PRESERVING TRAJECTORY PRIVACY IN DRIVING DATA RELEASE

Yi Xu, Chong Xiao Wang, Yang Song, Wee Peng Tay

School of Electrical and Electronic Engineering, Nanyang Technological University

ABSTRACT

Real-time data transmissions from a vehicle enhance road safety and traffic efficiency by aggregating data in a central server for data analytics. When drivers share their instantaneous vehicular information for a service provider to perform a legitimate task, a curious service provider may also infer private information it has not been authorized for. In this paper, we propose a privacy preservation framework based on the Hilbert Schmidt Independence Criterion (HSIC) to sanitize driving data to protect the vehicle's trajectory from adversarial inference while ensuring the data is still useful for driver behavior detection. We develop a deep learning model to learn the HSIC sanitizer and demonstrate through two datasets that our approach achieves better utility-privacy trade-offs when compared to three other benchmarks.

***Index Terms*—** Driver behavior detection, trajectory privacy, data sanitization.

1. INTRODUCTION

Recent years have seen a dramatically accelerating pace in the development of intelligent transport systems (ITS) [1], at the heart of which is the information fusion and knowledge exchange between road users. What comes with the innovative services provided by ITS are potential privacy attacks as there is nothing stopping a data recipient from abusing the data for illegitimate purposes. For example, in traffic monitoring systems, individual users send anonymized personal location traces continuously to a data aggregator to aid in traffic state estimation. However, an adversary may link an anonymous GPS trace to a particular person provided additional knowledge of the person's residence or working location [2]. In such cases, employing data privacy [3, 4] to remove the personally identifiable information may not necessarily prevent the service provider from making statistical inferences about some hidden variables [5]. As a consequence, users may feel reluctant to share their driving data with any third-party due to privacy leakage concerns.

This research is supported by the Singapore Ministry of Education Academic Research Fund Tier 2 grant MOE-T2EP20220-0002 and A*STAR under its RIE2020 Advanced Manufacturing and Engineering (AME) Industry Alignment Fund – Pre Positioning (IAF-PP) (Grant No. A19D6a0053). The computational work for this article was partially performed on resources of the National Supercomputing Centre, Singapore (https://www.nscc.sg).

We consider the case where drivers share fine-grained driving data in either real-time or batch fashion with a service provider for it to perform driver behavior recognition. Many usage-based automotive insurance companies analyze the aggregate statistics of customers' driving behavior to determine whether premiums should be reduced or not [6]. Meanwhile, drivers hope that any entity, after acquiring their driving data, cannot make accurate inference about their trajectory since it can be linked to personable information. This can not be achieved by data encryption or hiding the driver identity. We resort to the notion of inference privacy [7–9] that sanitizes raw data to limit the amount of contained private information.

The works [10–15] have explored privacy preserving mechanisms for systems with known system models. They however are not applicable to our case in which the underlying data model is unknown. The reinforcement learning framework proposed in [9] that protects privacy in time-series data sharing is limited to applications with discrete data. Several data-driven approaches are proposed based on the generative adversarial networks (GAN) [16], including generative adversarial privacy (GAP) [17] and compressive privacy generative adversarial network (CPGAN) [18]. However, the quality of privacy quantification given by adversarial training is determined by the capability of the neural network chosen and trained to act as the adversary, which in practice cannot incorporate all possible adversarial strategies. This implies that such approaches cannot achieve a universal privacy guarantee. In this paper, we instead explicitly incorporate a privacy function in our learning architecture.

In this paper, we develop a privacy-preserving framework to protect the trajectory privacy for driving data release. Our main contributions are the following: (1) we propose to use a non-parametric function to quantify the level of privacy leakage. Thus, our framework does not invoke adversarial training unlike most of the data-driven approaches. (2) We ensure the sequence of sanitized data can be still used for the utility application of driver behavior detection. We verify our framework on two real-world datasets.

The rest of this paper is organized as follows. In Section 2, we present our framework and utility model. In Section 3, we detail the privacy function used for measuring the amount of privacy leakage. In Section 4, we test our framework on two real-world datasets and finally make conclusions in Section 5.

2. PROBLEM FORMULATION

In this section, we present our problem formulation and utility model for driver behavior recognition.

2.1. Overall framework

Let $\mathbf{p}_k$ be the coordinates of a vehicle's location at time $k \geq 1$, and $\mathbf{s}_k = \mathbf{p}_k - \mathbf{p}_{k-1}$ denote the change in the vehicle's location from time $k-1$ to time k. Inertial measurement units installed on the vehicle generate a data vector $\mathbf{x}_k$ at each time k as the measurement of this positional change. With the help of prior knowledge such as map information, the vehicle's moving trajectory can be estimated from the sequence of data $(\mathbf{x}_i)_{i\geq 1}$. To make it difficult for a curious service provider to make inferences about the trajectory, we transform each $\mathbf{x}_i$ using a sanitization function $f_\theta(\cdot)$ with tunable parameter θ, and release the sanitized data $\mathbf{z}_i = f_\theta(\mathbf{x}_i)$ to the service provider. In this process, we target at finding a sanitation parameter θ to reduce the correlation between $(\mathbf{z}_i)_{i\geq 1}$ and $(\mathbf{s}_i)_{i\geq 1}$, which is quantified by a function $g(\cdot,\cdot)$ detailed in Section 3. Suppose a finite sequence of data $(\mathbf{x}_i)_{i=k}^{k+T}$ is associated with one type of driver behavior c_k from a set of prescribed classes. In the meantime, the sanitized data sequence $(\mathbf{z}_i)_{i\geq 1}$ is required to be still useful for driver behavior recognition. The utility is measured by a loss function $\ell(\cdot,\cdot)$ in terms of a driver behavior detection model represented by a function $q_\phi(\cdot)$ with ϕ being its trainable parameters. Our objective is to minimize the weighted sum of utility model loss and the level of privacy leakage:

$$\min_{\theta,\phi} \sum_{k\geq 1} \ell\left(c_k, q_\phi \circ f_\theta\left((\mathbf{x}_i)_{i=k}^{k+T}\right)\right) + \gamma g\left((\mathbf{s}_i)_{i=k}^{k+T}, f_\theta\left((\mathbf{x}_i)_{i=k}^{k+T}\right)\right), \quad (1)$$

where $\circ$ denotes function composition and γ is a constant controlling the trade-off between privacy and utility. The sanitizer f_θ aims to remove sensitive statistical information to preserve trajectory privacy.

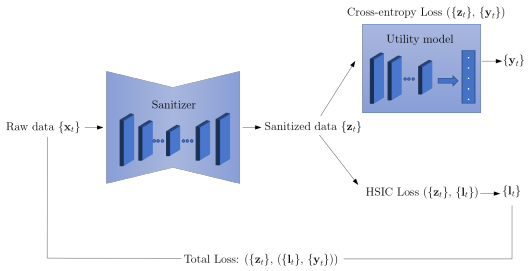

Fig. 1. Overall learning framework.

2.2. Utility model

Recurrent neural networks are capable of processing sequences of data. We build a driver behavior classification model using long short-term memory (LSTM) modules [19]. As shown in Fig. 2, our model consists of stacks of two LSTM units where $\mathbf{h}^1$ and $\mathbf{h}^2$ denote the hidden state vectors, also known as the output vectors of the two LSTM units, respectively, and $\mathbf{c}^1$ and $\mathbf{c}^2$ denote the cell state vectors of the two LSTM units, respectively. The decision is made at time $k+T$ by passing $\mathbf{h}_{k+T}^2$ to a multilayer perceptron (MLP) unit which is followed by a softmax activation function. The units with the same colour share their weights. The model input is a sequence of sanitized data $\mathbf{z}_i = f_\theta(\mathbf{x}_i)$ obtained from the sanitizer. The model output $\hat{\mathbf{y}}_k = q_\phi \circ f_\theta\left((\mathbf{z}_i)_{i=k}^{k+T}\right)$ is a vector indicating the likelihood of a driver's behavior belonging to each of the predefined classes. The model loss function ℓ is the cross-entropy loss defined as

$$\ell\left(c_k, q_\phi \circ f_\theta\left((\mathbf{x}_i)_{i=k}^{k+T}\right)\right) = -\log \hat{\mathbf{y}}_k(c_k),$$

where c_k is the driver behavior label and $\hat{\mathbf{y}}_k(c_k)$ is the predicted probability of driver behavior belonging to class c_k.

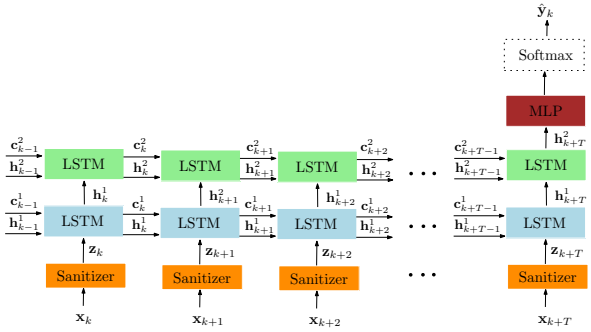

Fig. 2. Utility inference model.

3. PRIVACY METRIC: NORMALIZED HSIC

In this section, we present the Hilbert Schmidt Independence Criterion (HSIC) [20] and we propose a normalized HSIC that serves as the privacy function $g(\cdot,\cdot)$ in (1).

3.1. Normalized HSIC

We first give an overview of random elements in Hilbert spaces. More details can be found in [21]. Consider a Hilbert space $\mathcal{H}$ of functions from $\mathcal{X}$ to $\mathbb{R}$ with the inner product denoted as $\langle\cdot,\cdot\rangle_{\mathcal{H}}$. The Hilbert space $\mathcal{H}$ is a reproducing kernel Hilbert space (RKHS) if at each $x \in \mathcal{X}$, the point evaluation operator $L_x : f \mapsto f(x)$ for $f \in \mathcal{H}$ is a bounded linear functional, i.e., there exists a positive constant K such that $\|L_x(f)\|_{\mathcal{H}} \leq K\|f\|_{\mathcal{H}}$ for all $f \in \mathcal{H}$. RKHSs have the following reproducing property: for each $x \in \mathcal{X}$, there exists a unique element $\Phi_x \in \mathcal{H}$ such that

$$f(x) = \langle \Phi_x, f\rangle_{\mathcal{K}}, \forall f \in \mathcal{H}.$$

We call Φ_x a feature mapping of x in $\mathcal{H}$. Since $\Phi_x \in \mathcal{H}$, we have $\Phi_x(y) = \langle \Phi_x, \Phi_y \rangle_{\mathcal{H}}$ for any $y \in \mathcal{X}$. The kernel associated with $\mathcal{H}$ is a function $\mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}$ defined as

$$k(x, y) = \langle \Phi_x, \Phi_y \rangle_{\mathcal{H}},$$

for $x, y \in \mathcal{X}$.

Now we consider random elements $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$ that are jointly distributed according to some probability density function. Let $\mathcal{H}$ be a RKHS of functions $\mathcal{X} \mapsto \mathbb{R}$ and $\mathcal{L}$ be a RKHS of functions $\mathcal{Y} \mapsto \mathbb{R}$, with their respective kernels denoted as $h(\cdot, \cdot)$ and $l(\cdot, \cdot)$. Given $f \in \mathcal{H}$ and $g \in \mathcal{K}$, we are interested in computing the correlation between $f(X)$ and $g(Y)$. By defining a tensor product $f \otimes g : \mathcal{L} \mapsto \mathcal{H}$ as

$$(f \otimes g)h = f \langle g, h \rangle_{\mathcal{L}},$$

for $f \in \mathcal{H}$ and $g, h \in \mathcal{L}$, we have

$$\begin{aligned} &\operatorname{corr}(f(X), g(Y)) \\ &= \mathbb{E}[f(X)g(Y)] - \mathbb{E}[f(X)]\,\mathbb{E}[g(Y)] \\ &= \langle f, C_{XY} g \rangle_{\mathcal{H}}, \end{aligned}$$

where $C_{XY} = \mathbb{E}[\phi(X) \otimes \phi(Y)] - \mathbb{E}[\phi(X)] \otimes \mathbb{E}[\phi(Y)]$ is a linear operator from $\mathcal{L}$ to $\mathcal{H}$. Similarly, the variances of $f(X)$ and $g(Y)$ can be written as

$$\begin{aligned} \operatorname{var} f(X) &= \langle f, C_X f \rangle_{\mathcal{H}}, \\ \operatorname{var} g(Y) &= \langle g, C_Y g \rangle_{\mathcal{L}}, \end{aligned}$$

where C_X and C_Y are linear operators defined as

$$\begin{aligned} C_X &= \mathbb{E}[\phi(X) \otimes \phi(X)] - \mathbb{E}[\phi(X)] \otimes \mathbb{E}[\phi(X)], \\ C_Y &= \mathbb{E}[\phi(Y) \otimes \phi(Y)] - \mathbb{E}[\phi(Y)] \otimes \mathbb{E}[\phi(Y)]. \end{aligned}$$

Assuming $\mathcal{H}$ and $\mathcal{L}$ are both separable, i.e., they have complete orthonormal systems, the Hilbert-Schmidt norm of C_{XY} is defined to be

$$\|C_{XY}\|_{\mathrm{HS}}^2 \triangleq \sum_{i \geq 1, j \geq 1} \langle u_i, C_{XY} v_j \rangle_{\mathcal{H}}^2.$$

where $\{u_i\}_{i \geq 1}$ and $\{v_i\}_{i \geq 1}$ are arbitrary orthonormal bases of $\mathcal{H}$ and $\mathcal{L}$, respectively (it can be shown that the defintion is independent of the choice of orthonormal bases). Therefore, $\|C_{XY}\|_{\mathrm{HS}}^2$ can be interpreted as the sum of squared correlations with respect to (w.r.t.) the orthonormal functions in the Hilbert spaces. We propose the following normalized HSIC:

$$\overline{\mathrm{HSIC}}(X, Y) = \frac{\|C_{XY}\|_{\mathrm{HS}}^2}{\left\|C_X^{1/2}\right\|_{\mathrm{HS}}^2 \left\|C_Y^{1/2}\right\|_{\mathrm{HS}}^2}, \tag{2}$$

where $C_X^{1/2}$ denotes the square root of the linear operator C_X. Next we show the normalized HSIC (2) is closely related to the maximal correlation [22] defined as

$$\begin{aligned} \rho(X, Y) = \sup_{f \in \mathcal{H}, g \in \mathcal{L}} & \mathbb{E}[f(X)g(Y)] \\ \text{s.t. } & \mathbb{E}[f(X)] = 0, \mathbb{E}[g(Y)] = 0, \\ & \mathbb{E}\left[f^2(X)\right] = 1, \mathbb{E}\left[g^2(Y)\right] = 1. \end{aligned}$$

Any $f \in \mathcal{H}$ and $g \in \mathcal{L}$ can be expressed as

$$f = \sum_{i \geq 1} \alpha_i u_i \text{ and } g = \sum_{i \geq 1} \beta_i v_i,$$

where $\alpha_i = \langle f, u_i \rangle_{\mathcal{H}}$ and $\beta_i = \langle g, v_i \rangle_{\mathcal{L}}$. Thus $\rho^2(X, Y)$ can be written as

$$\begin{aligned} \rho^2(X, Y) &= \sup_{f \in \mathcal{H}, g \in \mathcal{K}} \frac{\operatorname{corr}(f(X), g(Y))^2}{\operatorname{var} f(X) \operatorname{var} g(Y)} \\ &= \sup_{\substack{\{\alpha_i\}_{i \geq 1} \\ \{\beta_j\}_{j \geq 1}}} \frac{\left(\sum_{i \geq 1, j \geq 1} \alpha_i \beta_j \langle u_i, C_{XY} v_j \rangle_{\mathcal{H}}\right)^2}{\sum_{i \geq 1} \alpha_i^2 \langle u_i, C_X u_i \rangle_{\mathcal{H}} \sum_{j \geq 1} \beta_j^2 \langle C_Y v_j, v_j \rangle_{\mathcal{L}}}. \end{aligned}$$

The normalized HSIC in (2) is then given by

$$\lim_{m,n \to \infty} \frac{\frac{1}{m^2 n^2} \sum_{i \geq 1, j \geq 1} \langle u_i, C_{XY} v_j \rangle_{\mathcal{H}}^2}{\frac{1}{m^2} \sum_{i \geq 1} \langle u_i, C_X u_i \rangle_{\mathcal{H}} \frac{1}{n^2} \sum_{j \geq 1} \langle v_j, C_Y v_j \rangle_{\mathcal{L}}}.$$

From the above two expressions, it can be concluded that

$$0 \leq \overline{\mathrm{HSIC}}(X, Y) \leq \rho^2(X, Y) \leq 1.$$

From $\|\phi(x) \otimes \phi(y)\|_{\mathrm{HS}} = \|\phi(x)\|_{\mathcal{H}} \|\phi(y)\|_{\mathcal{L}}$, we have

$$\begin{aligned} \|C_{XY}\|_{\mathrm{HS}}^2 &= \mathbb{E}_{X,X',Y,Y'}[K(X, X')K(Y, Y')] \\ &\quad + \mathbb{E}_{X,X'}[K(X, X')]\,\mathbb{E}_{Y,Y'}[K(Y, Y')] \\ &\quad - 2\mathbb{E}_{X,Y,}\left[\mathbb{E}_{X'}[K(X, X')]\,\mathbb{E}_{Y'}[K(Y, Y')]\right], \\ \left\|C_X^{1/2}\right\|_{\mathrm{HS}}^2 &= \mathbb{E}[K(X, X)] - \mathbb{E}[K(X, X')], \\ \left\|C_Y^{1/2}\right\|_{\mathrm{HS}}^2 &= \mathbb{E}[K(Y, Y)] - \mathbb{E}[K(Y, Y')], \end{aligned} \tag{3}$$

where X' and Y' are copies of X and Y, respectively.

3.2. Privacy metric

For a given set of data $(\mathbf{z}_i)_{i=1}^T$ and $(\mathbf{s}_i)_{i=1}^T$ with $\mathbf{z}_i = f_\theta(\mathbf{x}_i)$, we let $\mathbf{K}$ and $\mathbf{L}$ be $T \times T$ matrices with the entry at the ith column and jth row being $[\mathbf{K}]_{i,j} = h(\mathbf{z}_i, \mathbf{z}_j)$ and $[\mathbf{L}]_{i,j} = l(\mathbf{s}_i, \mathbf{s}_j)$, respectively. Replacing the expectations with sample averages in (3) and substituting them into (2), we obtain the estimate of the normalized HSIC and use it as the privacy function:

$$g\left((\mathbf{s}_i)_{i=1}^T, f_\theta((\mathbf{x}_i)_{i=1}^T)\right) = \frac{\operatorname{Tr}(\mathbf{KHLH})}{\operatorname{Tr}(\mathbf{KH})\operatorname{Tr}(\mathbf{LH})}, \tag{4}$$

where $\mathbf{H} = \mathbf{I} - T^{-1}\mathbf{1}\mathbf{1}^\intercal$ is a centering matrix and $\operatorname{Tr}(\cdot)$ denotes the matrix trace operator. When the kernels are differentiable w.r.t. to their inputs, this privacy function is differentiable w.r.t. the sanitization parameter θ.

4. EXPERIMENTS

In this section, we present experimental results on two real-world datasets in terms of the utility-privacy trade-offs between trajectory inference and driver behavior recognition.

The UAH-DriveSet [23] and Honda Driver Behavior (HDB) dataset [24] are used for our experiments. The vehicular sensor data $\mathbf{x}_i$ includes speed, course, accelerations in three directions, roll, pitch and yaw. Each finite sequence of data $(\mathbf{x}_i)_{i=k}^{k+T}$ is labeled with one of three types of driver behaviors, namely normal, drowsy and aggressive for UAH-DriveSet, and aggressive action, non-aggressive action and non-action for HDB dataset. The GPS measurements in UAH-DriveSet serve as the ground truth. We use the estimated trajectory as reference in HDB dataset.

The sanitizer is an auto-encoder with two dense layers with 7 units and 8 units, respectively. The utility model is described in Section 2, where a time-distributed layer with 16 neurons is used at first, followed by LSTM layers. Sanitized data is compressed into lower dimension before feeding into the Gaussian kernel with bandwidth 1 for computing the normalized HSIC (4). This forms a composition kernel. We set $T = 64$ and batch size for training to be 1000. We use Adam optimizer with learning rate 0.0001. Data are normalized into a range of 0 to 1.

We compare our N-HSIC method with the state-of-the-art differential privacy (DP) [4], GAP [17] and CPGAN [18] in terms of utility-privacy trade-offs. We use the Laplace mechanism for DP. GAP and CPGAN are based on adversarial training. The privacy function for GAP is the mean square error (MSE-L) from the adversarial trajectory inference model, while the privacy function for CPGAN is the mean square reconstruction error of the sanitized data (MSE-R). The model loss of driver behavior recognition is used for measuring utility.

Table 1. Utility-privacy trade-offs for N-HSIC, DP, GAP and CPGAN ($\gamma = 1$).

Model	Classifying Acc.(%)	MSE-R	MSE-L
UAH-DriveSet			
DP	41.80	0.0782	0.0581
GAP	91.80	0.0454	0.0662
CPGAN	82.50	0.0598	0.0817
N-HSIC	**96.10**	**0.1157**	**0.1088**
HDB dataset			
DP	61.30	**0.0192**	0.0352
GAP	**96.90**	0.0075	0.0197
CPGAN	88.80	0.0055	0.0233
N-HSIC	96.60	0.0022	**0.0589**

From Table 1, it is observed that our proposed N-HSIC method achieves better or comparable utility compared to GAP, while enjoying better location privacy implied by MSE-L. It should be highlighted that the same adversarial inference model is used for training and testing GAP, while N-HSIC does not include adversarial training. This demonstrates that the normalized HSIC is a superior privacy leakage measure. CPGAN protects the raw data well at the cost of less utility. DP adds excessive amount of noise to the raw data and hence greatly impairs utility.

In Fig. 3, a vehicle travels along a highway, makes a u-turn at the end of the highway, and then returns to its starting location (route in red color). We use the extended Kalman filter and map-matching to estimate the vehicle's trajectory from the sanitized data as well as from the raw data. The trajectory estimated from the raw data (in black color) is close to the ground-truth route, while the sanitized trajectory largely deviates from the ground-truth. This demonstrates that our method protects the trajectory privacy well even if an adversary has the map information. The degree of deviation is higher when the trade-off parameter γ is larger. The driver behavior recognition accuracy from the sanitized data is 89.1% ($\gamma = 20$) and 86.8% ($\gamma = 40$), which are maintained at a high level.

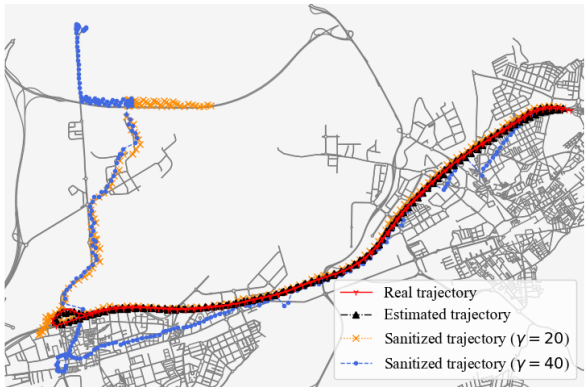

Fig. 3. Trajectory inference on UAH-DriveSet.

5. CONCLUSION

In this paper, we have presented a privacy-preserving framework based on the normalized HSIC as a privacy metric. We have developed a deep learning architecture to learn the sanitizer in the context of protecting trajectory information in ITS applications, while maintaining the utility of driver behavior recognition. Numerical experiments on two real datasets demonstrates that our approach achieves better or comparable utility-privacy trade-off.

6. REFERENCES

[1] J. Zhang, F. Y. Wang, K. Wang, W. H. Lin, X. Xu, and C. Chen, "Data-driven intelligent transportation systems: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 1624–1639, Dec. 2011.

[2] X. Pan, J. Xu, and X. Meng, "Protecting location privacy against location-dependent attacks in mobile services," *IEEE Trans. Knowl. Data Eng.*, vol. 24, no. 8, pp. 1506–1519, Aug. 2012.

[3] D. Rebollo-Monedero, J. Forné, and J. Domingo-Ferrer, "From t-closeness-like privacy to postrandomization via information theory," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 11, pp. 1623–1636, Nov. 2010.

[4] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Foundations and Trends® in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, Aug. 2014.

[5] M. Sun and W. P. Tay, "On the relationship between inference and data privacy in decentralized IoT networks," *IEEE Trans. Inf. Forensics Security*, vol. 15, no. 1, pp. 852–866, Dec. 2020.

[6] J. L. Yin and B. H. Chen, "An advanced driver risk measurement system for usage-based insurance on big driving data," *IEEE Trans. Intell. Veh.*, vol. 3, no. 4, pp. 585–594, Dec. 2018.

[7] F. P. Calmon and N. Fawaz, "Privacy against statistical inference," in *Proc. Allerton Conf. on Commun., Control and Computing*, Monticello, IL, Oct. 2012.

[8] L. Sankar, S. R. Rajagopalan, and H. V. Poor, "Utility-Privacy tradeoffs in databases: An information-theoretic approach," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 6, pp. 838–852, Jun. 2013.

[9] E. Erdemir, P. L. Dragotti, and D. Gündüz, "Privacy-Aware time-series data sharing with deep reinforcement learning," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 389–401, Jul. 2020.

[10] M. Sun, W. P. Tay, and X. He, "Toward information privacy for the Internet of things: A nonparametric learning approach," *IEEE Trans. Signal Process.*, vol. 66, no. 7, pp. 1734–1747, Apr. 2018.

[11] C. X. Wang and W. P. Tay, "Data-Driven privacy with domain regularization," in *Proc. IEEE Global Commun. Conf.*, Taipei, Taiwan, Dec. 2020.

[12] C. X. Wang, Y. Song, and W. P. Tay, "Arbitrarily strong utility-privacy tradeoff in multi-agent systems," *IEEE Trans. Inf. Forensics Security*, pp. 1–1, 2020.

[13] S. Y. Kung, "Compressive privacy: From information/estimation theory to machine learning," *IEEE Signal Process. Mag.*, vol. 34, no. 1, pp. 94–112, Jan. 2017.

[14] Y. Song, C. X. Wang, and W. P. Tay, "Privacy-Aware Kalman filtering," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Calgary, Canada, Apr. 2018.

[15] ——, "Compressive privacy for a linear dynamical system," *IEEE Trans. Inf. Forensics Security*, vol. 15, no. 1, pp. 895–910, Dec. 2020.

[16] I. J. Goodfellow, J. P. Abadie, M. Mirza, B. Xu, D. W. Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Int. Conf. on Neural Information Processing Systems*, Montreal, Canada, Dec. 2014.

[17] C. Huang, P. Kairouz, and L. Sankar, "Generative adversarial privacy: A data-driven approach to information-theoretic privacy," in *Proc. Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, USA, USA, Oct. 2018.

[18] B. O. Tseng and P. Y. Wu, "Compressive privacy generative adversarial network," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 2499–2513, Jan. 2020.

[19] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[20] A. Gretton, O. Bousquet, A. Smola, and B. Schölkopf, "Measuring statistical dependence with hilbert-schmidt norms," in *Proc. Int. Conf. Algorithmic Learning Theory*, Singapore, Oct. 2005.

[21] K. Muandet, K. Fukumizu, B. Sriperumbudur, and B. Schölkopf, *Kernel Mean Embedding of Distributions: A Review and Beyond*. Now Foundations and Trends, 2017.

[22] A. Rényi, "On measures of dependence," *Acta Math. Hung.*, vol. 10, pp. 441–451, Sep. 1959.

[23] E. Romera, L. M. Bergasa, and R. Arroyo, "Need data for driver behaviour analysis? presenting the public uah-driveset," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2016, pp. 387–392.

[24] J. Ferreira, E. Carvalho, B. V. Ferreira, C. de Souza, Y. Suhara, A. Pentland, and G. Pessin, "Driver behavior profiling: An investigation with different smartphone sensors and machine learning," *PLoS one*, vol. 12, no. 4, p. e0174959, 2017.