Incentive Dynamics in Interdependent Network Security
Or: Buying a Raft and Out-Running a Bear

John Chuang
UC Berkeley

Plenary Lecture at GameNets 2011, Shanghai China, April 18 2011.

1. Rational Decision-Making in Information Security

Are we investing too little in security? Are we investing too much? This pair of questions, first posed by Ross Anderson and Hal Varian in 2002 [1], may appear deceptively straightforward but lie at the heart of rational decision-making in information security.

In the wake of the recent earthquake-tsunami-nuclear tragedies in Japan, we hear the following questions being asked by citizens and experts alike: How secure are our nuclear power plants? Should the sea walls be built even higher? What is an acceptable level of risk, and how much are we willing to pay for it? Just how secure is secure enough?

German sociologist Ulrich Beck declared two decades ago, in 1992, that we are living in a "risk society" [2]. Not only do we face risks from natural hazards that are beyond human control, such as earthquakes and tsunamis, we also make conscious decisions to create, negotiate, and accept risks from man-made hazards, including radiation from nuclear power plants, chest x-rays, flying in an airplane controlled by computers, driving a car, and even shopping online. We even have a quantification of risk as expected loss, $risk = p \bullet L$, i.e., the product of the probability of event and the magnitude of loss per event. For many public safety and occupational safety scenarios, the losses are often measured in lives lost. Yet, the risk mitigation alternatives incur costs measured in dollars. Therefore, our governments compute, in real dollar figures, the *Value of Statistical Life* (VSL) [3], and use it to regulate acceptable levels of radiation exposure, to mandate installation of front passenger airbags, etc.

As societies and as individuals, we routinely perform cost-benefit analysis when it comes to security decision-making. When we are prompted to type in our credit card number online, we weigh the benefits of buying that pair of jeans against the risk of identity theft. As Chief Security Officers (CSO) of a corporation, we weigh the cost of implementing a new security mechanism against the possible mitigation of a security breach. As a nation, we debate the appropriate level of security expenditure for defending the cyber and physical infrastructures against various threats [4].

As risk managers, we face a myriad of challenges. First, we often lack the data to properly assess risks, especially for events with very low probability. This is exacerbated by the fact that computer security is a moving target – new systems introduce new vulnerabilities and new threats that can quickly render previous assessments obsolete. How do we account for the "black swans" [5] or "unknown unknowns" [6] of computer security? Second, we are subject to numerous behavioral biases that keep us from investing in the security of our systems, including our tendency towards risk-seeking behavior when faced with prospects of loss [7], and the effect of hyperbolic discounting that makes us shun short-term investments that have long-term payoffs. On top of all these, we have a hard time internalizing the fact that our action affects the security outcomes of others, and the actions of others can affect our security outcome. As a result, we are often unable to coordinate our actions in situations of *interdependent security* (IDS).

2. Interdependent Security

There is a common adage in system security that "a system is only as secure as its weakest link."

From the attacker's perspective, it is obvious that he/she should find and exploit the weakest link in a system. From the defender's perspective, therefore, it is imperative that they identifies and strengthens the weakest link, followed by the next weakest link, and so on.

However, what happens if the system is a large-scale federated system where each component, or each link, is owned and operated by a different defender? How would, and how should, the defenders coordinate their response given their shared fate?

In the canonical example of interdependent security (IDS), we can visualize a game where defenders are each responsible for erecting a section of a defensive wall. It takes effort for each defender to build up its section of the wall, but the effort pays off for all the defenders.

Depending on the nature of the attack, the effectiveness of the wall may be determined by its minimum height, average height, or maximum height. In the Economics literature on the private provisioning of public goods, these correspond to the *weakest link*, *best shot*, and *total effort* games. In network security, perimeter defense is a good example of a weakest link game. Censorship resistance a la WikiLeaks is a good example of a best shot game. Supporting anonymous communications via a peer-to-peer style mix-net like the Tor network is a good example of a total effort game.

We can characterize the utility function for player *i* as:

$$U_i = M_i - p*L*(1-H(e_i,e_{-i})) - b*e_i \qquad (1)$$

where M_i is the player's initial endowment, $p*L$ is the risk, e_i is the protection level by the player, e_{-i} is the protection levels by the other players, and b is the per unit cost of protection. Most importantly, the amount of risk mitigation offered by the protection function $H(e_i,e_{-i})$ is dependent on not just the protection level chosen by player i, e_i, but also the protection levels chosen by the other players, e_{-i}.

For the weakest link game, H is the minimum across all protection levels. For the best shot game, it is the maximum. For the total effort game, it is the sum or the average of all protection levels. (In the absence of any interdependency, H is simply e_i, the protection level chosen by the player.)

No interdependency: $H(e_i,e_{-i}) = e_i$
Weakest Link: $H(e_i,e_{-i}) = \min(e_i,e_{-i})$
Best Shot: $H(e_i,e_{-i}) = \max(e_i,e_{-i})$
Total Effort: $H(e_i,e_{-i}) = \text{sum}(e_i,e_{-i})$

For all three cases, and other variants such as weaker link or better shot, the bottom line remains unchanged -- free-riding is a rational strategy, and it leads to a suboptimal under-provisioning of the public good. By observing the public goods characteristics of interdependent security, Hal Varian correctly predicts that individual defenders will under-invest in security [8].

3. Protection vs. Insurance

We can make an important insight by distinguishing between different types of security investments. Kunrether and Heal made the distinction between protection and insurance [9], which we apply to our context of information security [10].

By protection, we are referring to security mechanisms such as firewalls, anti-virus software, regular system patching practices, that reduce the probability of loss, p.

By insurance, we refer to security investments that allow the defender to reduce the magnitude of loss, L, in the event that a loss occurs. Regular backup of data is an example of insurance. Purchasing a cyber-insurance policy is another example.

In the presence of both protection and insurance options, the player's utility function becomes:

$$U_i = M_i - p*L*(1-H(e_i,e_{-i}))*(1-s_i) - b*e_i - c*s_i \qquad (2)$$

where s_i is the insurance level chosen by the player, and *c* is the per unit cost of insurance.

Now, the defender is faced with a budget allocation problem. How should the defender allocate its budget between protection (e_i) and insurance (s_i)? Both options offer to mitigate the risk, but there is an important distinction between them.

For protection, its effectiveness is dependent not just on the investment by player *i*, but on the investments by the other players as well, and by the interdependency characteristic of the attack, which may or may not be known to the defenders beforehand. For insurance, on the other hand, its effectiveness is dependent only on the investment by the player *i*, and unaffected by the investments by the other players.

If investing in protection is like building a high defensive wall against incoming water, then investing in insurance is like buying a personal life raft. Now, each defender is faced with the problem of allocating resources between a public good (protection) and a private good (insurance).

Applying game theoretic techniques to the problem, we see a nuanced picture for each of the canonical games, as compared to the protection-only scenario.

For all three games, there exist protection equilibria, insurance equilibria, as well as passivity equilibria (where defenders invests in neither protection nor insurance), depending on the cost of protection *b*, cost of insurance *c*, and level of risk $p*L$.

Focusing on the protection equilibria, we find that it is possible to achieve full protection equilibrium for the total effort game. The protection equilibrium is more stable for heterogeneous populations, but is harder to achieve as the number of defenders, *N*, increases.

For the best shot game, pure symmetric equilibria do not exist. However, asymmetric protection equilibria (where one defender invests in full protection and all others free-ride) or mixed protection equilibria can be achieved. Of course, coordination among the defenders becomes more difficult with increasing *N*.

For the weakest link game, there exist multiple unstable protection equilibria. As the number of defenders increases, the protection equilibria collapse to either full insurance or passivity. In the weakest link game, all it takes is one defender to defect, switching from protection to insurance, to trigger defection by all other players. Protection equilibria are also harder to achieve with heterogeneous populations [11].

Overall, the news is not good. While insurance may mitigate the loss of an individual, it does nothing to improve the security of the system. The availability of cheap insurance options may lead to poorer security outcomes overall.

In the laboratory setting, we observe that even under a limited information environment, it is possible for players to learn the interdependent game structure and respond accordingly [12].

Next, we consider the possibility of having a mixture of expert and naive players, where expert players understand the interdependent nature of the game, but the naive players are unaware of any of the weakest link, best shot, or total effort dynamics, and makes decisions in a myopic manner [13]. Do the presence of experts improve the security outcome? We find that if the expert players behave in an individually rational way, they will lead to poorer security outcomes than if all players were naive. On the other hand, if these same experts collaborate with each other (via side payments), they can yield superior security outcomes in almost all cases we considered. (An exception is in the case of the total effort game with high protection costs, where a population consisting entirely of naive players outperforms a group of experts, even if they are collaborative.)

Finally, we consider a coordinating role that might be played by intermediaries such as Internet Service Providers (ISPs) and security technology vendors, and social planners such as governmental cyber-security policy units. In particular, we focus on the use of financial incentives, in the form of either pay-for-outcome or pay-for-effort, to nudge end users towards more secure outcomes [14]. For example, under a pay-for-outcome scheme, an ISP can offer its subscribers a rebate at the end of each month if the subscriber's machine was scanned and found to be free of any malware. For an example of a pay-for-effort scheme, the government might offer to subsidize the cost for consumers to purchase anti-virus software. We find that both of these approaches can be effective, and their relative cost-effectiveness depends on the same set of game parameters we have discussed, such as the loss probability p, loss magnitude L, protection cost b, and insurance cost c.

4. From Weakest Link to Weakest Target

We have seen that in interdependent security (IDS) games, coordination becomes more and more difficult with increasing number of defenders. As a result, the larger the network, the less likely the defenders will invest in protection as a public good.

If we scale the weakest link game to the scale of the Internet, where we have hundreds of millions of users and billions of devices, we might conclude that no user will bother with protecting their systems.

Of course, this is not what we observe in reality. While games like weakest link may adequately describe perimeter defense involving a small number of defenders, it is hard to imagine millions of defenders building a single perimeter defense. Therefore, we need to go back to the drawing board and consider a different class of IDS games that are appropriate for Internet-scale security scenarios.

The place to start is to consider the types of Internet-scale attacks that are prevalent today, e.g., spam, phishing, and botnets. In each of these attacks, the attacker is not looking to compromise all defenders at the same time, but instead is only looking to a small vulnerable subset of the population. Instead of looking for the weakest link into a single system, the attacker is looking for the *weakest targets* out of the entire population.

In the case of email spam, the spammers send out tens of billions of messages every day, with the hope that a small fraction of the messages make it into email inboxes that have porous or inactivated spam filters. In the case of phishing, the attackers send out billions of messages, with the hope that a small fraction of individuals click through to the phishing sites and enter their personal or confidential information. In the case of botnets, the attackers recruit machines as bots by spreading malware that exploits known vulnerabilities, and infects those machines are do not have the most up-to-date system patches.

In each of these cases, the defenders that are the weakest targets, e.g., email filters that are not up-to-date, careless users, machines with unpatched vulnerabilities, are vulnerable to compromise and loss, whereas other defenders are left unscathed. This is an important departure from the canonical weakest link, best shot, and total effort games, where either all defenders are compromised, or none of the defenders are.

Mathematically, we can model this weakest target game with the same utility functions as (1) or (2), together with the following protection function:

$$H(e_i,e_{-i}) = 0 \text{ if } e_i = \min(e_i,e_{-i})\text{; 1 otherwise}$$

For each of the players with a protection level, e_i, that is lowest amongst all players, it faces the unmitigated risk of $p*L$. All of the remaining players will suffer no loss.

What is the dynamics of this game? The weakest target game is analogous to the story of the hikers and the bear. The story goes, as follows:

> Two hikers come upon an angry bear. The first says, "I'm glad I wore my running shoes." The second says, "You can't outrun the bear." The first says, "I don't have to outrun the bear, I just have to outrun you."

Essentially, the players are investing in protection with the hope of coming in slightly ahead of the weakest targets, so as to avoid the loss. The larger the number of players, the more willing are the players to take the calculated gamble, instead of giving up on protection and opting for insurance. In fact, due to the strategic uncertainty of protection investments, players may even protect at a level that is higher than the socially optimal level [10].

As we see, there is an importance difference between the weakest target game and the weakest link game, as they scale up in N, the number of players. In the weakest link game, we saw that a larger population size leads to a shift from protection to insurance.

Now, in the weakest target game, we see the opposite occurring. As the population size increases, players shift away from insurance towards protection.

5. How Secure is Secure?

We close by returning to the initial pair of questions: "Are we investing too little in security? Are we investing too much?" To answer these questions, we started with a risk management approach, quantifying the security risks and analyzing defenses as risk mitigation options, while acknowledging the numerous challenges of this approach. Scaling up to multiple defenders, we considered the inter-dependent nature of information security, which highlights the public goods characteristics of security investments. However, we recognized that security investments might be of different forms, e.g., protection versus insurance, and this creates new incentive dynamics involving tradeoffs between public goods and private goods. Finally, we saw that interdependency does not have to imply "fate-sharing". Interdependency does not mean all users will realize identical losses. Rather, different users may adopt different strategies to mitigate their risks -- some through protection, some through insurance -- and realize different outcomes.

The weakest target game is, I believe, just one of many new interdependent security games that we, as a community, will develop to model different security dynamics at the Internet scale. Other interdependent security games might consider other forms of externalities such as correlated failures or cascaded failures.

Other directions for research might include the study of parallel or disjoint games, hybrid games, or even the reverse engineering of games. There are many calls for designing for resiliency. How is investing in resiliency different from investing in protection or insurance? In what ways is the communication of interdependent security risks similar to or different from traditional risk communication? Finally, a more nuanced understanding of interdependent security will allow us, as system designers or social planners, to be better informed about possible interventions, whether in the form of financial instruments, policy instruments, or design principles.

References

[1] First Workshop on Economics and Information Security, Berkeley CA, 2002.

[2] Ulrich Beck. *Risk Society: Towards a New Modernity*. London: Sage, 1992.

[3] Schelling, Thomas C. *Value of life.* The New Palgrave: A Dictionary of Economics. First Edition. Eds. John Eatwell, Murray Milgate and Peter Newman. Palgrave Macmillan, 1987.

[4] John Mueller and Mark Stewart. *Terror, Security and Money: Balancing the Risks, Benefits and Costs of Homeland Security*. Oxford University Press, September 2011.

[5] Nassim Taleb. *The Black Swan: The Impact of the Highly Improbable*. Random House, 2007.

[6] Donald Rumsfeld. U.S. Department of Defense briefing, February 12, 2002.

[7] Daniel Kahneman and Amos Tversky. Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, XLVII, 263-291, 1979.

[8] Hal Varian. System Reliability and Free Riding. In L. Camp and S. Lewis (ed.), Economics of Information Security (Advances in Information Security, Volume 12), pages 1–15. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2004.

[9] H. Kunreuther and G. Heal. Interdependent security. *Journal of Risk and Uncertainty*, 26(2–3):231–249, Mar. 2003.

[10] J. Grossklags, N. Christin, J. Chuang. Secure or Insure? A Game-Theoretic Analysis of Information Security Games. Proceedings of the 17th International World Wide Web Conference (WWW'08), April 2008.

[11] Grossklags, Christin, Chuang , 2009. J. Grossklags, N. Christin, J. Chuang. Security and Insurance Management in Networks with Heterogeneous Agents. Proceedings of ACM E-Commerce Conference (EC'08), July 2008.

[12] J. Grossklags, N. Christin, J. Chuang. Predicted and Observed User Behavior in the Weakest-Link Security Game. Proceedings of the 2008 USENIX Workshop on Usability, Psychology, and Security (UPSEC'08), April 2008.

[13] B. Johnson, J. Grossklags, N. Christin, J. Chuang. Are Security Experts Useful? Bayesian Nash Equilibria for Network Security Games with Limited Information. Proceedings of 15th European Symposium on Research in Computer Security (ESORICS'10), September 2010.

[14] J. Grossklags, S. Radosavac, A. Cardenas, J. Chuang. Nudge: Intermediaries' Role in Interdependent Network Security. Proceedings of 3rd International Conference on Trust and Trustworthy Computing (Trust'10), June 2010.