An Enhanced Seamless Localization Framework Using Spatial-temporal Uncertainty Predictor Under Obscured Indoor and Outdoor Scenes

Yue Yu[1], Shiyu Bai[1], Jingxian Wang[1], Kexin Tang[1], Sheng Bao[1], Liang Chen[2], Ruizhi Chen[2]

1 The Hong Kong Polytechnic University, LSGI, Hong Kong, China - {michael-yue.yu, shiyu.bai, jing-xian.wang, 21053105g, sheng-s.bao}@polyu.edu.hk
2 Wuhan University, LIESMARS, Wuhan, China - {l.chen, ruizhi.chen}@whu.edu.cn

Keywords: Uncertainty Modelling, Seamless Localization, Multi-source Fusion, GNSS, Wi-Fi, Obscured Indoor and Outdoor Scenes.

Abstract:

Uncertainty modelling is regarded as one of the core components in the field of urban navigation, that can affect the performance of indoor and outdoor location information acquisition, especially under obscured scenarios. Existing multi-source fusion-based seamless positioning algorithms are subjected to random and highly dynamic human motion characteristics and changeable observation errors caused by the dynamic occlusions of human and buildings under urban scenes, which lead to the insufficient spatiotemporal correlation and poor accuracy of final multi-source fusion structure. To fill in this gap, this paper proposes an enhanced seamless localization framework using spatial-temporal uncertainty predictor under obscured indoor and outdoor scenes (ESL-STUP), that takes into account both temporal correlation and spatial correlation of trajectories provided by Wi-Fi, GNSS, and sensor-originated motion information. An iPDR-based trajectory estimation structure is proposed, using the integration of INS/PDR mechanizations, magnetic observations, and deep-learning based speed estimation to enhance the performance of traditional PDR algorithm. A period of human motion features extracted from hybrid location sources are modelled instead of only one or adjacent location points to realize time-varying measured uncertainty errors prediction, and the predicted uncertainty errors of different indoor and outdoor location sources are integrated with iPDR to realize robust seamless positioning performance. Comprehensive experiments indicate that compared with existing multi-source fusion-based seamless positioning structure, the proposed ESL-STUP realizes much better performance under different scenes.

1. Introduction

Urban navigation, a field that requires accurate and reliable location information, has been a subject of extensive research due to its importance in various applications. A critical component of this field is uncertainty modelling, which significantly influences the performance of indoor and outdoor location information acquisition (Liu et al. 2022; Shi et al. 2022; Yu et al. 2022). This is particularly true in obscured scenarios where the dynamic nature of urban environments, characterized by random and highly dynamic human motion and changeable observation errors caused by dynamic occlusions of humans and buildings, poses significant challenges.

In outdoor environments, the Global Positioning System (GPS) modules integrated with Internet of Things (IoT) terminals have been widely used for pedestrian and vehicle navigation. However, in urban areas, the original satellite signals are often obscured by high-rise buildings, leading to significant Non-Line-of-Sight (NLOS) errors or interruptions in location information. Therefore, in practical navigation scenarios, it is crucial to evaluate the availability of the received GPS signals. Moreover, the performance of relative positioning algorithms based on integrated sensors is affected by cumulative errors and magnetic interference, and thus cannot be used alone over continuous periods (Wen et al. 2021; Wen et al. 2019).

The work in (Yu et al. 2021) improved the positioning accuracy and stability of a single global navigation satellite source and achieved seamless positioning by integrating Wi-Fi Fine Time Measure (FTM), GNSS, and terminal sensors. The work in (Wu et al. 2015) proposed a new method to reduce GPS bias and improve the GPS positioning performance reported by smartphones by extracting human motion features. The experiment proved that the accuracy of this method is 30% higher than that of the GPS method. Work (Zhang et al. 2020) applied reinforcement learning to improve GPS positioning accuracy without requiring any other device parameters or reference information. The confidence-based reward mechanism has an efficient training stage and more accurate error prediction results, which can improve the accuracy by 50% in the final positioning stage.

In indoor scenarios, how to effectively estimate and eliminate the positioning source errors that affect the accuracy of pedestrian trajectories has become a hot topic in current research. In terms of enhancing the accuracy of inertial odometer. In (Luo et al. 2020), authors proposed a deep learning model-assisted Pedestrian Dead Reckoning (PDR) algorithm that identifies complex motion patterns of pedestrians from the smartphone's Inertial Measurement Unit (IMU). The experimental results show that this algorithm can effectively handle data under complex motion patterns. (Zhang et al. 2018) proposed a PDR indoor positioning algorithm that combines online sequential extreme learning machine. By incorporating pedestrian's motion habits into the extreme learning process, it can predict the pedestrian's position without considering the handheld posture. A large number of experimental results have proven the effectiveness of this algorithm under different postures and the practicality of indoor positioning. (Wang et al. 2022) proposed a pedestrian trajectory estimation method based on learning-based inertial odometer and magnetometer. This method can estimate the bias of the magnetometer online, rather than relying on pre-calibrated magnetometer measurements. The test results show that even using calibrated magnetometer data, this method can achieve better positioning performance than other methods.

The seamless localization is aimed at providing an accurate and concrete location result between indoor and outdoor environment. In order to realize seamless positioning, different location sources should be combined together to provide a more robust location performance.

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-4-2024-493-2024 | © Author(s) 2024. CC BY 4.0 License.

(Weyn et al. 2009) fused GPS, Wi-Fi, and global system for mobile communications (GSM) signals in the measurement models of particle filters with different weights to achieve a seamless indoor and outdoor positioning without distinguishing the users’ indoor and outdoor status.

(Sakamoto et al. 2012) presented an indoor/outdoor seamless robot navigation method by using pseudo-satellites and the indoor messaging system. The evaluation showed that this method can achieve centimeter level to decimeter-level positioning accuracy. (Jiang et al. 2021) combined GNSS, INS and UWB by several enhanced tightly-coupled integration models to improve the seamless localization performance. In which GNSS/INS/UWB combination is applied under satellite-weak areas to increase the continuity of the total framework and the final experiments verified the centimeter-level accuracy of proposed system.

The other high-precision based location sources such as Wi-Fi Fine Time Measurement (FTM), UWB can also be applied for seamless localization purpose combined with the global navigation satellite system (GNSS) signal and built-in sensors (Di et al. 2020; Wan et al. 2022). Except for the wireless location sources indoors, in recent years, the visual positioning system is expected to provide accurate indoor and outdoor mapping and localization performance aiming at IoT terminals.

(Qin et al. 2018) developed a robust indoor and outdoor mapping and localization system (VINS-Mono) based on the monocular camera and inertial sensors, which contains a tightly-coupled visual-inertial odometry (VIO) and a loop detection module. In addition, the proposed system can be applied in both PCs platform and IOT platforms such as Android and IOS terminals. (Dong et al. 2018) proposed the scalable and cost-efficient visual navigation system (ViNav) which contains the functions of indoor mapping, localization, and navigation using IoT terminals based visual and inertial sensors data. Structure-from-motion (SfM) and image fingerprinting are applied to combine the crowdsourced visual data for pedestrian path recognition and barometer features are extracted and integrated for multi-floor navigation.

For the uncertainty error prediction of indoor location sources such as Wi-Fi fingerprinting, researchers have found that the effective prediction of Wi-Fi-reported location error makes a significant improvement in final fusion phase. (Li et al. 2019) proposed an uncertainty prediction method towards the Wi-Fi fingerprinting based positioning method, that extracts the observation features provided by Wi-Fi RSSI fingerprinting, machine learning is applied to autonomously learn and predict the matching error of Wi-Fi fingerprinting. (Yu et al. 2021) develops a multilayer perceptron (MLP) for uncertainty error estimation of Wi-Fi fingerprinting, and the Wi-Fi fingerprinting reported location and human motion information are considered and modelled as the input vector of MLP network, and which realizes the accuracy enhancement of final multi-source fusion. While the existing uncertainty error modelling algorithms are still subjected to random and highly dynamic human motion characteristics and changeable observation errors caused by the dynamic occlusions of human and buildings under urban scenes, which lead to the insufficient spatiotemporal correlation and poor accuracy of final multi-source fusion structure

In light of these challenges and the identified research gap, this paper proposes an enhanced seamless localization framework using a spatial-temporal uncertainty predictor under obscured indoor and outdoor scenes (ESL-STUP). This innovative framework takes into account both the temporal correlation and spatial correlation of trajectories, a significant advancement over existing models. Detailed contributions are described as follows:

1) An iPDR-based trajectory estimation structure is proposed, which integrates Inertial Navigation System (INS), magnetic observations, Pedestrian Dead Reckoning (PDR) mechanization, and deep-learning based speed estimation to realize more accurate sensor-based positioning magnetic interference contained urban scenes.

2) This paper develops a novel attentive-LSTM network for the uncertainty modelling of hybrid location sources using a period of human motion features extracted from different indoor and outdoor location sources including inertial sensors, GNSS, and Wi-Fi, and realizes the prediction of time-varying uncertainty errors of each.

3) The predicted uncertainty errors of different indoor and outdoor location sources are then integrated with iPDR to realize robust seamless positioning performance. This novel approach significantly improves the accuracy and reliability of location information acquisition in both indoor and outdoor environments, even under obscured scenarios. The effectiveness of the proposed ESL-STUP is validated through comprehensive experiments, which demonstrate its superior performance compared to existing multi-source fusion-based seamless positioning structures under different urban scenes.

Organized for clarity and depth, the paper unfolds as follows: Section 2 indicates the proposed uncertainty predictor enhanced seamless positioning structure. Section 3 describes the experimental results. Section 4 concludes this paper and indicates the future work.

2. Uncertainty Predictor Enhanced Seamless Positioning Structure

In this section, the methodology of the proposed uncertainty predictor enhanced seamless positioning structure, that includes the proposed iPDR-based trajectory estimation algorithm, attentive-LSTM based uncertainty error predictor, and final multi-source fusion structure.

2.1 iPDR-based Trajectory Estimation

The proposed iPDR combines the advantages of traditional positioning models with data-driven methods, providing more observations and constraints to reduce the cumulative error of inertial sensors and improve the performance of crowdsourced daily life trajectories collected from public users. In this case, the INS mechanization is adopted as the state equation:

$$\delta \mathrm{x} = \left[\left(\delta \mathrm{p}^n\right)_{1\times3} \quad \left(\delta v^n\right)_{1\times3} \quad \phi_{1\times3} \quad \left(\varepsilon_g^b\right)_{1\times3} \quad \left(\nabla_a^b\right)_{1\times3} \right]^T \tag{1}$$

where $\delta \mathrm{p}^n$, δv^n, and $\phi_{1\times3}$ are the state vectors in the navigation coordinate system, including three-dimensional position, speed, and attitude; ε_g^b and ∇_a^b are the biases of the inertial sensors. The state update equation of the overall framework is modeled as:

$$\begin{bmatrix} \delta \dot{p}^n \\ \delta \dot{v}^n \\ \dot{\phi} \\ \dot{\varepsilon}_g^b \\ \dot{\nabla}_a^b \end{bmatrix} = \begin{bmatrix} \delta v^n \\ f_{sf}^n \times \phi + C_b^m (\nabla_a^b + w_\nabla) \\ -C_b^n (\varepsilon_g^b + w_\varepsilon) \\ \varepsilon_g^b / \tau_{bg} + w_{\varepsilon g} \\ \nabla_a^b / \tau_{ba} + w_{\nabla a} \end{bmatrix} \tag{2}$$

where f_{sf}^{n} represents the acceleration conversion data from the carrier coordinate system to the navigation coordinate system; τ_{bg} and τ_{ba} represent the relevant parameters; $w_{\varepsilon g}$ and $w_{\nabla a}$ are the noise of the measurements w_{ε} and w_{∇}.

To implement the proposed method for estimating pedestrian trajectories, we use the Long Short-Term Memory (LSTM) to predict walking speed based on features extracted from sensors. Before inputting the built-in sensor data into the designed deep learning framework, the sensor data is pre-processed and simplified to enhance the network's ability to adapt to human motion features. The proposed trajectory estimator uses 9-axis sensor data to update the attitude quaternion in real-time. Finally, using the updated quaternion, the inertial sensor data is transformed from the carrier coordinate system to the navigation coordinate system. Formula (3) is applied to transform the vector in the carrier coordinate system into variables under n-frame:

$$AG_{INS} = q \times AG_{IMU} \times q^{-} \tag{3}$$

where AG_{INS} refers to the transformed IMU data, which is further modeled as the input vector of the trajectory estimator.

To improve the prediction effect of the deep learning architecture, this study adopts a sliding window method. The transformed IMU data is divided into discrete sequences through a fixed-size sliding window, with the sequence window dimension (n) being 100 frames (1s) and the stride being 10 frames. The deep neural network predicts the speed vector for each sequence, and the linking module combines these predictions to generate a position. The real-time position is refreshed based on the previous 200 frames, using overlapping windows for prediction correction. The output frequency is reduced to 20Hz, followed by low-pass filtering, and the controller performs additional operations on the position extrapolated from the speed vector to generate a seamlessly reconstructed predicted trajectory.

2.2 Attentive-LSTM based Uncertainty Error Predictor

This paper introduces an attentive-LSTM network design that combines the more sophisticated Long Short-Term Memory (LSTM) 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-LSTM for handling the input sequence $X = [x_1, x_2, \ldots, x_T]$. This transformed sequence is then passed through a multi-head attention module. The output with refined attention is subsequently inputted into a LSTM module for capturing the temporal dependencies, and the final step involves mapping the LSTM’s output to a one-dimensional vector via another fully connected layer, which serves as the prediction for the estimated uncertainty 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:

$$A = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right), \tag{4}$$

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:

$$X' = \text{MultiHead}(Q,K,V)\text{Concat}(\text{head}_1,\ldots,\text{head}_H)W^O, \tag{5}$$

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.

Then the pre-processed data by multi-head attention layer will be passed to the LSTM layer, and the description of LSTM structure can be found in (Liu et al. 2022). Finally, the output of the LSTM layer is passed through a fully connected layer to obtain the estimated uncertainty error. The loss function is defined as the mean squared error between the predicted error and the ground truth.

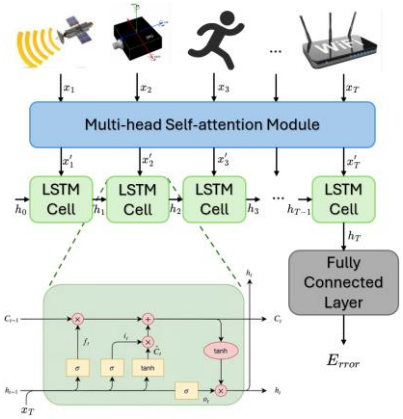

Figure 1. The Structure of Proposed Attentive-LSTM.

In this work, instead of detecting specific indoor/outdoor positioning scenarios and corresponding location sources, an error prediction model based on Attentive-LSTM is adopted to provide indiscriminate uncertainty error prediction results of different location sources. Three different sets of features are extracted from the corresponding location sources for learning. The basic feature group extracts common features from location sources based on GNSS, Wi-Fi fingerprinting, and MEMS sensors, effectively covering indoor and outdoor scenarios:

(1) The stride length and heading vector calculated during the adjacent Wi-Fi fingerprint recognition originating location or the GNSS location reported by the terminal:

$$Feature_{MEMS}(\text{k}) = \begin{pmatrix} S_1^{W/G} & \varpi_1^{W/G} \\ S_2^{W/G} & \varpi_2^{W/G} \\ \ldots & \ldots \\ S_M^{W/G} & \varpi_M^{W/G} \end{pmatrix} \tag{6}$$

where, M represents the number of steps during the adjacent collection positions; $S_M^{W/G}$ and $\varpi_M^{W/G}$ are the corresponding stride length and heading values during the adjacent reported positions.
(2) The relative movement distances calculated by the position sources based on Wi-Fi/GNSS and MEMS sensors, respectively:

$$\begin{cases} Dis_{\text{k}}^{MEMS} = \sqrt{\left(\sum_{\text{i}=1}^{M} S_{\text{i}} \cdot \cos(\varpi_{\text{i}})\right)^2 + \left(\sum_{\text{i}=1}^{M} S_{\text{i}} \cdot \sin(\varpi_{\text{i}})\right)^2} \\ Dis_{\text{k}}^{Wi-Fi/GNSS} = \sqrt{\left(\text{x}_G^W(k) - \text{x}_G^W(k-1)\right)^2 + \left(y_G^W(k) - y_G^W(k-1)\right)^2} \end{cases} \tag{7}$$

where $\left\{\text{x}_G^W(k), y_G^W(k)\right\}$ indicates the Wi-Fi/GNSS reported location.
(3) Real-world sampling rate and relative speed:

$$\begin{cases} Speed_{\text{k}}^{MEMS} = Dis_{\text{k}}^{MEMS} / \tau_k^{MEMS} \\ Speed_{\text{k}}^{Wi-Fi/GNSS} = Dis_{\text{k}}^{Wi-Fi/GNSS} / \tau_k^{Wi-Fi/GNSS} \end{cases} \tag{8}$$

where $Speed_{\mathrm{k}}^{MEMS}$ and $Speed_{\mathrm{k}}^{Wi-Fi/GNSS}$ represent the relative speed reference estimated by MEMS sensors and Wi-Fi/GNSS-based positioning sources; τ_k^{MEMS} and $\tau_k^{Wi-Fi/GNSS}$ represent the actual sampling rates of Wi-Fi and GNSS.

For indoor scenarios where GNSS is denied, Wi-Fi fingerprinting is considered the only absolute positioning source. Due to the low sampling rate and RF characteristics of Wi-Fi fingerprinting, specific features are added in addition to the basic features to improve error prediction performance, including:

(1) The generated reference heading values are estimated by Wi-Fi/MEMS-based positioning sources:

$$\begin{cases} \vartheta_{virtual}^{MEMS}(\mathrm{k}) = \arccos\left(\dfrac{\sum_{i=1}^{M} S_i \cdot \sin(\varpi_{\mathrm{i}})}{\sum_{i=1}^{M} S_i \cdot \cos(\varpi_{\mathrm{i}})}\right) \\ \vartheta_{virtual}^{Wi-Fi}(\mathrm{k}) = \arctan\left(\dfrac{x_G^W(k) - x_G^W(k-1)}{y_G^W(k) - y_G^W(k-1)}\right) \end{cases} \tag{9}$$

(2) The real-time received signal strength indication (RSSI) and the average RSSI difference value of the matched nearest database reference vector:

$$\begin{aligned} &\Delta RSSI(\mathrm{k}) = \frac{\sum_{\mathrm{j}=1}^{\beta}\left(RSSI_{\mathrm{j}}^{\mathrm{scanned}} - RSSI_{\mathrm{j}}^{\mathrm{reference}}\right)}{\beta} \\ &\text{s.t. :}\min\left(RSSI_k^{\mathrm{scanned}} - RSSI^{reference}\right) \end{aligned} \tag{10}$$

where $RSSI_{\mathrm{j}}^{\mathrm{scanned}}$ indicates the smartphone-reported RSSI vector, $RSSI_{\mathrm{j}}^{\mathrm{reference}}$ is the database-reported RSSI vector.

For outdoor spaces, GNSS is considered the only absolute position source. Therefore, the following improved input features are added for model-based transfer learning:

(1) The average value of a set of Signal-to-Noise Ratio (SNR) parameters: The SNR feature can describe the quality of a specific satellite signal, and this work averages it as an input value for the training stage:

$$SNR_{\mathrm{mean}}(k) = \frac{\sum_{\eta=1}^{\alpha} SNR_{\eta}(k)}{\alpha} \tag{11}$$

where α indicates the number of observed satellites, $SNR_{\eta}(k)$ is SNR value of each satellite.

(2) The ratio of precise satellites to received satellites: At least 4 satellites are required for each positioning stage. There is also a relative correspondence between the number of satellites received and the satellites with a high signal-to-noise ratio:

$$\chi_{Ratio}(k) = \frac{\delta_e(k)}{\delta_{total}(k)} \tag{12}$$

where $\delta_{total}(k)$ the number of observed satellites, $\delta_e(k)$ represents the qualified satellites with a high signal-to-noise ratio.

2.3 Structure of Proposed ESL-STUP

In this paper, the uncertainty error prediction results of indoor and outdoor location sources provided by the Attentive-LSTM method are further applied to the measurement error reference in the multi-source fusion stage, which can also realize the autonomous switching of indoor and outdoor scenes, replacing the traditional scene detection-based methods. In order to suppress the cumulative error of the inertial navigation system, the speed, position, and magnetic vector originated from iPDR are integrated with attentive-LSTM provided uncertainty errors of different indoor and outdoor location sources:

$$\delta z^n = \begin{bmatrix} \delta z_v^n \\ \delta z_r^n \\ \delta z_m^n \end{bmatrix} = \begin{bmatrix} v_{PDR}^n - v_{INS}^n \\ r_{PDR}^n - r_{INS}^{\mathrm{n}} \\ C_{\mathrm{n,t}}^b \cdot m_t^b - M^n \end{bmatrix} \tag{13}$$

where δz_v^n, δz_r^n, and δz_m^n are the observation residual vectors of position, speed, and magnetic field, v_{PDR}^n and r_{PDR}^n represent the starting position and speed of PDR, v_{INS}^n and r_{INS}^{n} represent the starting position and speed of INS, m_t^b is the magnetic vector extracted in the QSMF cycle, $C_{\mathrm{n,t}}^b$ is the current attitude matrix, M^n represents the reference magnetic vector within the identified QSMF period (Renaudin et al. 2019).

Finally, a loosely coupled fusion structure was developed for observation data based on GNSS and Wi-Fi fingerprinting, which skips the IO detection process:

$$\begin{cases} \delta z_r^n = r_{GNSS/Wi-Fi}^n - r_{Sensor}^n \\ \delta z_v^n = v_{GNSS/Wi-Fi}^n - v_{Sensor}^n \end{cases} \tag{14}$$

where $r_{GNSS/Wi-Fi}^n$ and $v_{GNSS/Wi-Fi}^n$ indicate the GNSS and Wi-Fi originated location and speed.

In this paper, an uncertainty error assessment model based on Attentive-LSTM is used to provide corresponding weights for different types of location sources. The predicted positioning error is modelled as a Gaussian distribution and combined with the above multi-source fusion model:

$$\begin{cases} p(x_k^{GNSS/Wi-Fi} \mid x_k^j) \sim N_x(0, \sigma_{GNSS/Wi-Fi}^2) \\ p(y_k^{GNSS/Wi-Fi} \mid y_k^j) \sim N_y(0, \sigma_{GNSS/Wi-Fi}^2) \end{cases} \tag{15}$$

where $\sigma_{GNSS/Wi-Fi}^2$ represents the reference error variance of the Wi-Fi and GNSS reported results, which is further used as an adaptive parameter for the multi-source fusion stage.

3. Experimental Results

3.1 Experimental Dataset Collection

Our datasets contain the different location sources including Wi-Fi fingerprinting, sensors-originated motion, and GNSS, collected under different indoor and outdoor environments. The Wi-Fi fingerprinting dataset is generated based on the open-source dataset provided by IPIN-2018 (Renaudin et al. 2019) and IPIN-2020 (Potortì et al. 2021) international indoor positioning competitions, track 3, and the GNSS dataset is collected in typical indoor and outdoor scenes in the Hong Kong Polytechnic University. The sensors and GNSS datasets are collected with the assistance of Lidar-based Simultaneous localization and mapping (SLAM) system (Bao et al. 2022), which can provide centi-meter level location reference, and the collected GNSS trajectories and SLAM-provided reference trajectories is compared in Figure.2.

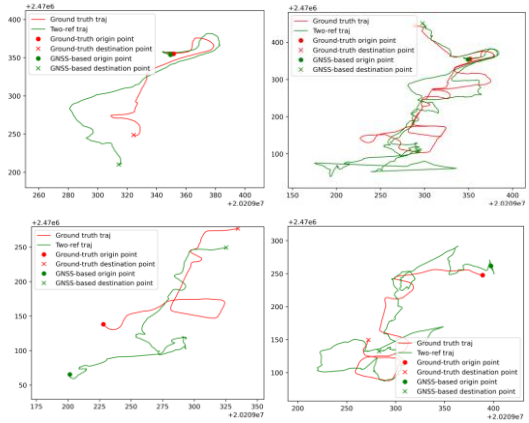

Figure 2. GNSS Dataset.

3.2 Model Training and Baseline Methods

To evaluate the performance of proposed Attentive-LSTM based uncertainty error prediction model, we choose existing deep-learning models including Bi-directional Long Short-Term Memory (Bi-LSTM) (Yu et al. 2023), Gate Recurrent Unit (GRU) (Zhou et al. 2022), and LSTM (Liu et al. 2022) for accuracy comparison.

For fairness, all models were implemented using PyTorch and trained on the same hardware server. 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 500 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

This paper proposes Attentive-LSTM network to evaluate the uncertainty error of different indoor and outdoor location sources, and the performance of Attentive-LSTM network is evaluated using generated GNSS and Wi-Fi fingerprinting datasets, respectively. We choose average prediction error for accuracy comparison under two different datasets, and the prediction results are described in Figure 3:

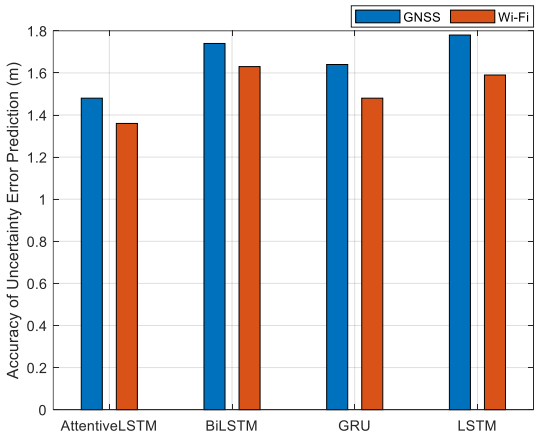

Figure 3. Performance of Uncertainty Error Prediction

It can be found from Figure 3 that the proposed Attentive-LSTM network proves comparable performance under GNSS and Wi-Fi fingerprinting datasets, and outperforms than the existing Bi-LSTM, GRU, and single LSTM networks. The average uncertainty error prediction results under GNSS dataset using four different networks are 1.48 m, 1.74 m, 1.64 m, 1.78 m, and the average uncertainty error prediction results under Wi-Fi fingerprinting dataset using four different networks are 1.36 m, 1.63 m, 1.48 m, 1.59 m, respectively.

3.4 Performance Evaluation of Proposed ESL-STUP

The proposed ESL-STUP structure is evaluated in a typical office environment contains both indoor and outdoor areas. The test route of pedestrian is described in Figure 4, and the estimated trajectories and positioning accuracy provided by iPDR, Wi-Fi fingerprinting, GNSS, and final ESL-STUP are compared in Figure 5 and Figure 6.

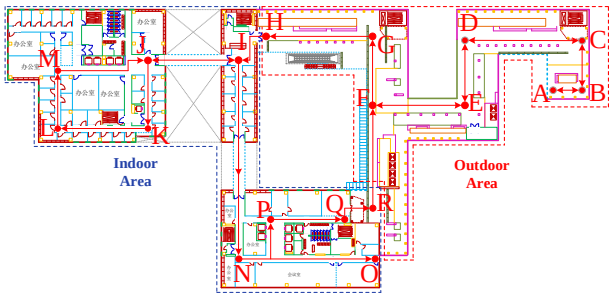

Figure 4. Test Route and Environment

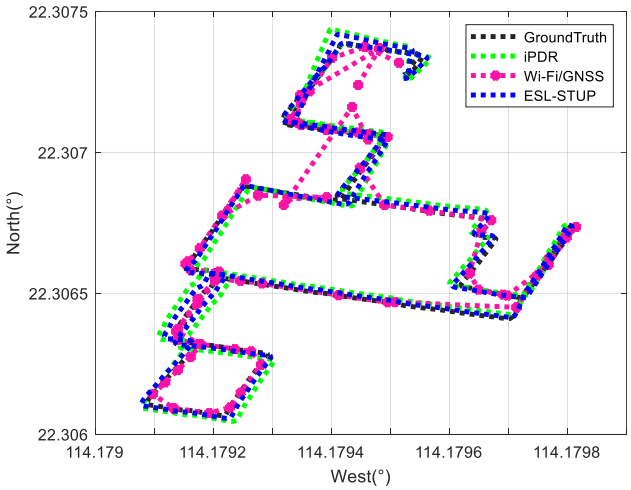

Figure 5. Trajectory Comparison

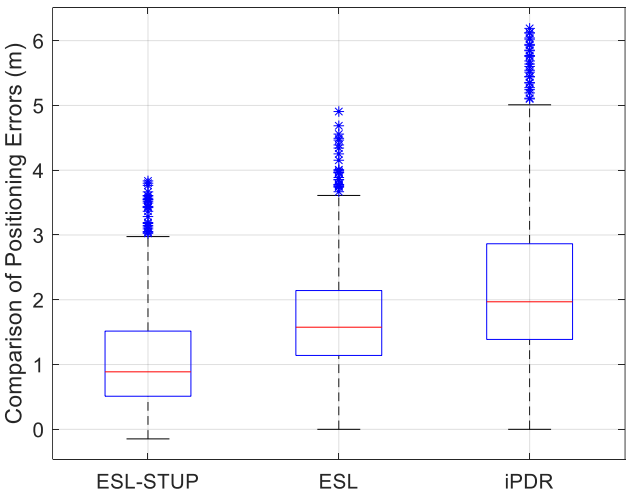

Figure 6. Positioning Accuracy Comparison

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-4-2024-493-2024 | © Author(s) 2024. CC BY 4.0 License.

It can be found from Figure 5 and Figure 6 that the proposed iPDR realizes high-precision performance of trajectory estimation, and the final proposed ESL-STUP further improves the performance of iPDR algorithm, using the predicted uncertainty error of different location sources to adjust the weight of each in the procedure of fusion phase, which outperforms ESL algorithm without uncertainty prediction of different location sources. The estimated positioning accuracy of proposed iPDR, ESL, and ESL-STUP reaches 2.86 m, 2.14 m, and 1.52 m respectively.

4. Conclusion

In conclusion, this paper introduces an advanced seamless localization framework, termed enhanced seamless localization with spatial-temporal uncertainty predictor (ESL-STUP), designed to operate effectively in environments where indoor and outdoor scenes are obscured. The ESL-STUP framework incorporates both temporal and spatial correlations of trajectory data sourced from Wi-Fi, GNSS, and sensor-derived motion information. It features an innovative trajectory estimation approach based on iPDR structure. This structure synergizes the mechanics of INS/PDR, magnetic field measurements, and a deep learning algorithm for precise speed estimation, thereby augmenting the accuracy of conventional PDR methods.

The framework models a sequence of human motion characteristics drawn from a combination of location data sources, rather than relying solely on singular or adjacent positional readings. This approach enables the prediction of time-varying measurement uncertainties. By integrating these predicted uncertainty errors from various indoor and outdoor location sources with the iPDR system, ESL-STUP achieves a robust and seamless positioning capability.

Extensive experimental evaluations demonstrate that the ESL-STUP framework significantly outperforms existing multi-source fusion-based seamless positioning systems, delivering superior accuracy across diverse environmental conditions.

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., Xiang, H., & Yu, Y. (2022). A systematic mapping framework for backpack mobile mapping system in common monotonous environments. *Measurement*, 197, 111243.

Di Pietra, V., Dabove, P., & Piras, M. (2020). Loosely coupled GNSS and UWB with INS integration for indoor/outdoor pedestrian navigation. *Sensors*, 20(21), 6292.

Dong, J., Noreikis, M., Xiao, Y., & Ylä-Jääski, A. (2018). ViNav: A vision-based indoor navigation system for smartphones. *IEEE Transactions on Mobile Computing*, 18(6), 1461-1475.

Jiang, W., Cao, Z., Cai, B., Li, B., & Wang, J. (2021). Indoor and outdoor seamless positioning method using UWB enhanced multi-sensor tightly-coupled integration. *IEEE Transactions on Vehicular Technology*, 70(10), 10633-10645.

Li, Y., Gao, Z., He, Z., Zhuang, Y., Radi, A., Chen, R., & El-Sheimy, N. (2019). Wireless fingerprinting uncertainty prediction based on machine learning. *Sensors*, 19(2), 324.

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

Luo, Y., Guo, C., Su, J., Guo, W., & Zhang, Q. (2020). Learning-based complex motion patterns recognition for pedestrian dead reckoning. *IEEE Sensors Journal*, 21(4), 4280-4290.

Potortì, F., Torres-Sospedra, J., Quezada-Gaibor, D., Jiménez, A. R., Seco, F., Pérez-Navarro, A., ... & Oh, H. L. (2021). Off-line evaluation of indoor positioning systems in different scenarios: The experiences from IPIN 2020 competition. *IEEE Sensors Journal*, 22(6), 5011-5054.

Qin, T., Li, P., & Shen, S. (2018). Vins-mono: A robust and versatile monocular visual-inertial state estimator. *IEEE transactions on robotics*, 34(4), 1004-1020.

Renaudin, V., Ortiz, M., Perul, J., Torres-Sospedra, J., Jiménez, A. R., Perez-Navarro, A., ... & Park, Y. (2019). Evaluating indoor positioning systems in a shopping mall: The lessons learned from the IPIN 2018 competition. *IEEE Access*, 7, 148594-148628.

Sakamoto, Y., Ebinuma, T., Fujii, K., & Sugano, S. (2012, May). GPS-compatible indoor-positioning methods for indoor-outdoor seamless robot navigation. In 2012 *IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)* (pp. 95-100). IEEE.

Shi, W., Yu, Y., Liu, Z., Chen, R., & Chen, L. (2022). A deep-learning approach for modelling pedestrian movement uncertainty in large-scale indoor areas. *International Journal of Applied Earth Observation and Geoinformation*, 114, 103065.

Wan, Q., Duan, X., Yu, Y., Chen, R., & Chen, L. (2022). Self-calibrated multi-floor localization based on wi-fi ranging/crowdsourced fingerprinting and low-cost sensors. *Remote Sensing*, 14(21), 5376.

Wang, Y., Kuang, J., Li, Y., & Niu, X. (2022). Magnetic field-enhanced learning-based inertial odometry for indoor pedestrian. *IEEE Transactions on Instrumentation and Measurement*, 71, 1-13.

Wen, W., Pfeifer, T., Bai, X., & Hsu, L. T. (2021). Factor graph optimization for GNSS/INS integration: A comparison with the extended Kalman filter. *NAVIGATION: Journal of the Institute of Navigation*, 68(2), 315-331.

Wen, W. W., Zhang, G., & Hsu, L. T. (2019). GNSS NLOS exclusion based on dynamic object detection using LiDAR point cloud. *IEEE transactions on intelligent transportation systems*, 22(2), 853-862.

Weyn, M., Klepal, M., & Widyawan. (2009, September). Adaptive motion model for a smart phone based opportunistic localization system. *In International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments* (pp. 50-65). Berlin, Heidelberg: Springer Berlin Heidelberg.

Wu, C., Yang, Z., Xu, Y., Zhao, Y., & Liu, Y. (2014). Human mobility enhances global positioning accuracy for mobile phone

localization. *IEEE Transactions on Parallel and Distributed Systems*, 26(1), 131-141.

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

Yu, Y., Chen, R., Chen, L., Li, W., Wu, Y., & Zhou, H. (2021). A robust seamless localization framework based on Wi-Fi FTM/GNSS and built-in sensors. *IEEE Communications Letters*, 25(7), 2226-2230.

Yu, Y., Chen, R., Chen, L., Li, W., Wu, Y., & Zhou, H. (2021). H-WPS: Hybrid wireless positioning system using an enhanced wi-fi FTM/RSSI/MEMS sensors integration approach. *IEEE Internet of Things Journal*, 9(14), 11827-11842.

Yu, Y., Yao, Y., Liu, Z., An, Z., Chen, B., Chen, L., & Chen, R. (2023). A Bi-LSTM approach for modelling movement uncertainty of crowdsourced human trajectories under complex urban environments. *International Journal of Applied Earth Observation and Geoinformation*, 122, 103412.

Zhang, J., Yu, L., Li, X., Zhang, C., Shi, T., Wu, X., ... & Wu, G. (2020). Exploring annual urban expansions in the Guangdong-Hong Kong-Macau Greater Bay Area: Spatiotemporal features and driving factors in 1986–2017. *Remote Sensing*, 12(16), 2615.

Zhang, M., Wen, Y., Chen, J., Yang, X., Gao, R., & Zhao, H. (2018). Pedestrian dead-reckoning indoor localization based on OS-ELM. *IEEE Access*, 6, 6116-6129.

Zhou, J., Qin, Y., Chen, D., Liu, F., & Qian, Q. (2022). Remaining useful life prediction of bearings by a new reinforced memory GRU network. *Advanced Engineering Informatics*, 53, 101682.