The Workshops of the Thirtieth AAAI Conference on Artificial Intelligence
Artificial Intelligence for Cyber Security: Technical Report WS-16-03

Inter-Temporal Incentives in Security Information Sharing Agreements — Position Paper

Parinaz Naghizadeh and Mingyan Liu
Department of EECS, University of Michigan, Ann Arbor

Abstract

In recent years, sharing of security information among organizations has been proposed as a method for improving the state of cybersecurity. However, despite its benefits, this disclosure entails additional costs for the reporting entity. In this paper, we take a game theoretic approach to understanding firms' incentives for participating in information sharing agreements given such costs. We present a repeated game formulation of security information sharing games. Our approach proposes the use of inter-temporal incentives (i.e., conditioning future cooperation on the history of past interactions) to support firms' cooperation on information sharing.

Background and Motivation

Improving the ability of analyzing cyber-incidents, and ensuring that the results are shared among organizations and authorities in a timely manner, has received increased attention in the recent years by governments and policy makers, as it can lead to a better protection of the national infrastructure against potential cyber-attacks, allow organizations to invest in the most effective preventive and protective measures, and protect consumer rights.

In the US, improving information sharing is listed as one of President Obama's administration's priorities on cybersecurity. Currently, most of the existing laws require organizations to only report to an authority, with a few other also mandating notification of the affected individuals (e.g., HIPAA); see (Laube and Böhme 2015) for a summary of prominent existing laws. However, most recently, President Obama signed Executive Order 13691 on "Promoting Private Sector Cybersecurity Information Sharing", encouraging companies to share cybersecurity information with one another, in addition to the federal government. Motivated by these initiatives, in this paper, we are interested in information sharing agreements *among firms*. Examples of existing information sharing organizations or initiatives of this type include: Information Sharing and Analysis Centers (ISACs), Information Sharing and Analysis Organizations (ISAOs), and the United States Computer Emergency Readiness Team (US-CERT). Currently, joining and reporting in all these information sharing organizations is voluntary.

Problem motivation Despite the introduction of information sharing laws and agreements, both anecdotal and empirical evidence point out that security breaches remain vastly under-reported. These observed disincentives can be primarily explained by analyzing the associated economic impacts. (Campbell et al. 2003; Cavusoglu, Mishra, and Raghunathan 2004) conduct event-study analyses of market reaction to breach disclosures, both demonstrating a drop in market values following the announcement of a breach. Additionally, breach disclosure may lead to a loss of customer confidence and competitive advantage, as well as bureaucratic burdens.

On the other hand, firms do benefit from having access to other firms' security information, as they can prevent similar attacks and invest in the best security measures by leveraging other firms' experience. As a result, an outcome in which firms fully disclose their security information is beneficial to all participants. Given these potential disclosure benefits and costs, and the evidence of under-reporting of security information, it is clear that we need a better understanding of firms' incentives for participating in information sharing organizations, as well as the economic incentives that could lead to voluntary cooperation by firms.

A game-theoretic approach We present a game-theoretic study of information sharing agreements among firms, in order to better understand firms' (dis)incentives for fully and honestly disclosing security breaches and existing flaws, given their potential disclosure benefits and costs. We first show that in a one stage information sharing game among rational firms, the disclosure cost acts as a deterrent, leading to a lack of shared information. Existing research has proposed audits and sanctions (e.g. by the government), or additional economic incentives (e.g. taxes and rewards) as remedies for encouraging disclosure, see e.g. (Gordon, Loeb, and Lucyshyn 2003; Laube and Böhme 2015).

In this work, we present a different approach for providing such incentives. We propose the study of a *repeated game* framework, therefore allowing for firms' future disclosure decisions to be dependent on the history of their interactions with other firms in the agreement. We first illustrate how cooperation can be incentivized with positive probability in a two-stage information sharing game. We then introduce infinitely repeated games, and briefly discuss potential approaches for sustaining cooperative equilibria.

	1	0
1	$G-L, G-L$	$-L, G$
0	$G, -L$	$0, 0$

Table 1: Info. sharing game

	G	B
G	h,h	0,0
B	0,0	ℓ,ℓ

Table 2: Partnerships

The information sharing game

Consider two firms who have agreed to share their security information through an information sharing agreement. Nevertheless, each firm has a choice as to whether (fully and honestly) disclose her information. We denote the decision of firm i by $r_i \in \{0, 1\}$, indicating (partially) concealing and (fully) disclosing, respectively. A choice of $r_i = 1$ results in *disclosure costs* $L > 0$ for firm i. We assume this choice benefits the other firm j by helping her improve her state of security, yielding an *information gain* $G > 0$ for firm j. Assume $G > L$. The payoff matrix of the *information sharing* game among the firms is given in Table 1. This game is therefore an instance of the *prisoner's dilemma*: the only Nash equilibrium of the one stage game is for neither firm to disclose her security information. However, a repeated game formulation can leverage firms' interest in maintaining a good reputation in the future to sustain cooperation among participants.

A two stage game First, consider a two-stage interaction among the firms, with first and second stage payoffs given by Tables 1 and 2, respectively. The second stage game captures decisions on a subsequent business partnership, with G (B) denoting a high (low) profit partnership, where $h > \ell > 0$.

To condition future behavior on past actions, we assume each firm can only imperfectly assess the honesty and comprehensiveness of the other's report. In particular, following the first stage, firm i forms a *belief* b_i about firm j's report, by monitoring firm j's externally observed security posture. We let $b_i = 1$ indicate a belief of full disclosure, and $b_i = 0$ otherwise. We assume the belief b_i of firm i is imperfect, private, and independent of firm j's belief b_j about firm i. Formally, we assume the distribution:

$$\pi_i(b_i|r_j) = \begin{cases} \epsilon, & \text{for } b_i = 0, r_j = 1 \\ 1-\epsilon, & \text{for } b_i = 1, r_j = 1 \\ \alpha, & \text{for } b_i = 0, r_j = 0 \\ 1-\alpha, & \text{for } b_i = 1, r_j = 0 \end{cases}$$

with $\epsilon \in (0, 1/2)$ modeling missed detection by firm j, and $\alpha \in (0, 1)$ as the accuracy of firm i's monitoring technology.

Pure strategy and mixed equilibria Ideally, we would like to identify a pure strategy equilibrium that supports $(r_i, r_j) = (1, 1)$ in the first period, by conditioning the second stage partnership on the first stage decisions. Nevertheless, it can be shown that (Mailath and Samuelson 2006), as firm i's belief about firm j's action in the second period is independent of i's observed signal, it is not sequentially rational for firm i to consider her signal in the second period. Therefore, with pure strategies, inter-temporal incentives can not be used to coordinate on $(r_i, r_j) = (1, 1)$.

We next consider an alternative strategy profile in which firms randomize their actions in the first period. Formally, suppose in the first period, firm i plays $r_i = 1$ with probability β, and $r_i = 0$ otherwise. In the second period, this firm will play H if and only if she has played $r_i = 1$ in the first period, and she has a belief $b_i = 1$ about firm j. It is possible to solve the equilibrium conditions to find a $\beta > 0$ as a function of the monitoring parameters α and ϵ. Therefore, inter-temporal incentives lead to full disclosure $(r_i = 1, r_j = 1)$ emerging with positive probability.

The infinitely repeated game Next, consider the stage game of Table 1 repeated infinitely. A longer history of play can allow for more elaborate strategies; e.g., non-disclosure periods that start after a certain number of suspected deviations, or that last only for a certain number of rounds. Therefore, one may expect the possibility of supporting cooperation with similar (or better) results, compared to the two-stage game, by considering longer lasting interactions.

(Compte 2002) shows a negative result in this game when trigger strategies are used: even if firms' signals about others' actions are highly informative, full cooperation on information disclosure can not be supported. There exist, however, alternative approaches that may help support (some) cooperation, including: allowing firms to communicate (cheap talk) (Compte 1998), availability of public actions (e.g., announcing sanctions) in addition to (private) disclosure decisions (Park 2011), or almost public monitoring, i.e., independent private monitoring with signals that are sufficiently correlated (Mailath and Samuelson 2006). A detailed analysis of these approaches, and their implications on the role of authorities as facilitators of public monitoring or communication, is the main direction of future work.

Acknowledgments

This work is supported by the Department of Homeland Security (DHS) via contract number HSHQDC-13-C-B0015.

References

Campbell, K.; Gordon, L. A.; Loeb, M. P.; and Zhou, L. 2003. The economic cost of publicly announced information security breaches: empirical evidence from the stock market. *Journal of Computer Security*.

Cavusoglu, H.; Mishra, B.; and Raghunathan, S. 2004. The effect of internet security breach announcements on market value: Capital market reactions for breached firms and internet security developers. *Int. Journal of Elec. Commerce*.

Compte, O. 1998. Communication in repeated games with imperfect private monitoring. *Econometrica* 597–626.

Compte, O. 2002. On failing to cooperate when monitoring is private. *Journal of Economic Theory* 102(1):151–188.

Gordon, L. A.; Loeb, M. P.; and Lucyshyn, W. 2003. Sharing information on computer systems security: An economic analysis. *Journal of Accounting and Public Policy*.

Laube, S., and Böhme, R. 2015. The economics of mandatory security breach reporting to authorities. In *Workshop on the economics of information security (WEIS)*.

Mailath, G. J., and Samuelson, L. 2006. *Repeated games and reputations*, volume 2. Oxford university press Oxford.

Park, J. 2011. Enforcing international trade agreements with imperfect private monitoring. *The Rev. of Econ. Studies*.