Privacy-Constrained Communication

Farhad Farokhi ∗ Girish Nair ∗

∗Department of Electrical and Electronic Engineering,
University of Melbourne, Parkville, VIC 3010, Australia
{e-mail: {ffarokhi,gnair}@unimelb.edu.au}

Abstract: A game is introduced to study the effect of privacy in strategic communication between a well-informed sender and a receiver. The receiver wants to accurately estimate a random variable. The sender, however, wants to communicate a message that balances a trade-off between providing an accurate measurement and minimizing the amount of leaked private information, which is assumed to be correlated with the to-be-estimated variable. The mutual information between the transmitted message and the private information is used as a measure of the amount of leaked information. An equilibrium is constructed and its properties are investigated.

Keywords: Data privacy; Information Theory; Communication Systems; Statistical inference

1. INTRODUCTION

Participatory and crowd-sensing technologies rely on honest data from recruited users to generate estimates of variables, such as traffic condition. Providing accurate information by the users can undermine their privacy. For instance, a road user that provides her start and finish points as well as travel time to a participatory-sensing app can significantly improve the quality of the traffic estimation at the expense of exposing her private life. Therefore, she can benefit from providing “false” information, not to deceive the system but to protect her privacy. The amount of the deviation from the truth is determined by the value of privacy that can vary across the population. To better understand this effect, we use a game-theoretic framework to model the conflict of interest and to study the effect of privacy in strategic communication.

We develop a model in which the receiver is interested in estimating a random variable. To this aim, it asks a better-informed sender to provide a measurement. The sender wants to find a trade-off between her desire to provide an accurate measurement of the variable while minimizing the amount of leaked private information, which is potentially correlated with that variable or its measurement. We assume that the sender has access to a possibly noisy measurement of the variable and a perfect measurement of her private information. We use mutual information between the communicated message and the private information to capture the amount of the leaked information. We present a numerical algorithm for finding an equilibrium (i.e., policies from which no one has an incentive to unilaterally deviate) of the presented game when the random variables are discrete. In the continuous case, a fundamental bound on the estimation error is provided for Gaussian random variables and the equilibrium over affine policies are shown to achieve the bound when the emphasis on the privacy grows.

The problem considered in this paper is close, in essence, to the idea of differential privacy and its application in estimation and signal processing, e.g. (Dwork, 2008; Friedman and Schuster, 2010; Huang et al., 2014; Le Ny and Pappas, 2014). Those studies rely on adding noise, typically Laplace noises, to guarantee the privacy of the users by making the outcome less sensitive to local parameter variations. Various studies were devoted to finding “optimal” noise distribution in differential privacy (Soria-Comas and Domingo-Ferrer, 2013; Geng and Viswanath, 2014) or other variants such as Lipschitz privacy (Koufogiannis et al., 2015). Contrary to these studies, we study privacy-constrained communication using game theory by explicitly modeling the conflict of interest between the senders and the receiver stemming from the privacy constraint. Further, we show that in the case of continuous random variables even when the emphasis on the privacy grows, the sender either does not add noise to the transmitted messages or the intensity of the noise does not grow with the value of privacy, which are not the case in the differential privacy literature.

In the information theory literature, wiretap channels have been studied heavily dating from the pioneering work in (Wyner, 1975). In these problems, the sender wishes to devise encoding schemes to create a secure channel for communicating with the receiver while hiding her data from an eavesdropper. This is a secrecy problem. In contrast, other studies have considered the privacy problem in which the masking or equivocation of information corresponds to either the intended primary receiver rather than an eavesdropper (Sankey et al., 2013; Courtade et al., 2012) or a secondary receiver with as much information as the primary one (Yamamoto, 1983, 1988). Information-theoretic guarantees on the amount of leaked private information when utilizing the differential privacy framework was given in (Alvim et al., 2011; du Pin Calmon and Fawaz, 2012). Privacy-aware machine learning was discussed in (Wainwright et al., 2012), where the mutual information between the transmitted message and the original data points is used to measure and constrain the loss of privacy. The contributions of the paper, which sets it apart from the above-mentioned studies, is that we use a strategic game to model the interactions between the sender and the receiver. As we see shortly, there exists an equilibrium in which the players cooperate and thus the problem transforms into a team play; however, there are

* The work of F. Farokhi was supported by a McKenzie Fellowship, ARC grant LP130100605, a grant from Melbourne School of Engineering. The work of G. Nair was supported by ARC grants DP140100819 and FT140100527.
The idea of strategic communication has been studied in the economics literature in the context of cheap-talk games (Crawford and Sobel, 1982) in which well-informed senders communicate with a receiver that makes a decision regarding the society’s welfare. In those games, the sender(s) and the receiver have a clear conflict of interest, which results in potentially dishonest messages. Contrary to those studies, here, the conflict of interest is motivated by the sense of privacy of the sender, which changes the form of the cost functions. Cheap-talk games were recently adapted to investigate privacy in communication and estimation (Farokhi et al., 2015). Contrary to these studies, we consider both discrete and continuous random variables. Privacy constrained information processing with an entropy based privacy measure was studied by (Akay et al., 2015). Contrary to these studies, we consider both discrete and continuous random variables with arbitrary distributions. For the continuous Gaussian random variables, we also present a fundamental bound on the variance of the estimation error at the equilibrium, which is missing from the literature.

The rest of the paper is organized as follows. In Section 2, the problem formulation for discrete random variables is introduced and the equilibria of the game are constructed. The results are extended to continuous random variables in Section 3. The paper is concluded in Section 4.

### 2. DISCRETE RANDOM VARIABLES

We consider strategic communication between a sender and a receiver as depicted in Fig. 1. The receiver wants to have an accurate measurement of a discrete random variable $X \in \mathcal{X}$, where $\mathcal{X}$ denotes the set of all the possibilities. To this aim, the receiver deploys a sensor (which is a part of the sender) to provide a measurement of the variable. The measurement is denoted by $\hat{Z} \in \mathcal{X}$. The sender also has another discrete random variable denoted by $W \in \mathcal{W}$, which is correlated with $X$ and/or $Z$. This random variable is the sender’s private information, i.e., it is not known by the receiver. The sender wants to transmit a message $Y \in \mathcal{Y}$ that contains useful information about the measured variable while minimizing the amount of the leaked private information (note that, because of the correlation between $W$ and $X$ and/or $Z$, an honest report of $Z$ may shine some light on the realization of $W$).

Throughout this paper, for notational consistency, we use capital letters to denote the random variables, e.g., $X$, and small letters to denote a value, e.g., $x$. Assumption 1. The discrete random variables $X, Z, W$ are distributed according to a joint probability distribution $p: \mathcal{X} \times \mathcal{X} \times \mathcal{W} \rightarrow [0, 1]$, i.e., $P[X = x, Z = z, W = w] = p(x, z, w)$ for all $(x, z, w) \in \mathcal{X} \times \mathcal{X} \times \mathcal{W}$.

The conflict of interest between the sender and the receiver can be modelled and analysed as a game. This conflict of interest can manifest itself in the following ways:

1) In participatory-sensing schemes, the sender’s measurement of the state potentially depends on the way that the sender experiences the underlying process or services. For instance, in traffic estimation, the sender’s measurement is fairly accurate on the route that she has travelled and, thus, an honest revelation of $Z$ provides a window into the life of the commuter. However, the underlying state $X$ is not related to the private information of sender $W$ since she is only an infinitesimal part of the traffic flow. In such a case, we have
\[
\mathbb{P}\{X = x, Z = z, W = w\} = \mathbb{P}\{Z = z | X = x, W = w\}\mathbb{P}\{X = x, W = w\} = \mathbb{P}\{Z = z | X = x, W = w\}\mathbb{P}\{X = x\}\mathbb{P}\{W = w\},
\]
where the second equality follows from independence of random variables $W$ and $X$.

2) In many services, such as buying insurance coverage or participating in polling surveys, an individual should provide an accurate history of her life or beliefs. In these cases, the variable $X$ highly depends on the private information of the sender $W$ (if not equal to it). In such cases, the measurement $Z$ may not contain any error as well.

In what follows, the privacy game is properly defined.

#### 2.1 Receiver

The receiver constructs its best estimate $\hat{X} \in \mathcal{X}$ using the conditional distribution $\mathbb{P}\{X = \hat{x} | Y = y\} = \beta_{\hat{x}y}$ for all $(\hat{x}, y) \in \mathcal{X} \times \mathcal{Y}$. The matrix $\beta = \{\beta_{\hat{x}y}(\hat{x}, y) \in \mathcal{X} \times \mathcal{Y} \}$ is the policy of the receiver with the set of feasible policies defined as
\[
B = \left\{(\beta, \beta_{\hat{x}y} \in [0, 1], \forall (\hat{x}, y) \in \mathcal{X} \times \mathcal{Y}) \land \sum_{\hat{x} \in \mathcal{X}} \beta_{\hat{x}y} = 1, \forall y \in \mathcal{Y}\right\}.
\]

The receiver prefers an average measurement of the variable $X$. Therefore, the receiver wants to minimize the cost function $\mathbb{E}[d(X, \hat{X})]$ with the mapping $d: \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$ being a measure of distance between the entries of set $\mathcal{X}$. An example of such a distance is
\[
d(x, \hat{x}) = \begin{cases} 0, & x = \hat{x}, \\ 1, & x \neq \hat{x}. \end{cases}
\]

When using the distance mapping in (1), the term $\mathbb{E}[d(X, \hat{X})]$ becomes the probability of error at the receiver. The results of this paper are valid irrespective of the choice of this mapping.

#### 2.2 Sender

The sender constructs its message $y \in \mathcal{Y}$ according to the conditional probability distribution $\mathbb{P}\{Y = y | Z = z, W = w\} = \alpha_{yzw}$ for all $(y, z, w) \in \mathcal{Y} \times \mathcal{X} \times \mathcal{W}$. Therefore, the tensor $\alpha = \{\alpha_{yzw}(y, z, w) \in \mathcal{Y} \times \mathcal{X} \times \mathcal{W} \}$ denotes the policy of the sender. The set of feasible policies is given by
\[
A = \left\{\alpha: \alpha_{yzw} \in [0, 1], \forall (y, z, w) \in \mathcal{Y} \times \mathcal{X} \times \mathcal{W}
\land \sum_{y \in \mathcal{Y}} \alpha_{yzw} = 1, \forall (z, w) \in \mathcal{X} \times \mathcal{W}\right\}.
\]

The sender wants to minimize $\mathbb{E}[d(X, \hat{X})] + \varrho I(Y; W)$, with $I(Y; W)$ denoting the mutual information between random variables $Y$ and $W$ (Thomas and Cover, 2006), to strike a balance between transmitting useful information about the measured variable and minimizing the amount of the leaked private information. In this setup, the privacy ratio $\varrho$ captures the sender’s emphasis on protecting her privacy. For small $\varrho$, the sender provides a fairly honest measurement of the state. However, as $\varrho$ increases, the sender provides a less relevant message to avoid revealing
her private information through the communicated message.

2.3 Equilibria

The cost function of the sender is equal

\[ U(\alpha, \beta) = E(d(X, \hat{X})) + g(Y; W), \]

where

\[ I(Y; W) = \sum_{y \in \mathcal{Y}} \sum_{w \in \mathcal{W}} P(Y = y, W = w) \times \log \left( \frac{P(Y = y, W = w)}{P(Y = y)P(W = w)} \right) \]

with \( P\{Y = y\} = \sum_{w \in \mathcal{W}} \sum_{x \in \mathcal{X}} P(x, z, w), \) \( P\{W = w\} = \sum_{z \in \mathcal{Z}} \sum_{x \in \mathcal{X}} p(x, z, w), \) and

\[ P\{Y = y, W = w\} = \sum_{z \in \mathcal{Z}} P\{Y = y\} P\{W = w\} \times P\{W = w, Z = z\} \]

\[ = \sum_{z \in \mathcal{Z}} \sum_{x \in \mathcal{X}} \alpha_{yzw} p(x, z, w). \]

Moreover, we have

\[ \mathbb{E}\{d(X, \hat{X})\} = \sum_{x \in \mathcal{X}} \sum_{\hat{x} \in \mathcal{X}} d(x, \hat{x}) P(\hat{X} = \hat{x}, X = x) \]

with

\[ P\{\hat{X} = \hat{x}, X = x\} = \sum_{y \in \mathcal{Y}} P\{\hat{X} = \hat{x}, X = x, Y = y\} \]

\[ = \sum_{y \in \mathcal{Y}} P\{\hat{X} = \hat{x}|X = x, Y = y\} P\{X = x, Y = y\} \]

\[ = \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} \beta_{xy} \sum_{w \in \mathcal{W}} P\{Y = y\} P\{X = x, Z = z, W = w\} \times P\{W = w, Z = z\} \]

\[ = \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} \sum_{z \in \mathcal{Z}} \beta_{xy} \alpha_{yzw} p(x, z, w). \]

Following these calculations, we can define mappings \( \xi : A \times B \rightarrow \mathbb{R} \) and \( \zeta : A \rightarrow \mathbb{R} \) such that

\[ \xi(\alpha, \beta) = \mathbb{E}\{d(X, \hat{X})\} \]

\[ = \sum_{x \in \mathcal{X}} \sum_{\hat{x} \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \sum_{z \in \mathcal{Z}} \sum_{w \in \mathcal{W}} d(x, \hat{x}) \beta_{xy} \alpha_{yzw} p(x, z, w) \]

and

\[ \zeta(\alpha) = I(Y; W) \]

\[ = \sum_{y \in \mathcal{Y}} \sum_{w \in \mathcal{W}} \left( \sum_{x \in \mathcal{X}} \sum_{z \in \mathcal{Z}} \alpha_{yzw} p(x, z, w) \right) \times \log \left( \frac{\sum_{w \in \mathcal{W}} \sum_{x \in \mathcal{X}} \sum_{z \in \mathcal{Z}} \alpha_{yzw} p(x, z, w) P\{W = w\}}{P\{W = w\}} \right). \]

Therefore, we can rewrite the costs of the sender and the receiver, respectively, as \( U(\alpha, \beta) = \xi(\alpha, \beta) + g(\alpha) \) and \( V(\alpha, \beta) = \zeta(\alpha) \). Now, we can properly define the equilibrium of the game.

**Definition 1. (Nash Equilibrium):** A pair \((\alpha^*, \beta^*) \in A \times B\) constitutes a Nash equilibrium of the privacy game if \((\alpha^*, \beta^*) \in N \) with

\[ N = \{(\alpha, \beta) \in A \times B| U(\alpha, \beta) \leq U(\alpha^*, \beta'), \forall \alpha' \in A \]

\[ V(\alpha, \beta) \leq V(\alpha, \beta'), \forall \beta' \in B\}. \]

**Algorithm 1** The best-response dynamics for learning an equilibrium.

**Require:** \( \alpha^0 \in A, \beta^0 \in B \)

1: for \( k = 1, 2, \ldots \)

3: \( \alpha^k \leftarrow \arg \min_{\alpha \in A} U(\alpha, \beta^{k-1}) \)

4: \( \beta^k \leftarrow \beta^{k-1} \)

5: else

6: \( \beta^k \leftarrow \arg \min_{\beta \in B} V(\alpha^{k-1}, \beta) \)

7: \( \alpha^k \leftarrow \alpha^{k-1} \)

8: end if

9: end for

As all signalling games (Sobel, 2012), the privacy game admits a family of trivial equilibria known as babbling equilibria in which the sender’s message is independent of the to-be-estimated variable and the receiver discards sender’s message.

**Theorem 1. (Babbling Equilibria):** Let \( \alpha^* \in A \) be such that \( \alpha^*_{yzw} = 1/|Y| \) for all \( y, z, w \in Y \times X \times W \). Further, let \( \beta^* \in B \) be such that \( \beta^*_{xy} = 1 \) for \( x \in \arg \max_{x \in X} \sum_{y \in Y} \sum_{w \in W} p(x, w). \) Then, \((\alpha^*, \beta^*)\) constitutes an equilibrium.

**Proof.** See Appendix A.

The messages passed at a babbling equilibrium are meaningless and do not contain any information. In what follows, we propose methods for capturing other equilibria of the game.

**Theorem 2.** Any \((\alpha^*, \beta^*) \in \arg \min_{(\alpha, \beta) \in A \times B} \{\xi(\alpha, \beta) + g(\zeta(\alpha))\}\) constitutes an equilibrium of the game.

**Proof.** See Appendix B.

Theorem 2 paves the way for constructing numerical methods to find an equilibrium of the game. This can be done by employing the various numerical optimization methods to minimize the function \( \xi(\alpha, \beta) + g(\zeta(\alpha)) \). However, this is a difficult task as this function is not convex (it is only convex in each variable separately and not in both variables simultaneously).

**Remark 1.** Theorem 2 shows that, for at least one equilibrium of the game, the sender and the receiver cooperate and the problem transforms into a team play. However, it should be noted that there are other equilibria, e.g., those captured by Theorem 1, in which the sender and the receiver do not cooperate.

We can simplify the construction of an equilibrium of the game for the special case where the transmitted message \( Y \) and the to-be-estimated variable \( X \) span over the same set.

**Theorem 3.** Assume that \( Y = X \). Let \( \beta' \) be such that \( \beta'_{xy} = 1 \) if \( x = y \) and \( \beta'_{xy} = 0 \) if \( x \neq y \). Moreover, let \( \alpha' \in \arg \min_{\alpha \in A} \{\xi(\alpha, \beta') + g(\zeta(\alpha))\} \). Then, \((\alpha', \beta')\) constitutes an equilibrium of the game.

**Proof.** See Appendix C.

**Remark 2.** The proof of Theorem 3 reveals that the sender’s policy is the solution of the optimization problem \( \min_{\alpha \in A} \mathbb{E}\{d(X, Y)\} + g(\alpha) \). This problem is equivalent to solving \( \min_{\alpha \in A} \{\xi(\alpha, \beta') + g(\zeta(\alpha))\} \) if \( \beta' \) is an appropriate function of \( \beta \). Therefore, intuitively, the sender aims at providing an accurate measurement of the state \( X \) while bounding the amount of the leaked information.
Algorithm 2 The best-response dynamics for learning an equilibrium.

Require: $\alpha^0 \in A$, $\beta^0 \in B$

1: for $k = 1, 2, \ldots$ do
2: \hspace{1em} if $k$ is even then
3: \hspace{2em} $\alpha' \in \arg\min_{\alpha' \in A} U(\alpha, \beta^{k-1})$
4: \hspace{2em} if $U(\alpha', \beta^{k-1}) - U(\alpha^k, \beta^{k-1}) > \epsilon$ then
5: \hspace{3em} $\alpha^k \leftarrow \alpha'$
6: \hspace{2em} else
7: \hspace{3em} $\alpha^k \leftarrow \alpha^{k-1}$
8: \hspace{2em} end if
9: \hspace{2em} $\beta^k \leftarrow \beta^{k-1}$
10: \hspace{1em} else
11: \hspace{2em} $\beta' \in \arg\min_{\beta' \in B} V(\alpha^{k-1}, \beta)$
12: \hspace{2em} if $V(\alpha^{k-1}, \beta') - V(\alpha^{k-1}, \beta^k) > \epsilon$ then
13: \hspace{3em} $\beta^k \leftarrow \beta'$
14: \hspace{2em} else
15: \hspace{3em} $\beta^k \leftarrow \beta^{k-1}$
16: \hspace{2em} end if
17: \hspace{2em} $\alpha^k \leftarrow \alpha^{k-1}$
18: \hspace{2em} end if
19: \hspace{1em} end for

For more general cases, we can use a distributed learning algorithm to recover an equilibrium. An example of such a learning algorithm is the iterative best-response dynamics. Following this, we can construct Algorithm 1 to recover an equilibrium of the game distributedly. To present our results, we need to introduce a more practical notion of equilibrium.

Definition 2. ($\epsilon$-Nash Equilibrium): For all $\epsilon > 0$, a pair $(\alpha^*, \beta^*) \in A \times B$ constitutes an $\epsilon$-Nash equilibrium of the privacy game if $(\alpha^*, \beta^*) \in \mathcal{N}_\epsilon$, with

$$
\mathcal{N}_\epsilon = \{ (\alpha, \beta) \in A \times B \mid U(\alpha, \beta) \leq U(\alpha', \beta) + \epsilon, \forall \alpha' \in A, V(\alpha, \beta) \leq V(\alpha, \beta') + \epsilon, \forall \beta' \in B \}.
$$

This notion of equilibrium means that each player cannot gain by more than $\epsilon$ from unilaterally changing her actions, which is a practical notion if the act of changing her actions has “some cost” for the player. Now, we are ready to prove that Algorithm 1 can extract an $\epsilon$-Nash equilibrium.

Theorem 4. For $\{(\alpha^k, \beta^k)\}_{k \in \mathbb{N}}$ generated by Algorithm 1 and all $\epsilon > 0$, there exists $K_\epsilon \in \mathbb{N}$ such that $(\alpha^k, \beta^k) \in \mathcal{N}_\epsilon$ for all $k \geq K_\epsilon$.

Proof. See Appendix D.

Theorem 4 shows that Algorithm 1 converges to an $\epsilon$-Nash equilibrium, for any $\epsilon > 0$, in a finite number of iterations. We can slightly tweak Algorithm 1 to also present bounds on the required number of iterations to extract an $\epsilon$-Nash equilibrium.

Theorem 5. For $\{(\alpha^k, \beta^k)\}_{k \in \mathbb{N}}$ generated by Algorithm 2 and all $\epsilon > 0$, $(\alpha^k, \beta^k) \in \mathcal{N}_\epsilon$ for all $k \geq 2 + \Psi(\alpha^0, \beta^0)/\epsilon$.

Proof. See Appendix E.

Example 1. Consider an example with $\mathcal{X} = \mathcal{W} = \mathcal{Y} = \{1, \ldots, 5\}$. Assume that $\mathcal{Z} = \mathcal{X}$, i.e., the sender has access to the perfect measurement of $X$. Moreover, let

$$
\begin{pmatrix}
0.14 & 0.02 & 0.01 & 0.01 & 0.02 \\
0.02 & 0.14 & 0.02 & 0.01 & 0.01 \\
0.01 & 0.02 & 0.14 & 0.02 & 0.01 \\
0.01 & 0.01 & 0.02 & 0.14 & 0.02 \\
0.02 & 0.01 & 0.01 & 0.02 & 0.14 \\
\end{pmatrix}
$$

This distribution implies that there is a reasonable correlation between $X$ and $W$. Therefore, for high enough $\rho$, we expect a good estimation quality (at the receiver) since, otherwise, the receiver can recover $X$, which carries a significant amount of information about $W$. For this example, we use the results of Theorem 3 to extract a nontrivial equilibrium for each $\rho$. Fig. 2 (top) illustrates the estimation error $\mathbb{E}\{d(X, \hat{X})\}$ as a function of the privacy ratio $\rho$. As we expect, by increasing $\rho$, the sender puts more emphasis on protecting her privacy rather than providing a useful measurement to the receiver, and therefore, $\mathbb{E}\{d(X, \hat{X})\}$ increases. Fig. 2 (bottom) shows the mutual information $I(Y; W)$ as a function of the privacy ratio $\rho$. Evidently, with increasing $\rho$, the amount of leaked information about the private information of the sender decreases. In both figures, there seems to be sudden change when the privacy ratio passes the critical value $\rho = 0.38$. That is, for all $\rho < 0.38$, the truth-telling seems to be an equilibrium of the game; however, if $\rho > 0.38$, the sender adds false reports to protect her privacy.

3. CONTINUOUS RANDOM VARIABLES

In this section, we consider continuous random variables $X, Z \in \mathcal{X} \subseteq \mathbb{R}^n$ and $W \subseteq \mathcal{W} \subseteq \mathbb{R}^m$. The following assumption is made.

Assumption 2. The continuous random variables $X, Z, W$ are distributed according to a joint probability density function $p : \mathcal{X} \times \mathcal{X} \times \mathcal{W} \to \mathbb{R}_{\geq 0}$, i.e., $P\{X \in \mathcal{X}, Z \in \mathcal{Z}, W \in \mathcal{W}\} = \int_{x \in \mathcal{X}} \int_{z \in \mathcal{Z}} \int_{w \in \mathcal{W}} p(x, z, w) \text{d}w \text{d}z \text{d}x$ for all Lebesgue-measurable sets $\mathcal{X}, \mathcal{Z}, \mathcal{W}$.

The receiver constructs its best estimate $\hat{X} \in \mathcal{X}$ using the conditional density function $\beta(\cdot|Y) : \mathcal{X} \to \mathbb{R}_{\geq 0}$, i.e., $P\{X \in \mathcal{X} | Y\} = \int_{x \in \mathcal{X}} \beta(x|Y) \text{d}x$ for all Lebesgue-measurable sets $\mathcal{X} \subseteq \mathcal{X}$. Let us denote the set of all such policies by $B$. Similarly, the receiver prefers an accurate measurement of the variable $X$. Therefore, the receiver wants to minimize the cost function $\mathbb{E}\{d(X, \hat{X})\}$, where

$$
d(x, \hat{x}) = \|x - \hat{x}\|^2_2.
$$

The sender constructs its message $y \in \mathcal{Y} \subseteq \mathbb{R}^n$ according to the conditional probability density function $\alpha(\cdot|Z, W) : \mathcal{X} \times \mathcal{W} \to \mathcal{Y}$, i.e., $P\{Y \in \mathcal{Y} | Z, W\} = \int_{y \in \mathcal{Y}} \alpha(y|Z, W) \text{d}y$ for
Proof. Theorem 6. We can prove similar results as in the previous section.

Theorem 2. (Nash Equilibrium): A pair \((a^*, b^*)\) constitutes a Nash equilibrium of the privacy game if \((a^*, b^*) \in \mathcal{N}_{\text{af}}\) with

\[
\mathcal{N}_{\text{af}} = \{(a, b) \in A \times B \mid U(a, b) \leq U(a', b), \forall a' \in A, V(a, b) \leq V(a', b'), \forall b' \in B\}.
\]

Theorem 8. There exists an equilibrium in affine policies in which the sender employs the policy

\[
Y = K_1 Z + K_2 W + N,
\]

and the receiver employs the policy

\[
\hat{X} = (V_{XZ} K_1^T + V_{XW} K_2^T) Y,
\]

where

\[
\begin{bmatrix}
K_1^T \\
K_2^T
\end{bmatrix} \in \arg \min_{K_t \in \mathbb{R}^{(n_x+n_y) \times n_y}} K_t + \frac{1}{2} \log(\det(I - K_t^T P K_t)),
\]

s.t. \(K_t^T E K_t \leq I\),

\[
Q = \begin{bmatrix} V_{ZZ} & V_{ZV} & V_{ZW} \\ V_{VZ} & V_{WW} \end{bmatrix},
\]

\[
P = \begin{bmatrix} V_{ZV} & V_{WW} \\ V_{ZW} & V_{WW} \end{bmatrix},
\]

\[
E = \begin{bmatrix} V_{ZZ} & V_{ZW} \\ V_{VZ} & V_{WW} \end{bmatrix},
\]

and

\[
V_{NN} = I - \begin{bmatrix} K_1^T \\
K_2^T
\end{bmatrix}^T E \begin{bmatrix} K_1^T \\
K_2^T
\end{bmatrix}.
\]

Proof. See Appendix G.

The optimization problem in Theorem 8 is non-convex. However, we can solve this problem explicitly for scalar random variables.

Corollary 9. Let \(n_y = n_x = n_w = 1\) and \(V_{ZW} V_{WW} \neq V_{ZZ} V_{WW}\). There exists an equilibrium in affine policies in which the sender employs the policy

\[
Y = K_1 Z + K_2 W,
\]

and the receiver employs the policy

\[
\hat{X} = (V_{XZ} K_1^T + V_{XW} K_2^T) Y,
\]

where

\[
\begin{bmatrix}
K_1 \\
K_2
\end{bmatrix} = \frac{1}{\sqrt{V_{ZZ} + V_{WW} \xi^2 + 2 V_{ZW} \xi}} \begin{bmatrix} 1 \\
\xi
\end{bmatrix},
\]

with \(\xi \in \mathbb{R}_{>0}\) denoting an eigenvector of

\[
\begin{bmatrix} V_{ZZ} & V_{ZV} & V_{ZW} \\ V_{VZ} & V_{WW} \end{bmatrix},
\]

\[
\begin{bmatrix} V_{ZV} & V_{WW} \\ V_{ZW} & V_{WW} \end{bmatrix} - \frac{g/2}{V_{ZZ} - V_{ZW} V_{WW}} \begin{bmatrix} V_{ZW} V_{WW} & V_{ZW} \\ V_{ZW} & V_{WW} \end{bmatrix}.
\]

Proof. See Appendix H.

Consider a numerical example with \(n_y = n_x = n_w = 1\) such that \(V_{XX} = V_{WW} = V_{ZZ} = 1.0, V_{XW} = V_{ZW} = 0.4, \) and \(V_{ZZ} = 1.5\). Figure 3 shows the expected estimation error \(\mathbb{E}(d(X, \hat{X}))\) and mutual information \(I(Y; W)\) at the extracted equilibrium in affine policies versus the privacy ratio \(\rho\). The red line illustrates the bound in Theorem 7. As we can see, the upper bound is achieved with an affine policy. Therefore, at least for large \(g\), the equilibrium in affine policies is also an equilibrium of the game. Interestingly, there is no additive noise at this equilibrium.
4. CONCLUSION

We developed a game-theoretic framework to investigate the effect of privacy in the quality of the measurements provided by a well-informed sender to a receiver. We used a privacy ratio to model the sender’s emphasis on protecting her privacy. Equilibria of the game were constructed. Future work can focus on extending the results to multiple sensors.

REFERENCES


Appendix A. PROOF OF THEOREM 1

If the receiver does not use the transmitted message $Y$, the sender’s best policy is to minimize $I(Y; W)$, which is achieved by employing a uniform distribution on $Y$ (Thomas and Cover, 2006). Furthermore, if the sender’s message is independent of $(Z, W)$, the receiver’s best policy is to set her estimate to be equal to the element with maximum ex ante likelihood.

Appendix B. PROOF OF THEOREM 2

The first step is to show that $\Psi(\alpha, \beta) = \xi(\alpha, \beta) + \varphi(\beta)$ is a potential function for the game; see (Monderer and Shapley, 1996) for more information. This is evident from that
\[ V(\alpha, \beta) - V(\alpha, \beta') = \xi(\alpha, \beta) - \xi(\alpha, \beta') = \Psi(\alpha, \beta) - \Psi(\alpha, \beta'), \]

and

\[ U(\alpha, \beta) - U(\alpha', \beta) = \xi(\alpha, \beta) + \varphi(\alpha) - \xi(\alpha', \beta) - \varphi(\alpha') = \Psi(\alpha, \beta) - \Psi(\alpha', \beta), \]

for all \( \alpha, \alpha' \in A \) and \( \beta, \beta' \in B \). The rest follows from Lemma 2.1 in (Monderer and Shapley, 1996).

Appendix C. PROOF OF THEOREM 3

Note that \( \beta' \) means that \( X = Y \) with probability one, i.e., no data precessing is performed at the receiver. Clearly, if \( X = Y, \) the sender finds \( \alpha' \) so that \( Y \) minimizes \( \mathbb{E}[d(X,Y)] + g(W;Y) \). By definition, this is equivalent of saying that \( \alpha' \in \arg \min_{\alpha \in A} \{ \xi(\alpha, \beta) + \varphi(\alpha) \} \). In the rest of the proof, we show that the best response of the receiver is to use \( \beta' \). Do this by reductio ad absurdum.

Assume that there exists \( X \) constructed according to the conditional distribution \( \mathbb{P}[X = \hat{x}|Y = y] = \beta_{xy} \), for all \( \hat{x}, y \in X \), such that \( \mathbb{E}[d(X,\bar{X})] \geq \mathbb{E}[d(X,Y)] \) (because otherwise the receiver sticks to \( \beta' \)). Following the data processing inequality from Theorem 2.8.1 (Thomas and Cover, 2006, p. 34), we have \( I(W;X) \leq I(W;Y) \). This shows that \( \mathbb{E}[d(X,\bar{X})] + g(W;\bar{X}) < \mathbb{E}[d(X,Y)] + g(W;Y) \). This is evidently in contradiction with the optimality of \( \alpha' \).

Appendix D. PROOF OF THEOREM 4

The proof is done by reductio ad absurdum. To do so, assume that there exists an increasing subsequence \( \{k_i\} \in \mathbb{N} \) such that \( (\alpha^{k_i}, \beta^{k_i}) \notin \mathcal{N}, \forall i \in \mathbb{N} \). If \( (\alpha^{k_i}, \beta^{k_i}) \notin \mathcal{N}, \) for some \( k_i \geq 3, \) at least one of the following cases hold.

- Case 1 (\( \exists \alpha' \in A : U(\alpha, \beta, k) > U(\alpha', \beta, k) + \epsilon \) and \( k \) is even): This means that \( \beta^k = \beta^k - 1 \). Thus, we know that there exists \( \alpha' \in A \) such that \( U(\alpha', \beta^{k-1}) > U(\alpha, \beta^{k}) - \epsilon \). This is in contradiction with Line 3 of Algorithm 1 and, thus, will never occur.

- Case 2 (\( \exists \alpha' \in A : U(\alpha, \beta, k) > U(\alpha', \beta, k) + \epsilon \) and \( k \) is odd): In this case, we have

\[ \Psi(\alpha^{k+1}, \beta^{k+1}) - \Psi(\alpha^k, \beta^k) = U(\alpha^{k+1}, \beta^{k+1}) - U(\alpha^k, \beta^k) \leq \alpha'. \beta^k - U(\alpha', \beta^k) \]

Line 3 in Algorithm 1 and

\[ \geq \alpha'. \beta^k - U(\alpha', \beta^k) \]

- Case 3 (\( \exists \beta' \in B : V(\alpha, \beta, k) > V(\alpha, \beta', k) + \epsilon \) and \( k \) is even): In this case, we have

\[ \Psi(\alpha^{k+1}, \beta^{k+1}) - \Psi(\alpha^k, \beta^k) = U(\alpha^{k+1}, \beta^{k+1}) - U(\alpha^k, \beta^k) \leq \alpha'. \beta^k - U(\alpha', \beta^k) \]

Line 6 in Algorithm 1 and

\[ \geq \alpha'. \beta^k - U(\alpha', \beta^k) \]

- Case 4 (\( \exists \beta' \in B : V(\alpha, \beta, k) > V(\alpha, \beta', k) + \epsilon \) and \( k \) is odd): This means that \( \alpha^k = \alpha^{k-1} \). Further, we know that there exists \( \beta' \in A \) such that \( V(\alpha^{k-1}, \beta') < V(\alpha^{k-1}, \beta^k) - \epsilon \). This is in contradiction with Line 6 of Algorithm 1 and, thus, will never occur.

From combining Cases 1–4, we know that if \( (\alpha^k, \beta^k) \notin \mathcal{N}, \) for some \( k \geq 3, \) then \( \Psi(\alpha^{k+1}, \beta^{k+1}) - \Psi(\alpha^k, \beta^k) < \epsilon \). Note that, in general, by construction of Line 3 in Algorithm 1, if \( k \) is an even number, we get

\[ \Psi(\alpha^k, \beta^k) - \Psi(\alpha^{k-1}, \beta^{k-1}) = U(\alpha^k, \beta^{k-1}) - U(\alpha^{k-1}, \beta^{k-1}) \leq 0. \]

Similarly, by construction of Line 6 in Algorithm 1, if \( k \) is an odd number, we have

\[ \Psi(\alpha^k, \beta^k) - \Psi(\alpha^{k-1}, \beta^{k-1}) = V(\alpha^{k-1}, \beta^k) - V(\alpha^{k-1}, \beta^{k-1}) \leq 0. \]

Therefore, we can deduce that

\[ \lim_{k \to \infty} \Psi(\alpha^k, \beta^k) = \Psi(\alpha^0, \beta^0) + \sum_{r \in \mathbb{N}} (\Psi(\alpha^{r+1}, \beta^{r+1}) - \Psi(\alpha^{r}, \beta^{r})) \]

which is in contradiction with \( \lim_{k \to \infty} \Psi(\alpha^k, \beta^k) \) exists, because, by (D.1) and (D.2), \( (\Psi(\alpha^k, \beta^k)) \in \mathbb{N} \) is a monotone decreasing sequence and is lower bounded by zero.

Appendix E. PROOF OF THEOREM 5

Algorithm 2 makes sure that \( (\alpha^k, \beta^k) = (\alpha^{k-1}, \beta^{k-1}) \) if \( (\alpha^k, \beta^k) \in \mathcal{N}, \) Therefore, there exists \( K \) such that \( (\alpha^k, \beta^k) \notin \mathcal{N}, \) for \( k \leq K - 1 \) and \( (\alpha^k, \beta^k) \in \mathcal{N}, \) for \( k \geq K \). Now, following the reasoning of the proof of Theorem 4, we can see that

\[ \Psi(\alpha^K, \beta^K) = \Psi(\alpha^0, \beta^0) + \sum_{r=2}^{K-1} (\Psi(\alpha^{r+1}, \beta^{r+1}) - \Psi(\alpha^{r}, \beta^{r})) \leq \Psi(\alpha^0, \beta^0) - (K-2)\epsilon. \]

The last inequality follows from that \( \Psi(\alpha^k, \beta^k) \) is a decreasing sequence. Noting that \( \Psi(\alpha^K, \beta^K) \geq 0 \) because of the properties of the mutual information and expected estimation error, we can see that \( K \leq 2 + \Psi(\alpha^0, \beta^0)/\epsilon. \)

Appendix F. PROOF OF THEOREM 7

First, define \( \bar{Z} = Z - V_{XX}^{-1}V_{XZ} W \). Note that \( X, \bar{Z}, W \) are still jointly distributed Gaussian variables. Therefore, \( \mathbb{E}[\bar{W}] = 0 \) implies that \( Z \) and \( W \) are independent random variables. Therefore, \( I(\bar{Z};W) = 0. \) Now, let \( Y = \bar{Z} \). We can show that

\[ \min_{\beta \in B} \mathbb{E}[\|X - \bar{X}\|_2^2] = \mathbb{E}[\|X - \mathbb{E}[X|\bar{Z}]\|_2^2] = V_{XX} - V_{XZ}V_{ZZ}^{-1}V_{ZX}, \]

where

\[ V_{ZZ} = V_{ZZ} - V_{ZW}V_{WZ}^{-1}V_{WZ}, \]

\[ V_{XZ} = V_{XZ} - V_{ZW}V_{WZ}^{-1}V_{WZ}. \]

Let \( \alpha^* \) be such that \( \mathbb{P}(Y = \bar{Z}) = 1. \) Note that
Furthermore, note that $\min_{(\alpha, \beta) \in A \times B} [\xi(\alpha, \beta) + \phi(\alpha)] \leq \min_{\beta \in B} [\xi(\alpha^*, \beta) + \phi(\alpha^*)]$

$$= E[\|X - E(X|Z)\|^2] + I(Z; W)$$

$$= V_{XX} - V_{XZ}V_{ZZ}^{-1}V_{ZX}.$$

Appendix G. PROOF OF THEOREM 8

First, note that

$$E \left\{ \begin{bmatrix} Y \\ W \end{bmatrix} \right\} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

and

$$E \left\{ \begin{bmatrix} Y \\ W \end{bmatrix} \begin{bmatrix} Y \\ W \end{bmatrix}^\top \right\} = \begin{bmatrix} V_{YY} & V_{YW} \\ V_{YW}^\top & V_{WW} \end{bmatrix},$$

where

$$V_{YY} = K_1^1V_{ZZ}K_1^1 + K_1V_{ZZ}K_2^\top + K_2V_{WW}K_1^\top + V_{NN},$$

$$V_{YW} = K_1V_{ZW} + K_2V_{WW}.$$ Without loss of generality, we can assume that $V_{YY} = I$ because otherwise the receiver can pre-scale the received message and work with $\hat{Y} = V_{YY}^{-1/2}Y$ that has the same amount of information as $Y$ (there is one-to-one mapping between $Y$ and $\hat{Y}$). Following equation (3) in (Hershey and Olsen, 2007) and noting the relationship between the mutual information and the Kullback-Leibler divergence (Thomas and Cover, 2006), we can see that

$$I(Y; W) = \frac{1}{2} \log \left( \frac{\det(V_{YY}) \det(V_{WW})}{\det\left( V_{YY}V_{YW}^\top \right)} \right)$$

$$= \frac{1}{2} \log \left( \frac{\det(V_{YY} - V_{YW}V_{WW}V_{YW}^\top)}{\det(V_{YY})} \right)$$

$$= \frac{1}{2} \log \left( \frac{n_y}{\det(V_{YY} - V_{YW}V_{WW}V_{YW}^\top)} \right)$$

$$= \frac{1}{2} \log(n_y) - \frac{1}{2} \log(\det(V_{YY} - V_{YW}V_{WW}V_{YW}^\top)),$$

where

$$V_{YY} - V_{YW}V_{WW}V_{YW}^\top = I - (K_1V_{ZW} + K_2V_{WW})V_{WW}^{-1}(V_{WZ}K_1^\top + K_2V_{WW}K_2^\top)$$

$$= I - K_1V_{ZW}V_{WW}^{-1}V_{WZ}K_1^\top - K_1V_{ZW}K_2^\top - K_2V_{WW}K_2^\top$$

$$= I - \begin{bmatrix} K_1^\top & K_2^\top \end{bmatrix} \begin{bmatrix} V_{ZW} & V_{ZW} \\ V_{WZ} & V_{WW} \end{bmatrix} \begin{bmatrix} K_1 \\ K_2 \end{bmatrix}.$$ Furthermore, note that

$$E[\|X - \hat{X}\|^2] = \text{trace}(V_{XX} - V_{XY}V_{YY}^{-1}V_{YX})$$

$$= \text{trace}(V_{XX}) - \text{trace}\left( \begin{bmatrix} K_1^\top & K_2^\top \end{bmatrix} Q \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} \right),$$

where

$$Q = \begin{bmatrix} V_{ZZ} & V_{ZZ}V_{ZW} \\ V_{ZW}V_{XZ} & V_{WW}V_{XW} \end{bmatrix}.$$

The rest of the proof follows from simple algebraic manipulations.

Appendix H. PROOF OF COROLLARY 9

First, we prove that the equilibrium cannot be a point such that

$$K^\top EK < 1,$$

where

$$K = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix}.$$ This part of the proof is done by reductio ad absurdum. If the optimal solution is inside the constraint set (i.e., $K^\top EK < 1$), the following identity should hold

$$-QK + \frac{\theta}{2(1 - K^\top PK)}PK = 0.$$ Therefore,

$$\left( \frac{\theta}{2(1 - K^\top PK)} - Q \right)K = 0.$$ For this to happen, there should exists $\lambda > 0$ such that

$$(\lambda P - Q)K = 0.$$ Now, note that

$$P = \frac{1}{V_{WW}} \begin{bmatrix} V_{ZW} \\ V_{WW} \end{bmatrix} \begin{bmatrix} V_{ZW} & V_{WW} \end{bmatrix}$$

$$Q = \begin{bmatrix} V_{XX} \\ V_{WW} \end{bmatrix} \begin{bmatrix} V_{XX} & V_{WW} \end{bmatrix}.$$ This only possible if

$$\begin{bmatrix} V_{ZW} \\ V_{WW} \end{bmatrix} \parallel \begin{bmatrix} V_{XX} \\ V_{WW} \end{bmatrix}.$$ This contradicts the assumption of the theorem that $V_{ZW}V_{WW} \neq V_{XX}.$ Now, we can form the Lagrangian

$$\mathcal{L} = -K^\top QK - \frac{\theta}{2} \log(1 - K^\top PK) + \mu(K^\top EK - 1).$$ Setting the derivative of $\mathcal{L}$ with respect to $K$ equal to zero results in

$$-QK + \frac{\theta}{2(1 - K^\top PK)}PK + \mu EK$$

$$= \begin{bmatrix} 1 & \mu K_2^2 E - Q + \frac{\theta}{2} \frac{V_{ZZ} - V_{ZW}V_{WW}^{-1}V_{WZ}}{V_{WW}} \end{bmatrix} \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} = 0,$$

where the first equality follows from that

$$P = E - \begin{bmatrix} V_{ZZ} - V_{ZW}V_{WW}^{-1}V_{WZ} \\ 0 \end{bmatrix}.$$ Therefore, $\mu K_2^2$ needs to be an eigenvalue of

$$\begin{bmatrix} V_{ZZ} & V_{ZW} \\ V_{ZZ} & V_{WW} \end{bmatrix}^{-1} \begin{bmatrix} V_{XX} & V_{XX} \\ V_{XX} & V_{WW} \end{bmatrix} - \frac{\theta}{2} \begin{bmatrix} V_{ZZ} - V_{ZW}V_{WW}^{-1}V_{WZ} \\ V_{WW} \end{bmatrix}.$$ Let $[\xi]$ denote the corresponding eigenvector. Evidently, $K_2/K_1 = \xi.$ Satisfying the constraint $K^\top EK = 1$ gives

$$\begin{bmatrix} K_1 \\ K_2 \end{bmatrix} = \pm \frac{1}{\sqrt{V_{ZZ} + V_{WW}\xi^2 + 2V_{ZW}\xi}} [\xi].$$ This concludes the proof.