Game Theoretic Modelling of Cyber Security Information Sharing Schemes

Carlos Cid

Royal Holloway University of London

based on joint work with A. Davidson, G. Fenn, A. Khouzani and V. Pham

29 Feb 2016

Game Theory & Cyber Security

- Topic of this talk: game theoretic modelling of cyber security problems – more specifically cyber security information sharing.
- Game Theory: studies and analyses mathematical models of multi-party interaction of decision-makers.
- assumptions:
  - players *consistently* pursue *well-defined* objectives (i.e. they are *rational*)
  - players take into account knowledge and/or expectations of the other players' preferences and behaviour (i.e. they reason *strategically*)
- game theory has broad applicability in areas such as Economics, Political Sciences and Biology.
  - more recently, game theoretical techniques have also become a popular tool for the analysis and study of problems of relevance to the cyber security community.

Game Theoretic Modelling I

- to model a certain interaction scenario as a game, we need to carefully define:
  - set of **players**
  - the game **rules**: actions available, either simultaneous or sequential playing, level of information available, cooperative or non-cooperative, etc.
  - the *consequences* of a particular *outcome*
- Game Theory: searches for and studies the properties of the **solutions** of such games.

Game Theoretic Modelling II

- game theoretic models may have two goals:
  - to be descriptive (important for *scientists*)
  - to be prescriptive (important for *consultants*)
- game theoretic modelling is almost certainly an *art*:
  - well-constructed models aim to capture the *essence* of the multi-player interaction situation in order to obtain *non-trivial insights* about the decision-making process.
    - it should avoid unnecessary complications ...
    - we then use available game theoretic analytic tools ...
    - to draw interesting lessons.
  - It may not (and in fact, should not) aim to be a faithful representation of reality; too faithful models are typically too complex – and thus intractable in practice – and often of little use.
- Because of this, users should be made aware of the *limitations* of game theoretic models!

Strategic Games

Simultaneous move game (*strategic* game).

- a (finite) set of players N
- sets A_i of *actions* available for each player i.
  - set of outcomes $A = A_1 \times \ldots \times A_N$.
- payoff (or utility) function $u_i : A \to \mathbb{R}$ for each player i.
  - given outcome $a = (a_1, a_2, \ldots, a_N)$, player i *receives* $u_i(a)$.
  - rationality: an action profile is preferred over another if it yields a higher payoff
    (i.e. player i prefers a to b if $u_i(a) > u_i(b)$).

Equilibrium

Classic solution concept: Nash Equilibrium.

Definition (*Nash Equilibrium*)

A action profile $\overline{a} = (\overline{a}_1, \ldots, \overline{a}_N)$ *is a Nash Equilibrium if for every player* $i \in N$, *we have*
$u_i(\overline{a}) = u_i(\overline{a}_1, \ldots, \mathbf{\overline{a}_i}, \ldots, \overline{a}_N) \geq u_i(\overline{a}_1, \ldots, \overline{a}_{i-1}, \mathbf{a_i}, \overline{a}_{i+1}, \ldots, \overline{a}_N)$
for all $a_i \in A_i$.

i.e., no player benefits from *unilateral* change of strategy.

NB: there are several other forms of game, and solution notions:
extensive form games (sequential move games), **Bayesian games** (incomplete information), dominant and dominated strategies, Mixed-strategy, Correlated equilibrium, Evolutionary games, best responses, rationalizable strategies, etc

Prisoner's Dilemma I

Two suspects are interrogated in separate cells. They can decide to confess or not confess their participation in the crime – no communication allowed. Prison sentences will be given depending on their actions.

	Don't Confess	Confess
Don't Confess	(-1,-1)	(-4,0)
Confess	(0,-4)	(-3,-3)

What is the equilibrium?

Prisoner's Dilemma II

Two suspects are interrogated in separate cells. They can decide to confess or not confess their participation in the crime – no communication allowed. Prison sentences will be given depending on their actions.

	Don't Confess	Confess
Don't Confess	(4,4)	(1,5)
Confess	(5,1)	(2,2)

What is the equilibrium?
The Nash Equilibrium provides the lowest *social welfare*, and is the only outcome which is not *Pareto efficient*!

Prisoner's Dilemma: an influential example, perhaps as an illustration that people's (rational) pursuit of their own best interests may lead to outcomes that are bad for **all** players ...

Multiple Equilibria

Battle of the Sexes: should husband/wife go to a football match or to the movies?

	Football	Cinema
Football	(2,1)	(0,0)
Cinema	(0,0)	(1,2)

they should perhaps coordinate, but have conflicting interests.

Coordination Game: should two friends go to football or cricket?

	Football	Cricket
Football	(2,2)	(0,0)
Cricket	(0,0)	(1,1)

they have mutual interest in reaching one of the outcomes.

Mixed-Strategy Equilibrium

Matching Pennies

	Heads	Tails
Heads	(1,-1)	(-1,1)
Tails	(-1,1)	(1,-1)

what is the equilibrium?

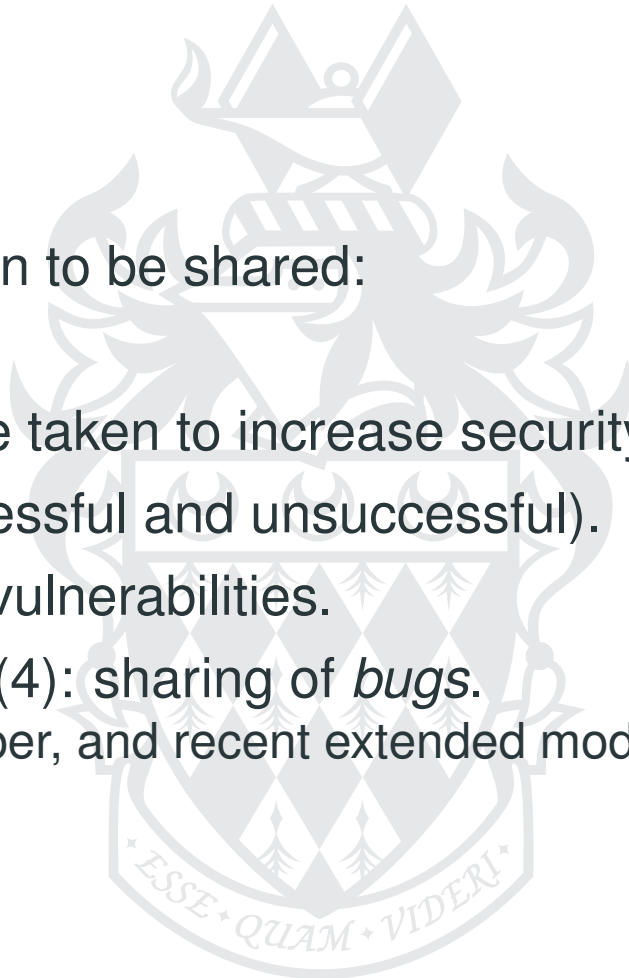
no pure strategy equilibrium, but rather randomisation between the available actions...

ok, after this brief overview, let's move to the research problem of this talk: **sharing of cybersecurity information**.

Cybersecurity Information Sharing

- exchange of *information* is one of the key factors in enhancing the effectiveness of cybersecurity measures. Initiatives include:
  - UK Cyber Security Information Partnerships (CISPs), part of CERT-UK
  - US “Information Sharing and Analysis Centers”
  - Feb 2015 US Executive Order, giving rise to Cybersecurity Information Sharing Act (Dec 2015)
  - private sector: Soltra, ThreatStream and Facebook?s ThreatExchange.
  - development of standards for representing and exchanging threat information (TAXII, STIX and CybOX).
- these are the platforms – incentives can’t however be ignored in the presence of strategic and competing profit maximising entities.
- in this work we consider relevant economics aspects in cyber security information sharing schemes.

Cybersecurity Information Sharing



Some types of information to be shared:

1. *threat intelligence*.
2. what measures to be taken to increase security.
3. past incidents (successful and unsuccessful).
4. discovered security vulnerabilities.

In this work we focus on (4): sharing of *bugs*.
(based on GameSec'14 paper, and recent extended model (under review)).

Model: high level overview I

1. companies may **invest** to discover security vulnerabilities in a specific popular *platform*:
   - number of bugs is unknown.
   - more investment and effort increase the chances of finding them.
   - companies will patch all the vulnerabilities found.
   - the bugs not found are potentially *exploitable*.

Model: high level overview II

2. when a bug is exploited there are **losses** and **gains**:
   - direct losses/gains: an attack will negatively affect the compromised company; it will also benefit the one which patched the vulnerability (customers will migrate to a safer company).
     - thus discovering a bug in a common platform may give a company a *competitive edge* .
   - indirect losses: a successful attack may affect all firms (customers and investors may lose confidence in the whole "sector" of the economy, and seek alternative "safer" options).
   - these effects create opposing incentives for *sharing* the findings.

Model: high level overview III

3. companies decide **sharing** of information:
   - on one hand, sharing information translates to a more effective discovery process due to its probabilistic nature; hence it encourages investment.
   - on the other hand, there can be a tendency of *free-riding* on the discovery investments of other companies.
   - further complicating the problem is *uncertainty* and *information asymmetry*: uncertainties about the total number of bugs, and the other company's number of findings.

Contributions of This Work I

1. model the interdependent **security research investment** and **information sharing decisions** of two strategic and competing firms as a two stage **Bayesian game**.
2. fully determine the **Perfect Bayesian Equilibria** of the game in closed-form.
   - sharing strategies are unique, dominant, and in the simple forms of “full-sharing” or “no sharing”, determined by the competitive nature of the findings.
3. derive the **investments** of the firms knowing their subsequent sharing strategies.
   - “full sharing” leads to free-riding and inefficiently low investments; “no sharing” is also inefficient by preventing mutual benefit of sharing, double-efforts and over-investment.
4. provide a lightweight **mediation** mechanism free of monetary-transfers that enables (partial) sharing of findings when firms fail to achieve sharing on their own.

Contributions of This Work II

1. In an extended model, we consider the case in which bugs can be assigned to different severity levels, and analyse how this affects the trading dynamics.
2. The extended model emphasises the importance of fair trades in the presence of a mediator.
3. We also show how we can remove the need of a trusted third-party mediator, by proposing a cryptographic protocol that securely performs private set operations.

Our Model

2-stage Bayesian game between two firms, in which each decides:

1. how much to **invest** in finding vulnerabilities on a common platform.
2. and subsequently, how many of the found bugs to **share**.

Vulnerabilities, Investment and Sharing

1. **vulnerabilities**
   - the platform has an unknown number of vulnerabilities, described by the random variable B, with mean value λ.
2. **investment**
   - firms decide, simultaneously, how much to invest, and make it publicly known.
   - investment c_i determines probability p_i that each bug will be found, i.e. $\pi_i(c_i) = p_i$.
   - given the properties of pi_i, we model each firm's investing strategy as their *choice* of p_i.
3. **sharing**
   - after investments are made, each firm privately and independently discovers some bugs N_i, some in common (this is the players *types*)
   - state of the world: $\Omega = (B, N_i, N_j, N_{ij})$
   - strategy for sharing: $s_i(p_j, n_i) \leq n_i$ bugs to share.

strategies for the whole game: $\sigma_i = (p_i, s_i)$ and $\sigma_j = (p_j, s_j)$

Model: distribution of bugs

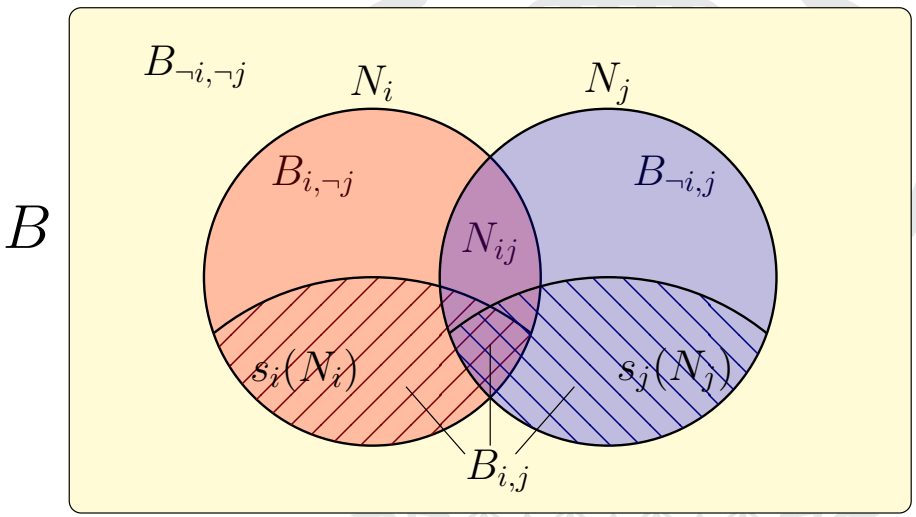


Figure: Venn diagram illustration of the sets of bugs

Model: utilities

losses:

- direct loss $\ell > 0$: affecting only the compromised firm.
- market shrinkage $\tau \geq 0$: all players lose, due to loss of confidence in the whole sector.
- competitive loss $\delta \geq 0$: when only one firm is compromised, it loses δ and the competitor gains δ.

cost of each bug to players:

1. bug is known to both players: $(0, 0)$
2. bug is known to i, but not to j: $(\delta - \tau, -\delta - \tau - \ell)$
3. vice-versa: $(-\delta - \tau - \ell, \delta - \tau)$
4. bug is known to neither: $(-\tau - \ell, -\tau - \ell)$

recall: known bugs are patched, and unpatched bugs are exploited with probability 1.

Model: recall notation...

Parameter	Definition
B, b	Random variable/realisation for the total number of bugs
N_i, n_i	Random variable/realisation for number of bugs discovered by i
N_{ij}	Random var. for number of common bugs discovered by both
s_i	Action of player i: how many discovered bugs to share
λ	Expected number of the total number of bugs
p_i, p_j	Probability that each bug is discovered by player i,j
u_i, u_j	Expected utilities of player i, j
c_i, c_j	Discovery investment cost of player i,j
ℓ	Direct loss upon exploitation of an bug by attackers
δ	Loss (gain) in utility of the player who is the only one attacked (not attacked) – the market competition effect
τ	Loss in utility of both players if a bug is exploited in either one of them – the market section shrinkage effect
$p = \pi(c)$	The relation relating the level of investment c to the discovery probability of a bug p; we use $p = \pi(c) = 1 - e^{-\theta c}$.

Model: utility functions

Expected utility of player i given a realisation of the state of the world $\omega = (b, n_i, n_j, n_{ij})$, $\sigma_i = (p_i, s_i)$ and $\sigma_j = (p_j, s_j)$:

Info Sharing Utility

$$u_i(\omega, \sigma_i, \sigma_j) = -c_i + 0 \cdot \mathbb{E}(B_{i,j}) + (\delta - \tau) \cdot \mathbb{E}(B_{i,\neg j}) + \\ + (-\delta - \tau - \ell) \cdot \mathbb{E}(B_{\neg i,j}) + (-\tau - \ell) \cdot \mathbb{E}(B_{\neg i,\neg j})$$

where $c_i = \pi^{-1}(p_i)$.

Analysis of the Game

We model the security decisions of the two firms as a 2-stage Bayesian game (*investment* then *sharing*):

- since the investment decisions are announced before sharing, each Bayesian game in the second stage is a proper sub game of the whole game.
- we will therefore use backward induction and first analyse the second stage (construct $((p_i, s_i), (p_j, s_j))$ such that s_i, s_j form a Bayesian Nash Equilibrium of the sharing game induced by the choices of p_i, p_j.
- this will determine the utility of players for each choice of (p_i, p_j), allowing to build a simple strategic-form game with actions p_i, p_j corresponding to the first stage.
- a Nash Equilibrium for this game will lead to a proper Perfect Bayesian Equilibrium for the whole game.

2nd stage: sharing bug discoveries

We obtain the following result for the 2nd stage Bayesian game:

Proposition

*If $\delta < \tau$, the unique dominant pure B.N.E. of the second stage of the game is **sharing all the discovered bugs**. If $\delta > \tau$, it is **sharing no information at all**. When $\delta = \tau$, any strategy pair becomes a B.N.E. These hold irrespective of the distribution of the total number of bugs.*

While the proposition is intuitive, it is surprising that the strategies are fully determined by the *relative* values of δ, τ (the market competition effect, market section shrinkage effect, respect.)

1st stage: investment for bug discovery

The firms decide their investment strategies, which will determine the probabilities of finding the vulnerabilities.

- we use $p = \pi(c) = 1 - e^{-\theta c}$.
  - θ measures the efficiency of investment in the firms.
- we look for the solution (p_i, p_j), considering the dominant strategies in the second stage.
- we can treat the first stage as one shot game of investment (for the different cases relating δ, τ).

1st stage: investment for bug discovery

Working with the best responses p^{BR}, for example, when $\delta < \tau$.

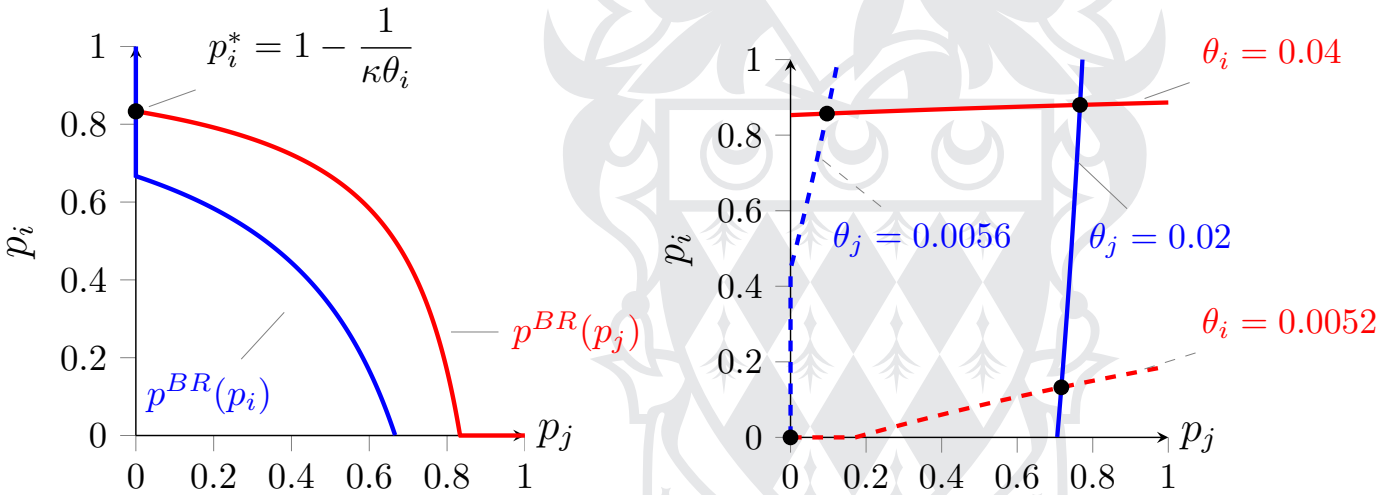


Figure: (a) Example best response curves for the case of $\delta < \tau$. In the figure $\theta_i > \theta_j$; (b) similarly for the case of $\delta < \tau$ and different θ_is and θ_js.

Analysis: PBE for the 2-stage game

Proposition

If $\delta < \tau$ and $\theta_i > \theta_j$, the Perfect Baysesian Equilibrium (PBE) of the two-stage game is that ***only the more efficient firm invests in discovery of the bugs – to achieve discovery probability of*** $[1 - (\kappa\theta_i)^{-1}]^+$ ***– and all the findings are then shared.***

Proposition

When $\delta > \tau$, the Perfect Bayesian Equilibrium (PBE) of the security information sharing game is unique, in which ***both of the firms may invest – to achieve discovery probabilities*** (p_i^*, p_j^*) ***provided in closed form – and none of the consequent findings are shared.***

Comparison with socially optimal outcomes

We also consider the comparison between the PBEs and the socially optimal outcomes:

- $\delta < \tau$: solution has **free-riding**, leading to under investment.
- $\delta > \tau$: solution has **no sharing**, leading to overinvestment and duplicate effort.

Mediation: encouraging information sharing

Trying to prevent some of the social inefficiencies, we introduce a mediator implementing *matched sharing*, which operates in two steps:

1. each player submits its set of found bugs to the mediator, along with a specification of a "threshold": the maximum number of bugs it is willing to exchange with competitor.
2. subsequently, the mediator marks the bugs that are exclusive to each player, and the information of a bug is transferred from player *i* to player *j* iff
   a) there is an exclusive bug to *match*, i.e., to transfer from player *j* to *i*, and
   b) if the total number of bugs transferred so far does not exceed either one of the players' requested maximum threshold.

Note that the mediator is **not** a strategic player, and its behaviour is known and trusted by both players.

Mediation: encouraging information sharing

In this new game, we have the following result:

Proposition

The weakly dominant pure Bayesian Nash Equilibrium of the second stage of the game is ***instructing the mediator to share the maximum number of exclusive bugs.*** *This holds irrespective of the distribution of the total number of bugs, or correlation in the discovery of bugs.*

Extended Model

1. In an extended model, we consider the case in which bugs can be assigned to different severity levels, and analyse how this affects the trading dynamics.
2. The extended model emphasises the importance of fair trades in the presence of a mediator.
3. We also show how we can remove the need of a trusted third-party mediator, by proposing a cryptographic protocol that securely performs private set operations.

Conclusions

We model the strategic decision of investment for discovery of security vulnerabilities, and subsequently sharing the findings by two competing firms as a 2-stage Bayesian game.

This simple modelling provides us with some interesting insights on how incentives and the nature of the sector may affect the firms behaviours:

- we show that sharing all findings becomes a dominant strategy when security tends to behave as a common good, i.e., when the common losses as a result of security attacks outweigh the resulting competitive gains.
  - this in turn leads to free-riding of the less efficient firm, and the under-investment of the more efficient firm.

Conclusions

- we also establish that when security effectively becomes a competitive advantage, i.e., when there is a net positive gain when a competitor is the sole victim of an attack, then sharing no information becomes the dominant strategy, with negative implication on the social efficiency.
- finally, we propose a monetary-free lightweight mediation mechanism that (partially) enables sharing of the found vulnerabilities in cases where they fail to achieve any sharing on their own.
- In an extended model (in which bugs can be assigned to different severity levels), we consider the enhanced role of the mediator, and show how we can remove the need of a trusted third-party mediator, by proposing a cryptographic protocol that securely performs private set operations.

Conclusions

Our models and construction can be seen as a framework that can be built upon to design practical and secure schemes for cybersecurity information sharing between competing entities.

- addressing the conflicting incentives for sharing information, and providing cryptographic guarantees on the security of their private knowledge.

Future Directions

This is an interesting topic, allowing several possible extensions:

- investigating the behaviour of risk-averse players – as opposed to risk-neutral in this work.
- other types of “security information” to share, e.g., past incidents of attacks and losses.
- investigating other means of encouraging sharing, e.g. like “bargaining”, a generalisation of the “matched sharing”’, “joint research ventures”, “exchange market”, etc.

Thank you!

Questions?