Efficient Indoor Localization with Multiple Consecutive Geomagnetic Sequences

Hui Zhuang, Tao He, Qun Niu and Ning Liu*
Sun Yat-sen University
{zhuangh8,hetao23}@mail2.sysu.edu.cn, {niuq3,liuning2}@mail.sysu.edu.cn

2022 International Conference on Computer Communications and Networks (ICCCN) | 978-1-6654-9726-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCCN54977.2022.9868904

***Abstract*—Geomagnetism-based indoor localization has great social and commercial value due to its pervasiveness and independence from extra infrastructure. To improve the distinguishability of geomagnetic signals as location clues, geomagnetic sequences are usually taken as input. Although longer input sequence can provide higher localization accuracy, it suffers from high response time in practice. To address the above, we first utilize short geomagnetic sequences as input, alleviating high response time, and propose an efficient single position estimation model, taking advantage of modified transformer to estimate position for each independent short sequence. Noticing the temporal dependency and the spatial consistency constraint during continuous positioning, we further propose a joint position estimation model to capture the correlations among consecutive short sequences, achieving higher accuracy with multiple short sequences. We have conducted extensive experiments in two typical trial sites, a narrow office area and a spacious parking lot. Experimental results show that the proposed approach outperforms state-of-the-art competing schemes, and the localization error is reduced by more than 32% with shorter geomagnetic sequences.**

I. Introduction

With the rapid development of Internet of Things (IoT), indoor localization techniques play an increasingly important role, empowering a wide range of applications in commercial marketing [1], assisted living [2] and crowd monitoring [3].

Suffering from the loss of satellite signals by obstructions from buildings in the indoor environment, the mature satellite-based localization techniques (like GPS) are usually invalidated indoors. Therefore, researchers have studied a plenty of other signals for indoor localization, e.g., Wi-Fi [4], ultra-wide band (UWB) [5], Bluetooth [6], geomagnetism [7] and vision [8]. Limited to the signal characteristics, most of these signals require the deployment of extra infrastructure, incurring extra deployment cost [9]. Benefiting from the pervasiveness and independence from extra infrastructure, geomagnetism has attracted attention to indoor localization. Geomagnetism-based approaches achieve high localization accuracy with negligible system deployment cost.

Intuitively, some geomagnetism-based indoor localization approaches directly leverage the single geomagnetic signal observation as the location clue to target position [10, 11]. However, a single geomagnetism signal observation obtained by the device usually consists of only three component values in the axes, which lack enough distinguishability as location clues especially in large indoor scenes. As a result, these approaches are prone to being inefficient and are easily affected by the noise of the indoor environment, leading to large localization errors consequently.

In order to improve the discrimination of location features, some researchers propose making use of multiple continuous geomagnetic signal observations, a geomagnetic sequence, for localization [12–14]. The approaches leverage the geomagnetic sequence as input and extract temporal correlations among continuous geomagnetic signal observations to generate location feature for positioning. Compared with the single geomagnetic signal observation, the geomagnetic sequence composed of multiple continuous geomagnetic signals carries more distinguishing positioning features under temporal constraint, thus being capable of providing more accurate localization.

Naturally, more distinguished location features can be extracted when using longer geomagnetic sequences. And some recent studies using geomagnetic sequences have been able to achieve high positioning accuracy with longer input sequences [7, 15]. These methods typically use geomagnetic sequences collected over a period of time to determine a final location. However, it takes a long time for the user to collect enough geomagnetic signals to get the target position at cold-start conditions, which is not user-friendly. And it is also difficult to ensure that the user has the patience to collect a sufficient length sequence of geomagnetic signals in a real random walk scenario. In addition, using long input sequences brings a great challenge for establishing a geomagnetic sequence dataset since the user usually walks randomly in real-world scenarios. Therefore, the shorter sequence is more adaptable in practice. However, as mentioned above, using shorter sequence faces the problem that how to obtain location features with high distinguishability, which are directly related to the performance of localization.

To address the above challenges, we propose an efficient localization approach for continuous positioning scenes (most common in reality). We first take short geomagnetic sequences as input and employ state-of-the-art transformer [16] as a basis, devising a single position estimation model to efficiently extract temporal features of geomagnetic sequence for coarse positioning estimation as the first step. On this basis, considering the consistency constraint of continuous positioning, we further propose a joint position estimation model utilizing multiple consecutive sequences, which combines the temporal dependency of consecutive geomagnetic observations and

978-1-6654-9726-8/22/$31.00 ©2022 IEEE

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on September 14,2022 at 11:18:13 UTC from IEEE Xplore. Restrictions apply.

the spatial consistency constraint among multiple consecutive sequences. With efficient location extraction and consistency constraint, the proposed approach achieves high localization accuracy even for short sequences. In summary, we make the following contributions:

- We propose a single position estimation model based on transformer to efficiently extract temporal features of independent short sequence for localization, and efficiently improve localization accuracy especially with short input sequence, compared with other existing approaches.
- Considering the consistency constraint during continuous positioning, we devise a joint position estimation model. We first encode this global constraint with a multi-layer structure, then jointly predict the positions for consecutive short sequences with a serial decoder. And the predictions from the single position estimation model are employed to alleviate the accumulated error of serial decoding.
- We have conducted extensive experiments in two typical trial sites. Experimental results show that the proposed approach achieves higher accuracy compared with other state-of-the-art competing schemes, when using different lengths of geomagnetic sequences. And the improvement effect is more significant for the shorter sequences.

In addition to the geomagnetic signal, the proposed approach can also be applied to other types of sequential signal, such as visible light [17] and radio signal [18], for localization.

The remainder of the paper is structured as follows: We review the existing work related to our approach in Section II and explain the calibration process of geomagnetic signals in Section III. Then, we elaborate on our approach in section IV. We present illustrative experimental results in Section V and conclude in Section VI.

II. Related Work

We review related work as follows. Some researchers use discrete geomagnetic signals collected at user locations to match with geomagnetic fingerprint database to achieve localization. In many works, discrete geomagnetic signals need to be fused with other signals to solve the problem of insufficient discrimination of positioning features. For example, the work [19] proposes a fingerprint localization algorithm combining channel state information (CSI) and geomagnetic field intensity, and leverages a multi-scale K-nearest Neighbor (MDS-KNN) method to achieve fingerprint matching localization. The positioning system LMDD [20] proposes a detection framework based on the majority voting model, which fuses geomagnetic signals with other signals.

Some researchers use continuous geomagnetic signals collected during movement to improve the discrimination of positioning features. Travi-Navi [21] leverages dynamic time warping (DTW) for localization, which considers both stretching and squeezing sequences to align them. Magicol [22] designs a bidirectional particle filter, which uses DTW to update the weight of particles. However, such methods require a lot of sequence matching work, leading to high computational cost. Some recent approaches employ neural networks to process sequential inputs for localization. Recently, some researches [23, 24] leverage RNN to localize with geomagnetic sequence. DeepML [12] uses LSTM to extract the fusion location features of geomagnetic sequence signals and visible light signals for indoor localization. MAIL [7] leverages the attention mechanism to extract multi-scale features of geomagnetic sequence to predict user position. However, most of these work require long geomagnetic sequences as input, which is not applicable in practice.

Transformers [16] are a recently developed class of deep learning models, which are suitable for processing time series data. Transformers perform feature extraction of contextual information based on an attention mechanism, and multi-headed attention allows it to consider information from different subspace representations. Due to its efficient parallel computing capability and superior performance, transformers are widely used in many fields, such as NLP [25] and time series processing [26].

III. Calibration of Geomagnetic Signals

When the user holds the device in different poses, the device may collect different geomagnetic signals even in the same position. We calibrate the geomagnetic signals collected in the user coordinate system to the standard coordinate system according to the offset of the device.

The device sensor collects sensor signals $v = \{b, a\}$, where $b = [mag_x, mag_y, mag_z]^T$ indicates the raw 3-axis geomagnetic signal magnitude, and $a = [a_x, a_y, a_z]^T$ indicates acceleration vector. We assume ϕ, θ,ψ denote corresponding the roll angle, pitch angle, and yaw angle respectively,and ϕ, θ,ψ can be calculated as:

$$\phi = arctan(\frac{a_y}{a_z}), \tag{1}$$

$$\theta = arctan(\frac{a_x}{a_z}), \tag{2}$$

$$\psi = arctan(\frac{a_y cos\phi - a_z sin\theta}{a_x cos\theta + a_y sin\theta sin\phi + a_z sin\theta cos\phi}). \tag{3}$$

The rotation matrix A from the user coordinate system to the standard coordinate system can be calculated as:

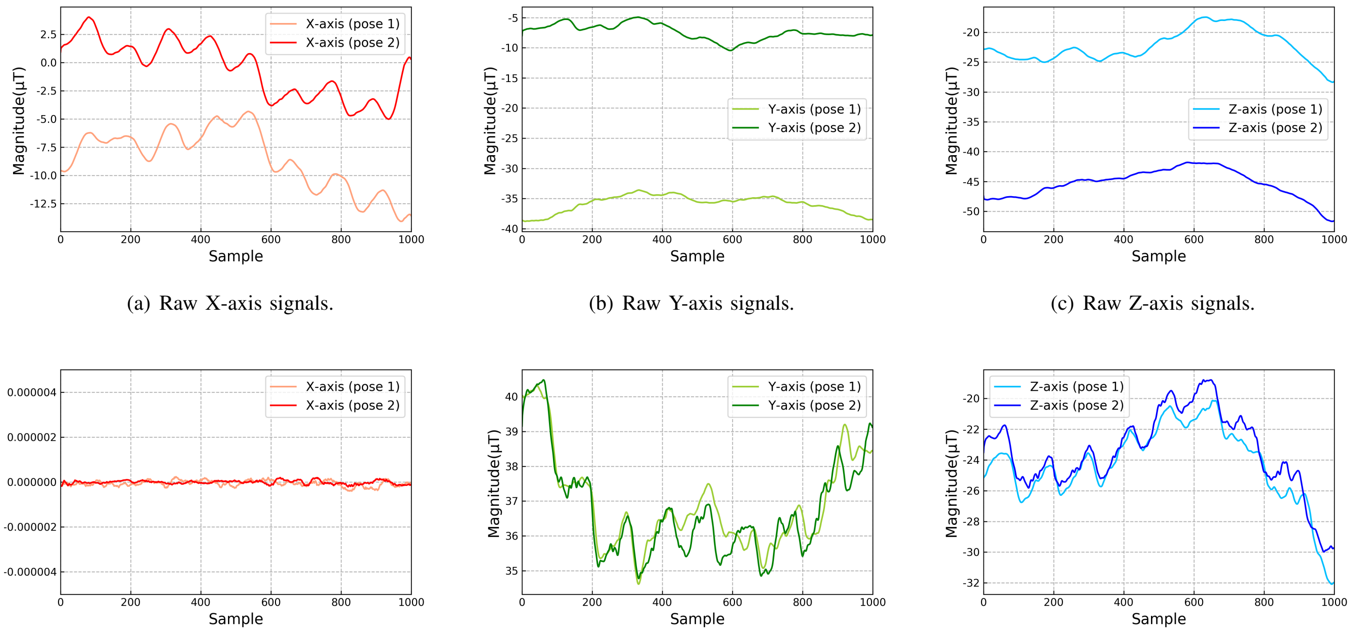
$$A(\phi, \theta, \psi) = \begin{bmatrix} cos\psi & -sin\psi & 0 \\ sin\psi & cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$\begin{bmatrix} cos\theta & 0 & sin\theta \\ 0 & 1 & 0 \\ -sin\theta & 0 & cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\phi & -sin\phi \\ 0 & sin\phi & cos\phi \end{bmatrix}. \tag{4}$$

The geomagnetic signals are calibrated using the rotation matrix A:

$$b_{calibrated} = A(\phi, \theta, \psi) \cdot b. \tag{5}$$

Fig. 1 shows the raw geomagnetic signals collected with different device holding postures and the signals after calibration. It can be seen that the magnitude of calibrated geomagnetic

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on September 14,2022 at 11:18:13 UTC from IEEE Xplore. Restrictions apply.



(a) Raw X-axis signals.

(b) Raw Y-axis signals.

(c) Raw Z-axis signals.

(d) Calibrated X-axis signals.

(e) Calibrated Y-axis signals.

(f) Calibrated Z-axis signals.

Fig. 1: Raw geomagnetic signals collected in different device poses and the signals after calibration.

signals collected from different poses is similar. We notice that the calibrated geomagnetic signals along the X axis are weak, resulting in the geomagnetic signals in this direction hardly fluctuating even if the position changes.

In addition, we calculate the dimensionless magnitude of $b_{calibrated}$, as $mag_{norm} = \|b_{calibrated}\|_2$, to replace the calibrated geomagnetic signals along the X-axis. Then calibrated geomagnetic signals can be modeled as:

$$b_{calibrated} = [mag_y, mag_z, mag_{norm}]^T . \quad (6)$$

For convenience, the calibrated geomagnetic signals below are represented by b.

IV. The Design of the Approach

In this section, we elaborate the core design of the proposed approach. First of all, we overview the overall structure in Section IV-A. Then we present the single position estimation for the single independent sequence in Section IV-B, which is a preliminary work. Finally, we elaborate the details of the joint position estimation model in Section IV-C.

A. Overview

The overall framework of the proposed approach is presented in Fig. 2. Firstly, continuous geomagnetic signals are collected by the mobile device while users walk, and collected data is calibrated in real time. For the convenience of representation, Fig. 2 shows a 1-element sequence (actually the geomagnetic sequence is a multivariate sequence). Then continuous geomagnetic signals are segmented into multiple consecutive geomagnetic sequences with specific overlapping by sliding windows. These short geomagnetic sequences are fed into the single position estimation model from ①. With extracted temporal features, the model estimates the corresponding single position independently. Meanwhile, these consecutive short geomagnetic sequences are taken as input to the joint position estimation model from ② to extract the spatio-temporal continuity among them. We also utilize the coarse position produced by the single position estimation model to further enhance the spatial constraint and alleviate the accumulated localization error. Finally, the joint position estimation model outputs the more accurate consecutive estimate positions.

B. Single position estimation for single independent sequence

The positioning approaches using continuous geomagnetic signals usually estimate the single position for a single sequence. It refers to calculating a representative position of the track (usually the position at the end of the track) using geomagnetic signals collected along the track.

We leverage the sliding window to generate many geomagnetic sequences. The continuous geomagnetic signals are segmented into multiple consecutive geomagnetic sequences with specific overlapping by sliding windows. The resulting geomagnetic sequences are expressed as $S = [s_1, s_2, \ldots]$. Each movement of the sliding window produces a snapshot, corresponding to a geomagnetic sequence. We record these sequences and the order they were generated.

In this part, we take the geomagnetic sequence as the sample, and the end position of the track corresponding to the sequence as the positioning target. Due to the temporal continuity of geomagnetic sequence, time series representation models can

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on September 14,2022 at 11:18:13 UTC from IEEE Xplore. Restrictions apply.

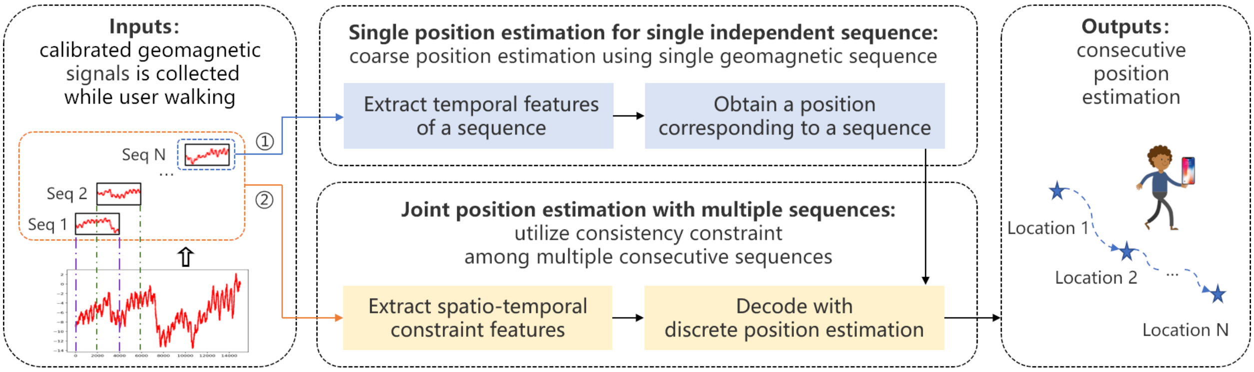


Fig. 2: Overall the proposed approach.

be used to extract temporal features from the geomagnetic sequences. We employ the encoder part of the transformer [16] architecture to extract temporal features and modify the model for regression tasks.

In particular, each sample is a multivariate time series of length w and three different variables, constitutes a sequence of w feature vector b_t: $s \in \mathbb{R}^{w\times 3} = [b_1, b_2, ..., b_w], b_t \in \mathbb{R}^3 = [mag_y, mag_z, mag_{norm}]$. Then s is denoised by outlier cleaning and smoothing filtering. The feature vector b_t is linearly projected onto a d dimensional vector space to obtain the feed-forward output u_t, where d is the dimension of internal representation vectors of the transformer model. In order to make it aware of the sequential nature, it is necessary to add positional encodings $W_{pos} \in \mathbb{R}^{w\times d}$ to the the feed-forward output $U \in \mathbb{R}^{w\times d} = [u_1, u_2, ..., u_w]$: $U' = U + W_{pos}$. We apply the sinusoidal encodings which were originally proposed by [16]. The resulting vector U' is fed into a stack of four decoder blocks. Each encoder block consists of a self-attention layer and a feed-forward layer. Each layer is followed by a normalization layer. The encoder blocks produce a d dimensional vector. The final encoded d dimensional vectors corresponding to all time steps are reshaped into a 1 dimensional vector of length $w \cdot d$, which is linearly mapped to an estimated position.

We define the $L2$ loss function e as the overall mean localization error during training stage:

$$e = \left\| \hat{l} - c \right\|_2, \tag{7}$$

where $\hat{l}$ is the ground truth location corresponding to each one and c is the estimated position.

C. Joint position estimation with multiple sequences

We notice that when the user continues positioning, the observation information (e.g. geomagnetic sequence) seen at each moment is of temporal continuity, and the position corresponding to each moment is of spatial continuity (Fig. 3). We propose to utilize multiple consecutive sequences for joint position estimation. The temporal dependency of consecutive geomagnetic observations and the spatial consistency constraint among multiple consecutive sequences can be used to further improve the localization accuracy.

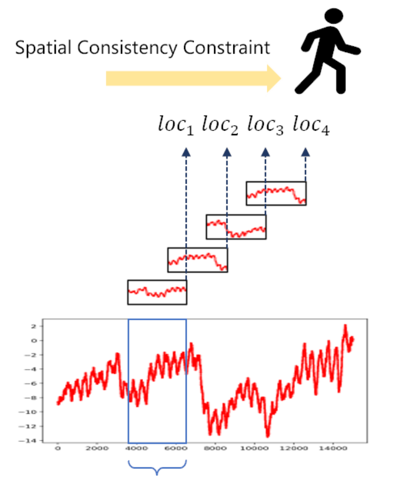


Fig. 3: Temporal and spatial continuity during continuous positioning.

We propose the joint position estimation model to capture spatio-temporal continuity, as shown in Fig. 4. The input of the model is several geomagnetic sequences arranged in the order generated by the sliding window, which represent the observation information at several consecutive spatial positions. The output of the model is several consecutive estimated positions.

1) Spatio-temporal constraint features extraction: During continuous positioning, assuming that M geomagnetic sequences are input to the network, we express the input as $q = \{s_1, s_2, ..., s_M\}$, which contains M geomagnetic sequences corresponding to consecutive M spatial positions.

We devise a multilayer transformer-encoder to extract the spatio-temporal constraint features among M consecutive sequences. The temporal features of the single geomagnetic sequence in q is extracted by a low-layer transformer-encoder respectively. And the low-layer transformer-encoder outputs M 1-dimensional vectors of length $w \cdot d$, which is presented as $F = \{f_1, f_2, ..., f_M\}$.

In order to establish connections between the observation

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on September 14,2022 at 11:18:13 UTC from IEEE Xplore. Restrictions apply.

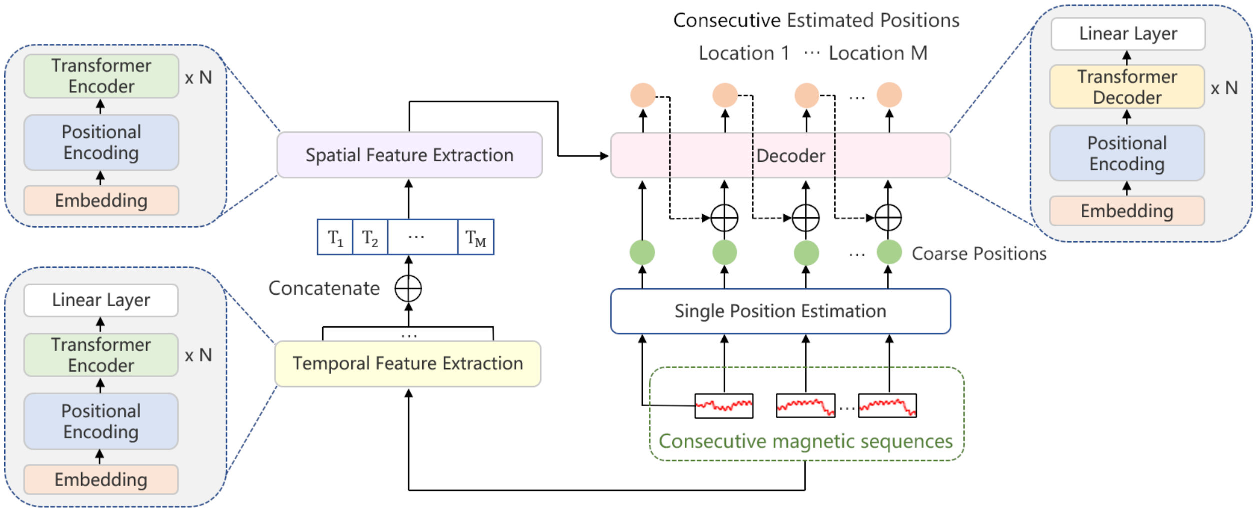


Fig. 4: Architecture of the joint position estimation model.

information of different spatial positions, the low-layer feature representation f corresponding to M positions is taken as time steps, concatenated into a feature sequence vector $\bar{F} \in \mathbb{R}^{M\times(w\cdot d)} = [f_1, f_2, ..., f_M]$ with spatial continuity. Then we employ a high-level transformer-encoder to extract the spatial features. Different from the low-layer transformer-encoder is used to extract the temporal features from the single sequence, the high-level transformer-encoder extracts the spatial consistency constraint features among multiple consecutive sequences. We input $\bar{F}$ into the high-layer transformer-encoder, which produces the global embedding representation $Z \in \mathbb{R}^{M\times d}$ as the spatio-temporal constraint features.

2) Continuous position estimation: We implement continuous positions estimation based on the decoder part of the transformer architecture. It is similar to the transformer-encoder, the transformer-decoder is composed of a feed-forward input layer, a positional encoding layer, and a stack of four decoder blocks. The decoder input begins with a start symbol, and the feed-forward input layer maps the decoder input to a d dimensional input representation vector. In addition to the self-attention layer and the feed-forward layer, the decoder blocks inserts a third layer to apply self-attention mechanisms over the global embedding representation Z. Finally, there is an output layer that maps the output of the last decoder block as the estimated positions corresponding to multiple continuous moments.

In particular, instead of using a fixed symbol (e.g. the dummy start in NLP) as the initial input of the decoder, we take the first element s_1 of q and employ the single position estimation model (section IV-B) to estimate a coarse position c_1 as the initial input of the decoder. Starting with the possible position provides the coarse range estimation for subsequent predictions and enhances spatial constraints.

The joint position estimation model works in parallel in the training stage, and the loss function ε consists of the localization error $L2$ loss and the displacement variance loss. The displacement variance represents the variance of the distance between adjacent localization locations. And ε is defined as follows:

$$\varepsilon = \frac{1}{M}\sum_{i=1}^{M}\left\|\hat{l}_i - l_i\right\|_2 + \frac{1}{M-1}\sum_{i=1}^{M-1}\left(\|l_i - l_{i+1}\|_2 - \overline{\Delta l}\right)^2, \tag{8}$$

$$\overline{\Delta l} = \frac{1}{M-1}\sum_{i=1}^{M-1}\|l_i - l_{i+1}\|_2\,, \tag{9}$$

where M is the number of estimated position, $\hat{l}_i$ is ground truth location corresponding to each sequence, l_i is the estimated position and $\overline{\Delta l}$ is the average adjacent location distance.

In the prediction stage, the joint position estimation model outputs M consecutive estimated positions serially. The serial output results in cumulative errors increasing with the number of predicted rounds. We utilize independent position estimation for the single geomagnetic sequence to alleviate cumulative errors. In each round of prediction, the weighted result of the previous predicted position and the current coarse position is used as input to the decoder. Specifically, the coarse position of the current round is obtained by single position estimation (section IV-B) for the geomagnetic sequence corresponding to the current moment. The input for round i of the decoder is represented as follows:

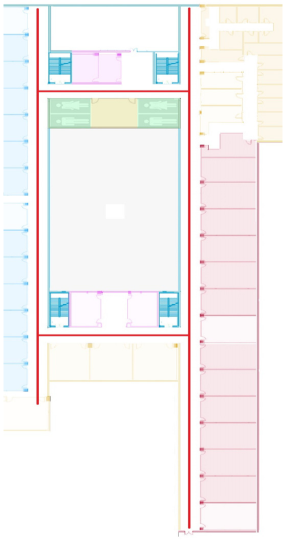
$$x_i = \alpha \cdot l_{i-1} + (1-\alpha)\cdot c_i, \tag{10}$$

where l_{i-1} denotes the predicted position at the previous round, c_i denotes the coarse position at the current round and α denotes the ratio.

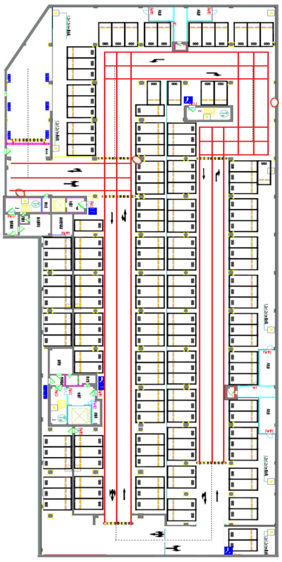
V. Illustrative Experimental Results

We present detailed experimental settings and comparison schemes in Section V-A and illustrative experimental results in Section V-B.

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on September 14,2022 at 11:18:13 UTC from IEEE Xplore. Restrictions apply.



(a) Office area.



(b) Parking lot.

Fig. 5: Floorplans of the trial sites.

A. Experimental Settings and Comparison Schemes

1) *Dataset and experimental setting:* We conduct the experiment in two typical trial sites. One is a narrow office area, the other is a spacious parking lot. The site plans are shown in Fig. 5. The office area covers around 2054 m^2 and the parking lot covers around 2560 m^2. To build datasets, we develop an Android application to collect sensor signals, including geomagnetic signal data and acceleration sensor data. The application continuously records signals along the path as the surveyor walks along the survey path. The sequence of signal values collected from each path can be expressed as $v = \{v_1, v_2, ...\}$, where $v_i = \{a_i, b_i\}$, b_i is the raw 3-axis geomagnetic signal magnitude and a_i represents the acceleration vector. Through the data calibration processing, the calibrated geomagnetic signals can be represented as $b = [mag_y, mag_z, mag_{norm}]^T$. The sampling frequency of signals is 50Hz, and the moving step size of the sliding window is set to 50 for segmenting the geomagnetic sequence.

We also leverage sliding windows to construct multiple consecutive geomagnetic sequences as inputs to the joint position estimation model. The length of the sliding window M depends on the number of geomagnetic sequences input, each of which corresponds to the observation information at a spatial position. The consecutive geomagnetic sequences is represented as $Q = \{q_1, q_2, ..., q_{n-M+1}\}, q_k = \{s_k, ..., s_{k+M-1}\}$.

For training dataset, we design dense survey path in trial sites. And for testing dataset, volunteers are asked to walk through a number of randomly paths. Then geomagnetic signal sequences in each path will be segmented into geomagnetic sequence samples by sliding windows. We have collected 699 and 1390 training sequences in the office area and the parking lot, respectively. For localization, we have collected another 285 and 410 magnetic sequences, respectively. We use training samples from different trial sites to train our model individually and evaluate its performance with test samples from the corresponding trial site. The baseline parameters in our experiment is presented in TABLE. I. We implemented our network with PyTorch[1] and employ the Adam optimizer to update weights.

TABLE I: Baseline parameters in experiment.

Parameters	Value
epoch	1500
batch size	32
initial learning rate	0.0005
learning rate decay	0.0002
dim.model	64
num.heads	8
num.encoder blocks	4
num.decoder blocks	4
num.FFW	256

2) *Comparison schemes:* Our approach will be compared with the following state-of-the-art geomagnetic localization approaches:

- Magicol [22] vectorizes collected sequential geomagnetism based on user's steps and designed an enhanced bi-directional particle filter to estimate the position. It employed DTW to compare geomagnetic sequences with a database to update particles.
- MAIL [7] employs recurrent neural network and the attention mechanism to extract multi-scale features of geomagnetic long sequence to infer current location.
- BiLSTM [27] has been used to extract geomagnetic sequence features for position estimation in recent works [12, 15]. In our experiment, we build a 4-layer BiLSTM network as a comparison scheme.
- Transformer-encoder [16] is an advanced sequence feature extraction structure. We leverage the encoder part of the transformer for location estimation as a comparison scheme (section IV-B).

Compared with the one-to-one of the above approaches (input a single geomagnetic sequence and output a single estimated position), our approach is many-to-many (input multiple geomagnetic sequences and output multiple estimated positions). The length of the single geomagnetic sequence in each set of inputs in our approach is equal to the length of a single input in comparison schemes. In practical localization, our approach reuses the sequence generated at the past moment, and the system only needs to input the newly generated sequences.

B. Experimental Results

We compare our approach with state-of-the-art competing schemes. Fig. 6 shows the CDF of localization error in the office area. The length of sequence is set to 400. Our approach considers the position information of three moments and uses three sequences as inputs. Fig. 6 demonstrates that our approach

[1] https://pytorch.org/

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on September 14,2022 at 11:18:13 UTC from IEEE Xplore. Restrictions apply.

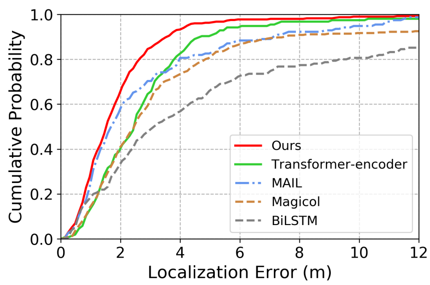


Fig. 6: CDF of location error in the office area.

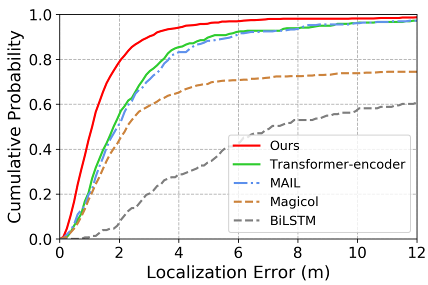


Fig. 7: CDF of location error in the parking lot.

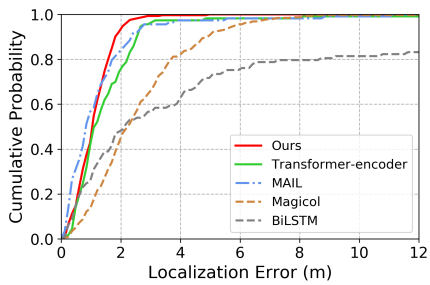


Fig. 8: CDF of location error using longer sequences in the office area.

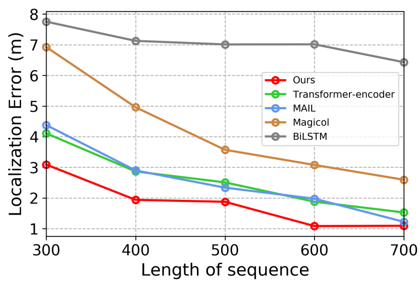


Fig. 9: Mean localization error with different lengths of geomagnetic sequences.

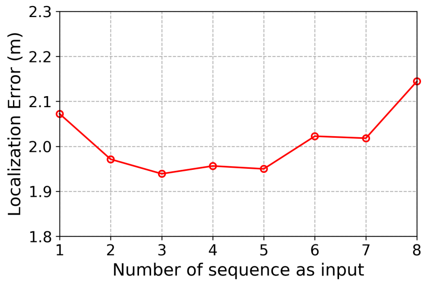


Fig. 10: Mean localization error with different numbers of geomagnetic sequences as input.

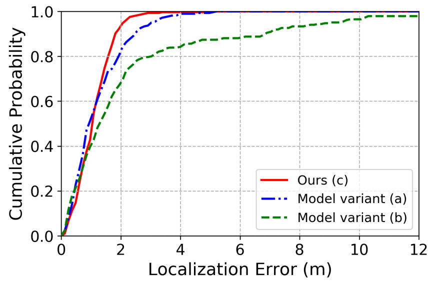


Fig. 11: CDF of localization error of model's variants.

achieves higher accuracy than comparison schemes. Our approach takes into account the geomagnetic information seen at multiple moments and constructs a spatio-temporal connection between those, thus is able to achieve higher overall accuracy.

Fig. 7 shows the CDF of localization error in the parking lot. The length of sequence is 200 and three sequences are used as inputs in our approach. It can be seen that our approach is also able to achieve higher localization accuracy. And some comparison schemes perform worse compared to the office area. This is because the parking lot is more spacious and the discrimination of geomagnetic signals is less. Our approach also achieves sufficient localization accuracy in more spacious trial sites.

Fig. 8 shows the CDF of localization error using longer sequences in the office area. The length of geomagnetic sequence is 700. It can be seen that our approach and some comparison schemes achieve high accuracy when using long sequences for positioning. However, MAIL and transformer-encoder have long tails compared with our approach. The reason is that our approach takes into account the constraint relationship among the consecutive estimation positions, which can avoid large errors compared with the individual position estimation at each moment.

We conducted experiments on different lengths of geomagnetic sequences in the office area. TABLE. II illustrates the mean localization accuracy of different lengths. Fig. 9 shows the change of the mean localization error with different lengths of geomagnetic sequences. It can be seen that the mean localization error increases with the decrease of sequence length. Our approach outperforms state-of-the-art comparison schemes in mean localization error, especially for shorter input sequences. Compared with these comparison schemes, the proposed approach is able to reduce the localization error by around 10% when the length of sequence reaches 700, and by more than 32% when the length decreases to 400.

TABLE II: Mean localization error with different lengths.

	300	400	500	600	700
BiLSTM	7.735	7.126	7.011	7.014	6.431
Magicol	6.926	4.960	3.572	3.078	2.595
MAIL	4.379	2.899	2.337	1.977	1.222
Transformer-encoder	4.111	2.866	2.508	1.879	1.527
Ours	**3.091**	**1.939**	**1.875**	**1.083**	**1.092**

Fig. 10 shows the change of the mean localization error with different numbers of geomagnetic sequences as input. We conducted experiments with sequences of length 400. It can be seen that there is no significant difference in the mean localization error between 2 and 5 geomagnetic sequences. The best effect can be achieved when three geomagnetic sequences are input. After reaching the balance point, the mean localization error increases as the number of input geomagnetic sequences increases. This is because the increase in the number of input sequences leads to more serial position predictions, and the longer serial prediction sequence causes greater cumulative errors.

Fig. 11 illustrates the CDF of localization error of model's

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on September 14,2022 at 11:18:13 UTC from IEEE Xplore. Restrictions apply.

variants. For the decoder module of the model, we evaluate (a) using only the fixed start symbol, (b) using the coarse position instead of the fixed start symbol as the initial input of the decoder, (c) fusing the coarse position in each round predict on the basis of (b). It shows that (a) only using a fixed start symbol has a poor positioning effect, while (b) using the coarse position as the start symbol greatly improves the positioning effect. This is because the spatial constraint is enhanced by the coarse estimation of the starting position, thus the overall localization accuracy is improved. In addition, we add the coarse position to each round of prediction, which further improves the positioning effect. We speculate that the serial prediction of multiple positions will lead to cumulative errors with the increase of rounds, while the individual estimation of each position is able to alleviate the accumulation of localization error.

VI. Conclusion

The localization accuracy of geomagnetic sequences is closely related to the length of the sequences. The shorter sequence may lead to larger localization errors. To address the above, first we devise the single position estimation model for independent short sequence. On this basis, we propose the joint position estimation model with multiple consecutive sequences, which combines the temporal dependency of consecutive geomagnetic observations and the spatial consistency constraint among multiple consecutive sequences to achieve high localization accuracy for the multiple short sequences. We have conducted extensive experiments in two typical trial sites: a narrow office area and a spacious parking lot. Experimental results show that our approach achieves higher localization accuracy compared with other state-of-the-art competing schemes with different sequence length. Furthermore, the localization error is reduced by more than 32% with short geomagnetic sequences.

Acknowledgment

This work is supported in part by the National Natural Science Foundation of China under Grants 61972433 and 62102459, and in part by Guangdong Basic and Applied Basic Research Foundation under Grant 2021A1515012242.

References

[1] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.

[2] B. Li, J. P. Muñoz, X. Rong, Q. Chen, J. Xiao, Y. Tian, A. Arditi, and M. Yousuf, "Vision-based mobile indoor assistive navigation aid for blind people," *IEEE Transactions on Mobile Computing*, vol. 18, no. 3, pp. 702–714, 2019.

[3] T. Kulshrestha, D. Saxena, R. Niyogi, and J. Cao, "Real-time crowd monitoring using seamless indoor-outdoor localization," *IEEE Transactions on Mobile Computing*, vol. 19, no. 3, pp. 664–679, 2020.

[4] D. Vasisht, S. Kumar, and D. Katabi, "Decimeter-level localization with a single wifi access point," in *Proceedings of the 13th Usenix Conference on Networked Systems Design and Implementation*, ser. NSDI'16. USA: USENIX Association, 2016, p. 165–178.

[5] B. Großiwindhager, M. Stocker, M. Rath, C. A. Boano, and K. Römer, "Snaploc: An ultra-fast uwb-based indoor localization system for an unlimited number of tags," in *2019 18th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, 2019, pp. 61–72.

[6] P. Kriz, F. Maly, and T. Kozel, "Improving indoor localization using bluetooth low energy beacons," *Mobile Information Systems*, vol. 2016, 2016.

[7] Q. Niu, T. He, N. Liu, S. He, X. Luo, and F. Zhou, "MAIL: Multi-scale attention-guided indoor localization using geomagnetic sequences," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 2, pp. 1–23, 2020.

[8] A. Kendall, M. Grimes, and R. Cipolla, "Posenet: A convolutional network for real-time 6-dof camera relocalization," in *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 2938–2946.

[9] S. He and K. G. Shin, "Geomagnetism for smartphone-based indoor localization: Challenges, advances, and comparisons," *ACM Computing Surveys*, vol. 50, no. 6, 2017.

[10] J. Chung, M. Donahoe, C. Schmandt, I. J. Kim, P. Razavai, and M. Wiseman, "Indoor location sensing using geomagnetism," *MobiSys'11 - Compilation Proceedings of the 9th International Conference on Mobile Systems, Applications and Services and Co-located Workshops*, pp. 141–154, 2011.

[11] K. P. Subbu, B. Gozick, and R. Dantu, "LocateMe: Magnetic-fields-based indoor localization using smartphones," *ACM Transactions on Intelligent Systems and Technology*, vol. 4, no. 4, 2013.

[12] X. Wang, Z. Yu, and S. Mao, "Indoor Localization Using Smartphone Magnetic and Light Sensors: a Deep LSTM Approach," *Mobile Networks and Applications*, vol. 25, no. 2, pp. 819–832, 2020.

[13] H. Wu, S. He, and S. H. Chan, "A Graphical Model Approach for Efficient Geomagnetism-Pedometer Indoor Localization," *Proceedings - 14th IEEE International Conference on Mobile Ad Hoc and Sensor Systems, MASS 2017*, pp. 371–379, 2017.

[14] Q. Niu, N. Liu, J. Huang, Y. Luo, S. He, T. He, S.-H. G. Chan, and X. Luo, "DeepNavi: A Deep Signal-Fusion Framework for Accurate and Applicable Indoor Navigation," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 3, pp. 1–24, 2019.

[15] N. Liu, T. He, S. He, and Q. Niu, "Indoor localization with adaptive signal sequence representations," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 11, pp. 11 678–11 694, 2021.

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on September 14,2022 at 11:18:13 UTC from IEEE Xplore. Restrictions apply.

[16] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, ser. NIPS'17. Red Hook, NY, USA: Curran Associates Inc., 2017, p. 6000–6010.

[17] S. Shao, A. Khreishah, and I. Khalil, "Retro: Retroreflector based visible light indoor localization for real-time tracking of iot devices," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, 2018, pp. 1025–1033.

[18] Q. Niu, Y. Nie, S. He, N. Liu, and X. Luo, "Recnet: A convolutional network for efficient radiomap reconstruction," in *2018 IEEE International Conference on Communications (ICC)*, 2018, pp. 1–7.

[19] X. Huang, S. Guo, Y. Wu, and Y. Yang, "A fine-grained indoor fingerprinting localization based on magnetic field strength and channel state information," *Pervasive and Mobile Computing*, vol. 41, pp. 150–165, 2017.

[20] L. Gong, Y. Zhao, C. Xiang, Z. Li, C. Qian, and P. Yang, "Robust light-weight magnetic-based door event detection with smartphones," *IEEE Transactions on Mobile Computing*, vol. 18, no. 11, pp. 2631–2646, 2019.

[21] Y. Zheng, G. Shen, L. Li, C. Zhao, M. Li, and F. Zhao, "Travi-Navi: Self-Deployable Indoor Navigation System," *IEEE/ACM Transactions on Networking*, vol. 25, no. 5, pp. 2655–2669, 2017.

[22] Y. Shu, C. Bo, G. Shen, C. Zhao, L. Li, and F. Zhao, "Magicol: Indoor Localization Using Pervasive Magnetic Field and Opportunistic WiFi Sensing," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 7, pp. 1443–1457, 2015.

[23] H. J. Bae and L. Choi, "Large-scale indoor positioning using geomagnetic field with deep neural networks," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, 2019, pp. 1–6.

[24] H. J. Jang, J. M. Shin, and L. Choi, "Geomagnetic field based indoor localization using recurrent neural networks," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 2017, pp. 1–6.

[25] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: pre-training of deep bidirectional transformers for language understanding," *CoRR*, vol. abs/1810.04805, 2018. [Online]. Available: http://arxiv.org/abs/1810.04805

[26] G. Zerveas, S. Jayaraman, D. Patel, A. Bhamidipaty, and C. Eickhoff, *A Transformer-Based Framework for Multivariate Time Series Representation Learning*. New York, NY, USA: Association for Computing Machinery, 2021, p. 2114–2124. [Online]. Available: https://doi.org/10.1145/3447548.3467401

[27] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

Authorized licensed use limited to: SUN YAT-SEN UNIVERSITY. Downloaded on September 14,2022 at 11:18:13 UTC from IEEE Xplore. Restrictions apply.