Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Information Value on Private State Inference in Network Systems

Hao Jiang, Xuda Ding, Jianping He, Yunfeng Peng

The Dept. of Automation, Shanghai Jiao Tong University, and the Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai, China 200240
(e-mail: mouse826612011@sjtu.edu.cn, dingxuda@sjtu.edu.cn, jphe@sjtu.edu.cn, pengyf@sjtu.edu.cn).

Abstract: In network systems, neighboring nodes usually need to exchange and update their state information iteratively to achieve a global computation and control goal. Considering the nodes' states may include some sensitive/private information, e.g., location and income, different random mechanizes have been proposed to preserve the privacy of the states. However, no matter what type of random mechanisms is used, the eavesdropping attacker can infer/estimate a node's state based on the information it holds, and the estimation depends on the available information. The relationship between the estimation and the information is a critical and open issue. Therefore, in this paper, we investigate how to obtain the optimal estimation of a node's state with available information and how to quantify the value of the information in the state inference. First, we exploit a utility function to quantify the utility of the estimation accuracy, and then the optimal estimation and information value are defined to depict the estimation and quantify the information, respectively. Next, the optimal estimation under different settings of the noise and utility function is provided. Lastly, we obtain some essential properties of information value and analyze the value of state outputs in distributed algorithms.

Keywords: Distributed algorithm, Noise adding process, Optimal estimation, Data privacy, Average consensus.

1. INTRODUCTION

In network systems, nodes cooperatively achieve some global computation or control goals, e.g., data aggregation and formation control, by local information exchanging (Olfati-Saber et al. (2007); Rajagopalan and Varshney (2006)). In this process, each node iteratively communicates with its neighbor nodes to obtain their states and then updates its state using the designed distribution algorithms. A distributed algorithm provides the state update rule for each node in the network system to reach a global goal. Most popular distributed algorithms include distributed estimation, statistics, control and optimization algorithms, etc (Blondel et al. (2005); Olfati-Saber et al. (2007); Ren et al. (2007)). Due to the distributed nature, these algorithms have strong robustness and scalability, and have been widely used in network systems (Liu et al. (2017); Pasqualetti et al. (2010); Zhao et al. (2016)).

In many networks, the nodes' states may include some sensitive or private information. To preserve the state privacy, nodes may not be willing to share real state with their neighbors during and will use processed data for communication. A widely used approach is adding random noises to the real state for data exchanging during communication. However, using processed data for data exchanging will affect the performance of the distributed algorithm directly. Thus, how to carefully design the rule for the data processing considering the tradeoff between the performance of the distributed algorithm and the privacy preservation, has attracted attention recently. For instance, the authors in (Huang et al. (2012, 2019); Imtiaz and Sarwate (2018); Le Ny and Pappas (2013); Manitara and Hadjicostis (2013); Mo and Murray (2016); Nozari et al. (2017)) aimed to design the noise adding a mechanism for the average consensus, such that the privacy of nodes' initial states is preserved while the "average consensus" can still be achieved. It is proved that the average consensus is achieved in probability sense if the noises are variance decaying and zero-sum.

However, considering the privacy disclosure in network systems, the eavesdropping attacker can still estimate the nodes' states based on different kinds of information it holds, e.g., the prior knowledge of the node's initial state, etc. In (He et al. (2018)), it has shown that no matter what type of the noise distribution it is, there is a positive probability that an estimated value of the original data is close to the real data, where the estimation accuracy is less than a given small constant. The authors have considered how to estimate the node's state with only local information outputs under the distributed algorithm optimally so that the probability is minimized. Obviously, the state inference highly depends on the information available. However, how to define the optimal estimation under different available information sets and quantify the value of the information for private state inference are critical issues, especially

considering the information dynamics caused by the action of the distributed algorithm.

Inspired by the pioneers' work on the information value (Fogel and Huang (1982); Howard (1966)), in this paper, we define the information value on private state inference considering distributed algorithms in network systems. Information value describes the expected utility of the optimal estimation under the given information, where the utility is modeled as a function of the estimation accuracy, and the optimal estimation is defined as the estimate achieving the maximum utility. Thus, the information value links the optimal estimation and the utility of the estimation. We then study the optimal estimation under certain conditions, the basic properties of the information value, and the value of information outputs in the distributed algorithm. The main contributions of this paper are summarized as follows.

- To the best of our knowledge, this is the first work to investigate the information value considering the private state inference in network systems. We introduce a novel information value definition to quantify the contribution of the information on the state inference and the corresponding utility. Based on the definition, it is not difficult to calculate the privacy lost when a specified information set is released.
- We first obtain the optimal estimation and its closed-form of expression under different conditions. Then, we prove several important properties of the information value function. Lastly, the value of information outputs in distributed algorithms is investigated. The obtained results provide the basis for further privacy analysis and design of distributed algorithms.

The remainder of this paper is organized as follows. Section 2 provides the preliminaries and formulates the problem. The main results on optimal estimation and information value are introduced in Section 3. Lastly, the conclusions are given in Section 4.

2. PRELIMINARIES AND PROBLEM FORMULATION

A network system is described as a weighted, undirected and connected graph, denoted by $G = (\mathcal{V}, \mathcal{E})$, with node set $\mathcal{V}$ of cardinality n ($n \geq 2$) and edge set $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$. An edge $(i, j) \in \mathcal{E}$ exists if and only if (iff) node i can communicate with node j. $\mathcal{N}_i = \{j | (i, j) \in E\}$ is defined as the neighbor set of node i, and assume $i \notin \mathcal{N}_i$. Let $x(0) \in \mathcal{R}^n$ be the initial state vector of nodes, where $x_i(0) \in x(0)$ is an initial scalar and private state of node i, e.g., the sensing data or the location of the node. For simplicity, we let $x_i = x_i(0)$ and $\hat{x}_i = \hat{x}_i(0)$ for $i \in \mathcal{V}$, where $\hat{x}_i$ is the estimate of x_i.

2.1 Privacy-Preserving Distributed Algorithm

In network systems, the neighboring nodes communicate with each other periodically for data exchanging. Based on the obtained information, they update states following the rule of a designed distributed algorithm, to obtain the statistics (e.g., mean/avarage, maxi/minimum value, variance, etc.) of all nodes' initial states in a distributed way. Due to the sensitive information of the initial states, nodes are may not be willing to release theirs real states to neighboring nodes directly. Thus, to preserve the privacy, a widely used approach is utilizing a randomized mechanism to process real states, and then using the randomized/processed states for data exchanging during the communication, e.g., adding random noises to the real states for communication (Manitara and Hadjicostis (2013); Nozari et al. (2017)).

Define $x_i^+(k)$ to be the communication data of node i in each iteration k. $x_i^+(k)$ is generated by applying the randomized mechanism $\mathcal{M}_i$ on $x_i(k)$, which is given by

$$\mathcal{M}_i(x_i(k)) = \mathsf{g}_i(x_i(k)) + \theta_i(k), \tag{1}$$

i.e., $x_i^+(k) = \mathcal{M}_i(x_i(k))$, where $\mathcal{M}_i$ is the randomized mechanism, $\mathsf{g}_i : \mathcal{R} \to \mathcal{R}$ is an invertible function, and $\theta_i(k) \in \mathcal{R}$ is a random variable. If node i receives the information from neighboring nodes, it updates the state with the following equation,

$$x_i(k+1) = \mathsf{h}_i(x_i^+(k), x_{j_1}^+(k), x_{j_2}^+(k), ..., x_{j_{|\mathcal{N}_i|}}^+(k)), \tag{2}$$

where the state-transition function, $\mathsf{h}_i : \mathcal{R} \times \mathcal{R} \times ... \times \mathcal{R} \to \mathcal{R}$, depends on $x_i^+(k)$ and $x_j^+(k)$ for $j \in N_i$ only. The dynamic (2) is a typical privacy-preserving distributed algorithm (PPDA), since in each iteration, the state updating is only using the local information and the communication data are processed by a randomized mechanism (1). For example, many existing privacy-preserving consensus algorithms, e.g., (He et al. (2018); Manitara and Hadjicostis (2013); Mo and Murray (2016); Nozari et al. (2017)), can be modeled by (1) and (2).

2.2 Important Definitions

Suppose that there is an attack node s (eavesdropper) in the network. Node s will infer/estimate the initial state of node i with information set it holds, denoted by $\mathcal{I}_s$, and node s may or may not be a node in V.

It is noted that for node s, the higher estimation accuracy means the better privacy attack. In the view of normal nodes, a more accurate estimation may cause higher privacy lost. Hence, we introduce a utility function to evaluate the quality of the state inference, which is modeled as a function of the estimation accuracy, denoted by $\mathsf{u}(\hat{x}_i - x_i)$. Suppose that $\mathsf{u}(\cdot) \geq 0$, and it is a continuous and integrable function, and is an increasing function when $\hat{x}_i - x_i < 0$ and decreasing function when $\hat{x}_i - x_i > 0$. Clearly, $\mathsf{u}(0)$ is of the maximum value. Then, $-\mathsf{u}(\hat{x}_i - x_i)$ is viewed as the privacy lost function of node i given estimation $\hat{x}_i$.

During the estimation, x_i is not available to node s, and thus $\hat{x}_i - x_i$ is unknown and $\mathsf{u}(\cdot)$ cannot be calculated directly. Hence, the expectation of $\mathsf{u}(\cdot)$ is used to calculate the utility of an estimation given by node s. Given $\mathcal{I}_s$ and $\hat{x}_i$, the expectation of $\mathsf{u}(\cdot)$ is calculated by

$$\mathbf{E}\left[\mathsf{u}(\hat{x}_i - x_i)|\mathcal{I}_s\right] = \int \mathsf{f}_{x_i|\mathcal{I}_s}(z)\mathsf{u}(\hat{x}_i - z)\mathrm{d}z, \tag{3}$$

where $\mathsf{f}_{x_i|\mathcal{I}_s}(z)$ is the probability density function (PDF) of possible values of x_i given condition $\mathcal{I}_s$. When an information set $\mathcal{I}_i$ of node i is released to node s, then node s holds $\mathcal{I}_s \cup \mathcal{I}_i$ and the PDF of the possible values of x_i is changed to $\mathsf{f}_{x_i|\mathcal{I}_s \cup \mathcal{I}_i}(z)$, which results in the change of the expectation of $\mathsf{u}(\cdot)$.

Definition 2.1. [Optimal Estimation] We say $\hat{x}^*_{i|\mathcal{I}_s}$ is the optimal estimation of x_i under $\mathcal{I}_s$, if

$$\hat{x}^*_{i|\mathcal{I}_s} = \arg\max_{\hat{x}_i \in \mathcal{R}} \mathbf{E}\left[\mathsf{u}(\hat{x}_i - x_i)|\mathcal{I}_s\right]. \tag{4}$$

To quantify the value of an information set in the sense of state inference, we give the following definitions.

Definition 2.2. [Information Value] Let $\mathcal{I}_i$ be an information set of node i, and its value is defined by

$$\mathsf{V}(\mathcal{I}_i) = \mathbf{E}\left[\mathsf{u}(\hat{x}^*_{i|\mathcal{I}_i} - x_i)|\mathcal{I}_i\right], \tag{5}$$

which is the expectation of $\mathsf{u}(\cdot)$ under $\hat{x}^*_{i|\mathcal{I}_i}$ and $\mathcal{I}_i$.

Definition 2.3. [Relative Information Value] Suppose that $\mathcal{I}_i$ is released to node s who already holds $\mathcal{I}_s$. The relative information value is defined by $\mathsf{V}_r(\mathcal{I}_i|\mathcal{I}_s)$, satisfying

$$\mathsf{V}_r(\mathcal{I}_i|\mathcal{I}_s) = \mathsf{V}(\mathcal{I}^i_s) - \mathsf{V}(\mathcal{I}_s)$$

where $\mathcal{I}^i_s = \mathcal{I}_s \cup \mathcal{I}_i$.

In the above definition, the information value is evaluated by node s. When the information $\mathcal{I}_i$ is released to node s, for normal node i, the relative privacy lost is given by

$$\mathsf{L}(\mathcal{I}_i) = -\mathsf{u}(\hat{x}^*_{i|\mathcal{I}^i_s} - x_i) + \mathsf{u}(\hat{x}^*_{i|\mathcal{I}_s} - x_i).$$

Note that the information theory of entropy developed by Shannon provides a quantitative measure on the amount of information involved in any communication (Howard (1966)). However, the entropy cannot indicate how the information contribute the estimation of states, which can be depicted by information value defined in this paper.

2.3 Problem Formulation

In network systems, there are mainly four kinds of information, i.e., the information of the distributed algorithm (e.g., the updating rule, the noise distribution, etc.), the priori knowledge (e.g., the distribution of the possible values of nodes' states), the topology information, the observable state information of nodes. In this paper, we focus on the value of the later two kinds.

We define the topology information and the observable state information set, respectively, as

$$\mathcal{I}_g = \cup_{i\in\mathcal{V}}\mathcal{I}^i_g = \cup_{i\in\mathcal{V}}\{\mathcal{N}_i \cup i\} \tag{6}$$

and

$$\mathcal{I}_o(k) = \cup_{i\in\mathcal{V}}\mathcal{I}^i_o(k) = \cup_{i\in\mathcal{V}}\{x^+_i(0), ..., x^+_i(k)\} \tag{7}$$

until iteration k. Then, we have $\{x^+_i(k)\} = \mathcal{I}^i_o(k) - \mathcal{I}^i_o(k-1)$ and $\{x^+(k)\} = \mathcal{I}_o(k) - \mathcal{I}_o(k-1)$. Let $\mathcal{I}_o(\infty) = \lim_{k\to\infty}\mathcal{I}_o(k)$. The objective of this paper is to provide a theoretical framework to quantify the information value in the sense of state interference, considering PPDA (2) in network systems. Four issues will be investigated: i) how to choose the $\mathsf{f}_{x_i|\mathcal{I}_s}(z)$; ii) how to obtain the optimal estimation given the utility function and the available information set; iii) what are the basic properties of the information value function given in Definition 2.3; iv) the values of the topology information $\mathcal{I}_g$ and the observable information outputs $\mathcal{I}_o(k)$.

3. MAIN RESULTS

From Definition 2.3, it is not difficult to see that information value depends on the optimal estimation, the utility function, and the effect of the information on the estimation and utility function. In the following subsections, we discuss important properties of the optimal estimation and the information value function, and then investigate the value of the information outputs of the algorithm.

3.1 Possible State Values

During the state inference or inference, we assume that

(1) if node s has no information of node i, the possible value set of x_i is viewed $[-M, M]$, where M is a large positive constant. Then, for the state inference,

$$\Pr\{\hat{x}_i = x_i\} = \begin{cases} \dfrac{1}{2M}, & \text{if } x_i \in [-M, M], \\ 0, & \text{otherwise;} \end{cases}$$

(2) if node s knows $x_i \in [a, b]$, then for the state inference, we have $\Pr\{\hat{x}_i = x_i\} = \frac{1}{b-a}$, where $\hat{x}_i \in [a, b]$.

3.2 Optimal Estimation

Note that if two estimates satisfies $\hat{x}^1_i - x_i = x_i - \hat{x}^2_i$, then they have the same estimation accuracy as the accuracy is decided by $|\hat{x}_i - x_i|$. Hence, we can assume the utility function $\mathsf{u}(\cdot)$ is symmetric, i.e., $\mathsf{u}(z) = \mathsf{u}(-z)$. Meanwhile, the symmetric also holds for many widely used noises' PDFs, e.g., Gaussian and Laplace distribution, etc. Then, we provide a theorem as follows.

Theorem 3.1. Suppose that both $\mathsf{f}_{x_i|\mathcal{I}_s}$ and $\mathsf{u}(\cdot)$ are symmetric and unimodal functions, and we have

$$\hat{x}^*_{i|\mathcal{I}_s} = \arg\max_{z\in\mathcal{R}} \mathsf{f}_{x_i|\mathcal{I}_s}(z).$$

Proof. Since $\mathsf{u}(z)$ is a symmetric and unimodal function, and $\mathsf{u}(0)$ has the maximum value, we have $\mathsf{u}(z) = \mathsf{u}(-z)$. Giving any two estimates $\hat{x}_i(1)$ and $\hat{x}_i(2)$ and assuming $\hat{x}_i(1) < \hat{x}^*_{i|\mathcal{I}_s} < \hat{x}_i(2))$, it follows from (3) and (4) that

$$\begin{aligned}
&\mathbf{E}\left[\mathsf{u}(\hat{x}_i(2) - x_i)|\mathcal{I}_s\right] - \mathbf{E}\left[\mathsf{u}(\hat{x}^*_{i|\mathcal{I}_s} - x_i)|\mathcal{I}_s\right] \\
&= \int \mathsf{f}_{x_i|\mathcal{I}_s}(z)\left[\mathsf{u}(z - \hat{x}_i(2)) - \mathsf{u}(z - \hat{x}^*_{i|\mathcal{I}_s})\right]\mathrm{d}z \\
&= \left(\int_{-\infty}^{\hat{x}^*_{i|\mathcal{I}_s}} + \int_{\hat{x}^*_{i|\mathcal{I}_s}}^{\frac{\hat{x}^*_{i|\mathcal{I}_s}+\hat{x}_i(2)}{2}} + \int_{\frac{\hat{x}^*_{i|\mathcal{I}_s}+\hat{x}_i(2)}{2}}^{\hat{x}_i(2)} + \int_{\hat{x}_i(2)}^{+\infty}\right) \\
&\quad \mathsf{f}_{x_i|\mathcal{I}_s}(z)\left[\mathsf{u}(z - \hat{x}_i(2)) - \mathsf{u}(z - \hat{x}^*_{i|\mathcal{I}_s})\right]\mathrm{d}z.
\end{aligned} \tag{8}$$

Note that

$$\begin{aligned}
&\int_{-\infty}^{\hat{x}^*_{i|\mathcal{I}_s}} \mathsf{f}_{x_i|\mathcal{I}_s}(z)\left[\mathsf{u}(z - \hat{x}_i(2)) - \mathsf{u}(z - \hat{x}^*_{i|\mathcal{I}_s})\right]\mathrm{d}z \\
&= \int_{-\hat{x}^*_{i|\mathcal{I}_s}}^{+\infty} \mathsf{f}_{x_i|\mathcal{I}_s}(-z)\left[\mathsf{u}(-z - \hat{x}_i(2)) - \mathsf{u}(-z - \hat{x}^*_{i|\mathcal{I}_s})\right]\mathrm{d}z \\
&= \int_{\hat{x}_i(2)}^{+\infty} \mathsf{f}_{x_i|\mathcal{I}_s}(\hat{x}_i(2) + \hat{x}^*_{i|\mathcal{I}_s} - z) \times \left[\mathsf{u}(z - \hat{x}^*_{i|\mathcal{I}_s}) - \mathsf{u}(z - \hat{x}_i(2))\right]\mathrm{d}z.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
&\left(\int_{-\infty}^{\hat{x}^*_{i|\mathcal{I}_s}} + \int_{\hat{x}_i(2)}^{+\infty}\right) \mathsf{f}_{x_i|\mathcal{I}_s}(z)\left[\mathsf{u}(z - \hat{x}_i(2)) - \mathsf{u}(z - \hat{x}^*_{i|\mathcal{I}_s})\right]\mathrm{d}z \\
&= \int_{\hat{x}_i(2)}^{+\infty} \left(\mathsf{f}_{x_i|\mathcal{I}_s}(\hat{x}_i(2) + \hat{x}^*_{i|\mathcal{I}_s} - z) - \mathsf{f}_{x_i|\mathcal{I}_s}(z)\right) \\
&\quad \times \left[\mathsf{u}(z - \hat{x}^*_{i|\mathcal{I}_s}) - \mathsf{u}(z - \hat{x}_i(2))\right]\mathrm{d}z < 0.
\end{aligned} \tag{9}$$

Due to the following two reasons. On one hand, since $\mathsf{f}_{x_i|\mathcal{I}_s}(z)$ is a symmetric and unimodal function with its maximum value at $z = \hat{x}^*_{i|\mathcal{I}_s}$, $\mathsf{f}_{x_i|\mathcal{I}_s}(z)$ is a decreasing function for $z \geq \hat{x}^*_{i|\mathcal{I}_s}$ and the following equation holds,

$$\mathsf{f}_{x_i|\mathcal{I}_s}(2\hat{x}^*_{i|\mathcal{I}_s} - z) = \mathsf{f}_{x_i|\mathcal{I}_s}(z).$$

It follows that

$$\begin{aligned}&\mathsf{f}_{x_i|\mathcal{I}_s}(\hat{x}_i(2) + \hat{x}^*_{i|\mathcal{I}_s} - z) - \mathsf{f}_{x_i|\mathcal{I}_s}(z)\\ =&\mathsf{f}_{x_i|\mathcal{I}_s}(\hat{x}^*_{i|\mathcal{I}_s} + z - \hat{x}_i(2)) - \mathsf{f}_{x_i|\mathcal{I}_s}(\hat{x}^*_{i|\mathcal{I}_s} + z - \hat{x}^*_{i|\mathcal{I}_s}) > 0\end{aligned}$$

due to $0 \leq z - \hat{x}_i(2) < z - \hat{x}^*_{i|\mathcal{I}_s}$ when $z \geq \hat{x}_i(2)$. On the other hand, noting that $\mathsf{u}(z)$ is a decreasing function for $z \geq 0$, we have

$$\begin{aligned}&\mathsf{u}(z - \hat{x}^*_{i|\mathcal{I}_s}) - \mathsf{u}(z - \hat{x}_i(2))\\ =&\mathsf{u}(z - \hat{x}_i(2) + \hat{x}_i(2) - \hat{x}^*_{i|\mathcal{I}_s}) - \mathsf{u}(z - \hat{x}_i(2)) < 0\end{aligned}$$

for $z \geq \hat{x}_i(2)$. The above two equations yield (9).

Similarity, one notes that

$$\begin{aligned}&\int_{\hat{x}^*_{i|\mathcal{I}_s}}^{\frac{\hat{x}^*_{i|\mathcal{I}_s}+\hat{x}_i(2)}{2}} \mathsf{f}_{x_i|\mathcal{I}_s}(z)\left[\mathsf{u}(z - \hat{x}_i(2)) - \mathsf{u}(z - \hat{x}^*_{i|\mathcal{I}_s})\right]\mathrm{d}z\\ =&\int_{-\frac{\hat{x}^*_{i|\mathcal{I}_s}+\hat{x}_i(2)}{2}}^{-\hat{x}^*_{i|\mathcal{I}_s}} \mathsf{f}_{x_i|\mathcal{I}_s}(-z)\\ &\times\left[\mathsf{u}(-z - \hat{x}_i(2)) - \mathsf{u}(-z - \hat{x}^*_{i|\mathcal{I}_s})\right]\mathrm{d}z\\ =&\int_{\frac{\hat{x}^*_{i|\mathcal{I}_s}+\hat{x}_i(2)}{2}}^{\hat{x}_i(2)} \mathsf{f}_{x_i|\mathcal{I}_s}(\hat{x}^*_{i|\mathcal{I}_s} + \hat{x}_i(2) - z)\\ &\times\left[\mathsf{u}(\hat{x}^*_{i|\mathcal{I}_s} - z) - \mathsf{u}(\hat{x}_i(2) - z)\right]\mathrm{d}z\end{aligned} \tag{10}$$

Then, based on $\mathsf{u}(-z) = \mathsf{u}(z)$, one infers that

$$\begin{aligned}&\left(\int_{\hat{x}^*_{i|\mathcal{I}_s}}^{\frac{\hat{x}^*_{i|\mathcal{I}_s}+\hat{x}_i(2)}{2}} + \int_{\frac{\hat{x}^*_{i|\mathcal{I}_s}+\hat{x}_i(2)}{2}}^{\hat{x}_i(2)}\right)\mathsf{f}_{x_i|\mathcal{I}_s}(z)\\ &\times\left[\mathsf{u}(z - \hat{x}_i(2)) - \mathsf{u}(z - \hat{x}^*_{i|\mathcal{I}_s})\right]\mathrm{d}z\\ =&\int_{\frac{\hat{x}^*_{i|\mathcal{I}_s}+\hat{x}_i(2)}{2}}^{\hat{x}_i(2)} \left(\mathsf{f}_{x_i|\mathcal{I}_s}(z) - \mathsf{f}_{x_i|\mathcal{I}_s}(\hat{x}^*_{i|\mathcal{I}_s} + \hat{x}_i(2) - z)\right)\\ &\times\left[\mathsf{u}(z - \hat{x}_i(2)) - \mathsf{u}(z - \hat{x}^*_{i|\mathcal{I}_s})\right]\mathrm{d}z < 0.\end{aligned} \tag{11}$$

This is because that when $z \in [\frac{\hat{x}^*_{i|\mathcal{I}_s}+\hat{x}_i(2)}{2}, \hat{x}_i(2)]$, $\mathsf{f}_{x_i|\mathcal{I}_s}(z)$ is a decreasing function but $\mathsf{f}_{x_i|\mathcal{I}_s}(\hat{x}^*_{i|\mathcal{I}_s} + \hat{x}_i(2) - z)$ is an increasing function. It means that

$$\begin{aligned}&\mathsf{f}_{x_i|\mathcal{I}_s}(z) - \mathsf{f}_{x_i|\mathcal{I}_s}(\hat{x}^*_{i|\mathcal{I}_s} + \hat{x}_i(2) - z)\\ <&\mathsf{f}_{x_i|\mathcal{I}_s}(\frac{\hat{x}^*_{i|\mathcal{I}_s} + \hat{x}_i(2)}{2}) - \mathsf{f}_{x_i|\mathcal{I}_s}(\hat{x}^*_{i|\mathcal{I}_s} + \hat{x}_i(2) - \frac{\hat{x}^*_{i|\mathcal{I}_s} + \hat{x}_i(2)}{2})\\ =&\mathsf{f}_{x_i|\mathcal{I}_s}(\frac{\hat{x}^*_{i|\mathcal{I}_s} + \hat{x}_i(2)}{2}) - \mathsf{f}_{x_i|\mathcal{I}_s}(\frac{\hat{x}^*_{i|\mathcal{I}_s} + \hat{x}_i(2)}{2}) = 0,\end{aligned}$$

and $\mathsf{u}(z - \hat{x}_i(2))$ is an increasing function but $\mathsf{u}(z - \hat{x}^*_{i|\mathcal{I}_s})$ is a decreasing function. Thus, one infers that

$$\begin{aligned}&\mathsf{u}(z - \hat{x}_i(2)) - \mathsf{u}(z - \hat{x}^*_{i|\mathcal{I}_s})\\ >&\mathsf{u}(\frac{\hat{x}^*_{i|\mathcal{I}_s} + \hat{x}_i(2)}{2} - \hat{x}_i(2)) - \mathsf{u}(\frac{\hat{x}^*_{i|\mathcal{I}_s} + \hat{x}_i(2)}{2} - \hat{x}^*_{i|\mathcal{I}_s})\\ =&\mathsf{u}(\frac{\hat{x}^*_{i|\mathcal{I}_s} - \hat{x}_i(2)}{2}) - \mathsf{u}(-\frac{\hat{x}^*_{i|\mathcal{I}_s} - \hat{x}_i(2)}{2}) = 0.\end{aligned}$$

Substituting the results of (9) and (11) into (8), it yields

$$\mathbf{E}\left[\mathsf{u}(\hat{x}_i(2) - x_i)|\mathcal{I}_s\right] - \mathbf{E}\left[\mathsf{u}(\hat{x}^*_{i|\mathcal{I}_s} - x_i)|\mathcal{I}_s\right] < 0.$$

For the same reason, we can also obtain that

$$\mathbf{E}\left[\mathsf{u}(\hat{x}_i(1) - x_i)|\mathcal{I}_s\right] - \mathbf{E}\left[\mathsf{u}(\hat{x}^*_{i|\mathcal{I}_s} - x_i)|\mathcal{I}_s\right] < 0.$$

Therefore, the proof is completed.

Consider the case that node s has no information of node i or knows $x_i \in [a, b]$. In this case, $\mathsf{f}_{x_i|\mathcal{I}_s}(z)$ is of an uniform distribution based on the assumption given in Section 2.2. Then, we can obtain a theorem as follows.

Theorem 3.2. If $\mathsf{f}_{x_i|\mathcal{I}_s}(z)$ is an uniform distribution function with domain $\mathcal{X}_i$, then we have

$$\hat{x}^*_{i|\mathcal{I}_s} = \arg\max_{\hat{x}_i \in \mathcal{X}_i} \oint_{\mathcal{X}_i} \mathsf{u}(\hat{x}_i - z)\mathrm{d}z.$$

Especially, if $\mathsf{u}(\cdot)$ is a symmetric and unimodal function and $\mathcal{X}_i = [a, b]$, we have $\hat{x}^*_{i|\mathcal{I}_s} = \frac{a+b}{2}$.

Proof. Since $\mathsf{f}_{x_i|\mathcal{I}_s}(z)$ is an uniform distribution function with domain $\mathcal{X}_i$, $\mathsf{f}_{x_i|\mathcal{I}_s}(z)$ is a positive constant for $z \in \mathcal{X}_i$ and equal 0 for otherwise. It follows that

$$\arg\max_{\hat{x}_i \in \mathcal{R}} \mathbf{E}\left[\mathsf{u}(\hat{x}_i - x_i)|\mathcal{I}_s\right] = \arg\max_{\hat{x}_i \in \mathcal{X}_i} \oint_{\mathcal{X}_i} \mathsf{u}(\hat{x}_i - z)\mathrm{d}z.$$

When $\mathsf{u}(\cdot)$ is symmetric and unimodal functions and $\mathcal{X}_i = [a, b]$, we have

$$\arg\max_{\hat{x}_i \in \mathcal{X}_i} \oint_{\mathcal{X}_i} \mathsf{u}(\hat{x}_i - z)\mathrm{d}z = \arg\max_{\hat{x}_i \in [a,b]} \int_a^b \mathsf{u}(\hat{x}_i - z)\mathrm{d}z.$$

If an estimate $\hat{x}_i \in [a, b]$ and $\hat{x}_i - \frac{a+b}{2} = \Delta_i > 0$, we have

$$\begin{aligned}&\int_a^b \left(\mathsf{u}(\hat{x}_i - z) - \mathsf{u}(\frac{a+b}{2} - z)\right)\mathrm{d}z\\ =&\int_{a-\Delta_i}^{b-\Delta_i} \mathsf{u}(\hat{x}_i - \Delta_i - z)\mathrm{d}z - \int_a^b \mathsf{u}(\frac{a+b}{2} - z)\mathrm{d}z\\ =&\left(\int_{a-\Delta_i}^{a} - \int_{b-\Delta_i}^{b}\right)\mathsf{u}(\frac{a+b}{2} - z)\mathrm{d}z\\ =&\left(\int_{a-\Delta_i}^{a} - \int_{a}^{a+\Delta_i}\right)\mathsf{u}(\frac{a+b}{2} - z)\mathrm{d}z < 0\end{aligned}$$

since $\mathsf{u}(\frac{a+b}{2} - z)$ is an increasing function when $z \in [a - \Delta_i, a + \Delta_i]$. If an estimate $\hat{x}_i \in [a, b]$ and $\hat{x}_i - \frac{a+b}{2} = \Delta_i < 0$, we can use the same approach to obtain that

$$\int_a^b \left(\mathsf{u}(\hat{x}_i - z) - \mathsf{u}(\frac{a+b}{2} - z)\right)\mathrm{d}z < 0.$$

Therefore, one infers that

$$\hat{x}^*_{i|\mathcal{I}_s} = \arg\max_{\hat{x}_i \in [a,b]} \int_a^b \mathsf{u}(\hat{x}_i - z)\mathrm{d}z = \frac{a+b}{2}.$$

The proof is completed.

Theorem 3.3. If $\mathsf{u}(z) = c$ for $z \in [-a, a]$ and $\mathsf{u}(z) = 0$ for $z \notin [-a, a]$, where $a, c > 0$, i.e., $\mathsf{u}(\cdot)$ is a constant function, we have

$$\hat{x}^*_{i|\mathcal{I}_s}(a) = \arg\max_{\hat{x}_i \in \mathcal{R}} \int_{\hat{x}_i - a}^{\hat{x}_i + a} \mathsf{f}_{x_i|\mathcal{I}_s}(z)\mathrm{d}z,$$

especially, we have

$$\lim_{a \to 0} \hat{x}^*_{i|\mathcal{I}_s}(a) = \arg\max_{z \in \mathcal{R}} \mathsf{f}_{x_i|\mathcal{I}_s}(z).$$

Proof. Since $\mathsf{u}(z)$ is a constant function, we have

$$\arg\max_{\hat{x}_i\in\mathcal{R}}\int \mathsf{f}_{x_i|\mathcal{I}_s}(z)\mathsf{u}(\hat{x}_i-z)\mathrm{d}z = \arg\max_{\hat{x}_i\in\mathcal{R}}$$
$$\left[\int_{\hat{x}_i-a}^{\hat{x}_i+a} c\cdot \mathsf{f}_{x_i|\mathcal{I}_s}(z)+\left(\int_{-\infty}^{\hat{x}_i-a}+\int_{-\infty}^{\hat{x}_i-a}\right)\mathsf{f}_{x_i|\mathcal{I}_s}(z)\cdot 0\right]\mathrm{d}z$$
$$=\arg\max_{\hat{x}_i\in\mathcal{R}}\int_{\hat{x}_i-a}^{\hat{x}_i+a}\mathsf{f}_{x_i|\mathcal{I}_s}(z)\mathrm{d}z=\hat{x}^*_{i|\mathcal{I}_s}(a).$$

Note that when a is small enough, one obtains

$$\int_{\hat{x}_i-a}^{\hat{x}_i+a}\mathsf{f}_{x_i|\mathcal{I}_s}(z)\mathrm{d}z\approx \mathsf{f}_{x_i|\mathcal{I}_s}(\hat{x}_i)\cdot 2a.$$

Thus, when $a\to 0$, we have

$$\arg\max_{\hat{x}_i\in\mathcal{R}}\int_{\hat{x}_i-a}^{\hat{x}_i+a}\mathsf{f}_{x_i|\mathcal{I}_s}(z)\mathrm{d}z=\arg\max_{\hat{x}_i\in\mathcal{R}}\mathsf{f}_{x_i|\mathcal{I}_s}(\hat{x}_i)\cdot 2a$$
$$=\arg\max_{z\in\mathcal{R}}\mathsf{f}_{x_i|\mathcal{I}_s}(z),$$

which completes the proof.

Note from the above theorem, the optimal estimation is the point that $\arg\min\Pr\{|\hat{x}_i-x_i|\le a\}$, which is equivalent to the optimal state inference defined in (He et al. (2018)), and thus the optimal state inference is viewed as a special case of the above theorem by $a=\epsilon$.

3.3 Information Value Function

In this subsection, we investigate the basic properties of the information value function and privacy lost function, and the relationship between them.

Theorem 3.4. The information value function $\mathsf{V}(\mathcal{I}_i)$ has the following properties:

(1) $\mathsf{V}(\mathcal{I}_i)\ge 0$ and $\mathsf{V}(\mathcal{I}_i)=0$ if and only if (iff)
$$\mathsf{f}_{x_i|\mathcal{I}_i}(z+\hat{x}^*_{i|\mathcal{I}_i})\mathsf{u}(-z)\equiv 0.$$
(2) $\mathsf{V}(\mathcal{I}_i)\le \mathsf{u}(0)$ and the equal sign holds iff
$$\mathsf{u}(\hat{x}^*_{i|\mathcal{I}_i}-z)=\mathsf{u}(0),\forall\ \mathsf{f}_{x_i|\mathcal{I}_i}(z)>0.$$

Proof. According to the definition of $\mathsf{V}(\mathcal{I}_i)$, we have

$$\mathsf{V}(\mathcal{I}_i)=\int \mathsf{f}_{x_i|\mathcal{I}_i}(z)\mathsf{u}(\hat{x}^*_{i|\mathcal{I}_i}-z)\mathrm{d}z=\int \mathsf{f}_{x_i|\mathcal{I}_i}(z+\hat{x}^*_{i|\mathcal{I}_i})\mathsf{u}(-z)\mathrm{d}z$$

Note that both $\mathsf{f}_{x_i|\mathcal{I}_i}(\cdot)$ and $\mathsf{u}(\cdot)$ are nonnegative functions. Then, it follows from the above equation that $\mathsf{V}(\mathcal{I}_i)\ge 0$. And, it is not difficult to follow that $\mathsf{V}(\mathcal{I}_i)=0$ iff $\mathsf{f}_{x_i|\mathcal{I}_i}(z+\hat{x}^*_{i|\mathcal{I}_i})\mathsf{u}(-z)=0$ holds for $\forall z\in\mathcal{R}$. Similarly, we have

$$\int \mathsf{f}_{x_i|\mathcal{I}_i}(z)\mathsf{u}(\hat{x}^*_{i|\mathcal{I}_i}-z)\mathrm{d}z\le\int \mathsf{f}_{x_i|\mathcal{I}_i}(z)\mathsf{u}_{\max}\mathrm{d}z$$
$$=\mathsf{u}(0)\int \mathsf{f}_{x_i|\mathcal{I}_i}(z)\mathrm{d}z=\mathsf{u}(0),$$

and

$$\int \mathsf{f}_{x_i|\mathcal{I}_i}(z)\mathsf{u}(\hat{x}^*_{i|\mathcal{I}_i}-z)\mathrm{d}z-\mathsf{u}(0)=0$$
$$\Leftrightarrow\int \mathsf{f}_{x_i|\mathcal{I}_i}(z)\left(\mathsf{u}(\hat{x}^*_{i|\mathcal{I}_i}-z)-\mathsf{u}(0)\right)\mathrm{d}z=0$$
$$\Leftrightarrow \mathsf{f}_{x_i|\mathcal{I}_i}(z)\left(\mathsf{u}(\hat{x}^*_{i|\mathcal{I}_i}-z)-\mathsf{u}(0)\right)=0,\forall z\in\mathcal{R}$$
$$\Leftrightarrow \mathsf{u}(\hat{x}^*_{i|\mathcal{I}_i}-z)-\mathsf{u}(0)=0,\forall \mathsf{f}_{x_i|\mathcal{I}_i}(z)>0.$$

Thus, the proof is completed.

Next, we consider the properties of the relative information value function. Note from the definition of $\mathsf{V}_r(\mathcal{I}_i|\mathcal{I}_s)$ that

$$\mathsf{V}_r(\mathcal{I}_i|\mathcal{I}_s)=\mathsf{V}(\mathcal{I}^i_s)-\mathsf{V}(\mathcal{I}_s)$$
$$=\int\left(\mathsf{f}_{x_i|\mathcal{I}^i_s}(z)\mathsf{u}(\hat{x}^*_{i|\mathcal{I}^i_s}-z)-\mathsf{f}_{x_i|\mathcal{I}_s}(z)\mathsf{u}(\hat{x}^*_{i|\mathcal{I}_s}-z)\right)\mathrm{d}z$$
$$=\int\left(\mathsf{f}_{x_i|\mathcal{I}^i_s}(z+\hat{x}^*_{i|\mathcal{I}^i_s})\mathsf{u}(-z)-\mathsf{f}_{x_i|\mathcal{I}_s}(z+\hat{x}^*_{i|\mathcal{I}_s})\mathsf{u}(-z)\right)\mathrm{d}z$$
$$=\int\left(\mathsf{f}_{x_i|\mathcal{I}^i_s}(z+\hat{x}^*_{i|\mathcal{I}^i_s})-\mathsf{f}_{x_i|\mathcal{I}_s}(z+\hat{x}^*_{i|\mathcal{I}_s})\right)\mathsf{u}(-z)\mathrm{d}z.$$

Theorem 3.5. The relative information value function $\mathsf{V}_r(\mathcal{I}_i|\mathcal{I}_s)$ has the following properties:

(1) $\mathsf{V}_r(\mathcal{I}_i|\mathcal{I}_s)\ge 0$ is not always true.
(2) $\mathsf{V}_r(\mathcal{I}_i|\mathcal{I}_s)\le \mathsf{u}(0)$ and the equal sign holds iff
$$\mathsf{u}(\hat{x}^*_{i|\mathcal{I}^i_s}-z)=\mathsf{u}(0),\forall\ \mathsf{f}_{x_i|\mathcal{I}^i_s}(z)>0,$$
and $\mathsf{V}(\mathcal{I}_s)=0$.

Proof. It is not difficult to see that $\mathsf{V}_r(\mathcal{I}_i|\mathcal{I}_s)\ge 0$ iff

$$\int\left(\mathsf{f}_{x_i|\mathcal{I}^i_s}(z+\hat{x}^*_{i|\mathcal{I}^i_s})-\mathsf{f}_{x_i|\mathcal{I}_s}(z+\hat{x}^*_{i|\mathcal{I}_s})\right)\mathsf{u}(-z)\mathrm{d}z\ge 0.$$

We then provide a counter-example to show that the above inequality may not be true, i.e., $\mathsf{V}_r(\mathcal{I}_i|\mathcal{I}_s)$ could be a negative value in some cases. For example, if $\mathsf{f}_{x_i|\mathcal{I}^i_s}(z+\hat{x}^*_{i|\mathcal{I}^i_s})$ and $\mathsf{f}_{x_i|\mathcal{I}_s}(z+\hat{x}^*_{i|\mathcal{I}_s})$ are uniform distributions with domain $[-a,0]$ and $[-a,a]$, and $\mathsf{u}(\cdot)$ is a constant function with domain $[0,a]$. Clearly, we have

$$\int\left(\mathsf{f}_{x_i|\mathcal{I}^i_s}(z+\hat{x}^*_{i|\mathcal{I}^i_s})-\mathsf{f}_{x_i|\mathcal{I}_s}(z+\hat{x}^*_{i|\mathcal{I}_s})\right)\mathsf{u}(-z)\mathrm{d}z$$
$$=-\int_0^a\mathsf{f}_{x_i|\mathcal{I}_s}(z+\hat{x}^*_{i|\mathcal{I}_s})\mathsf{u}(-z)\mathrm{d}z<0,$$

in this case, which means that $\mathsf{V}_r(\mathcal{I}_i|\mathcal{I}_s)<0$.

Then, we prove the second property. From Theorem 3.4, one infers that

$$\mathsf{V}_r(\mathcal{I}_i|\mathcal{I}_s)=\mathsf{V}(\mathcal{I}^i_s)-\mathsf{V}(\mathcal{I}_s)\le\mathsf{V}(\mathcal{I}^i_s)\le\mathsf{u}(0).$$

Then, we have

$$\mathsf{V}_r(\mathcal{I}_i|\mathcal{I}_s)=\mathsf{u}(0)\Leftrightarrow\mathsf{V}(\mathcal{I}_s)=0\ \text{and}\ \mathsf{V}(\mathcal{I}^i_s)=\mathsf{u}(0),$$

where

$$\mathsf{V}(\mathcal{I}^i_s)=\mathsf{u}(0)\Leftrightarrow\mathsf{u}(\hat{x}^*_{i|\mathcal{I}^i_s}-z)=\mathsf{u}(0),\forall\ \mathsf{f}_{x_i|\mathcal{I}^i_s}(z)>0.$$

We thus have completed the proof.

From the above theorem, one sees that the relative information value may be negative, i.e., $\mathsf{V}_r(\mathcal{I}_i|\mathcal{I}_s)<0$, since more information may cause worse estimation to node s.

3.4 The Value of Information Outputs

In this subsection, we assume that both g_i and h_i are known to the attack node s, and for simplicity we let $\mathsf{g}_i(x_i(k))=x_i(k)$. We also assume that the initial states of nodes are independent of each other, i.e., there are no relationship among nodes' initial states.

Theorem 3.6. Consider PPDA, under $\mathcal{I}^i_o(0)$, the optimal estimation of x_i satisfies

$$\hat{x}^*_i=\arg\max_{\hat{x}_i\in\mathcal{R}}\int\mathsf{f}_{\theta_i(0)}(x^+_i(0)-z)\mathsf{u}(\hat{x}_i-z)\mathrm{d}z;\qquad(12)$$

and the corresponding value is

$$\mathsf{V}(\mathcal{I}^i_o(0))=\int\mathsf{f}_{\theta_i(0)}(x^+_i(0)-z)\mathsf{u}(\hat{x}^*_i-z)\mathrm{d}z,\qquad(13)$$

where $\mathsf{f}_{\theta_i(0)}$ is the PDF of $\theta_i(0)$.

Proof. With $\mathcal{I}_o^i(0)$, we have $x_i^+(0) = x_i(0) + \theta_i(0)$, where $x_i^+(0)$ is released. It follows that the possible value of $x_i(0)$ has the same distribution with that of $x_i^+(0) - \theta_i(0)$. Since $x_i^+(0)$ is fixed, the PDF of $x_i^+(0) - \theta_i(0)$ is $\mathsf{f}_{\theta_i(0)}(x_i^+(0) - z)$. Thus, we have

$$\mathsf{f}_{x_i|\mathcal{I}_o^i(0)}(z) = \mathsf{f}_{\theta_i(0)}(x_i^+(0) - z).$$

Then, substituting the above equation into (4) and (5), we obtain (12) and (13), respectively.

Theorem 3.7. Consider PPDA, suppose that $\mathcal{I}_o(k) \cup \mathcal{I}_g$ is available to node s, then we have

$$\hat{x}_i^* = \arg\max_{\hat{x}_i \in \mathcal{R}} \int \mathsf{f}_{\theta_i(0)|\theta_i(1),\cdots,\theta_i(k)}(x_i^+(0) - z)\mathsf{u}(\hat{x}_i - z)\mathrm{d}z; \quad (14)$$

and

$$\mathsf{V}(\mathcal{I}_o(k) \cup \mathcal{I}_g) = \int \mathsf{f}_{\theta_i(0)|\theta_i(1),\cdots,\theta_i(k)}\mathsf{u}(\hat{x}_i^* - z)\mathrm{d}z, \quad (15)$$

where $\mathsf{f}_{\theta_i(0)|\theta_i(1),\cdots,\theta_i(k)}$ is the PDF of $\theta_i(0)$ under the condition that $\theta_i(1), \cdots, \theta_i(k)$ are given.

Proof. When node s has the information $\mathcal{I}_o(k) \cup \mathcal{I}_g$, then the variables in (2) and their values are known to node s in each iteration. Hence, node s can obtain $x_i(\ell+1)$ from (2) for $\forall \ell = 0, 1, ..., k-1$. Then, node s gets the values of $\theta_i(\ell+1)$ for $\forall \ell = 0, 1, ..., k-1$ using (1).

Given the condition that $\theta_i(1), \cdots, \theta_i(k)$, the PDF of $\theta_i(0)$ is changed to $\mathsf{f}_{\theta_i(0)|\theta_i(1),\cdots,\theta_i(k)}(z)$. By using similar analysis of Theorem 3.6, we obtain that

$$\mathsf{f}_{x_i|\mathcal{I}_o^i(0)}(z) = \mathsf{f}_{\theta_i(0)|\theta_i(1),\cdots,\theta_i(k)}(x_i^+(0) - z),$$

and substituting it into (4) and (5) gives the results (14) and (15), respectively.

From the above theorem, it is observed that if $\theta_i(0)$ is independent of $\theta_i(1), ..., \theta_i(k)$, we have

$$\mathsf{f}_{\theta_i(0)|\theta_i(1),\cdots,\theta_i(k)}(z) = \mathsf{f}_{\theta_i(0)}(z)$$

the optimal estimation $\hat{x}_i^*$ will not change and

$$\mathsf{V}(\mathcal{I}_o^i(0)) = \mathsf{V}(\mathcal{I}_o(k) \cup \mathcal{I}_g).$$

It means that the topology information and the information output of nodes in PPDA after iteration 1 are valueless in this case. However, if $\theta_i(0), \cdots, \theta_i(k)$ are correlated, the relationship among them may change the optimal estimation $\hat{x}_i^*$, and the value of information output would not be zero.

4. CONCLUSIONS

In this paper, we investigated the problem of the optimal estimation and information quantification considering private state interference over privacy preserving distributed algorithms. We introduced the novel definitions of optimal estimation, information value and information relative value, respectively, based on the utility function of the state inference accuracy. A theoretical framework was provided for the optimal estimation and information values, where the closed-form expression of the optimal estimation and the properties of the information value function were obtained. Since the state inference can be viewed as the optimal estimation and the information value denotes the privacy lost, the proposed framework provides the foundations to the further privacy analysis and algorithm design. More fundamental theoretical analysis on information value and the application development will be considered in our future works.

REFERENCES

Blondel, V.D., Hendrickx, J.M., Olshevsky, A., and Tsitsiklis, J.N. (2005). Convergence in multiagent coordination, consensus, and flocking. In *Proceedings of the 44th IEEE Conference on Decision and Control*, 2996–3000.

Fogel, E. and Huang, Y.F. (1982). On the value of information in system identification—bounded noise case. *Automatica*, 18(2), 229–238.

He, J., Cai, L., and Guan, X. (2018). Preserving data-privacy with added noises: Optimal estimation and privacy analysis. *IEEE Transactions on Information Theory*, 64(8), 5677–5690.

Howard, R.A. (1966). Information value theory. *IEEE Transactions on Systems Science and Cybernetics*, 2(1), 22–26.

Huang, Z., Mitra, S., and Dullerud, G. (2012). Differentially private iterative synchronous consensus. In *Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society*, 81–90.

Huang, Z., Hu, R., Guo, Y., Chan-Tin, E., and Gong, Y. (2019). Dp-admm: Admm-based distributed learning with differential privacy. *IEEE Transactions on Information Forensics and Security*, 15, 1002–1012.

Imtiaz, H. and Sarwate, A.D. (2018). Differentially private distributed principal component analysis. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2206–2210.

Le Ny, J. and Pappas, G.J. (2013). Differentially private filtering. *IEEE Transactions on Automatic Control*, 59(2), 341–354.

Liu, J., Mou, S., and Morse, A.S. (2017). Asynchronous distributed algorithms for solving linear algebraic equations. *IEEE Transactions on Automatic Control*, 63(2), 372–385.

Manitara, N.E. and Hadjicostis, C.N. (2013). Privacy-preserving asymptotic average consensus. In *2013 European Control Conference (ECC)*, 760–765.

Mo, Y. and Murray, R.M. (2016). Privacy preserving average consensus. *IEEE Transactions on Automatic Control*, 62(2), 753–765.

Nozari, E., Tallapragada, P., and Cortés, J. (2017). Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design. *Automatica*, 81, 221–231.

Olfati-Saber, R., Fax, J.A., and Murray, R.M. (2007). Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1), 215–233.

Pasqualetti, F., Carli, R., Bicchi, A., and Bullo, F. (2010). Distributed estimation and detection under local information. *IFAC Proceedings Volumes*, 43(19), 263–268.

Rajagopalan, R. and Varshney, P.K. (2006). Data aggregation techniques in sensor networks: A survey.

Ren, W., Beard, R.W., and Atkins, E.M. (2007). Information consensus in multivehicle cooperative control. *IEEE Control Systems Magazine*, 27(2), 71–82.

Zhao, C., He, J., Cheng, P., and Chen, J. (2016). Consensus-based energy management in smart grid with transmission losses and directed communication. *IEEE Transactions on Smart Grid*, 8(5), 2049–2061.