An Evaluation Framework for Future Privacy Protection Systems: A Dynamic Identity Ecosystem Approach

David Liau, Razieh Nokhbeh Zaeem and K. Suzanne Barber
The Center For Identity, The University of Texas at Austin. U.S.A.

Keywords: Privacy Protection, Personal Identity Information, Stochastic Game, Policy Evaluation, Identity Ecosystem.

Abstract: Today, more than ever, everyday authentication processes involve combinations of Personally Identifiable Information (PII) to verify a person's identity. Meanwhile the number of identity thefts is increasing dramatically compared to the past decades. As a response to this phenomenon, numerous privacy protection regulations, and identity management frameworks and companies thrive luxuriantly.
In this paper, we leverage previous work in the Identity Ecosystem, a Bayesian network mathematical representation of a person's identity, to create a framework to evaluate identity protection systems. After reviewing the Identity Ecosystem, we populate a dynamic version of it and propose a protection game for a person's PII given that the owner and the attacker both gain some level of control over the status of other PII within the dynamic Identity Ecosystem. Next, We present a game concept on the Identity Ecosystem as a single round game with complete information. We then formulate a stochastic shortest path game between the owner and the attacker on the dynamic Identity Ecosystem. The attacker is trying to expose the target PII as soon as possible while the owner is trying to protect the target PII from being exposed. We present a policy iteration algorithm to solve the optimal policy for the game and discuss its convergence. Finally, an evaluation and comparison of identity protection strategies is provided given that an optimal policy is used against different protection policies. This study is aimed to understand the evolutionary process of identity theft and provide a framework for evaluating different identity protection strategies and in future privacy protection system.

1 INTRODUCTION

In the year of 2018, General Data Protection Regulation(GDPR) had been enforced by the European Union(EU) for regulation of data protection and privacy security for individuals. The discussion on privacy protection had once again become one of the emerging topics in the society. Personally Identifiable Information, known as PII, often refers to all the information in real world that relates to a person. Multiple chosen PII attributes are used in a group for authentication or authorization. As the number of ways that PII is used for authentication and authorization increases, so does the number of thefts against identity. Identity theft can generally be defined as any unauthorized use of a person's identity. In the latest government-published statistics [Harrell, 2017], three general types of incidents of Identity theft are included: unauthorized use or attempted use of an existing account, unauthorized use or attempted use of personal information to open a new account, and misuse of personal information for a fraudulent purpose. According to the document, identity theft had already affected 16.7 million people in the U.S. in the year of 2017.

The study of identity theft [Berghel, 2012] with the concept of PII, including protection/prevention [Shah and Okeke, 2011], management [Yuan Cao and Lin Yang, 2010] [Khattak et al., 2010], recovery [Goode and Lacey, 2017], or risk of exposure [Delaitre, 2006], had been raised to focus throughout the past decade. Companies like Discover, LifeLock and IdentityForce nowadays start providing identity protection services to both individual and business as counter measures against identity theft incidents. It is of our interest to develop a universal framework to compare different identity protection policies with the presence of malicious attackers. In this work, we proposed an identity protection evaluation benchmark that serve as an analytic tool to provide benchmark for different identity protection policies.

From previous analysis of the Identity Threat Assessment and Prediction (ITAP) model [Lacey et al., 2016], which was built from real world identity theft data by the Center for Identity at the University of Texas at Austin, the project established an evidence-

Liau, D., Zaeem, R. and Barber, K.
An Evaluation Framework for Future Privacy Protection Systems: A Dynamic Identity Ecosystem Approach.
DOI: 10.5220/0008913501360143
In *Proceedings of the 12th International Conference on Agents and Artificial Intelligence (ICAART 2020) - Volume 1*, pages 136-143
ISBN: 978-989-758-395-7; ISSN: 2184-433X

based understanding of how one personal identity information(PII) attribute is used to access another. This particular evolutionary process within owner's PII collective results in identity theft incidents that directly causes identity owner's loss of properties. To formally establish understanding within a person's PII collective, previous work from Zaeem *et al.* [Zaeem et al., 2016] has establish the Identity Ecosystem, a Bayesian network representation of a person's identity, to study the identity theft as well as other personal identity related issues. For instance, one could analyze the security level to an authentication method utilizing the power of the Identity Ecosystem [Chang et al., 2018]. The Ecosystem features different relationships between PII attributes to answer three main queries of the real world: The risk of exposure of a certain PII, the cause of exposure, and the cost/liability issue.

The Identity Ecosystem grants us the ability to calculate the probability of exposure of PII given the status of the Bayesian network. On the other hand, stochastic games, introduced by Shapley [Shapley, 1953], have been studied for more than half a century since the original paper. In the original paper, the game was constructed with two players, some finite set of game states and finite actions to both players associated with every state of the game. The transition from a state of the game to another follows a distribution that is controlled by both players through their actions played in the previous state. It also brings in the concept of a "terminal state" which the game would transit into with some positive probability at each state. The game is proved to be in equivalent form to a infinite-horizon game with discounted cost. Later on, many extensions and variants of the original work have been developed and studied [Kumar and Shiau, 1981].

In this paper, we apply our novel dynamic approach to the Identity Ecosystem to create an evaluation platform for future privacy protection system. In addition to that, as part of the evaluation framework, we provided the stochastic game based approach with the dynamic Identity Ecosystem to generate the minimax response to general attacker strategies. Moreover, the method is also used to provided survival evaluation to existing privacy protection systems by generating effective attack strategies against a person's identity. We provide a policy iteration algorithm as well as a running example of the algorithm for the minimax strategy. In Section 2, we cover related academic work. Then we formally explain the idea behind our dynamic Identity Ecosystem and the gaming setup for the stochastic game in Section 3. We provide the basic policy iteration algorithm setup for the specific problem and the running example for the algorithm in Section 4. The result from the algorithm is also illustrated so that readers with minimal understanding of general Markov decision processes can follow the paper without difficulty. Finally, conclusions are given in Section 5.

2 RELATED WORK

In this section, we briefly review the several important concepts and background knowledge behind the evaluation tool we are proposing in the paper. In addition, it is worth mention that this work is an extension of previous work [Liau et al., 2019].

2.1 Bayesian Networks

In order to capture the inter-relationship within a person's PII collective, the Identity Ecosystem treats a person's identity as a Bayesian network. Previous work Zaeem *et al.* [Zaeem et al., 2016] established an example of such probabilistic graphical model with real world security breach data. The work consider each PII as a node in the Bayesian network—a random variable with some distribution given the status information from other related PII. The directed edges in the Bayesian network model represents the casual dependencies between random variables.

2.2 Game Theory in Network Security

Game theory has been widely used to study network security and related topics [Roy et al., 2010] in the past decades. Certain literature can be found [Roy et al., 2010] that apply game theory in general. For static game setup, [Carin et al., 2008] provides a good example about how to apply basic game theory to information warfare. For dynamic games, [Manshaei et al., 2013] surveyed several different concepts of applying game theory to network security problems. Game theory is often used to prove that a protocol is optimized considering all the participants involved in the system. In our work, we study the case where the malicious users, having some partial control over the network, are trying to acquire a specific PII attribute.

The problem of protecting PII that we tackle is very similar to the problems like detection of credit card fraud. In this type of problem, the owner of the credit card is the one that actually has some partial observation of the current status of PII. The owner would like to detect the credit card fraud and take counter measures even before the credit card fraud incident takes place. A concrete example can be found

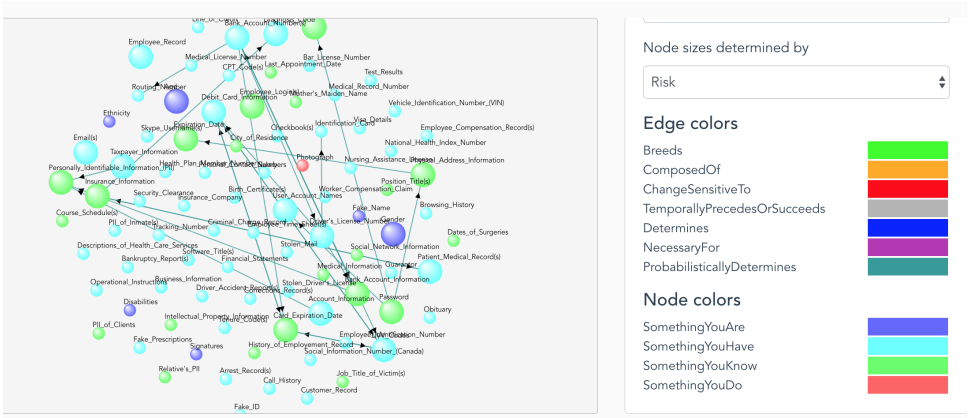

Figure 1: A snapshot of the Identity Ecosystem. In this particular example, the size of the node is determined by the risk of exposure and different colors are used to distinguish the types of PII. It also has the ability to filter and display only related nodes and edges to a specific PII.

in many articles like [Panigrahi et al., 2009] in which a hybrid method to the detection of credit card fraud was studied.

3 MATHEMATICAL REPRESENTATION

In this section, we go through the mathematical representation used throughout the paper including the setup for a dynamic Identity Ecosystem. Then, we provide a single round game example for demonstration of the idea behind our work. Finally, we introduce the concept of treating identity protection as a stochastic game.

3.1 Background: The Identity Ecosystem

The original Identity Ecosystem in [Zaeem et al., 2016] is a Bayesian model of PII attributes and their relationships. Our version of the Identity Ecosystem, as shown in Fig. 1, model is populated with real-world data from approximately 6,000 reported identity theft and fraud cases. We leverage this populated model to provide unique, research-based insights into the variety of PII, their properties, and how they interact. Informed by the real-world data, it enables the investigation of the ecosystem of identifiable information in which criminals compromise PII and misuse them.

As an example query, the identity ecosystem implementation is used to predict future risk and losses of losing a given set of PII and the liability associated with its fraudulent use. In the Bayesian model, each PII (e.g., Social Security Number) is modeled as a graph node. Probabilistic relationships between these attributes are modeled as graph edges. We leverage this Bayesian Belief Network with Gibb's Sampling to approximate the posterior probabilities of the model, assuming the given set of PII attributes is compromised. In addition, once more information about the victim or the incident is available, the Ecosystem is able to refine the predicted risk and value to reflect the new information and converge to the risk and value in the real world. Note that in general discussion, the number for PII attributes that a person could have is a finite number.

Utilizing a probabilistic graphical model to represent the instances of these complicated relationships fits the purpose of understanding different aspects of a person's identity. The effect of PII exposure, for example, is different depending on the status of other PII of a person. For instance, the exposure or theft of a person's Social Security Number (SSN) could result in credit card fraud, identity fraud in mobile devices or even unauthorized access to a person's bank accounts. The count on PII attributes that are at risk of exposure if a person's SSN is leaked may also be different depending on the status of other PII. When the SSN is leaked due to some incident, the impact on personal identity for persons who share their birthday publicly are more severe than those who do not.

3.2 PII Dynamics within Identity Ecosystem

We establish our system dynamic over the Ecosystem to construct our identity protection system evaluation tool. Suppose there is an attacker who has some ability to expose/acquire some of the unexposed PII as he/she wishes. The random variable turns from 0 to 1 when a person's PII is exposed by the attacker. PII

can also become exposed accidentally rather then being intentionally breached by the attacker. On the other hand, the owner of the PII can also actively make some of the exposed PII unexposed if it is allowed to do so (e.g., changing a phone number). The random variable then turns from 1 to 0 when the owner takes action to protect the PII. Now we formally define the state of the network as a $N \times 1$ vector $S_V := [V_1, V_2, \cdots, V_N]$.

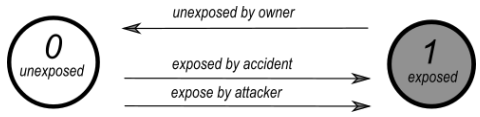

Figure 2: Node Model of PII. The status of the PII can change due to owner protective action (e.g. changing passport) or attacker exposure (e.g. phishing). Another case is by exposure by accident (e.g. subscription).

This setup forms a basic game if we associate a proper cost function to it. Consider two players, one is the owner, who tries to minimize the probability of a certain PII being exposed while the other is the attacker, who tries to maximize the probability of a certain PII being exposed. Without the lose of generality, we consider the basic case where each player can only choose one PII to change.

3.3 Single Round Game with Complete Information

To illustrate our idea of the basic game setup, let us consider a similar but different example where assuming the game will be played only once. In this scenario, we consider a single round, complete information setup, meaning that both the attacker (denoted as malicious player$M$) and the owner (denoted as player U) have the same amount of knowledge as the other one. Let S_U and S_M be the collections of PII that the owner and the attacker can change, respectively. Moreover, the value of each random variable in S_U has to be 1 while the value of each random variable in S_M has to be 0. V_D is the target PII which the owner wants to protect while the attacker wishes to expose. Suppose there is a strictly positive cost function $f : V_1 \times V_2 \times \cdots \times V_N \to \mathbf{R}$ when someone wants to change the value of the node from 0 to 1 and also another strictly positive bounded cost function $g : V_1 \times V_2 \times \cdots \times V_N \to \mathbf{R}$ if someone wants to change the value of the node from 1 to 0.

Recall that in the original Identity Ecosystem, a total of N PII attributes are deployed on the graph to form a Bayesian Network. Link e_{ij} in the Bayesian network is directional, representing a direct influence of V_i to V_j. A *graph* G is defined as a collection of nodes and directional links. Define also a *path* from V_i to V_j as a sequence of nodes and directed links that starts at i and ends at j following the orientation of the graph, i.e.,

$$(V_i \to V_j) = \{V_i, (V_i, V_k), V_k, \cdots, (V_m, V_j), V_j\}$$

where $k, m \in [0, N]$. An action of player i on identity j is defined as A^i_j meaning that player i is trying to change the value of V_j. There exists two sets S_U and S_M denoting the set of PII that the owner and the attacker can change. The following is common properties of the sets

1. $S_U \cap S_M = \emptyset$
2. $\forall V_j \in S_U, \quad V_j = 1$
3. $\forall V_j \in S_M, \quad V_j = 0$
4. $V_D \notin S_M$
5. A action A^i_j is valid if and only if $V_j \in S_i$

Now recall the function $P_D : V_1 \times V_2 \times \cdots V_j \cdots \times V_N \to [0,1], V_j \in V \setminus V_D$ from our ecosystem which gives us the probability that V_D is exposed under the state S_V. Denote the value of PII as V'_j after an action had been taken by both of the players, then the utility function of player i with action A^i_j in the single round game is given as

$$\begin{cases} r_U(A^U_j, A^{-U}) = P_D(V_1, \cdots, V_N) - P_D(V'_1, \cdots, V'_N) \\ r_M(A^M_j, A^{-M}) = P_D(V'_1, \cdots, V'_N) - P_D(V_1, \cdots, V_N) \end{cases}$$

where A^{-i} is the action taken by another player in the game.

Next we provide a basic example of the single round game calculation.

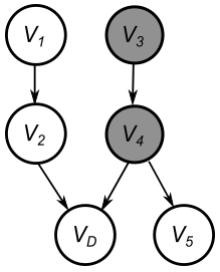

Figure 3: Basic Example for a single round game with complete information. The color indicates whether the specific PII is exposed (grey) or not. The increase or decrease in exposure probability is calculated with the action players choose and the Bayesian inference from the graph.

Consider a smaller instance of the ecosystem in Figure 3. In the example, white vertices are the PII attributes that are unexposed and the grey ones are exposed. Now we can give the game with the following strategic form in Table 1. In this example, the

probability of V_D being exposed is 0.5 for the initial setup $S_V = [0,0,1,1,0]$. For instance, to calculate the difference in probability for action pair (A_U^4, A_M^1), we uses Bayesian inference to calculate the probability of exposure for state $S_V^{'} = [1,0,1,0,0]$ which is 0.4.

Table 1: The strategic form of the basic example in Fig 3.

	None	A_1^M	A_2^M	A_5^M
None	$(0,0)$	$(-0.41, 0.41)$	$(-0.42, 0.42)$	$(0,0)$
A_3^U	$(0.15, -0.15)$	$(-0.4, 0.4)$	$(0.25, -0.25)$	$(0.1, -0.1)$
A_4^U	$(0.25, -0.25)$	$(0.1, -0.1)$	$(-0.3, 0.3)$	$(0.2, -0.2)$

The calculation of the example should give us one unique mixed strategy Nash equilibrium with player M playing A_1^M with probability $8/21$ and A_2^M with $13/21$ while A_5^M strictly dominated by the two strategy. Player U should be playing A_3^U for $11/21$ and A_4^U for $10/21$.

Note that in this case, since both of the parties are only considering the probability difference of the actions, it is a zero sum game in any case under the setup.

3.4 Stochastic Game

Given the discussion in the previous section, we now present the proposed dynamic Identity Ecosystem as follows. The whole dynamic Ecosystem can be considered as a stochastic game, which conducted with many rounds of modified single round version in the previous section. In this the dynamic Ecosystem, each round consists of three phases: decision, exposure and resolve. In the decision phase, we play the game as a single round game where the owner and the attacker each choose certain PII to protect and expose. However, the payoff function here is replaced with a constant for both attacker and owner for every round they played, which we shall explain after setting up the rest of the game. In the exposure phase, each unexposed PII attributes in the network V_j is going to be exposed with a probability given by a strictly positive function $P_j : V_1 \times V_2 \times \cdots V_k \cdots \times V_N \to (0,1], V_k \in V \setminus V_j$ representing accidental exposure. We setup this phase to capture the real world phenomenon that personal PII often get exposed unintentionally by neither the owner nor the attacker. In the last phase, all PII in the Bayesian network are updated with the new status simultaneously and if the target PII is exposed, the game ends in that round. Otherwise the game continues until the target PII (e.g., password, credit card information) is exposed. Different from the single round game, the utility of the owner is now to prolong the game while attacker is trying to end the game as fast as possible. Note that since the objective for both player now is directly connected to the number of rounds that the games is being played, it is convenient for our discussion to setup the payoff function as a constant that both player would receive at each round.

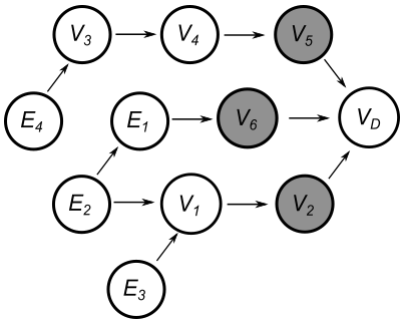

Figure 4: Partial Network of the identity Ecosystem. Note that in this example, the attacker cannot directly expose PII V_4 before PII V_3 is exposed. The figure also illustrates the initial status of the PII in our example.

4 ALGORITHM RESULT

Here we utilize the result from Patek and Bertsekas [Patek and Bertsekas, 1999]. From our setting for the game and above lemmas, assumption $\mathbb{SSP}$ and assumption $\mathbb{R}$ is satisfied thus suggests that there is a unique fix point for the game and the convergence of policy iteration to the fixed point.

Since the system is constructed on real world identity stories, it suggests that the game always ends with the target PII being exposed which is a direct result from the PII exposure probability function P_j which we stated as a strictly positive function if we look at our setup of the game. The interpretation of this corollary is consistent with the construction of the Identity Ecosystem. The Identity Ecosystem is constructed from two main sources, one from manually listing PII relationships and the other from the Identity Threat Assessment and Prediction (ITAP) project at the Center for Identity at the University of Texas at Austin. All of the PII being studied from these two data sources are guaranteed to have some level of exposure since the ITAP project generates the Identity Ecosystem data from fraud cases that took place in the real world. Thus we take the assumption that PII exposure probability function P_j is strictly positive. Define the action chosen by attacker and owner in state i as u^i and v^i respectively. Then we define the policy as the set of actions that specify the player action in every single stage and state the game is being played.

Algorithm 1: Policy Iteration.

State Set $\mathcal{X} := V_1 \times V_2 \times \cdots \times V_N$
State $x, y \in \mathcal{X}$
Actions $A^U(x) = \{V_i, \ldots, V_j\}, A^U(x) : \mathcal{X} \Rightarrow A^U(x)$
Actions $A^M(x) = \{V_l, \ldots, V_k\}, A^M(x) : \mathcal{X} \Rightarrow A^M(x)$
Policy $\nu(x) : \mathcal{X} \Rightarrow \bigcup_{x \in \mathcal{X}} A^U(x)$
Policy $\mu(x) : \mathcal{X} \Rightarrow \bigcup_{x \in \mathcal{X}} A^M(x)$
Cost function $g : \mathcal{X} \times A^U(x) \Rightarrow \mathbf{R}^+$
Transition probability $p_{xy}(A_i^U, A_l^M) = \mathbb{P}(y|x, A_i^U, A_l^M)$

1: **procedure** POLICY ITERATION($\mathcal{X}$, A^U, A^M, g)
2: Initialize $J, J' : \mathcal{X} \to \mathbb{R}_0^+$ arbitrarily
3: **while** J is not converged **do**
4: **for** $x \in \mathcal{X}$ **do**
5: $J(x, \nu(x), \mu(x)) \leftarrow \sum_{y \in \mathcal{X}} p_{xy}(\nu(x), \mu(x))[J(y) + g(y, \nu(x), \mu(x))]$
6: **end for**
7: **end whilereturn** J
8: Improve Policy Set $(\nu(x), \mu(x))$
9: **for** $x \in \mathcal{X}$ **do**
10: $\nu(x)', \mu(x)' \leftarrow \min_{\mu(x)} \max_{\nu(x)} [\mathbb{E}(g(x, \nu(x), \mu(x))) + J(x, \nu(x), \mu(x))]$
11: **end for**
12: $\nu(x), \mu(x) \leftarrow \nu'(x), \mu'(x)$
13: **if** Policy Set $\nu(x), \mu(x)$stable **then return** $(\nu(x), \mu(x))$
14: **else return** to line 3
15: **end if**
16: **end procedure**

4.1 Policy Iteration Algorithm

Consider a partial Identity Ecosystem which we run our algorithm on as in Figure 4. The network structure is referenced from the original Identity Ecosystem but modified and trimmed to better demonstrate how a miniaturized system work. There are 11 nodes including the target PII, 6 controllable PII and 4 evident PII. Note that in this case, we have information about the evident PII which either they are exposed or not in the initial condition. In our experience, when we try to estimate the exposure risk on a person's Identity system, we can actually gather information about the status of some portion of the PII. The evident PIIs $\{E_1, E_2, E_3, E_4\}$ represent the PII that we have the status information. We then assign an initial action to play for the owner along with assigning an initial value to all of the states. We choose the cost function for both the attacker and the owner to be a constant function 1 to simulate the situation in which both the attacker and the owner shall try they best to end or prolong the game without the consideration of costs. Then the attacker calculates its best strategy in the sense of minimizing the payoff of owner. After that the ecosystem runs for one step to determine the accidental exposure of PII for each of the states. Combining the action choice of the attacker, the action choice of the owner and the result from accidental exposure, the Identity Ecosystem generates the transition distribution to all the other states connected under A^U and A^M. The expected cost value is then calculated accordingly before the algorithm enters another iteration.

Note that the algorithm is a greedy algorithm in which the convergence to the unique fixed point is guaranteed while the complexity is high. For each step the algorithm runs, it evaluates the current policies with respect to all the states that the Identity Ecosystem has. While properties of different PII prevent the system from visiting all the possible states, there are still $O(2^N)$ states in general for each evaluation. Within the scope of this paper methods like general Stationary Iterative Method (SIM) [Bertsekas and Tsitsiklis, 1996] or Krylov Subspace Method [Senda et al., 2014] can be applied to the problem to improve the performance of the algorithm.

As our example in Fig. 4, eleven PII attributes are presented including the target PII. We use a uniform cost setup for all the actions that are available to both the attacker and owner, which means that the attacker and owner have no concern about the cost of achieving their goal while trying their best to expose or defend the target PII.

Nodes $V_1 \cdots V_6$ are representing different PII that the owner has. We use another practical setup here to demonstrate the flexibility of the Identity Ecosystem and the algorithm: for PII V_3 and PII V_4, the attacker cannot directly expose V_4 if V_3 was not exposed previously either accidentally or by the attacker.

Table 2: Strategy profile for Identity Ecosystem.

	V_1	V_2	V_3	V_4	V_5	V_6
status	unexposed	exposed	unexposed	unexposed	exposed	exposed
Defend	—	0	—	—	1	0
Attack	0	—	1	0	—	—

For example, at the initial state, the best strategy for the owner is to unexpose V_5 with probability 1 while for the attacker the best strategy is to expose V_3 with probability 1.

The game then enters the next state where the transition is the result of combining actions of the owner and the attacker as well as accidental exposure. When

Table 3: Strategy profile for Identity Ecosystem for another state

	V_1	V_2	V_3	V_4	V_5	V_6
status	unexposed	exposed	exposed	exposed	exposed	unexposed
Defend	—	0	0	1	0	—
Attack	0	—	—	—	—	1

the algorithm converges to its optimal, for every state such strategy profile is provided. The strategy profile in this example is deterministic because of the setup of the exposure probability function. The direct result from our setup is that there exist strong dominant strategies in most of the states. If non-linearity was included in the setup, a general probability distribution among the actions is expected.

4.2 The Evaluation Framework

What can these strategy profiles tell us? The strategy profile from the algorithm result is the minimax solution to the stochastic game problem. The solution solves the Bellman equation which it does not solve for the best strategy against some specific predefined attack strategies. It is actually a conservative solution that tries to maximize the individual payoff among all of the possible strategies that the opponent might play. We utilize this property to generate a minimax strategy profile for both the PII owner and the malicious attacker. While the game is being played, the payoff that both player receive is the same in absolute value, which directly create the incentive for the thief to end the game as soon as possible. This generate a minimax strategy profile that incline with that goal.

On the other hand, given that we have the strategy profile from the algorithm, we can interpret existing privacy protection policies to the system as the owner's strategy to complete against the minimax strategy attacker strategy we have. Based on the number of rounds that the target PII or owner had survive, it is possible to benchmark different identity protection strategies in a numeric way.

Take the partial dynamic Identity Ecosystem in Fig. 4 as an running example, nowadays credit card companies like Discover monitor several crucial PII like Social Security Number in the internet to determine whether the PII is exposed or not instead of actively providing instructions to protect person identity. This type of identity management framework was mentioned in [Mashima and Ahamad, 2008]. We can translate this type of strategies as "passive" strategies into the dynamic Identity Ecosystem. This type of strategies hand-pick PII from past experience and monitor the status of them. For instance, if we create a strategy with similar idea and run the partial system with owner strategy as:

1. Create a watchlist consist of V_5 and V_6.
2. Monitor any of the PII within the watchlist, if any of them get exposed, try to unexposed them as soon as possible.
3. If multiple PII within the list is exposed, randomly choose one of them to unexpose at that round.

Then run the system against the minimax strategy. As the experiment result, we compare the "passive monitor" strategy against randomly selection and minimax strategy. The result is shown in table 4. We run the system 50 times for the sample mean of how many rounds specific strategy survive.

Table 4: Rounds of Survival(Mean) for different Owner Strategies.

	Passive Monitor	Random Selection	Minimax Strategy
number of rounds	5.32	8.72	9.34

We can see from the result that the passive monitor strategy does not perform well in the evaluation. One of the reason is that it only monitor 2 PII attributes out of the 10 within our system. In the ITAP data we have, the Ecosystem utilize 627 PII attributes to define a person's Identity status. In our survey process, most of the protection system does not monitor PII in this scale, as the passive monitor strategy in our experiment. The random selection performs close to the minimax strategy since the system we have is a miniaturized Identity Ecosystem and the probability of exposure assigned to the example does not fluctuate much among the actions that the attacker can take in many of the rounds.

5 CONCLUSION

Revisiting the idea of studying personal identity in the Identity Ecosystem, we introduced our dynamic mechanism for capturing the evolutionary process of identity theft. We also applied game theory to understand the strategies to effectively protect/attack the system. In addition, we presented a complete algorithm by treating the interaction between the owner of identity and the attacker as a stochastic shortest path game and applied it to a partial Identity Ecosystem to demonstrate how strategy profiles work. Finally, we interpreted one of the common identity protection mechanisms from the real world and utilized our generated strategy profile against it. We also provided a new methodology for comparison among different identity protection strategies.

In future work, we wish to evaluate our framework on a full-size Identity Ecosystem. Furthermore, it is particularly an interesting topic to evaluate and compare more existing identity protection strategies with the system we present. Meanwhile, the time complexity of the proposed algorithm can be optimized for certain network topology once more mature results about the dynamic Identity Ecosystem are studied. Finally, we hope that results from this paper can provide

some insight about applicable strategies of protecting the integrity of one's personal identity.

REFERENCES

Berghel, H. (2012). Identity theft and financial fraud: Some strangeness in the proportions. *Computer*, 45(1):86–89.

Bertsekas, D. P. and Tsitsiklis, J. N. (1996). *Neuro-Dynamic Programming*. Athena Scientific, 1st edition.

Carin, L., Cybenko, G., and Hughes, J. (2008). Cybersecurity strategies: The queries methodology. *Computer*, 41(8):20–26.

Chang, K. C., Zaeem, R. N., and Barber, K. S. (2018). Enhancing and evaluating identity privacy and authentication strength by utilizing the identity ecosystem. In *Proceedings of the 2018 Workshop on Privacy in the Electronic Society*, pages 114–120. ACM.

Delaitre, S. (2006). Risk management approach on identity theft in biometric systems context. In *First International Conference on Availability, Reliability and Security (ARES'06)*, pages 4 pp.–982.

Goode, S. and Lacey, D. (2017). Designing and evaluating two interventions to improve identity theft recovery outcomes. In *2017 IEEE International Symposium on Technology and Society (ISTAS)*, pages 1–6.

Harrell, E. (2017). Victims of identity theft. *Bureau of Justice Statistics*.

Khattak, Z. A., Sulaiman, S., and Manan, J. A. (2010). A study on threat model for federated identities in federated identity management system. In *2010 International Symposium on Information Technology*, volume 2, pages 618–623.

Kumar, P. and Shiau, T. (1981). Zero-sum dynamic games. In LEONDES, C., editor, *Advances in Theory and Applications*, volume 17 of *Control and Dynamic Systems*, pages 345 – 378. Academic Press.

Lacey, D., Zaiss, J., and Barber, K. S. (2016). Understanding victim-enabled identity theft. In *2016 14th Annual Conference on Privacy, Security and Trust (PST)*, pages 196–202.

Liau, D., Zaeem, R. N., and Barber, K. S. (2019). Evaluation framework for future privacy protection systems. In *International Conference on Privacy, Security and Trust (PST)*, pages 339–341.

Manshaei, M. H., Zhu, Q., Alpcan, T., Bacşar, T., and Hubaux, J.-P. (2013). Game theory meets network security and privacy. *ACM Comput. Surv.*, 45(3):25:1–25:39.

Mashima, D. and Ahamad, M. (2008). Towards a user-centric identity-usage monitoring system. In *2008 The Third International Conference on Internet Monitoring and Protection*, pages 47–52.

Panigrahi, S., Kundu, A., Sural, S., and Majumdar, A. (2009). Credit card fraud detection: A fusion approach using dempstershafer theory and bayesian learning. *Information Fusion*, 10(4):354 – 363. Special Issue on Information Fusion in Computer Security.

Patek, S. and Bertsekas, D. (1999). Stochastic shortest path games. *SIAM Journal on Control and Optimization*, 37(3):804–824.

Roy, S., Ellis, C., Shiva, S., Dasgupta, D., Shandilya, V., and Wu, Q. (2010). A survey of game theory as applied to network security. In *2010 43rd Hawaii International Conference on System Sciences*, pages 1–10.

Senda, K., Hattori, S., Hishinuma, T., and Kohda, T. (2014). Acceleration of reinforcement learning by policy evaluation using nonstationary iterative method. *IEEE Transactions on Cybernetics*, 44(12):2696–2705.

Shah, M. and Okeke, R. I. (2011). A framework for internal identity theft prevention in retail industry. In *2011 European Intelligence and Security Informatics Conference*, pages 366–371.

Shapley, L. S. (1953). Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10):1095–1100.

Yuan Cao and Lin Yang (2010). A survey of identity management technology. In *2010 IEEE International Conference on Information Theory and Information Security*, pages 287–293.

Zaeem, R. N., Budalakoti, S., Barber, K. S., Rasheed, M., and Bajaj, C. (2016). Predicting and explaining identity risk, exposure and cost using the ecosystem of identity attributes. In *2016 IEEE International Carnahan Conference on Security Technology (ICCST)*, pages 1–8.