediary, the hacker is left to attack only those users who are not part of the infomediary's clientele. The last term corresponds to the total payment made by all subscribing users to the infomediary. Substituting for different parameters, one can compute the total user loss as a function of p_h and p_s.

Similarly, we compute the total industry loss as the loss incurred by the users in addition to the profit/loss incurred by the infomediary. Equation 18 corresponds to the loss incurred by the users. This combined with the infomediary's profit equates to

$$IL = K_h \left(\int_0^{\bar{\theta}} \theta^2 \, d\theta \right) + K_p \left(\int_0^{(\bar{\theta} - N_c)} \theta^2 \, d\theta \right) + K_r \, p_h \tag{19}$$

When we compute the industry profits, the term $N_c \; p_s$ which appears in equation 18 does not appear in equation 19. This is because $N_c \; p_s$ is the transfer of rent between the users and the infomediary. Thus, the only remaining term is the expected payment made by the infomediary for vulnerability disclosure and it appears in equation 19. Substituting for different parameters, one can compute the total industry loss as a function of p_h and p_s. Note that we will be using these welfare parameters for comparing the following mechanisms.

3.2 Market-based Mechanism

A classic example of this type of infomediary is iDefense (www.iDefense.com). It purchases vulnerability information and notifies its clients about the vulnerability. Recall that the infomediary charges each of its clients a one-time fee of p_s and pays a reward of p_h for every vulnerability reported. These variables are the outcomes of the expected profit maximization function given by

$$\max_{p_s, p_h} N_c \; p_s - K_r p_h \tag{20}$$

The first term in the expression corresponds to the revenue that the infomediary generates from charging its clients p_s. The second term is the cost it incurs to pay for each vulnerability reported.

We substitute for N_c from equation 2. Therefore, the objective function is:

$$\max_{p_s, p_h} \left(\overline{\theta} - \sqrt{\frac{p_s}{K_p}} \right) p_s - K_r \, p_h \tag{21}$$

To find the optimal p_s, we differentiate the expected profit expression by p_s and set it to zero.

$$p_s^* \;=\; \frac{4 \, K_p \overline{\theta}^2}{9} \tag{22}$$

We set $p_s = p_s^*$ in equation 21, differentiate the expression with respect to p_h and equate it to zero to obtain the optimal p_h.

$$p_h^* = \frac{32M^2(108\gamma M^2 - 4\gamma M\overline{\theta}^3 - \overline{\theta}^5 + \gamma\overline{\theta}^5)}{-3456M^3 + 1728\gamma M^3 + 432M^2\overline{\theta}^2 - 108\gamma M^2\overline{\theta}^2 - 16M\overline{\theta}^5 + 4\gamma M\overline{\theta}^5 + \overline{\theta}^7} \tag{23}$$

As γ increases, p_h decreases. $p_h > 0$ only for $\gamma < \frac{\overline{\theta}^5}{108 \; M^2 - 4 \; M \; \overline{\theta}^3 + \overline{\theta}^5}$. In other words, the value for γ has to be lesser than the threshold value for this market to survive. Beyond this threshold value, the infomediary has no incentive to fund vulnerability disclosure and the market fails. The welfare-parameters for this mechanism – the total user loss, UL_{MKT}, and the total industry loss, IL_{MKT} – can be obtained by substituting p_h^* and p_s^* in equations 19 and 18.

3.3 CERT-type Mechanism

In this mechanism scheme, $p_h = 0$ i.e., no monetary incentives are provided to the identifier, and $p_s = 0$ i.e., users are not charged any subscription fees for vulnerability notification. All users are notified about the vulnerability. By virtue of our derivation in section 3.1, when $p_h = 0$, $\alpha = 0$ i.e., the benign identifier does not exert any effort at all. But, vulnerabilities are still discovered by

a benign identifier with a probability of γ. We assume that the benign identifier always reports the vulnerability to the infomediary.

For this mechanism, we are interested in computing the total industry loss, IL_{CERT}, and the loss incurred by the users, UL_{CERT}. But in this case, both losses are equal to one another since there is no transfer of payment. For $p_h = 0$ and $p_s = 0$, the losses are given by

$$UL_{\text{CERT}} = IL_{\text{CERT}} = \frac{1}{48\ M}(2-\gamma)((2-\gamma)\overline{\theta}^2 + 8\ M\gamma)\overline{\theta}^3 \tag{24}$$

3.4 Consortium Mechanism

Although currently, no equivalent framework exists, one can imagine Information Sharing and Analysis Centers (ISAC)[7] to execute such a mechanism. In this mechanism, the infomediary is assumed to maximize the welfare generated for all its clients. It does so by charging each client a one-time fee of p_s that is sufficient enough to pay for the vulnerabilities reported:

$$N_c\ p_s = K_r\ p_h \tag{25}$$

The left hand side of this expression is the income to the infomediary and the right hand side is the expected payment made by the infomediary for the vulnerability discovery. Substituting for N_c, we have

$$\left(\overline{\theta} - \sqrt{\frac{p_s}{K_p}}\right) p_s = K_r\ p_h$$

[7]The US federal government, under Presidential Decision Directive NSC-63, has encouraged the establishment of industry based Information Sharing and Analysis Centers (ISACs) to promote the disclosure and sharing of security information among firms. Currently, these ISACs are focused on gathering, analyzing and sharing information related to actual, as well as unsuccessful attempts at, security breaches. We envision their role to be broader for this model.

This is a constraint to the following objective function optimized by the infomediary:

$$\max_{p_h} \left(\int_{\overline{\theta}-N_c}^{\overline{\theta}} \theta^2 d\theta \right) K_p - N_c \, p_s$$

The first term corresponds to the loss prevented by this framework whereas the second term corresponds to the cost incurred by all clients.

We express this constrained optimization function used by the infomediary in the following manner:

$$\max_{p_h, p_s, L} \left(\int_{\overline{\theta}-N_c}^{\overline{\theta}} \theta^2 d\theta \right) K_p - N_c \, p_s + L(K_r \, p_h - N_c p_s) \tag{26}$$

where L is the Lagrange variable. We solve this function using Kuhn-Tucker method to obtain

$$p_s^* = K_p \left(\frac{2(L-1)\overline{\theta}}{(3L-2)} \right)^2 \tag{27}$$

$$p_h^* = \frac{32 r^2 \overline{\theta}^2 (48\gamma(2-3L)^2 r^2 - L(-3+4L)(1+\gamma(-1+4r))\overline{\theta})}{48(2-3L))^2 r^2 (4-\gamma+16(-2+\gamma)r) + L(-3+4L)(1+4(-4+\gamma)r)\overline{\theta}} \tag{28}$$

where we let r represent the ratio $\frac{M}{\overline{\theta}^2}$. L becomes a solution to a fifth order polynomial equation. Out of the five possible solutions, we are interested only in the value of L which makes $0 < N_c < \overline{\theta}$. That corresponds to $1 < L < 2$. For a given value of γ, M and $\overline{\theta}$, one can compute the optimal p_s and p_h values numerically. Table 1 shows these values for $M = 1$ and $\overline{\theta} = 1$ for different values of γ. Notice that the optimal value for p_h decreases with γ. Also for $M = 1$ and $\overline{\theta} = 1$, we observe that the infomediary has no incentive to fund vulnerability discovery for $\gamma \geq 0.0225$,

Similar to the CERT-framework, the industry loss IL_{MSP} and the loss incurred by the users UL_{MSP} are identical i.e., $UL_{\text{MSP}} = IL_{\text{MSP}}$. This is so because the transfer of payment is equal to the

γ	N_c	p_h	K_r	p_s
0.0025	0.592326629	0.014154304	0.007595424	0.000181501
0.005	0.633957278	0.012340251	0.009069785	0.000176549
0.0075	0.674749699	0.010477494	0.01052353	0.000163408
0.01	0.715589755	0.008583861	0.011963563	0.000143506
0.0125	0.757443127	0.006675769	0.013396185	0.000118069
0.015	0.801649409	0.004772678	0.014828781	8.82822E-05
0.0175	0.850522102	0.002904378	0.016272552	5.55711E-05
0.02	0.909855062	0.001141046	0.017753763	2.22694E-05
0.0225	na			

Table 1: Optimal parameter values for a Consortium Mechanism.

loss incurred by the infomediary. Since p_s and p_h are analytically intractable, one can compute these losses numerically.

3.5 Federally-Funded Social Planner

This subsection deals with the case when the infomediary pays for vulnerability disclosure but charges nothing to the users i.e., $p_s = 0$. The price that it is willing to pay for the vulnerabilities is a solution to the following optimization function

$$\max_{p_h} \left(\int_0^{\overline{\theta}} \theta^2 d\theta \right) K_p - K_r \, p_h \tag{29}$$

Solving this optimization function, we obtain the optimal p_h^*.

$$p_h^* = \frac{32M^2(48\gamma M^2 - 4\gamma M\overline{\theta}^3 - \overline{\theta}^5 + \gamma\overline{\theta}^5)}{-1536M^3 + 768\gamma M^3 + 192M^2\overline{\theta}^2 - 48\gamma M^2\overline{\theta}^2 - 16M\overline{\theta}^5 + 4\gamma M\overline{\theta}^5 + \overline{\theta}^7} \tag{30}$$

In this expression, p_h decreases with an increase in γ. $p_h > 0$ only for $\gamma < \frac{\overline{\theta}^5}{48\ M^2 - 4M\ \overline{\theta} + \overline{\theta}^5}$. Beyond this, the performance will be similar to that of CERT. Corresponding to these prices, we evaluate the total user loss, UL_{FED}, and total industry loss, IL_{FED}.

3.6 Comparison

We use results from the previous subsections to compare the four mechanisms. We study the sensitivity of the total user loss and the total industry loss to variations in each of γ, M and $\overline{\theta}$. We also provide intuitions for our observations. For simplicity of exposition, we will refer to the total industry loss as TIL and the total user loss TUL for the rest of this section.

Figure 2 shows the variation of the TIL to $\overline{\theta}$ when M is set to 1 and γ is set to 0. For the same values of M and γ, the variation of TUL to $\overline{\theta}$ is shown in Figure 3. Note that the lower the loss, the better the mechanism is. As $\overline{\theta}$ increases, i.e., as the maximum loss suffered by the user increases, the losses – TIL and TUL – increase under all four cases. This result conforms to our intuition. Next, when comparing the designs, we observe that the Federally-Funded mechanism – the design which maximizes welfare for the industry – performs better than all other mechanisms as expected. Also, at the set values of γ and M, the market-based mechanism and the consortium mechanism incur lesser losses than a CERT-type mechanism in both Figure 2 and Figure 3. We provide intuitions for this result while discussing the sensitivity of TUL and TIL to γ.

Figure 4 and Figure 5 respectively show the sensitivity of TIL and TUL to M for a set value of $\gamma = 0$ and $\overline{\theta} = 6$. We observe that as M increases, both TUL and TIL decrease. Intuitively as the cost of discovering vulnerabilities increases, the optimal effort level exerted by the hacker decreases which in turn, decreases the loss suffered by users. Note that similar to the earlier case, the Federally-Funded mechanism performs the best. Also, observe that the market-based mechanism and the consortium mechanism incur lesser losses than a CERT-type mechanism.

Next, consider Figure 6 and Figure 7 which respectively show the variations of TIL and TUL to γ. Recall that γ is the random independent probability that the vulnerability is discovered without any effort. As γ increases, we observe the losses to increase. This increase is explained as

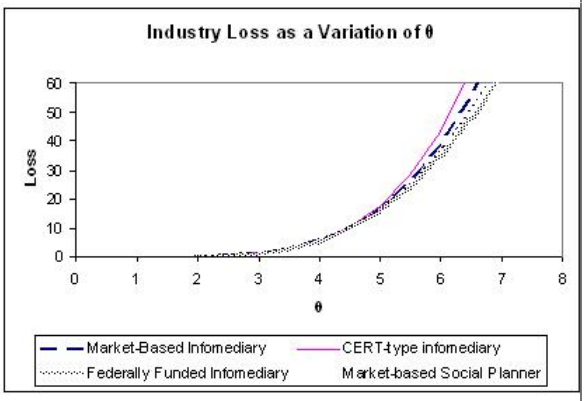

Figure 2: Total Industry Loss for different Mechanism Framework as $\overline{\theta}$ changes. $M = 1$ and $\gamma = 0$ for this analysis.

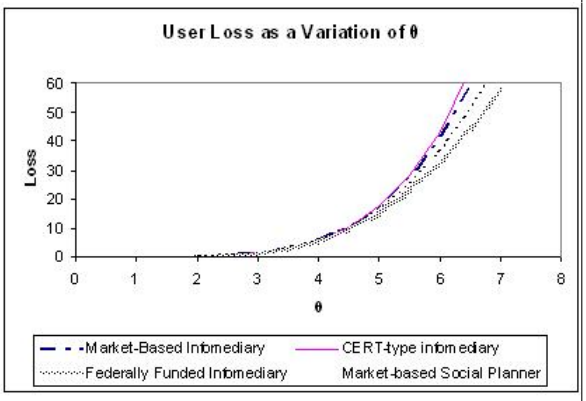

Figure 3: Total User Loss for different Mechanism Framework as $\overline{\theta}$ changes. $M = 1$ and $\gamma = 0$ for this analysis.

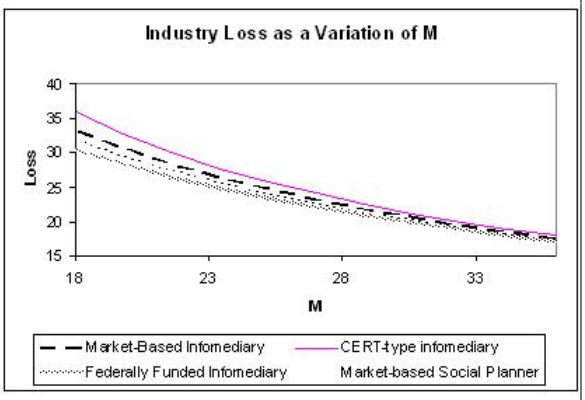

Figure 4: Total Industry Loss for different Mechanism Framework as M changes. $\overline{\theta} = 6$ and $\gamma = 0$ for this analysis.

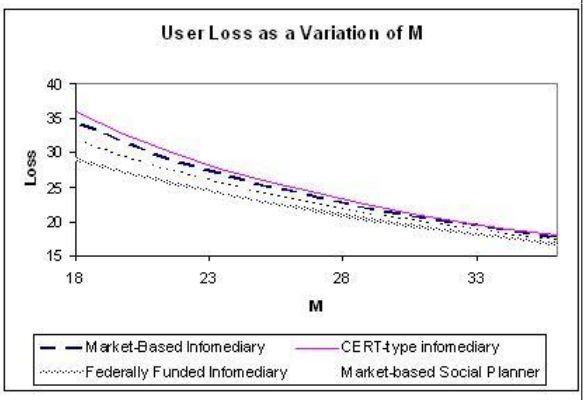

Figure 5: Total User Loss for different Mechanism Framework as M changes. $\overline{\theta} = 6$ and $\gamma = 0$ for this analysis.

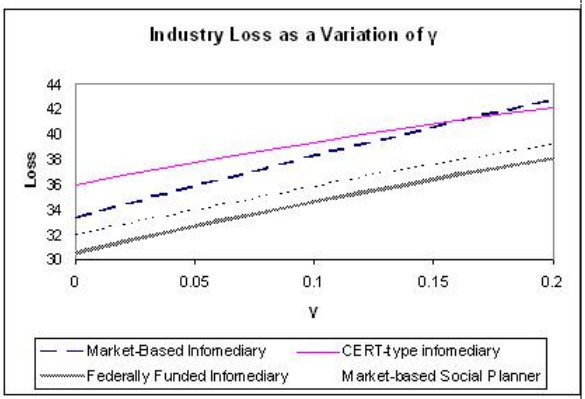

Figure 6: Total Industry Loss for different Mechanism Framework as γ changes. $\overline{\theta} = 6$ and $M = 18$ for this analysis.

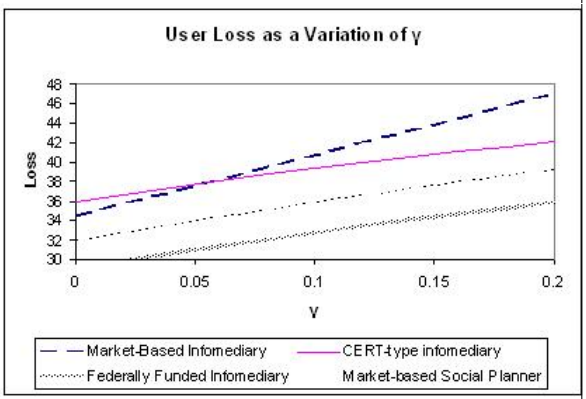

Figure 7: Total User Loss for different Mechanism Framework as γ changes. $\overline{\theta} = 6$ and $M = 18$ for this analysis.

the outcome of three factors, two of which aid the increase:

1. For a given $\overline{\theta}$, p_h, p_s and M, we observe that K_p, K_h and K_r increase as γ increases. Since K_p and K_h are the two main attributes that determine the loss, this contributes to an increase in the loss incurred.

2. We also know that as γ increases, p_h^* decreases. This is true for all other mechanisms except the CERT-type one. Intuitively, if the random probability that the vulnerability is identified increases, the incentive to fund vulnerability discovery decreases. Since p_h – the incentive to fund vulnerability discovery – decreases, the effort exerted by the benign identifier decreases and this results in higher losses for both the users and the industry.

3. This is the only factor that does not aid the increase in the loss and it is an indirect effect of point # 1. Notice that K_p increases as γ increases. But we know that when K_p increases, N_c – the number of subscribers to the infomediary's service – increases. This contributes to the decrease in user losses and hence, the industry losses also.

The effect of the first two factors outweigh the effect of the last factor and therefore we see an increase TUL and TIL as γ increases.

Other observations can also be explained using the same three factors. Point # 2 states that as γ increases, the optimal reward price set, p_h^*, by all other mechanisms except the CERT-type one decreases. This implies that there is a limit for γ in each of those mechanisms beyond which, the mechanism itself fails. This limit corresponds to the condition $p_h^* < 0$. Within its valid range of γ, the Federally-Funded mechanism performs better than all other mechanisms on both the welfare-parameters. This is not surprising, considering the objective function optimized by that infomediary. Beyond the valid range of γ, p_h^* for the Federally-Funded infomediary is equal to 0

i.e., it functions similar to the CERT-type mechanism.

Also, we observe that the market-based mechanism and the consortium mechanism perform better than their CERT-type counterpart for lower values of γ. It is the other way for higher values of γ. To provide intuitions for this result, we first explain why the slopes in the figure – $\frac{\partial TIL}{\partial \gamma}$ or $\frac{\partial TUL}{\partial \gamma}$ – are different for these mechanisms. After that, we explain why the CERT-type mechanism performs worse than the other types for lower values of γ. Let us consider the three points mentioned earlier. Under the CERT-type mechanism, point # 2 does not exist at all since p_h is always set to 0. Since one of the factors that serves to increase the loss incurred does not exist, the slope of CERT-Type mechanism's TUL and TIL curves are lesser than those of the other two mechanisms.

Since the CERT-type mechanism does not fund vulnerability disclosure, no vulnerability is identified when $\gamma = 0$. In contrast, other mechanisms fund vulnerability disclosure which serves to improve TUL and TIL. The losses prevented because of funding vulnerability discovery are found to increase with $\overline{\theta}$ and decrease in M (see Figure 4, Figure 2). This explains why at lower values of γ, the market-based mechanism and consortium mechanism perform better than a CERT-type one. When this is combined with the reasoning behind why the loss rate with respect to γ is different, we are able to explain the other observation i.e., why the performance of the CERT-type mechanism improves relatively as γ increases.

4 Conclusion

In conclusion, we have studied an important real-world problem related to software vulnerability disclosure. Specifically, we use game-theory to compare the welfare-effects of different mechanisms. Based on our analysis, we observe that a Federally-Funded Social Planner performs better

than all other mechanisms. We also find that under certain conditions, the performance of market-based mechanism is better than the CERT-type one and vice-versa. Intuitions provided for these results can aid a policy-maker in understanding the implications of existing disclosure mechanisms. This understanding is critical especially given the current scenario where practitioners in different capacities are proposing new frameworks. We intend to extend this work by comparing other disclosure frameworks proposed by different practitioners.

References

Arora, A., Caulkins, J.P. & Telang, R. (2003). Provision of Software Quality in the Presence of Patching Technology, Carnegie Mellon University, working paper.

Banker, R., Davis, G. & Slaughter, S. (1998). Software Development Practices, Software Engineering Complexities, and Software Maintenance. *Management Science*, **44**, 433–450.

Dasgupta, P. & Stiglitz, J. (1980). Uncertainty, Industrial Structure, and the Speed of R&D. *Bell Journal of Economics*, **11**, 1–8.

Du, W. & Mathur, A. (1998a). Categorization of Software Errors that led to Security Breaches. In *21st National Information Systems Security Conference, Crystal City, VA*, 392–407.

Du, W. & Mathur, A. (1998b). Vulnerability Testing of Software System Using Fault Injection. Tech. rep., Department of Computer Science, Purdue University, Reference: Coast TR 98-02.

e Week (2003). CERT, Feds Consider New Reporting Process, `http://www.eweek.com/article2/0,3959,970574,00.asp`.

Gordon, L.A. & Loeb, M.P. (2002). The Economics of Information Security Investment. *ACM Transactions on Information and System Security*, **5**.

Krishnan, M.S., Kriebel, C.H., Kekre, S. & Mukhopadhyay, T. (2000). An Empirical Analysis of Productivity and Quality in Software Products. *Management Science*, **46**, 745–59.

Krsul, I., Spafford, E. & Tripunitara, M. (1998). Computer Vulnerability Analysis. Tech. rep., Department of Computer Science, Purdue University, `citeseer.nj.nec.com/krsul98computer.html`.

Reinganum, J. (1982). A Dynamic Game of R&D: Patent Protection and Competitive Behavior. *Econometrica*, **48**, 671–688.

Security-Focus (2003). Security Research Exemption to DMCA Considered, `http://www.securityfocus.com/news/4729`.

Varian, H.R. (2000). Managing Online Security Risks. *The New York Times*, `http://www.nytimes.com/library/financial/columns/060100econ-scene.html`.

ZD-Net (2003). Trusted Computing Comes with a Warning, `http://www.zdnet.com.au/newstech/security/story/0,2000048600,20273805,00.htm`.