Using Game Theory to Analyze Risk to Privacy: An Initial Insight

Lisa Rajbhandari, Einar Snekkenes

Using Game Theory to Analyze Risk to Privacy: An Initial Insight

Lisa Rajbhandari and Einar Arthur Snekkenes

Norwegian Information Security Lab, Gjøvik University College
Gjøvik, Norway
{lisa.rajbhandari,einar.snekkenes}@hig.no

Abstract. Today, with the advancement of information technology, there is a growing risk to privacy as identity information is being used widely. This paper discusses some of the key issues related to the use of game theory in privacy risk analysis. Using game theory, risk analysis can be based on preferences or values of benefit which the subjects can provide rather than subjective probability. In addition, it can also be used in settings where no actuarial data is available. This may increase the quality and appropriateness of the overall risk analysis process. A simple privacy scenario between a user and an online bookstore is presented to provide an initial understanding of the concept.

Keywords: game theory, privacy, risk analysis

1 Introduction

Every individual has a right to the privacy of their personal information. People are dependent on information technology in their daily lives, which includes the risk of their personal information being misused, stolen or lost. The personal information of an individual might be collected and stored by government agencies, businesses and other individuals. These organizations and individuals might have the incentive to misuse such gained information, at least from the perspective of the individual.

In [1], Anderson stated that individuals produce as well as consume commodity information. There are growing problems of identity theft, tracking of an individual, personal information being used as a commodity and so on. Thus, there is a necessity to protect the privacy of information, perform risk analysis and evaluation for proper protection of the entire identity management infrastructure. According to the guidelines of ISO/IEC 27005, for information security risk assessment, risk identification, estimation and evaluation are necessary tasks [2].

In this paper, we suggest that instead of a classical risk analysis like Probabilistic Risk Analysis (PRA), we use a game theory based approach. The distinction between using PRA and Game theory for general risk analysis are shown in Table 1.

Table 1. Comparison of general Risk Analysis steps: Using PRA and Game theory

	Classical Risk Analysis	*Our proposal*
Risk Analysis	**PRA**	**Game theory**
Collect data	Ask for subjective probability or historical data	Ask for preferences or benefits
Compute risk	Compute risk (e.g. expected value)	Compute probability and expected outcome (e.g. mixed strategy Nash equilibrium)
Evaluate	Decide what to do	Decide what to do

In PRA, the risk level is estimated by studying the likelihood and consequences of an event and assigning the probabilities in a quantitative or qualitative scale. Moreover, it can be considered as a single person game because the strategies followed by the opposing player or adversary are not considered. In [3], Bier has stated the challenges to PRA as subjective judgment and human error and performance.

With game theory we can consider settings where no actuarial data is available. Moreover, we do not have to rely on subjective probabilities. By obtaining the preferences or benefits from the subjects, we can compute the probabilities and outcomes to determine the risk. We propose that it can be used in studying and evaluating the behavior of the players in privacy scenarios. It also allows for better audit as the outcomes can be verified at each incident.

In this paper, we will focus on two important issues - 'suitability of game theory for privacy risk analyses' and 'how the payoffs of the players are calculated'.

2 Overview of Game Theory

Game theory is a branch of applied mathematics proposed by John von Neumann and Oskar Morgenstern in paper [4]. It has been used in many fields [5] like economics, political science, evolutionary biology, information security and artificial intelligence. It is the study of the strategic interactions among rational players and their behavior [6], [7], which can be modeled in the form of a game.

For a game, the required four components are: the players, their strategies, payoffs and the information they have [8]. The players are the ones whose actions affect each other's payoffs. Whereas, a strategy is a plan of action that the player can take in response to the opponent's move. It is impossible to act against all the defensive attacks at all times [9]. Thus, it is important to find out the preferred strategies of the players.

The payoff of a particular player is affected by both the actions taken by him and the other player. The thing that matters is that the value of the payoff should be consistent throughout the game. According to Auda, besides ordering the preferences of the payoffs, the players can *'also express the ratio of the differences of the preferences'* on an interval scale called utility [10].

Players make decisions based on the gained information. According to the information they have, the game can be categorized into a perfect/imperfect game and a complete/incomplete game. A complete information game is one in which the players know about the strategies and payoffs of one another (vice versa for the incomplete information game). While, a game where at least a player has no knowledge about the previous actions, as a minimum of one other player is called the imperfect information game (vice versa for the perfect information game).

After determining the components, the game can be represented in the normal (strategic or matrix) form or extensive (tree) form. The normal form is usually used to represent static situations in which the players choose their actions simultaneously and independently. On the other hand, the extensive form is usually used to represent dynamic situations in which the players have some information about the choices of other players.

In a game, each player chooses among a set of strategies to maximize their utilities or payoffs based on the information they have. The choice of strategies can be determined by equilibrium solutions such as the pure strategy Nash equilibrium [11] and the mixed strategy Nash equilibrium, both named after John Forbes Nash.

A strategy profile is a Nash equilibrium if each player's chosen strategy is the best response to the strategies of the other [6] and the players have no incentive to unilaterally deviate. A mixed strategy is a probability distribution (mixing or randomization) of the players' pure strategies so that it makes the other player indifferent between their pure strategies [6], [12]. The equilibrium gives the outcome of the game [8].

3 Why Game Theory?

Today, whenever we have to provide our personal information for instance, while purchasing a ticket online, most of us wonder and are concerned about our information being collected. Some questions that usually pop into our minds are - 'Is our information being stored and if so, to what extent? Who gets access to the stored information? Are all the insiders having access to the stored information 'good'?' If we ask people how often they face risks by providing their personal information, like a credit card number, the answer would be in terms of probability which would be rather vague. However, if we ask them how much they would benefit by providing it, we can have an appropriate answer, for example, in terms of monetary values or time. Thus, with game theory, we can ask expressive questions that the people can answer. Based on these data, risk analysis can be carried out.

In addition, we can perform risk analysis more accurately if we place the situation in the form of a 'game'. If we consider a game of poker, the players are rational. They not only think about their action but also what the other players will do in return to their own particular move. Kardes and Hall state that Probabilistic Risk Analysis (PRA) does not consider the strategies of the

adversary and thus, suggest using the game theoretic approach [13]. In the real world, we have to plan our moves considering the moves of the others, especially if the opponent is an adversary. By using game theory, we can find out how the players choose their strategies in different situations of interdependence. For instance, let us consider zero-sum and cooperative games. In a zero-sum game, gain to one player is a loss to another. Thus, each player takes individual actions to maximize their own outcome. In cooperative games, players cooperate or negotiate their actions for the benefit of each other.

Moreover, the adversary usually does not give up when his attempts have been defended; he rather uses different strategies. In a game theoretic setting, the benefits are based on outcomes and the incentives of the players are taken into account. Thus, game theory helps to explore the behavior of real world adversaries [9].

4 Scenario and Game Formulation

In this section, we will look at the scenario between a user and an online bookstore and the steps used to formulate the scenario in a game theoretic setting.

4.1 Scenario

The user subscribes to a service from an online bookstore. The online bookstore collects and stores additional private information such as book and magazine preferences. These preferences can then be used according to the privacy policy of the online bookstore to provide customized purchase recommendations.

When these recommendations are selected, they generate additional sales for the online bookstore. Also, these recommendations are beneficial to the user, as they save him valuable time. However, it is somewhat tempting for the online bookstore to breach the agreed privacy policy by providing these additional preferences of the user to third parties to be utilized for marketing. This third party marketing incurs additional costs for the user, mostly in terms of time wasting activities like advertisements. However, at the initial stage, the online bookstore cannot determine whether the given information is genuine or fake.

4.2 Game Formulation

For formulating the game, we take into account of the following assumptions- it is a complete information game but of imperfect information. The game is of complete information as we assume both the user and the online bookstore know about the strategies and outcomes of one other. It is of imperfect information as we stipulate that they have no information about the previous action taken by the other player when they have to make their decision. Moreover, it is a one shot game between the user and the online bookstore as a single interaction between them is considered and their actions are taken to be simultaneous and independent.

We now explain the strategies of the players and how the data are collected to estimate the payoffs. We then represent the game in the normal form.

Players. It is a two player game, between the user and the online bookstore. We assume that both the players are intelligent and rational. They have the incentive to win or to optimize their payoffs.

Strategies. We will use a set of simple strategies for this two player non-cooperative game between a user and an online bookstore. The user has the choice to either provide his genuine or fake personal information, knowing the possibility of his information being sold. The strategies of the user are given by {GiveGenuineInfo, GiveFakeInfo}.

The online bookstore either exploits the personal information of the user by selling it to third parties or does not exploit and uses it for its own internal purpose, given by {Exploit, NotExploit}.

Payoff. For obtaining the payoffs of the players, we collect the data and then estimate it.

1. **Data collection:** We have assumed the values for the user and the online bookstore as shown in Table 2. However, it can be collected by conducting experiments and surveys. The values are in hours, reflecting saved or lost time. A positive value represents the hours saved while a negative value represents the hours lost. The profit corresponds to the hours of work saved. The variables 'a' to 'h' are used to represent the cells. The values of the user and the online bookstore are explained below.

Table 2. Assumed saved or lost hours for the user and online bookstore.

	For user		**For online bookstore**	
Information provided by the user	**Genuine**	**Fake**	**Genuine**	**Fake**
The online bookstore usage of information for its internal purpose	(a) 1	(b) 0.1	(c) 1	(d) -0.01
The online bookstore usage of information by selling it to third parties	(e) -1	(f) -0.01	(g) 0.5	(h) -0.2

For the user : If the online bookstore uses the user's preferences and personal information for its own internal purpose according to the policy, we assume

that the user saves an equivalent of an hour, if he had provided genuine personal information, and 0.1 hours if he had given fake information. However, if the online bookstore sells the information to third parties, the user wastes time dealing with the sale attempts. Thus, we assume that the user loses an hour if he had provided genuine personal information, and 0.01 hours if he had provided fake information.

For the online bookstore : Similarly, when the bookstore uses the user's personal information for its own internal purpose according to the policy, we assume it saves an hour if the user had provided genuine information whereas, it loses 0.01 hours dealing with the fake information of the user. However, when it violates the privacy policy and sells the information to third parties, it saves 0.5 hours in case the user had provided genuine information and loses 0.2 hours in case of fake information. We stipulate that it is not possible to assess if private information is fake or not before the information sale is finalized.

2. **Estimation:** We can represent the game in a two players normal form as shown in Fig. 1. The first value of each cell given by x_{ij} is the payoff of the user while the second value given by y_{ij} is the payoff of the online bookstore. Here, $i = 1$ to n, $j = 1$ to n and $n =$ number of players. Each value of the cell is explained and estimated below, along with stating how much each of the players influences the outcome.

   The payoffs are in utility, estimated using the assumed values of hours from Table 2. We have to keep in mind that when the online bookstore exploits the information, it uses the information for its own internal purpose as well as to gain profit by selling it to third parties.

User \ Online bookstore	Exploit	NotExploit
GiveGenuineInfo	x_{11}, y_{11}	x_{12}, y_{12}
GiveFakeInfo	x_{21}, y_{21}	x_{22}, y_{22}

Fig. 1. Normal form representation of the scenario.

The first strategy profile **(GiveGenuineInfo, Exploit)** states that the user provides the personal information genuinely, while the online bookstore exploits it by selling it to third parties. Now, we will calculate the values of the payoff for this particular strategy profile-
x_{11}: Even though the user will benefit from the service, he will have to waste time dealing with advertisements and sale attempts by third parties,

incurred from the exploitation by the online bookstore. The user payoff is obtained by adding the cell values of online bookstore usage of information for its internal purpose (a) and usage by selling it to third parties (e). As mentioned earlier, when the bookstore exploits the information, it uses the data for its own purpose and also sells it to the third parties. Thus, the user's payoff is given by: $x_{11} = a + e = 0$.

y_{11}: The online bookstore will be able to utilize and exploit the user's personal information both for legitimate and unauthorized usage. However, it will not lose time. Thus, y_{11} is obtained by summing the cell values of online bookstore usage of information for its internal purpose (c) and usage by selling it to third parties (g) i.e. $y_{11} = c + g = 1.5$.

The payoffs for the strategy profile **(GiveGenuineInfo, NotExploit)** is estimated as given below-
x_{12}: As the user provides his genuine information, he will receive a customized service in accordance with the agreed privacy policy and save time by utilizing the recommendations. Thus, x_{12} equals the cell value 'a', which is obtained as the user provides genuine information and the online bookstore uses the information for its internal purpose i.e. $x_{12} = a = 1$.

y_{12}: The online bookstore will be able to utilize personal data to offer an improved service in accordance with the agreed privacy policy and will not lose time. Thus, y_{12} equals the cell value 'c', which is obtained as the user provides genuine information and the online bookstore uses the information for its internal purpose i.e. $y_{12} = c = 1$.

The payoffs for the strategy profile **(GiveFakeInfo, Exploit)** is estimated as given below-
x_{21}: The online bookstore will try to exploit the data, but later on, will discover that the data was incorrect. The user will only have some limited benefits from the service but will not lose time dealing with the sale attempts from third parties. Thus, x_{21} is obtained by summing the cell values of the online bookstore usage of information for its internal purpose (b) and usage by selling it to third parties (f) i.e. $x_{21} = b + f = 0.09$.

y_{21}: The fake data provided by the user may only be discovered by the online bookstore at a later stage, for example, at the time when the fake information is to be used to generate profit. The online bookstore will receive limited benefits from the interaction and lose time dealing with the fake data. Thus, the online bookstore's payoff is obtained by summing the cell values of the online bookstore usage of information for its internal purpose (d) and usage by selling it to third parties (h) i.e. $y_{21} = d + h = -0.21$.

The payoffs for the strategy profile **(GiveFakeInfo, NotExploit)** is estimated as given below-

x_{22}: The online bookstore will try to use this fake information given by the user to provide a customized service. However, the user will not receive any benefits and saves less time, as the improved service generated from fake data will be irrelevant. Thus, the user's payoff equals the cell value 'b', which is obtained as the user provides the fake information and the online bookstore uses the information for its internal purpose i.e. $x_{22} = b = 0.1$.

y_{22}: As the user provides fake information, the online bookstore will not be able to provide a customized service, resulting in reduced future sales. Moreover, it will lose time dealing with the fake data. Thus, the online bookstore's payoff equals the cell value 'd', which is obtained as the user provides the fake information and the online bookstore uses the information for its internal purpose i.e. $y_{22} = d = -0.01$.

The normal form representation with the estimated payoffs is given in Fig. 2.

		Online bookstore	
	User	*q*	*1-q*
		Exploit	**NotExploit**
p	**GiveGenuineInfo**	0 , 1.5	1 , 1
1-p	**GiveFakeInfo**	0.09 , -0.21	0.1 , -0.01

Fig. 2. Normal form representation of the scenario with estimated payoffs.

5 Game Solution

5.1 Pure/Mixed Strategy Nash Equilibrium

Using the above payoffs, we found that the game has no pure strategy Nash equilibrium as the players do not agree on a particular strategy profile. However, we can always find the mixed strategy Nash equilibrium.

For obtaining the mixed strategy Nash equilibrium, we will use the calculation as explained in [6](p. 123).We assume that the user plays the strategies GiveGenuineInfo and GiveFakeInfo with probabilities p and 1-p respectively, for $(0 \leq p \leq 1)$. After the calculation, we get $p = 0.29$. Thus, the user plays with the mixed strategy $(0.29, 0.71)$. This means that the user provides genuine information with a 0.29 probability and fake information with a 0.71 probability when playing this game.

Similarly, assume that the online bookstore plays the strategies Exploit and NotExploit with probabilities q and $1 - q$ respectively, for $(0 \leq q \leq 1)$. After

the calculation, we obtain the mixed strategy as $(0.91, 0.09)$ for the strategy profile $(\text{Exploit}, \text{NotExploit})$. Hence, with this mixed strategy we can know the probabilities with which each of the players will choose a particular strategy.

5.2 Expected Outcome

We can represent the game in the normal form with the matrix A. Then, a_{ij} represents each cell of the matrix. In case of a two player game, the expected outcome of the game using mixed strategy to each player is given by

$$\sum_{i=1}^{k} \sum_{j=1}^{l} p_i \, q_j \, a_{ij} \quad . \tag{1}$$

where,
i- number of strategies of player 1 (user) $(1 \leq i \leq k)$,
j- number of strategies of player 2 (online bookstore) $(1 \leq j \leq l)$,
p_i- probabilities with which player 1 plays each of his strategies
$(0 \leq p_i \leq 1)$, $\sum p_i = 1$,
q_j- probabilities with which player 2 plays each of his strategies
$(0 \leq q_j \leq 1)$, $\sum q_j = 1$.

By using (1) and substituting the values of p and q, the expected outcome of the game for the user and the online bookstore is 0.09 and 0.28 respectively. We can conclude that by playing this game, the online bookstore benefits more than the user.

The overall values of the expected outcome that the players get by playing each of the strategies can also be estimated which are given in Fig. 3. The benefit to each of the player can be based on these outcomes.

Expected outcome			0.25	0.03	Sum: 0.28
			q = 0.91	1-q = 0.09	
		Online bookstore / User	Exploit	NotExploit	
0.03	p = 0.29	GiveGenuineInfo	0, 1.5	1, 1	
0.06	1-p = 0.71	GiveFakeInfo	0.09, -0.21	0.1, -0.01	
Sum: 0.09					

Fig. 3. Normal form representation along with the probabilities and expected outcomes.

6 Discussion

We formulated the scenario in the form of a strategic game. We used the concept of mixed strategy Nash equilibrium to compute the probabilities with which the players play each of their strategies, the expected outcome the players gain by playing each of the strategies and also the expected outcome of the game for each player.

Risk analysis can then be based on these computed probabilities and outcomes. However, the following two issues needs to be considered-

The first issue is the preference of the players. It is important to understand the uncertainties related to the preferences of the players in any game. The players might think differently, which may lead them to choosing a different strategy than the equilibrium. Some of the questions that need to be taken into account are -

1. Does the user know what the online bookstore prefers and how he orders the preferences and vice versa?
2. What are the consequences if the two players play 'different games' i.e. it differs in the perception of outcome?

The second is obtaining appropriate data by conducting experiments, interviews and surveys.

In addition, this scenario in a real world situation is usually of partial information. The user knows the exact value of his own 'saved/lost' time. The online bookstore knows the distribution of saved/lost time of the user from the population of all users. However, the online bookstore cannot guess the exact value because, at a given instant, it does not know with which user it is playing the game while the saved/lost time of the online bookstore is known by all users.

7 Conclusion

We can conclude that, with game theory, risk analysis can be based on the computed expected outcomes and probabilities rather than relying on subjective probability. For demonstrating this, we have considered a simple scenario between the online bookstore and its user. Moreover, we have explained how the data can be collected for estimating the payoffs.

The present study provides a starting point for further research. We will conduct a survey for gathering data as the next step. Further, the main objective of the research will be to incorporate the use of game theory in real world privacy scenarios besides the theoretical details.

Acknowledgment. The work reported in this paper is part of the PETweb II project sponsored by The Research Council of Norway under grant 193030/S10.

References

[1] Anderson, H.: The privacy gambit: Toward a game theoretic approach to international data protection. Vanderbilt Journal of Entertainment and Technology Law **9**(1) (2006)

[2] ISO/IEC 27005: Information technology -security techniques -information security risk management (2008)

[3] Bier, V.: Challenges to the acceptance of probabilistic risk analysis. Risk Analysis **19** (1999) 703–710

[4] von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press (1944) Princeton, NJ.

[5] Shoham, Y.: Computer science and game theory. Commun. ACM **51** (2008) 74–79

[6] Watson, J.: Strategy : An Introduction to Game Theory. 2nd edn. W. W. Norton & Company (2008)

[7] Ross, D.: Game Theory. The Stanford Encyclopedia of Philosophy (2010) http://plato.stanford.edu/archives/fall2010/entries/game-theory/.

[8] Rasmusen, E.: Games and Information: An Introduction to Game Theory. 4th edn. Wiley-Blackwell (2006) Indiana University.

[9] Fricker, J.R.: Game theory in an age of terrorism: How can statisticians contribute?, Springer (2006)

[10] Auda, D.: Game theory in strategy development of reliability and risk management. In: Reliability and Maintainability Symposium, 2007. RAMS '07. Annual. (2007) 467–472

[11] Nash, J.: Equilibrium points in n-person games. Proceedings of the National Academy of Sciences of the United States of America **36** (1950) 48–49

[12] Fudenberg, D., Tirole, J.: Game theory. MIT Press, Cambridge, MA (1991)

[13] Kardes, E., Hall, R.: Survey of literature on strategic decision making in the presence of adversaries. CREATE Report (2005)