Personalized Location Intelligence Privacy Protection Technology Based on Differential Privacy Mechanism in Different Data Publishing Scenarios

Hongying Tan

Department of Artificial Intelligence, Chongqing Industry and Trade Polytechnic, Chongqing, 408000, China
E-mail: tanhy23hongying@163.com

Received 16 June 2025; Accepted 24 June 2025

Abstract

In the current era of mobile Internet, personalized location services have significantly enhanced user experience. However, they have concomitantly given rise to substantial challenges related to the protection of location privacy. The study aims to develop a location privacy protection model for static and dynamic data publishing scenarios by deeply analyzing the differential privacy mechanism and its application on mobile aggregated data features. The study designs a protection model using diverse perturbation mechanisms for static data and constructs a set of protection models for dynamic data by combining prediction, adaptive adjustment, and data grouping and merging. The experiments were validated using the Chengdu cab trajectory dataset (real data) and the simulated synthetic population movement dataset (simulated generated data). The experimental results showed that the static model successfully reduced the trajectory recovery rate to 18% and achieved the lowest mean absolute error of 0.8 and the lowest mean relative error of 0.45

Journal of Cyber Security and Mobility, Vol. 14_4, 901–926.
doi: 10.13052/jcsm2245-1439.1446
© 2025 River Publishers

on the simulated dataset. The dynamic model achieved a minimum mean absolute error of 0.047 and a minimum mean relative error of 20 with a fixed privacy budget. The above results demonstrate the potential of the two models in improving the efficacy of location privacy protection and point out the direction for the future development of privacy protection techniques in personalized location services.

Keywords: Differential privacy, data publishing, personalized services, location intelligence, trajectory recovery.

1 Introduction

With the popularity of mobile Internet and smart devices, personalized location services have become an indispensable part of daily life, and at the same time bring the risk of leaking users' location privacy [1]. Although existing privacy protection technologies such as anonymization, perturbation, and Differential Privacy (DP) have achieved certain results, there are still many limitations in practical applications. For example, it is difficult to prevent data from being maliciously inferred or restored in location sharing; The protection of user privacy is incomplete in third-party service scenarios [2]. DP, as a method that can provide mathematically provable privacy protection strength, has become an important research direction in privacy protection in recent years due to its ability to effectively resist background knowledge attacks and inference attacks. Compared with traditional methods, DP can strictly control the probability of personal user information leakage while publishing statistical data. Therefore, this study chooses DP as the core technology for Location Privacy Protection (LPP). Although there have been some studies using DP mechanisms for location data protection in recent years, most of the work has mainly focused on static data scenarios. There is a lack of in-depth consideration of personalized protection requirements in dynamic data stream environments. In addition, existing research often has shortcomings in balancing privacy protection intensity and data availability. For example, introducing too much noise leads to data distortion, or unreasonable allocation of privacy budget reduces the effectiveness of protection. To address these gaps and limitations, the study innovatively proposes two new LPP models. The motivation of this study is to address the issue of difficulty in balancing privacy protection and personalized service needs in current personalized location services. It attempts to explore how to enhance the protection of personal trajectories and movement patterns while

improving data availability, especially in dynamic real-time data streaming environments. For static data scenarios, a DP processing method that dynamically adjusts the noise intensity based on time region division and data discretization is proposed. A privacy budget adaptive allocation strategy combining real-time prediction and sampling point changes was designed for dynamic data scenarios. This strategy can flexibly control the amount of noise injection based on data changes, improving the balance between data utility and privacy protection. The contributions of the study are threefold. First, a dynamic data privacy protection model combining prediction, adaptive adjustment, and group merging is proposed, which achieves a lower value of mean error under the condition of fixed privacy budget. Second, a time zone static data protection model based on differential perturbation is designed, which significantly improves the trajectory recovery speed. Third, extensions are made for the application of DP mechanism in dynamic and static location services, which improve the balanced effect of data availability and privacy protection.

This research is divided into five sections. The first section is a brief introduction to the article as a whole. The second section is an analysis and summary of other scholars' research. The third section describes how the two privacy-preserving models are constructed, while the fourth section tests the performance of the two models. The last section is a summary.

2 Related Works

The rise of personalized location services enables users to enjoy more convenient life services and personalized recommendation experiences, but this also brings the risk of personal Location Information Privacy (LIP) leakage [3, 4]. To solve this problem, many industry scholars have successively proposed the privacy protection model using dynamic data and the privacy protection model using static data. For the dynamic data privacy protection model, Liu et al. [5] proposed a data aggregation framework that combines intra-group effects. This new framework could effectively handle large-scale location information accumulation problems and was more feasible compared to other methods. Experimental results showed that the novel framework could handle large-scale location information accumulation problems well and was feasible. To improve the integrity of source LPP in wireless sensor networks, Chen et al. [6] proposed a novel source location protection mechanism after combining it with the sector phantom routing method. Experimental results showed that the stability and accuracy of this

strategy were much higher than other methods of the same type. To improve the location privacy of users when using social networks, Huo et al. [7] proposed a novel user information protocol protection mechanism after combining it with the clustered dichotomous graph grid algorithm. Experimental results showed that the proposed new mechanism could effectively improve the user's personal privacy security. Zhang et al. [8] found that the existing privacy protection schemes cannot simultaneously protect the user's location privacy or query privacy. Therefore, the research team proposed a privacy protection scheme for location services that combines DP. Experimental results showed that the strength of user privacy protection was significantly enhanced by this new scheme. Parashar et al. [9] found that maintaining the location privacy of gait datasets in deep learning pipelines becomes more difficult. Therefore, the research team proposed a dynamic location dataset protection method incorporating texture modification techniques. The experimental results showed that the dataset privacy under this method was well guaranteed with a high similarity index.

For static data privacy preserving models, DP mechanisms are the simplest and fastest way to deal with them today. Yin et al. [10] proposed an adaptive batch method combining DP to address the issue of personal privacy islands in electroencephalography. This method could effectively improve the electroencephalograph islanding problem in imbalanced datasets. Cheng et al. [11] found that neural network models can inadvertently leak personal privacy during large-scale training. Therefore, the team proposed an improved DP random gradient descent algorithm to address privacy protection issues in deep learning tasks. The denoising effect of this algorithm was excellent, while the cost of privacy protection was low. Fernandez et al. [12] found that privacy protection in the process of using smart meter data for load forecasting is challenging. Therefore, a privacy protection technology was proposed by combining DP and security aggregation. This technology could ensure the security of privacy while achieving high-level information sharing. Qingkui et al. [13] found that local update files uploaded by edge servers may pose a risk of privacy leakage. Therefore, they proposed a protection technology combined with DP to enhance the privacy and confidentiality of raw data before transmission to edge servers through DP. This new technology had superior security performance compared to gradient purification. To optimize the location privacy of end-access users, Zoltánádám Mann et al. [14] proposed a novel LPP approach after studying the graph team combined with edge computing and cloud resources. The experimental results showed that the approach was faster and more cost-efficient in processing. Dave M et al.

proposed a video surveillance system based on distributed edge fog nodes in response to the problem that information collected by security surveillance at home or office may be used to initiate cyber fraud. The system performed reversible fuzz testing on privacy-sensitive objects detected in captured video streams, resulting in a solution that consumes fewer resources in terms of performance and provides better privacy protection [15]. Gulati S et al. implemented a federated learning environment consisting of multiple clients using MobileNetV2 as the backbone deep learning model. The model was trained on a combination of 2 datasets obtained from the Kaggle repository, including color fundus images labeled as diabetic retinopathy, diabetic macular edema, and normal cases. Experimental results showed that the raw data of the model were never shared with the central server, which improved detection efficiency and effectiveness while ensuring data privacy and security [16]. The comprehensive analysis of the common themes, challenges and advances in the above references is shown in Table 1.

Table 1 shows that the adoption of DP and other privacy-preserving techniques can significantly enhance the effectiveness of privacy preservation in dealing with large-scale location information accumulation. This improves the integrity of source location privacy preservation in wireless sensor networks, location privacy of social network users, and location service query privacy preservation. Nonetheless, there is still a research gap in the existing studies on the application of combining DP with personalized location privacy preservation, which suggests that future research can explore new models and methods for combining the two in different data dissemination scenarios.

3 Construction of Personalized LPP Model Combining DP Mechanism in DDPS

For the design of the LIP protection model under Dynamic Data Publishing Scenarios (DDPS), the first section first introduces DP and Crowd Movement Characteristics (CMC). The second and third sections respectively introduce how protection models are constructed in Static Data Publishing Scenarios (SDPS) and DDPS. All parameter symbols in the method are defined as shown in Table 2.

3.1 DP Mechanism and Crowd Movement Characteristics

The DP mechanism protects personal privacy by introducing noise to the data. It can also offer statistical insights into the movement patterns and trends of a

Table 1 Synthesis of the literature

References	Research Direction	Used Methods	Main Findings	Limitations and Challenges	Research Gap
Liu X et al. [5], Parashar A et al. [9]	Dynamic Data Privacy Protection Model	Group effect aggregation; Texture modification	Effective handling of location information accumulation.	Real-time vs. privacy-preserving balance.	Incorporate individualized service needs.
Chen Y et al. [6]	Source LPP in wireless sensor networks	Sector phantom routing	Improve source LPP stability and accuracy.	Complex network environment applications.	More protection strategies.
Huo Y H et al. [7]	Social network user location privacy	Bipartite graph clustering algorithm	Enhancing location privacy for social network users.	Maintaining privacy efficiency in open environments.	Long-term protection mechanisms.
Zhang Q et al. [8], Gulati S et al. [16]	Location service query privacy protection	DP enhancement; Federated learning	Protect user location and query privacy.	Handling high-dimensional data and complex queries.	Comprehensive protection programs.
Yin K et al. [10], Fernandez J D et al. [12], Qingkui Z et al. [13], Zoltán ádám Mann et al. [14]	Static data privacy protection model	Adaptive batch processing; Secure aggregation; Privacy-enhanced federated learning	DP is valid in static data protection.	Increase the level of protection without sacrificing data utility.	Widely used technologies.
Cheng X F et al. [11], Dave M V et al. [15]	LPP in deep learning	Rényi DP SGD; Reversible blurring via edge nodes	Providing privacy protection in deep learning tasks.	Cost control.	More efficient protection algorithms.

Table 2 Parameter symbol

Notation	Account	Notation	Account
D	Randomized data set	u_m	Number of users at base station m.
M	The Laplace mechanism	E_i	Feedback error
ε	Scale parameter	θ_1	Integral gain factor
Δf	Sensitivity	θ_D	Differential gain factor
f	Functional characterization of the dataset	θ_p	Proportional Gain Factor
$Range$	The output set of the query function	D_i	Data situation of sampling points under time point i.
r	Output item	D_j	Latest release of sampling perturbation data
Δq	Sensitivity of the scoring function	C_l	Remaining capacity of current sampling point
i	First time division point	C	Current number of sampling points
j	Second time division point	S‘	Total number of sampling points to be taken
$aveDis$	Average of one-paradigm distances at adjacent times	S_l'	Remaining volume of total sample points to be taken
$simVar$	Data Dispersion at Time i and Time j.	ε_j	Privacy budget values assigned to sampling point j
L_1	The first data to introduce Laplace noise	u_g	Aggregated data values after noise addition is complete
S	Total number of time periods divided in a day	u_i	Per-position estimates of linear regression predictions
H_0	Histogram of raw data	x	Received perturbation data
H_1	Histogram of data after introducing noise	β	Learning rate
H_2	Histogram after the action of the reprocessing mechanism	θ	Parameters
U_m	Hypothetical function for region m.	/	/

crowd without disclosing the precise location of an individual [17]. Figure 1 is the schematic diagram of the DP mechanism.

In Figure 1, before collecting raw data, DP processing can be used to perturb the raw data, such as adding noise. It can achieve the effect of hiding the original data while avoiding attacks from other speculators. There are two common mechanisms for implementing DP, namely the Laplace mechanism and the exponential mechanism [18]. The Laplacian mechanism protects data

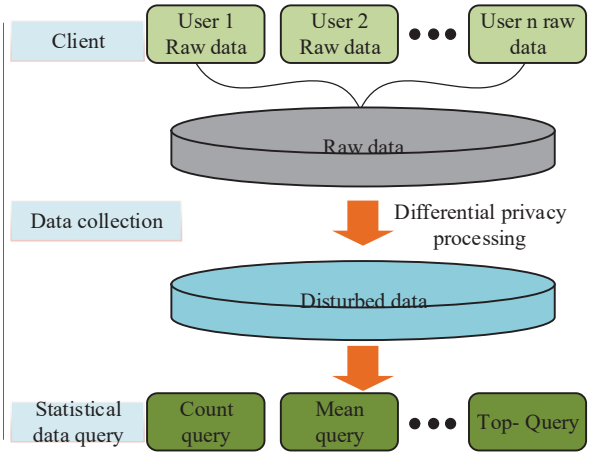

Figure 1 Schematic diagram of DP mechanism.

privacy by adding noise with Laplacian distribution to the query result. In this study, the Laplace mechanism is mainly chosen as the core privacy protection tool. This choice is mainly due to the fact that the cabin trajectory data and simulated crowd motion data used are both continuous numerical data. The Laplace mechanism can directly add noise to numerical statistical results, which not only protects privacy but also maintains the continuity characteristics of the data. The exponential mechanism, on the other hand, is usually suitable for scenarios where the output is discrete choice items, and its application to the dataset of this study will lead to over-discretization of the data, affecting the accuracy of the subsequent analysis. The formula of Laplace distribution is shown in Equation (1).

$$M(D) = f(D) + Lap\left(\frac{\Delta f}{\varepsilon}\right) \tag{1}$$

In Equation (1), D represents a random dataset. M represents the algorithm. ε represents the scale parameter. Δf represents sensitivity. f represents the functional representation of the dataset. If M satisfies the DP of ε, the larger the Δf or the smaller the ε, and the greater the Laplace noise. The exponential mechanism, on the other hand, protects data privacy by adding noise with an exponential distribution to the query results. The calculation formula for the index mechanism is shown in Equation (2).

$$M(D,q)\left\{r|\Pr[r \in Range] \propto \exp\left[\frac{\varepsilon q(D,q)}{2\Delta q}\right]\right\} \tag{2}$$

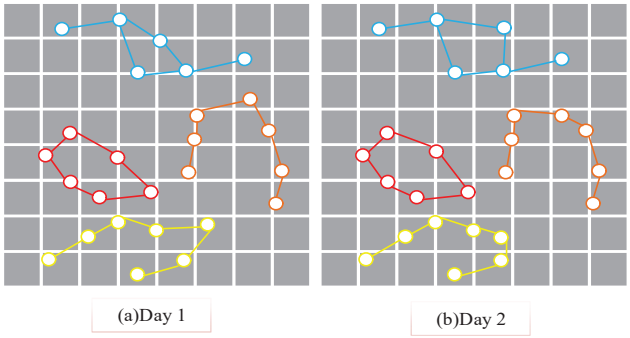

Figure 2 Characteristic map of crowd movement pattern for 2 consecutive days.

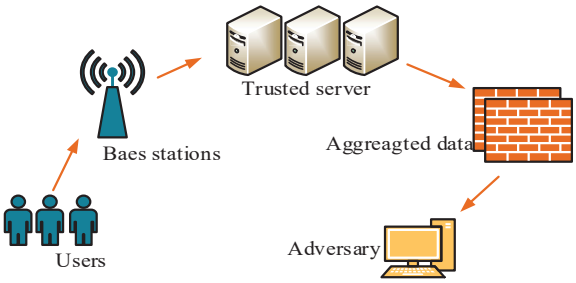

Figure 3 Mobile aggregated data system model.

In Equation (2), $Range$ represents the output set of the query function. r represents the output item. Δq represents the sensitivity of the scoring function. Similarly, if M satisfies the DP of ε, the smaller the ε, the smaller the probability of selecting data. Figure 2 is an illustration of CMC.

Figures 2(a) and 2(b) show the movement trajectories of five random users on the first and second days. The two day movement trajectories of these 5 users are roughly similar, indicating that CMC has certain spatio-temporal regularity, periodicity, and uniqueness.

3.2 Construction of LPP Model under SDPS

SDPS refers to selecting a fixed period for data filtering and crowd movement location analysis in a specific population environment. Figure 3 shows the user mobile aggregation data publishing system in this scenario.

In Figure 3, to safeguard the privacy and security of the user's location data, this study analyzes the length of time a user stays at a particular location at different times of the day. It aims to cope with attack methods that exploit

the regularity and uniqueness of user movements. The specific manifestation is frequent movement during the day, while the range of movement at night is relatively small. A novel LIP protection model under SDPS is proposed based on the DP mechanism for this time segmentation. The calculation formula for the time zone division of this model is Equation (3).

$$U(i,j) = \log_a(j-i+1) \times \frac{aveDis}{simVar} \tag{3}$$

In Equation (3), i and j represent the 1st and 2nd time division points. $aveDis$ represents the average value of a norm distance between adjacent times. $simVar$ represents the degree of data dispersion within time i and time j. To calculate $aveDis$, Equation (4) is further introduced to calculate the average distance between two neighboring sampling points, defined as follows:

$$aveDis = \frac{1}{j-+1} \sum_{t=i}^{j-1} L_1(D_t, D_{t+1}) \tag{4}$$

In Equation (4), L_1 represents the first data to introduce Laplace noise. The larger the $aveDis$ value, the faster the movement of aggregated data changes, making it more suitable to use direct interference for data published during that time period. To reflect the consistency of the data changes over the time period, a defined data similarity variance $simVar$ over the time period is introduced as a measure of the degree of change.

$$simVar = \frac{1}{j-i+1} \sum_{i=1}^{n} |aveDis - L_1(d_t, D_{t+1})| \tag{5}$$

When the variance value of $simVar$ is relatively small and the average value is relatively large, it indicates that the distance of a norm within that time period is too large. According to the division criteria, when the changes within a time period are small, direct perturbation should be used. Therefore, the calculation formula for the variance threshold condition satisfied by further establishing the time period division is shown in Equation (6).

$$\log_a S = \begin{cases} simVar \geq 1 \\ j-i+1 \leq S \end{cases} \tag{6}$$

In Equation (6), S represents the total number of time periods divided in a day. After segmentation, a noise smoothing mechanism is designed to

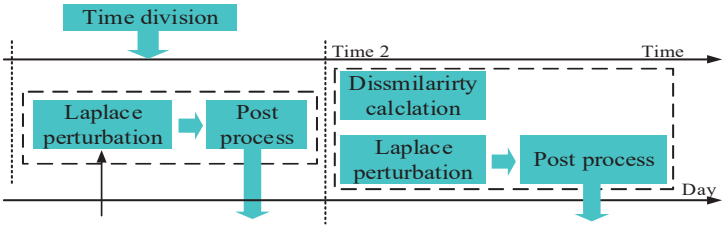

Figure 4 A model framework for protecting location privacy in SDPS.

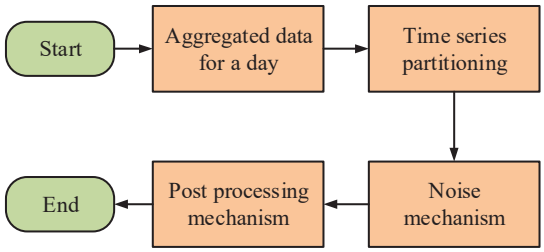

Figure 5 Flow of the static data privacy protection model.

enhance the practicality of post-processing data with noise. Noise smoothing requires controlling the variance between the perturbed data and the original data within a certain range, and the calculation formula is shown in Equation (7).

$$\|H_1 - H_0\|_1 \leq \|H_2 - H_0\|_1 \tag{7}$$

In Equation (7), H_0, H_1, and H_2 represent the histograms of the original data, the data after introducing noise, and the processing mechanism, respectively. In summary, the model framework is shown in Figure 4, which combines time zone division, noise mechanism, and post-processing mechanism.

In Figure 4, by dividing the data distribution characteristics of the mobile aggregation dataset into time regions, different noise mechanisms are introduced to interfere with the data in different time regions. Then, the post-processing mechanism is further introduced to improve the data utility after noise processing. The flow of the static data privacy protection model is shown in Figure 5.

In Figure 5, the time division mechanism is used to calculate the data that needs to be published within each time period, and then Laplace noise is introduced to interfere with the data, smoothing the noisy data after interference. For data within a non-predetermined time period, threshold judgment

is first performed to determine whether to introduce noise for perturbation based on the judgment results.

3.3 LPP Model Construction under DDPS

To ensure the real-time nature of data and prevent leakage issues during data transmission and sharing, this study conducts in-depth research on the LIP data real-time publishing module for each user. Assuming that the statistical information of mobile users in each base station in the city is published in real-time. The real-time data flow publishing process is shown in Figure 6.

Figure 6(a) demonstrates the budget allocation process, which mainly includes four modules: adaptive sampling, packet merging, post-processing, and privacy budget allocation. The number of sampling points is dynamically adjusted by the adaptive sampling module. DP noise processing and reasonable budget allocation are introduced after packet merging to ensure flexibility and real-time privacy protection. The data prediction process is shown in Figure 6(b). On the basis of perturbed data after publication, data denoising and time series prediction are carried out sequentially to correct the bias caused by noise and improve the accuracy and stability of subsequent data utilization. Predictive data indicate that by using this data, the number of denoising time periods for users in the next few days, i.e. the number of sampling points, can be calculated. The calculation formula for the prediction module is shown in Equation (8).

$$U_m(x) = \theta_1 + \theta_2 x \tag{8}$$

In Equation (8), U_m represents the assumed function of the m region. θ represents the parameter. x represents the received perturbation data.

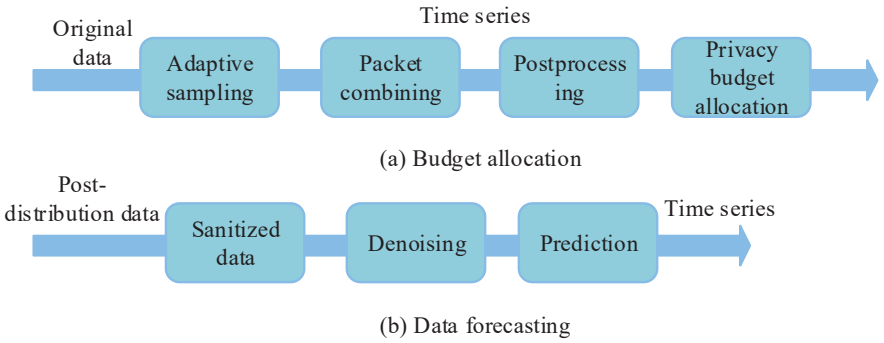

Figure 6 Data stream real-time publishing model.

During the training process of the measurement model, the gradient descent method is used to update the parameter θ, which is calculated as shown in Equation (9).

$$\theta = \theta + \beta(x)^U(x \cdot \theta - u_m) \tag{9}$$

In Equation (9), β represents the learning rate. u_m represents the number of users in the m-th base station. The prediction error metric is calculated using the error term in the manner of Equation (10).

$$\delta_i = \theta_p \times E_i + \theta_1 \times \sum_{n=t-1}^{t} + \theta_D \times E_{i+1} \tag{10}$$

In Equation (10), E_i represents the feedback error. θ_1, θ_D, and θ_p represent the integral gain factor, differential gain factor, and proportional gain factor, respectively. The expression for calculating E_i is shown in Equation (11).

$$E_i = |D_i - D_j| \tag{11}$$

In Equation (11), D_i represents the sampling point data at time point i. D_j represents the latest released sampled perturbation data. In summary, the adaptive sampling module can achieve reasonable allocation of privacy budget in disturbed data. The schematic diagram of the privacy budget allocation module combined with adaptive allocation is shown in Figure 7.

In Figure 7, compared to the traditional PID control allocation method, the dynamic data intelligence privacy budget allocation method proposed in this study can flexibly adjust the allocation of intelligence privacy budget based on the number of current sampling points. The benefit of this approach is its ability to dynamically adjust and allocate privacy budgets based on the

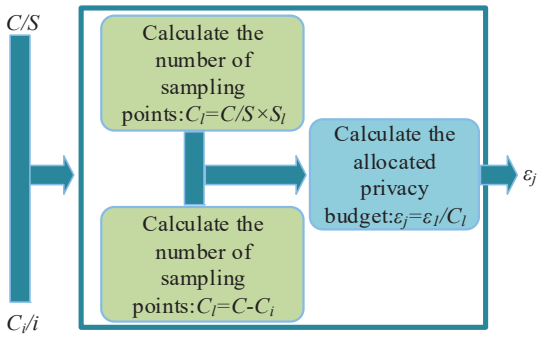

Figure 7 Dynamic data privacy budget allocation diagram.

sampling rate capacity promptly [19–21]. To realize the dynamic allocation of the privacy budget, the following parametric relation equation is introduced.

$$C_l = \frac{C}{S'} \times S'_l \tag{12}$$

In Equation (12), C_l represents the remaining amount of the current sampling point. C represents the current number of sampling points. S' represents the total number of sampling points to be collected. S'_l is the remaining quantity of the total sampling points to be collected. Based on this, the privacy budget allocation for each sampling point is defined as shown in Equation (13).

$$\varepsilon_j = \frac{\varepsilon_1}{C_l} \tag{13}$$

In Equation (13), ε_j represents the privacy budget value assigned to sampling point j at time point t. Ultimately, the aggregation of perturbation data is processed using linear weighting, calculated as shown in Equation (14).

$$u_g = \frac{u_i}{l} \tag{14}$$

In Equation (14), u_g represents the aggregated data value after denoising is completed. u_i represents the estimated value of each position predicted by linear regression. The flow of the dynamic data privacy protection model is shown in Figure 8.

In Figure 8, under DDPS, this study extracts data features based on users' daily location information and predicts aggregated data through a prediction module. Then, the adaptive allocation module is used to determine whether each segment of data is subjected to noise processing and allocate an appropriate privacy budget.

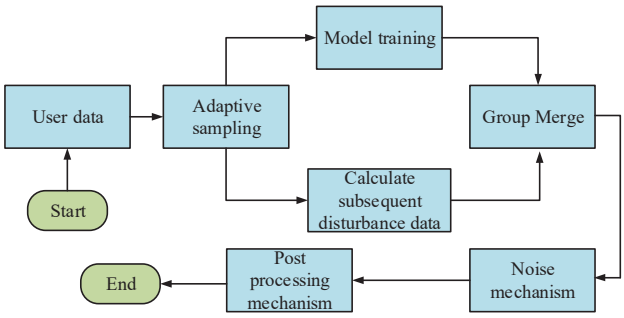

Figure 8 Process for dynamic data privacy protection model.

4 Performance Testing of LPP Model

The first section of the study tested the static data perturbation model performance using trajectory recovery rate, Mean Absolute Error (MAE), Mean Relative Error (MRE), and privacy budget. The second section then tested the dynamic data perturbation model using privacy budget, MAE, MRE, temporal resolution, and spatial resolution as metrics. Both models were compared with traditional and advanced data modification algorithms to validate the effectiveness of the models. Cross-validation techniques were used in the experiments to ensure the stability and reliability of the results. Data preprocessing included cleaning invalid data, standardizing coordinate and timestamp formats, and calibrating trajectory data using GIS techniques. For static data protection, different perturbation parameters, such as Laplace noise ratio and distribution, were set according to the dataset characteristics. For dynamic data protection, the parameters of the prediction module integrated the historical data to maximize accuracy, and the parameters of the adaptive module and group merging module were adjusted according to the characteristics of the dynamic data stream. The computational formula of MAE is shown in Equation (15).

$$MAE = \frac{1}{n}\sum_{i=1}^{1}|y_i - \hat{y}_i| \tag{15}$$

In Equation (15), n denotes the number of samples. y_i denotes the actual observation. $\hat{y}_i$ denotes the actual counterpart. The MRE is calculated as shown in Equation (16).

$$MRE = \frac{1}{n}\sum_{i=1}^{1}\frac{|y_i - \hat{y}_i|}{|y_i|} \tag{16}$$

In Equation (16), all algebraic meanings are consistent with previous explanations.

4.1 Performance Testing of LPP Model Under SDPS

Experiments were conducted equipped with Intel®CoreTMi7-9700, CPU@ 3.00GHz×16, and GPU of NVIDIA GeForce RTX 3060Ti. The trajectory data of nearly 10,000 taxis in Chengdu from October 2022 to November 2023 were included. This dataset contained a total of approximately 24 million location points, with a total trajectory mileage of 10 million kilometers

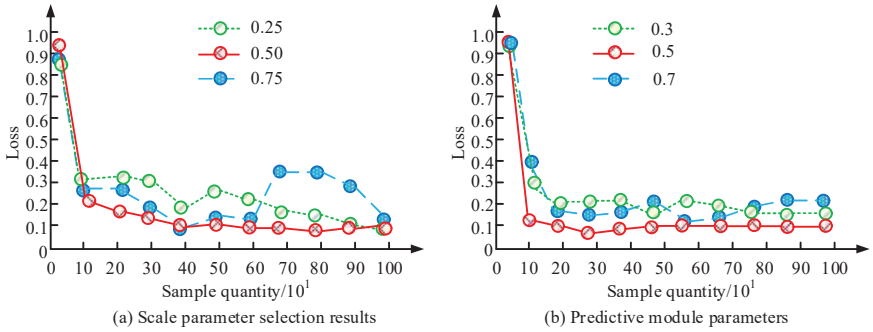

Figure 9 Hyperparameter selection test results.

and an average sampling interval of 4 minutes. The second dataset was the urban population movement trajectory generated by software simulation, which included the stopping points of nearly 150,000 people in Chengdu within a day, with an average of 3.4 per person. The study first tested the hyperparameter selection of the scale parameters and the parameters of the prediction module, and the results are shown in Figure 9.

Figure 9(a) shows the test results of scale parameter selection and Figure 9(b) shows the test results of prediction module parameter selection. From Figure 9(a), when the scale parameter is set to 0.5, the overall loss value of the model is the smallest and the convergence speed is the fastest at different sampling times. In contrast, when the value of the scale parameter is 0.25 or 0.75, the initial loss value decreases slower and fluctuates when the number of samples is higher. This indicates that too small or too large a scale parameter can lead to a decrease in the stability of the model. Therefore, the scale parameter is finally set to 0.5. From Figure 9(b), in the parameter test of the prediction module, when the parameter value is 0.5, the loss function also shows the best convergence and the smallest stable fluctuation range. When the parameter takes the value of 0.3 or 0.7, the loss value fluctuates more after the number of samples is increased, which affects the overall prediction accuracy, although the initial decline rate is faster. Taking the model performance and stability into account, the parameter of the prediction module is finally selected as 0.5. Using trajectory recovery rate as a reference indicator, the static data privacy protection model and traditional data modification algorithms proposed in the study, such as hash algorithm, Gaussian algorithm, and Poisson algorithm, are tested on two datasets. These algorithms have a wide range of applications in the LPP field, representing typical methods in this field and serving as benchmarks for evaluating the performance and

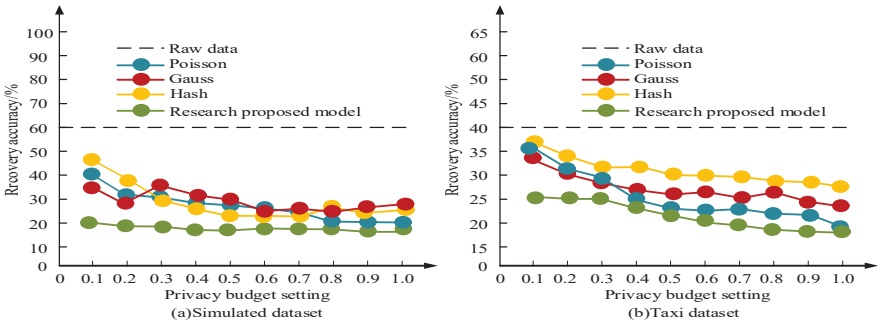

Figure 10 The trajectory recovery rate of different models.

effectiveness of proposed privacy protection models in existing technologies. Among them, the trajectory recovery rate is used to evaluate how well the DP mechanism protects the trajectory data. A lower trajectory recovery rate indicates that it is more difficult to recover the original trajectory from the published privacy-protected data. The test results are shown in Figure 10.

Figures 10(a) and 10(b) show the position trajectory recovery results of different models in simulated data and taxi data. From Figure 10, the hash algorithm has the highest trajectory recovery rate and the recovered trajectory is closest to the original data in both datasets, indicating that its privacy preserving ability is weak. In contrast, the proposed static data privacy protection model exhibits the lowest trajectory recovery rate on both datasets, where the lowest recovery rate is 18% on the simulated dataset. The lower the recovery rate, the greater the difficulty for attackers to infer the original trajectory from perturbed data, and the better the privacy protection effect. Therefore, the results validate the effectiveness of the proposed model in enhancing the privacy protection effect of static data locations. It is worth noting that under the condition of fixed privacy budget ε, the difference in the recovery rate reflects the difference in the trade-off between perturbation strength and data availability among the models. The privacy budget is set to a fixed value, while MAE and MRE are used as reference indicators to explore the impact of threshold changes on model testing. Both MAE and MRE are used to assess the degree of difference between the privacy-protected data and the original data. As depicted in Figure 11, a smaller value indicates a more effective DP mechanism in safeguarding the original data.

Figures 11(a) to 11(d) show the MAE and MRE test results of the four models in the simulation dataset and taxi dataset. In Figure 11, the MAE and MRE values of all four models gradually decrease as the threshold increases,

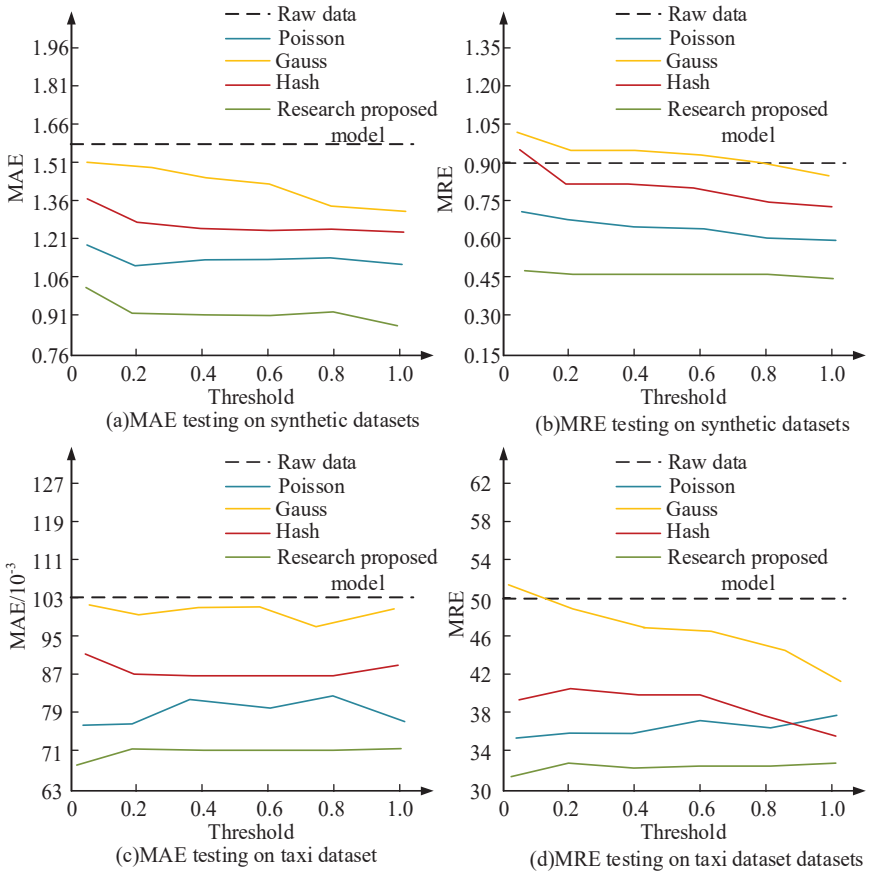

Figure 11 MAE and MRE test results under different data.

and there are some data fluctuations. Among them, the proposed static data privacy protection model performs optimally on both datasets, with the lowest MAE value on the simulation dataset of about 0.8 and the lowest MRE value of about 0.45. The lowest MAE value on the cab dataset is about 0.067 and the lowest MRE value is about 32. This indicates that when the privacy budget ε is high, injecting data noise is reduced. The protection strength of the model for raw data decreases, but the data utility increases and the error metric (MAE/MRE) decreases. While in the case of a small privacy budget, the error metric becomes larger due to the large amount of noise injection, reflecting the typical trade-off between privacy protection and data utility. The threshold is set to a fixed value, and MAE and MRE are also used as

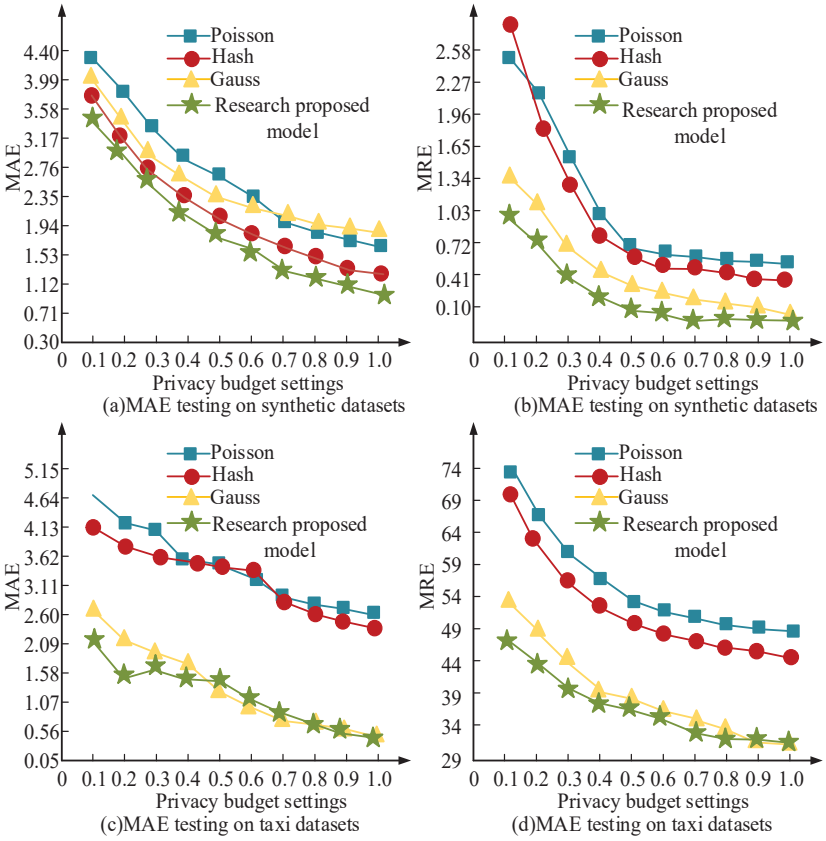

Figure 12 Privacy prediction results of different models in different datasets.

reference indicators to continue exploring the impact of privacy budgeting on the four intervention models, as shown in Figure 12.

Figures 12(a) to 12(d) show the privacy budget and MAE, privacy budget, and MRE test results of four models in the simulation dataset and taxi dataset. From Figure 12, the MAE and MRE values of all models continue to decrease as the privacy budget ε gradually increases. This trend is consistent with the DP theory, which states that as the privacy budget increases, the allowed level of privacy leakage also increases. This correspondingly reduces the injected data noise, resulting in lower data errors and higher data availability in the model output. Under the condition of fixed threshold, the lowest MAE value is 1.12 and the lowest MRE value is 0.05 on the simulated dataset. On the cab dataset, the lowest MAE value is 0.56 and the lowest MRE

Table 3 Indicator test results of different models

Data Set	Modle	MAE	MRE	Resource Consumption Rate/%	Privacy Protection Intensity/%
Simulated dataset	GAN	1.08	1.08	57.6	89.7
	HE	1.02	0.58	54.3	88.4
	VAE	0.96	0.27	37.2	86.2
	Our modle	0.89	0.15	31.8	92.1
Taxi dataset	GAN	0.89	36.5	27.6	81.5
	HE	0.71	39.6	25.8	85.4
	VAE	0.64	31.2	24.5	87.6
	Research model	0.58	29.2	22.1	91.3

value is 31. The difference in the performance of the different models with different values of ε mainly stems from the differences in the design of their respective perturbation mechanisms. The proposed static data privacy-preserving model under most privacy budget configurations maintains a lower error, reflecting better stability and data utility balancing ability. The study introduces the latest methods for static data privacy protection such as Generative Adversarial Networks (GAN), Homomorphic Encryption (HE), and Variational Auto-encoders Encoder (VAE) for test. To ensure the fairness of model performance comparison, all models (including the research methods, GAN, HE, and VAE) are implemented and tested in the same experimental environment. The experimental datasets include the urban cab trajectory dataset and the urban crowd movement trajectory dataset. The test results are shown in Table 3.

From Table 3, the values of the test metrics of the four models under the two types of datasets did not differ much, but in general, the research model was better. In particular, under the simulation dataset, the privacy protection strength of the proposed model could reach up to 92.1% and the MRE could reach down to 0.15. While under the cab dataset, the MAE of the proposed model could reach down to 0.58 and the resource consumption rate could reach down to 22.1%. The above data showed the superiority and effectiveness of the proposed method among many state-of-the-art methods.

4.2 Performance Testing of LPP Model Under DDPS

The proportional parameters θ_1, θ_D, and θ_p of the control mechanism in the adaptive module took values of 0.2, 0.1, and 0.7, respectively. The spatial

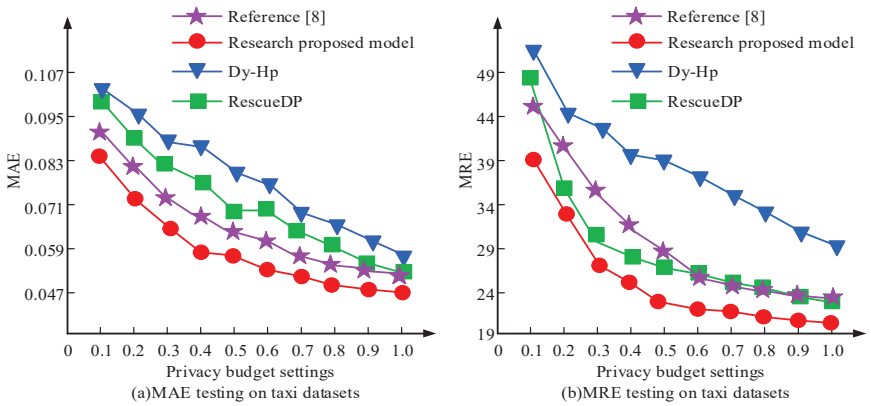

Figure 13 The impact of privacy prediction on data utility on taxi datasets.

resolution was set to 500 m * 500 m, and data sampling and aggregation statistics were conducted every half hour. Using MAE and MRE as reference indicators, the dynamic data privacy protection model proposed in the study was compared with the linear regression data protection model (Dy-HP), infinite flow protection model (ResueDP), and similar dynamic data protection approaches in the literature [8]. Among them, Dy-HP utilized historical data to build a linear model through which the user's movement trajectory or location information could be predicted. While RescueDP responded to different privacy needs and data sensitivities by dynamically adjusting the privacy protection level in the data stream. In addition, the privacy budget was used to measure the privacy protection strength of the DP mechanism. A larger privacy budget meat more privacy leakage was allowed, and the privacy budget under this dataset was useful for each model data. The test results are shown in Figure 13.

Figures 13(a) and 13(b) show the relationship curves between privacy budget and MAE, privacy budget and MRE for three models. The MAE and MRE values of the research model were much lower than those of the other two models. When the privacy budget was 1, the minimum MAE was about 0.047, and the minimum MRE was about 20. The reason might be that the research model could associate the geographic location information of mobile aggregated data through grouping and merging, fully considering the statistical data of neighboring regions. On the contrary, the RescueDP model only considered the similarity of regions with smaller statistics for denoising. The test results of the methods in the literature [8] for MAE and MRE were comparable to those of the proposed model. This was because the proposed

Table 4 Test results for different temporal and spatial resolutions

Spatial Resolution	Time Resolution	MAE	MRE
/	15min	0.13	35
/	30min	0.07	25
/	45min	0.03	23
0.5 × 0.5 km	/	0.05	21
1 × 1 km	/	0.06	25
1.5 × 1.5 km	/	0.07	30

model utilized feature extraction and thresholding optimization techniques based on the literature [8] to enhance the prediction efficiency of dynamic data. In addition, the study continued to test the three identified temporal and spatial resolutions using the number of sampling points as a reference metric. The temporal resolutions were 15 min, 30 min, and 45 min, and the spatial resolutions were 0.5 × 0.5 km, 1 × 1 km, and 1.5 × 1.5 km. The temporal resolution was used to measure the granularity or resolution of the data release in the time dimension. Higher temporal resolution indicated shorter time intervals for data release, while lower temporal resolution indicated longer intervals. Spatial resolution was used to measure the granularity or resolution of data distribution in the spatial dimension. A smaller spatial extent was indicated by a higher spatial resolution, and vice versa for a larger spatial extent. The test results are shown in Table 4.

In Table 4, the data for all the metrics performed well with 45 min as the temporal resolution. This data illustrated that with a fixed privacy budget and spatial resolution, the smaller the temporal resolution, the greater the total number of subdivided time periods in a day, and the greater the total number of sampling points. Therefore, the average amount of privacy budget allocated to each sampling point decreased, and the noise injected into the sampled data increased, making the utility performance of the final noise-processed data degrade. In addition, when the spatial resolution was 0.5 × 0.5 km, the MAE and MRE values were relatively smallest, with minimum values of 0.55 and 21, respectively. This data illustrated that when the privacy budget was determined with a fixed temporal resolution, the larger the spatial resolution was, the lesser the number of user statistics in the region, which resulted in the abatement of the resistance to noise in each region. On the contrary, the smaller the resolution, the more resistant to noise each region was, and the less help was needed from the grouping and merging module. Therefore, the utility of the noisy data was better at this point.

5 Conclusion

With the rapid development of mobile Internet technology, the release of personalized location data may pose a potential threat to the privacy of individuals. Given this, the study took DDPS as the starting point and proposed two different scenarios of location privacy intervention protection models in combination with DP mechanisms. The results showed that the static data model had a minimum trajectory recovery rate of 18%, and the dynamic data model had a minimum MAE of about 0.047. In addition, the significant reduction of the trajectory recovery rate and the data optimization of the MAE and MRE reflected the effectiveness of the model for privacy protection, and these metrics were the key to the success of the LPP scheme. For the dynamic data privacy preserving model, the study found that a temporal resolution of 45 minutes and a spatial resolution of 0.5×0.5 km were optimal, and these findings provided important insights for understanding and improving real-time location data processing. The determination of temporal and spatial resolution not only affected the efficiency of data processing but also directly correlated to the strength of privacy protection and data utility. Despite confirming the superiority and feasibility of the two types of models, they provide effective theoretical support for location data distribution methods with good performance and privacy preservation. However, the centralized DP mechanism used in the study still has its limitations, especially in scenarios where the service provider is not fully trusted. Therefore, future research directions should include exploring the application of localized DP, which not only further enhances the privacy protection of data, but also may overcome some of the limitations in centralized processing. For example, how to maximize privacy protection while maintaining data utility, and how to design efficient localized data processing algorithms. In summary, the results not only demonstrate the effectiveness of these models but also provide valuable insights and directions for future research, especially in further exploring localized DP mechanisms.

Acknowledgment

The research is supported by the Scientific and Technological Research Project of Chongqing Municipal Education Commission, “Research on Personal Location Differential Privacy Protection under Big Data Intelligence” (No. KJQN202303602).

References

[1] Shin D, Hwang Y. The effects of security and traceability of blockchain on digital affordance. Online Information Review, 2020, 44(4):913–932.
[2] Tran A T, Luong T D, Karnjana J, Van-Nam H. An efficient approach for privacy preserving decentralized deep learning models based on secure multi-party computation. Neurocomputing, 2021, 422(1):245–262.
[3] Wu J, Wang T, Li G, Yu K, Luo C. Secure storage scheme of trajectory data for digital tracking mechanism. International journal of intelligent systems, 2022, 37(12):12490–12510.
[4] Cheng Z, Ye D, Zhu T, Zhou W, Yu P, Zhu C. Multi-agent reinforcement learning via knowledge transfer with differentially private noise. International Journal of Intelligent Systems, 2022, 37(1):799–828.
[5] Liu X, Chen Y. Group effect-based privacy-preserving data aggregation for mobile crowdsensing. Computer Networks, 2023, 222(2):1–18.
[6] Chen Y, Sun J, Yang Y, Li T, Niu X, Zhou H. PSSPR: A source location privacy protection scheme based on sector phantom routing in WSNs. International Journal of Intelligent Systems, 2022, 37(2):1204–1221.
[7] Huo Y H, Chen B, Tang J, Zeng Y. Privacy-preserving point-of-interest recommendation based on geographical and social influence. Information Sciences, 2021, 543(3):202–218.
[8] Zhang Q, Zhang X, Wang M, Li X. DPLQ: Location-based service privacy protection scheme based on differential privacy. IET information security, 2021, 15(6):442–456.
[9] Parashar A, Shekhawat R S. Protection of gait data set for preserving its privacy in deep learning pipeline. IET Biometrics, 2022, 11(6):557–569.
[10] Yin K, Ding Z, Yang X, Tan Z, Zhu R, Ji R, Wang Z, Yin G. Sleep Staging Method for Imbalanced EEG Data Based on Differential Privacy Federated Learning. International Journal on Artificial Intelligence Tools, 2022, 31(6):2240018–2240039.
[11] Cheng X F, Yao Y Q, Zhang L, Liu A, Li Z. An improved stochastic gradient descent algorithm based on Rényi differential privacy. International journal of intelligent systems, 2022, 37(12):10694–10714.
[12] Fernandez J D, Potenciano S, Lee C M, Rieger A, Fridgen G. Privacy-preserving federated learning for residential short-term load forecasting. Applied energy, 2022, 326(15):174–187.
[13] Qingkui Z, Liwen Z, Zhuotao L, Huakun H, Jungyoon K. Privacy-Enhanced Federated Generative Adversarial Networks for Internet of Things. The Computer Journal, 2022, 65(11):2860–2869.

[14] Zoltán ádám Mann, Metzger A, Prade J. Cost-Optimized, Data-Protection-Aware Offloading Between an Edge Data Center and the Cloud. IEEE transactions on services computing, 2023, 16(1):206–220.
[15] Dave M, V. Rastogi, M. Miglani, P. Saharan, and N. Goyal. "Smart fog-based video surveillance with privacy preservation based on blockchain," Wirel. Pers. Commun., vol. 124, no. 2, pp. 1677–1694, November, 2022, DOI: 10.1007/s11277-021-09426-8.
[16] S. Gulati, K. Guleria, and N. Goyal. "Privacy-preserving and collaborative federated learning model for the detection of ocular diseases," Int. J. Math. Eng. Manag. Sci., vol. 10, no. 1, pp. 218–248, February, 2025, DOI: 10.33889/IJMEMS.2025.10.1.013.
[17] Song Z, Wu K, Shao J. Destination prediction using deep echo state network. Neurocomputing, 2020, 406(17):343–353.
[18] Y. Dang, "Research on Collaborative Positioning Algorithm of Wireless Sensor Network Security Under Strong Topology Relation", JCSANDM, vol. 14, no. 01, pp. 25–46, Feb. 2025.
[19] J. K. Dawson, F. Twum, J. B. H. Acquah, and Y. M. Missah, "Cryptographic Solutions for Data Security in Cloud Computing: A Run Time Trend-based Comparison of NCS, ERSA, and EHS", JCSANDM, vol. 13, no. 02, pp. 265–282, Feb. 2024.
[20] Wan W, Cai M. Phone-vehicle trajectory matching framework based on ALPR and cellular signalling data. IET Intelligent Transport Systems, 2021, 15(1):107–118.
[21] Hidayat I, Ali M Z, Arshad A. Machine Learning-Based Intrusion Detection System: An Experimental Comparison. Journal of Computational and Cognitive Engineering, 2022, 2(2):88–97.

Biography

Hongying Tan obtained a bachelor's degree in computer science and technology from Chongqing University in 2008 and a master's degree in library, information and archive management from the same university in 2014. She is currently employed at Chongqing Industry and Trade Polytechnic, and her research focuses mainly on big data analysis technology, and artificial intelligence. She has led the research project "Archives Management Based on Big Data Technology" at Chongqing Industry & Trade Polytechnic, and participated in the horizontal project "Research on Key Technologies for Remote Intelligent Recognition and Monitoring of Public Passenger Vehicle Driving Behavior" funded by the Chongqing Municipal Science and Technology Commission. She has also co-authored one professional textbook, and published six research papers, two of which are in core journals indexed by Beijing University.