Advanced Trajectory Privacy Protection with Attention Mechanism and Auxiliary Classifier Generative Adversarial Networks

Jihwan Shin
Department of Artificial Intelligence Application
Kwangwoon University
Seoul, South Korea
shinjihwan1997@kw.ac.kr

Yeji Song
Department of Artificial Intelligence Convergence
Kwangwoon University
Seoul, South Korea
yeah9song@kw.ac.kr

Yoo-Young Cheong
Department of Artificial Intelligence Application
Kwangwoon University
Seoul, South Korea
yycheong@kw.ac.kr

Jinhyun Ahn
Department of Management Information Systems
Jeju National University
Jeju, South Korea
jha@jejunu.ac.kr

Taewhi Lee
Smart Data Research Section
Electronics and Telecommunications Research Institute
Daejeon, South Korea
taewhi@etri.re.kr

Dong-Hyuk Im
School of Information Convergence
Kwangwoon University
Seoul, South Korea
dhim@kw.ac.kr

***Abstract*—Advancements in smartphone technology have led to an increase in the usage of various location-based services (LBS). This resulted in a lot of trajectory data being generated. LBSs provide personalized services for users through continuous queries. However, there is a problem that the user's sensitive information can be inferred through such continuous queries. Although various methods have been proposed to protect personal information, traditional personal information protection methods cannot provide absolute personal information protection when user location information itself, such as LBS, is required. In particular, if sensitive points visited by the user are exposed, this may lead to additional information leakage. Therefore, this paper guarantees anonymity by protecting sensitive points visited by users through the class conditional synthesis of ACGAN, and proposes a synthetic trajectory generation model for generating highly useful trajectory data through the combination of attention mechanisms. Furthermore, the usability and anonymity aspects of the synthetic trajectory data generated by the proposed are compared with those of existing models to verify the performance of the proposed.**

Keywords—generative adversarial network, auxiliary classifier GAN, location based system, trajectory privacy protection, privacy, attention mechanism

I. INTRODUCTION

Advancements in smartphone technology have led to an increase in the usage of various location-based services (LBSs). Users submit queries related to their location to the LBS server, and the LBS provides different types of personalized services, such as in providing restaurant recommendations, directions, and traffic notifications [1, 2]. However, if the LBS server is hacked, the private and personal information of users may be leaked. For example, attackers track and analyze users' trajectory data through continuous queries to infer their home and office addresses. Furthermore, other personal information such as specific disease types can be inferred by checking whether users have visited certain locations such as hospitals [3]. Trjaectory data can be highly utilized in various fields, but since trajectory data contains sensitive information, there is a high risk of personal information leakage, and many studies are being conducted to ensure personal information protection while maintaining the usefulness of data.

Methods for de-identifying data include k-anonymity and l-diversity. K-anonymity is a data protection method that makes it indistinguishable from at least k-1 individuals when disclosing personal information [4]. Wang *et al.* [5] proposed a positional recombination mechanism LRM that captures the probabilistic and geographical features of the trajectory and satisfies k-anonymity while maintaining the availability of data. L-diversity is a method of protecting both sensitive and general attributes by ensuring that personal information cannot be identified from at least L other individuals in the dataset based on sensitive attributes [6]. Temuujin *et al.* [7] proposed an efficient l-diversity anonymization algorithm that can more efficiently protect privacy by identifying the limits of data privacy for dynamically evolving datasets. Jeon *et al.* [8] proposed an l-diversity anatomy de-identification method for resource description frameworks (RDFs) that overcomes the limitations of k-anonymity and ensures stronger privacy.

However, the above de-identification methods cannot guarantee absolute privacy protection. K-anonymity is a privacy method that does not take into account the attacker's background knowledge, and l-diversity only works when there are at least l unique values for each sensitive property in the dataset.

For this reason, when dealing with meaningful data, such as trajectory data, many studies aim to maintain both usefulness and anonymity by generating synthetic data using generative models. Rao *et al*. [9] proposed LSTM-TrajGAN, a synthetic trajectory generation model that preserves temporal and spatial information of trajectory data by combining LSTM (Long Short-Memory) and GAN. Shin *et al*. [10] proposed a synthetic trajectory generating model TCAC-GAN that combines LSTM and ACGAN to randomly change points with sensitive categories. Song *et al*. [11, 12] proposed a synthetic trajectory generation model that protects sensitive points by allocating conditions that should not be included in the outputs generated by GAN. Xiong et al. [13] proposed ADGAN(Auto-Driving GAN), a GAN-based image-to-image conversion method to prevent the problem of leaking the location and trajectory of the vehicle through the camera of an autonomous vehicle.

In this paper, we propose a synthetic trajectory generation model that can efficiently protect sensitive points through the synthesis of class conditional synthesis of ACGAN and increase usefulness through combination with attention mechanism. We also measure the usefulness and anonymity of synthetic trajectory data generated for performance demonstration and compare them with previous research models.

II. Background

GAN is a generation model that is in the spotlight in the field of image generation [14]. Fig. 1 shows the GAN architecture. The GAN consists of two models: the generator and the discriminator. The generator is responsible for generating the image, and the discriminator is responsible for determining the authenticity of the image. The object function of the GAN model is expressed shown in Eq. (1):

$$\min_G \max_D V(D,G) = E_{x\sim P_{data}(x)}[logD(x)] + E_{z\sim p_z(z)}\left[\log\left(1-D\left(G(z)\right)\right)\right] \quad (1)$$

D represents the discriminator, and G represents the generator. x represents real data, and $x\sim P_{data}(x)$ represents data sampled from the probability distribution of the entire real data. $D(x)$ uses real data as input of the discriminator to determine real data. z represents noise, and $z\sim p_z(z)$ represents random noise sampled from Gaussian distribution. $G(z)$ represents fake data generated by a generator that took noise as an input, and $D(G(z))$ represents a discriminator that took fake data as an input. G aims to minimize the Eq. (1), and D aims to maximize it.

ACGAN is a class conditional image synthesis model that uses improved training methods for image synthesis [15]. The Fig. 2 shows the ACGAN architecture. ACGAN consists of two models: generator and discriminator. The generator generates a class conditional image through a label, and the discriminator determines the authenticity of the image and class prediction. The log likelihood of the correct source L_s is expressed shown in Eq. (2):

$$L_s = E[logP(S=real|X_{real})] + E\left[logP\left(S=fake\middle|X_{fake}\right)\right] \quad (2)$$

L_s represents the sum of estimates that distinguish between real data and fake data. The log likelihood of the correct class L_c is expressed shown in Eq. (3):

$$L_c = E[logP(C=c|X_{real})] + E\left[logP\left(C=c\middle|X_{fake}\right)\right] \quad (3)$$

L_c represents the sum of the estimates that distinguish between the real class and fake class. D is trained to maximize the sum of Eq. (2) and Eq. (3). This means that from the perspective of the D, the object is to determine the source of the data well and to predict the class well. G is trained to maximize the difference between Eq. (3) and Eq. (2). This means that from the perspective of the G, the object is for the D to predict the class well and not determine the source.

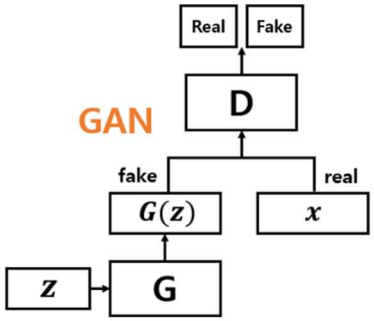

Fig. 1. GAN architecture

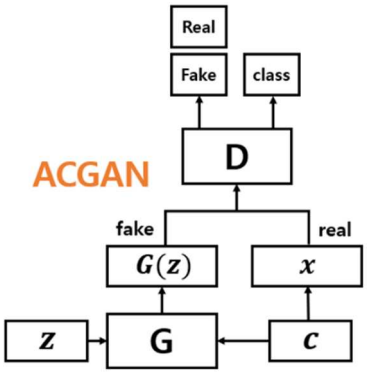

Fig. 2. ACGAN architecture

III. METHODS

A. Generate labels

The generated labels are used as input for the proposed model along with the original trajectory data. The trajectory data used in the model consists of location, date, time, and category data points. Location data consists of the latitude and longitude; date data indicate the day of the week denoted by a number: 0–6. Time data indicates the hour of the day denoted by a number: 0–24. Category data indicate the properties of points denoted by a number: 0–9. Because the labels generated are used to hide the category value of sensitive points, they are also represented by a number 0–9. Fig. 3 shows an example of a label generated for hiding sensitive points. Assuming that a point of trajectory data has four points, as shown in Fig. 3, each point contains location, date, time, and category information. The third point is assumed to be a sensitive point with the hospital category. The label of the third point is then randomly changed to a value between 0 and 9. Accordingly, the generator approximates the trajectory data to a point with the category value corresponding to the label, thereby protecting the sensitive point.

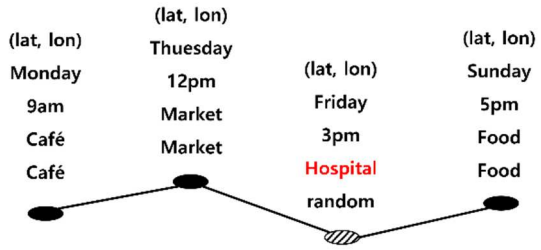

Fig. 3. Example of encoding process.

B. Trajectory encoding process

Fig. 4 shows an example of encoding one of the points. If we consider trajectory data with four points, as shown on the left, then we normalize the location data using the deviation value from the centroid (star) of all points to each point. This process allows the deviation between points to be learned more effectively. Additionally, date data are one-hot encoded using a dimension of 7, time data using a dimension of 24, and category and label data using dimensions of 10.

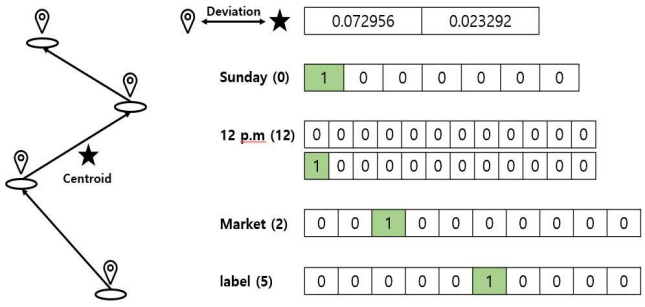

Fig. 4. Example of generating a label.

C. Trajectory generation

Fig. 5 shows the generator architecture. A generator consists of four layers: embedding, feature fusion, attention and LSTM, and decoding. The embedding layer vectorizes the trajectory data input using a multilayer perceptron (MLP). Location data are embedded into a 64-dimension vector, and other data types are embedded into a vector with the dimension of the vocab. The feature fusion layer converges feature vectors corresponding to one point into a 100-dimension vector. This process facilitates temporal and spatial learning between feature vectors. The attention and LSTM layer focuses on a specific location of a previous layer when modeling each location point of the trajectory. Accordingly, trajectory modeling can be more accurate and practical, and the spatial relation between points can be represented more effectively. Other feature vectors in addition to the location vector are modeled while maintaining their spatial characteristics through the many-to-many LSTM layer. Finally, the decoding layer decodes the synthetic trajectory data. The location vector is decoded into latitude and longitude using the tanh activation function, and other features are decoded using the softmax activation function.

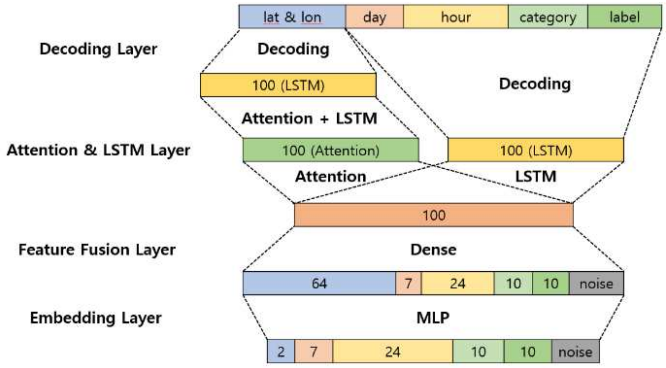

Fig. 5. Generator architecture.

The synthetic trajectory data generated by the generator and original trajectory data are used as input for the discriminator. Fig. 6 shows the discriminator architecture. The discriminator consists of four layers: embedding, feature fusion, LSTM, and classification and prediction. The overall process is identical to the generator; however, the discriminator discerns the authenticity of the trajectory data and predicts classes. Therefore, noise is not added to the output of the embedding layer and only the many-to-many LSTM layer is used instead of the attention layer. The last hidden state in the LSTM layer discerns the authenticity of the trajectory using the sigmoid activation function. In addition, the final output of the LSTM predicts the class (label) of the trajectory using the softmax activation function.

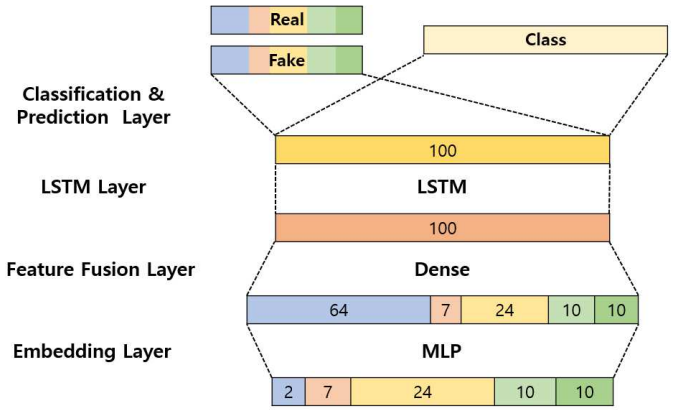

Fig. 6. Discriminator architecture.

IV. Experiments

A. Experiment design

The hyperparameters set for the model training were 2,000 epochs, a batch size of 256, learning rate of 0.01, and Adam optimizer. The LSTM-TrajGAN and TCAC-GAN models were used to compare their performances.

B. Datasets

This study used Foursquare's weekly trajectory dataset for New York City (NYC) [16]. Table I summarizes the Foursquare NYC weekly trajectory dataset. The dataset's attributes comprise the User ID, Trajectory ID, Location (Latitude and Longitude), Day, Hour, Category, Price Tier, Rating, and Weather; we removed the Price Tier, Rating, and Weather attributes. In addition, we did not encode the User or Trajectory IDs because they indicate only the user and trajectory, respectively, of the data point. The location data that deviated from the latitude and longitude ranges of NYC were removed. Consequently, the entire dataset contained 193 users, 3,079 trajectories, and 66,962 points. Two-thirds of the dataset was used as training data and one-third as test data.

TABLE I. SUMMARY OF THE FOURSQUARE NYC WEEKLY TRAJECTORY DATASET

Attribute	Number/Range
Trajectory ID	3,079
User ID	193
Latitude	(40.550852, 40.988332)
Longitude	(-74.269644, -73.685767)
Hour	24
Day	7
Category	10

C. Usefulness measurement

The usefulness of the three models was measured using the Hausdorff distance. The Hausdorff distance was used to measure the similarity between two sets of data, based on the distance between the points of the original trajectory data and the trajectory data generated by the model. The measured items include the MIN, MAX, AVG, and MEAN; lower values indicate a higher similarity between the original and synthetic trajectory data. Table 2 presents the Hausdorff distances measured by the three models. The proposed model achieved lower values for all measurements compared with LSTM-TrajGAN and TCAC-GAN. This implies that the synthetic trajectory data generated by the proposed model are more similar to the original data compared with that of the other models, and more useful.

TABLE II. HAUSDORFF DISTANCE RESULTS

	LSTM-TrajGAN	TCAC-GAN	Proposed
MIN	0.006839	0.004439	**0.002554**
MAX	0.062886	0.052982	**0.051066**
AVG	0.020647	0.017155	**0.015668**
MEAN	0.019752	0.015978	**0.015228**

D. Anonymity measurement

The trajectory-user linking (TUL) test was performed to measure the anonymity of the three models. The TUL test measures the prediction accuracy of a user's trajectory in which a lower accuracy indicates that the possibility of identifying a user is low, and a higher accuracy implies that the possibility of identifying a user is high. Therefore, the lower the accuracy, the higher the anonymity. The measurement items of the TUL test include ACC@1, ACC@5, Macro-Precision, Macro-Recall, and Macro-F1 Score. ACC@1 means Top1-Accuracy, and an index that is calculated when the highest value in the output of softmax is the correct answer. ACC@5 means Top5-Accuracy, calculating the proportion of predicted classes among the five upper classes in the softmax output. Table 3 presents the TUL accuracy. The Macro-Precision of the proposed model is slightly higher than that of TCAC-GAN. This indicates that the attention mechanism was applied to a point approximated to the label, which resulted in a slight decrease in the average precision. However, the proposed model has a lower TUL accuracy than other models in all items excluding the Macro-Precision. This implies that the proposed model can effectively prevent the identification of a user's trajectory.

TABLE III. TUL-TEST RESULTS

	LSTM-TrajGAN	TCAC-GAN	Proposed
ACC@1	0.406037	0.308666	**0.256086**
ACC@5	0.651412	0.587147	**0.515093**
Macro-P	0.345144	0.230609	**0.251950**
Macro-R	0.390910	0.256103	**0.243969**
Macro-F1	0.374779	0.293933	**0.206121**

V. Conclusion

The growing scope of LBSs has enabled users to utilize more useful services, resulting in the creation of a large amount of user trajectory data. However, the trajectory data extensively contain sensitive information, which highlights the importance of protecting personal information. However, traditional data protection methods cannot provide absolute privacy protection. Therefore, this study examined a model capable of protecting

personal information of users while maintaining the usefulness of the trajectory data. The proposed model, which is a synthetic trajectory generation model based on ACGAN, was experimentally proven to control the output of sensitive and insensitive points and be capable of executing more practical trajectory modeling using an attention mechanism. Using the trajectory data generated by the proposed model can help protect personal information from attackers and heighten the usefulness of the trajectory data. As future research, we plan to conduct to introduce differential privacy, a mathematically provable privacy protection method, when conducting statistical queries using the trajectory data generated by the proposed model.

ACKNOWLEDGEMENT

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2021-0-00231, Development of Approximate DBMS Query Technology to Facilitate Fast Query Processing for Exploratory Data Analysis, 50%). This work was also supported by the National Research Foundation of Korea(NRF) granted by the Korea government(MSIT) (No.NRF-2021R1F1A1054739, 50%).

REFERENCES

[1] K. G. Shin, X. Ju, Z. Chen, and X. Hu, "Privacy protection for users of location-based services," *IEEE Wireless Communications*, vol. 19, no. 1, pp. 30-39, 2012.

[2] B. Niu, X. Zhu, Q. Li, J. Chen, and H. Li, "A novel attack to spatial cloaking schemes in location-based services," *Future Generation Computer Systems*, vol. 49, pp. 125-132, 2015.

[3] R. H. Hwang, Y. L. Hsueh, and H. W. Chung, "A novel time-obfuscated algorithm for trajectory privacy protection," *IEEE Transactions on Services Computing*, vol. 7, no. 2, pp. 126-139, 2013.

[4] L. Sweeny, "k-anonymity: A model for protecting privacy," *International journal of uncertainty, fuzziness and knowledge-based systems*, vol. 10, no. 5, pp. 557-570, 2002.

[5] Y. Wang, M. Li, S. Luo, Y. Xin, H. Zhu, Y. Chen, and Y. Yang, "LRM: a location recombination mechanism for achieving trajectory k-anonymity privacy protection," *IEEE Access*, vol. 7, pp. 182886-182905, 2019.

[6] 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.

[7] O. Temuujin, J. Ahn, and D. H. Im, "Efficient L-Diversity Algorithm for Preserving Privacy of Dynamically Published Datasets," *IEEE Access*, vol. 7, no. 1, pp. 122878-122888, 2019.

[8] M. Jeon, O. Temuujin, J. Ahn, and D. H. Im, "Distributed L-diversity using spark-based algorithm for large resource description frameworks data," *The Journal of Supercomputing*, vol. 77, pp. 7270-7286, 2021.

[9] J. Rao, S. Gao, Y. Kang, and Q. Huang, "LSTM-TrajGAN: A Deep Learning Approach to Trajectory Privacy Protection," *11th International Conference on Geographic Information Science (GIScience2021)*, 2020.

[10] J. Shin, Y. Song, J. Ahn, T. Lee, and D. H. Im, "TCAC-GAN: Synthetic Trajectory Generation Model Using Auxiliary Classifier Generative Adversarial Networks for Improved Protection of Trajectory Data," *IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp. 314-315, 2023.

[11] Y. Song, J. Shin, J. Ahn, T. Lee, and D. H. Im, "Except-Condition Generative Adversarial Network for Generating Trajectory Data," *International Conference on Database and Expert Systems Applications*, pp. 289-294, 2023.

[12] Y. Song, J. Shin, J. Ahn, T. Lee, and D. H. Im, "Generating Labeled Multiple Attribute Trajectory Data with Selective Partial Anonymization based on Exceptional Conditional Generative Adversarial Network,", *IEEE Access*, 2023.

[13] Z. Xiong, W. Li, Q. Han, and Z. Cai, "Privacy-preserving auto-driving: a GAN-based approach to protect vehicular camera data," *IEEE international Conference on Data Mining (ICDM)*, pp. 668-677, 2019.

[14] I. Goodfellow et al., "Generative adversarial nets," *Proceedings of the 27th International Conference on Neural Information Processing Systems*, pp. 2672-2680, 2014.

[15] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier GANs," *Proceedings of the 34th International Conference on Machine Learning*, pp. 2642-2651, 2017.

[16] D. Yang, D. Zhang, V. W. Zheng, and Z. Yu, "Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, no. 1, pp. 129-142, 2015.