Privacy Protection Method for Vehicle Trajectory Based on DP-BiLSTM

Caiyun Chen[1], Lianfen Huang[2,3,4], Zufeng Weng[2], Wu Zheng[1]*

[1] School of Artificial Intelligence and Computer Science, Xiamen University Tan Kah Kee College, China
[2] School of Electronic Science and Technology, Xiamen University Tan Kah Kee College, China
[3] The Key Laboratory of Intelligent Manufacturing Equipment and Industrial Internet Technology, Fujian Provincial Universities, China
[4] Department of Informatics and Communication Engineering, Xiamen University, China
chency@xujc.com, lfhuang@xmu.edu.cn, wengzf@xujc.com, zhengwu@xujc.com

Abstract

In the context of the rapid advancement of Internet of Vehicles (IoV) technologies, the accumulation of large scale vehicle trajectory data has emerged as a crucial cornerstone for intelligent transportation services. However, this development also engenders increasingly severe risks of privacy disclosure and poses substantial challenges to data security governance. To tackle the problems of high privacy disclosure risks associated with trajectory data and attain a balance between the intensity of privacy protection and data utility, this research presents a differential privacy-bidirectional Long Short-Term Memory (DP-BiLSTM) trajectory privacy protection approach. This method adopts the BiLSTM autoencoder as the core framework, capturing bidirectional temporal dependencies within trajectory data via forward and backward Long Short-Term Memory (LSTM) layers. After trajectory reconstruction, differential privacy protection is implemented through post processing, and the privacy budget ε is employed to flexibly adjust the trade-off between the degree of privacy protection and data utility. Experimental results from the T-Drive trajectory dataset demonstrate that the DP-BiLSTM method exhibits outstanding performance under various privacy budget settings, attaining a comprehensive score of 0.6773, which is significantly superior to benchmark models. The DP-BiLSTM method provides a promising solution for trajectory privacy protection in the field of intelligent transportation.

Keywords: Vehicle trajectory, Differential privacy, BiLSTM, Privacy budget

1 Introduction

With the rapid advancement of intelligent vehicle technology, vehicle road cooperation, location based services, and urban computing, vehicle trajectory data has emerged as a core fundamental data resource underpinning intelligent navigation, traffic flow prediction, road network planning, driving behavior analysis, and public security prevention and control [1-2]. Vehicle trajectories are composed of continuous spatiotemporal points, containing multi-dimensional information such as longitude, latitude, speed, and timestamps, which can directly or indirectly correlate with sensitive private information, including users' personal identities and travel habits. Relevant research has indicated that even when trajectory data is desensitized through operations such as de-identification and generalization, attackers can still achieve user re-identification via methods such as background knowledge mining, spatiotemporal correlation analysis, and trajectory clustering, leading to large-scale privacy leakage [3]. To effectively tackle such severe trajectory privacy risks, researchers have put forward a diverse range of privacy protection technologies for vehicle trajectory data [4-5]. In recent years, the integration of deep learning and differential privacy has become a research focus in the field of privacy computing [6-7]. However, existing methods fail to fully capture bidirectional trajectory context and achieve fine-grained privacy control, while direct noise injection damages spatiotemporal continuity.

To address these issues, this paper proposes a differential privacy-bidirectional Long Short-Term Memory (DP-BiLSTM) trajectory privacy protection model. The main contributions are as follows: Firstly, a bidirectional autoencoder for DP-BiLSTM trajectory privacy protection is constructed. It utilizes bidirectional Long Short-Term Memory (BiLSTM) to simultaneously extract the forward historical information and backward context information of trajectories, thereby improving the feature representation ability and reconstruction accuracy. Secondly, differential privacy is utilized to attain trajectory protection via a post-processing strategy subsequent to trajectory reconstruction. This strategy encompasses risk assessment, dynamic privacy budget allocation, and Laplace noise injection, guaranteeing the privacy security of trajectory data release. Thirdly, a multi-dimensional evaluation system for trajectory privacy protection is established to comprehensively assess the overall performance of the model in accordance with core indicators. Finally, comprehensive experiments are conducted on the actual T-Drive dataset to verify the proposed privacy protection method. The findings demonstrate its effectiveness in

*Corresponding Author: Caiyun Chen; Email: chency@xujc.com
DOI: https://doi.org/10.70003/160792642026112706007

achieving a balance between privacy and utility, presenting a practical privacy protection solution for intelligent transportation and Internet of Vehicles (IoV) scenarios.

2 Related Work

2.1 Trajectory Privacy Protection Technology

Trajectory privacy protection technologies have developed into three major mainstream categories: anonymization, data perturbation, and deep learning based protection. Anonymization and generalization methods are the traditional core techniques for trajectory privacy protection. These methods reduce the individual identifiability of trajectory data through operations such as suppression, blurring, and merging [8-9]. Typical mechanisms encompass k-anonymity and l-diversity, these methods are characterized by their simplicity in implementation and low computational cost, and have been extensively utilized in low dimensional trajectory data scenarios [4, 10-11]. Nevertheless, in high dimensional IoV trajectory data, they are highly dependent on data generalization and aggregation, which readily leads to significant information loss. Furthermore, they are unable to withstand background knowledge attacks and correlation attacks, resulting in limited applicability in scenarios demanding strong privacy protection.

Perturbation based privacy protection methods revolve around differential privacy, which masks the impact of an individual sample on the output by introducing controllable noise into data or model outputs [12-13]. Differential privacy trajectory protection is typically implemented through a fixed noise strategy, where the same noise intensity is uniformly applied to all trajectory points [5, 14]. Although it is straightforward to implement, this method fails to account for the spatiotemporal heterogeneity of trajectories. Consequently, it leads to an excessive presence of noise in low sensitivity locations and insufficient protection in high sensitivity locations, rendering it difficult to achieve a balance between privacy strength and data utility.

Trajectory privacy protection methods based on deep learning leverage models such as autoencoders, generative adversarial networks, recurrent neural networks, and federated learning to accomplish trajectory feature learning and data reconstruction. This approach mitigates the risk of privacy leakage while maintaining statistical characteristics [6-7, 15-17]. These approaches can more effectively preserve the temporal structure and spatial distribution of trajectories. Nevertheless, the majority of models lack rigorous privacy proofs, which makes it difficult to quantify the privacy strength. Additionally, unidirectional network structures have insufficient capabilities for modeling temporal dependencies, resulting in limited reconstruction accuracy.

2.2 Time Series Modeling and Differential Privacy

LSTM networks address the issues of gradient vanishing and gradient explosion in traditional Recurrent Neural Networks (RNNs) via gating structures, and are capable of effectively capturing long term dependencies in sequential data [18]. Consequently, LSTM is extensively employed in the temporal feature mining of IoV trajectory data, and can efficiently extract crucial information such as position variations, speed fluctuations, and temporal evolution patterns of trajectories. BiLSTM is an advanced variant of unidirectional LSTM, which consists of forward and backward LSTM layers. It is capable of simultaneously processing sequential data from both forward and reverse directions, and enhances feature representation by fully exploiting comprehensive contextual information [19]. Leveraging this potent bidirectional modeling capacity, the BiLSTM achieves excellent performance in tasks including trajectory prediction, data completion, and anomaly detection, thereby positioning it as the optimal method for the temporal modeling of IoV trajectories.

Differential privacy is distinguished by rigorous mathematical definitions and composable characteristics, making it a dependable privacy protection framework. Its core mechanism entails the incorporation of controlled noise perturbation to achieve the privacy goal of facilitating data utility while precluding individual identification. Consequently, the integration of differential privacy into deep learning has emerged as a significant research direction for privacy preserving modeling. Deep learning training frameworks that incorporate differential privacy attain model privacy protection via gradient clipping, noise injection, and momentum optimization [13]. Although some researchers have proposed models integrating differential privacy and deep learning for IoV data privacy protection [20], these schemes do not comprehensively consider the inherent characteristics of IoV trajectories, such as spatiotemporal correlation, continuous dependence, and uneven distribution of sensitive information. When directly applied to trajectory privacy protection, these schemes often lead to excessive loss of data utility and unbalanced privacy strength.

In the research regarding the integration of time series modeling and differential privacy, some researchers have proposed the DP-LSTM model and successfully applied it to non-IoV scenarios, such as stock prediction and action recognition [21]. Additionally, certain studies integrate differential privacy and LSTM to improve the quality of IoV data [22]. Nevertheless, these studies do not focus on the core objectives of privacy protection and are incapable of achieving a balance between privacy security and data utility necessary for the publication of IoV trajectories. Generally speaking, existing studies exhibit obvious limitations. Some studies only focus on the temporal modeling of trajectories without taking privacy requirements into account. Other studies merely integrate differential privacy with unidirectional LSTM, failing to fully utilize the bidirectional context information of trajectories. This results in inadequate precision in feature extraction. Consequently, it is extremely challenging for them to fulfill the dual requirements of stringent privacy protection and high data utility in IoV scenarios.

In conclusion, there is an immediate requirement to establish a privacy protection mechanism that not only aligns with the spatiotemporal and continuous properties of IoV trajectories but also offers strict and provable

privacy protection while maintaining high data utility. To address this gap, this paper presents the DP-BiLSTM privacy protection method. By thoroughly integrating the advantages of BiLSTM in bidirectional temporal modeling and the robust protection ability of differential privacy, it effectively surmounts the limitations of prior studies and comprehensively demonstrates the necessity and advancement of this research.

3 DP-BiLSTM Trajectory Model

3.1 Model Framework

The DP-BiLSTM trajectory privacy protection model employs a bidirectional autoencoder architecture. The input consists of a three dimensional, normalized IoV trajectory sequence, specifically encompassing longitude, latitude, and speed. The output is a trajectory sequence after privacy protection. The framework of the DP-BiLSTM Model is illustrated in Figure 1. Initially, the BiLSTM encoder extracts the bidirectional temporal dependency features of trajectories and generates a latent representation characterized by high robustness. Subsequently, the LSTM decoder receives the latent representation and progressively reconstructs the trajectory sequence. Ultimately, the differential privacy processing module performs risk assessment, dynamic privacy budget allocation, and Laplace noise injection on the reconstructed trajectory to achieve provable differential privacy protection.

This implementation positions privacy protection operations within the post processing phase of trajectory data. By means of dynamic risk assessment and adaptive noise addition, the model can maximally retain the temporal structure, spatial distribution, and statistical characteristics of IoV trajectories during the privacy protection process. Consequently, an optimal equilibrium between privacy strength and data availability can be attained.

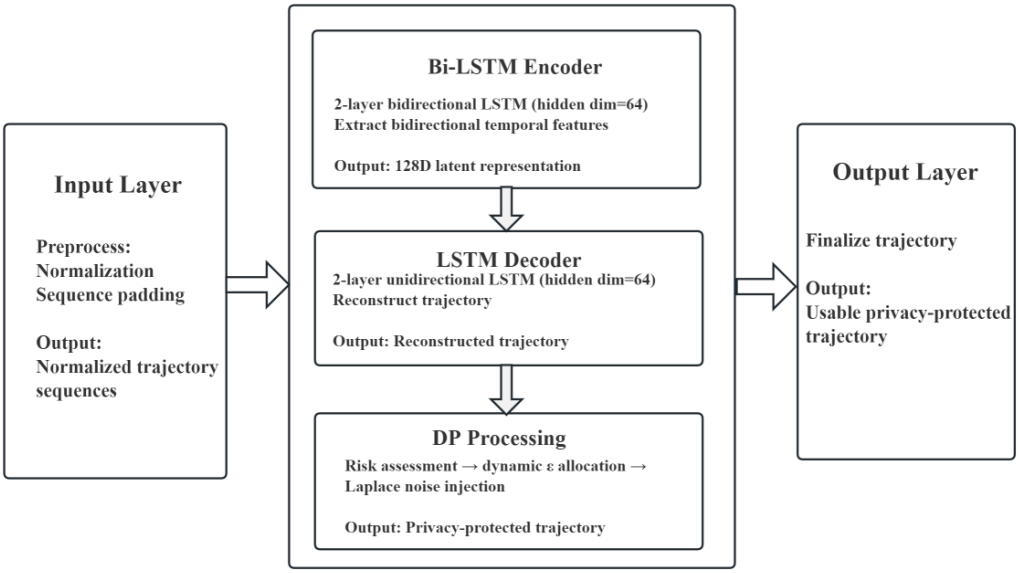

Figure 1. Framework of the DP-BiLSTM model

3.2 DP-BiLSTM Trajectory Encoder

The DP-BiLSTM encoder serves as the core element for feature extraction within the model. Its main function is to simultaneously capture the forward historical information and backward contextual information of trajectories, explore the bidirectional temporal dependencies of IoV trajectorie. The forward LSTM network processes the trajectory sequentially from the starting point to the ending point, focusing on learning historical dependencies. The backward LSTM network processes the trajectory in a reverse sequential order, emphasizing the capture of future contextual information. The calculation and concatenation formulas for the hidden states are presented as follows:

$$\overrightarrow{h_t} = LSTM\left(x_t, \overrightarrow{h_{t\text{-}1}}\right) \tag{1}$$

$$\overleftarrow{h_t} = LSTM\left(x_t, \overleftarrow{h_{t+1}}\right) \tag{2}$$

$$h_t = \left[\overrightarrow{h_t}, \overleftarrow{h_t}\right] \tag{3}$$

Where x_t represents the input of the trajectory feature at time, t, $\overrightarrow{h_t}$ denotes the hidden state of the forward LSTM network, $\overleftarrow{h_t}$ signifies the hidden state of the backward LSTM network, and h_t represents the final concatenated hidden representation.

3.3 Differential Privacy Module

The differential privacy module functions as the core privacy preserving component of the model. It is designed to provide provable differential privacy during the trajectory post processing stage. The module mainly consists of three essential operations: trajectory point risk assessment, dynamic privacy budget allocation, and Laplace noise injection, which jointly guarantee the privacy and security of IoV trajectory data.

The risk assessment operation calculates a risk score for each trajectory point by taking into account speed, location sensitivity, and temporal factors, thereby providing a basis for dynamic privacy budget allocation. The dynamic privacy budget allocation module non-linearly adjusts the privacy budget ε according to the risk score: a higher risk corresponds to a smaller ε with stronger noise. Subsequent to the allocation of the privacy budget, the incorporation of Laplace noise imparts controlled noise to the trajectory points. The magnitude of this noise is jointly determined by the dimensional sensitivity and the privacy budget, and its calculation formula is as follows:

$$Scale = \frac{C \cdot \Delta}{\varepsilon} \tag{4}$$

Where C is the clipping threshold, Δ is the global sensitivity of each trajectory dimension, including longitude, latitude, and speed. Additionally, ε represents the privacy budget. Encompassing a comprehensive spectrum of privacy protection intensities spanning from high to low levels, thereby fulfilling the privacy requirements of diverse IoV scenarios.

3.4 LSTM Trajectory Decoder

The LSTM decoder serves as the core element for trajectory reconstruction within the model. Consisting of a two layer unidirectional LSTM, it is designed to receive the latent representation output by the encoder. Subsequently, it gradually reconstructs the trajectory sequence, maps the features back to the original three dimensional input space, and generates the reconstructed IoV trajectory.

As trajectory reconstruction is essentially a regression task with continuous spatiotemporal variables, we adopt the Mean Squared Error (MSE) as the loss function to quantify the numerical discrepancy between the original trajectory and the reconstructed trajectory. Model parameters are optimized by minimizing the loss function to achieve high precision trajectory reconstruction. The loss function is defined as follows:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{5}$$

Where y_i represents the original trajectory point, $\hat{y}_i$ is the reconstructed trajectory point output by the model, and N denotes the length of a single trajectory.

4 Experimental Results and Analysis

4.1 Experimental Steps

The model utilizes a trajectory autoencoder composed of a BiLSTM encoder and a unidirectional LSTM decoder, which can efficiently learn and reconstruct the spatiotemporal patterns of trajectories. The core algorithm is presented as follows: during the training procedure, the standard gradient descent algorithm is employed to ensure the model's performance. Following the training phase, trajectory privacy protection is accomplished via post processing. The original data is fed into the trained model for forward propagation and subsequently processed by the differential privacy module to acquire the protected trajectory. Simultaneously, this method supports dynamic privacy budget allocation, which adaptively modifies the protection intensity in accordance with the risk score of each trajectory point and implements enhanced protection for sensitive locations, thereby enhancing the overall data availability.

The trajectory pre-processing for the DP-BiLSTM model is based on the Microsoft T-Drive dataset, which contains one week's GPS trajectory data of over 10,000 taxis in 2008. A series of pre-processing operations are conducted to obtain a high quality trajectory dataset that is clean, complete, dimension - consistent, and standardized. This step effectively eliminates scale differences and outlier interference, providing reliable input for model training and ensuring stability and convergence speed.

As illustrated in Figure 2, the feature encoding module utilizes a BiLSTM encoder to convert the normalized trajectory sequences into high dimensional latent representations. It extracts comprehensive bidirectional spatiotemporal features that characterize IoV trajectories and preserves the core statistical and temporal patterns, thus laying a solid foundation for subsequent reconstruction and privacy protection. The trajectory reconstruction module employs the LSTM decoder to transform the latent representation back into a trajectory sequence.

Trajectory privacy protection constitutes the core procedure of the method. It realizes differential privacy via post processing. Specifically, after reconstruction, each trajectory point undergoes risk assessment, dynamic privacy budget allocation, and Laplace noise injection. Controllable noise is introduced without undermining the core features of the trajectory, thus protecting the privacy of individual samples and attaining provable differential privacy. The protection intensity can be flexibly adjusted through the privacy budget ε. The final output is a privacy protected trajectory that maintains spatiotemporal continuity, spatial distribution, and data utility, optimizing its applicability for intelligent transportation analysis, location based services, and other applications while providing strong privacy protection.

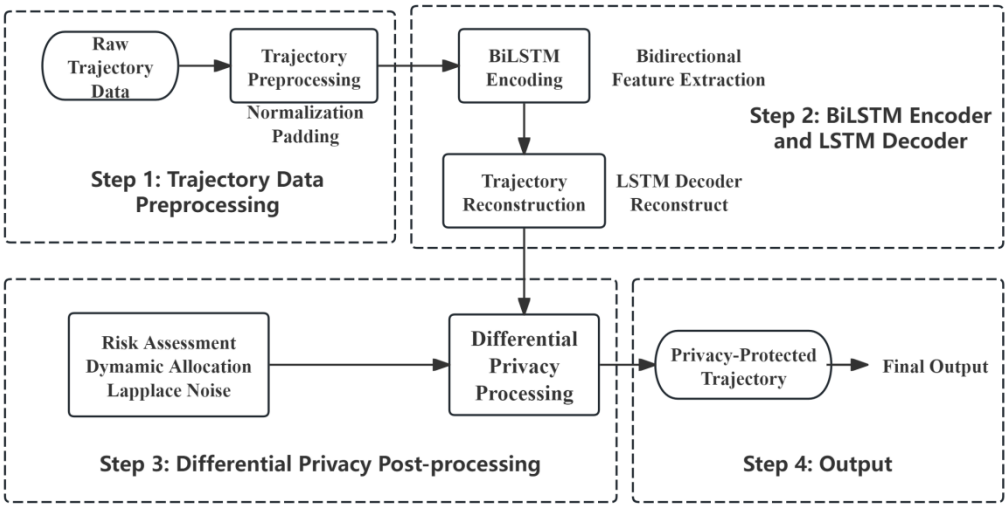

Figure 2. Flowchart of the DP-BiLSTM trajectory privacy protection method

Algorithm 1. The DP-BiLSTM trajectory privacy protection method

```
// 1. Data Preprocessing
Function preprocess(trajectory):
    return standardize_and_pad(trajectory, max_len=100)

// 2. BiLSTM Autoencoder Architecture
Class BiLSTM_Autoencoder:
   -Encoder: 2-layer Bidirectional LSTM (hidden_dim=64)
   -Decoder: 2-layer LSTM (hidden_dim=128)
   -Output: Dense layer (output_dim=3)

// 3. Standard Model Training
Function train_model(D, epochs=50, batch_size=128):
    for each epoch:
        for each batch:
            Forward pass to compute reconstructed trajectory
            Calculate MSE loss
            Backward pass with gradient clipping (max_norm=1.0)
            Update parameters
    return trained model

// 4. Differential Privacy Protection (Post-processing)
Function protect_trajectory(trajectory, ε_max=2.0, ε_min=0.1):
    protected = []
    for point in trajectory:
        risk = calculate_risk_score(speed, location_sensitivity, time)
        ε = ε_min + (ε_max-ε_min)*exp(-3*risk)  // Dynamic privacy budget
        Add Laplace noise to longitude, latitude, and speed (scale=sensitivity/ε)
        protected.append(noisy_point)
    return protected

// 5. Privacy-Protected Trajectory Generation
Function generate_private_trajectory(model, raw_trajectory):
    processed = preprocess(raw_trajectory)
    reconstructed = model(processed)  // Model reconstruction
    trajectory = inverse_standardize(reconstructed)
    return protect_trajectory(trajectory)  // Post-processing with noise addition
```

4.2 Performance Metrics Analysis

This paper establishes a multi-dimensional evaluation system for trajectory privacy protection. Four core metrics, namely Root Mean Square Error (RMSE), Query Error (QE), Jensen-Shannon divergence (JS), and Data Utility, are selected to comprehensively evaluate the reconstruction accuracy, query availability, distribution consistency, and practical application value of the model. The RMSE quantifies the numerical discrepancy between the original and reconstructed trajectories. A lower RMSE value indicates a higher level of reconstruction accuracy and a greater consistency between the protected and original trajectories. The QE assesses the accuracy of protected trajectories in typical IoV query tasks. A smaller QE value signifies a higher query availability and a stronger practical utility. The JS divergence measures the spatiotemporal distribution similarity between the original and protected trajectories. A lower JS divergence value implies a higher distribution consistency and a better preservation of statistical characteristics. Utility comprehensively reflects the usability of protected trajectories in practical IoV tasks, such as traffic analysis and path recovery. A larger utility value represents a higher data utility. The data utility is defined as:

$$Utility = \frac{1}{1+PE} \times \omega_{PE} + \frac{1}{1+SE} \times \omega_{SE} + LS \times \omega_{LS} \quad \textbf{(6)}$$

where PE represents the mean Euclidean distance between the original and protected trajectory points; SE denotes the mean absolute difference in speed; LS is defined as the ratio of the minimum value between the original length and the protected length to the maximum value; and the coefficients set in this paper are $\omega_{PE} = 0.4$, $\omega_{SE} = 0.4$, $\omega_{LS} = 0.2$.

To facilitate the quantitative comparison of the overall performance, the four metrics are normalized using the min-max normalization method and subsequently integrated to compute a comprehensive score. A higher score indicates superior overall performance. The formula is presented as follows:

$$Score = \frac{1}{4}\left[\left(1-\widetilde{RMSE}\right)+\left(1-\widetilde{QE}\right)+\left(1-\widetilde{JS}\right)+\widetilde{Utility}\right] \quad \textbf{(7)}$$

$$FinalScore = \frac{\sum_{i=1}^{n} w_i \cdot Score_i}{\sum_{i=1}^{n} w_i} \quad \textbf{(8)}$$

Where $\tilde{}$ denotes the min-max normalization $\tilde{X} = \frac{X - \min(X)}{\max(X) - \min(X)}$. w_i represents the weight of the i-th privacy budget, and $Score_i$ signifies the comprehensive score under the i-th privacy budget.

Experiments are conducted on the Windows operating system using Python 3.8 or a higher version. The key libraries include NumPy, Matplotlib, and the deep learning framework PyTorch for model construction and training. It is recommended to use a computer with at least 8 GB of RAM to support the computation. The privacy budget ε is set at five levels: 0.1, 0.5, 1.0, 1.5, and 2.0, which span the privacy requirements from high to low. Experiments are performed under each setting to analyze the influence of the privacy budget on performance and verify the model's adaptability under different protection intensities.

Table 1. Results of DP-BiLSTM under different privacy budgets

Privacy budget ε	RMSE	QE	JS	Utility	Score
0.1	0.1340	0.0513	0.6558	0.3624	0.4436
0.5	0.0427	0.0080	0.6501	0.4227	0.6901
1.0	0.0226	0.0066	0.6567	0.4636	0.6910
1.5	0.0162	0.0040	0.6477	0.4901	0.7332
2.0	0.0127	0.0031	0.6138	0.5155	0.8289

As shown in Table 1, the four core metrics exhibit distinct trends, when $\varepsilon = 0.1$, the noise intensity reaches its maximum, providing the highest level of privacy protection. Meanwhile, the model still maintains basic reconstruction and query capabilities, making it suitable for IoV scenarios with strict privacy requirements. As ε the value increases, the noise intensity decreases correspondingly, resulting in a gradual reduction in the strength of privacy protection. Simultaneously, the RMSE and QE demonstrate a continuous downward trend, the JS remains at a low level, and the data utility improves significantly. When $\varepsilon = 2.0$, the noise intensity is minimized, resulting in a moderate level of privacy protection. Under this configuration, the model achieves the highest data utility while retaining reliable privacy protection performance, which satisfies the requirements of conventional IoV applications. This trend reveals a clear privacy-utility trade-off, a smaller ε enhances privacy protection but reduces data utility, while a larger ε improves utility with a milder privacy guarantee. The stable JS values across budgets indicate consistent preservation of trajectory distribution characteristics.

4.3 Model Comparison and Analysis

To verify the superiority of the proposed DP-BiLSTM model, four representative models are selected for comparison: LSTM, DP-LSTM, BiLSTM, and DP-BiLSTM. These models include both non-privacy protected temporal models and privacy preserving models based on differential privacy, and all are tested under the

same experimental conditions and evaluation criteria.

The LSTM is a traditional unidirectional LSTM autoencoder without privacy protection, which only employs historical information for modeling. The DP-LSTM implements post processing differential privacy with the same noise strategy based on the unidirectional LSTM. The Bidirectional BiLSTM is a BiLSTM autoencoder without privacy protection, which leverages contextual information. The proposed DP-BiLSTM integrates the BiLSTM and post processing differential privacy to attain a more optimal equilibrium between privacy and utility.

To ensure fair and credible comparisons, all models (LSTM, DP-LSTM, BiLSTM, and DP-BiLSTM) adopt identical core parameters. All models use a 2-layer structure with a hidden dimension of 64, a dropout rate of 0.2, and an input dimension of 3. Training settings are unified: Adam optimizer with a learning rate of 0.001, 50 training epochs, a batch size of 16, MSE loss function, and gradient clipping with max_norm=1.0. Differential privacy models (DP-LSTM, DP-BiLSTM) share the same privacy budget settings and Laplace noise injection strategy. This consistent parameter setup eliminates confounding variables, ensuring that performance differences stem solely from architectural and privacy mechanism distinctions.

To verify the stability and efficacy of model training, the loss variation during the entire training process is presented in Figure 3, all models converge within 50 epochs. The loss undergoes a rapid decrease before reaching a stable state, without divergence or severe overfitting, which validates the reliability of the training procedure. The final convergence loss of DP-BiLSTM is slightly higher than that of DP-LSTM, mainly due to its more complex network structure and the accompanying additional optimization difficulties. Nevertheless, DP-BiLSTM maintains high convergence accuracy and generalization ability, effectively capturing bidirectional temporal dependencies while achieving robust privacy protection and reliable data utility.

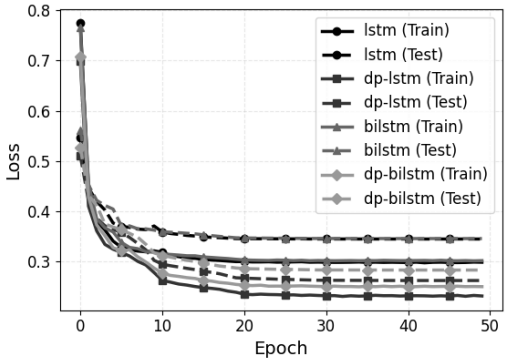

Figure 3. Training loss of different models

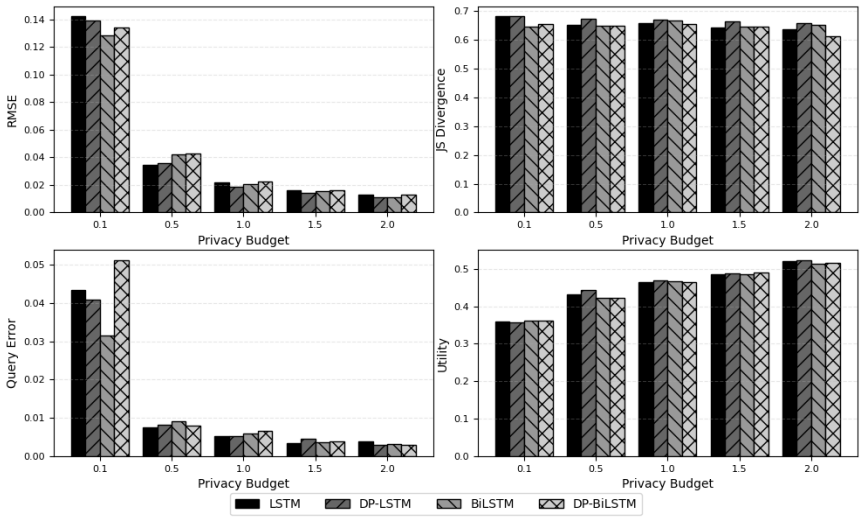

Figure 4. Performance comparison of different models

To investigate the influence of the privacy budget, the DP-BiLSTM model is evaluated at five levels: 0.1, 0.5, 1.0, 1.5, and 2.0, and the results are presented in Figure 4. Although the DP-BiLSTM model does not outperform other models in each individual metric, it achieves the optimal overall performance according to the comprehensive score. As shown in Table 2, the proposed DP-BiLSTM model attains the highest comprehensive score of 0.6773, exceeding the LSTM, DP-LSTM, and BiLSTM models. The comparison indicates that bidirectional models demonstrate significantly better performance compared to unidirectional models, validating the importance of bidirectional context in trajectory modeling. With differential privacy, the DP-BiLSTM model outperforms the DP-LSTM model, suggesting that the bidirectional structure is more resistant to noise. DP-BiLSTM's marginal advantage over BiLSTM (+0.0012) and substantial lead over DP-LSTM (+0.0506) confirm that integrating bidirectional modeling with post-processing differential privacy optimizes both privacy robustness and data utility.

Table 2. Final scores of different models

Model	LSTM	DP-LSTM	BiLSTM	DP-BiLSTM
Final score	0.6625	0.6267	0.6761	0.6773

The DP-BiLSTM model achieves the optimal balance between privacy protection and data utility. Based on the comprehensive experimental results, the privacy budget ε directly determines the equilibrium between privacy intensity and data utility, which can be flexibly chosen according to practical IoV scenarios. The model based on the BiLSTM can comprehensively capture contextual information and exhibits significant advantages in reconstruction accuracy compared with unidirectional models. Applying post processing differential privacy to reconstructed trajectories preserves spatiotemporal structures more effectively than directly adding noise to raw data. The DP-BiLSTM model attains the best overall performance, suggesting that it can provide strict privacy protection while maximizing the application value of trajectory data.

5 Conclusion

In light of the critical problems of a high risk of privacy leakage in IoV trajectory data and the difficulty of traditional protection methods in achieving a balance between privacy preservation and data availability, this paper proposes a DP-BiLSTM privacy protection method for vehicle trajectories. By integrating the temporal modeling capability of BiLSTM and the strict protection capability of differential privacy, this method adopts the BiLSTM autoencoder as its core and is composed of three crucial modules: the trajectory encoder, the trajectory decoder, and the differential privacy processor. It utilizes BiLSTM to extract the bidirectional temporal characteristics of trajectories and realizes differential privacy through post processing subsequent to trajectory reconstruction. The protection intensity can be flexibly regulated via the privacy budget ε.

Experimental results on the T-Drive dataset demonstrate that the DP-BiLSTM model can adapt to trajectory privacy requirements in different scenarios by tuning ε. It achieves a comprehensive score of 0.6773, which outperforms comparative models such as LSTM, DP-LSTM, and BiLSTM. The proposed method can provide provable privacy protection while maximizing the availability of trajectory data, thus achieving an optimal balance between privacy preservation and data utility. This privacy protection method can offer a reliable solution for the secure publication of IoV trajectories, intelligent transportation analysis, and other related scenarios.

6 Acknowledgments

The work was supported in part by the Projects of Fujian Provincial Department of Education JAT251256 and JZ250072, in part by the Fujian Youth Talent Cultivation Program under Grant 2026350800, and in part by the National Natural Science Foundation of Xiamen under Grant 3502Z202674025, in part by the National Natural Science Foundation of China under Grant 62371406.

References

[1] C. Liu, Z. Xiao, W. Long, T. Li, H. Jiang, K. Li, Vehicle trajectory data processing, analytics, and applications: A survey, *ACM Computing Surveys*, Vol. 57, No. 9, pp. 1-36, September, 2025.
https://doi.org/10.1145/3715902

[2] R. S. Rathore, C. Hewage, O. Kaiwartya, J. Lloret, In-vehicle communication cyber security: challenges and solutions, *Sensors*, Vol. 22, No. 17, Article No. 6679, September, 2022.
https://doi.org/10.3390/s22176679

[3] E. Zavvos, E. H. Gerding, V. Yazdanpanah, C. Maple, S. Stein, M. C. Schraefel, Privacy and Trust in the Internet of Vehicles, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 23, No. 8, pp. 10126-10141, August, 2022.
https://doi.org/10.1109/TITS.2021.3121125

[4] Y. Li, B. Wang, Q. Liu, X. Zheng, J. Li, Y. Wang, J. Xi, W. Qin, LPPS-IKHC: Location Privacy-Preserving Scheme using Improved k-anonymity and Hybrid Cache for IoV, *IEEE Transactions on Vehicular Technology*, Vol. 74, No. 8, pp. 12864-12878, August, 2025.
https://doi.org/10.1109/TVT.2025.3553550

[5] Y. Xu, Y. Xu, Z. Xu, Trajectory Protection with Individual Semantic Utility under Local Differential Privacy, *IEEE Internet of Things Journal*, Vol. 13, No. 10, pp. 19817-19834, May, 2026.
https://doi.org/10.1109/JIOT.2025.3628849

[6] H. Khan, G. G. Tejani, R. AlGhamdi, S. Alasmari, N. K. Sharma, S. K. Sharma, A secure and efficient deep learning-based intrusion detection framework for the internet of vehicles, *Scientific Reports*, Vol. 15, No. 1, Article No. 12236, April, 2025.
https://doi.org/10.1038/s41598-025-94445-9

[7] M. Ma, X. Han, S. Liang, Y. Wang, L. Jiang, Connected vehicles ecological driving based on deep reinforce learning: Application of Web 3.0 technologies in traffic optimization, *Future Generation Computer Systems*, Vol. 163, Article No. 107544, February, 2025.
https://doi.org/10.1016/j.future.2024.107544

[8] C. Li, M. Dong, Y. Fu, F. R. Yu, N. Cheng, Integrated sensing, communication, and computation for IoV: Challenges and opportunities, *IEEE Communications Surveys & Tutorials*, Vol. 28, pp. 1136-1168, September, 2025.
https://doi.org/10.1109/COMST.2025.3612388

[9] L. Sweeney, k-anonymity: A model for protecting privacy, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol. 10, No. 5, pp. 557-570, October, 2002.
https://doi.org/10.1142/S0218488502001648

[10] O. Abul, F. Bonchi, M. Nanni, Never walk alone:

Uncertainty for anonymity in moving objects databases, *2008 IEEE 24th International Conference on Data Engineering*, Cancun, Mexico, 2008, pp. 376-385. https://doi.org/10.1109/ICDE.2008.4497446

[11] A. Machanavajjhala, D. Kifer, J. Gehrke, M. Venkitasubramaniam, l-diversity: Privacy beyond k-anonymity, *ACM Transactions on Knowledge Discovery from Data*, Vol. 1, No. 1, Article No. 3-es, March, 2007. https://doi.org/10.1145/1217299.1217302

[12] C. Dwork, Differential privacy: A survey of results, *International Conference on Theory and Applications of Models of Computation*, Xi'an, China, 2008, pp. 1-19. https://doi.org/10.1007/978-3-540-79228-4_1

[13] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, L. Zhang, Deep learning with differential privacy, *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, Vienna, Austria, 2016, pp. 308-318. https://doi.org/10.1145/2976749.2978318

[14] M. E. Gursoy, L. Liu, S. Truex, L. Yu, Differentially private and utility preserving publication of trajectory data, *IEEE Transactions on Mobile Computing*, Vol. 18, No. 10, pp. 2315-2329, October, 2019. https://doi.org/10.1109/TMC.2018.2874008

[15] H. Li, S. Sahrani, M. R. Sarker, Y. Xiao, State-of-the-Art on IoV-Based Deep Learning Framework for Enhanced Driving Behavior Recognition: Recent Progress, Technology Updates, Challenges, and Future Direction, *IEEE Access*, Vol. 13, pp. 135969-135989, July, 2025. https://doi.org/10.1109/ACCESS.2025.3594091

[16] R. Kumar, P. Kumar, R. Tripathi, G. P. Gupta, N. Kumar, P2SF-IoV: A privacy-preservation-based secured framework for Internet of Vehicles, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 23, No. 11, pp. 22571-22582, November, 2022. https://doi.org/10.1109/TITS.2021.3102581

[17] M. Han, K. Xu, S. Ma, A. Li, H. Jiang, Federated learning-based trajectory prediction model with privacy preserving for intelligent vehicle, *International Journal of Intelligent Systems*, Vol. 37, No. 12, pp. 10861-10879, August, 2022. https://doi.org/10.1002/int.22987

[18] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Computation*, Vol. 9, No. 8, pp. 1735-1780, November, 1997. https://doi.org/10.1162/neco.1997.9.8.1735

[19] L. Astuti, Y.-C. Lin, ABi-LSTM: Attention-Based Bidirectional-LSTM for Multiclass Traffic Agents Trajectory Prediction, *IEEE Transactions on Vehicular Technology*, Vol. 74, No. 10, pp. 15603-15618, October, 2025. https://doi.org/10.1109/TVT.2025.3572329

[20] N. Fan, J. Liu, S. Zhao, Y. Dai, W. Fan, TLPP: Deep-learning-based two-layer privacy preserving mechanism for protecting vehicle trajectory data, *IEEE Internet of Things Journal*, Vol. 11, No. 22, pp. 36084-36098, November, 2024. https://doi.org/10.1109/JIOT.2024.3439393

[21] K. Hu, F. Zheng, L. Weng, Y. Ding, J. Jin, Action recognition algorithm of spatio–temporal differential LSTM based on feature enhancement, *Applied Sciences*, Vol. 11, No. 17, Article No. 7876, September, 2021. https://doi.org/10.3390/app11177876

[22] F. Khanmohammadi, R. Azmi, Time-series anomaly detection in automated vehicles using D-CNN-LSTM autoencoder, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 25, No. 8, pp. 9296-9307, August, 2024. https://doi.org/10.1109/TITS.2024.3380263

Biographies



Caiyun Chen received the MS degree in Electronic and Communication Engineering from Xiamen University, Xiamen, China, in 2019. She is currently a lecturer at Xiamen University Tan Kah Kee College. Her current research interests include artificial intelligence and information security.



Lianfen Huang (Member, IEEE) received the BS degree in radio physics and the PhD degree in communication engineering from Xiamen University, Xiamen, China, in 1984 and 2008, respectively. Her current research interests include wireless communication, wireless network, and signal processing.



Zufeng Weng received the Ph.D. degree from Xiamen University. He is currently a lecturer at Xiamen University Tan Kah Kee College. His research focuses on artificial intelligence, machine learning, and image processing.



Wu Zheng is a Lecturer at the School of Artificial Intelligence and Computer Science, Xiamen University Tan Kah Kee College. He received his M.S. degree in Control Theory and Control Engineering from Xiamen University.