Privacy-preserving of system model with perturbed state trajectories using differential privacy: With application to a supply chain network

Nandakumar, Lakshminarayanan; Ferrari, Riccardo M.G.; Keviczky, Tamas

DOI
10.1016/j.ifacol.2019.12.173

Publication date
2019

Document Version
Final published version

Published in
IFAC-PapersOnLine

Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnline CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 52-20 (2019) 309–314

Privacy-Preserving of System Model with Perturbed State Trajectories using Differential Privacy: With application to a Supply Chain Network ⋆

Lakshminarayanan Nandakumar * **Riccardo Ferrari** ** **Tamas Keviczky** ***

* *CGI Nederland B.V., (e-mail: srinath0393@gmail.com).*
** *Delft Center for Systems and Control, 2628CD, Delft, The Netherlands, (e-mail: R.Ferrari@tudelft.nl)*
*** *Delft Center for Systems and Control, 2628CD, Delft, The Netherlands, (e-mail: T.Keviczky@tudelft.nl)*

Abstract: Releasing state samples generated by a dynamical system model, for data aggregation purposes, can allow an adversary to perform reverse engineering and estimate sensitive model parameters. Upon identification of the system model, the adversary may even use it for predicting sensitive data in the future. Hence, preserving a confidential dynamical process model is crucial for the survival of many industries. Motivated by the need to protect the system model as a trade secret, we propose a mechanism based on differential privacy to render such model identification techniques ineffective while preserving the utility of the state samples for data aggregation purposes. We deploy differential privacy by generating noise according to the sensitivity of the query and adding it to the state vectors at each time instant. We derive analytical expressions to quantify the bound on the sensitivity function and estimate the minimum noise level required to guarantee differential privacy. Furthermore, we present numerical analysis and characterize the privacy-utility trade-off that arises when deploying differential privacy. Simulation results demonstrate that through differential privacy, we achieve acceptable privacy level sufficient to mislead the adversary while still managing to retain high utility level of the state samples for data aggregation.

Keywords: Differential Privacy, State Trajectories, Model Parameters, Data Aggregation

1. INTRODUCTION

Innovative business models are the key to success in any industry Tohidi and Jabbari (2012). Companies make significant capital investments to develop innovative models for improving the performance of their existing systems Bottegal et al. (2017). Many companies restrict the use of their ideas by filing patents or by hiding certain features of their model and capitalize on it to generate revenues and profits for their business He et al. (2008). Thus, protecting a model as a confidential trade secret play an important role in the growth and innovation of a company. For example, beverage companies successfully ensure that the syrup formula cannot be reverse engineered by using their beverage available in the market. However, when dealing with dynamical systems, the problem of protecting the model becomes challenging since system identification techniques can be employed for identifying a black-box system model or to extract the parameters of a grey-box system. Furthermore, with the advent of machine learning, data mining techniques can be applied for system identification Pillonetto (2016); Saitta et al. (2006). Such techniques can help competitors unravel a company's trade secret. This in turn advocates for privacy-preserving solutions while disclosing information associated with the model to a third party for statistical purposes.

One of the most important quantities to be computed in surveys and audits is the aggregate information. Data aggregation enables performing data mining applications for understanding important phenomena, such as traffic congestion patterns, influenza outbreaks, etc. Fan and Xiong (2014). For example, consider a case where an external party wants to periodically know the number of products produced and sold in a particular time period, and use it to compute the average for statistical analysis. Companies sometimes voluntarily release this information in the form of their production report, sales report etc, and in some cases, they are mandated to release even the amount of raw materials used in the production of their product. For example, in the case of beverage industries, the main raw material is water, and the amount of water used in production must be disclosed to a governmental body for keeping a check on the underground water levels. However, as illustrated in the next section, periodically releasing this information may also be used to identify sensitive model parameters. Thus, there is a fundamental need to preserve the system model as well as share the

⋆ This work was supported by the TU Delft Safety and Security Institute under the DSyS Grant.

10.1016/j.ifacol.2019.12.173

information generated by the model with reasonable utility levels.

2. EXISTING WORK AND OUR CONTRIBUTIONS

There has been a large body of work done in the statistics and database literature on disclosure limitation and privacy-preserving publication of data Dwork and Roth (2014). The recently proposed formulation of privacy by Dwork Dwork (2006) called Differential Privacy (DP) has been adopted as a standard definition of privacy in many applications offering quantitative privacy guarantees. Originally, differential privacy was proposed for a static system as a measure of maximizing the accuracy of queries from statistical databases while minimizing the probability of identifying the individuals. In recent years, differential privacy has gained a significant amount of attention in the context of dynamical systems and control where researchers have used it for a diverse set of objectives such as control Huang et al. (2014); Wang et al. (2014), consensus Nozari et al. (2017); Mo and Murray (2017), optimization Hale and Egerstedty (2015); Nozari et al. (2016); Ling et al. (2016) and distributed fault diagnosis Rostampour et al. (2018). Although differential privacy has made its way to systems and control, very little work has been done in utilizing differential privacy for protecting the system model. To the best of our knowledge, only Bottegal et al. (2017), Le Ny and Pappas (2013); Katewa et al. (2015) discuss the problem of model-preservation in the context of differential privacy. In Bottegal et al. (2017), differential privacy was explored for designing output noises for preserving the model. In Katewa et al. (2015) differential privacy was used to protect the consensus network topology from an eavesdropper who may have an unauthorized access to the central estimator. They present a mechanism where each agent in the network adds differential private noise to its output, and transmits it to the central estimator to estimate the topology matrix and its eigenvalues. However, the authors in Katewa et al. (2015) do not define any utility function and characterize the privacy-utility trade-offs. In Le Ny and Pappas (2013), the authors present several perturbation techniques to release a model describing the dynamics of a large group of users responding to a common single input signal and producing a single output signal. However, their approach assumes a *trusted* data aggregator which receives confidential scalar model parameters from the other participants and then uses it to publish a SISO transfer function describing the relationship between common input and aggregate output. These gaps in the present state-of-the-art and a strong fundamental need to protect the system model without a trusted aggregator motivates us to explore this problem in-depth using differential privacy.

The major contributions of this paper are as follows:

- We propose a novel differential privacy mechanism to preserve the system model (system matrix A) privacy while releasing the state sequences for data aggregation *without* a trusted intermediary aggregator. We also derive an analytical expression to estimate the minimum noise level required to guarantee differential privacy.
- Furthermore, we define a utility function in our problem setup, and characterize the resulting privacy-utility trade-off using numerical simulations. We also analyze the effect of the DP mechanism w.r.t the various privacy design parameters.

3. MOTIVATING EXAMPLE

Consider an example of supply chain economics depicted in Figure 1 which involves three different parties,: Supplier (S), Producer (P) and Retailer (R). S purchases the quantity $u(k)$ of raw materials at each day k and discards a fraction δ_1 of raw materials when shipping a fraction α_1 to P. P transforms these raw materials into finished products and sells a fraction of α_2 to R while discarding a fraction of δ_2 due to faults, low quality etc. Finally R returns a fraction β_3 of defective products every day, and sells a fraction γ_3 to customers. This supply chain model can be recast into a discrete-time linear state space equation as follows:

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \underbrace{\begin{bmatrix} 1-\alpha_1-\delta_1 & 0 & 0 \\ \alpha_1 & 1-\alpha_2-\delta_2 & \beta_3 \\ 0 & \alpha_2 & 1-\beta_3-\gamma_3 \end{bmatrix}}_{A} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u(k). \tag{1}$$

The state $x_1(k)$ represents the amount of raw material in S, $x_2(k)$ and $x_3(k)$ denotes the products in P and R respectively. The output $y(k)$ represents the products sold to customers. The system model matrix A contains information about the percentage of products discarded in each chain, the percentage of defective products returned by the retailer to producer, and typically companies would like to keep such internal information private when disclosing the state vector $x(k)$ to any external organization for survey and auditing purposes. Exposing the system matrix might damage the reputation of each party in the supply chain and may even result in the breach of trust among the customers. For example, consider exposing the information δ_1 and β_3 to the public. A higher δ_3 implies that a large percentage of raw material supplies have been discarded due to poor quality. Higher β_3 implies that the percentage of defective products produced is high. Information such as the percentage of products discarded in each chain or percentage of defective products may also give insight into sensitive information such as the quality and efficiency of the production machine, thinking pattern behind rejecting products etc. Furthermore, if a competitor gets hold of the supply chain model i.e. the A matrix of the target company, then it could very likely predict the amount of production of the target company with high accuracy and beat them to the market. Thus, the system matrix A must be protected while releasing the information of the state vectors for data aggregation purposes.

Now, let us discuss the scenario of releasing the state samples after a certain time duration of T days without perturbing the state samples, i.e., no privacy. These state vectors are typically required in surveys and audits for measuring the aggregate amount of products from the supplier, producer and retailer side for a period of time. Consider the input $u(k)$ to be a *Dirac delta function* of magnitude C i.e. $u(k) = C\delta$. This means once for every T days, the supplier S purchases a raw material of quantity C. Thus for $k \in (0, T]$, the state space equation in (1) reduces to

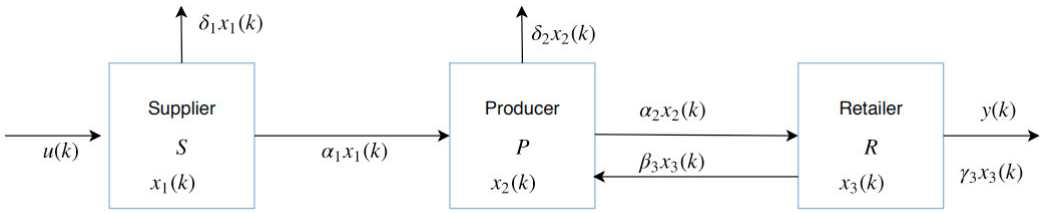

Fig. 1. Internal flow of information between the Supplier, Producer and Retailer in a simple supply chain model.

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \underbrace{\begin{bmatrix} 1-\alpha_1-\delta_1 & 0 & 0 \\ \alpha_1 & 1-\alpha_2-\delta_2 & \beta_3 \\ 0 & \alpha_2 & 1-\beta_3-\gamma_3 \end{bmatrix}}_{A} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix}. \tag{2}$$

Let the external query at a given time index k be

$$Q(k) = x(k),$$

Upon repeating this query for a period T, the adversary arrives at the following relation

$$\underbrace{[x_1 \ x_2 \ \cdots x_T]}_{\mathbf{X}_f} = A \ \underbrace{[x_0 \ x_1 \ \cdots x_{T-1}]}_{\mathbf{X}_p}. \tag{3}$$

Equation (3) can be solved for A by

$$\bar{A} = \mathbf{X}_f \, \mathbf{X}_p^T \, (\mathbf{X}_p \, \mathbf{X}_p^T)^{-1}. \tag{4}$$

Without the presence of noise, the estimate $\bar{A} = A$ for sufficient time samples which means the adversary could easily infer the sensitive information about the model, thus resulting in a privacy breach.

4. PROBLEM SETUP

In this paper an autonomous linear time-invariant system with perfect measurement will be considered

$$\begin{cases} x(k+1) & = Ax(k) \\ y(k) = x(k) & \end{cases}, \tag{5}$$

where $x(0) \neq 0$ is assumed to be publicly known. The notation $x_A(k)$ will be used to represent the value of x generated by a particular system matrix A at time instant k. The sequence up to time T is denoted by $\mathbf{X}_A[0:T]$, while $\|.\|_p$ represents the $p-$norm of a vector or the induced $p-$norm of a matrix with $p \in [1, \infty)$. $Lap(0,b)^n$ denotes n dimensional Laplace distribution with *i.i.d.* components, each with a probability density function $p(x) = \frac{1}{2b} e^{-\frac{x}{b}}$.

4.1 Differential Privacy for System Model Identification

Differential privacy is a method of introducing randomness or noise into a particular system such that the adversary cannot uniquely identify the data to be protected while at the same time computing the query from the data with considerable amount of utility Dwork and Roth (2014). In this case, the data to be protected is the model or the system matrix A while the query is

$$Q(k) = x(k). \tag{6}$$

The noise is calibrated according to the sensitivity of the query and added to the state vectors $x(k)$ as given by (7). These perturbed samples will be transmitted to an external entity (potential adversary) who wants to compute the aggregate of all the state vectors up to time T. To design the noise $\eta(k) \in \mathbb{R}^n$ in $\mathcal{M}(k)$ given by (7),

$$\mathcal{M}(k): \qquad \tilde{x}(k) = x(k) + \eta(k), \tag{7}$$

we first present the following definitions:

Definition 1. (β *Adjacency*) : Two state matrices A and A' are β adjacent (denoted by Adj^β) if for some $\beta \geq 0$,

$$\text{Adj}^\beta \stackrel{\text{def}}{=} \|A - A'\|_2 \leq \beta \, . \tag{8}$$

Remark 1. Adjacency in differential privacy captures the quantity to be hidden. Contrary to the standard definition of adjacency used in DP for static and dynamic systems Ny and Pappas (2014); Le Ny and Pappas (2013); Fan and Xiong (2014), where adjacency is defined w.r.t. changes in only component i while keeping the other components $j \neq i$ unchanged, our definition allows changes that can possibly affect various components of the state matrix A. While other p-norms are possible in (8), we decided to follow the definition given in Katewa et al. (2015), where the privacy of topology in consensus networks is addressed.

Definition 2. (*Sensitivity*) : The *sensitivity* $\Delta(x)$ represents the maximum possible difference between two state vectors generated by any two β-adjacent state matrices starting from the same initial condition $x(k-1)$. In this paper, we define sensitivity in terms of the L_1 norm

$$\begin{aligned} \Delta(x) &= \left\| x_A(k) - x_{A'}(k) \right\|_1, \\ &= \left\| Ax(k-1) - A'x(k-1) \right\|_1. \end{aligned} \tag{9}$$

Figure 2 illustrates the adjacency relationship and sensitivity.

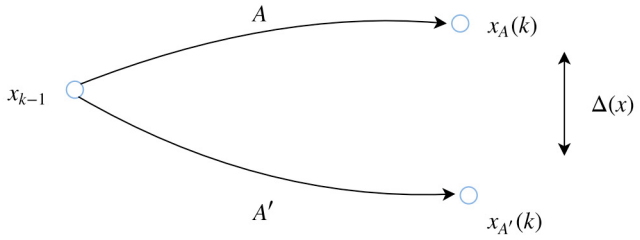

Fig. 2. Illustration of the Adjacency Relationship and Sensitivity.

Definition 3. (*Finite Time* ϵ *Differential Privacy*): Given $\epsilon \geq 0$, the mechanism $\mathcal{M}$ given in (10)

$$\mathcal{M}: \qquad \tilde{\mathbf{X}}[0:T] = \mathbf{X}[0:T] + \eta[0:T], \tag{10}$$

preserves ϵ-differential privacy up to time T if for any two β-adjacent state matrices A and A', and for any $R \subseteq \text{range}(\mathcal{M})$ the following relationship is satisfied:

$$\Pr\left[\tilde{\mathbf{X}}_A[0:T] \in R\right] \leq e^{\epsilon} \Pr\left[\tilde{\mathbf{X}}_{A'}[0:T] \in R\right]. \quad (11)$$

Definition 3 says that if the state matrix changes from A to an A' that is β- adjacent, then the corresponding state trajectory statistics change at most by a factor of e^{ϵ}, where ϵ quantifies the *privacy loss*. Clearly, when the privacy loss is minimal, the utility of the resulting state samples becomes minimal as well. Thus, to quantify the degree of utility of the resulting state samples, we introduce the following definition:

Definition 4. (*Utility*) : Utility $\mathcal{U}$ is defined as

$$\mathcal{U} = \mathbb{E}\left[1 - \frac{\left\|X_{avg} - \tilde{X}_{avg}\right\|_1}{2\max(\|X_{avg}\|_1, \|\tilde{X}_{avg}\|_1)}\right], \quad (12)$$

where $X_{avg} = \frac{1}{T}\sum_{k=0}^{T} x(k)$ and $\tilde{X}_{avg} = \frac{1}{T}\sum_{k=0}^{T} \tilde{x}(k)$.

Remark 2. By design, it holds $0 \leq \mathcal{U} \leq 1$, with $\mathcal{U}=1$ when $X_{avg} = \tilde{X}_{avg}$, and $\mathcal{U}=0$ when $X_{avg} = -\tilde{X}_{avg}$.

4.2 Adversarial Estimate

Upon receiving the perturbed state samples $\tilde{x}(k)$, the adversary will be able to obtain an estimate denoted by $\hat{A}$ as follows:

$$\hat{A} = \underset{\tilde{A} \in R^{n \times n}}{\arg\min} \left\|\tilde{\mathbf{X}}_f - \tilde{A}\tilde{\mathbf{X}}_p\right\|, \quad (13)$$

where $\tilde{\mathbf{X}}_f = \tilde{\mathbf{X}_A}[1:T]$ and $\tilde{\mathbf{X}}_p = \tilde{\mathbf{X}_A}[0:T-1]$. The DP-mechanism in (7) will ensure that the state samples generated by A and $A' \in Adj^{\beta}$ are almost equally likely i.e., the state samples generated by Adj^{β} state matrices are *statistically not very different* Katewa et al. (2015). Hence by observing these state samples, the adversary will not be able to distinguish between A and A' with high confidence level while retaining the necessary information to compute the query with reasonable utility. Thus, the privacy of the state matrix A is preserved which results in the estimation error

$$\mathbf{E} = \mathbb{E}\left[\left\|A - \hat{A}\right\|_2\right]. \quad (14)$$

The term $\left\|A - \hat{A}\right\|_2$ is defined as the perturbation norm.

4.3 Research Problem

With the above problem setup: design the noise $\eta(k)$ in (7) satisfying the DP-definition in (11) for a given ϵ, β, and characterize the resulting trade-off between the privacy and utility levels.

5. PROPOSED SOLUTION

In this section, we present a noise adding DP-mechanism to protect the privacy of the system matrix A. The most common way to implement the DP mechanism is to add noise generated according to the Laplacian distribution based on the sensitivity Δ of the system Dwork and Roth (2014). We consider the sensitivity from the system matrix A to the state vector $x(k)$ since the goal is to protect the A matrix. To design $\eta(k)$, we first need to estimate the sensitivity up to time T given as:

$$\Delta(T) = \max_{A,A':Adj^{\beta}} \left\|\mathbf{X}_A[0:T] - \mathbf{X}_{A'}[0:T]\right\|_1. \quad (15)$$

Once $\Delta(T)$ is obtained, the following theorem stated in Ny and Pappas (2014) provides a sufficient condition for the noise design.

Theorem 1. The mechanism $\mathcal{M}$ in (10) is ϵ- differentially private up to time T if $\eta(k)$ is white Laplacian noise with distribution $\eta(k) \sim Lap(0,b)^n$ and $b \geq \frac{\Delta(T)}{\epsilon}$.

Proof 1. The proof follows from Theorem 2 in Ny and Pappas (2014).

Remark 3. Notice that the noise design parameter b is inversely proportional to ϵ. Thus as ϵ decreases, the noise parameter b increases resulting in a flat tail Laplacian distribution curve i.e. the probability of picking a random number close to zero is very low and hence a higher noise level is generated. As a result, when ϵ decreases, the privacy level increases and vice-versa. Also, notice the noise parameter b is directly proportional to the sensitivity $\Delta(T)$. Thus the lower the sensitivity, the lower the noise that needs to be added and vice-versa.

Intuitively, if the sensitivity is low, then for two different $\beta-$adjacent matrices, the change in the corresponding state trajectories will not be large, and hence the level of noise required to make the two state trajectories *"statistically not very different"* will also be small Katewa et al. (2015). Since sensitivity is crucial in designing a DP-induced noise, we may try to calculate $\Delta(T)$. However it is difficult to obtain analytical expression for $\Delta(T)$. Hence, we propose the following theorem instead that characterizes the upper bound on the sensitivity. Through this upper bound, we obtain the minimum noise level required to ensure DP.

Theorem 2. The sensitivity $\Delta(T)$ is upper bounded by

$$\Delta(T) \leq \sqrt{n}\,\beta \left\|x(0)\right\|_1 \sum_{k=0}^{T} \left\|A^k\right\|_1 \quad (16)$$

Proof 2. To obtain the upper bound for the sensitivity function $\Delta(T)$ given by (15), let us consider two measurements x_A and x'_A produced by β adjacent state matrices A and A'

$$\begin{aligned}\left\|x_A(k+1) - x_{A'}(k+1)\right\|_1 &= \left\|Ax(k) - A'x(k)\right\|_1 \\ &\leq \left\|A - A'\right\|_1 \left\|x(k)\right\|_1 \\ &\leq \sqrt{n}\left\|A - A'\right\|_2 \left\|x(k)\right\|_1 \\ &\leq \sqrt{n}\,\beta \left\|x(0)\right\|_1 \left\|A^k\right\|_1\end{aligned}$$

$$\begin{aligned}\left\|\mathbf{X}_A[0:T] - \mathbf{X}_{A'}[0:T]\right\|_1 &= \sum_{k=0}^{T} \left\|x_A(k+1) - x_{A'}(k+1)\right\|_1 \\ &\leq \sqrt{n}\,\beta \left\|x(0)\right\|_1 \sum_{k=0}^{T} \left\|A^k\right\|_1\end{aligned} \quad (17)$$

Remark 4. From (17), we can see that the sensitivity $\Delta(T)$ is bounded for a finite value of T. Also, when A is a Schur matrix, $\Delta(T)$ is bounded for all values of T. Anyway,

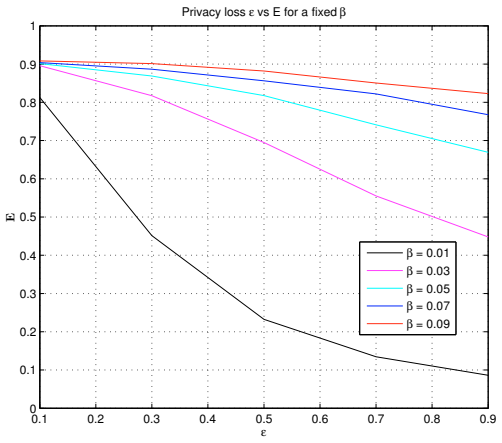

Fig. 3. Illustration of the DP Mechanism simulated for various values of β and ϵ

$\Delta(T)$ is a monotone function of T, which formalizes the intuitive fact that given a longer observation horizon T, and adversary can get a more accurate estimate of the model. Also, we must note that $\Delta(T)$ can be very large in case A is not or close to not being Schur. The noise $\eta(k)$ can be generated by setting b as follows

$$b = \frac{\Delta(T)}{\epsilon} = \sqrt{n}\ \frac{\beta}{\epsilon}\ \|x(0)\|_1 \sum_{k=0}^{T} \|A^k\| \qquad (18)$$

Note that the ratio $\lambda \stackrel{\text{def}}{=} \frac{\beta}{\epsilon}$ represents the *privacy level* Katewa et al. (2015) whereas ϵ represents the *privacy loss* with β being another privacy design parameter. If β increases for a fixed ϵ, then DP is ensured for a larger set of state matrices. Similarly if ϵ decreases, then the privacy level increases.

6. SIMULATION RESULTS

We consider the same supply chain example as explained in Section 3. The system matrix A is taken as

$$A = \begin{bmatrix} 0.16 & 0 & 0 \\ 0.8 & 0.25 & 0.01 \\ 0 & 0.7 & 0.19 \end{bmatrix}.$$

We set $T = 15$ days and $x_0 = [1000\ 0\ 0]^T$.

6.1 System Matrix Estimation

The adversary obtains an estimate of the system matrix $\hat{A}$ given by (13). Figure 3 shows the estimation error $\mathbf{E}$ simulated for various values of β and ϵ. It is clear that as ϵ increases the estimation error tends to zero i.e., the adversarial estimate approaches the true value, and hence there is no privacy. It is evident that when β is increased for a fixed ϵ, DP is ensured for a larger set of state matrices. As β increases, the noise added also increases due to the rise in sensitivity, and hence, the estimation error is found to be higher for larger β's as we increase ϵ. Next, we simulated for multiple noise realizations and calculated the sample mean to approximate the expected error $\mathbf{E}$ given in (14). Figure 4 shows the variation of the expected error w.r.t.

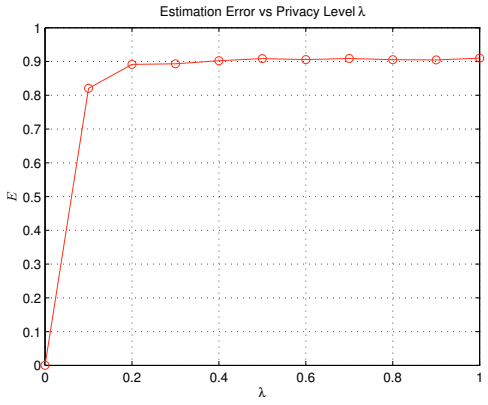

Fig. 4. State Matrix Estimation Error.

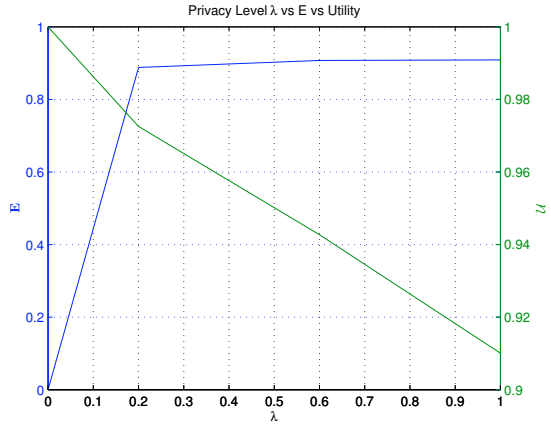

Fig. 5. Privacy-Utility Trade-off Characterization

the privacy level λ. Naturally, when $\lambda = 0$ i.e. no privacy, the expected error $\mathbf{E} = 0$. As expected, the estimation accuracy degrades with increase in the privacy level.

6.2 Privacy vs Utility

We now characterize the privacy-utility trade-off by calculating the estimation error $\mathbf{E}$ and utility $\mathcal{U}$ for various values of the privacy level λ. Figure 5 depicts the plot between the estimation error and utility for various values of λ. This plot helps in choosing an appropriate privacy and utility level for this system. For example, when $\lambda = 0.6$ we have,

$$X_{avg} = \begin{bmatrix} 79.3651 \\ 85.6429 \\ 74.0124 \end{bmatrix}, \quad \tilde{X}_{avg} = \begin{bmatrix} 95.9388 \\ 81.4923 \\ 83.1509 \end{bmatrix}, \quad \mathcal{U} = 0.9427.$$

$$\hat{A} = \begin{bmatrix} 0.0014 & -0.0019 & -0.0040 \\ 0.0031 & 0.0030 & -0.0033 \\ -0.0064 & 0.0036 & 0.0024 \end{bmatrix}, \quad \mathbf{E} = 0.9075.$$

Notice the difference between the actual A matrix and the matrix $\hat{A}$ obtained by the adversary from the perturbed state trajectories using (4). Let us now compare both the matrices element-wise. From A, we see that $a_{12} = 0$, which means in the actual supply chain model (see Figure 1), there was no *backward communication* between the producer P and supplier S. Similarly, $a_{13} = 0$ and $a_{31} = 0$ i.e., there was no communication link between the

supplier S and retailer R in the actual model. However, the presence of non-zero components in $\hat{a}_{12}$, $\hat{a}_{13}$, $\hat{a}_{31} \neq 0$ from $\hat{A}$ obtained by the adversary suggests him/her otherwise. The adversary certainly cannot deduce with high confidence level regarding the presence or absence of *backward communication* between different parties with $\mathbf{E} \approx 0.91$, thereby, preserving the privacy of the given model. Note that as emphasized in Remark 1, this was possible because our definition of adjacency allowed changes in multiple components of the A matrix. Next, analyzing the eigenvalues of A, we have:

$$\gamma = [0.1311 \;\; 0.3089 \;\; 0.1600]^T ,$$

whereas the adversarial estimated eigenvalues are:

$$\hat{\gamma} = [-0.0032 \;\; 0.0050 + 0.0036i \;\; 0.0050 - 0.0036i]^T .$$

The adversary does not obtain any information regarding the system properties from the estimated eigenvalues and can even be misled to falsely conclude the presence of oscillation in the system due to complex poles, whereas, no such oscillatory behavior is present in the original system. Thus, in this case, through differential privacy, we are able to mislead the adversary in several directions while still managing to retain a high data utility ($\mathcal{U} \approx 0.95$).

7. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed a differential privacy mechanism to protect the system matrix. The proposed mechanism adds synthetic noise generated according to the Laplacian probability distribution and prevents an external adversary from uniquely identifying the system matrix when accessing the state samples to compute the aggregate information. We derived an analytical bound on the sensitivity function and calculated a sufficient noise level required to ensure DP. Simulation results validate our DP approach and show that the expected error of the system matrix estimation increases leading to an increase in the privacy level. Furthermore, we characterized the resulting trade-off between the privacy and utility level using empirical evidence. Using this characterization, we saw how differential privacy aids in the process of introducing uncertainty and ambiguity in the adversarial mind while still retaining higher levels of utility. Future work may involve designing an asymptotically decaying noise to retain even higher levels of data utility and providing mathematical proofs that the resulting mechanism preserves differential privacy. We also plan to extend this framework to non-autonomous systems with the goal of protecting both the system and input matrices.

REFERENCES

Bottegal, G., Farokhi, F., and Shames, I. (2017). Preserving privacy of finite impulse response systems. *IEEE Control Systems Letters*, 1(1), 128–133.

Dwork, C. (2006). Differential privacy. In M. Bugliesi, B. Preneel, V. Sassone, and I. Wegener (eds.), *Automata, Languages and Programming*, 1–12. Springer Berlin Heidelberg, Berlin, Heidelberg.

Dwork, C. and Roth, A. (2014). The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3–4), 211–407.

Fan, L. and Xiong, L. (2014). An adaptive approach to real-time aggregate monitoring with differential privacy. *IEEE Transactions on Knowledge and Data Engineering*, 26(9), 2094–2106.

Hale, M.T. and Egerstedty, M. (2015). Differentially private cloud-based multi-agent optimization with constraints. In *2015 American Control Conference (ACC)*, 1235–1240. doi:10.1109/ACC.2015.7170902.

He, X., Zhang, F., and Adam, N. (2008). Towards ranking the importance of patents. In *2008 IEEE Symposium on Advanced Management of Information for Globalized Enterprises (AMIGE)*, 1–5.

Huang, Z., Wang, Y., Mitra, S., and Dullerud, G.E. (2014). On the cost of differential privacy in distributed control systems. In *Proceedings of the 3rd International Conference on High Confidence Networked Systems*, HiCoNS '14, 105–114. ACM, New York, NY, USA.

Katewa, V., Chakrabortty, A., and Gupta, V. (2015). Protecting privacy of topology in consensus networks. In *2015 American Control Conference (ACC)*, 2476–2481.

Le Ny, J. and Pappas, G.J. (2013). Privacy-preserving release of aggregate dynamic models. In *Proceedings of the 2Nd ACM International Conference on High Confidence Networked Systems*, HiCoNS '13, 49–56.

Ling, Q., Xu, W., Wang, M., and Li, Y. (2016). *Distributed Constrained Optimization Over Cloud-Based Multi-agent Networks*, 91–102. Springer International Publishing.

Mo, Y. and Murray, R.M. (2017). Privacy preserving average consensus. *IEEE Transactions on Automatic Control*, 62(2), 753–765.

Nozari, E., Tallapragada, P., and Corts, J. (2016). Differentially private distributed convex optimization via objective perturbation. In *2016 American Control Conference (ACC)*, 2061–2066.

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

Ny, J.L. and Pappas, G.J. (2014). Differentially private filtering. *IEEE Transactions on Automatic Control*, 59(2), 341–354.

Pillonetto, G. (2016). The interplay between system identification and machine learning. *CoRR*, abs/1612.09158. URL http://arxiv.org/abs/1612.09158.

Rostampour, V., Ferrari, R., Teixeira, A.M., and Keviczky, T. (2018). Differentially-private distributed fault diagnosis for large-scale nonlinear uncertain systems. *IFAC-PapersOnLine*, 51(24), 975–982.

Saitta, S., Raphael, B., and Smith, I.F.C. (2006). Combining two data mining methods for system identification. In I.F.C. Smith (ed.), *Intelligent Computing in Engineering and Architecture*, 606–614. Springer Berlin Heidelberg, Berlin, Heidelberg.

Tohidi, H. and Jabbari, M.M. (2012). Innovation as a success key for organizations. *Procedia Technology*, 1, 560 – 564.

Wang, Y., Huang, Z., Mitra, S., and Dullerud, G.E. (2014). Entropy-minimizing mechanism for differential privacy of discrete-time linear feedback systems. In *53rd IEEE Conference on Decision and Control*, 2130–2135.