A Mean-Field Stackelberg Game Approach for Obfuscation Adoption in Empirical Risk Minimization

Jeffrey Pawlick
New York University Tandon School of Engineering
Department of Electrical and Computer Engineering
Email: jpawlick@nyu.edu

Quanyan Zhu
New York University Tandon School of Engineering
Department of Electrical and Computer Engineering
Email: quanyan.zhu@nyu.edu

Abstract—Data ecosystems are becoming larger and more complex due to online tracking, wearable computing, and the Internet of Things. But privacy concerns are threatening to erode the potential benefits of these systems. Recently, users have developed obfuscation techniques that issue fake search engine queries, undermine location tracking algorithms, or evade government surveillance. Interestingly, these techniques raise two conflicts: one between each user and the machine learning algorithms which track the users, and one between the users themselves. In this paper, we use game theory to capture the first conflict with a Stackelberg game and the second conflict with a mean field game. We combine both into a dynamic and strategic bi-level framework which quantifies accuracy using empirical risk minimization and privacy using differential privacy. In equilibrium, we identify necessary and sufficient conditions under which 1) each user is incentivized to obfuscate if other users are obfuscating, 2) the tracking algorithm can avoid this by promising a level of privacy protection, and 3) this promise is incentive-compatible for the tracking algorithm.

Index Terms—Mean-Field Game, Stackelberg Game, Differential Privacy, Empirical Risk Minimization, Obfuscation

I. INTRODUCTION

We often hear that data is the new oil. On the Internet, websites sell user information to third-party trackers such as advertising agencies, social networking sites, and data analytic companies [12]. In the Internet of things (IoT), devices such as smartwatches include accelerometers, heart rate sensors, and sleep trackers that measure and upload data about users' physical and medical conditions [16]. At a larger scale, smart grid and renewable energy also stand to benefit from developments in networks of sensors and actuators [2].

While these technologies promise positive impacts, they also threaten privacy. Specifically, wearable computing and IoT devices collect sensitive information such as health and location data [1]. In addition, the pervasiveness of tracking allows learners to infer habits and physical conditions over time. For instance, tracking algorithms may predict "a user's mood; stress levels; personality type; bipolar disorder; demographics" [15]. These are unprecedented degrees of access to user information.
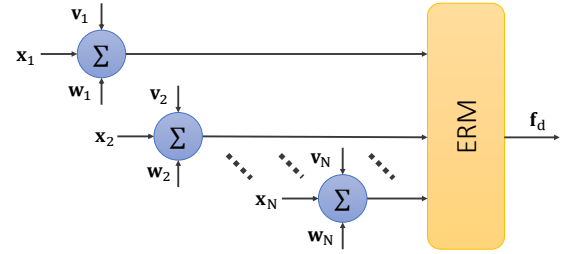

Figure 1. Data flow in the obfuscation-tracking model. Users $1, \ldots, N$ have data $\mathbf{x}_i$ with labels y_i. They add noise $\mathbf{v}_i \sim \mathcal{V}_i$ to $\mathbf{x}_i$, and the learner promises noise $\mathbf{w}_i \overset{\text{i.i.d.}}{\sim} \mathcal{W}$. Noise degrades accuracy but improves privacy. The users and the learner have misaligned incentives.

Interestingly, users have recently begun to take privacy into their own hands using basic tools for *obfuscation*. Obfuscation is "the deliberate addition of ambiguous, confusing, or misleading information to interfere with surveillance and data collection" [3]. In a signal processing sense, obfuscation provides noise.

Finn and Nissenbaum describes two recent obfuscation technologies: *CacheCloak* and *TrackMeNot* [3]. *TrackMeNot* is a browser extension that generates randomized search queries in order to prevent trackers from assembling accurate profiles of their users [11]. In the realm of the IoT, *CacheCloak* provides a way for a user to access location-based services without revealing his or her exact geographical position [13]. The app predicts multiple possibilities for the path of a user, and then retrieves location-based information for each path. An adversary tracking the requests is left with many possible paths rather than a unique one. of these obfuscation technologies see them as a way for users to resist tracking and put pressure on machine learning algorithms to guarantee some privacy protection.

In this paper, we use game theory to identify conditions under which the threat of user obfuscation motivates machine learners to promise privacy protection. We construct a bi-level framework to model this interaction. In the user level,

a large number of users play a mean-field game (MFG) (*c.f.* [4]) to decide whether to use obfuscation. In the learner level, a machine learner plays a Stackelberg game (SG) [17] to decide whether to promise some level of privacy protection in order to avoid obfuscation by the users.

Related work includes research in *privacy markets*, in which a learner pays users to report data truthfully [10], [18]. Our paper differs by allowing the learner to promise privacy protection. In [7], [6], users play a multiple person, prior-commitment game, which determines how much they obfuscate. In these papers, the learner calculates the average of a dataset, while in our framework a learner can use empirical risk minimization to compute more general statistics. Finally, [14] considers a Stackelberg game, but unlike the present paper, it does not include a mean-field interaction among the users. This interaction captures a cascading effect by which many users may rapidly adopt obfuscation technology.

II. MODEL

Figure 1 depicts an interaction between a set of users $i \in \mathbb{S} = \{1, \dots, N\}$ and a learner L. Users submit possibly-perturbed data to L, and L releases a statistic or predictor $\mathbf{f}_d$ of the data. Assume that the data generating process is a random variable $\mathcal{Z}$ with a fixed but unknown distribution. Denote the realized data by $\mathbf{z}_i \overset{\text{i.i.d.}}{\sim} \mathcal{Z}$, $i \in \mathbb{S}$. Each data point is composed of a feature vector $\mathbf{x}_i \in \mathbb{R}^d$ and a label $y_i \in \{-1, 1\}$. The goal of the learner L is to predict y_i given $\mathbf{x}_i$, based on the trained classifier or predictor $\mathbf{f}_d$.

We investigate whether it is advantageous for L to promise some level of privacy protection in order to avoid user obfuscation [1]. L adds noise with the same variance to each data point $\mathbf{x}_i$. For $i \in \mathbb{S}$, $k \in 1, \dots, d$, L draws $w_i^{(k)} \overset{\text{i.i.d}}{\sim} \mathcal{W}$, where $\mathcal{W}$ is a mean-zero Gaussian random variable with standard deviation σ_L. While DP often considers Laplace noise, we use Gaussian noise for reasons of mathematical convenience. Knowing σ_L, each user adds noise $v_i^{(k)} \overset{\text{i.i.d}}{\sim} \mathcal{V}_i$, $k \in 1, \dots, d$, where $\mathcal{V}_i$ is Gaussian with variance σ_S^i. It is also convenient to define $\bar{\sigma}_S^2 = \frac{1}{N} \sum_{i=1}^{N} (\sigma_S^i)^2$, the average variance of the perturbations of every user, and $(\bar{\sigma}_S^{-i})^2 = \frac{1}{N} \sum_{j=1}^{N} (\sigma_S^j)^2 - \frac{1}{N} (\sigma_S^i)^2$, the average variance of the perturbations of every user *other than* i. The perturbed data points are given by $\tilde{\mathbf{x}}_i = \mathbf{x}_i + \mathbf{v}_i + \mathbf{w}_i$, $i \in \mathbb{S}$.

A. Empirical Risk Minimization

Empirical risk minimization (ERM) refers to one popular family of machine learning. In ERM, L calculates a value of an output $\mathbf{f}_d \in \mathbf{F}$ that minimizes the empirical risk, *i.e.*, the total penalty due to imperfect classification of the realized data. Define a *loss function* $l(\tilde{\mathbf{z}}_i, \mathbf{f})$, which expresses the penalty due to a single perturbed data point $\tilde{\mathbf{z}}_i$ for the output

---

$\mathbf{f}$. L obtains $\mathbf{f}_d$ given by Eq. 1, where $\rho \geq 0$ is a constant and $R(\mathbf{f})$ is a regularization term to prevent overfitting:

$$\mathbf{f}_d = \arg\min_{\mathbf{f} \in \mathbf{F}} \rho R(\mathbf{f}) + \frac{1}{N} \sum_{i=1}^{N} l(\tilde{\mathbf{z}}_i, \mathbf{f}), \qquad (1)$$

Expected loss provides a measure of the accuracy of the output of ERM. Let $\mathbf{f}^*$ denote the $\mathbf{f}$ which minimizes the expected loss for unperturbed data:

$$\mathbf{f}^* = \arg\min_{\mathbf{f} \in \mathbf{F}} \mathbb{E}\{\rho R(\mathbf{f}) + l(\mathcal{Z}, \mathbf{f})\}. \qquad (2)$$

In Definition 1, $\mathbf{f}^*$ forms a reference to which the expected loss of the perturbed classifier $\mathbf{f}_d$ can be compared.

Definition 1. (ϵ_g-Accuracy) Let $\mathbf{f}_d$ and $\mathbf{f}^*$ denote the perturbed classifier and the classifier which minimizes expected loss, respectively. Let ϵ_g be a positive scalar. We say that $\mathbf{f}_d$ is ϵ_g-accurate if it satisfies

$$\mathbb{E}\{\rho R(\mathbf{f}_d) + l(\mathcal{Z}, \mathbf{f}_d)\} \leq \mathbb{E}\{\rho R(\mathbf{f}^*) + l(\mathcal{Z}, \mathbf{f}^*)\} + \epsilon_g. \quad (3)$$

Lemma 1 obtains ϵ_g as a function of the obfuscation levels.

Lemma 1. *(Accuracy Level) If L perturbs with variance σ_L^2, user $i \in \mathbb{S}$ perturbs with $(\sigma_S^i)^2$, and the other users perturb with $(\bar{\sigma}_S^{-i})^2$, then the difference ϵ_g in expected loss between the perturbed classifier and the population-optimal classifier is on the order of*

$$\epsilon_g\left(\sigma_L, \bar{\sigma}_S^{-i}, \sigma_S^i\right) \propto \frac{1}{\rho^2 N}\left(\sigma_L^2 + \frac{N-1}{N}\left(\bar{\sigma}_S^{-i}\right)^2 + \frac{1}{N}\left(\sigma_S^i\right)^2\right).$$

B. Differential Privacy

Using differential privacy (DP), a machine learning agent promises a bound ϵ_p on the maximum information leaked about an individual. Let $\mathcal{A}(*)$ denote an algorithm and D denote a database. Let D' denote a database that differs from D by only one entry (*e.g.*, the entry of the user under consideration). Let c be some set among all possible sets C in which the output of the algorithm $\mathcal{A}$ may fall. Then Definition 2 quantifies privacy using the framework of DP [5], [8].

Definition 2. (ϵ_p-Privacy) - An algorithm $\mathcal{A}(B)$ taking values in a set C provides (ϵ_p, δ)-differential privacy if, for all D, D' that differ in at most one entry, and for all $c \in C$,

$$\mathbb{P}\{\mathcal{A}(D) \in c\} \leq \exp\{\epsilon_p\} \mathbb{P}\{\mathcal{A}(D') \in c\} + \delta. \qquad (4)$$

For a cryptographically-small δ, the degree of randomness determines the privacy level ϵ_p. Lower values of ϵ_p correspond to more privacy. That randomness is attained through the noise added in the forms of $\mathcal{V}$ and $\mathcal{W}$.

Lemma 2. *(Privacy Level) If L adds noise with variance σ_L^2 and user $i \in \mathbb{S}$ perturbs with variance $(\sigma_S^i)^2$, then the user obtains differential privacy level $\epsilon_p \in (0, 1)$ on the order of*

$$\epsilon_p\left(\sigma_L, \sigma_S^i\right) \propto \left(\sigma_L^2 + \left(\sigma_S^i\right)^2\right)^{-1/2}. \qquad (5)$$

---

[1] L can accomplish this by collecting data at low resolution. This is consistent with the spirit of DP, in which a learner publishes ϵ_p.
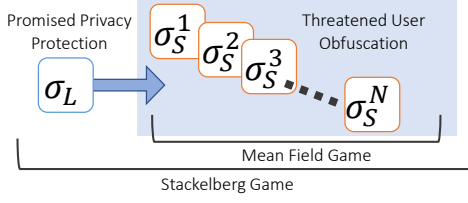
Figure 2. Bi-level structure of the strategic interaction. Users may adopt obfuscation technologies in a cascading manner. This is modeled by an MFG. To avoid this, a learner can proactively add noise to their data. His interaction with the users in modeled by an SG.
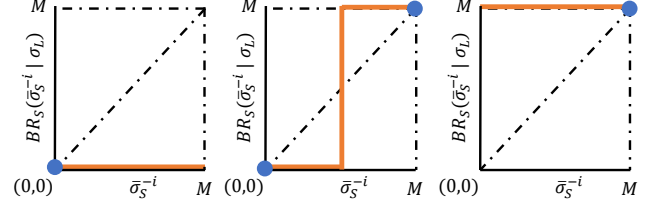

Figure 3. Best response mappings (orange) for each user i against the other users $-i$. MFG equilibria occur at the intersections (blue circles) of the mappings with the identity mapping.

C. Bi-Level Game

Let $\mathbb{R}_M$ denote a subset $[0, M]$ of the non-negative real numbers, and let M be arbitrarily large[2]. Let $\sigma_L \in \mathbb{R}_M$ denote the noise variance added by the learner. If the users are not satisfied with this level of privacy protection, they may add noise with variances $\sigma_S^i \in \mathbb{R}_M$, $i \in \mathbb{S}$.

Define a utility function by $U_L : \mathbb{R}_M^2 \to \mathbb{R}$ such that $U_L(\sigma_L, \bar{\sigma}_S)$ gives the utility that L receives for using noise σ_L^2 while the users add an average noise of $\bar{\sigma}_S^2$. Also define utility functions $U_S^i : \mathbb{R}_M^3 \to \mathbb{R}$ such that user $i \in \mathbb{S}$ receives utility $U_S^i(\sigma_L, \bar{\sigma}_S^{-i}, \sigma_S^i)$ for obfuscating with variance $(\sigma_S^i)^2$ while the other users obfuscate with average variance $(\bar{\sigma}_S^{-i})^2$ and L perturbs with σ_L^2. U_L and U_S^i, $i \in \mathbb{S}$, are given by

$$U_L(\sigma_L, \bar{\sigma}_S) = A_L \exp\left\{-\epsilon_g\left(\sigma_L, \bar{\sigma}_S^{-i}, \sigma_S^i\right)\right\} - C_L \mathbf{1}_{\{\sigma_L > 0\}},$$

$$U_S^i\left(\sigma_L, \bar{\sigma}_S^{-i}, \sigma_S^i\right) = A_S^i \exp\left\{-\epsilon_g\left(\sigma_L, \bar{\sigma}_S^{-i}, \sigma_S^i\right)\right\} - P_S^i\left(1 - \exp\left\{-\epsilon_p\left(\sigma_L, \sigma_S^i\right)\right\}\right) - C_S^i \mathbf{1}_{\{\sigma_S^i > 0\}},$$

where A_L (resp. A_S^i) gives the maximum benefit to the learner (resp. to each user) for output accuracy, P_S^i gives the maximum privacy loss to each user, and C_L (resp. C_S^i) gives the flat cost of perturbation for the learner (resp. to each user).

D. Equilibrium Requirements

Chronologically, L first promises perturbation σ_L, and then the users choose obfuscation σ_S^i, $i \in \mathbb{S}$. The solution, however, proceeds backwards in time.

1) Mean-Field Game: Given the promised σ_L, the group of users plays a MFG in which each user best responds to the average perturbation of the other users[3]. Consider symmetric utility functions for the users[4]. Let $BR_S : \mathbb{R}_M \to \mathbb{R}_M$ denote a best response function, such that

$$BR_S\left(\bar{\sigma}_S^{-i} \mid \sigma_L\right) = \arg\max_{\sigma_S^i \in \mathbb{R}_M} U_S^i\left(\sigma_L, \bar{\sigma}_S^{-i}, \sigma_S^i\right) \quad (6)$$

[2]This rigorously deals with large perturbation variances.

[3]This is a strategic interaction, because each user would prefer to protect her own privacy while making use of accurate data from the other users.

[4]That is, $A_S^i = A_S$, $P_S^i = P_S$, and $C_S^i = C_S$, $i \in \mathbb{S}$.

gives the set of best responses for user $i \in \mathbb{S}$ to the average perturbation $\bar{\sigma}_S^{-i}$ of the other users, given that the learner has promised σ_L^2. Then the equilibrium of the MFG is $\sigma_S^{1*} = \sigma_S^{2*} = \ldots = \sigma_S^{N*}$ (which is also equal to $\bar{\sigma}_S^*$) which satisfies the fixed-point equation

$$\bar{\sigma}_S^* \in BR_S\left(\bar{\sigma}_S^* \mid \sigma_L\right). \quad (7)$$

Now define a mapping $\Gamma : \mathbb{R}_M \to \mathbb{R}_M$ such that $\Gamma(\sigma_L)$ gives the $\bar{\sigma}_S^*$ which satisfies Eq. (7) given[5] σ_L. We say that, by promising σ_L, L *induces* $\bar{\sigma}_S^* = \Gamma(\sigma_L)$.

2) Stackelberg Game: Since L promises σ_L^2 before the users obfuscate, L is a Stackelberg leader, and the users are collectively a Stackelberg follower which plays $\Gamma(\sigma_L)$. The optimality equation for L is

$$\sigma_L^* \in \arg\max_{\sigma_L \in \mathbb{R}_M} U_L\left(\sigma_L, \Gamma(\sigma_L)\right). \quad (8)$$

Definition 3. (Perfect Bayesian Nash Equilibrium) A perfect Bayesian Nash equilibrium (PBNE) (*c.f.*, [9]) of the overall game is $(\sigma_L^\dagger, \sigma_S^{1\dagger}, \sigma_S^{2\dagger}, \ldots, \sigma_S^{N\dagger})$ such that $\bar{\sigma}_S^\dagger = \sigma_S^{1\dagger} = \sigma_S^{2\dagger} = \cdots = \sigma_S^{N\dagger}$, and

$$\bar{\sigma}_S^\dagger = \Gamma\left(\sigma_L^\dagger\right) = BR_S\left(\bar{\sigma}_S^\dagger \mid \sigma_L^\dagger\right), \quad (9)$$

$$\sigma_L^\dagger \in \arg\max_{\sigma_L \in \mathbb{R}_M} U_L\left(\sigma_L, \Gamma(\sigma_L)\right). \quad (10)$$

III. Mean Field Game Analysis

First, Lemma 3 solves for BR_S.

Lemma 3. *(Best Response)* Define $\mathbf{AC}(\sigma_L, \bar{\sigma}_S^{-i}) \triangleq A_S \exp\{-\epsilon_g(\sigma_L, \bar{\sigma}_S^{-i}, 0)\} + C_S$ and $\mathbf{P}(\sigma_L) \triangleq P_S(1 - \exp\{-\epsilon_p(\sigma_L, 0)\})$. Then BR_S is given by

$$BR_S\left(\bar{\sigma}_S^{-i} \mid \sigma_L\right) = \begin{cases} 0, & \text{if } \mathbf{P}(\sigma_L) < \mathbf{AC}\left(\sigma_L, \bar{\sigma}_S^{-i}\right) \\ M, & \text{if } \mathbf{P}(\sigma_L) > \mathbf{AC}\left(\sigma_L, \bar{\sigma}_S^{-i}\right) \\ [0, M], & \text{if } \mathbf{P}(\sigma_L) = \mathbf{AC}\left(\sigma_L, \bar{\sigma}_S^{-i}\right) \end{cases}.$$

Figure 3 depicts Lemma 3. Users with low privacy sensitivity (left) never obfuscate, while users with high privacy sensitivity (right) always obfuscate. Importantly, users with moderate privacy sensitivity (center) cascade: each user i

[5]We will apply a selection criteria to ensure there is only one $\bar{\sigma}_S^*$.

obfuscates if $\bar{\sigma}_S^{-i}$ is high. Theorem 1 states that the MFG equilibria occur at the fixed points of the best response mappings[6].

Theorem 1. *(MFG Equilibrium) Given a promised privacy protection level $\sigma_L^\dagger$, Eq. (9) is satisfied by the symmetric strategies $\sigma_S^{\dagger 1} = \ldots = \sigma_S^{\dagger N} = \bar{\sigma}_S^\dagger$, where $\bar{\sigma}_S^\dagger = \Gamma\left(\sigma_L^\dagger\right)$*

$$= \begin{cases} 0, & \text{if } \mathbf{P}(\sigma_L) < \mathbf{AC}(\sigma_L, M) < \mathbf{AC}(\sigma_L, 0) \\ \{0, M\} & \text{if } \mathbf{AC}(\sigma_L, M) \le \mathbf{P}(\sigma_L) \le \mathbf{AC}(\sigma_L, 0) \\ M & \text{if } \mathbf{AC}(\sigma_L, M) < \mathbf{AC}(\sigma_L, 0) < \mathbf{P}(\sigma_L) \end{cases}.$$

In the middle case, U_S^i is higher for $\bar{\sigma}_S^\dagger = 0$ than for $\bar{\sigma}_S^\dagger = M$. Therefore, we select $\bar{\sigma}_S^\dagger = 0$ and write $\Gamma(\sigma_L) = M\mathbf{1}_{\{\mathbf{P}(\sigma_L) > \mathbf{AC}(\sigma_L, 0)\}}$.

IV. STACKELBERG GAME

Next, L chooses σ_L in order to maximize $U_L(\sigma_L, \Gamma(\sigma_L))$.

A. Status Quo Equilibrium

Lemma 4 gives a solution in which L does not perturb.

Lemma 4. *(Status Quo SG Solution) If $P_S - C_S < A_S$, then $\Gamma_L(0) = 0$. In this case, the optimal $\sigma_L^\dagger = 0$, for which L receives his maximum possible utility: $U_L(0,0) = A_L$.*

$P_S - C_S < A_S$ holds if users are willing to suffer a total loss of privacy in order to obtain complete accuracy. We have called this the *status quo* because it seems to represent the current preferences of many users.

B. Equilibrium Outside of the Status Quo

Consider $P_S - C_S > A_S$. Define $\tau \in \mathbb{R}_M$ such that $\mathbf{P}(\tau) = \mathbf{AC}(\tau, 0)$. By promising to perturb with at least τ, L is able to induce $\Gamma(\tau) = 0$, *i.e.*, to make it incentive-compatible for the users to not obfuscate. But we must analyze whether promising τ is incentive-compatible for L. Since the analytical expression for τ is cumbersome, define an approximation $\hat{\tau} > \tau$, where $\hat{\tau}^2 = 1/\ln\{P_S/(P_S - C_S)\}$. Next, define $\kappa \triangleq 1/(\rho^2 N)$. Then $U_L(\sigma_L, \Gamma(\sigma_L))$ is

$$\approx \begin{cases} 0, & \text{if } \sigma_L^2 = 0 \\ -C_L, & \text{if } 0 < \sigma_L^2 < \hat{\tau}^2 \\ A_L \exp\left\{-\kappa \sigma_L^2\right\} - C_L, & \text{if } \hat{\tau}^2 \le \sigma_L^2 < M \end{cases}.$$

U_L is maximized by either 0 or $\hat{\tau}$ according to Theorem 2.

Theorem 2. *(SG Equilibrium) For $P_S - C_S > A_S$, the perturbation promise which satisfies Eq. (10) is*

$$\sigma_L^\dagger = \begin{cases} 0, & \text{if } \frac{1}{\rho^2 N} > \ln\left\{\frac{A_L}{C_L}\right\} \ln\left\{\frac{P_S}{P_S - C_S}\right\} \\ \hat{\tau}, & \text{if } \frac{1}{\rho^2 N} < \ln\left\{\frac{A_L}{C_L}\right\} \ln\left\{\frac{P_S}{P_S - C_S}\right\} \end{cases}. \quad (11)$$

Theorem 2 shows that high costs C_S of user perturbation incentivize L to promise privacy protection, because users

[6]Mixed strategies are omitted due to limited space.

Table I
EQUILIBRIUM RESULTS OF THE BI-LEVEL GAME

Parameter Regime	$\bar{\sigma}_S^\dagger$	$\sigma_L^\dagger$
1) $P_S - C_S < A_S$	0	0
2) $P_S - C_S > A_S \ \cap \ \frac{1}{\rho^2 N} > \ln\left\{\frac{A_L}{C_L}\right\} \ln\left\{\frac{P_S}{P_S - C_S}\right\}$	M	0
3) $P_S - C_S > A_S \ \cap \ \frac{1}{\rho^2 N} < \ln\left\{\frac{A_L}{C_L}\right\} \ln\left\{\frac{P_S}{P_S - C_S}\right\}$	0	$\hat{\tau}$

easily decide not to obfuscate. On the other hand, high privacy sensitivity P_S decreases L's incentive to add noise. Somewhat surprisingly, high accuracy sensitivity A_L leads L to promise privacy protection[7].

C. Summary of Results

Table I summarizes the results of the overall game. The equilibrium strategies $\bar{\sigma}_S^\dagger$ and $\sigma_L^\dagger$ satisfy Definition 3. Equilibrium 1 is the *status quo* equilibrium in which users submit unperturbed data and L does not protect it. This equilibrium achieves complete accuracy at the cost of complete loss of privacy. In Equilibrium 2, users obfuscate as much as possible. L lacks incentive to promise privacy protection, so he does not perturb. He receives zero utility, making machine learning useless. Equilibrium 3 is the best equilibrium. In this scenario, *the threat* of user obfuscation convinces L to promise privacy protection $\sigma_L^2 = \hat{\tau}^2$. The users accept this level, and do not adopt obfuscation.

V. DISCUSSION OF RESULTS

Privacy skeptics argue that users are not willing to pay for privacy protection. This is captured by $P_S - C_S < A_S$, which leads to Equilibrium 1. But as obfuscation technologies such as *TrackMeNot* [11] and *CacheCloak* [13] continue to develop, the cost C_S of obfuscation will decrease, and the awareness P_S of privacy concerns will increase. Both will lead to $P_S - C_S > A_S$. In Equilibrium 3, obfuscation motivates the learner to promise some level of privacy protection. Nevertheless, technologists should be careful about the effects of obfuscation. In the case of Equilibrium 2, users perturb their data as much as possible, but this only decreases the opportunities for meaningful analysis and discourages machine learning. Our work lays a foundation for identifying the scenarios in which adoption of obfuscation is beneficial.

REFERENCES

[1] Internet of things: Privacy and security in a connected world. Technical report, Federal Trade Commission, January 2015.
[2] R. Baheti and H. Gill. Cyber-physical systems. *The impact of control technology*, 12:161–166, 2011.
[3] Finn Brunton and Helen Nissenbaum. *Obfuscation: A User's Guide for Privacy and Protest*. MIT Press, 2015.
[4] Patrick Chan and Ronnie Sircar. Bertrand and cournot mean field games. *Applied Mathematics & Optimization*, 71(3):533–569, 2015.
[5] Kamalika Chaudhuri, Claire Monteleoni, and Anand D Sarwate. Differentially private empirical risk minimization. *The J. of Machine Learning Research*, 12:1069–1109, 2011.

[7]Accuracy sensitivity increases his sensitivity to user obfuscation.

[6] Michela Chessa, Jens Grossklags, and Patrick Loiseau. A game-theoretic study on non-monetary incentives in data analytics projects with privacy implications. In *Computer Security Foundations Symposium (CSF), 2015 IEEE 28th*, pages 90–104. IEEE, 2015.

[7] Michela Chessa, Jens Grossklags, and Patrick Loiseau. A short paper on the incentives to share private information for population estimates. In *Financial Cryptography and Data Security*, pages 427–436. Springer, 2015.

[8] Cynthia Dwork. Differential privacy. In *Automata, languages and programming*, pages 1–12. Springer, 2006.

[9] D. Fudenberg and J. Tirole. *Game theory*, volume 393. 1991.

[10] Arpita Ghosh and Aaron Roth. Selling privacy at auction. *Games and Economic Behavior*, pages 334–346, 2015.

[11] Daniel C. Howe and Helen Nissenbaum. TrackMeNot: Resisting surveillance in web search. *Lessons from the Identity Trail: Anonymity, Privacy, and Identity in a Networked Society*, 23:417–436, 2009.

[12] Jonathan R. Mayer and John C. Mitchell. Third-party web tracking: Policy and technology. In *Security and Privacy (SP), 2012 IEEE Symposium on*, pages 413–427. IEEE, 2012.

[13] Joseph Meyerowitz and Romit Roy Choudhury. Hiding stars with fireworks: location privacy through camouflage. In *Proceedings of the 15th annual Intl. Conf. on Mobile computing and networking*, pages 345–356. ACM, 2009.

[14] Jeffrey Pawlick and Quanyan Zhu. A stackelberg game perspective on the conflict between machine learning and data obfuscation. In *IEEE Workshop on Information Forensics and Security*, 2016.

[15] Scott R. Peppet. Regulating the Internet of things: First steps toward managing discrimination, privacy, security and consent. *Tex. L. Rev.*, 93:85, 2014.

[16] Melanie Swan. Sensor mania! the Internet of things, wearable computing, objective metrics, and the quantified self 2.0. *J. of Sensor and Actuator Networks*, 1(3):217–253, November 2012.

[17] Heinrich Von Stackelberg. *Marktform und gleichgewicht*. J. Springer, 1934.

[18] David Xiao. Is privacy compatible with truthfulness? In *Proceedings of the 4th Conf. on Innovations in Theoretical Computer Science*, pages 67–86. ACM, 2013.