

BmmW: A DNN-based Joint BLE and mmWave Radar System for Accurate 3D Localization

Peizheng Li^{**}, Jagdeep Singh^{**}, Han Cui[†], Carlo Alberto Boano[‡]

^{*} Bristol Research & Innovation Laboratory, Toshiba Europe Ltd., UK; [†] University of Bristol, UK;

[‡] Institute of Technical Informatics, Graz University of Technology, Graz, Austria.

Email: peizheng.li, jagdeep.singh@toshiba-bril.com, hancui9140@gmail.com, cboano@tugraz.at

Abstract—Bluetooth Low Energy (BLE) has emerged as one of the reference technologies for the development of indoor localization systems, due to its increasing ubiquity, low-cost hardware, and to the introduction of direction-finding enhancements improving its ranging performance. However, the intrinsic narrowband nature of BLE makes this technology susceptible to multipath and channel interference. As a result, it is still challenging to achieve decimetre-level localization accuracy, which is necessary when developing location-based services for smart factories and workspaces. To address this challenge, we present BmmW, an indoor localization system that augments the ranging estimates obtained with BLE 5.1’s constant tone extension feature with mmWave radar measurements to provide real-time 3D localization of a mobile tag with decimetre-level accuracy. Specifically, BmmW embeds a deep neural network (DNN) that is jointly trained with both BLE and mmWave measurements, practically leveraging the strengths of both technologies. In fact, mmWave radars can locate objects and people with decimetre-level accuracy, but their effectiveness in monitoring stationary targets and multiple objects is limited, and they also suffer from a fast signal attenuation limiting the usable range to a few metres. We evaluate BmmW’s performance experimentally, and show that its joint DNN training scheme allows to track mobile tags in real-time with a mean 3D localization accuracy of 10 cm when combining angle-of-arrival BLE measurements with mmWave radar data. We further evaluate a variant of BmmW, named BmmW-LITE, that is specifically designed for single-antenna BLE devices (i.e., that avoids the need of bulky and costly multi-antenna arrays). Our results show that BmmW-LITE achieves a mean 3D localization accuracy of 36 cm, thus enabling accurate tracking of objects in indoor environments despite the use of inexpensive single-antenna BLE devices.

Index Terms—Bluetooth 5.1, Angle of Arrival/Departure, mmWave, Heatmap, Feature Fusion, Deep Neural Networks.

I. INTRODUCTION

Accurate localization and tracking is one of the fundamental pillars of 6G networks. In the research realm “Beyond 5G”, the concept of Integrated Sensing and Communication (ISAC) has been widely discussed [1], with the aim of providing high-quality wireless connectivity and seamless location awareness in both indoor and outdoor environments.

The integration of these capabilities in off-the-shelf IoT devices can enable the design of location-aware applications that can be highly beneficial to our society at large. For example, the ability to accurately track people indoors allows the creation of targeted services improving the quality of life for elderly people or vulnerable groups [2]. Similarly, autonomous robot navigation and asset tracking in smart warehouses and workspaces are essential services to provide organizations with

a way to efficiently manage, monitor, and track their resources, towards higher productivity and lower operating costs [3].

Achieving a reliable and accurate tracking of moving targets in such autonomous settings is of utmost importance, especially where the safety of people/workers or the optimal utilization of moving targets are key considerations [3]. However, despite the wide range of optical and radio frequency (RF) technologies available, achieving decimetre-level accuracy in indoor settings with minimal costs is still a key challenge. On the one hand, existing communication technologies deployed in smart industries (e.g., Wi-Fi [4], Bluetooth [5], and RFID [6]) are typically narrowband and hence inherently susceptible to multipath fading, which limits the achievable ranging accuracy. On the other hand, retrofitting or replacing existing devices and installations with more accurate technologies may require significant investments and is labor-intensive – consider, for example, the installation of optical systems [7], [8] or of an anchor infrastructure of ultra-wideband (UWB) devices [9].

Combining the benefits of different technologies. In this work, we study how to increase the localization performance of BLE devices. We focus on BLE, as it is arguably the most commonly used IoT technology for indoor localization due to its wide availability and relatively cheap hardware [10]. Moreover, BLE is already ubiquitous in industrial settings, but its *narrowband* nature makes it hard to carry out decimetre-level localization due to strong multipath effects in indoor environments, which calls for novel solutions. In contrast to BLE devices, mmWave radar sensors have a bandwidth of a few gigahertz, which enables them to sense objects with a resolution of a few cm. Such inexpensive sensors ($\approx 10\$$), combined with Multiple Input Multiple Output (MIMO) techniques, can be used for the *passive* localization of objects [11]. They can hence be used *alongside* BLE devices to improve their location estimates without the need of additional infrastructure. However, mmWave radar sensors are known to suffer from a fast signal attenuation that limits their usable range, and are not too effective in the presence of stationary targets and multiple objects [12]. For this reason, we investigate an approach that can *effectively combine* the strength of BLE and mmWave radar measurements while overcoming their individual shortcomings. Specifically, we study the adoption of a DNN to fuse signal features from both BLE and mmWave sources. We do so, as machine learning has emerged in recent

^{*}Both authors contributed equally to this research.

years as a promising solution to tackle the multipath complexity and to model the mapping between radio signal features and target locations [13], [14]. Especially DNNs, with their powerful function approximation and design flexibility, have achieved remarkable success in various domains, including image processing [15], gaming [16], and language models [17].

Contributions. We present BmmW, an indoor localization system that leverages the strengths of BLE and mmWave radar technology to provide real-time 3D localization and tracking of a mobile tag with decimetre-level accuracy. BmmW builds upon BLE 5.1’s Angle-of-Arrival (AoA) direction-finding enhancement, which uses a multi-antenna array to measure the phase difference of the received signal at multiple antennas and translate it into angular information. Specifically, BmmW uses this angular information along with mmWave radar measurements as a feature for training a DNN model. To improve the effectiveness of the data fusion scheme and DNN model, we present a novel mmWave radar signal processing and corresponding signal heatmap generation method that converts irregular radar point clouds to regular probability distributions of target locations. We also design and implement a variant of BmmW, named BmmW-LITE, which uses raw IQ samples from a *single-antenna* BLE device along with mmWave radar measurements. We evaluate both BmmW and BmmW-LITE in real-world environments, demonstrating that mobile tags can be tracked in real-time with a mean 3D localization accuracy of 10 cm and 36 cm, respectively. Such accuracy is 80% superior to that of classical BLE localization methods [5] – albeit at a price of an increased computational complexity – and is especially remarkable for BmmW-LITE, as the latter does not require the use of bulky and costly multi-antenna arrays.

Paper outline. The remainder of this paper is structured as follows: Sec. II provides an overview of the employed technologies and of related work. In Sec. III, we describe the inner details and implementation of BmmW. The experimental extraction of DNN model features is presented in Sec. IV, and the corresponding localization evaluation and results obtained with a mobile target are discussed in Sec. V. Finally, the paper concludes with a summary of our findings in Sec. VI.

II. BACKGROUND AND RELATED WORK

In this section, we first provide details of BLE and mmWave technologies, as well as of NN-based data fusion methods. We then describe the most related works to BmmW, from those leveraging multiple technologies for indoor localization, to those performing multi-sensor fusion to increase its accuracy.

A. Bluetooth Low Energy

BLE is widely used in smart industries and workspaces due to the ubiquity and low cost of its devices. However, the limited bandwidth, susceptibility to multipath, and the use of the crowded 2.4 GHz band make it challenging to attain a high localization accuracy. In fact, several works have shown that received signal strength information (RSSI) alone can hardly achieve sub-metre accuracy in real-world settings [18]–[23]. To improve accuracy, direction-finding enhancements

were introduced in the BLE 5.1 specifications, namely support for Angle of Arrival (AoA) and Angle of Departure (AoD). Several researchers have conducted empirical studies to evaluate the accuracy of BLE’s AoA technique [24], [25], but mainly for static targets and using only a limited number of tested locations. Recently, Pau et al. [26] have used a hybrid solution based on both RSSI and AoA information that results in an average distance error of 0.7 m. However, the authors tested only a few locations and did not describe in detail how these were chosen nor the impact of dense multipath on RSSI-based measurements. To improve the accuracy of the AoA technique, researchers have proposed algorithms based on non-linear recursive least square and unscented Kalman filters to reduce multipath and antenna switching errors [27]. Unfortunately, the choice of channel greatly impacts the direction-finding techniques [24], [25], and the angular error spreads more at lower frequencies [5]. An empirical study using software-defined radios demonstrated the effectiveness of the AoA technique in achieving sub-metre accuracy when tracking moving targets [5]; however, measurements were carried out in an outdoor environment with few multipath reflections, and information from several packets was averaged to obtain the reported accuracy. Therefore, despite recent advances, achieving sub-metre accurate localization for mobile targets using BLE remains an open challenge. Moreover, the use of the direction finding enhancements introduced in Bluetooth 5.1 requires costly and bulky (typical size exceeding 15×15 cm) antenna arrays that are hard to find on the market [28]. With BmmW, we train a neural network (NN) with both BLE and mmWave radar measurements in order to develop a decimetre-level localization system for mobile targets that is robust to dense multipath effects. Moreover, with BmmW-LITE, we also eliminate the need for multi-antenna design considerations.

B. mmWave Radar

mmWave technology commonly refers to the use of RF signals above 60 GHz. The use of sensors operating at these frequencies has shown great promise for high-precision indoor tracking applications. These sensors typically have a wide bandwidth of a few gigahertz and implement the Frequency Modulated Continuous Wave Radar (FMCW) approach, enabling them to sense objects with a high distance resolution of a few centimetres. When combined with MIMO antennas, whose size is typically very small (e.g., 2×2 mm [29]), mmWave radars can operate as 3D imaging sensors, accurately detecting the 3D coordinates of objects and generating point clouds that encode their spatial shape [12]. Many researchers have used commercial mmWave radars as a low-cost and easy-to-deploy solution, and have shown that they can locate people within a 20 cm error when the target is within the effective detection range of the radar. For example, authors in [30] and [31] use radars from Texas Instruments that have integrated transmitters and receivers on a single chip. Zhao et al. [30] use neural networks to process mmWave radar data and propose a human identification and tracking system with a 16 cm median error, but that can only detect one person at a time. Cui et al. [31]

have shown that a single radar can have a high false alarm rate: while this can be significantly improved by fusing information from multiple radars, a minimum distance (15 cm) between people is necessary to correctly discern multiple individuals. Wu et al. [32] have used separate mmWave transmitters and receivers, and designed a novel system that can locate multiple people simultaneously with a 10 cm error. However, they have also shown that the accuracy would decrease significantly to more than 30 cm when the person is more than 1.5 m away, as well as when there are more than two people present. Hence, while mmWave radars offer several advantages over other tracking technologies (such as not requiring tags or smart devices to be carried by subjects), the problem of failing to distinguish between multiple targets and being prone to clutters in the environment and occlusion on the line of sight are intrinsic weaknesses of this technology. Additionally, mmWave signals attenuate quickly through the air and rely on Doppler detection, which can reduce their range of view and limit their effectiveness in monitoring stationary targets [12]. To address these challenges and enhance localization performance, BmmW proposes a joint tracking system that combines information from BLE and mmWave technologies: by leveraging the strengths of both technologies, BmmW improves the accuracy and reliability of indoor tracking in challenging environments.

C. NN-based Data Fusion

Data fusion is a process dealing with the correlation and integration of data and information from multiple sources in order to produce more consistent, accurate, and useful combined information than that provided by any individual data source. Data fusion can be found in many applications, such as health monitoring, video/audio processing, and communication. For instance, a sensor data fusion system is presented in [33] for healthcare applications. A multi-modal fusion module is proposed in [34] to fuse the audio stream, video stream, and speaker embedding stream to realise the speaker separation.

Algorithms for data fusion can be very diverse, e.g., weighted average, K-means, and Bayesian inference, just to name a few¹. Also NNs, by their design flexibility, have gained a lot of attention for data fusion. Michelsanti et al. [36] have summarised the main NN-based fusion techniques used in audio-visual systems (e.g., concatenation, addition, product). Based on similar NN design principles, these methods can be extended to the fusion of radio features for localization. This paper adopts the concatenation-based NN for radio feature fusion, wherein a two-head input NN is designed for BLE and mmWave feature processes, and then the processed features are concatenated together for the following joint training.

D. Exploiting Multiple Communication Technologies and Multi-Sensor Data Fusion for Indoor Localization

Several research works have proposed the combination of multiple communication technologies to improve the accuracy of indoor localization. For instance, Liu et al. [37] fuse Wi-Fi, inertial sensors, and BLE beacons for indoor localization. However, this method requires users to carry a smartphone with extra sensors, which may not be practical in some situations.

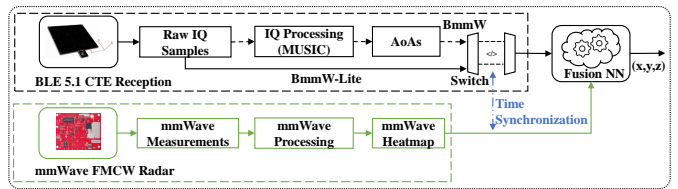


Figure 1: Overview of BmmW, an indoor localization system leveraging BLE 5.1 and mmWave measurements to jointly train a DNN and predict the 3D coordinates of a mobile tag.

Bala et al. [38] combine UWB and BLE signals to provide real-time location updates, but require installation of UWB and BLE devices throughout the indoor environment, which can be expensive and time-consuming. Jeong et al. [39] propose a machine learning-based fusion that requires a large amount of training data to predict the user’s location accurately. Istomin et al. [40] propose a dual-radio protocol to enable energy-efficient and accurate social contact detection, leveraging a narrowband radio (BLE) and an UWB radio. Zhang et al. [41] propose a system that fuses Wi-Fi and Bluetooth fingerprints in edge computing. Several other works [42] have also used multi-sensor fusion to increase the accuracy of indoor localization (e.g., IMU data, etc.). However, these approaches are highly dependent on the quality of the radio signal, which can be affected by factors such as signal interference and the number of access points in the environment. The main limitations of these approaches are related to the complexity and cost of the technology, the need for user participation, and environmental factors that affect the accuracy of the positioning system. In contrast to these approaches, BmmW does not require additional hardware on the target device and leverages angular information, which is less affected by multipath.

III. BMMW: DESIGN AND IMPLEMENTATION

This section describes the design and implementation of BmmW, further providing the technical details of its components, including BLE, mmWave, and fusion DNN model.

A. Overview of BmmW

The structure of the proposed joint localization framework is shown in Fig. 1. BmmW’s foundation lies in the reception of BLE 5.1’s constant tone extension (CTE) and mmWave FMCW radar measurements. The *reception* and *processing* of the BLE CTE packets are selective. In BmmW, the raw IQ samples are collected from BLE anchors with *multiple* antennas, and are processed by the Multiple Signal Classification (MUSIC) algorithm for AoA estimation, which serves as the default BLE feature for the proposed fusion model. Instead, BmmW-LITE accepts the raw IQ samples from a *single* antenna as the BLE feature, in order to save the feature process time and reduce the cost as well as the computational complexity. The switch in Fig. 1 represents the above BLE feature selection process. The techniques used to extract the relevant ‘features’ from BLE will be detailed in Sec. III-B. For the mmWave measurement, the critical step is the generation of the heatmap, which aims to overcome the irregularity of the radar point cloud: details

¹A comprehensive overview of data fusion algorithms is presented in [35].

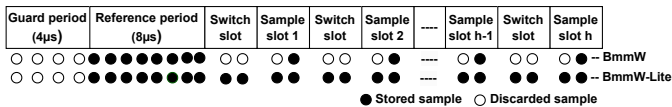


Figure 2: CTE packet & sampling structure for ‘multiple-antenna’ (BmmW) & ‘single-antenna’ (BmmW-LITE) system.

about this component will be discussed in Sec. III-C. Please note that, for accurate location prediction, the BLE features stream should be synchronized with the mmWave heatmap stream according to the recorded timestamps. Ultimately, the synchronized BLE/mmWave features are fed into the fusion DNN model for 3D location estimation. The architecture of such DNN model is detailed in Sec. III-D.

B. BLE Direction Finding

AoA and AoD are the direction-finding enhancements introduced in the BLE 5.1 standard [43]. The key concept behind these techniques involves measuring the phase difference of the received waveform across multiple antennas and determining the direction of the signal from the computed phase difference.

Constant tone extension. To perform these techniques, the Bluetooth SIG added a new field called *Constant Tone Extension (CTE)* at the end of a Bluetooth packet. The CTE field lasts for $16 \mu\text{s}$ to $160 \mu\text{s}$ and consists solely of binary ones that serve as a 250 kHz frequency offset to the unmodulated carrier. The purpose of the CTE is to provide a constant wavelength signal that can be used for IQ sampling and phase detection of the incoming RF signal. Unlike other parts of the Bluetooth packet, the CTE field is not subject to whitening processes nor included in the CRC [43] calculation. It contains a $4 \mu\text{s}$ guard band, $8 \mu\text{s}$ reference period and $148 \mu\text{s}$ band for IQ sampling. According to the standard, it is assumed that only one transmitter is active during the CTE phase difference measurement. Specifically, in the AoA technique, the receiver is equipped with multiple antennas that are controlled using an RF switch. By determining the phase difference observed at these multiple antennas, the receiver can determine the direction of the transmitter. Conversely, in the AoD technique, the transmitter is equipped with multiple antennas and transmits the signal over these antennas in a time-division manner. The receiver, which has a single antenna, estimates the direction of the transmitter based on the time-multiplexed signal. Additionally, given the known separation between the antennas, the AoA and AoD can be calculated using Eq. 1:

$$\begin{aligned} \theta_A &= \arccos((\phi\lambda)/(2\pi D)) \\ \theta_D &= \arcsin((\phi\lambda)/(2\pi D)) \end{aligned} \quad (1)$$

where λ is the wavelength, ϕ is the phase difference, and D is the distance between adjacent antennas (11.8 mm in BmmW) in an antenna array. However, to determine the true AoA, MUSIC is run on the obtained angles, a detailed process outlined in [44].

CTE sampling. The sampling band of $148 \mu\text{s}$ in the CTE field does not necessarily equate to collecting the maximum possible number of samples due to the duration of the switching slots, making it difficult for classical AoA/AoD techniques to utilize the entire band. While the BLE 5.1 standard defines the

sampling period, manufacturers have the flexibility to develop their own methods for efficient utilization. One such method, proposed by Silabs [44] and adopted by BmmW, allows to collect a maximum of 74 samples with a $1 \mu\text{s}$ switching slot, as illustrated in Fig. 2. However, this technique discards the remaining 74 samples during antenna switching slots, and the samples are reduced to half when the switching slot is $2 \mu\text{s}$. To optimally utilize the CTE sampling band, we propose BmmW-LITE: a *single-antenna system*, illustrated in Fig. 2, that eliminates the need for multiple antenna switching, hence allowing a single antenna to collect all IQ samples within the $148 \mu\text{s}$ CTE band. We utilize the collected raw IQ samples as a ‘feature’ to train the NN model. As the CTE is a constant frequency tone, the model can learn channel impairments. Additionally, these samples are jointly trained with the info from the mmWave technology, as described in Sec. III-D.

C. mmWave Radar Heatmap Construction

The principle of mmWave radars and the underlying FMCW approach have been documented in detail in the literature, e.g., in [12]. The radar sends mmWave signals to detect objects in the scene, receives the reflections, and uses a set of signal processing techniques (including a set of Fourier transform and peak detection) to produce a point cloud that encodes the 3D location and shape of the objects. However, the density and the accuracy of the point cloud are prone to noise and can have arbitrary population, making them unsuitable to be processed by a neural network directly. Additionally, commercial off-the-shelf mmWave radars are commonly designed for automotive driving applications and suffer from a lack of elevation resolution. For instance, the TI AWR series features eight virtual antennas along the azimuth direction, but only two along the elevation direction [45], rendering it unsuitable for numerous geometry-based solutions due to its unbalanced resolution. As a result, since the elevation information may prove less dependable, we propose relying solely on 2D information (i.e., azimuth and depth) from the radar, and converting it into heatmaps. These heatmaps encode the probability distribution of the person’s position in the scene, offering a feasible alternative. The heatmap approaches deliver low computational complexity and facilitate the NN’s ability to extract features by compressing the feature space.

The complete mmWave data processing chain involves three parts. The first part follows the standard FMCW approach as described in [12] and is completed by the radar on-chip processors. It processes the raw mmWave RF signal and generates a 3D point cloud representing the scene. Then, a clutter removal stage is introduced by recording the detected point cloud from the empty environment and subtracting it from the actual experimental data. Finally, a heatmap representing a probability distribution on the person’s location in the area is generated, where the probabilities are calculated based on the strength of the mmWave signal (i.e., the population of the point cloud in each unit region):

$$\text{Heatmap}(P) = \begin{cases} (\sum_{p \in P} G(p))/|P|, & \text{if } |P| \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

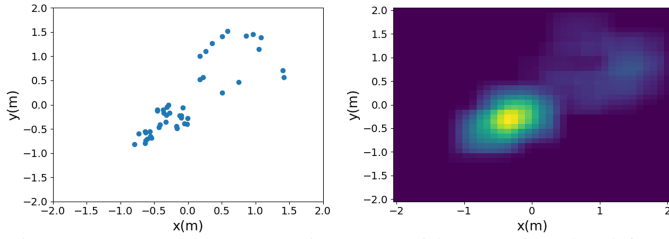


Figure 3: Heatmap construction e.g., with 3D point cloud from the radar (top view) (left), and constructed 2D heatmap (right).

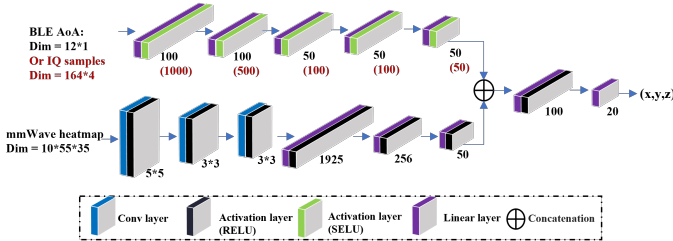


Figure 4: The concatenated two-branch NN architecture employed for feature fusion, where the top branch is the FCN aiming to process the BLE features, and the bottom branch is designed as the CNN to handle the mmWave heatmaps.

where P is a set representing a point cloud with a population of $|P|$. $G(p)$ creates a 2D Gaussian kernel at point p . The resulting heatmap, illustrated in Fig. 3, indicates a higher probability for points with a larger population, such as the person cluster in the middle, compared to the clutter on the left.

D. Fusion Neural Network Design

Fig. 4 shows the architecture of the designed fusion NN, which contains two branches for different inputs. The top branch is designed as fully connected network (FCN) for the processing of the BLE feature, while the bottom branch is a convolutional NN (CNN), taking the mmWave heatmap as the input feature. After the convolution, this feature is flattened to a linear layer and eventually coated with the output of the top NN branch. The output of this NN is the (x, y, z) coordinate of the estimated location. The neuron numbers of each linear layer and the involved activation functions are annotated in Fig. 4. Please note that, for BmmW and BmmW-LITE, the neuron number of each layer in FCN is different due to the diverse dimensions of the training features. Specifically, in BmmW, the estimated azimuth angle θ , elevation angle ϕ , and corresponding estimated distance to the objective of all anchors consist of the ultimate training feature, which is a 3×4 dimensional vector. Hence, the number of neurons of each linear layer is 100, 100, 50, 50, and 50. Instead, in BmmW-LITE, the training features are IQ samples with dimensions of 164×4 . Hence, the number of neurons of each linear layer is set to 1000, 500, 100, 100, and 50. Adam [46] is selected as the optimizer of the entire NN. The initial learning rate lr is the default value, while it decays according to $lr = lr * 0.3^{epoch//20}$, where $epoch$ indicates the index of current training epoch, and $//$ represents the discard remainder operation. The mean absolute error is the selected loss function for NN training, which is defined as

$loss = mean(|(\hat{x}, \hat{y}, \hat{z}) - (x, y, z)|)$, with (x, y, z) being the ground truth coordinate and $(\hat{x}, \hat{y}, \hat{z})$ the estimated coordinate.

E. Implementation of BLE and mmWave Data Branches

1) *BLE Data Branch*: The collection of BLE features is performed using Silabs EFR32xG22 boards [28], which serve as anchor nodes, and the implementation is performed by employing the direction-finding solution provided by Silabs' AoA implementation [44], [47]. The IQ sampling capability is added to this stack to collect the raw measurements, as detailed in [44]. The EFR32xG22 boards are equipped with a 4×4 antenna array, but as we do not need to switch between antennas in BmmW-LITE, we used only a single antenna for IQ collection. The EFR32BG22 Thunderboard, detailed in [48], is used as a target device that operates in the *connectionless-mode* sending the CTE packets in periodic advertisements [44]. For the BmmW, the BLE stack implementation is the same, but we used all the antennas on the array board to determine the phase difference of the incoming signal, as antenna switching is necessary to analyze the AoA technique [43].

2) *mmWave Radar Data Branch*: We use one IWR1843 radar from Texas Instruments, which operates at 77 GHz to 81 GHz. We define the size of the heatmap to be 55×35 pixels to cover the $5.5 \text{ m} \times 3.5 \text{ m}$ region (see Sec. IV-A), so that the resolution of the heatmap would be 0.1 m per pixel. To convert a point cloud to a heatmap, we project the point cloud to 2D, define a Gaussian kernel of radius 4 for each point, and add all the kernels together to the heatmap based on their 2D coordinates. The resulting heatmap is normalized between 0 and 1, where the value represents the radar's confidence in the target's location. During data recording, we first operate the radar towards the empty scene and record the background reflection from clutters. This reflection is converted to a clutter heatmap and subtracted from the actual data recording. Then, we operate the radar at 25 frames/sec, and stack 10 neighbouring frames into one data instance to reduce the weights of outliers. Each data instance is converted to a heatmap of the defined size. The heatmaps are fed into one branch of the NN and fused with the BLE data branch.

IV. REAL-TIME DATA COLLECTION

This section describes the scenario employed to collect real-time data for training and testing our neural network model.

A. Experimental Setup

We build a real-time indoor testbed for joint data collection from BLE and mmWave, employing four BLE anchors at the corner of a $5.5 \text{ m} \times 3.5 \text{ m}$ area and placing the mmWave transceiver in the middle between two BLE anchor nodes, as presented in Fig. 5. The BLE anchors nodes are placed at a height of 2 m from the ground, while the mmWave transceiver is placed at 0.8 m. A highly accurate Optitrack system consisting of eight cameras is erected around the edge of the site to gather ground truth coordinates of the mobile target. We first calibrate the Optitrack and achieve a localization accuracy of 0.44 mm.

We perform a total of 12 trials for data collection, with a person holding a BLE tag in its hand. In each trial, the person follows a random path in the experimental area. Furthermore,

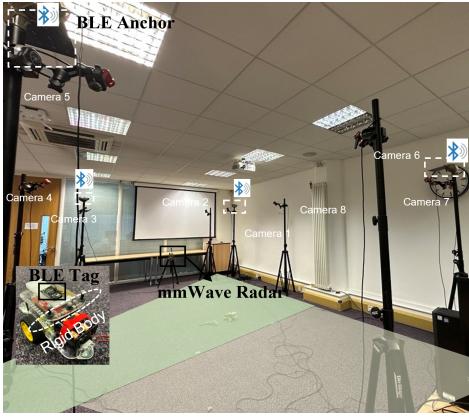


Figure 5: Experimental testbed used to simultaneously collect the real-time data from both BLE & mmWave sources with a moving target. The target, held by a human subject, follows a random trajectory within the testbed arena.

we perform the trials with three different people having different heights for the robustness of data collection. During this process, the BLE anchors continuously receive IQ packets from the asset tag (i.e., the target beacon), while the mmWave simultaneously collects the reflected signals of the moving person. Eventually, we collect 12 sessions through this experiment, with each session recording two minutes of data from BLE, mmWave radar, and Optitrack system. The total number of collected data, after synchronization, is around 66000.

B. Dataset Collection

The methodology for dataset collection is illustrated in Fig. 5. Three data streams (from BLE, mmWave, and OptiTrack) are synchronized according to the *UTC time stamping* for training the NN. For one sample, the BLE IQ features form a $164 * 4$ vector consisting of amplitude and phase information extracted from the recorded I/Q samples. The BLE AoA features for one sample consist of a $3 * 4$ vector with estimated azimuth angles, elevation angles, and corresponding distances. As discussed in Sec. III-E2, one mmWave sample contains a $55 * 25 * 10$ array as the feature. To evaluate the effects of the mmWave signal in different regions, the entire experimental area is divided into two parts: the *mmWave strong* area, which is the square area in front of the mmWave board (with a size of $3 \text{ m} \times 3 \text{ m}$, marked in light green); and the *mmWave weak* area, which is far from the mmWave board (marked in light yellow). This division is necessary because $\approx 48\%$ of the 66000 data instances of the mmWave radars fail to detect the target person at all, confirming the argument that the mmWave radar alone has a limited range/angle of view, thus emphasizing the importance of combining multiple technologies in BmmW.

V. EXPERIMENTAL EVALUATION

This section describes the localization performance metric and demonstrates the achieved 3D localization for a mobile target.

A. Evaluation Metric

The model evaluation is conducted on a server with Intel 2 E5-2640v4 CPU, and 2 RTX 2080Ti GPU. The splitting of

the training and test dataset follows the 80%-20% principle. For the model training and test, the batch size is set to 100, and early stopping is adopted, with the stopping patience equal to 10. Meanwhile, in order to eliminate the effect of data splitting in model performance, the 5-fold cross-validation scheme is taken for every evaluation. The model performance evaluation criteria in this paper is the Mean Localization Error (MLE), which is defined as the averaged Euclidean distance between the predicted location and the ground truth location among all test samples, as shown in Eq. 3:

$$MLE = \sum_{i=1}^N \sqrt{(\hat{x}_i - x_i)^2 + (\hat{y}_i - y_i)^2 + (\hat{z}_i - z_i)^2} / N \quad (3)$$

where N represents the number of test data. *It is worth mentioning that the NN predictions $(\hat{x}, \hat{y}, \hat{z})$ are raw predictions without any further processing like smoothing or filtering.*

B. Results

We evaluated the performance of BmmW and BmmW-LITE with different numbers of BLE anchors, ranging from 1 to 4 represented as BLE* k , where k indicates the number of BLE devices used. The results of our evaluation are presented in Tab. I, which shows the comparison of BmmW and BmmW-LITE against results obtained without fusion with mmWave radar. Feature *fusion* provides clear benefits for both methods across all scenarios, with the highest accuracy gain of 53.91% achieved in the case of three BLE anchors. The highest accuracy achieved is 0.09 m and 0.341 m for BmmW and BmmW-LITE, respectively, which is 80% and 60% higher than that of classical BLE localization methods [5]. BmmW provides significantly higher localization accuracy, especially with an increase in BLE anchors. Moreover, by comparing different rows in Table I, it is evident that the performance improvement of the fusion NN model in the mmWave strong region is greater than that in the mmWave weak region. This is due to the decay of mmWave signals with increasing detect distance. Although the mmWave radar may fail to detect the person around half of the time, that information can still be helpful as it indicates that the person may not be in the mmWave strong area. The highest improvement in BmmW-LITE is observed with 1 BLE anchor fused with mmWave heatmap, which reduces error by 40.07%. Notably, even when using a single BLE anchor, the fusion model achieves sub-metre accuracy in all testing areas, with a maximum error of 0.73 m.

In addition, we use the Cumulative Distribution Function (CDF) to statistically evaluate the localization performance of BmmW and BmmW-LITE in different scenarios, as shown in Fig 6. The CDF results show that BmmW achieves almost 90% localization accuracy within 0.5 m in all scenarios. Especially with the ‘BLE*4+mmWave’ scenario shown in Fig. 6b (which corresponds to the fusion of BLE and mmWave measurements when using 4 BLE anchors), the CDF curve is extremely steep, demonstrating highly-accurate predictions. BmmW-LITE achieves 60% localization accuracy under 50 cm across all scenarios, and up to 90% when using four BLE anchors. Furthermore, to visualize the 2D tracked trajectory, we selected random predicted locations from the test set

Table I: The MLE (in metres) in the test set under different scenarios using BmmW model and BmmW-LITE model.

Area	Method	BLE*1		BLE*2		BLE*3		BLE*4	
		BmmW	BmmW-LITE	BmmW	BmmW-LITE	BmmW	BmmW-LITE	BmmW	BmmW-LITE
Entire	BLE alone	0.663	0.914	0.398	0.727	0.291	0.664	0.222	0.582
	Fusion BLE & mmWave	0.643	0.731	0.320	0.578	0.134	0.428	0.107	0.369
	Gains of Fusion over BLE	3.05%	19.94%	19.58%	20.57%	53.91%	35.55%	51.62%	36.53%
mmWave Strong	BLE alone	0.420	0.642	0.233	0.501	0.213	0.517	0.181	0.473
	Fusion BLE & mmWave	0.331	0.385	0.227	0.341	0.191	0.398	0.090	0.362
	Gains of Fusion over BLE	21.20%	40.07%	2.48%	31.95%	10.27%	22.97%	50.26%	23.40%
mmWave Weak	BLE alone	0.646	0.704	0.427	0.658	0.253	0.547	0.191	0.478
	Fusion BLE & mmWave	0.524	0.656	0.426	0.502	0.242	0.464	0.148	0.402
	Gains of Fusion over BLE	18.95%	6.87%	0.28%	23.67%	4.37%	15.28%	22.40%	15.90%

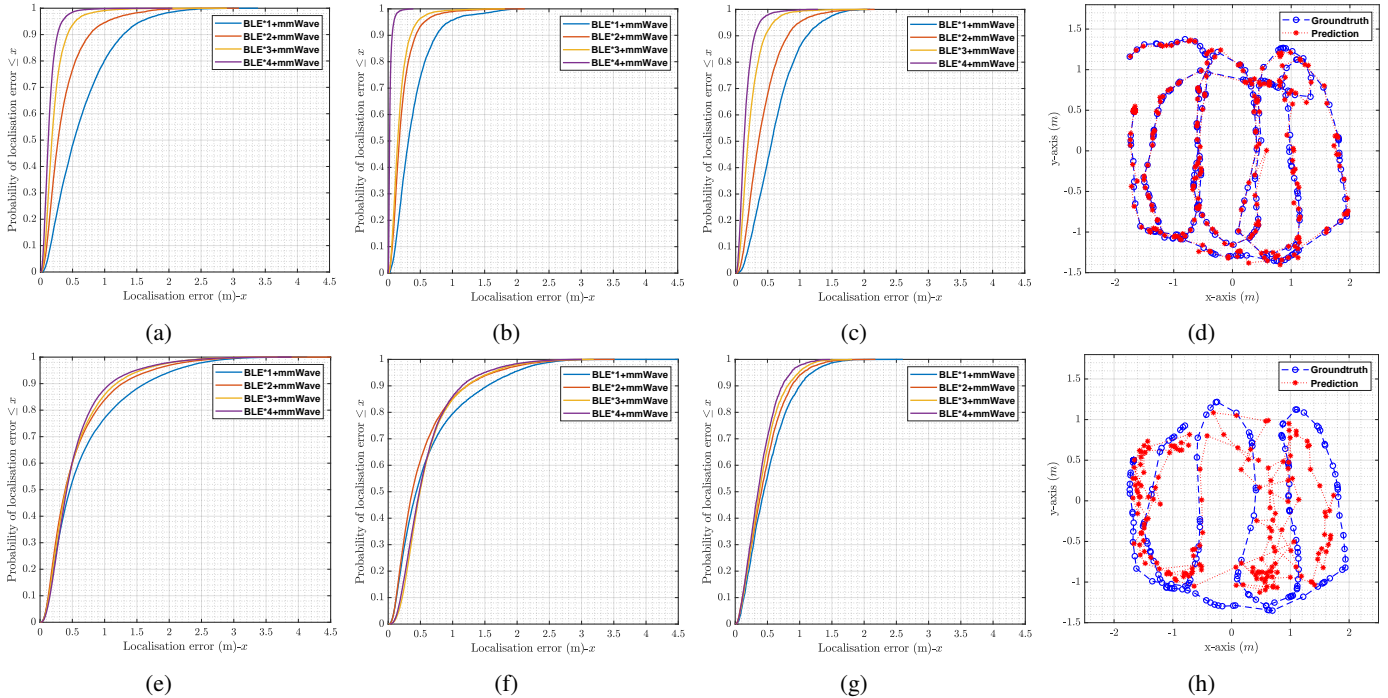


Figure 6: The Mean Localization Error (MLE) CDF of the BmmW and BmmW-LITE models in an entire area (a,e); in the mmWave strong area (b, f); and in the mmWeak area (c, g). Comparison of the ground truth locations and NN-predicted locations in a part of the test set for BmmW (d) and BmmW-LITE (h).

Table II: % increment in MLE for BmmW w.r.t. BmmW-LITE.

Areas	BLE*1	BLE*2	BLE*3	BLE*4
Entire Room	12.12%	44.65%	68.65%	70.92%
mmWave Strong	14.01%	33.27%	51.94%	75.12%
mmWave Weak	20.1%	15.12%	47.74%	63.15%

and compared them with the ground truth locations. This comparison is shown in Fig. 6d and 6h for the “Entire room with BLE*4” scenario with BmmW and BmmW-LITE, respectively. BmmW’s estimated trajectory closely matches the ground truth trajectory. BmmW-LITE’s estimated trajectory also matches the ground truth, but there are some out-of-the-box predictions at certain locations hindering its performance. These outliers can be removed in a post-processing step, if necessary.

C. Discussion and Future Work

Complexity. In BmmW, a mobile tag continuously sends CTE packets: the latter are sent to a central server for NN model training and validation. Despite complexity concerns for real-world and real-time deployments, please note that the model has a rather small size, and can be deployed on common edge

devices [49], [50] after network compression and pruning, allowing inference time to approach real-time levels.

Accuracy vs costs. Tab. II quantitatively compares the performance of BmmW and BmmW-LITE. Even though the use of raw IQ measurements in BmmW-LITE causes a loss in accuracy, it requires less computational efforts. In fact, BmmW involves the use of AoA measurements obtained by running the MUSIC algorithm, which has a complexity of $O(N^3)$, where N is the number of antennas [51]. Moreover, BmmW requires multiple bulky antenna arrays, whereas BmmW-LITE offers a less complex, cost-efficient solution. Still, both methods offer superior performance than the state-of-the-art [5], [6].

Scalability. However, the addition of more BLE anchors increases the coverage area, but also incurs extra costs. We tested BmmW with one office environment, the training and testing were performed in the same environment, which may lead to the unsatisfied generalisation ability of the NN. Hence, more diverse data collections from different dynamic & large indoor environments should be implemented in future work.

Clock-drift. Multi-modality sensing models may experience

clock drift caused by differences in sampling rates between the modalities. BLE boards, on the other hand, have a predictable curve of clock drift which can be used to mitigate this issue [52]. Additionally, addressing the discrepancy in sensing frequency between diverse sensors can be investigated in future studies.

VI. CONCLUSION

Our paper introduces BmmW, a novel localization system that combines the strengths of BLE 5.1 direction-finding and mmWave radar technology through a DNN-based fused model, achieving decimetre-level accuracy. We present two methods for incorporating BLE data into the NN model: BmmW utilizes ranging data, while BmmW-LITE uses raw IQ measurements. Experimental results show that both BmmW and BmmW-LITE sustain decimetre-level accuracy, with a mean localization error of only 10 and 36 cm, respectively, an improvement of 80% and 60% compared to classical BLE localization methods. Moreover, 90% of the error is under 50 cm for both approaches, making it suitable for mobile applications. BmmW outperforms BmmW-LITE due to its additional processing and data filtration, but BmmW-LITE offers a computationally and cost-efficient system that eliminates the need for bulky multi-antenna arrays. By combining BLE and mmWave radar technology, BmmW overcomes the limitations of conventional techniques and offers a practical solution for high-accurate localization in ‘Beyond 5G’ wireless communication systems.

ACKNOWLEDGEMENTS

This work has been funded by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska Curie grant agreement ENLIGHT’EM No 814215.

REFERENCES

- [1] F. Liu *et al.*, “Integrated Sensing and Communications: Towards Dual-functional Wireless Networks for 6G and Beyond,” *IEEE JSAC*, 2022.
- [2] J. Zhao *et al.*, “Research and Implementation of Autonomous Navigation for Mobile Robots Based on SLAM Algorithm under ROS,” *Sensors*, 22.
- [3] Y. D. Yasuda *et al.*, “Autonomous Visual Navigation for Mobile Robots: A Systematic Literature Review,” *ACM CSUR*, 2020.
- [4] C. S. Álvarez-Merino *et al.*, “WiFi FTM, UWB and Cellular-based Radio Fusion for Indoor Positioning,” *Sensors*, vol. 21, no. 21, 2021.
- [5] M. Cominelli *et al.*, “Dead on Arrival: An Empirical Study of The Bluetooth 5.1 Positioning System,” in *Proc. of the 13th WiNTECH Workshop*, 2019.
- [6] L. Mo *et al.*, “UHF RFID Indoor Localization Algorithm Based on BP-SVR,” *IEEE Journal of Radio Frequency Identification*, 2022.
- [7] J. Singh *et al.*, “HueSense: Featuring LED Lights Through Hue Sensing,” in *Proc. of the 1st ACM MORSE Workshop*, 2023.
- [8] J. Singh *et al.*, “Detecting and Controlling Smart Lights with LiTalk,” in *Proc. of the 1st ACM MORSE Workshop*, 2023.
- [9] D. Lymberopoulos *et al.*, “The Microsoft Indoor Localization Competition: Experiences and Lessons Learned,” *IEEE Sig. Proc. Mag.*, 2017.
- [10] ABI Research, “2021 Market Update,” 2021.
- [11] Z. Gao *et al.*, “Integrated Sensing and Communication with mmWave massive MIMO: A Compressed Sampling Perspective,” *IEEE Transactions on Wireless Communications*, 2022.
- [12] H. Cui *et al.*, “Millimetre-wave Radar for Low-Cost 3D Imaging: A Performance Study,” *arXiv preprint 2301.13553*, 2023.
- [13] P. Li *et al.*, “Wireless Localisation in WiFi using Novel Deep Architectures,” in *Proc. of the 25th ICPR Conf.*, 2021.
- [14] P. Li *et al.*, “Deep Transfer Learning for WiFi Localization,” in *Proc. of the IEEE RadarConf Conf.*, 2021.
- [15] K. He *et al.*, “Deep Residual Learning for Image Recognition,” in *Proc. of the IEEE CVPR Conf.*, 2016.
- [16] D. Silver *et al.*, “Mastering Chess and Shogi by Self-play with a General Reinforcement Learning Algorithm,” *arXiv preprint 1712.01815*, 2017.
- [17] OpenAI, “ChatGPT: Optimizing Language Models for Dialogue,” 2023.
- [18] X. Li *et al.*, “Design and Implementation of Indoor Positioning System based on iBeacon,” in *Proc. of the ICALIP Conf.*, 2016.
- [19] A. Dasgupta *et al.*, “An Internet of Things Platform with Google Eddystone Beacons,” *IJSEA*, vol. 9, no. 6, 2016.
- [20] Y. Wang *et al.*, “RSSI-Based Bluetooth Indoor Localization,” in *Proc. of the 11th MSN Conf.*, 2015.
- [21] Z. Hajiakhondi *et al.*, “Bluetooth Low Energy-based AoA Estimation in Presence of Rayleigh Fading,” in *Proc. of the IEEE SMC Conf.*, 2020.
- [22] X.-Y. Lin *et al.*, “A Mobile Indoor Positioning System based on iBeacon Technology,” in *Proc. of the 37th IEEE EMBC Conf.*, 2015.
- [23] M. Ji *et al.*, “Analysis of Positioning Accuracy Corresponding to the Number of BLE Beacons in Indoor Positioning System,” in *Proc. of the 17th ICACT Conf.*, 2015.
- [24] H. Ye *et al.*, “A Method of Indoor Positioning by Signal Fitting and PDDA Algorithm using BLE AOA Device,” *IEEE Sensors Journal*, 2022.
- [25] M. Qian *et al.*, “Performance Analysis of BLE 5.1 New Feature Angle of Arrival for Relative Positioning,” *The Intl. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 46, 2022.
- [26] G. Pau *et al.*, “Bluetooth 5.1: An Analysis of Direction Finding Capability for High-Precision Location Services,” *Sensors*, vol. 21, no. 11, 2021.
- [27] S. He *et al.*, “Multi-antenna Array-based AoA Estimation Using BLE for Indoor Positioning,” in *Proc. of the 7th ICCO Conf.*, 2021.
- [28] Silicon Labs, “BG22 Bluetooth Dual Polarized Antenna Array Radio Board.”
- [29] Texas Instruments, “xWR1843 Evaluation Module Single-Chip mmWave Sensing Solution – User’s Guide SPRUM4B,” May 2020.
- [30] P. Zhao *et al.*, “mID: Tracking and Identifying People with Millimeter Wave Radar,” in *Proc. of the 15th DCOSS Conf.*, 2019.
- [31] H. Cui *et al.*, “High Precision Human Detection and Tracking Using Millimeter-wave Radars,” *IEEE Aerosp. Electron. Syst. Mag.*, 2021.
- [32] C. Wu *et al.*, “mmTrack: Passive Multi-Person Localization Using Commodity mmWave Radio,” in *Proc. of the INFOCOM Conf.*, 2020.
- [33] J. Vitola *et al.*, “A Sensor Data Fusion System Based on k-nearest Neighbor Pattern Classification for Structural Health Monitoring Applications,” *Sensors*, vol. 17, no. 2, 2017.
- [34] R. Gu *et al.*, “Multi-modal Multi-channel Target Speech Separation,” *IEEE Journal of Selected Topics in Signal Proc.*, vol. 14, no. 3, 2020.
- [35] R. King *et al.*, “Application of Data Fusion Techniques and Technologies for Wearable Health Monitoring,” *Medical Engineering & Physics*, 2017.
- [36] D. Michelsanti *et al.*, “An Overview of Deep-Learning-based Audio-Visual Speech Enhancement and Separation,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, 2021.
- [37] J. Liu *et al.*, “Fusing WiFi, Inertial Sensors and Bluetooth Beacons for Accurate Indoor Positioning,” *Sensors*, vol. 21, no. 1, 2021.
- [38] P. Balasubramanian *et al.*, “Multi-technology based Indoor Localization using UWB and BLE Signals,” *JAIHC*, vol. 12, no. 2, 2021.
- [39] S. Jeong *et al.*, “Robust Indoor Positioning System with UWB, Bluetooth, and Machine Learning,” *JAIHC*, vol. 12, no. 7, 2021.
- [40] T. Istomin *et al.*, “Janus: Efficient and Accurate Dual-radio Social Contact Detection,” *ACM IMWUT*, vol. 5, no. 4, 2021.
- [41] K. Ma *et al.*, “A Differentially Private Indoor Localization Scheme with Fusion of WiFi and Bluetooth Fingerprints in Edge Computing,” *IEEE Access*, vol. 9, 2021.
- [42] S. Deb *et al.*, “Indoor Localization using Magnetometer and Wi-Fi Fingerprinting Techniques,” *JAIHC*, vol. 11, no. 3, 2020.
- [43] S. Bluetooth, “Bluetooth Core Specification v5.1,” 2019.
- [44] Silicon Labs, “QSG175: Direction Finding Solution Quick-Start Guide.”
- [45] S. Rao, “MIMO radar,” tech. rep., Texas Instruments, 2017.
- [46] D. P. Kingma *et al.*, “Adam: A Method for Stochastic Optimization,” *arXiv preprint 1412.6980*, 2014.
- [47] Silicon Labs, “AN1297: Custom Direction-Finding Solutions using the Silicon Labs Bluetooth Stack.”
- [48] Silicon Labs, “SLTB010A EFR32BG22 Thunderboard Kit.”
- [49] NVIDIA, “Advanced AI Embedded Systems,” 2023.
- [50] L. Mocerino *et al.*, “CoopNet: Cooperative Convolutional Neural Network for Low-Power MCUs,” in *Proc. of 26th ICECS Conf.*, 2019.
- [51] J. Kim, D. Oh, S. Park, S. Lee, and Y. Jung, “Design of MUSIC-based DoA Estimator for Bluetooth Applications,” *Journal of IKEE*, vol. 24, no. 1, pp. 339–346, 2020.
- [52] M. Baddeley *et al.*, “Understanding Concurrent Transmissions: The Impact of Carrier Frequency Offset and RF Interference on Physical Layer Performance,” *Cornell University – arXiv:2304.00371*, 2023.