Time-series Anonymization of Tabular Health Data using Generative Adversarial Network

Atiye Sadat Hashemi, Kobra Etminani, Amira Soliman, Omar Hamed, Jens Lundström
Center for Applied Intelligent Systems Research (CAISR), Halmstad University, Sweden
{atiye-sadat.hashemi, kobra.etminani, amira.soliman, omar.hamed, jens.r.lundstrom}@hh.se

2023 International Joint Conference on Neural Networks (IJCNN) | 978-1-6654-8867-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/IJCNN54540.2023.10191367

Abstract—**Data anonymization has been used as a fundamental tool in various domains, e.g. healthcare, to alter personal data such that individuals can no longer be identified directly or indirectly in a way to enable broader sharing of data. For example, data perturbation techniques add noise to original data allowing individual record confidentiality while maintaining high-quality data for analytical purposes. In this paper, we propose a perturbation technique for anonymizing longitudinal tabular data such as electronic health records (EHRs). Our model starts by learning a latent space of original data to better capture temporal trends, then employs a generative adversarial network together to train a perturbation generator. During model training, a time-supervised loss function for handling sequence-dependent noise, together with the adversarial unsupervised, anonymization, and reconstruction loss functions are utilized. To evaluate our model quantitatively, we use multiple evaluation metrics for the fidelity, utility, and identifiability of generated data, in addition, the model is evaluated qualitatively by visualizing generated and original data. The results confirm that our model preserves the privacy of the original data and generates a perturbed version with high fidelity and utility compared to some state-of-the-art techniques.**

Index Terms—**generative adversarial networks, anonymization, synthetic data, data perturbation, EHR.**

I. Introduction

In recent years, advances in artificial intelligence (AI) and deep learning (DL) present an opening to optimize different research areas using big datasets [1]. Researchers have shown that the analysis of data with advanced data mining tools can assist healthcare organizations in improving service quality, diagnosis strategies, and developing personalized treatment [2, 3]. Although, access to high-quality big datasets for improving DL-based models is a big challenge, more specifically in highly sensitive applications such as healthcare systems where maintaining data privacy is a necessity [4]. Writing ethical applications for requesting access to sensitive data is not only time-consuming but due to privacy issues, there are many restrictions on the use of data even after access. Data anonymization is a principal tool for minimizing privacy risk, especially risks related to the reidentification of individuals, for broader sharing of data [5].

There are several techniques to anonymize sensitive data such as data suppression, data generalization, data permutation, and data perturbation [6]. The output of these methods is known as anonymous data. The necessity of data anonymization differs from application to application and can be listed as some essential groups such as protecting privacy, improving data accessibility, improving machine-learning models for research, and product development [7]. Additionally, generating a synthetic version of the original data by approximating the underlying distributions introduces the possibility of sharing the synthetic version to be used for further purposes [6].

There are two main types of data perturbation techniques, the probability distribution and the value distortion approaches [8]. The first method takes the data and replaces it from the same distribution sample or the distribution itself. For instance, in a health database that contains a patient's age, address, phone number, and historical medical information, the transmitter can scramble the patient's age so that they won't match the details. The latter approach, meanwhile, uses several additive noises or other randomization processes. Each data point can thus be anonymized by perturbation. On the other hand, our proposed anonymization model aims to perturb data embedding space. By employing a generative model with an autoencoder network, we aim to perturb all attributes' latent space and generate an anonymous dataset.

In this work, we propose a data perturbation scheme to anonymize longitudinal data, specifically health data in the form of electronic health records (EHRs) [9]. Fig. 1 illustrates a patient's EHR in a trajectory from birth to death, while each visit might contain different information such as diagnosis, medications, lab values, radiography, electrocardiogram (ECG), hospitalization data, and clinical notes. In this paper, we consider a limited part of EHRs as time series tabular data with attributes like patient's age, lab values, and important dates, which can be extended.

To perturb time series data, such as a patient record that includes the sequence of visits, we need to pay attention to data temporality by including recurrent neural networks (RNNs) in model architecture. Time-GAN [10], as an RNN-based generative adversarial network, has been a movement using GANs for sequential data generation. In this paper, we leverage Time-GAN for generating universal anonymous perturbation for our time series EHR data.

The contribution of this work is as follows:

- We present a unifying framework for anonymizing time-series tabular data. Our framework so-called time series anonymization perturbation generative adversarial network (TAP-GAN) can generate high-fidelity anonymous time-series data. To improve the security aspect of data anonymization, our method trains a perturbation generator to add perturbation to the latent space instead of the original space. After that, a recovery network is employed to generate the respective anonymous data set.

- While we focus on anonymizing patients' electronic health records (EHRs) in our paper, we test our proposed model with different numbers of RNN layers. Results show that shallow recurrent neural networks (i.e., only one RNN layer), suit better for this application.

In the following, Section II covers a background of data anonymizing as well as an overview regarding GANs as state-of-the-art generative models in privacy-preserving. The mathematical notation of our problem is stated in Section III. Section IV comprises the proposed method. Section V covers the experimental setup and results. Our conclusions are drawn in the final section.

II. LITERATURE REVIEW

The field of data anonymization investigates different approaches for generating anonymous high-quality application-based data. In this section, we review this area as well as the tool of GANs for privacy-preserving.

A. Anonymization

Anonymization, which by applying some operations on original data tries to effectively protect data privacy without degrading its utility [11], is a possible solution for privacy-preserving data publishing (PPDP) [12]. In recent years, PPDP has received significant consideration, and various traditional and AI approaches have been pursued to prevent identity disclosures through anonymization techniques such as data perturbations [13]. Regulations such as the general data protection regulation (GDPR) [14] require data anonymization or removal of personal information before processing any knowledge extraction task or query. Moreover, many techniques have been proposed for different types of data (tabular, graph, ...) to authorize data publishing [12, 15, 16, 17]. In this paper, we consider time-series tabular data perturbation tasks where we need time-dependent data perturbation [18].

Some renowned anonymization models proposed in the literature are k-anonymity [19], l-diversity [20], and t-closeness [21], and their different modified versions [12]. By dividing the attributes of data into direct identifiers, quasi-identifiers, and sensitive attributes, these methods leverage generalization or suppression techniques for privacy-preserving. They define an equivalence class (EC) of an anonymized table to be a set of records that have the same values for the quasi-identifiers (QIs). The k- anonymity model requires that each EC has at least k records to protect against disclosure. The l-diversity modifies k-anonymity in a way to be sure that there are at least l-distinct values of sensitive attributes in each EC. In t-closeness, an EC has t-closeness if the distance between the distribution of a sensitive attribute in this class and the distribution of the attribute in the whole data is less than or equal to t as a threshold. A tabular dataset is considered to have t-closeness if all ECs have t-closeness.

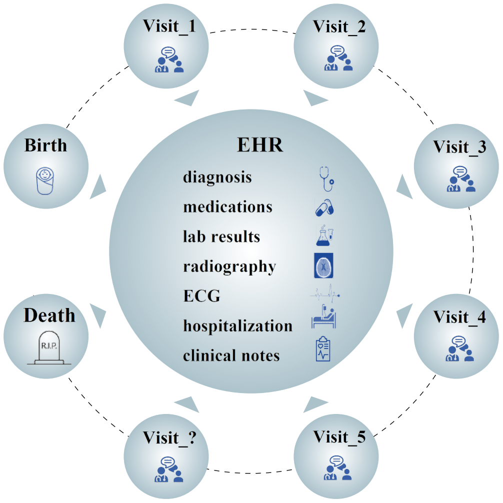


Fig. 1: Patient trajectory electronic health records from birth to death.

On the other hand, some other methods like data perturbation and synthetic data generation try to disturb or generate all the attributes in a dataset [22]. It has been proven that sometimes it is necessary to perturb even nonconfidential attributes to avoid biased responses to queries [16]. While synthetic data generators try to mirror the statistical properties of data, the perturbation method attempts to preserve data privacy by modifying the values of attributes using a randomized process [2]. It is achieved through two possible approaches known as value-class membership and value distortion. In this work, the emphasis is on the latter. In value distortion approaches, the owner of a dataset provides a perturbation p, a random value drawn from a certain distribution, which can be added to the original data. The Uniform and Gaussian distributions are the most commonly used distributions. It is important to carefully select the value p used to modify the original values since if p is too small, the data will not be sufficiently anonymized, and if it's too large, the data may not be usable anymore.

In recent years, several deep learning-based methods are applied for privacy-preserving as well [23]. In the next section, we review GAN-based methods.

B. Generative Adversarial Networks for Privacy Preserving

Generative models are one of the most promising approaches to analyzing and understanding the treasure trove of data [24]. GANs are a deep learning-based generative model based on the game theoretic scenario in which the generator must compete against a discriminator model. The generator directly produces fake data, and the discriminator attempts to distinguish between real and fake generated data.

Among all the different applications of GANs [24, 25], from the privacy aspect, they can be used for preserving data privacy as well as model privacy. In this paper, we consider data privacy preservation by GANs. In these scenarios, the generator is designed as a function to hide private information and/or generate data trained by one or more discriminators for privacy-preserving data generation. On the other hand, the discriminator is employed to ensure data similarity so that the generated privacy-preserving data is still usable in real applications but is hard to be distinguished from the real data.

Recently, the use of GANs for anonymizing different data types (image, video, tabular data, etc.) is increasing [26, 27, 28]. Feutry et al., utilized deep convolutional GANs for image anonymization [26]. Tieu et al., proposed a Spatio-temporal generative adversarial network to generate anonymized gaits that appear natural as a means of preventing people in a video from being identified by a gait recognition system [27]. There are several GAN-based anonymizing approaches considering tabular data as well, like Health-GAN, Med-GAN, Time-GAN [10], P-GAN [29], PP-GAN [30], AnomiGAN [31], and HCGAN [32] where some of them focused on health data but not the time-series aspect of EHRs. In Med-GAN and Health-GAN, authors leverage a different version of GANs known as Wasserstein GAN with gradient penalty and boundary-seeking GAN together with an autoencoder for generating categorical data. In AnomiGAN, the authors employed a target classifier together with a GAN to score each generator's output in order to create high-fidelity synthetic data. They also use a privacy parameter that controls the privacy level. In HCGAN [32], for considering privacy-sensitive aspects of health data, the original data is not used in the training phase. They identify QI features and apply f-differential privacy on them (instead of generalization and suppression anonymization techniques). Then, the trained model generates synthetic data which is not memorizing original data points. Since it has been proved that the gradient parameters of these models can remember the training data, PP-GAN adds well-designed noise to the model gradients during the training procedure. In ADS-GAN [33], after discussing the limitation of differential privacy in healthcare, a measurable definition for identifiability is proposed for generating synthetic data through conditional GANs. In P-GAN, a privacy-preserving generative adversarial network is proposed to model Electronic Health Records. One of the most significant limitations of all these proposed GAN models is their inability to deal with time-series data.

Time-GAN [10] is a framework for time-series synthetic data generation that combines an RNN-based GAN with an autoencoder. For generating more realistic data, Time-GAN combines the flexibility of the unsupervised GAN approach with the control afforded by the supervised loss function. While we also consider time series data in this paper, we make use of RNN-based GAN to generate perturbation for time-series data anonymization. Recently, researchers have also applied transformers for data anonymization, e.g., TTS-GAN [34] consists of two transformers encoder-based in the role of a generator and a discriminator for generating various lengths of time-series data.

III. PROBLEM FORMULATION

The goal is to utilize training data $\mathcal{X}^{train}$ to learn a perturbation generator g_p that can anonymize the original data while the quality and privacy of generated data are preserved.

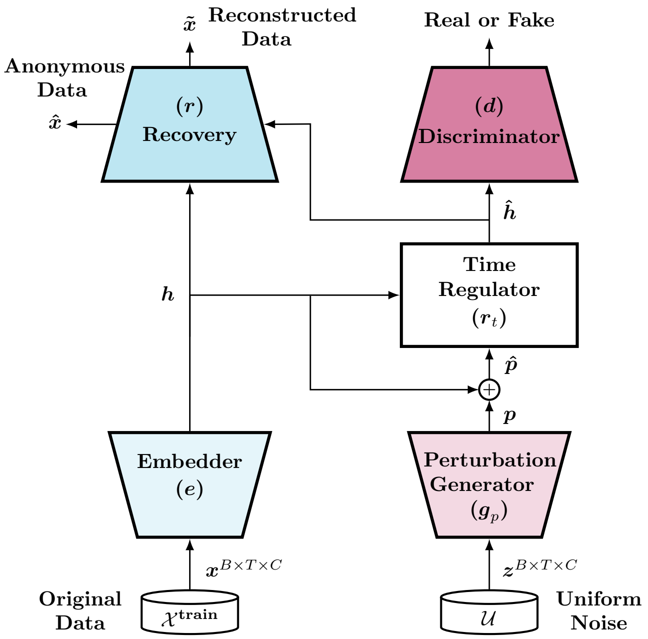


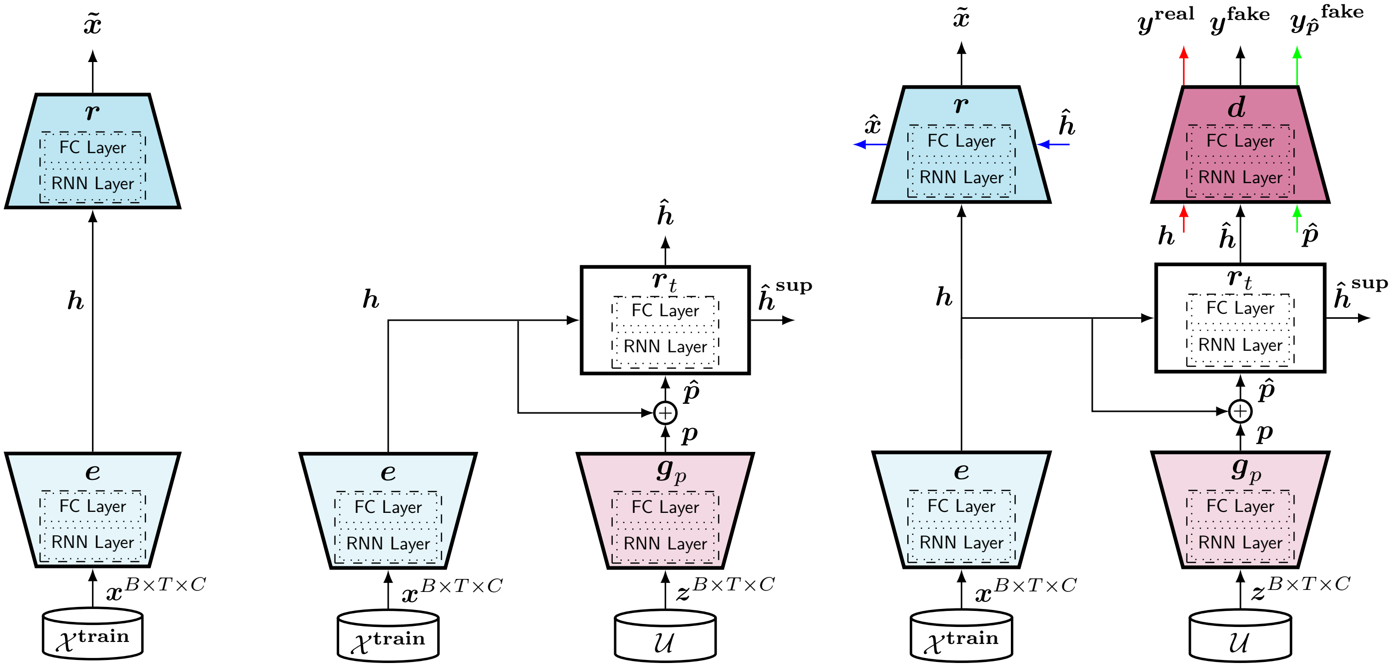
Fig. 2: Our proposed approach to anonymize time series data.

We denote the training dataset with N samples $\mathcal{X}^{train} = \{x_{n,1:T_n}\}_{n=1}^{N}$ as the standard discrete-time setting for tabular time series. Hereinafter, the subscripts n are omitted for simplicity. Let original input $x \in \mathcal{X}^{train}$ be indexed by time steps t in rows, i.e., the full trajectory is $x_{t,1:T}$, where T shows the sequence length. Moreover, the input x includes C columns of attributes while each feature can be continuous or categorical. Then x^{B*T*C} is a batch of original input where B, T, and C, represents the batch size, the sequence length, and the number of input features, respectively. The original data can be transformed from real space to latent space using an embedder network. The latent space of x^{B*T*C} is stated as h^{B*T*H_d}. The dimension of data in latent space is considered as $B * T * H_d$ where B, T, and H_d, represent the batch size, the sequence length, and hidden dimensions respectively. The final output of the model is the anonymous data $\hat{x}$ (with the same dimensions as the original data).

IV. PROPOSED METHOD

Fig. 2 illustrates our proposed model in the training scheme. The model architecture consists of five network components so-called **embedder**, and **recovery** (as the autoencoding part) together with a **perturbation generator**, **time regulator**, and **discriminator** (as the adversarial part). The key intuition of using an autoencoder besides GAN is that model learns to encode features and generate latent representations which we aim to anonymize them. At the core of our proposed method, there is a perturbation generator beside a time regulator that employs other networks to generate perturbations for conducting data anonymization across time.

A multi-dimensional input $z \in [0,1]^{B*T*C}$ sampled from a uniform distribution is fed to the perturbation generator g_p. The network g_p outputs the anonymization perturbation



(a) Phase 1

(b) Phase 2

(c) Phase 3

Fig. 3: TAP-GAN training scheme. (a): Phase 1 includes the Autoencoder training step, (b): Phase 2 is the Time Regulator training step, and (c): Phase 3 is the general training step where all the networks are involved during the training.

p. From the other part, the input $x \in \mathcal{X}^{train}$ feeds the embedder e, to result in the latent space representation of original data points as h. In the training phase, the resulting anonymizing perturbation p is added to the representation of original data h to create $\hat{p}$. The generated $\hat{p}$ is then fed to a time regulator network r_t to learn the temporal aspect of input data through a supervised procedure. Therefore, for each $h \in \mathcal{X}^{train}$ there exists p that can anonymize it through $h+p$. Then the discriminator d, in an adversarial way, tries to recognize the real and anonymous inputs. Finally, a network so-called recovery is employed to produce the anonymized data points $\hat{x}$. The network training is performed through 3 phases as follows.

A. Training Phase-1

In this step, the embedder e and the recovery r would be trained through a reconstruction loss function. Fig. 3a illustrates this process as well as the architecture of each network. The loss function for training these two networks is defined as follows.

$$\mathcal{L}_{phase1} = MSE(x, \tilde{x}) = Mean(\sum_{t=1}^{T} \|x_t - \tilde{x}_t\|_2) \quad (1)$$

where MSE stands for the mean square error, and $Mean$ is the average on N. Also, x is the original input, and $\hat{x}$ is the reconstructed input which is the output of recovery r.

B. Training Phase-2

Fig. 3b shows the networks which are involved in the second training phase. In this step, the embedder's weights are frozen, and the generator and time regulator are the only networks that will be trained through the following loss function.

$$\mathcal{L}_{phase2} = MSE(h, \hat{h}^{sup}) = Mean(\sum_{t=1}^{T} \|h_t - \hat{h}_t^{sup}\|_2) \quad (2)$$

where h is the original input latent space, and $\hat{h}^{sup}$ is the time regulator's output when its input is h. It is the time-supervised loss function that we apply for learning the time-dependent data perturbation.

C. Training Phase-3

Fig. 3c illustrates the final phase of model training, in which the parameters of all networks will be updated in each iteration using the following loss functions. The GAN adversarial loss function, for optimizing the generator and discriminator weights, as well as the reconstruction, the anonymization, and the supervised loss function, are involved in joint training. First, the generator and time regulator weights are updated through a combination of the adversarial, the supervised, and the anonymization loss functions as follows.

$$\begin{aligned}\mathcal{L}_{phase3_a} &= CE(y^{fake}, y^{ones}) + CE(y_{\hat{p}}^{fake}, y^{ones}) \\ &+ Mean(\sum_{t=1}^{T} \|h_t - \hat{h}_t^{sup}\|_2) + Mean(\sum_{t=1}^{T} \|x_t - \hat{x}_t\|_2)\end{aligned} \quad (3)$$

where CE stands for the cross-entropy loss function. y^{fake} and $y_{\hat{p}}^{fake}$ are the output of the discriminator when the inputs are $\hat{h}$ and $\hat{p}$, respectively. The y^{ones} is a vector filled by one as the ground truth labels with the same dimension of y^{fake} and $y^{fake}_{\hat{p}}$. By using the adversarial loss function, $CE(y^{fake}, y^{ones}) + CE(y^{fake}_{\hat{p}}, y^{ones})$, we aim to make the generator powerful to perturb the latent space while the quality of data should be preserved as well. $Mean(\sum_{t=1}^{T} \|h_t - \hat{h}^{sup}_t\|_2)$ is the time supervised loss function, and $Mean(\sum_{t=1}^{T} \|x_t - \hat{x}_t\|_2)$ shows the anonymization loss function in which the $\hat{x}$ is the anonymous data.

Second, the embedder and recovery networks are trained again through the reconstruction and time-supervised loss functions as follows.

$$\mathcal{L}_{phase3_b} = Mean(\sum_{t=1}^{T} \|x_t - \tilde{x}_t\|_2) + Mean(\sum_{t=1}^{T} \|h_t - \hat{h}^{sup}_t\|_2) \tag{4}$$

Third, the discriminator would be also trained through adversarial loss function, i.e.,

$$\mathcal{L}_{phase3_c} = CE(y^{real}, y^{ones}) + CE(y^{fake}, y^{zeros}) + CE(y^{fake}_{\hat{p}}, y^{zeros}) \tag{5}$$

where y^{real}, y^{fake}, $y^{fake}_{\hat{p}}$ are the output of the discriminator when the inputs are h, $\hat{h}$, and $\hat{p}$, respectively. Also, y^{zeros} and y^{ones} show the ground truth labels filled by Zero and One, respectively. In the discriminator training phase, the ground truth for input h is considered as a vector filled with Ones since h is the original latent space and we aim to give the true information regarding h to the discriminator. These phases together help us to reach a perturbation generator for anonymizing data points. The next section covers the experimental setup and the results of the proposed method.

V. EXPRIMENTAL SETUP AND RESULTS

In this section, we present our experimental setup for training and evaluating the proposed framework.

A. Model Architecture

As in Fig. 3 is illustrated, the architecture of all five models, e, r, g_p, r_t, d, in TAP-GAN consist of an RNN layer (which can be an LSTM, GRU, or an LSTM layer normalization) followed by a fully connected layer. We used LSTM layers in the experiments.

B. Dataset

We utilized the Medical Information Mart for Intensive Care (MIMIC-IV) [35], a freely available public dataset created at the laboratory for computational physiology in the USA, as our training EHR data. While it contains de-identified data for over 60,000 patients admitted to ICUs, we focused on the Heart Failure (HF) cohort of the MIMIC-V dataset. For this aim, all admissions that include ICD codes related to the HF cohort are retrieved [35]. Then, patients are obtained from these admissions. Finally, all the admissions for these HF patients (the patients' trajectory) are extracted. The HF Cohort dataset includes 19989 admissions for 9133 different patients. For simplicity, we consider the length of the sequence fixed where there are 5 visits for each HF patient. The attributes of each visit include patient age, length of stay in the hospital, time from the first diagnosis as the HF, time from the previous admission, Potassium, and Sodium lab values.

To test the performance of TAP-GAN on a monotonic function, we simulate multivariate sinusoidal sequences of different phases $\phi \sim \mathcal{U}[0, 1]$ and frequencies $f \sim \mathcal{U}[-\pi, \pi]$ as well. Considering 5 dimensions, $x_i(t) = sin(2\pi f t + \phi)$ for each $i \in \{1, 2, 3, 4, 5\}$. The summary statistics of the two used time-series datasets are shown in Table 1.

C. Evaluation Metrics

Several evaluation metrics can be used to measure the quality of the generated anonymous data. However, the assessment process is inherently tricky since different evaluation metrics yield different trade-offs. The three important aspects are fidelity, utility, and privacy aspects [36].

Fidelity: means generated samples should be distributed to cover the real data and they should be indistinguishable from the real data as well. For fidelity, in addition to PCA, t-SNE, and pairwise correlation plots, we calculate the discriminative score (to evaluate the classification accuracy between original and anonymous data using post-hoc network).

Utility: aims to state that samples should be as useful as the real data when used for the same application (i.e. train-on-synthetic data, test-on-real data). To assess the utility, we report the predictive score of original and anonymous data in which we use Post-hoc RNN architecture to predict one step ahead. This score is expressed by mean absolute error, so the lower the better.

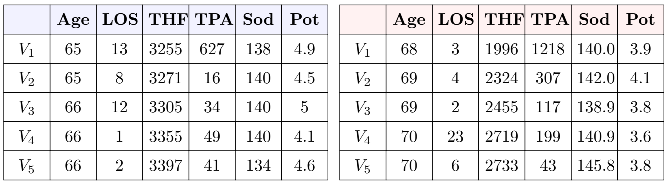
Privacy: Evaluating privacy is one of the most challenging tasks in data anonymization. This is because there is a wide vision for privacy preservation based on vague regulations in GDPR and HIPPA. To assess the privacy provided in generated data we utilize the identifiability metric proposed by [33]. The anonymous dataset $\mathcal{X}'$ is ϵ-identifiable from original dataset $\mathcal{X}$ if:

$$\frac{1}{N}[\mathcal{P}(\hat{d}_i < d_i)] < \epsilon \tag{6}$$

where N is the total number of instances in the dataset, and $\mathcal{P}$ represents the probability. $\hat{d}_i$, and d_i are the minimum weighted Eclidean distance between each instance in $\mathcal{X}$ and the other original observations in $\mathcal{X}$, and the minimum weighted Eclidean distance between each instance in $\mathcal{X}$ and the generated observations in $\mathcal{X}'$, respectively. Also, $\epsilon \in [0, 1]$ where 0 would represent a complete non-identifiable generated data and 1 would represent a completely identifiable dataset.

D. Experimental Result

Fig. 4 shows one example of a patient EHR in the MIMIC-IV dataset and an anonymous sample generated through TAP-



	Age	LOS	THF	TPA	Sod	Pot
V_1	65	13	3255	627	138	4.9
V_2	65	8	3271	16	140	4.5
V_3	66	12	3305	34	140	5
V_4	66	1	3355	49	140	4.1
V_5	66	2	3397	41	134	4.6

(a) Original

	Age	LOS	THF	TPA	Sod	Pot
V_1	68	3	1996	1218	140.0	3.9
V_2	69	4	2324	307	142.0	4.1
V_3	69	2	2455	117	138.9	3.8
V_4	70	23	2719	199	140.9	3.6
V_5	70	6	2733	43	145.8	3.8

(b) Anonymous

Fig. 4: An example of original patient EHR, and generated anonymous EHR, when there are 5 visits for a patient.

TABLE I: The summary statistics of the two used time-series datasets.

Dataset	Dimension	S-Length	#Training set	#Test set
Sins	5	5	8000	2000
HF MIMIC-VI	6	5	3448	871

TABLE II: Result while the different number of RNN layers in each network in TAP-GAN is employed. Results are reported on the HF MIMIC-VI dataset while the sequence length is 5. We report results for the discriminative score, predictive score, and identifiability between original and generated anonymous data. Best results are printed in boldface.

Metric	Number of RNN layers in each network				
	5	4	3	2	1
discriminative score	0.387	0.382	0.167	0.126	**0.114**
predictive score	0.386	0.387	0.383	0.351	**0.333**
identifiability	**0.371**	0.378	0.376	0.380	0.409

GAN. The six attributes in these EHRs are patient Age, length of stay in the hospital (LOS), time from the first diagnosis as the HF (THF), time from the previous admission (TPA), Sodium (Sod), and Potassium (Pot). In Fig. 5, the visualization with pairwise correlation plots for attributes in our MIMIC-HF cohort is illustrated to provide insights into the relationships between attributes. Comparing the number of RNN layers in the networks, we observe that anonymous data generated by 1-layer LSTM in TAP-GAN show markedly better overlap with the original EHR than other states. Fig. 6, where the PCA and t-SNE on HF-MIMIC dataset are illustrated, confirms that 1 layer LSTM works better. We analyze the performance of our anonymous data perturbation method in terms of the discriminative score, predictive score, and identifiability measure in Table II. The result shows the trade-off between the data usefulness and the privacy aspect, which is discussed in detail in the research community as well [33, 37]. Moreover, the comparison of these three metrics between TAP-GAN and some other GAN-based methods in Table III, shows our method performs better than others.

Training of TAP-GAN, implemented in Tensorflow 1.15, was performed on 2 Nvidia GeForce RTX 2080.

TABLE III: Result on multiple time-series datasets for different methods. *The results are reported from the respective paper [10]. Best results are shown using boldface font.

Metric	Method	Sines	HF MIMIC-IV
Discriminative score	Wave-GAN*	0.27	–
	C-RNN-GAN*	0.22	–
	RC-GAN*	0.22	–
	Time-GAN	0.11	0.264
	TAP-GAN	**0.01**	**0.114**
Predictive score	Original	0.059	0.282
	Wave-GAN*	0.134	–
	C-RNN-GAN*	0.127	–
	RC-GAN*	0.097	–
	Time-GAN	0.074	0.371
	TAP-GAN	**0.061**	**0.333**
Identifiability	Time-GAN	**0.25**	0.540
	TAP-GAN	**0.25**	**0.409**

VI. Conclusion

This paper introduces TAP-GAN, a novel framework for anonymizing time series healthcare data. Though the proposed model is evaluated using the EHR dataset, the introduced architecture is suitable for any other time series dataset. Our model anonymizes longitudinal data by generating time series perturbations that are learned in the embedding space of original data to better maintain temporal dependency. A visual comparison and data point distributions mapped in two dimensions show the similarity of the original and the anonymous data while the identifiability score illustrates that privacy is preserved. Since LSTMs are not the best choice to capture long temporal relations, employing Transformers for EHRs anonymization will be the subject of our future research.

References

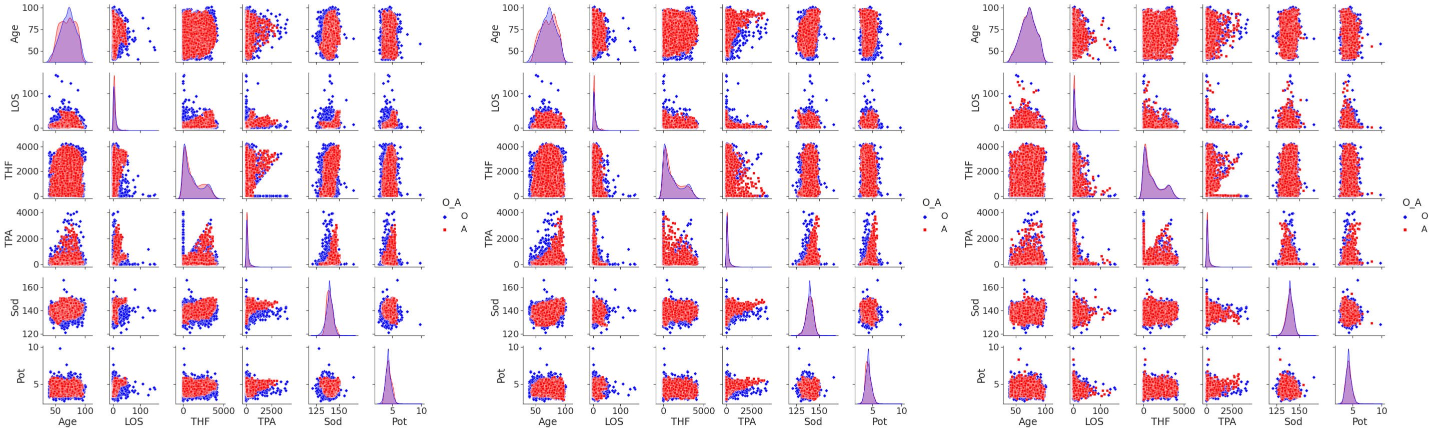
[1] H. Park, H. Bharadhwaj, and B. Y. Lim, "Hierarchical multi-task learning for healthy drink classification," in *2019 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2019, pp. 1–8.

[2] A. Soliman, J. R. Chang, K. Etminani, S. Byttner, A. Davidsson, B. Martínez-Sanchis, V. Camacho, M. Bauckneht, R. Stegeran, M. Ressner *et al.*, "Adopting transfer learning for neuroimaging: a comparative analysis with a custom 3d convolution neural network model," *BMC medical informatics and decision making*, vol. 22, no. 6, pp. 1–16, 2022.

[3] F. Mercaldo, F. Martinelli, and A. Santone, "A proposal to ensure social distancing with deep learning-based object detection," in *2021 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2021, pp. 1–5.

[4] B. Eze and L. Peyton, "Systematic literature review on the anonymization of high dimensional streaming datasets for health data sharing," *Procedia Computer Science*, vol. 63, pp. 348–355, 2015.

[5] Z. Zuo, M. Watson, D. Budgen, R. Hall, C. Kennelly, N. Al Moubayed *et al.*, "Data anonymization for pervasive health care: Systematic literature mapping study,"

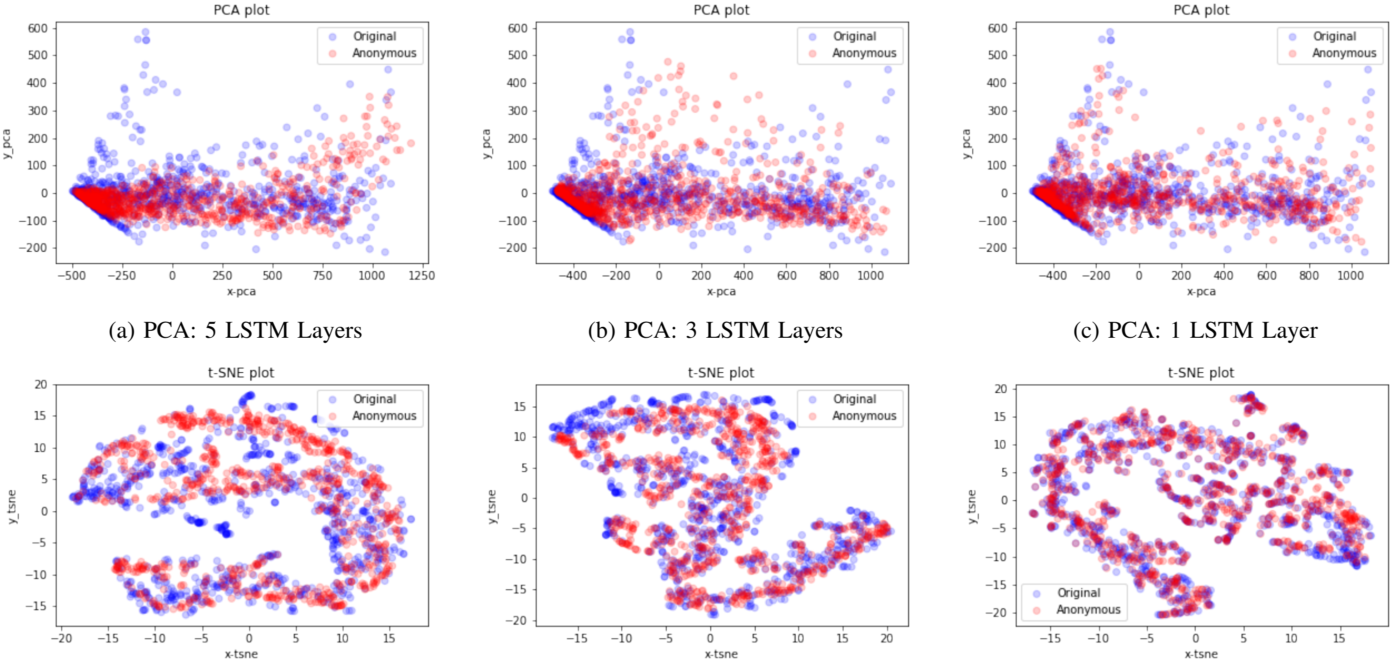


(a) 5 LSTM Layers

(b) 3 LSTM Layers

(c) 1 LSTM Layer

Fig. 5: Correlation pair plots for original and anonymous EHRs while a different number of LSTM layers is employed in TAP-GAN. ’O’ and ’A’ stand for original and anonymous data, respectively.



(a) PCA: 5 LSTM Layers

(b) PCA: 3 LSTM Layers

(c) PCA: 1 LSTM Layer

(d) t-SNE: 5 LSTM Layers

(e) t-SNE: 3 LSTM Layers

(f) t-SNE: 1 LSTM Layer

Fig. 6: PCA (1^{st} row) and t-SNE (2^{st} row) visualizations on HF MIMIC-IV while a different number of LSTM layers is employed in TAP-GAN.

JMIR medical informatics, vol. 9, no. 10, p. e29871, 2021.

[6] I. E. Olatunji, J. Rauch, M. Katzensteiner, and M. Khosla, “A review of anonymization for healthcare data,” *Big Data*, 2022.

[7] J. Li, B. J. Cairns, J. Li, and T. Zhu, “Generating synthetic mixed-type longitudinal electronic health records for artificial intelligent applications,” *arXiv preprint arXiv:2112.12047*, 2021.

[8] K. Muralidhar, R. Parsa, and R. Sarathy, “A general additive data perturbation method for database security,” *management science*, vol. 45, no. 10, pp. 1399–1415, 1999.

[9] A. Ashfaq, A. Sant’Anna, M. Lingman, and S. Nowaczyk, “Readmission prediction using deep learning on electronic health records,” *Journal of biomedical informatics*, vol. 97, p. 103256, 2019.

[10] J. Yoon, D. Jarrett, and M. Van der Schaar, “Time-series generative adversarial networks,” *Advances in neural information processing systems*, vol. 32, 2019.

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

[12] A. Majeed and S. Lee, "Anonymization techniques for privacy preserving data publishing: A comprehensive survey," *IEEE access*, vol. 9, pp. 8512–8545, 2020.

[13] A. Goncalves, P. Ray, B. Soper, J. Stevens, L. Coyle, and A. P. Sales, "Generation and evaluation of synthetic patient data," *BMC medical research methodology*, vol. 20, no. 1, pp. 1–40, 2020.

[14] C. J. Hoofnagle, B. van der Sloot, and F. Z. Borgesius, "The european union general data protection regulation: what it is and what it means," *Information & Communications Technology Law*, vol. 28, no. 1, pp. 65–98, 2019.

[15] B. C. Fung, K. Wang, R. Chen, and P. S. Yu, "Privacy-preserving data publishing: A survey of recent developments," *ACM Computing Surveys (Csur)*, vol. 42, no. 4, pp. 1–53, 2010.

[16] K. Muralidhar and R. Sarathy, "Security of random data perturbation methods," *ACM Transactions on Database Systems (TODS)*, vol. 24, no. 4, pp. 487–493, 1999.

[17] M. Hernandez, G. Epelde, A. Alberdi, R. Cilla, and D. Rankin, "Synthetic data generation for tabular health records: A systematic review," *Neurocomputing*, 2022.

[18] X. Ge, S. J. Binnie, D. Rocca, R. Gebauer, and S. Baroni, "turbotddft 2.0—hybrid functionals and new algorithms within time-dependent density-functional perturbation theory," *Computer Physics Communications*, vol. 185, no. 7, pp. 2080–2089, 2014.

[19] L. Sweeney, "k-anonymity: A model for protecting privacy," *International journal of uncertainty, fuzziness and knowledge-based systems*, vol. 10, no. 05, pp. 557–570, 2002.

[20] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, "l-diversity: Privacy beyond k-anonymity," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 1, no. 1, pp. 3–es, 2007.

[21] N. Li, T. Li, and S. Venkatasubramanian, "t-closeness: Privacy beyond k-anonymity and l-diversity," in *2007 IEEE 23rd international conference on data engineering*. IEEE, 2006, pp. 106–115.

[22] M. Ballout, M. Tuqan, D. Asmar, E. Shammas, and G. Sakr, "The benefits of synthetic data for action categorization," in *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2020, pp. 1–8.

[23] C. Yan, Z. Zhang, S. Nyemba, and B. A. Malin, "Generating electronic health records with multiple data types and constraints," in *AMIA annual symposium proceedings*, vol. 2020. American Medical Informatics Association, 2020, p. 1335.

[24] P.-C. Chang, Y.-Y. Tien, C.-L. Chen, L.-F. Chen, Y.-S. Chen, and H.-L. Chan, "Facial image reconstruction from functional magnetic resonance imaging via gan inversion with improved attribute consistency," in *2022 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2022, pp. 1–8.

[25] A. S. Hashemi and S. Mozaffari, "Secure deep neural networks using adversarial image generation and training with noise-gan," *Computers & Security*, vol. 86, pp. 372–387, 2019.

[26] C. Feutry, P. Piantanida, Y. Bengio, and P. Duhamel, "Learning anonymized representations with adversarial neural networks," *arXiv preprint arXiv:1802.09386*, 2018.

[27] N.-D. T. Tieu, H. H. Nguyen, H.-Q. Nguyen-Son, J. Yamagishi, and I. Echizen, "Spatio-temporal generative adversarial network for gait anonymization," *Journal of Information Security and Applications*, vol. 46, pp. 307–319, 2019.

[28] F. H. Foomani, D. Anisuzzaman, J. Niezgoda, J. Niezgoda, W. Guns, S. Gopalakrishnan, and Z. Yu, "Synthesizing time-series wound prognosis factors from electronic medical records using generative adversarial networks," *Journal of biomedical informatics*, vol. 125, p. 103972, 2022.

[29] R. Venugopal, N. Shafqat, I. Venugopal, B. M. J. Tillbury, H. D. Stafford, and A. Bourazeri, "Privacy preserving generative adversarial networks to model electronic health records," *Neural Networks*, vol. 153, pp. 339–348, 2022.

[30] Y. Liu, J. Peng, J. James, and Y. Wu, "Ppgan: Privacy-preserving generative adversarial network," in *2019 IEEE 25Th international conference on parallel and distributed systems (ICPADS)*. IEEE, 2019, pp. 985–989.

[31] H. Bae, D. Jung, H.-S. Choi, and S. Yoon, "Anomigan: Generative adversarial networks for anonymizing private medical data," in *PACIFIC SYMPOSIUM ON BIOCOMPUTING 2020*. World Scientific, 2019, pp. 563–574.

[32] R. Indhumathi and S. S. Devi, "Healthcare cramér generative adversarial network (hcgan)," *Distributed and Parallel Databases*, pp. 1–17, 2021.

[33] J. Yoon, L. N. Drumright, and M. Van Der Schaar, "Anonymization through data synthesis using generative adversarial networks (ads-gan)," *IEEE journal of biomedical and health informatics*, vol. 24, no. 8, pp. 2378–2388, 2020.

[34] X. Li, V. Metsis, H. Wang, and A. H. H. Ngu, "Tts-gan: A transformer-based time-series generative adversarial network," *arXiv preprint arXiv:2202.02691*, 2022.

[35] A. Johnson, L. Bulgarelli, T. Pollard, S. Horng, and L. Celi, "Mark," *R. MIMIC-IV (version 1.0). PhysioNet*, 2021.

[36] O. Mendelevitch and M. D. Lesh, "Fidelity and privacy of synthetic medical data," *arXiv preprint arXiv:2101.08658*, 2021.

[37] T. Stadler, B. Oprisanu, and C. Troncoso, "Synthetic data–anonymisation groundhog day," in *31st USENIX Security Symposium (USENIX Security 22)*, 2022, pp. 1451–1468.