Static vs. Dynamic Databases for Indoor Localization based on Wi-Fi Fingerprinting: A Discussion from a Data Perspective

Zhe Tang*†, Ruocheng Gu*, Sihao Li*†, Kyeong Soo Kim*, and Jeremy S. Smith†
*School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou 215123, China.
Email: {Zhe.Tang15, Ruocheng.Gu22, Sihao.Li19}@student.xjtlu.edu.cn, Kyeongsoo.Kim@xjtlu.edu.cn
†Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, L69 3GJ, U.K.
Email: {Zhe.Tang, Sihao.Li, J.S.Smith}@liverpool.ac.uk

Abstract—**Wi-Fi fingerprinting has emerged as the most popular approach to indoor localization because it does not require deployment of new infrastructure or the modification of existing systems but exploits Wi-Fi networks already deployed in most indoor environments. The use of machine learning algorithms, including deep neural networks (DNNs), has greatly improved the localization performance of Wi-Fi fingerprinting, but its success heavily depends on the availability of fingerprint databases composed of a large number of the received signal strength indicators (RSSIs) measured at reference points, the medium access control addresses of access points, and the other available measurement information. However, most fingerprint databases do not reflect well the time varying nature of electromagnetic interferences in the more complicated modern indoor environment due to the increase in Wi-Fi and Bluetooth equipment. This could result in significant changes in statistical characteristics of training/validation and testing datasets, which are often constructed at different times, and even the characteristics of the testing datasets could be different from those of the data submitted by users during the operation of localization systems after their deployment. In this paper, we consider the implications of time-varying Wi-Fi fingerprints on indoor localization from a data-centric point of view and discuss the differences between static and dynamic databases. As a case study, we have constructed a dynamic database covering three floors of the International Research building on the south campus of Xi'an Jiaotong-Liverpool University (XJTLU) based on RSSI measurements, over 44 days, and investigated the differences between static and dynamic databases in terms of statistical characteristics and localization performance. The analyses based on variance calculations and Isolation Forest show the temporal shifts in Wi-Fi RSSIs, which result in a noticeable trend of the increase in the localization error of a Gaussian process regression model with the maximum error of 6.65 m after 14 days of training without model adjustments. The results of the case study with the XJTLU dynamic database clearly demonstrate the limitations of static databases and the importance of the creation and adoption of dynamic databases for future indoor localization research and real-world deployment.**

Index Terms—**Indoor localization, Wi-Fi fingerprinting, database construction, dynamic database, static database.**

I. Introduction

Wi-Fi fingerprinting has become the most popular technology for indoor localization, which can leverage existing Wi-Fi network infrastructure to provide reliable indoor localization services. The use of machine learning algorithms—especially deep neural networks (DNNs)—has brought substantial improvement in the performance and scalability of indoor localization based on Wi-Fi fingerprinting [1], [2].

Complicated and time-varying indoor electromagnetic interferences, however, pose a serious challenge to the robustness of indoor localization models. As indoor electromagnetic interferences are highly time-varying [3], it is desirable to develop more robust indoor localization models based on databases that can take into account the time variability of Wi-Fi fingerprints over days, weeks, or even longer. Nevertheless, due to the higher labor cost of constructing such Wi-Fi fingerprint databases, researchers often end up with validating their algorithms based on closed, private databases typically covering a single floor with a smaller number of access points (APs) and reference points (RP) and neglecting temporal signal fluctuations.

A better alternative is to use well-known, publicly-available databases like those summarized in Table I, which enables fair comparison among indoor localization algorithms. In Table I, we classify the databases into two groups, i.e., dynamic and static databases. Dynamic databases are defined by their incorporation of temporal variations in received signal strengths (RSSs) or received signal strength indicators (RSSIs)[1] over a long period of time. Such databases provide the periodic measurements of RSSs/RSSIs over a predefined set of RPs to ascertain the continual relevance and timeliness of the recorded data. Static databases, on the other hand, are characterized by the absence of such periodic measurements. In this regard, databases with training and test datasets measured at different times (e.g., UJIIndoorLoc [4]) are considered as static databases. Note that the existence of auxiliary information on APs, such as channel specifications, are not considered in this classification.

In this paper, we consider the implications of time-varying Wi-Fi fingerprints on indoor localization from a data-centric point of view and discuss the limitations of static databases for

[1] As a relative indicator, RSSI has no unit unlike RSS whose typical unit is dBm.

TABLE I
PUBLICLY-AVAILABLE WI-FI FINGERPRINT DATABASES.

Database	No. of APs	Covered Area [m^2]	Category	Frequency Band	Year
UJIIndoorLoc [4]	520	108703	Static	2.4 GHz	2014
UJI LIB DB [5]	448	N/A	Dynamic	2.4 GHz	2018
Tampere University [6]	992	22570	Dynamic	2.4/5 GHz	2017
WI-FI RSSI Indoor Localization [7]	6	384	Static	2.4/5 GHz	2019
WI-FI Fingerprinting Radio Map Database [8]	695	717	Static	N/A	2023
Hybrid Dataset [9]	17	386	Static	2.4 GHz	2022
MTLoc [10]	3365	8350	Dynamic	2.4/5 GHz	2023

indoor localization based on a careful review of the existing publicly-available databases. We also present the results of systematic investigation of a dynamic database covering three floors of the International Research (IR) building on the south campus of Xi'an Jiaotong-Liverpool University (XJTLU) composed of RSSI measurements over 44 days. Our work, in this paper, highlights the importance of the creation and adoption of dynamic databases for indoor localization and provides researchers valuable guidelines for the construction of dynamic databases.

II. Issues in the Construction of Static Databases

In practice, static Wi-Fi fingerprint databases are constructed in such a way that the RSSI measurement times of their training, validation, and testing datasets are different from one another. The different RSSI measurement times may result in shifts in the statistical characteristics of their RSSI fingerprints due to changes in APs, time-varying wireless channels, and imperfect measurement practice, which could worsen the actual localization performance with the testing dataset of a model trained and fine-tuned with the training and validation datasets.

A. Changes in APs

The major causes for the changes in APs are as follows:

- **Mobile hotspots**: Nowadays, many mobile devices can work as Wi-Fi hotspots, which causes significant issues in Wi-Fi fingerprinting based on static databases. Due to their sporadic operations and moving with users, the detection of mobile hotspots highly depends on measurement time and location. Moreover, some hotspots can employ ephemeral MAC addresses upon each initiation, so the resulting RSSIs could be regarded as those from multiple, different APs. Proper handling of these mobile hotspots (e.g., filtering them out in the datasets), however, is quite challenging with conventional static databases.
- **Addition and failure of fixed APs**: As typical APs are deployed and installed in fixed and often difficult-to-access locations of a building (e.g., ceiling), the measurement of their RSSIs is confined to the neighborhood of their deployment. Still, APs could be replaced by new ones due to device malfunction or as part of network upgrade. Also, new APs could be deployed in addition to the existing ones. Unlike mobile hotspots, however, the changes in fixed APs are rare and hardly captured in a short time frame.
- **Network Maintenance**: In large-scale building complexes like shopping malls and office buildings, network maintenance may be done on a periodical basis to check and improve the functionality of the whole network infrastructure. During the maintenance, a network operator can change the attributes (e.g., service set identifiers (SSIDs) and channels) and the operation mode (e.g., active, standby, and sleep) of managed APs, which, again, would result in differences in the statistical characteristics of the RSSI fingerprints of the datasets.

B. Time-Varying Wireless Channels

In addition to the changes in APs, time-varying wireless channels affect RSSIs as well, which could result from the following phenomena:

- **Multipath Propagation**: It is likely that in an indoor environment, radio signals from an AP reaches the receiving antenna of a user device through multiple paths due to reflections on the surfaces of objects like furniture, walls, and ceilings. This multipath propagation causes multipath interference, which is a highly time-varying phenomenon and thereby results in time-varying wireless channels.
- **Dynamic Disturbance**: Another major cause of time-varying wireless channels is dynamic disturbance from devices like Bluetooth devices, microwave ovens, and wireless microphones operating in the same frequency bands as Wi-Fi APs, which negatively affects Wi-Fi RSSIs at receivers through co-channel interference.
- **Environmental Changes**: Environmental changes like those in atmospheric temperature and humidity result in the dispersion, diffraction, and absorption of electromagnetic signals, while lightning could cause electromagnetic pulse disruptions. Their combined effects make wireless channels time-varying, too.

C. Imperfect Measurement Practice

The measurement practice, too, could adversely affect the consistency in the statistical characteristics of RSSI fingerprints unless carefully planned and applied.

- **Measurement Devices**: The number of 5 GHz-enabled Wi-Fi devices has been continuously growing. For dual-band APs, a singular MAC address is used for both 2.4 GHz and 5 GHz bands. Most of the fingerprint databases constructed before the widespread adoption of 5 GHz devices, by the way, do not provide channel information. Also, the computation of RSSI values is not consistent among equipment manufacturers, most of which do not provide explicit information on it. Additionally, the computation of RSSI values on the same device could be different depending on the versions of firmware and operating systems.

- **Measurement RPs**: When the same RPs (e.g., those based on a fixed grid on a floor) are used for the construction of the datasets over a period, measuring RSSIs at the same RPs at different times could also make the acquired RSSIs time-varying (e.g., due to different poses and directions), which further exacerbates the issues of static databases already discussed.

III. A Case Study: XJTLU Dynamic Database

Having discussed the issues in the construction of conventional static fingerprint databases in Section II, here we demonstrate how a dynamic database can be utilized to systematically investigate the temporal aspects of Wi-Fi RSSI fingerprints and their impact on indoor localization performance through a case study based on the XJTLU dynamic database.

A. Experimental Setup

For this case study, we constructed a new dynamic Wi-Fi RSSI fingerprint database covering three floors of the IR building on the south campus of XJTLU.

Building a dynamic database requires repeated access to the same RPs over a long period of time (e.g., one month), so we adopted a hybrid measurement scheme; users carrying laptops and Android smartphones visited the assigned RPs and measured RSSIs on a daily basis, while Raspberry Pi Pico Ws mounted on the corridor walls measured RSSIs automatically every hour. Although the Raspberry Pi Pico W supports only the 2.4 GHz-band Wi-Fi, we selected it as an anchor device due to the right balance between power consumption and performance. Table II provides an overview of the XJTLU single-building multi-floor dynamic fingerprint database.

The distribution of RPs is shown in Fig 1, where the red markers indicate the RPs for Raspberry Pi Pico Ws. The spacing between RPs, which is about 3 m, is not strict because we put each RP in close proximity to a landmark of the building to ease the task of repetitive RSSI measurements without compromising the normal use of the building. The coordinates of an RP are relative to the reference point of RP No. 0 and given in meters. The number of RPs per floor is also summarized in Table II. Due to the differences in the floor structures, the numbers of APs for the three floors are different from one another. The average width of the building's corridors is around 1.3 m, and their walls are made of glass. The walls between the rooms, on the other hand, are of brick construction. In the centers of the 6th and the 7th floor, there is a corkscrew staircase connecting the two, while the center of the 8th floor is an open plan space.

Fig. 1. RP distribution on the 7th floor of the IR building; the RPs with Raspberry Pi Pico W are marked in red.

B. Statistical Characteristics

We undertook statistical analyses of the time slices of XJTLU dynamic database based on a machine learning algorithm called Isolation Forest [11] as well as conventional techniques to investigate temporal aspects of RSSI fingerprints.

1) RSSI Variation over Time: Fig 2 shows the time variation of the RSSIs from AP No. 79 measured at RPs No. 3 and 14 on the 7th floor of the IR building. This combination of an AP and RPs is taken as an example because the range of the RSSI values for this case—i.e., $[-110, -25]$[2]—is the largest.

Note that, though RPs No. 3 and 14 are located far from each other and thereby have different wireless channel characteristics, the two time series of RSSIs show similar patterns and that their statistical variances of 324.89 and 324.91 at RPs No. 3 and 14 are also close to each other. These imply that there is a potential correlation among RSSIs from the same AP over time, which could be exploited to improve the performance of indoor localization based on dynamic data bases.

2) RSSI Anomaly Detection based on Isolation Forest: We carried out RSSI time series anomaly detection based on the Isolation Forest algorithm [11] to demonstrate that, because RSSIs could be significantly abnormal at certain points of time, the mean RSSI value cannot properly represent the whole RSSI time series. The core idea is that normal samples can be easily separated by the decision tree of the Isolation Forest, while abnormal samples are more difficult to separate, and thus, their path lengths and isolation degrees become larger. In the Isolation Forest, an anomaly score $s(x,n)\in[0,1]$ indicates the degree of anomaly in the sample, which is defined as follows:

$$s(x,n) = 2^{-\frac{\mathrm{E}[h(x)]}{c(n)}}, \tag{1}$$

where $h(x)$ denotes the length of the path from the root node to the leaf node containing a sample x, and $c(n)$ is a function of the number of samples n; if s is close to 1, then the sample x is very likely to be an outlier [11].

[2]As in [2], we use -110 as an RSSI value to indicate no detection of an AP.

TABLE II
OVERVIEW OF THE XJTLU DYNAMIC DATABASE.

Building	Floor	RP Numbers	Devices		Period /day	AP	Samples	Covered Area $[\mathrm{m}^2]$
IR	6	0–27	6 Raspberry Pi Pico W	1 laptop/1 smartphone*	44	446	511237	1200
	7	0–34	6 Raspberry Pi Pico W					
	8	0–25, 35–46						

* The laptop and the smartphone are MacBook Pro and Xiaomi 13, respectively, and were used for all the floors.

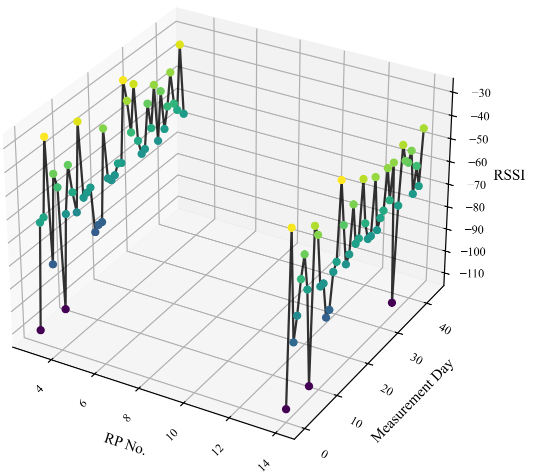

Fig. 2. Time variation of the RSSIs from AP No. 79 measured at RPs No. 3 and 14 on the 7th floor of the IR building.

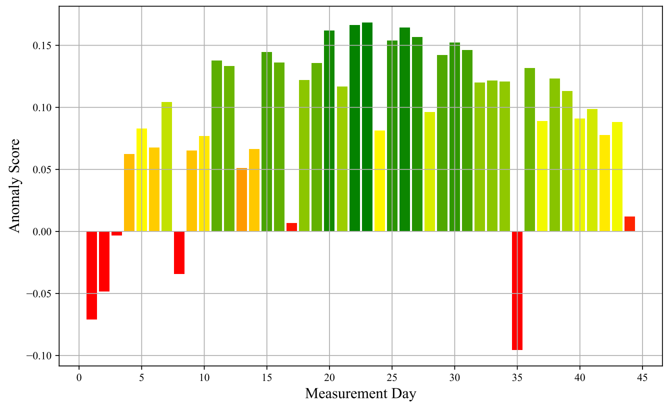

Fig. 3. Isolation Forest anomaly scores for the RSSIs from AP No. 79 measured at RP No. 14 on the 7th floor of the IR building.

TABLE III
HYPERPARAMETERS AND THEIR VALUES FOR ISOLATION FOREST.

Parameter Settings	Value	Description
n_estimators	100	number of random trees
contamination	0.10	percentage of anomalous data
max_samples	auto	number of samples to construct the subtree
max_features	1.0	constructing number of features for each subtree
random_state	42	random seed

We used the implementation of the Isolation Forest algorithm provided by the scikit-learn Python package with the parameter settings summarized in Table III. Note that, unlike the anomaly score defined in (1), the anomaly score returned from the scikit-learn implementation provides a negative value for outliers [12]. Fig 3 shows the anomaly scores for the RSSIs from AP No. 79 measured at RP No. 14. The higher, positive anomaly scores (i.e., the green bars) indicate that the RSSIs on the corresponding measurement days are likely to be normal given the whole RSSI time series, while the lower, negative anomaly score (i.e., the red bars) indicates that the RSSIs on the corresponding measurement days are likely to be anomalous (e.g., strong interference during the measurements).

The large fluctuation in RSSIs from a single AP measured at a single RP may not be directly related with the localization performance. Considering the possible correlation of the time variations of RSSIs measured at different RPs shown in Fig. 2, however, the impact of those fluctuations on the localization performance can't be simply ignored. Likewise, static databases based on those anomalous RSSI measurements could negatively affect the actual performance of indoor localization models trained based on them.

C. *Indoor Localization Performance*

We investigate the impact of temporal aspects of Wi-Fi RSSI fingerprints on indoor localization performance using DNN and Gaussian Process (GP) models, which are used for the classification of location labels and the regression of location coordinates, respectively. As we mainly focus on the *changes* of indoor localization performance over time with dynamic fingerprint databases, we selected simpler models.

Table IV summarized the network structure and the hyper-parameter values of the DNN model. With the data measured on the 7th floor of the IR building in June and July as a training set for a total of 24 days and those in August as a test set for a total of 20 days, we evaluated the classification performance of the DNN model trained with the training set against each daily time slice of the testing set without retraining. The localization classification accuracy over time is shown in Fig 4, where there is a significant drop in the classification accuracy on the 11th and 12th test days.

We also implemented a GP model based on GPy [13] and evaluated its regression performance with the same experimental settings as the DNN model; the details of the GP model are given in Table V. The minimum, the maximum, and the average of the localization errors for a single measurement day are $4.67\,\mathrm{m}$, $6.65\,\mathrm{m}$, and $5.65\,\mathrm{m}$, respectively. The regression

TABLE IV
HYPERPARAMETERS AND THEIR VALUES OF THE DNN MODEL.

Category	Details
General Settings	
Device	i9 12900k & RTX 4090
Random Seed	12345
Model Settings	
Epochs	SAE:* 30, CLS:† 30
Number of APs	465
Number of Classes	30 (RPs)
Hidden Units	512
SAE Output Units	64
Batch Size	20
SAE Model	
Structure	1. Linear layer: 465 ->116 (FEA_DIM/4), Activation: ELU‡ 2. Linear layer: 116 ->64 (SAE_DIM) 3. Decoder: 64 ->116, Activation: ELU 4. Decoder: 116 ->465
CLS Model	
Structure	1. Encoder (from SAE) 2. Linear layer: 64 ->512 (HID_DIM), Activation: ELU 3. Linear layer: 512 ->512 (HID_DIM), Activation: ELU 4. Linear layer: 512 ->512 (HID_DIM), Activation: ELU 5. Output layer: 512 ->30 (CLS_DIM), Batch normalization
SAE Training	
Loss Function	Mean Square Error (MSE)
Optimizer	Adam with learning rate = 1×10^{-4}
CLS Training	
Loss Function	Cross Entropy Loss
Optimizer	Adam with learning rate = 1×10^{-3}

* Stacked autoencoder.
† Classifier.
‡ Exponential linear unit.

TABLE V
HYPERPARAMETERS AND THEIR VALUES OF THE GP MODEL.

Category	Details
Kernel Function	Ornstein-Uhlenbeck (OU)
Variance	1
Length scale	100
Likelihood	Gaussian

error over time is shown in Fig 5, where we can observe a rough cyclic behavior with a long-term trend that the regression error goes down in the middle of the period of about 5 days and the average over that period slowly increases over time; Table VI confirms this group behavior that the average localization error of a group of 5 days increases over time, i.e., up to $0.62\,\mathrm{m}$ from the group of 1–5 days to that of 16–20 days. These results are consistent with the results of the

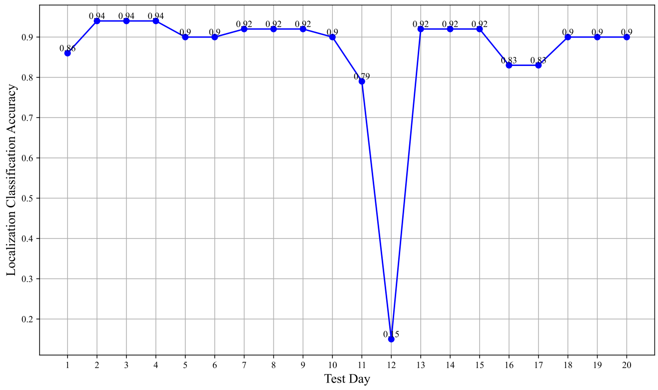

Fig. 4. DNN localization accuracy over time.

TABLE VI
AVERAGE LOCALIZATION ERRORS FOR A GROUP OF 5 TEST DAYS.

Test Day	All 20 days	1–5	6–10	11–15	16–20
Average Error [m]	5.65	5.37	5.36	5.86	5.99

analyses based on variance calculations and Isolation Forest in Section III-B2 that there are temporal shifts in Wi-Fi RSSIs.

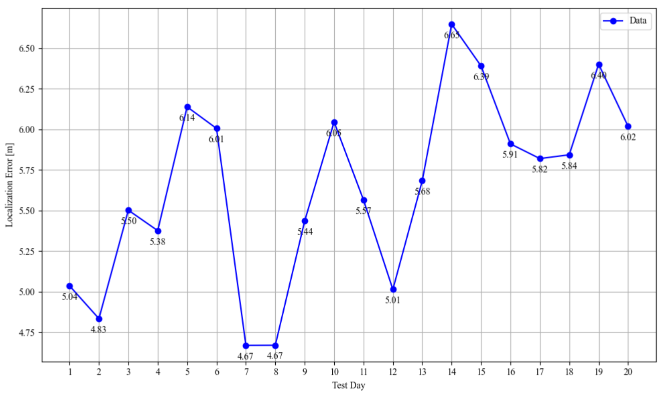

Fig. 5. GP localization error over time, where the red dashed line shows the results of curve fitting based on a 6th-order polynomial.

As shown in Fig 5, our case study with the XJTLU dynamic database clearly demonstrates that the localization error increases over time as the time gap between the measurement times of training and testing datasets widens, which implies that the model will eventually lose its location service ability without retraining with the updated training dataset.

IV. CHALLENGES AND OPPORTUNITIES

The discussions in Section II and the results from the case study in Section III provide compelling reasons for the employment of dynamic databases for indoor localization. However, there are several challenges specific to dynamic databases, which are to be addressed before their adoption.

One of the major challenges in constructing dynamic databases is to ensure the access to the same RPs accurately for repetitive RSSI measurements over a long period of time. Most static databases rely on global coordinates systems like the universal transverse mercator (UTM) or the world geodetic system 1984 (WGS 84) for specifying RP locations. The advantage of global coordinate systems is their compatibility and ease of conversion between them. Note that global coordinate systems require GPS devices, detailed floor maps of buildings, and special mapping software. The use of global coordinate systems would be even more challenging for large-scale indoor localization systems for building complexes such as hospitals, shopping malls, and transport hubs. In this regard, the use of a local coordinate system, together with distinct landmarks and anchor devices, could be a viable option reducing resource overhead and ensuring the accurate access to the same RPs repeatedly during the construction of dynamic databases.

Another major challenge is the higher labor costs incurred from repetitive data measurements over a long period of time. The use of hybrid networks with anchor devices reduce the labor cost in data measurements because anchor devices can automatically measure fingerprints and thereby facilitate the measurement cycle adjustment. Considering that one of the advantages of Wi-Fi fingerprinting is no requirement for the deployment of new infrastructure or special new user devices [14], we recommend to deploy only a limited number of anchor devices in certain RPs where Wi-Fi signals exhibit substantial fluctuations. As for anchor devices, we recommend those with a lightweight and straightforward architecture while capable of the reliable recording of signal variations and providing sufficient storage space to avoid frequent data transmissions from them.

V. Concluding Remarks

In this paper, we have investigated the implications of the time-varying nature of Wi-Fi RSSI fingerprints on indoor localization through the case study with the XJTLU dynamic database. For this case study, we have constructed the XJTLU dynamic database covering three floors of the IR building on the south campus of XJTLU, whose Wi-Fi fingerprint data were measured on a daily basis over a period of 44 days.

The experimental results, with the XJTLU dynamic database, show that the indoor localization performance of DNN and GP models become worse as the time difference between the training and the estimation increases; specifically, there is a noticeable trend of the increase in the localization error of a GP regression model with the maximum error of 6.65 m after 14 days of training without model adjustments. In fact, the analyses based on variance calculations and Isolation Forest indicate the temporal shifts in Wi-Fi RSSIs. In this regard, our results of the case study with the XJTLU dynamic database clearly demonstrate the limitations of static databases and the importance of the creation and adoption of dynamic databases for future indoor localization research and real-world deployments.

Based on our hands-on experience obtained during the construction of the XJTLU dynamic database, we have also proposed guidelines for the cost-effective construction of dynamic databases, where we encourage the adoption of local coordinates and the integration of hybrid network data, the latter of which enables researchers to calibrate their algorithms leveraging dynamic databases and to handle issues resulting from the differences between simple laboratory and more complicated deployment environments.

Acknowledgment

This work was supported in part by the Postgraduate Research Scholarship (under Grant PGRS1912001), the Key Program Special Fund (under Grant KSF-E-25), and the Research Enhancement Fund (under Grant REF-19-01-03) of Xi'an Jiaotong-Liverpool University.

References

[1] T. Chuenurajit, S. Phimmasean, and P. Cherntanomwong, "Robustness of 3D indoor localization based on fingerprint technique in wireless sensor networks," in *2013 10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, 2013, pp. 1–6.

[2] K. S. Kim, S. Lee, and K. Huang, "A scalable deep neural network architecture for multi-building and multi-floor indoor localization based on Wi-Fi fingerprinting," *Big Data Analytics*, vol. 3, no. 1, Apr 2018. [Online]. Available: http://dx.doi.org/10.1186/s41044-018-0031-2

[3] X. Gan, B. Yu, L. Huang, and Y. Li, "Deep learning for weights training and indoor positioning using multi-sensor fingerprint," in *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2017, pp. 1–7.

[4] J. Torres-Sospedra, R. Montoliu, A. Martínez-Usó, J. P. Avariento, T. J. Arnau, M. Benedito-Bordonau, and J. Huerta, "UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems," in *Proc. 2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2014, pp. 261–270.

[5] G. Mendoza-Silva, P. Richter, J. Torres-Sospedra, E. Lohan, and J. Huerta, "Long-Term WiFi Fingerprinting Dataset for Research on Robust Indoor Positioning," *Data*, vol. 3, no. 1, p. 3, Jan. 2018.

[6] E. S. Lohan, J. Torres-Sospedra, H. Leppäkoski, P. Richter, Z. Peng, and J. Huerta, "Wi-Fi crowdsourced fingerprinting dataset for indoor positioning," *Data*, vol. 2, no. 4, 2017. [Online]. Available: https://www.mdpi.com/2306-5729/2/4/32

[7] M. T. Hoang, X. Dong, T. Lu, B. Yuen, and R. Westendorp, "WiFi RSSI indoor localization," 2019. [Online]. Available: https://dx.doi.org/10.21227/1yd5-rn96

[8] A. M. Rajab and B. Wang, "Wifi fingerprinting radio map database for indoor localization," 2023. [Online]. Available: https://dx.doi.org/10.21227/sdpr-fh91

[9] A. N. Nor Hisham, Y. H. Ng, C. K. Tan, and D. Chieng, "Hybrid wi-fi and ble fingerprinting dataset for multi-floor indoor environments with different layouts," *Data*, vol. 7, no. 11, 2022. [Online]. Available: https://www.mdpi.com/2306-5729/7/11/156

[10] J. Wang, Z. Zhao, M. Ou, J. Cui, and B. Wu, "Automatic update for wi-fi fingerprinting indoor localization via multi-target domain adaptation," vol. 7, no. 2, jun 2023. [Online]. Available: https://doi.org/10.1145/3596239

[11] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation Forest," in *2008 Eighth IEEE International Conference on Data Mining*. IEEE, pp. 413–422. [Online]. Available: http://ieeexplore.ieee.org/document/4781136/

[12] "sklearn.ensemble.IsolationForest - Isolation Forest framework in Python." [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.IsolationForest.html

[13] "GPy - A Gaussian Process (gp) framework in Python." [Online]. Available: https://gpy.readthedocs.io/en/deploy/

[14] F. Wang, J. Feng, Y. Zhao, X. Zhang, S. Zhang, and J. Han, "Joint activity recognition and indoor localization with WiFi fingerprints," *IEEE Access*, vol. 7, pp. 80 058–80 068, 2019.