Sharing private information or data between individuals or organizations

Daniele Mellino Emanuele Mincato
{daniele.mellino, emanuele.mincato}@studenti.unipd.it

Abstract

In the modern society where digitization accelerates, firms have to deal with a huge amount of cyber-threats. One way to increase their defence is by enrolling in ISACs. Such centers increase information sharing and help firms to decrease cyber security risks, prevent attacks, and increase their overall resilience. However privacy risk and information disclosure concerns are still major challenges. That is why several firms still does not exploit this opportunity.

To prevent cyber-threats is key to understand how an attacker would play. Therefore in this paper we construct a bayesian game in which an attacker choose between three choices(no attack, small attack, big attack), not knowing if the target is in ISAC or not. Analyzing the cases in which a firm is or is not probable to be in ISAC, we show that the number of small attack decrease in the first scenario, assessing even more the needs of a firm to enroll in such organization.

1. Introduction

The frequency and complexity of cyber-attacks have increased with the significant growth of our daily life dependency to the cyberspace. To get ahead of the security threats, it is crucial to have a proactive security approach to prevent any dangers before they occur. Cybersecurity information sharing is a key factor in proactively defending against sophisticated cyber-attacks [1]. In addition it decrease the time and enhance the detection of malicious behaviour in the system.

In the last years governments and authorities around the world started a series of actions to encourage firms to share cybersecurity information such as vulnerabilities, security breaches, penetrations, etc. (for instance CISA in USA [6] or European legislations like the NIS Directive and the Cybersecurity Act [5]). Those actions promote the creation of Information Sharing and Analysis Centers (ISACs) which are non-profit organizations that provide a central resource for gathering information on cyber threats.

Using the shared information, ISACs can quickly alert the member firms about the physical and cyber-threats to help them protect their information systems as soon as possible. In addition, experts analyze the threats in order to recognize solutions that are shared among firms, thus they can get the latest techniques and the proven best practices to protect themselves against emerging and known security threats.

ISACs indeed provide value-added services for updating members by fusing and merging information and performing extra analyses on the participant's data. This way, the fused information that is shared growths over time and becomes increasingly dynamic and complex.

Sharing data and sensible information comes with a price. There is insufficient trust to these centers as well as the highly negative impact of privacy and information disclosure risks. In fact attackers might utilize the shared information for reconnaissance, the competitive organization might take advantage of the shared information which indirectly affects the organization's utility, and sensitive private information might leak out. That's why ISACs member can decide to apply the organization without sharing information (free riding).

To avoid this process and to finance the research and improvement inside the organization, ISACs usually require a periodic fee. This is not the only entry of the organization, in fact they usually receive governmental funds. To avoid the consequences of the untrusty feelings regarding those associations, membership is a private information. In this paper, we aim to asses that the presence of an ISACs centre reduces the number of small attacks to the network from an attacker point of view. In our work we consider a bayesian game in which an attacker does not know if his target is or is not in a ISAC.

In Section 2. we introduce some inpirational paper we read before the formulation of the game. In Section 3 we introduce the payoffs function and the proposed game model after discussing about the taken assumptions. In Section 4. we shows and debate the results of our work.

2. Related work

Cybersecurity information sharing have been studied extensively in the last years. The approach used in those study vary from the more economical one to more based game-theoretic approach. Here we resume the main paper from which we take inspiration.

Cavusoglu et al. [2] in their work build a sequential game in which consider a firm and an attacker. They showed how that the sequential game results in the maximum payoff to the firm, but requires that the firm move first before the hacker.

In the formulation of the mathematical model is important to consider the risk that a firm could have due to information leakage and privacy breech. This could incentivate firms to enroll in ISAC and not share data. The participation fee ought to prevent organizations to take advantage of the shared information without participating in sharing. Furthermore, the fee should be fair, such that the organizations' payment should be proportional to their benefits from the information. In [7] the authors build a coalition games and show a fee mechanism that avoid free riding aiming to increase the organization value.

ISACs should not limit their work to just share issue and signal possible threats. As shown from M.Ezhei and B.T. Ladani [4] these associations should work as enhancer, increasing the information value of the firms. In their work they build a differential game between n firms and n attackers in which they consider not only the privacy cost but also the impact of ISAC in changing the volume of information, which they consider linear.

None of the above works consider the process from an attacker points of view, neither consider the scenario of a bayesian game of incomplete information.

3. The proposed game model

3.1. Bayesian game theory

Bayesian game theory assumes that both the attacker and the defender do not know all the details about their opponent. Both players only disclose certain information about their assets, goals and working methods. Mostly this is done unwillingly but necessarily to progress the game [3]. There are some assumptions that have to be made when you want a mathematically correct way of defining a Bayesian game system. It needs to be assumed that the number of agents is known and that this is a fixed number. In our case there are just two player, the Firm and the Attacker. Also, the assumption that an agent's belief is posterior, i.e., a common prior can be determined based on individual private information. We discuss the value of *p* in the next section.

3.2. The Game

We consider a two-player Bayesian game represented by $\mathcal{G} =< (\mathcal{N}), (\mathcal{A}), (\Theta), p, (\mathcal{U}) >$. Where $\mathcal{N}$ = {F, A} represents the two players: Player F is the firm that has to protect his information, while player A is the attacker who is trying to harm them. Θ is the type space, in this game only Player F has his own type, so the type space corresponds to θ_1 = (θ_{11},θ_{12}) . In this formula, θ_{11} is the case when the firm joined ISACs while θ_{12} represent the case in which the firm did not join ISACs.
$\mathcal{A}$ is the set of sets of action (read as $\mathcal{A}$=$\{\mathcal{A}_1$,$\mathcal{A}_2\}$, where A_i is the set of actions available for player i), and their utilities are given by $\mathcal{U}$. In this case we have $\mathcal{A}_1 \in$ {ld,hd} and $\mathcal{A}_2 \in$ {na, la, ha}. Finally *p* is a common prior over all agents, usually expressing itself as a probability function over all type spaces, to indicate which set or action will be executed (according to Nature).
In our game *p* is the probability drawn by Nature, that the firm is in ISACs or not. His value is closely related to the reputation of ISACs. This mean that the value of *p* will increase as more firms decide to join the organisation and share the information with it. In this game the set of strategies for Player F are type-dependent, this means that his strategy set is equal to $\mathcal{S}_f$ = $\{xy|xy \in \{ld, hd\}\}$, where *x* is the action the player will take if he joined ISACs while *y* is the action he will take if he did not join ISACs. Instead the strategies of Player A are action-dependent, so his strategy set is equal to $\mathcal{S}_a$ = $\{ij|ij \in \{na, la, ha\}\}$, where *i* is the action he will take if player F chooses action *ld* while *j* is the actions he will take if Player F chooses action *hd*.
The possible actions of Player F are *ld* (low-defense) and *hd* (high-defense). In this game the 'level' of defense is intended as the amount of money the firm invest to protect their information, so invest more money will provide a stronger defense. Instead for Player A the possible actions are *na* (no-attack) ,*la* (low-attack) and *ha* (high-attack), so, similar to the firm, the attacker can choose the amount of money to invest to attack the firm, to try a weaker or a stronger attack, but he can also decide to not attack.

Since is a Bayesian game, the game consists of two 'stages'. In the first stage, the "Nature" takes action, and it determines the probability vector *p*(θ_{11},θ_{12}) . The second stage is the actual game. The two Players know the type that "Nature" had chosen for themselves (in this particular game only player F has types), but they did not know the type that "Nature" had chosen for the other player. However, the type space of each person and its probability distribution are common knowledge. So, according to the "Nature" action, this static game of incomplete information is converted into a complete but imperfect information game.

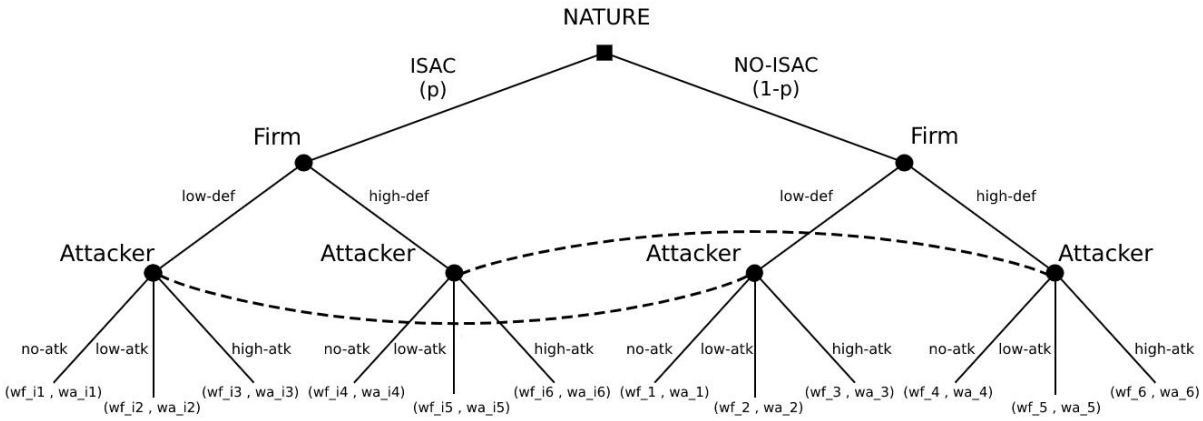

Figure 1. Extensive form of the game with imperfect information.

3.3. Model parameters

To model this game, and to be able to decide the utility of each strategy, we need to define different parameters, which are able to efficiently describe the interactions between the two Players. The most important variables are $v_d[i]$ and $v_a[i]$ ($i \in \{0, 1\}$), which respectively define the value of the information that the firm must protect and the gain that the attacker obtains if he obtains such information, both in the case the firm has a low defense (*i*=0) and when it has a high defense (i=1). The next thing we have defined is the cost to protect the information, $c_f[i]$ ($i \in \{0, 1\}$), and the cost of the attacker to attack it, $c_a[j]$ ($j \in \{0, 1, 2\}$). In this case $c_a[0]$ is equal to 0, since represent the case when the attacker decides not to attack. To define correctly the values of these variables we consider two assumptions:

1. The cost of the defense/attack is smaller than the value/gain of the information.

2. The cost of the defense is greater than the cost of the attack.

The first assumption is quite trivial since it makes no sense to invest more money for the defense/attack of the information than the actual value of them. The second assumption considers the fact that the investment for the defense of information is greater than the one of the attacker and that the economic availability of a company is usually greater than that of the attacker.

We then define $z_d[i]$ ($i \in \{0, 1\}$) and $z_a[j]$ ($j \in \{0, 1, 2\}$), which are respectively the efficiency of defense and attack, for each of their levels. Finally the last parameter we include to model this game is the *t* variable. This parameter appears only inside the Probability of break the firm defense and takes into account the increase in defense provided by ISACs. We define it as: $t = k \cdot p$. Where *k* is a constant value (to scale the variable) and *p* is the prior probability of the Nature. *p* is closely related to the reliability of ISACs, therefore as it becomes more established (*p* increases) it provides a greater increase in the defense of the company.

3.3.1 Probability of break the firm defense

To define the probability, of the success of the attack, we decide to modify the following security breach probability function [4]:

$$P = v_i \frac{(1 + \gamma c_i^a(t))^\phi}{(1 + \alpha z_i^a(t))^\beta}$$

The above function was introduced by Cavusoglu et al. (2008) based on the Gordon and Loeb model (2004). We have modified the formula so that it works with our parameters but at the same time maintains the same properties.

Thus in the end we got the following function and we use a normalized version of it for the probability:

$$Pb = \frac{(1 + z_a[i])^\phi}{(1 + z_f[j] + t)^\beta}$$

Assuming that $\beta > 1, 0 < \phi < 1$ ensures that higher attack (defense) enhances (reduces) the security breach probability.

The above function is considering the attack efficiency in the numerator, $z_a[i]$ (respect to the level of the attack), while in the denominator is considering the defense efficiency, $z_d[j]$, plus the contribution of ISACs (*t*). In the case the firm is not part of ISACs the *t* variable will be equal to zero. Furthermore if we have $z_a[0]$ the Probability will be equal to 0, since the attackers decides not to attack.

3.3.2 Privacy cost

Firms do not only suffer from security breaches' direct cost, but also respect privacy cost associated with information sharing [4]. This problem can arise if the organization, in this case ISACs, to which they share the data is not so reliable, for example it may itself have security breaches or voluntarily decide to share information with other firms. Furthermore the privacy cost directly depends on the value of information shared by the firm. We need to define a function that takes into account these aspect.

$$Pc = (Pb(0, j, z_m, z_d) - Pb(0, j, z_m, z_d, t)) \cdot v_d[j] \cdot (1.5 - p)$$

where $z_m = (z_a[1] + z_a[2])/2$

The difference in the probability of break the firm defense (without ISACs - with ISACs) is always positive, since ISACs always provides an increment to the defense of the firm. The value (1.5 - *p*) takes into account the reliable of ISACs. For example if ISACs is not a reliable organisation (like *p*=0.2) the value will be 1.3, while if ISACs is reliable (*p* = 0.8) it will be 0.7 . This means that as the value of *p* increases the organization becomes more trustworthy, and so the cost of the privacy for the firm will decrease.

3.4. Utilities

The last thing we need to model in our game are the utility function for each possible combination of action. We define the utilities for the firm (wf) and for the attacker (wa) in two different ways, based on the type of the firm. The utilities for the firm in the case is in ISACs are:

$$wfk_I = (1 - Pb(i, j, z_a, z_d, t)) \cdot v_d[j] - c_f[j] - Pc(j)$$

$$wak_I = Pb(i, j, z_a, z_d, t) \cdot v_a[i] - c_a[i]$$

and for the case the firm is not in ISACs:

$$wfk = (1 - Pb(i, j, z_a, z_d, t = 0)) \cdot v_d[j] - c_f[j]$$

$$wak = Pb(i, j, z_a, z_d, t = 0) \cdot v_a[i] - c_a[i]$$

where $k \in \{0 : 5\}; i \in \{0, 1, 2\}; j \in \{0, 1\}$

In these formula i represent the actions of the attacker (no-atk, low-atk, high-atk), j represent the actions of the firm (low-def, high-def), while k represent the possible combination of actions for the two Player in the case the firm is or is not in ISACs, since there are six possible combination for each branch it takes values from 0 to 5. We can notice that the main differences are that in the ISACs branch the *t* values inside the probability is different from 0 (is the increase of defence provides by ISACs) and that in the NO-ISACs branch there is not the Privacy cost for the firm. Moreover when we simulate the game we consider the following assumption on the utilities to select the right value for each parameter:

1. In the case of ISACs, if the attacker and the firm choose the same level of attack/defense (e.g. low attack, low def) the final utility strictly depends on the value of *p*. So if the *p* value is high enough the attack is unsuccessful, and vice versa if the *p* value is not high enough (attack successfully).

2. The attacker utility will be always zero if he decides to not attack, while the firm has to pay the cost of the defense and the privacy cost even if the attacker doesn't attack.

3.4.1 Expected Utility

Since we are dealing with a Bayesian game the above utilities are not the final one for the game. It is necessary to calculate the expected utility, which is nothing more than a weighted average of the previous ones, in the event that the firm has joined ISACs and have not done so.

$$\mathbf{E}(p) = p \cdot (w_{Ii}) + (1 - p) \cdot (w_i)$$

where w_{Ii} is the case when the firm is in ISACs and w_i is the other case. Since Player F has four possible strategies (two type and two actions) and player A has nine possible strategies (three action but are action dependent) there are 36 possible combination of strategies between the two players, so this means there are 36 different expected utility. For graphical reason we do not report here the table with all the possible combination of the expected utilities, but are pretty straightforward to calculate since we only have to substitute in the above formula the right combination of strategies in the case the firm is or is not in ISACs.

4. Results

In our work, we simulate the game through Python language, to compute the Nash Equilibria we use the tools in the library "Nashpy".

For reproducible purposes, in Table [1] we report the set of parameters with their relative value. In our simulation we consider different level of *p*. We found a threshold value $p_{thr} = 0.59$ such that if $p < p_{thr}$ the game have only one Bayesian Nash Equilibrium that correspond to (hh, HH). It is the scenario in which these kind of organization are not widespread, so in this case the cost of Privacy is greater than the increase in defense provided by ISACs. Here the firm will have to plan an high investment to have a chance of survive to the attacker.

If $p >= p_{thr}$ the game have four Bayesian Nash Equilibrium: $(nn, LL), (nl, LL), (nh, LL)$*and*$(hh, HH)$. This is the case in which is very common that a firm has enrolled in ISACs. Here we notice the addition of three more Nash equilibrium, respect to the previous one (hh, HH), which all have in common that the attacker prefers not to make any

Symbol	Meaning	Simulation's value
p	Nature probability that a firm is in ISAC	$(0.2, 0.8)$
ϕ	A coefficient that ensure attacker investment exhibits a diminishing marginal return	0.7
β	A coefficient that ensure security investment exhibits a diminishing marginal return	2
v_d	Firm's information value	$[19, 30]$
v_a	Attacker's profit	$[0, 10, 20]$
c_f	Firm security investment cost	$[4, 17]$
c_a	Attacker investment cost in response to v_a	$[0, 7, 10]$
t	Aggregated defense value in case each firm is in ISAC	0.19
z_f	Firm privacy efficiency	$[0.2, 0.85]$
z_a	Attacker privacy efficiency	$[0, 0.2, 0.55]$

Table 1. Table of parameters used for simulation.

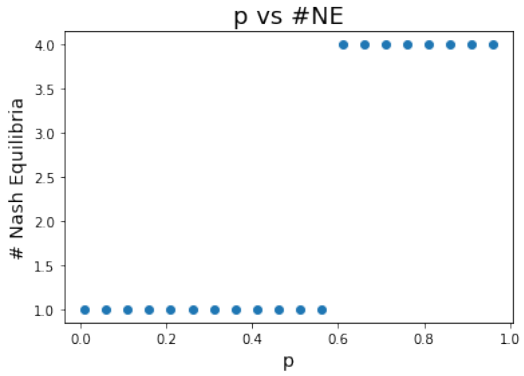

Figure 2. Number of BNE for different value of p

attack when the firm invests low capital in defense, since it suspects the firm to be in ISACs, therefore more difficult to succeed in the attack. Following this reasoning we can say that the adoption of ISACs centre could reduce the presence of low investment attack. As a result a firm that does not enroll in ISACs risk to undergo a strong attack and have a bad payoff. In figure 2 we can see the different number of Bayesian NE depending on the value of p, especially we can notice the presence of the threshold, where the number of BNE increase from one to four.

For the first case scenario (ISACs not reliable) with our choice of parameter and $p = 0.2$, the payoffs of the players are described in the following Tables [2], [3] and [4]

index	Ln	Ll	Lh	Hn	Hl	Hh
Firm	13.7	4.9	3.7	12.2	5.6	4.7
Attacker	0.0	-2.4	0.5	0.0	-4.8	-5.0

Table 2. Payoff values in case the Firm is in ISACs ($p = 0.2$)

index	Ln	Ll	Lh	Hn	Hl	Hh
Firm	15.0	5.3	3.97	13.0	5.82	4.84
Attacker	0.0	-1.89	1.61	0.0	-4.61	-4.56

Table 3. Payoff values in case the Firm is not in ISACs($p = 0.2$)

index	LL	LH	HL	HH
nn	[0,14.74]	[0,13.14]	[0,14.44]	[0,12.84]
nl	[0,14.74]	[-3.68,7.39]	[-0.96,13.12]	[-4.64,5.77]
nh	[0,14.74]	[-3.65,6.60]	[-1.00,12.94]	[4.81,4.81]
ln	[-1.99,5.22]	[-0.48,11.38]	[-1.51,6.68]	[0,12.84]
ll	[-1.99,5.22]	[-4.16,5.64]	[-2.47,5.35]	[-4.64,5.77]
lh	[-1.99,5.22]	[-4.12,4.85]	[-2.51,5.17]	[4.81,4.81]
hn	[1.38,3.92]	[0.09,11.15]	[1.29,5.62]	[0,12.84]
hl	[1.38,3.92]	[-3.59,5.40]	[0.33,4.29]	[-4.64,5.77]
hh	[1.38,3.92]	[-3.55,4.61]	[0.29,4.11]	[4.81,4.81]

Table 4. Expected payoff in case $p = 0.2$

Instead for the opposite case, where p is above the threshold and so organisation like ISACs are reliable, we decided to analyze the case where p = 0.8. The payoff of each combination of actions and the final expected payoff are shown in the following Tables [5], [6], [7].

index	Ln	Ll	Lh	Hn	Hl	Hh
Firm	12.6	6.1	5.2	11.5	6.3	5.6
Attacker	0.0	-3.6	-2.2	0.0	-5.3	-6.1

Table 5. Payoff values in case the Firm is in ISAC($p = 0.8$)

index	Ln	Ll	Lh	Hn	Hl	Hh
Firm	15.0	5.3	3.97	13.0	5.82	4.84
Attacker	0.0	-1.89	1.61	0.0	-4.61	-4.56

Table 6. Payoff values in case the Firm is not in ISAC($p = 0.8$)

Notice that when we are over the threshold the only Bayesian Perfect Equilibrium is (LL, nn), since is a pooling-equilibria (where, regardless of its type, player 1 chooses the same action) the system of belief of the two players are unchanged and are equal to the prior knowledge.

5. Conclusion

We have implemented a Bayesian game between a firm and an attacker taking into account information sharing, privacy cost and the diffusion of ISACs centres. We showed

index	LL	LH	HL	HH
nn	[0,13.11]	[0,12.71]	[0,12.21]	[0,11.81]
nl	[0,13.11]	[-0.92,11.27]	[-4.21,8.05]	[-5.13,6.21]
nh	[0,13.11]	[-0.91,11.07]	[-4.85,7.48]	[5.45,5.45]
ln	[-3.22,5.93]	[-2.84,7.47]	[-0.38,10.27]	[0,11.81]
ll	[-3.22,5.93]	[-3.76,6.03]	[-4.59,6.11]	[-5.13,6.21]
lh	[-3.22,5.93]	[-3.76,5.84]	[-5.28,5.54]	[5.45,5.45]
hn	[-1.41,4.95]	[-1.74,6.76]	[0.32,10.00]	[0,11.81]
hl	[-1.41,4.95]	[-2.68,5.32]	[-3.89,5.84]	[-5.13,6.21]
hh	[-1.41,4.95]	[-2.65,5.12]	[-4.53,5.28]	[5.45,5.45]

Table 7. Expected payoff in case $p = 0.8$

that when ISACs centres will be widespread an attacker will resort to a strategy in which it will avoid low-effort attack. In such a context the investment in privacy security will be key even more.

The next things we can do to improve and have a more complete study of the Game is to transform the discrete variables in continuous one. In the game we have z_f, c_f which represent the efficiency and the cost of the defense for the firm in two different case (low-defense, high-defense) and z_a, c_a that are the efficiency and the cost of the attack in three different case (no-attack, low-attack, high-attack). We can transform those variable in continuous one, defining a range for the possible values, and study how the best strategies of each player will change as the value of those variable is changing. Finally an interesting and correlated project idea is to formulate a coalitional-game between firms, to study which are the best condition to form a grand coalition, so that they will all share their information to ISACs [7].

References

[1] Sarah Brown, Joep Gommers, and Oscar Serrano. From cyber security information sharing to threat management. In *Proceedings of the 2nd ACM Workshop on Information Sharing and Collaborative Security*, WISCS '15, page 43–49, New York, NY, USA, 2015. Association for Computing Machinery.

[2] Huseyin Cavusoglu, Srinivasan Raghunathan, and Wei T. Yue. Decision-theoretic and game-theoretic approaches to it security investment. *Journal of Management Information Systems*, 25(2):281–304, 2008.

[3] Van Schooten van den Berg Dunnewind, Van der Veeken and Shinde. Applying bayesian game theory to analyse cyber risks of bank transaction systems. *Conference: 2016 International Conference on Computing, Analytics and Security Trends (CAST)*, 2016.

[4] Mansooreh Ezhei and Behrouz Tork Ladani. Information sharing vs. privacy: A game theoretic analysis. *Expert Systems with Applications*, 88:327–337, 2017.

[5] Dimitra Markopoulou, Vagelis Papakonstantinou, and Paul Hert. The new eu cybersecurity framework: The nis directive, enisa's role and the general data protection regulation. *Computer Law Security Review*, 35, 11 2019.

[6] 114th Cong. (2015) S.754. To improve cybersecurity in the united states through enhanced sharing of information about cybersecurity threats, and for other purposes.

[7] I. Vakilinia and S.Sengupta. A coalitional game theory approach for cybersecurity information sharing. *Conference: MILCOM 2017 - 2017 IEEE Military Communications Conference (MILCOM)*, 2017.