Protecting Sensory Data against Sensitive Inferences

Mohammad Malekzadeh
Queen Mary University of London
London, UK
m.malekzadeh@qmul.ac.uk

Richard G. Clegg
Queen Mary University of London
London, UK
r.clegg@qmul.ac.uk

Andrea Cavallaro
Queen Mary University of London
London, UK
a.cavallaro@qmul.ac.uk

Hamed Haddadi
Imperial College London
London, UK
h.haddadi@imperial.ac.uk

ABSTRACT

There is growing concern about how personal data are used when users grant applications direct access to the sensors of their mobile devices. In fact, high resolution temporal data generated by motion sensors reflect directly the activities of a user and indirectly physical and demographic attributes. In this paper, we propose a feature learning architecture for mobile devices that provides flexible and negotiable privacy-preserving sensor data transmission by appropriately transforming raw sensor data. The objective is to move from the current binary setting of granting or not permission to an application, toward a model that allows users to grant each application permission over a limited range of inferences according to the provided services. The internal structure of each component of the proposed architecture can be flexibly changed and the trade-off between privacy and utility can be negotiated between the constraints of the user and the underlying application. We validated the proposed architecture in an activity recognition application using two real-world datasets, with the objective of recognizing an activity without disclosing gender as an example of private information. Results show that the proposed framework maintains the usefulness of the transformed data for activity recognition, with an average loss of only around three percentage points, while reducing the possibility of gender classification to around 50%, the target random guess, from more than 90% when using raw sensor data. We also present and distribute MotionSense, a new dataset for activity and attribute recognition collected from motion sensors.

CCS CONCEPTS

• **Security and privacy**; • **Computing methodologies** → **Machine learning**; **Distributed computing methodologies**;

W-P2DS'18, April 23–26, 2018, Porto, Portugal

ACM ISBN 978-1-4503-5654-1/18/04…$15.00
https://doi.org/10.1145/3195258.3195260

KEYWORDS

Privacy, Sensor Data, Activity Recognition, Machine Learning, Time-Series Analysis

ACM Reference Format:
Mohammad Malekzadeh, Richard G. Clegg, Andrea Cavallaro, and Hamed Haddadi. 2018. Protecting Sensory Data against Sensitive Inferences. In *W-P2DS'18: 1st Workshop on Privacy by Design in Distributed Systems , April 23–26, 2018, Porto, Portugal.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3195258.3195260

1 INTRODUCTION

Smartphones and wearable devices are equipped with sensors such as accelerometers, gyroscope, barometer and light sensors that are directly accessed by applications (apps) to provide through a cloud service analysis and statistics about, for example, the activities of the user. However, by granting to these apps access to raw sensor data, users may unintentionally reveal information about gender, mood, personality, which is unnecessary for the specific services.

To address this problem, we introduce the Guardian-Estimator-Neutralizer (GEN) framework that, instead of granting apps direct access to sensors, is designed to share only a transformed version of the sensor data, based on the functions and requirements of each application and privacy considerations. The *Guardian* provides an inference-specific transformation, the *Estimator* guides the Guardian by estimating sensitive and non-sensitive information in the transformed data, and the *Neutralizer* is an optimizer that helps the Guardian converge to a near-optimal transformation function (see Figure 1).

Unlike privacy-preserving works that only hide users' identity by sharing population data using generative models for data synthesis [2, 9], our solution concerns sensitive information included in a single user's data. There are, however, some methods which transform only selected temporal sections of sensor data that correspond to predefined sensitive activities [11, 12], our framework enables concurrently eliminating private information from each section of data, while keeping the utility of shared data.

GEN is a feature learning and data reconstruction framework that helps to efficiently establish a trade-off between apps utility and user privacy. Specifically, in this paper, we instantiate the framework for an activity recognition application based on data recorded by the accelerometer and gyroscope of a smartphone. In the context of this application, we categorize information that can be inferred from sensor data into two types: information about a predefined set

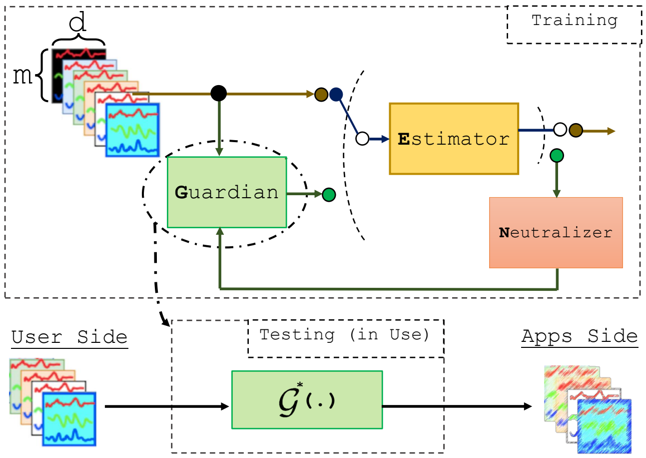

Figure 1: GEN Architecture: First, the Estimator is trained; then the Guardian is trained using the Estimator with the help of the Neutralizer.

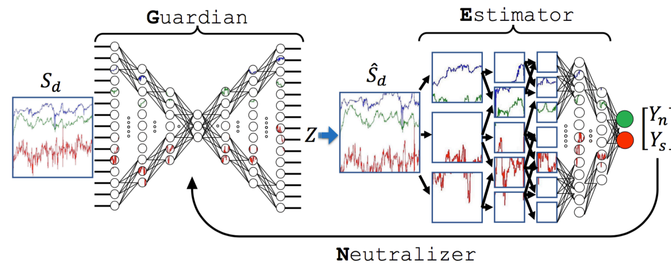

Figure 2: An instantiation of GEN for activity recognition from sensor data without revealing the gender information. The Guardian is an *autoencoder*. The Estimator is a *multi-task ConvNet*.

of activities of the user (*non-sensitive inferences*) and information about attributes of the user such as gender, age, weight and height (*sensitive inferences*).

Our goal is to establish a tradeoff between the ability of the apps to accurately infer non-sensitive information to maximize their *utility* and the reduction of revealed sensitive information to minimize the risk of *privacy* infringement. We show that GEN can accurately maintain the usefulness of the released (transformed) data for activity recognition while considerably reducing the risk of attribute recognition.[1]

2 PROBLEM DEFINITION

Let $X(t) = \Big(X_1(t), X_2(t), \ldots, X_m(t)\Big)$ be the recorded values of the m sensor-data components during a collection period of duration T, where $t \in \{1, 2, \ldots, T\}$. We assume the data to be synchronized and collected at the same frequency.

Let us consider a running window of duration d that contains consecutive values of $X(t)$ from time t to $t + d - 1$. Let $S_d(t)$ be the corresponding section of the time-series:

$$S_d(t) = X[t, t + d - 1] = \Big(X(t), X(t + 1), \ldots, X(t + d - 1)\Big),$$

where the value of d should be chosen such that the running window be large enough for making desired inferences by apps. However, in order to be computationally effective, it should not be chosen very large. For simplicity, we remove the index t, from $S_d(t)$, in the following.

We define two types of inference on each S_d: inference of **s**ensitive information, $I_{\mathbf{s}}(.)$, and inference of **n**on-sensitive information, $I_{\mathbf{n}}(.)$. Our goal is to find a transformation function, $\mathcal{G}^*(.)$, in a way that the transformed data $\hat{S}^*_d = \mathcal{G}^*(S_d)$ are such that $I_{\mathbf{s}}(\hat{S}^*_d)$ fails to reveal private information, whereas $I_{\mathbf{n}}(\hat{S}^*_d)$ generates inferences that are as accurate as $I_{\mathbf{n}}(S^*_d)$. Here, $\hat{S}_d$ is the transformation of corresponding S_d, and $\hat{S}^*_d$ is its optimal privacy-preserving transformation.

[1]The code and data used in this paper are publicly available at: https://github.com/mmalekzadeh/motion-sense

3 LEARNING THE INFERENCE-SPECIFIC TRANSFORMATION

We present the proposed framework that includes three components: the Guardian, the Estimator, and the Neutralizer (Figure 1), and discuss its instantiation for an activity recognition application (Figure 2).

The **Guardian**, which provides *inference-specific transformation*, is a feature learning framework that recognizes and distinguishes discerning features from data. In the specific implementation of this paper, we use a *deep autoencoder* [16] as Guardian. An autoencoder is a neural network that tries to reconstruct its input based on an objective function. Here, the autoencoder receives a section of *m-dimensional* time-series with length of d as input, and produces a time-series with the same dimensionality as the output; based on the Neutralizer's objective function, which is described below.

The **Estimator** quantifies how accurate an algorithm can be at making sensitive and non-sensitive inferences on the transformed data. In the specific implementation of this paper, we use a *multi-task convolutional neural network (MTCNN)* as Estimator [17]. The shape of input is similar to the Guardian and the shape of output depends on the number of activity classes. MTCNN has the ability to share learned representations from input between several tasks. More precisely, we try to simultaneously optimize a CNN with two types of loss function, one for sensitive inferences and another for non-sensitive ones. Consequently, MTCNN will learn more generic features, which should be used for several tasks, at its earlier layers. Then, subsequent layers, which become progressively more specific to the details of the desired task, can be divided into multiple branches, each for a specific task.

The **Neutralizer**, the most important contribution of this paper, is an optimizer that helps the Guardian find the optimal $\mathcal{G}^*(\cdot)$ for transforming each section S_d into $\hat{S}^*_d$ using as objective

$$\mathcal{G}^*(.) = \underset{\mathcal{G}(.) \in \mathcal{F}}{\text{argmin}} \Bigg(p\Big(I_{\mathbf{s}}\big(\mathcal{G}(S_d)\big)\Big) - p\Big(I_{\mathbf{n}}\big(\mathcal{G}(S_d)\big)\Big) \Bigg),$$

	MobiAct	MotionSense
#Males	32	16
#Females	14	10
#Features (m)	9	12
Sample Rate (Hz)	20	50

Table 1: Details of the MobiAct and MotionSense datasets.

where $p\Big(I_s(\cdot)\Big)$ and $p\Big(I_n(\cdot)\Big)$ are the probabilities of making sensitive and non-sensitive inferences, respectively, and the $\mathcal{F}$ is the set of all possible transformation functions for the Guardian. In the specific application of this paper the Neutralizer is a multi-task objective function used by backpropagation to update the weights of the Guardian (autoencoder). The $\mathcal{F}$ is also the set of all possible weight matrices for the selected autoencoder.

Particularly, we aim to transform each section S_d such that we can recognize an activity from $\hat{S_d}$ without revealing the gender of the user. For each section S_d, let $Y_{\mathbf{a}}(S_d)$ and $Y_{\mathbf{a}}(\hat{S_d})$ be the true and predicted class of activity, respectively, and $Y_{\mathbf{g}}(\hat{S}_d)$ be the predicted gender class. We define the Neutralizer's objective function as

$$\hat{S}_d^* = \underset{\hat{S}_d}{\operatorname{argmin}} \Bigg(\sum_{i=1}^{c} Y_{\mathbf{a}}^i(S_d) \log Y_{\mathbf{a}}^i(\hat{S}_d) + \mid (0.5 - Y_{\mathbf{g}}(\hat{S}_d)) \mid \Bigg), \quad (1)$$

where c is the number of activity classes. In the r.h.s. of the equation, the first part is a categorical cross entropy and the second part is our custom gender-neutralizer loss function. The constant 0.5 is the desired confidence for a gender predictor that will process the transformed data.

4 EXPERIMENTS

We validate the proposed framework on recognizing the following activities from smartphone motion sensors: *Downstairs, Upstairs, Walking, Jogging*. The non-sensitive inferences, $I_{\mathbf{n}}$, is the recognition of the activities, whereas the sensitive inference, $I_{\mathbf{s}}$, is the recognition of gender.

We aim to measure the trade-off between the utility of data for activity recognition and privacy, e.g. keeping gender secret. To this end, we first compare the accuracy of activity recognition and gender classification when a trained MTCNN has access to original data and to the corresponding transformed data. Then we try to measure the amount of sensitive information which is still available in the transformed data using different methods.

Model	Layer (Neurons | Kernel | Chance)
MTCNN	Inp(m, d) Conv(50: 1×5); Conv(50: 1×3) Dense(50); MP(1×2); DO(0.2) Conv(40: 1×5) Dense(40); MP(1×3); DO(0.2) Conv(20: 1×3); DO(0.2) Flatten; Dense(400); DO(0.4) OutA = Softmax(4); OutG = Sigmoid
AE	Inp($\lvert x\rvert$); Dense($\lvert x\rvert/2$); Dense($\lvert x\rvert/4$) Dense($\lvert x\rvert/8$) Dense($\lvert x\rvert/4$); Dense($\lvert x\rvert/2$); Out($\lvert x\rvert$)

Table 2: Structure of the hidden layers. The activation function for all the layers is "ReLU". Key – MP: MaxPooling; DO: DropOut; $|x| = m \times d$.

4.1 Datasets

We use two real-world datasets: MobiAct[2] and MotionSense[3]. The latter dataset is one of the contributions of this paper.

MobiAct [15] includes accelerometer, gyroscope and orientation data ($m = 9$) from a smartphone collected when data subjects performed 9 activities in 16 trials. A total of 67 participants in a range of gender, age, weight, and height collected the data with a Samsung Galaxy S3 smartphone (we use a subset of 48 subjects who have no missing data). Unlike other datasets, which require the smartphone to be rigidly placed on the human body and with a specific orientation, MobiAct attempted to simulate every-day usage of mobile phones where a smartphone is located with random orientation in a loose pocket chosen by the subject (Table 1).

MotionSense includes the accelerometer (acceleration and gravity), attitude (pitch, roll, yaw) and gyroscope data ($m = 12$) collected with an iPhone 6s kept in the participant's front pocket using SensingKit [10]. A total of 24 participants in a range of gender, age, weight, and height performed 6 activities in 15 trials in the same environment and conditions: downstairs, upstairs, walking, jogging, sitting, and standing. With this dataset, we aim to look for *personal attributes fingerprints* in time-series of sensor data, i.e. attribute-specific patterns that can be used to infer physical and demographic attributes of the data subjects in addition to their activities.

See http:github.com/mmalekzadeh/motion-sense for details on the methodology and the data (Table 1).

[2]publicly available at:
http://www.bmi.teicrete.gr/index.php/research/mobiact
[3]publicly available at:
http://github.com/mmalekzadeh/motion-sense

Setting	Dataset	Inf.	S_d	$\hat{S}_d$
Trial	**MotionSense**	$I_{\mathbf{a}}$	95.08	93.71
		$I_{\mathbf{g}}$	95.15	49.32
	MobiAct	$I_{\mathbf{a}}$	94.31	90.46
		$I_{\mathbf{g}}$	93.74	49.83
Subject	**MotionSense**	$I_{\mathbf{a}}$	86.33	85.19
		$I_{\mathbf{g}}$	75.35	52.16
	MobiAct	$I_{\mathbf{a}}$	70.49	65.01
		$I_{\mathbf{g}}$	66.18	45.54

Table 3: Activity recognition, $I_{\mathbf{a}}$, and gender classification, $I_{\mathbf{g}}$, accuracy for *original*, S_d, and *transformed*, $\hat{S}_d$, data in percent (%).

4.2 Experimental Setup

For each dataset, we consider two types of setting, namely Trial and Subject. In *Trial*, we keep 2/3 of trials for training and 1/3 of them for testing. For example, if there are 3 walking trials per participant, we keep the first two trials for training and the last one for testing. In *Subject* we keep data of 75% of all subjects for training and the data of remaining 25% subjects for testing. In the Subject setting, we report the average results of four selections for test dataset.

We train an MTCNN as the Estimator by considering two tasks: (i) activity recognition (4 classes) with *categorical cross-entropy* loss function [4], and (ii) gender classification (2 classes) with *binary cross-entropy* loss function. [4]. After training MTCNN, we freeze the weights of the MTCNN layers and attach the output of a deep autoencoder (AE) as the Guardian to the input of the MTCNN to build the GEN neural network. Finally, we compile GEN and set its loss function equals to the objective function of the Neutralizer in Equation (1). The deep network architectures are described in Table 2.

4.3 Transformation Efficiency

Table 3 shows that the Guardian produces time-series that keep the utility of non-sensitive inferences at a comparable level to the original ones (the average loss is three percentage points) while preventing sensitive inferences, as the gender classification accuracy decreases from more than 90% to near the target random guess (50%).

Cross-Dataset Validation. We also validate GEN in an ecosystem where edge users benefit from pre-trained models of a service provider. At the *cloud side* the Estimator (MTCNN) is trained on a public dataset, the MobiAct dataset in our case. At the *edge side*, the Guardian receives the trained Estimator and uses its locally (personally) defined Neutralizer to transform the user's data, the MotionSense dataset in our case.

The results show that the accuracy of the Estimator on raw data for I_a and I_g are 93.67% and 92.80%, respectively; whereas on transformed data are 90.92% and 51.93%, respectively. This shows an interesting property of GEN which makes it more applicable to deploy in edge devices.

The only concern here is whether users trust the pre-trained Estimator received from an untrusted service provider. User can verify the Estimator by running it on a publicly available dataset. We leave more investigation on this concern for future work.

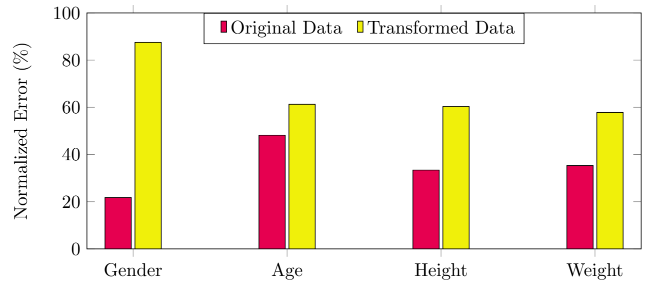

Figure 3: Error for gender is "classification error" and for the rest of attributes is "mean absolute error". All the values are divided by the error of a random estimator on the MotionSense dataset.

4.4 Measuring Information Leakage

We aim to experimentally quantify the amount of information about user's attributes that is still available in the transformed data.

Using Dynamic Time Warping. To measure the amount of residual attribute-information in sensor data, we chose[4] k-Nearest Neighbors (k-NN) with Dynamic Time Warping (DTW) [13]. We aim to verify whether a different algorithm will also fail to guess gender, even when adversaries get access to the entire time-series, and not just a section of it. To this end we build an $n \times n$ matrix D_l, where n is the number of subjects in the dataset. For each activity $a_l \in \{downstairs, upstairs, walking, jogging\}$, let $d_l(i, j)$ be the distance between the time-series of users u_i and u_j calculated by FastDTW [13]. Then, we calculate the final *distance* matrix D as the element-wise average of all the matrices D_l; $d(i, j) = \frac{1}{4}\sum_l d_l(i, j)$.

We calculate distance matrices D and $\hat{D}$ for the original time-series and the transformed series (the output of the Guardian) respectively. Then we *compare* the ability of the estimation based on these matrices. For each user u_i; $i \in \{1, \ldots, n\}$ (one out-of-sample), we *estimate* the value of each attribute $v_a(u_i)$; $a \in \{gender, age, weight, height\}$, using distance weighted k-NN based on matrix D, where the weight is:

$$w(i, j) = \frac{1}{d(i, j)^2}.$$

Figure 3 shows that the estimation error for gender classification approaches that of a random estimator after transformation. In this Figure, the error of a random estimator for gender is $\frac{N_f}{N_f + N_m} = \frac{10}{24}$

[4] k-NN with DTW outperforms other methods in time-series classification, except when considerable computation and implementation cost is acceptable for very small improvements [1].

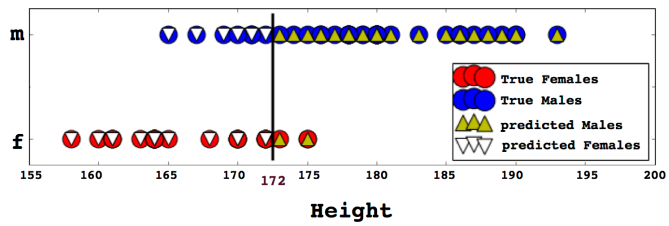

Figure 4: Dependencies between height and gender on the MotionSense and MobiAct datasets. A classification threshold of 172cm predicts gender with 84% accuracy.

and for the rest of attributes is considered as the half of the variation interval in dataset; e.g. $\frac{190-161}{2} = 14.5$ for height.

Thus the GEN eliminates similarities between same-gender time-series and an attacker cannot confidently use distance measures to make inference about gender. Interestingly, by eliminating gender information, we also partially eliminate information on other attributes, as there are dependencies between attributes. For example, the estimation error for height and weight increases by near 25% and 20%, respectively.

Height is indeed highly correlated with gender in both datasets (Figure 4): the prediction accuracy of gender-based on height only is 81%. However, gender prediction from both datasets using the MTCNN architecture is considerably better than that.

Using Supervised Learning. We explore learning gender discriminative features from *transformed* data. Figure 5 shows the training and validation accuracy of activity recognition and gender classification using supervised learning on transformed data. Gender-discriminative features in the transformed data are rare, even with a large number of epochs as in this experiment. GEN eliminates gender-related features and thus makes it is difficult for a classifier to train on them even when it has access to the labels of transformed data.

Although, with experiments in this section, we have shown an acceptable efficiency in eliminating sensitive information, it is highly desired to statistically prove the efficiency of the proposed solution. Generally, high temporal granularity of time-series and strong correlation between their samples make this task very challenging. We leave exploring this area to future research.

5 RELATED WORK AND DISCUSSION

Generative adversarial networks (GANs) [7] learn to capture the statistical distribution of data for synthesizing new samples from the learned distribution. In the GANs a discriminator model learns to determine whether a sample is from the model distribution (i.e. from the generator) or from the data distribution (i.e. from a real-world source). The discriminator aims to maximize an objective function in minimax game that the generator aims to minimize. GANs have also been applied for enhancing privacy [9, 14]. For example, to protect health records, synthetic medical datasets can be published instead of the real ones using generative models training on sensitive real-world medical datasets [3, 6]. To provide a formal privacy

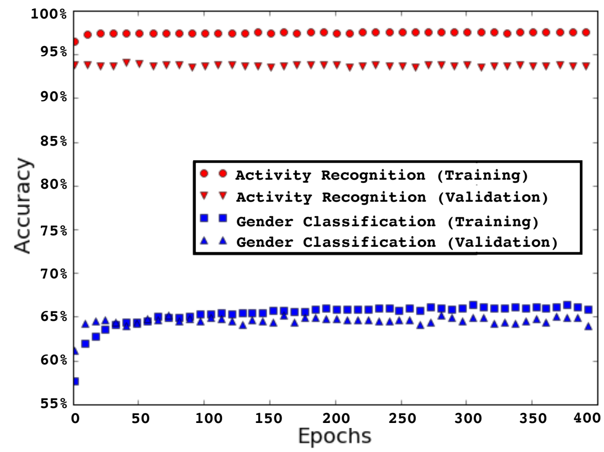

Figure 5: Activity and gender classification accuracy, on the MotionSense dataset in Trial setting, when the Estimator is trained on transformed data produced by the Guardian. Although activity-features can be easily learned, there is no useful discerning information about gender.

guarantee, [2] trains GANs under the constraint of differential privacy [5] to protect against common privacy attacks.

Although the architecture of our proposed framework looks similar to GANs, there are key structural and logical differences with other existing frameworks. First, the focus of existing works is mainly on protecting users' privacy against membership attack by releasing a synthetic dataset through differential privacy constraints. Instead, we consider a situation where a user wants to grant third parties access to sensor data that can be used to make both sensitive and non-sensitive inferences.

Second, the generator in GANs seeks to learn the underlying distribution of the data to produce realistic simulated samples from random vectors. Instead, the Guardian in GEN seeks to partition the underlying features of the data to reconstruct privacy-preserving outputs from real-world input vectors.

Finally, the minimax game in GANs is a two-player game between generator and discriminator (i.e. two models) that updates weights of both models in each iteration. Instead the minimax objective of GEN is a trade-off between utility and privacy that updates the weights of one only model (i.e. the guardian) in each iteration.

Previous works on data collected from embedded sensors of personal devices, such as [11, 12], consider temporal inferences on different activities over time (i.e. some sections of time-series corresponding to non-sensitive activities and some of them to sensitive ones). In this paper, for the first time, we concurrently consider both activity and attribute inferences on the same section of time-series.

Our framework is applicable in distributed environments: we have shown that the Estimator can be trained remotely (e.g. on a powerful system and with a large dataset) and edge devices just need to download the resulting trained model to use it as the Estimator part of their locally implemented GEN under user's control. For example, the Guardian can be trained in user side using individuals' personal data processing platforms, like Databox [8].

6 CONCLUSION

We proposed the GEN framework for locally transforming sensor data on mobile edge devices to respect functions and requirements of an application as well as user privacy. We evaluated the efficiency of the trade-off between utility and privacy GEN provides on real-world datasets of motion data.

Open questions to be explored in future work include providing theoretical bounds on the amount of sensitive information leakage after transformation and exploring dependencies between different attributes, e.g. co-dependence of gender and height. Finally, we will measure the costs and requirements for running GEN on edge devices.

ACKNOWLEDGMENTS

This work was kindly supported by the Life Sciences Initiative at Queen Mary University London and a Microsoft Azure for Research Award. Hamed Haddadi was partially funded by EPSRC Databox grant (Ref: EP/N028260/1).

REFERENCES

[1] A. Bagnall, J. Lines, A. Bostrom, J. Large, and E. Keogh. The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data Mining and Knowledge Discovery*, 31(3):606–660, 2017.

[2] B. K. Beaulieu-Jones, Z. S. Wu, C. Williams, and C. S. Greene. Privacy-preserving generative deep neural networks support clinical data sharing. *bioRxiv*, page 159756, 2017.

[3] E. Choi, S. Biswal, B. Malin, J. Duke, W. F. Stewart, and J. Sun. Generating multi-label discrete electronic health records using generative adversarial networks. *arXiv preprint arXiv:1703.06490*, 2017.

[4] F. Chollet et al. Keras. https://github.com/fchollet/keras, 2015.

[5] C. Dwork. Differential privacy: A survey of results. In *International Conference on Theory and Applications of Models of Computation*, pages 1–19. Springer, 2008.

[6] C. Esteban, S. L. Hyland, and G. Rätsch. Real-valued (medical) time series generation with recurrent conditional gans. *arXiv preprint arXiv:1706.02633*, 2017.

[7] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

[8] H. Haddadi, H. Howard, A. Chaudhry, J. Crowcroft, A. Madhavapeddy, D. McAuley, and R. Mortier. Personal data: thinking inside the box. In *Proceedings of The Fifth Decennial Aarhus Conference on Critical Alternatives*, pages 29–32. Aarhus University Press, 2015.

[9] C. Huang, P. Kairouz, X. Chen, L. Sankar, and R. Rajagopal. Context-aware generative adversarial privacy. *Entropy*, 19(12):656, 2017.

[10] K. Katevas, H. Haddadi, and L. Tokarchuk. Poster: Sensingkit: A multi-platform mobile sensing framework for large-scale experiments. In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*, pages 375–378. ACM, 2014.

[11] M. Malekzadeh, R. G. Clegg, and H. Haddadi. Replacement autoencoder: A privacy-preserving algorithm for sensory data analysis. *The 3rd ACM/IEEE International Conference on Internet-of-Things Design and Implementation, 17-20 April, 2018, Orlando, Florida, U.S.A. (to appear)*.

[12] N. Saleheen, S. Chakraborty, N. Ali, M. M. Rahman, S. M. Hossain, R. Bari, E. Buder, M. Srivastava, and S. Kumar. msieve: differential behavioral privacy in time series of mobile sensor data. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 706–717, 2016.

[13] S. Salvador and P. Chan. Toward accurate dynamic time warping in linear time and space. *Intelligent Data Analysis*, 11(5):561–580, 2007.

[14] A. Tripathy, Y. Wang, and P. Ishwar. Privacy-preserving adversarial networks. *arXiv preprint arXiv:1712.07008*, 2017.

[15] G. Vavoulas, C. Chatzaki, T. Malliotakis, M. Pediaditis, and M. Tsiknakis. The mobiact dataset: Recognition of activities of daily living using smartphones. In *ICT4AgeingWell*, pages 143–151, 2016.

[16] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th International Conference on Machine learning*, pages 1096–1103, 2008.

[17] J. Yang, M. N. Nguyen, P. P. San, X. Li, and S. Krishnaswamy. Deep convolutional neural networks on multichannel time series for human activity recognition. In *Proceedings of the 24th International Conference on Artificial Intelligence*, pages 3995–4001, 2015.