An Efficient Attentive-GRU Structure for Uncertainty Modelling of Crowdsourced Human Trajectories Under Building-obscured Urban Scenes

Yue Yu[1], Zhewei Liu[1], Kexin Tang[1], Yunjuan Sun[1], Shuyu Zhang[1], Liang Chen[2], Ruizhi Chen[2]

1 The Hong Kong Polytechnic University, LSGI, Hong Kong, China – {yue806.yu, jackie.zw.liu, 23036584r, yj.sun, lionel-shuyu.zhang}@connect.polyu.hk
2 Wuhan University, LIESMARS, Wuhan, China – {l.chen, ruizhi.chen}@whu.edu.cn

Keywords: Uncertainty Modelling, Human Movement Trajectory, Attentive-GRU, GNSS, Building-obscured Urban Scenes.

Abstract:

Uncertainty modelling is regarded as one of the core components in the field of human mobility analysis and urban navigation, that can affect the performance of human behaviour modelling and location information acquisition. Existing uncertainty modelling algorithms towards the human movement trajectory are subjected to random and highly dynamic human motion characteristics and sampling and observation errors of Global Navigation Satellite System (GNSS) signals caused by the occlusion of buildings in urban scenes, which lead to the insufficient spatiotemporal correlation and poor accuracy of uncertainty modelling. To fill in this gap, this paper proposes an efficient attentive-GRU structure for uncertainty modelling of crowdsourced human trajectories under building-obscured urban scenes, that takes into account both temporal correlation and spatial correlation of human-originated GNSS trajectories and related motion features. A period of human motion data is modelled instead of only one or adjacent location points to avoid interference factors caused by the obstruction of urban buildings, and time-varying measurement and sampling errors are further estimated and combined with comprehensive human motion features to improve the accuracy of final uncertainty modelling. Comprehensive experiments indicate that compared with existing uncertainty modelling methods including physical models and deep-learning models, the proposed attentive-GRU structure realizes much better performance under different accuracy indexes.

1. Introduction

Pedestrian trajectory data within urban contexts is undeniably pivotal for advancing human mobility analysis. It provides a rich tapestry of information, elucidating the spatial-temporal dynamics and intricate social interactions of individuals and communities. The advent of Micro-Electro-Mechanical Systems (MEMS) sensors and their integration into mobile devices have revolutionized the acquisition of pedestrian movement data (Yu et al., 2021). This technological leap has enabled a plethora of applications, from enhancing intelligent transportation systems (Zhu et al., 2019) to pioneering advances in smart healthcare (Pal et al., 2018). These applications leverage pedestrian trajectory data to gain insights into user behaviours (Liu et al., 2022a; Liu et al., 2022b), support epidemic prevention measures (Yang et al., 2023), and facilitate the development of intelligent logistics solutions (Liu et al., 2020; Shi et al., 2021). Through these lenses, pedestrian trajectories serve as a foundational element in understanding urban mobility, contributing significantly to the optimization of location-based services (LBS) and informing city planning and public policy (Liu et al., 2019; Yu et al., 2022; Liu et al., 2021).

Despite their invaluable contributions, the task of capturing accurate and reliable pedestrian trajectories in the complex and dynamic urban environment is fraught with challenges (Rajput 2024; Liu et al., 2024). The diversity of pedestrian movement patterns, coupled with the limitations and variability of data provided by different mobile sensing and positioning technologies, introduces a significant degree of uncertainty in the trajectory data collected. This movement uncertainty, if not accurately addressed and mitigated, can severely impact the reliability of analyses conducted, undermining efforts to derive actionable insights and make informed decisions (Downs et al., 2018; Shi et al., 2021).

At the heart of trajectory uncertainty are two primary sources: sampling error and measurement error. Sampling error emerges from the collection process itself, where gaps in the data sequence result from variable sampling rates, leaving segments of motion information unrecorded or cannot be acquired (Zheng et al., 2012). Measurement error, conversely, arises from the variability inherent in positioning methodologies, the changing environmental conditions under which data collection occurs, and discrepancies in hardware performance (Zheng, 2015; Zheng et al., 2014). Addressing these errors necessitates innovative approaches that go beyond traditional methodologies, which often rely on assumptions of constant velocity or distance, leading to potential overestimations of the actual path area (Miller, 1991; Kwan, 1998).

Recognizing the limitations of conventional models in accurately quantifying movement uncertainty, especially in the nuanced and unpredictable realm of urban pedestrian movement, researchers have sought to refine PPA estimations through adaptive models. These models, such as the Approximate Upper Bound (AUB) approach (Furtado et al., 2018) and adaptive speed control criteria (Liu et al., 2022), represent significant advancements towards achieving more accurate uncertainty region calculations. However, the unique complexities presented by indoor environments—where the majority of valuable public trajectory data is generated—demand a more nuanced understanding and innovative solutions. The scarcity of absolute location references, the variability of measurement errors in complex indoor spaces, and the inherent randomness of indoor pedestrian movement patterns present formidable challenges to uncertainty estimation and lead to difficulty in balancing temporal and spatial correlations (Li et al., 2021; Liu et al., 2020).

In light of these challenges and building upon the foundational work of preceding research, this paper proposes a novel and comprehensive methodology for estimating the uncertainty of crowdsourced GNSS trajectories, specifically in urban environments characterized by architectural obstructions and complex human motion. By harnessing the power of advanced deep learning techniques, for instance Long Short-Term Memory

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-2024-427-2024

(LSTM) (Shi et al., 2022) and Bidirectional LSTM (Bi-LSTM) (Yu et al., 2023) networks, we endeavour to transcend the limitations of existing models. Our approach meticulously integrates human motion data with GNSS information to offer a holistic and adaptive framework for urban pedestrian trajectory analysis, setting a new benchmark for accuracy and reliability in the field.

Specifically, the contributions of our study lie in the following aspect:

(1) Compared with the existing methods that only use a single location point or nearby points for uncertainty modelling, this paper takes into account the time correlation and spatial correlation of human trajectory data in the feature extraction stage, by extracting GNSS-originated locations within a period of time, and combines with the related characteristics of human movement information to overcome the interference factors caused by the obstruction of urban buildings.

(2) This paper develops a novel attentive-GRU structure to achieve efficient and autonomous learning of human trajectory data. In order to improve the efficiency of existing LSTM, this article uses the GRU network to achieve more efficient human trajectory sequence modelling, and uses the attention mechanism to improve the accuracy and stability of spatiotemporal features learning and presentation.

(3) This paper adopts the ground-truth trajectory provided by Simultaneous Localization and Mapping (SLAM) technology to generate training and testing datasets under building-obscured urban scenes. The generated training dataset includes time-varying measurement and sampling errors, and also collects and models complex human motion information to improve the accuracy of uncertainty error modelling. Comprehensive experiments indicate that compared with existing uncertainty modelling methods including physical models and deep-learning models, the proposed attentive-GRU structure realizes much better performance under different accuracy indexes.

Organized for clarity and depth, the paper unfolds as follows: The "Method" section delves into our proposed approach for uncertainty estimation. In the "Experiment" section, we detail our experimental design and disseminate our findings, illustrating the robustness and adaptability of our model across diverse urban landscapes. The "Conclusion" section encapsulates our contributions to urban pedestrian trajectory analysis, advocating for future research directions that leverage the transformative potential of deep learning to navigate the complexities of urban mobility patterns further.

2. Problem Statement and Methedology

In this section, the key definitions regarding our movement uncertainty modelling are given, followed by a detailed elaboration of the proposed feature construction and methodology.

2.1 Key definitions and problem statement

This study aimed at developing a comprehensive and dependable framework for predicting the uncertainty of pedestrian movements within the complex settings of urban environments. The framework takes into account the variability of human motion, along with errors in sampling and measurement, offering an adaptive approach for predicting uncertainty errors. This approach thoroughly addresses the dynamic errors emanating from various location sources and their temporal impacts. The key concepts pertinent to this subject are outlined as follows:

Definition 1. Space-time Location Point (STP): An STP represents a two-dimensional geographical point specified by a certain time interval, provided through diverse positioning methods and described as $STP_i=(x,y,t)$. Here, (x,y) specifies the geographical coordinates, and t is the associated timestamp.

Definition 2. Ground-truth Trajectory (GT): The GT, derived through SLAM technologies, outlines a sequence of STPs $\{GT.STP_1,GT.STP_2,...,GT.STP_n\}$ that denotes the actual path walked by a pedestrian, incorporating reference locations. This trajectory is obtained by integrating SLAM systems with open street maps, enabling the retrieval of reference trajectories with centimeter-level precision for comparison.

Definition 3. GNSS-based Trajectory (HT): In this study, to maintain trajectory continuity in areas where GNSS signals are unavailable, human motion detection is utilized. The HT is thus a series of indexed STPs $\{HT.STP_1,HT.STP_2,...,HT.STP_n\}$ that estimates an individual's path using human motion detection over periods when GNSS signals are absent, with each $HT.STP_n$ accurately reflecting a corresponding point in $GT.STP_n$.

Consequently, our movement uncertainty modelling can be illustrated as:

Problem Statement: Given a HT and its corresponding GT, we aimed to find a mapping relationship M from HT to Dis (HT, GT), i.e., M: HT -> Dis(HT.STP, GT.STP), where Dis(HT.STP, GT.STP) denotes the spatial distance between HT.STP and GT.STP

2.2 Trajectory Collection and Data Preprocessing

In this paper, we choose the typical building-obscured urban scenes to collect crowdsourced GNSS trajectories data with complex human motion information, which contains:

1) the GNSS trajectory HT is provided by crowdsourced smartphone-reported GNSS locations provided by different users, and the human motion information is modelled as the combination of handheld-assisted trajectory dead-reckoning, described in Figure.1 (Yu et al., 2023):

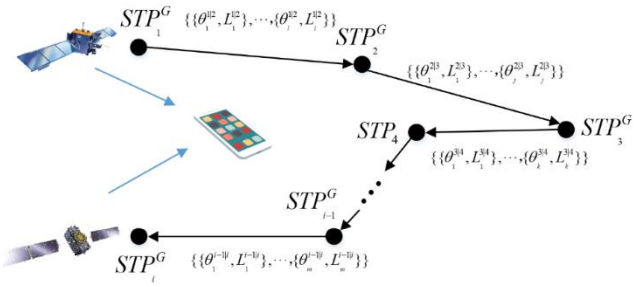

Figure 1. Description of GNSS Location and Human Motion (Yu et al., 2023)

After the handheld mode detection and conversion, the entire trajectory can be recreated as follows utilizing the motion data gathered throughout the trajectory period:

$$STP_k(\lambda_k,\psi_k)=\varpi_0+\sum_{k=0}^{n}\begin{bmatrix}\lambda_k cos(\psi_k)\\ \lambda_k sin(\psi_k)\end{bmatrix} \quad (1)$$

where ϖ_0 indicates the starting point of the selected trajectory, λ_k and ψ_k represent the heading and gait-length information calculated in reference (Yu et al., 2023).

2) the ground-truth trajectory GT is provided by the centi-meter level backpack SLAM product (Bao et al., 2022), that can provide high-accuracy reference trajectory for uncertainty error calculation. Finally, the generated real-world dataset contains the combination of following elements:

$$vector_{STP_k} = \left\{STP_k^{GT}(x), STP_k^{GT}(y), \lambda_k, \psi_k, STP_k^{HT}(x), STP_k^{HT}(y)\right\} \quad (2)$$

where $vector_{STP_k}$ indicates the basic element vector of training and test datasets. $STP_k^{GT}(x)$ and $STP_k^{GT}(y)$ are the ground-truth location, $STP_k^{HT}(x)$ and $STP_k^{HT}(y)$ are the GNSS location.

2.3 Feature Construction

This paper takes into account the time correlation and spatial correlation of human trajectory data in the feature extraction stage, by extracting GNSS-originated locations within a period of time, and combines with the related characteristics of human movement information to overcome the interference factors caused by the obstruction of urban buildings as follows:

1) Actual sampling interval of adjacent GNSS location T_i and GNSS reported location $\left\{x_G(k), y_G(k)\right\}$ at current timestamp.

2) Acquired gait-length and heading vector among two adjacent GNSS calculated locations:

$$\mathfrak{R}_{k|k-1} = \begin{pmatrix} \varsigma_1^{W/G} & \vartheta_1^{W/G} \\ \varsigma_2^{W/G} & \vartheta_2^{W/G} \\ \cdots & \cdots \\ \varsigma_N^{W/G} & \vartheta_N^{W/G} \end{pmatrix} \quad (3)$$

where N indicates the amounts of steps detected among adjacent GNSS locations, $\varsigma_N^{W/G}$ and $\vartheta_N^{W/G}$ indicate the corresponding step-length and heading vector.

3) Coordinate differences provided by human motion detection, GNSS, and hybrid value under x-axis and y-axis:

$$\begin{cases} \Delta X_k^M = \sum_{i=1}^{N} \varsigma_i \cdot \cos(\vartheta_i) \\ \Delta Y_k^M = \sum_{i=1}^{N} \varsigma_i \cdot \sin(\vartheta_i) \\ \Delta X_k^G = x_G(k) - x_G(k\text{-}1) \\ \Delta Y_k^G = y_G(k) - y_G(k\text{-}1) \\ \Delta X_k^{G-M} = x_G(k) - \sum_{i=1}^{N} \varsigma_i \cdot \cos(\vartheta_i) \\ \Delta Y_k^{G-M} = y_G(k) - \sum_{i=1}^{N} \varsigma_i \cdot \sin(\vartheta_i) \end{cases} \quad (4)$$

4) Euclidean distance of adjacent GNSS based location source and gait and heading originated based location source:

$$\begin{cases} D_k^M = \sqrt{\left(\sum_{i=1}^{N} \varsigma_i \cdot \cos(\vartheta_i)\right)^2 + \left(\sum_{i=1}^{N} \varsigma_i \cdot \sin(\vartheta_i)\right)^2} \\ D_k^{GNSS} = \sqrt{\left(x_G(k) - x_G(k\text{-}1)\right)^2 + \left(y_G(k) - y_G(k\text{-}1)\right)^2} \end{cases} \quad (5)$$

where $\left\{x_G(k), y_G(k)\right\}$ represents the GNSS-originated locations.

5) Estimated GNSS-originated speed and location observations and gait and heading originated observations:

$$\begin{cases} v_k^M = D_k^M / \varepsilon_k^{GNSS} \\ v_k^{GNSS} = D_k^{GNSS} / \varepsilon_k^{GNSS} \end{cases} \quad (6)$$

where v_k^M and v_k^{GNSS} represent the MEMS sensors and GNSS originated walking speeds. ε_k^{GNSS} indicates the sampling intervals of GNSS based location source.

2.4 Proposed Model

Prior studies (Shi et al. 2022; Yu et al. 2023) have highlighted the importance of extracting temporal features for predicting errors. Nonetheless, the Long Short-Term Memory (LSTM) networks previously utilized did not adequately capture long-range dependencies. Additionally, these networks were limited by their capacity to determine which points in time were most crucial for accurate forecasts. To overcome these shortcomings, our research introduces an innovative neural network design that combines the more sophisticated Gated Recurrent Unit (GRU) with a multi-head attention mechanism. This hybrid approach is engineered to more effectively recognize temporal relationships and the movement patterns of pedestrians.

Figure 1 illustrates our model Attentive-GRU for handling the input sequence $X = [x_1, x_2, \ldots, x_T]$, in which each $x_i = \left[delta_update_time, GPS_x, GPS_y, \ldots\right]$ is the feature vector at time step i, and T is the total number of time steps. Initially, we transform the input sequence into a higher dimensional space using a fully connected layer. This transformed sequence is then passed through a multi-head attention module. The output with refined attention is subsequently inputted into a Gated Recurrent Unit (GRU) module, which is responsible for capturing the temporal dependencies within the data. The final step involves mapping the GRU's output to a one-dimensional vector via another fully connected layer, which serves as the prediction for the estimated error.

Within the multi-head self-attention module, the initial step is to project the input sequence X onto three separate spaces to generate the query $Q = W_q X$, the key $K = W_k X$, and the value $V = W_v X$. Here, W_q, W_k, and W_v represent the respective weight matrices for the query, key, and value projections. Subsequent to these projections, we calculate the attention scores matrix A:

$$A = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right), \quad (7)$$

where d_k is the dimension of the key. The attention score matrix A represents the importance of each time step in the sequence. We then compute the output of the multi-head attention layer as:

$$X' = \text{MultiHead}(Q, K, V)\text{Concat}(\text{head}_1, \ldots, \text{head}_H)W^O, \quad (8)$$

where $\text{head}_i = \text{Attention}\left(QW_i^Q, KW_i^K, VW_i^V\right)$, and W^O is the output weight matrix.

The resulting output X' from the multi-head self-attention module is subsequently channeled into the GRU to discern the sequential dependencies. The GRU encompasses a reset gate r_t, an update gate z_t, and a candidate hidden state $\tilde{h}_t$. The update dynamics within the GRU are described by the equations below:

$$\begin{aligned} r_t &= \sigma(W_r[h_{t-1}, x'_t]), \\ z_t &= \sigma(W_z[h_{t-1}, x'_t]), \\ \tilde{h}_t &= \tanh(W_h[r_t \odot h_{t-1}, x'_t]), \\ h_t &= (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t, \end{aligned} \quad (9)$$

where σ is the sigmoid function, $\odot$ denotes element-wise multiplication, and W_r, W_z, and W_h are the weight matrices for the reset gate, update gate, and candidate hidden state,

respectively.

Finally, the output of the GRU module is passed through a fully connected layer to obtain the estimated error. The loss function is defined as the mean squared error between the predicted error and the ground truth error.

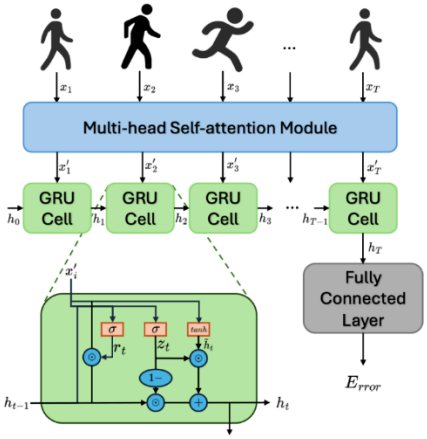

Figure 2. The architecture of proposed AttentiveGRU

3. Experimental Results

3.1 Experimental Dataset Collection

Our dataset was collected on the open campus of The Hong Kong Polytechnic University, where the received GNSS signals were significantly affected by reflections and occlusions, leading to time-varying sampling errors and large measurement inaccuracies. To obtain the ground-truth trajectories of individuals, we employed a SLAM system capable of surveying locations with centimeter-level precision. The GNSS-based trajectory dataset, denoted as HT, encompasses human movements and locations recorded using a variety of smartphones. Additionally, some representative trajectories within the constructed dataset are detailed in Figure. 3.

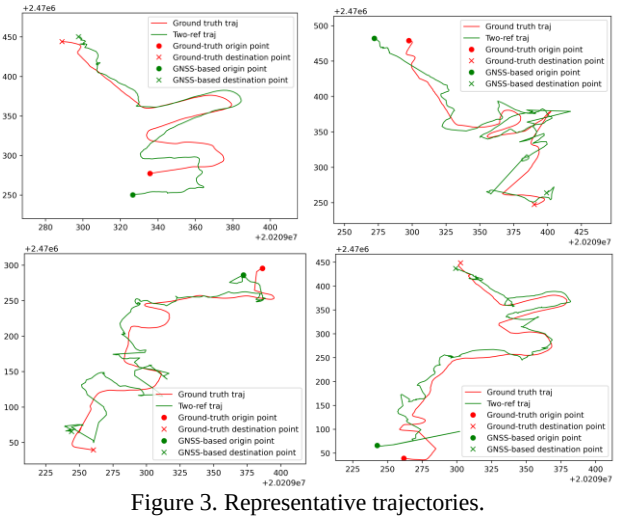

Figure 3. Representative trajectories.

In our study, a total of 15 trajectories were analyzed, covering an extensive dataset with a significant number of 7954 location points. These trajectories varied in length and duration, with an average trajectory length of 981.7 m and an average time period of 577.47 s. The dataset captured a wide range of human movements, including the largest walking speed recorded at 55.6 m/s. Additionally, the data was collected with an average sampling interval of 1.1 s. Of the collected 15 trajectories, three trajectories were designated for testing, while the remainder were used for training and validation during the training process.

3.2 Model Training and Baseline Methods

To evaluate our model, we used existing models LSTM and BiLSTM as baselines and added a GRU for comparison. In addition, we introduced an LSTM model with an attention mechanism to assess the impact of attention on error estimation. During the training of these models, we utilized the mean squared error as the loss function.

To prevent data leakage, we partitioned the dataset into training and testing sets based on estimates, with the training set containing 12 trajectories and the test set containing 3 trajectories. Moreover, we allocated 3 trajectories from the training set as a validation set and employed early stopping to avoid overfitting. Ultimately, we assessed the performance of the models by comparing their performance on the test set.

For fairness, all models were implemented using PyTorch and trained on an NVIDIA Tesla P100 GPU. We set the hyperparameters of the models as follows: the dimension of the hidden layer was 64, the number of attention heads was 8, the learning rate was set to 0.001, and the batch size was 32. Furthermore, we trained the models for 300 epochs and employed early stopping based on the validation set. Ultimately, we evaluated the performance of the models by comparing their results on the test set.

3.3 Performance Evaluation of Uncertainty Prediction

Our evaluation focuses on two main metrics: the average error (AvgErr) and the errors at specific percentiles (P25Err, P50Err, and P75Err). The average error provides a straightforward measure of overall accuracy, reflecting the average deviation of the estimated ages from their true values across all data points. We evaluate error variability and distribution using the 25th (P25Err), 50th (P50Err), and 75th (P75Err) percentile errors. P25Err represents the lower error range, P50Err the median, indicating average performance, and P75Err the higher error range, showing the spread of errors from best to worst cases. The results are given in Table 1.

Model	AvgErr	P25Err	P50Err	P75Err
Our model	1.35	0.39	0.84	1.53
AttentiveBiLSTM	1.46	0.38	0.83	1.57
AttentiveLSTM	1.48	0.40	0.89	1.67
BiLSTM	1.74	0.52	1.13	2.05
GRU	1.64	0.48	1.03	1.87
LSTM	1.78	0.53	1.15	2.10

Table 1. Errors of Uncertainty prediction by Respective Models

Our model demonstrates superior performance with the lowest average error of 1.35, indicating its effectiveness in accurately estimating ages across diverse datasets. Additionally, it shows commendable accuracy in the lower error range with a P25Err of 0.39, closely followed by the AttentiveBiLSTM model at 0.38. This highlights our model's precision in handling a significant portion of the data with minimal deviation from the actual ages.

At the median of the error distribution (P50Err), our model and the AttentiveBiLSTM model exhibit nearly identical performance, with errors of 0.84 and 0.83, respectively, suggesting both models maintain consistent accuracy across the

median of the datasets. However, our model slightly outperforms the AttentiveBiLSTM in the upper quartile (P75Err), with an error of 1.53 compared to 1.57, indicating better handling of outliers or more challenging age estimations.

Comparatively, existing deep-learning models such as BiLSTM, GRU, and LSTM show higher average errors (1.74, 1.64, and 1.78, respectively) and progressively larger errors across the 25th, 50th, and 75th percentiles. This suggests that while these models have been foundational in sequential data processing, their performance in age estimation tasks is surpassed by models incorporating attention mechanisms, like our model and the AttentiveBiLSTM. The AttentiveLSTM model, while slightly less accurate than our model and the AttentiveBiLSTM, still demonstrates competitive performance with an average error of 1.48 and shows a balanced error distribution across the percentiles.

In this analysis, respective models are used to prediction the movement uncertainty and their performance are further evaluated. A comparison of the generated uncertainty areas by existing physical models: UB (Li et al., 2018), AUB (Furtado et al., 2018), BAEE (Shi et al., 2021) are shown in Figure 4.

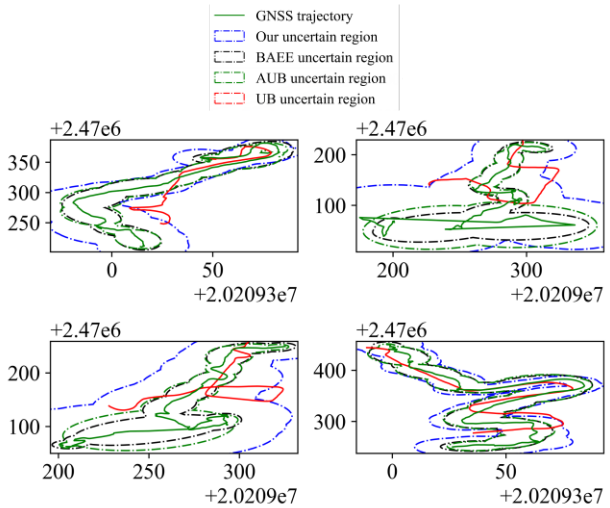

Figure.4 Comparison of Generated Uncertainty Areas

It can be from Figure.4 that the proposed Attention-GRU proves much better coverage performance compared with existing physical models, the generated uncertainty area can well cover the ground-truth walking trajectories of the human. Overall, our analysis underscores the efficacy of our model in age estimation, with it not only achieving the lowest average error but also exhibiting robustness across different portions of the error distribution, thereby confirming its reliability and accuracy in age estimation tasks across varied datasets.

Besides the accuracy metric, two additional metrics, coverage ratio and the mean absolute error (MAE), are further revised to respectively evaluate how well the ground-truth trajectories are covered by the uncertainty ratio. Coverage ratio measures the proportional coverage of the ground-truth trajectory points by the generated uncertain regions. MAE refers to the normalized equivalent error, calculated from the adaptive Euclidean distance based on different models. These metrics collectively provide a comprehensive assessment of the performance, ensuring not only accuracy but also thorough coverage and optimal distribution of the uncertain regions relative to the actual trajectories. The performance of proposed accuracy indexes coverage ratio and MAE is described in Table 2 and Table 3.

	Coverage Ratio			
Num	AttentiveGRU	BAEE	AUB	UB
T01_05	97.6%	50.2%	15.2%	100%
T02_05	99.5%	82.9%	59.5%	100%
T03_05	98.6%	82.1%	46.9%	100%
Average	98.6%	68.2%	38.2%	100%

Table 2. Coverage ratio of respective method on test dataset

	MAE (meter)			
Trajectory	AttentiveGRU	BAEE	AUB	UB
T01_05	1.20	31.95	32.39	29.23
T02_05	1.32	9.63	10.61	44.80
T03_05	1.49	8.25	9.66	45.68
Average	1.35	16.30	17.23	40.00

Table 3. MAE (meter) of respective method on test dataset

It can be found from Table 2 and Table 3 that the proposed AttentiveGRU model proves significant improvement of index MAE compared with three existing physical models, because they cannot take into account the estimation of positioning source measurement error under the task of human trajectory uncertainty modeling. In addition, the proposed AttentiveGRU also proves comparable performance of coverage ratio with UB, and improved performance with BAEE and AUB.

4. Conclusion

In conclusion, this paper presents a comprehensive analysis of GNSS signal-based trajectory uncertainty estimation in the context of an open campus environment. We have introduced a novel AttentiveGRU model that outperforms both physical models and deep-learning models in terms of accuracy and reliability, as evidenced by our extensive evaluations across multiple metrics. Our work not only advances the understanding of GNSS signal characteristics in complex urban settings but also provides a robust framework for accurate trajectory estimation.

Looking ahead, our future work will focus on several key areas. We plan to extend our research to more complex indoor and outdoor scenarios, where signal interference and multipath effects pose additional challenges. Additionally, we aim to design a more universal model that can adapt to varying environmental conditions and user behaviors. Optimizing training samples will also be a priority, ensuring that our model can learn from a diverse and representative dataset. Through these efforts, we aspire to further enhance the accuracy and generalizability of trajectory estimation techniques, contributing to the development of more sophisticated location-based services and urban navigation systems.

Acknowledgements

This work was supported by The Hong Kong Polytechnic University (P0045937); Open research Fund of State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University (23P03).

References

Bao S, Shi W, Chen P, et al. A systematic mapping framework for backpack mobile mapping system in common monotonous environments. *Measurement*, 2022, 197: 111243.

Downs, J., Horner, M., Lamb, D., et al., 2018. Testing time-geographic density estimation for home range analysis using an agent-based model of animal movement. *Int. J. Geograph. Inform. Sci*. 32 (7), 1505–1522.

Furtado, A.S., et al., 2018. Unveiling movement uncertainty for robust trajectory similarity analysis. *Int. J. Geograph. Inform. Sci*. 32 (1), 140–168.

Kwan, M.P., 1998. Space-time and integral measures of individual accessibility: a comparative analysis using a point-based framework. *Geogr. Anal*. 30 (3), 191–216.

Li H, Lu H, Shou L, et al. In search of indoor dense regions: An approach using indoor positioning data. *IEEE Transactions on Knowledge and Data Engineering*, 2018, 30(8): 1481-1495.

Li, Z., Zhao, X., Zhao, Z., et al., 2021. WiFi-RITA positioning: Enhanced crowdsourcing positioning based on massive noisy user traces. *IEEE Trans. Wirel. Commun*. 20 (6), 3785–3799.

Liu Z, Shi W, Yu Y, et al. A LSTM-based approach for modelling the movement uncertainty of indoor trajectories with mobile sensing data. *International Journal of Applied Earth Observation and Geoinformation*, 2022, 108: 102758.

Liu, T., Kuang, J., Ge, W., et al., 2020. A simple positioning system for large-scale indoor patrol inspection using foot-mounted INS, QR code control points, and smartphone. *IEEE Sens. J*. 21 (4), 4938–4948.

Liu, Z., Liu, J., Huang, X., Zhang, E., Chen, B., 2022. Measuring Chinese cities’ economic development with mobile application usage. *J. Geog. Sci*. 32 (12), 2415–2429.

Liu, Z., Wang, A., Weber, K., et al., 2022. Categorisation of cultural tourism attractions by tourist preference using location-based social network data: The case of Central, Hong Kong. *Tourism Manage*. 90, 104488.

Liu, C., Feng, Y., Lin, D., et al., 2020. Iot based laundry services: an application of big data analytics, intelligent logistics management, and machine learning techniques. *Int. J. Prod. Res*. 58 (17), 5113–5131.

Liu, Z., Felton, T. and Mostafavi, A., 2024. Interpretable machine learning for predicting urban flash flood hotspots using intertwined land and built-environment features. *Computers, Environment and Urban Systems*, 110, p.102096.

Liu, Z., Zhou, X., Shi, W. and Zhang, A., 2019. Recommending attractive thematic regions by semantic community detection with multi-sourced VGI data. *International Journal of Geographical Information Science*, 33(8), pp.1520-1544.

Liu, Z., Zhang, A., Yao, Y., Shi, W., Huang, X. and Shen, X., 2021. Analysis of the performance and robustness of methods to detect base locations of individuals with geo-tagged social media data. *International Journal of Geographical Information Science*, 35(3), pp.609-627.

Miller, H.J., 1991. Modelling accessibility using space-time prism concepts within geographical information systems. *Int. J. Geogr. Inform. Syst*. 5 (3), 287–301

Pal, D., Funilkul, S., Charoenkitkarn, N., et al., 2018. Internet-of-things and smart homes for elderly healthcare: An end user perspective. *IEEE Access*. 6, 10483–10496.

Rajput, A.A., Liu, C., Liu, Z. and Mostafavi, A., 2024. Human-centric characterization of life activity flood exposure shifts focus from places to people. *Nature Cities*, pp.1-11.

Shi, W., Liu, Z., An, Z., Chen, P., 2021. RegNet: a neural network model for predicting regional desirability with VGI data. I*nt. J. Geogr. Inf. Sci*. 35 (1), 175–192.

Shi, W., Chen, P., Shen, X., et al., 2021. An adaptive approach for modelling the movement uncertainty in trajectory data based on the concept of error ellipses. *Int. J. Geograph. Inform. Sci*. 35 (6), 1131–1154.

Shi W, Yu Y, Liu Z, et al. A deep-learning approach for modelling pedestrian movement uncertainty in large-scale indoor areas. *International Journal of Applied Earth Observation and Geoinformation*, 2022, 114: 103065.

Yang D, Shi W, Yu Y, et al. Analysis of the Spatial Distribution and Associated Factors of the Transmission Locations of COVID-19 in the First Four Waves in Hong Kong. *ISPRS International Journal of Geo-Information*, 2023, 12(3): 111.

Yu Y, Chen R, Chen L, et al. A robust seamless localization framework based on Wi-Fi FTM/GNSS and built-in sensors. *IEEE Communications Letters*, 2021, 25(7): 2226-2230.

Yu Y, Shi W, Chen R, et al. Map-assisted seamless localization using crowdsourced trajectories data and bi-lstm based quality control criteria. *IEEE Sensors Journal*, 2022, 22(16): 16481-16491.

Yu Y, Yao Y, Liu Z, et al. A Bi-LSTM approach for modelling movement uncertainty of crowdsourced human trajectories under complex urban environments. *International Journal of Applied Earth Observation and Geoinformation*, 2023, 122: 103412.

Yu Y, Zhang Y, Chen L, et al. Intelligent fusion structure for Wi-Fi/BLE/QR/MEMS sensor-based indoor localization. *Remote Sensing*, 2023, 15(5): 1202.

Zheng, K., et al., 2012. Reducing uncertainty of low-sampling-rate trajectories. In: *2012 IEEE 28th International Conference on Data Engineering. Arlington, Virginia*, pp. 1144–1155.

Zheng, Y., 2015. Trajectory data mining: an overview. *ACM Trans. Intell. Syst. Technol. (TIST)*. 6 (3), 29.

Zheng, Z., Rasouli, S., Timmermans, H., 2014. Evaluating the accuracy of GNSS-based taxi trajectory records. *Proc. Environ. Sci*. 22, 186–198.

Zhu, F., Lv, Y., Chen, Y., et al., 2019. Parallel transportation systems: Toward IoT enabled smart urban traffic control and management. *IEEE Trans. Intell. Transp. Syst*. 21 (10), 4063–4071.