Preprints.org

Wi-Fi RSS and RTT Indoor Positioning with Graph Temporal Convolution Network

Lila Rana [*] and Aayush Dulal

Article

Wi-Fi RSS and RTT Indoor Positioning with Graph Temporal Convolution Network

Lila Rana *[iD] and **Aayush Dulal**

University of Texas at Dallas
*　　Correspondence: lila.rana@utdallas.edu

Abstract

Indoor positioning using commodity Wi-Fi has gained significant attention; however, achieving sub-meter accuracy across diverse layouts remains challenging due to multipath fading and non-line-of-sight (NLOS) effects. In this work, we propose a hybrid Graph–Temporal Convolutional Network (GTCN) model that incorporates access point (AP) geometry through graph convolutions while capturing temporal signal dynamics via dilated temporal convolutional networks. The proposed model adaptively learns per-AP importance using a lightweight gating mechanism and jointly exploits Wi-Fi Received Signal Strength (RSS) and Round Trip Time (RTT) features for enhanced robustness.It is evaluated across four experimental areas such as lecture theatre, office, corridor, and building floor covering areas from 15×14.5 m^2 to 92×15 m^2. We further analyze the sensitivity of the model to AP density under both LOS and NLOS conditions, demonstrating that positioning accuracy systematically improves with denser AP deployment, particularly in large-scale mixed environments. Despite its high accuracy, the proposed GTCN remains computationally lightweight, requiring fewer than 10^5 trainable parameters and only tens of MFLOPs per inference, enabling real-time operation on embedded and edge devices.

Keywords: WiFi indoor positioning; WiFi RSS and RTT; temporal convolution network; graph convolution network

---

1. Introduction

Global Positioning System (GPS) [1] has been the state of the art when it comes to outdoor location estimation. In contrast, indoor positioning is a more difficult problem because of obstacles and environmental factors, which are both natural and human-developed. Wi-Fi positioning system (WPS) makes accurate indoor positioning possible by using wireless signals and access points [2–12]. These systems rely on the abilities of components to quantify things such as signal strength and signal travel time. In addition, it's important that the components of the system are economically viable and easy to deploy [13].

Most wireless indoor positioning systems can be classified as either range-based or fingerprint-based. Range-based methods rely on metrics such as received signal strength (RSS), round-trip time (RTT), and time of arrival (ToA). Fingerprinting methods work by building a radio map of signal features, which is used to compare with new measurements to estimate the closest matching position [14]. Proximity-based methods have also been proposed due to their simplicity; however, accuracy decreases during operation [15]. Use of Bluetooth [16], ultrasonic signals [17], radio frequency [18], and computer vision [19] has also been studied, but each has its disadvantages. The use of Bluetooth is limited by coverage problems due to its inability to scale to larger areas, and ultrasonic-based devices have a high production and deployment cost. In contrast, RFID-based technology is cheaper but, like Bluetooth, lacks coverage. ZigBee technology [20] is inexpensive and power consumption is low in comparison to other technologies, but has problems in maintaining stability during operation. Use of computer vision has been extremely common in numerous applications, including indoor positioning.

Despite its popularity, computer vision-based devices and algorithms leverage camera feed, which is prone to disturbance and obstacles. Because of these limitations in different technologies, Wi-Fi-based solutions have generated interest in the research and development of indoor positioning systems.

Using a fingerprinting-based method, the problem of indoor positioning can be broken down into two phases: the offline phase and the online phase. The offline phase is when signal fingerprints are obtained at known positions to build a dataset, and during online query, the received signal strength indicator (RSSI) vector [2,4,5,8] is measured and compared with entries in the database from the phase before. This step can be considered the training phase for simplicity. The training dataset is developed by collecting metrics such as channel state information (CSI) and RSS. This input feature is collected at each reference point (RP) from all access points (AP) [21]. Algorithms based on K-nearest neighbors (KNN) [5,22], Generative Adversarial Network (GAN) variants such as Conditional GAN (CGAN) [23], Self-Normalizing GAN (SNGAN) [24], and Conditional Tabular GAN (CTGAN) [25], as well as Gaussian Process Regression (GPR) [26], deep neural networks (DNN), and random forest (RF) [27] are popular approaches for Wi-Fi fingerprinting-based indoor positioning. All these are supervised learning based methods that require extensive data to be effectively trained. Generating a rich dataset every time is not practical, which brings forward the possibility of using unsupervised learning algorithms. Both these phases can be made more effective by leveraging the temporal and spatial patterns of signals, using additional sensors, and using clever sensor fusion techniques like Kalman Filters [28]. Unlike fingerprinting methods, ranging methods can predict the final location of the user using trilateration or multilateration techniques [2], [4].

The core part of range-based positioning is to generate a path loss model from the RSS and distance information, which makes them prone to larger position errors and high computational costs. To overcome this, channel state information (CSI) can be used instead of RSS, but the hardware requirements to generate and deliver CSI are of concern. CSI can provide more accurate and stable positioning results, but it needs special hardware to work. The fine time measurement (FTM) protocol in Wi-Fi,introduced in the IEEE 802.11-2016 standard [29], can also offer high-precision distance estimates between a Wi-Fi RTT-enabled smartphone and APs. This protocol measures the distance by sending ranging requests from the smartphone to the APs without needing to establish a connection. It gathers both RSS and RTT data within a short time. Unlike the time of flight (ToF) method, FTM does not need clock synchronization between the smartphone and the APs, which makes the system simpler and more dependable. The FTM protocol was later extended through successive IEEE amendments to enhance ranging accuracy and robustness. Specifically, the IEEE 802.11az amendment improved indoor positioning by introducing multi-user MIMO (up to 8×8) for better performance in NLOS environments. It also added support for 160 MHz channels, enhanced security through protected long training fields (LTF) and media access control (MAC) frames, enabled scalable multi-station ranging, and optimized power efficiency and location-based link adaptation. As a result, the ranging accuracy improved from the 1–2 m level of IEEE 802.11mc to below 1 m. The more recent IEEE 802.11bk amendment further extends these capabilities by allowing measurements over 320 MHz channels, achieving sub-decimeter (< 0.1 m) precision—often referred to as "product-on-shelf" accuracy [30].

While fingerprinting using RSS and RTT measurements is popular, they have limitations specific to them. This includes high sensitivity to environmental conditions, limited spatial resolution, and more common NLOS effects. NLOS effects can be identified and removed but if every valid signal is prone to NLOS conditions, removing them defeats the purpose. To mitigate these limitations, [31] proposes combining RTT and RSS measurements while generating a fingerprinting map, which results in improved position accuracy. In contrast to the conventional ToF methods that require clock synchronization between transmitter and receiver, the introduction of FTM enables high-precision ranging between smartphones and access points by measuring the RTT. This significantly reduces the time to obtain RTT and RSS readings within short sampling intervals, which makes this approach more practical and reliable [32]. Despite this development, RTT measurements and the resulting positioning estimates suffer greatly from NLOS conditions because of indoor obstacles and signal reflections.

To enhance positioning accuracy, the spatial relationships among APs can be effectively exploited by modeling their pairwise connectivity and optimizing AP placement. In [33], [34]centimeter-level indoor positioning accuracy was achieved through precise AP localization. For graph-based modeling, inter-AP edges are typically weighted using the inverse of their physical distance, allowing the network to capture spatial proximity and signal correlations across APs. In parallel, temporal dependencies between consecutive Wi-Fi scans can be modeled to mitigate the effects of noise and multipath fading. The study in [35] demonstrated these principles by integrating Graph Convolutional Networks (GCNs) and Temporal Convolutional Networks (TCNs) into a Graph–Temporal Convolutional Network (GTCN) for processing CSI data. However, because CSI extraction requires specialized hardware and firmware modifications, such approaches have limited deployability on commodity Wi-Fi devices.

Inspired by this concept, the present work introduces a lightweight hybrid GTCN that jointly models spatial dependencies among APs through graph convolutions and temporal dynamics of Wi-Fi RSS and RTT measurements via causal dilated convolutions. By fusing both modalities, the proposed framework captures complementary spatial–temporal patterns while remaining hardware-compatible and easily deployable. To the best of our knowledge, this is the first study to employ a TCN-based architecture for Wi-Fi RSS and RTT-based indoor positioning.

The main contributions of this paper are summarized as follows:

- We propose an enhanced Wi-Fi indoor positioning system that jointly utilizes RSS and RTT measurements. This hybrid approach exploits the complementary advantages of signal strength–based and time-of-flight–based ranging to achieve robust sub-meter accuracy across diverse environments, including LOS, NLOS, and mixed conditions.
- We introduce a computationally efficient GTCN that explicitly models spatial dependencies among APs using graph convolutions with inverse-distance edge weighting, while concurrently capturing causal temporal correlations between consecutive signal scans through dilated TCNs. This unified design enhances both spatial consistency and temporal robustness while remaining suitable for real-time embedded deployment.
- We systematically investigate the impact of AP density under both LOS and NLOS conditions to assess the scalability and robustness of the proposed model. This analysis highlights the adaptability of the proposed hybrid GTCN model under both sparse and dense AP deployments.

The remainder of this paper is organized as follows. Section 2 reviews the related work, followed by the system model and the proposed methodology in Section 3. Section 4 presents the experimental datasets, implementation details, and performance evaluation results. Finally, Section 5 concludes the paper and outlines potential directions for future research.

2. Related Work

In recent years, Wi-Fi RTT positioning has gained significant attention as a practical and standards-compliant approach for indoor positioning using IEEE 802.11mc and its successors. Early RTT-based works focused on improving ranging accuracy and mitigating multipath bias. A recent commercial-grade implementation, [36] demonstrates the maturity of RTT-based indoor positioning using only commodity smartphones and access points. They employs a Wi-Fi RTT positioning system using an Extended Kalman Filter (EKF) with random-walk and step-heading motion models, achieving sub-metre error across diverse walking datasets.[37] employed stack ensemble learning for indoor positioning, utilizing support vector regression and the XGBoost algorithm.

To improve robustness under mixed LOS and NLOS conditions, [38] introduces an LOS identification and range calibration framework that combines Gaussian Process Regression (GPR)-based scenario recognition with LOS distance correction, yielding 0.99 m RMSE. In harsh multipath environments, [39] extends this concept by employing GPR with Particle Swarm Optimization to model RTT measurement differences, reducing mean error by 68.5% compared to least squares. There are some other work works to enhance the performance in complex NLOS conditions, [40] proposes an RTT localization algorithm with LOS compensation and trusted NLOS recognition using a support

vector machine (SVM) classifier and Bayesian selection, achieving a 53% improvement over baseline least-squares methods. Complementary work [41] presents an automated site-survey approach that integrates GPS and smartphone sensors to determine AP coordinates for RTT-based localization, reducing pre-deployment effort while maintaining sub-meter accuracy.

Building on these RTT-based advancements for indoor positoning, researchers have increasingly explored hybrid RTT–RSS fusion methods that take advantage of the complementary properties of signal strength and time-of-flight measurements to improve positioning accuracy and robustness. The work in [42] introduces a hybrid Wi-Fi RTT–RSS ranging framework that compensates for transmitter-side clock skew and calibrates RTT offsets, achieving a mean positioning error of 1.43 m with 0.19 s update intervals. Similarly, [43] improves the classical RSS logarithmic path-loss model using quadratic fitting and fuses calibrated RTT and RSS via adaptive weighting, yielding notable accuracy gains in complex indoor layouts. The study in [44] further analyzes the joint contribution of RTT and RSS features in fingerprinting, demonstrating that coupling both measurements reduces network overhead while improving scalability and precision.

From a learning perspective, [45] proposes a deep fusion model (RRLoc) that integrates time-based and fingerprinting features via deep canonical correlation analysis, improving localization accuracy by over 250% relative to conventional fingerprinting and multilateration methods. A dynamic model-switching approach is proposed in [46], which adaptively selects the optimal positioning model per environment using weighted machine learning, yielding up to 1.8 m improvement over traditional methods. To address multipath-induced instability, [47] develops real-time NLOS/LOS identification algorithms based on short RSS/RTT sequences, achieving 96% discrimination accuracy with minimal latency.

Recent efforts have also explored multi-sensor fusion and optimization frameworks. [48] presents a tightly coupled integration platform that combines Wi-Fi RTT, RSS, and MEMS-IMU data using adaptive filtering, achieving a 20% improvement in accuracy and robustness over standard EKF approaches. A similar concept is extended in [49], where factor-graph optimization (FGO). Similarly, in [50], implicit connectivity is proposed for GNN-based cooperative localization filter variants. In [51], a particle-filter-based system fuses Wi-Fi RTT, RSS, PDR, and map constraints, using semi-parametric error models and collaborative optimization to achieve meter-level accuracy within 0.19 s. These studies collectively underscore the potential of RTT and RSS fusion, often complemented by inertial or map-based constraints, to enhance robustness against multipath and NLOS effects. However, most existing methods process spatial and temporal information independently, without explicitly modeling the spatial graph topology of APs or the temporal evolution of signal dynamics.

These studies collectively underscore the potential of RTT and RSS fusion often complemented by inertial or map-based constraints to enhance robustness against multipath and NLOS effects. However, most existing methods process spatial and temporal information independently, without explicitly modeling the spatial graph topology of APs or the temporal evolution of signal dynamics. While CSI-based deep architectures can learn such spatio–temporal correlations, they require hardware modifications that limit deployment feasibility on commercial devices.

In [52], the authors introduce a mobility-induced graph learning framework that models user movement as graphs to enhance Wi-Fi positioning accuracy, leveraging cross-graph and self-supervised learning to achieve robust performance without the need for labeled data. [53] applies graph neural networks to nonlinear regression in network localization, demonstrating their efficiency and robustness. [54] proposes a hybrid belief-propagation–GNN framework that improves cooperative localization consistency without increasing complexity. [55] presents a vision-assisted, privacy-preserving approach for generating data on large-scale Wi-Fi positioning. Similarly, [50] introduces implicit connectivity modeling and self-attention embeddings for cooperative localization, while [56] and [57] demonstrate the strength of temporal convolutional networks (TCNs) for sequential modeling tasks, including Bluetooth-based indoor localization and human activity recognition. In parallel, recent studies have explored the application of TCN in Wi-Fi-based sensing tasks such as human activity and interaction

recognition. These works demonstrate the strong temporal modeling capabilities of TCNs in extracting discriminative motion patterns from wireless signals. For instance,[58] proposed a lightweight mobile TCN (LM-TCN) that employs depthwise separable convolutions and gated residual mechanisms for low-power, memory-efficient human activity recognition across multiple indoor locations, achieving 95.2% accuracy while reducing computational cost to 6% of the baseline TCN. [59] introduced Wi-ATCN, an attentional temporal convolutional network that integrates self-attention with Wi-Fi CSI for human action prediction, improving performance stability across environments. More recently [60] proposed WiFi-TCN, which combines temporal convolution, data augmentation, and attention mechanisms for efficient human–human interaction recognition, achieving 99.4% accuracy and demonstrating high generalization across subjects and scenarios.

While these TCN-based frameworks have demonstrated outstanding results for recognition tasks, they primarily focus on classifying human motion using high-resolution CSI data. In contrast, the proposed hybrid GTCN framework adapts the temporal modeling strength of TCNs to the problem of Wi-Fi-based indoor positioning using low-dimensional RSS and RTT fingerprints. By jointly exploiting spatial graph reasoning among access points and causal temporal dependencies within signal sequences, our approach bridges the gap between lightweight fingerprinting and deep spatio–temporal learning, achieving robust and hardware-compatible sub-meter positioning performance.

3. System Model and Proposed Methodology

3.1. Preliminaries and Problem Formulation

This study addresses Wi-Fi indoor positioning using measurements obtained directly from IEEE 802.11-compliant devices. We use WiFi RSS and RTT measurements as input features, both of which are measurable from commodity APs without requiring physical-layer modifications or dedicated ranging hardware. The objective is to estimate user coordinates in real time with a computationally efficient model suitable for embedded and edge platforms.

Each environment is discretized into N_{RP} reference points (RPs). The coordinates of RP r are denoted by $\mathbf{y}_r = [x_r, y_r]^\top$ in meters, determined by a grid scale factor g derived from the floor layout. At each RP, Wi-Fi scans are collected from up to N APs. The t-th measurement from AP i at RP r is

$$\left(x_{r,t,i}^{\mathrm{rss}},\ x_{r,t,i}^{\mathrm{rtt}}\right),$$

where $x_{r,t,i}^{\mathrm{rss}}$ is the RSS (dBm) and $x_{r,t,i}^{\mathrm{rtt}}$ is the RTT (mm).

For each RP r, the RSS and RTT samples form

$$\mathbf{R}_r^{(\mathrm{rss})} \in \mathbb{R}^{T_r \times N}, \qquad \mathbf{R}_r^{(\mathrm{rtt})} \in \mathbb{R}^{T_r \times N},$$

where T_r is the number of scans at RP r. The complete fingerprint database is

$$\mathcal{D} = \left\{ \mathbf{R}_r^{(\mathrm{rss})},\ \mathbf{R}_r^{(\mathrm{rtt})},\ \mathbf{y}_r \right\}_{r=1}^{N_{\mathrm{RP}}}.$$

The database is segmented into temporal windows of fixed length T to construct training sequences $\mathbf{X} \in \mathbb{R}^{T \times N \times F}$, where F is the feature dimension. Each sequence corresponds to one RP label $\mathbf{y}_r$.

3.1.1. Signal Modeling and Objective

WiFi RSS and RTT measurements provide complementary information on the propagation channel. RSS reflects large-scale attenuation and shadowing, while RTT captures geometric propagation delay. Both are affected by multipath, hardware bias, and NLOS conditions, but their error characteristics are only partially correlated. Integrating WiFi RTT and RSS measurements enables the model to mitigate individual weaknesses, such as multipath distortion in RTT and shadowing in RSS, resulting in improved overall positioning accuracy.

For a receiver located at $\mathbf{p} = [x, y]^\top$, the RSS from AP i follows the log-distance model

$$\text{RSS}_i(\mathbf{p}) = P_0 - 10n \log_{10}\left(\frac{d_i(\mathbf{p})}{d_0}\right), \tag{1}$$

where P_0 is the received power (dBm) at reference distance d_0, n is the path-loss exponent, and $d_i(\mathbf{p}) = \|\mathbf{p} - \mathbf{p}_i^{(\text{AP})}\|_2$ is the AP–user distance. In practice, $\text{RSS}_i(\mathbf{p})$ exhibits stochastic deviations due to multipath fading, human blockage, and environmental dynamics.

The corresponding RTT measurement for AP i can be modeled as

$$t_{\text{RTT},i} = \frac{2d_i(\mathbf{p})}{c} + \Delta_i + \eta_i, \tag{2}$$

where c is the speed of light, Δ_i is an AP-dependent hardware bias, and η_i is zero-mean noise due to timestamp quantization and multipath. The induced distance estimate is

$$\tilde{d}_i(\mathbf{p}) = \frac{c\, t_{\text{RTT},i}}{2} = d_i(\mathbf{p}) + \frac{c}{2}\left(\Delta_i + \eta_i\right), \tag{3}$$

which typically overestimates $d_i(\mathbf{p})$ in NLOS conditions, as Δ_i captures the excess path length.

For each AP i and time t, we construct a composite feature vector

$$\mathbf{f}_{t,i} = \left[x_{t,i}^{\text{rss}},\ x_{t,i}^{\text{rtt}},\ m_{t,i}^{\text{rss}},\ m_{t,i}^{\text{rtt}},\ \boldsymbol{\phi}_{t,i}^{\text{stat}},\ \hat{x}_i,\ \hat{y}_i\right], \tag{4}$$

where $m_{t,i}^{\text{rss}}$ and $m_{t,i}^{\text{rtt}}$ are binary validity masks indicating whether the RSS and RTT measurements from AP i at time t are valid or missing/invalid. These masks prevent placeholder values (e.g., extremely low RSS or saturated RTT) from biasing training and allow the network to down-weight unreliable observations. $\boldsymbol{\phi}_{t,i}^{\text{stat}}$ denotes rolling statistical descriptors (mean, median, standard deviation, skewness, and kurtosis) computed over a causal sliding window of past RSS/RTT samples, and $(\hat{x}_i, \hat{y}_i)$ are normalized AP coordinates which is set to zero when it is unavailable.

Stacking features across APs yields

$$\mathbf{F}_t = \begin{bmatrix} \mathbf{f}_{t,1}^\top \\ \vdots \\ \mathbf{f}_{t,N}^\top \end{bmatrix} \in \mathbb{R}^{N \times F}, \tag{5}$$

and a spatio-temporal input tensor over T scans:

$$\mathbf{X}_{1:T} = \{\mathbf{F}_1, \ldots, \mathbf{F}_T\} \in \mathbb{R}^{T \times N \times F}. \tag{6}$$

Given the causal input $\mathbf{X}_{1:t}$, the model f_θ predicts

$$\hat{\mathbf{y}}_t = f_\theta(\mathbf{X}_{1:t}) = [\hat{x}_t, \hat{y}_t]^\top, \tag{7}$$

where $\hat{\mathbf{y}}_t$ depends only on observations up to time t.

The parameters θ are optimized by minimizing the total positioning loss

$$\mathcal{L}(\theta) = \frac{1}{N_s} \sum_{i=1}^{N_s} \left[\phi_\beta\left(\hat{\mathbf{y}}_i^{(n)} - \mathbf{y}_i^{(n)}\right) + w_E \|\hat{\mathbf{y}}_i - \mathbf{y}_i\|_2\right] + \mathcal{R}(\theta), \tag{8}$$

where N_s is the total number of training samples (temporal windows), $\phi_\beta(\cdot)$ denotes the Huber loss with threshold β, w_E weights the Euclidean positioning error, $(\cdot)^{(n)}$ represents normalized coordinates, and $\mathcal{R}(\theta)$ aggregates regularization terms for the adaptive graph and node gating.

The Huber loss [61] is defined as

$$\phi_\beta(a) = \begin{cases} \frac{1}{2}a^2, & \text{if } |a| \le \beta, \\ \beta(|a| - \frac{1}{2}\beta), & \text{otherwise,} \end{cases} \tag{9}$$

where $a = y - \hat{y}$ denotes the residual error and β is the transition parameter between the quadratic and linear regions. This formulation behaves like an ℓ_2 loss for small residuals and an ℓ_1 loss for large residuals,providing robustness against outlier in RSS and RTT measurements.

The optimal network parameters are obtained as

$$\theta^* = \arg\min_\theta \mathcal{L}(\theta), \tag{10}$$

subject to the causality constraint in (7). This formulation enables joint learning of spatial correlations among APs, temporal consistency in RSS/RTT measurements, and a stable mapping from features to user position. The learned mapping $f_\theta(\cdot)$ forms the basis of the proposed Graph–Temporal Convolutional Network (GTCN) described next.

3.2. Proposed Indoor Positioning Method

Inspired by recent GTCN-based CSI localization frameworks [35], we extend the graph temporal paradigm to WiFi RSS and RTT fingerprinting using only commodity Wi-Fi measurements. The proposed model consists of: (i) a GCN to capture spatial correlations among APs; and (ii) a TCN with causal, dilated convolutions to capture temporal dynamics. Their integration yields a unified, causal, and computationally efficient spatio-temporal architecture.

3.2.1. GCN: Spatial Correlation Modeling

We represent the AP layout as a weighted undirected graph $G = (V, E)$, where each node corresponds to an AP and edges encode spatial proximity. The adjacency matrix $\mathbf{A}$ is defined as

$$A_{ij} = A_{ji} = \begin{cases} 0, & i = j, \\ \dfrac{1}{d_{ij}}, & i \ne j, \end{cases} \tag{11}$$

where d_{ij} is the Euclidean distance between the i-th and j-th APs. This construction emphasizes closer AP pairs while suppressing weak long-range couplings.

To ensure numerical stability and include self-connections, we define

$$\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}, \qquad \tilde{\mathbf{D}} = \text{diag}\left(\sum_j \tilde{A}_{ij}\right), \tag{12}$$

and use the symmetrically normalized adjacency $\tilde{\mathbf{D}}^{-\frac{1}{2}}\tilde{\mathbf{A}}\tilde{\mathbf{D}}^{-\frac{1}{2}}$. Given an input feature matrix $\mathbf{X} \in \mathbb{R}^{N \times F_{\text{in}}}$ at a single time step (e.g., one row of $\mathbf{F}_t$ in (6)), a GCN layer produces

$$\mathbf{Y} = \sigma\left(\tilde{\mathbf{D}}^{-\frac{1}{2}}\tilde{\mathbf{A}}\tilde{\mathbf{D}}^{-\frac{1}{2}}\mathbf{X}\mathbf{W}\right), \tag{13}$$

where $\mathbf{W} \in \mathbb{R}^{F_{\text{in}} \times F_{\text{out}}}$ is a learnable weight matrix and $\sigma(\cdot)$ is a non-linear activation (ReLU). Stacking multiple GCN layers allows each AP to aggregate information from multi-hop neighbors, yielding spatially smoothed and topology-aware representations.

In the proposed framework, the GCN is applied to per-scan features across APs, producing spatially enriched node embeddings that serve as input to the temporal modeling stage.

3.2.2. TCN: Causal Sequence Modeling

To capture the temporal evolution of WiFi RSS and RTT fingerprints, we employ a TCN with dilated causal convolutions [62]. This design ensures that the output at time t depends only on $\{\mathbf{x}_\tau\}_{\tau \leq t}$, which is essential for real-time tracking.

For an input feature sequence $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T]^\top$, a dilated causal convolutional layer with kernel size k and dilation factor d computes the output sequence $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_T]^\top$ as

$$\mathbf{y}_t = \sum_{i=0}^{k-1} \mathbf{W}_i\, \mathbf{x}_{t-d\cdot i} + \mathbf{b}, \tag{14}$$

where $\mathbf{W}_i$ are learnable filters, $\mathbf{b}$ is a bias term, and indices with $t - d \cdot i < 1$ are handled via causal padding. Equivalently, this can be written compactly as

$$\mathbf{Y} = f(\mathbf{W} *_d \mathbf{X} + \mathbf{b}), \tag{15}$$

where $*_d$ denotes the dilated causal convolution.

Residual connections are incorporated to stabilize training and mitigate gradient degradation:

$$\mathbf{O} = \sigma(\mathbf{X} + F(\mathbf{X})), \tag{16}$$

where $F(\mathbf{X})$ denotes the sequence of operations inside a residual block (two dilated causal convolutions with normalization, activation, and dropout), and a 1×1 convolution is used on the skip path when channel dimensions differ.

By stacking residual blocks with dilation factors $d \in \{1, 2, 4\}$, the receptive field grows exponentially, enabling the TCN to capture both short-term variations and long-range dependencies in RSS/RTT sequences without violating causality.

3.2.3. Hybrid GTCN Model

In the proposed hybrid GTCN model, the GCN extracts spatial relationships among APs, while the TCN captures temporal dependencies in the sequence of WiFi RSS and RTT fingerprints. Together, they form the hybrid GTCN, which jointly learns spatial–temporal correlations for robust indoor positioning. At each time step, preprocessed WiFi RSS and RTT fingerprints are passed through stacked GCN layers that operate on the AP adjacency graph, as shown in Figure 1. These layers aggregate neighborhood information and generate topology-aware node embeddings. A weighted node pooling module then fuses the node features into a compact global representation of the wireless environment for that scan. The sequence of pooled embeddings across consecutive time windows is fed into a stack of dilated causal convolutions (the TCN block) to model temporal continuity and motion patterns, as shown in Figure 2. The final regression head maps the temporal feature vector to a 2-D position estimate $\hat{\mathbf{y}}_t = [\hat{x}_t, \hat{y}_t]^\top$.

The overall end-to-end process is illustrated in Figure 3. The pipeline begins with data collection from multiple APs, extraction of statistical WiFi RSS and RTT features, and formation of the spatial graph using AP geometry. These data feed the GCN layers (for spatial embedding) and the TCN layers (for temporal modeling) during offline phase. The model outputs the estimated position, which is compared against the ground-truth coordinates to compute the positioning error during online phase. The network parameters are updated according to the loss function in (8), combining Huber and Euclidean terms.
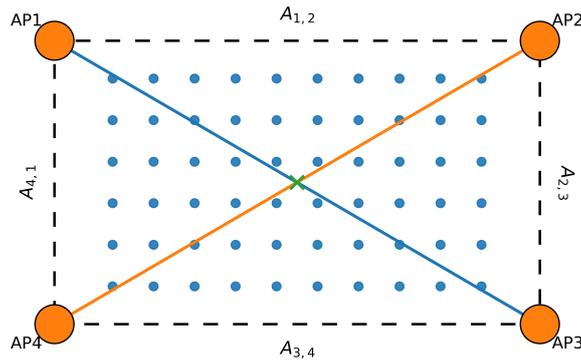
Figure 1. Graph representation of the AP topology. Nodes correspond to APs, and edge weights are defined as the inverse of the Euclidean distance between APs. This spatial graph underlies the GCN-based feature propagation in (13).
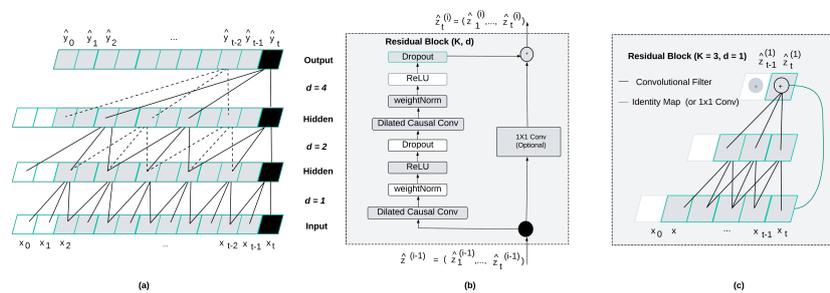

Figure 2. Temporal Convolutional Network (TCN) with dilated causal convolutions and residual connections. Dilation factors increase across layers, enlarging the receptive field while ensuring that each prediction depends only on current and past inputs.[62] (a) TCN layer structure (b) Residual block with dilated convolutions (c) Residual block with identity mapping.
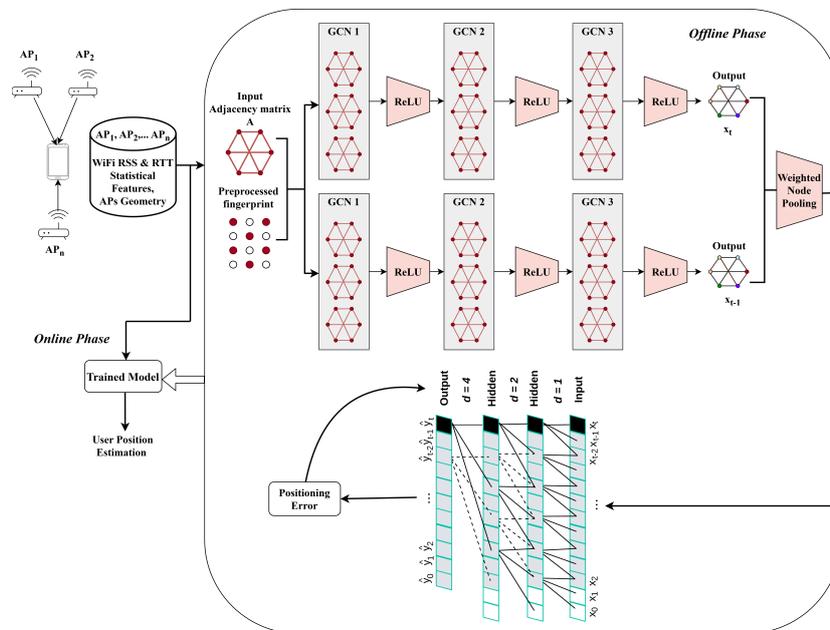

Figure 3. Overall architecture of the proposed hybrid GTCN model for Wi-Fi indoor positioning using RSS and RTT features. The GCN layers capture spatial dependencies among APs via the adjacency matrix, while the TCN layers exploit temporal dynamics through dilated causal convolutions. Weighted node pooling and residual connections ensure efficient feature fusion, leading to robust position estimation under multipath and NLOS conditions.

Algorithm 1: Proposed hybrid GTCN-Based Wi-Fi Indoor Positioning Method

---

Input: Raw Wi-Fi RSS and RTT fingerprint sequences; optional AP coordinates
Output: Estimated positions $\hat{\mathbf{y}} = [\hat{x}, \hat{y}]^{\top}$

1 **Initialization:** Initialize model parameters.
2 **Offline Phase:**

Step 1.   **Preprocessing:**

- Load raw Wi-Fi RSS and RTT scans for each environment.
- Compute per-AP statistics and normalize WiFi RSS and RTT measurements.
- Augment inputs with causal rolling-window statistics (mean, median, std, skewness, kurtosis)
- Encode AP coordinates (z-scored (x, y)) as node features if available.
- Group scans by reference point (RP) into sequences of length T.

Step 2.   **Graph Construction:**

- Build base adjacency $\mathbf{A}_{\text{base}}$ from inverse inter-AP distances.
- Learn symmetric residual adjacency $\mathbf{R}$ and mixing weight α.
- Form normalized adaptive graph:

$$\hat{\mathbf{A}} = \mathbf{D}^{-1/2}\big(\mathbf{A}_{\text{base}} + \alpha\mathbf{R} + \mathbf{I}\big)\mathbf{D}^{-1/2}.$$

Step 3.   **Spatial Encoding (GCN + Node Gating):**

- For each time step and AP, compute a gating weight via a lightweight MLP.
- Apply the gates and propagate features using residual GCN layers over $\hat{\mathbf{A}}$.

Step 4.   **Temporal Modeling (Causal TCN):**

- Aggregate gated node features across APs at each time step to obtain a compact sequence of embeddings.
- Feed this sequence into stacked dilated causal TCN blocks to model temporal dependencies without future leakage.

Step 5.   **Localization and Optimization:**

- Map the final TCN representation to normalized coordinates $\hat{\mathbf{y}}^{(n)} = [\hat{x}^{(n)}, \hat{y}^{(n)}]^{\top}$.
- Compute the Huber loss on normalized coordinates and the Euclidean error on denormalized coordinates.
- Update all parameters using AdamW with gradient clipping.

Online Phase:

- Apply the same preprocessing (using training statistics) to incoming RSS/RTT scans.
- Form input sequences of length T and pass them through the trained GTCN.
- For each sequence (or time step), output the estimated position $\hat{\mathbf{y}} = [\hat{x}, \hat{y}]^{\top}$.

---

4. Experimental Dataset and Results

4.1. Data Collection and Testbeds

The proposed GTCN model is evaluated using the publicly available Wi-Fi RSS/RTT dataset introduced by [46]. The dataset contains measurements from four representative indoor environments: *Lecture Theatre*, *Office*, *Corridor*, and *Building Floor*, each exhibiting different LOS and NLOS characteristics.

All measurements were collected with an LG G8X ThinQ smartphone equipped with IEEE 802.11mc-compliant hardware and a set of commercial APs. For each reference point (RP), RSS (dBm) and RTT (mm) were recorded simultaneously from all visible APs. Each testbed is discretized using a uniform 0.6×0.6 m grid, and the corresponding ground-truth coordinates are provided in the dataset.

At every RP, multiple Wi-Fi scans were performed to capture temporal variations: 60 scans per RP in the Lecture Theatre, Office, and Corridor, and 120 scans per RP in the larger Building Floor

environment. Data collection includes both LOS and obstructed NLOS links, producing a diverse set of propagation conditions suitable for benchmarking spatio temporal learning models.

Table 1 summarizes the four environments. The Reference Points (RPs) and Testing Points (TPs) are spatially disjoint in all cases, ensuring that the model is evaluated on unseen locations rather than memorized grid points. Figures 4–7 illustrate the spatial layouts and the disjoint training/testing partitions for all environments.

Table 1. Overview of the Wi-Fi RSS/RTT fingerprinting datasets used for evaluation.

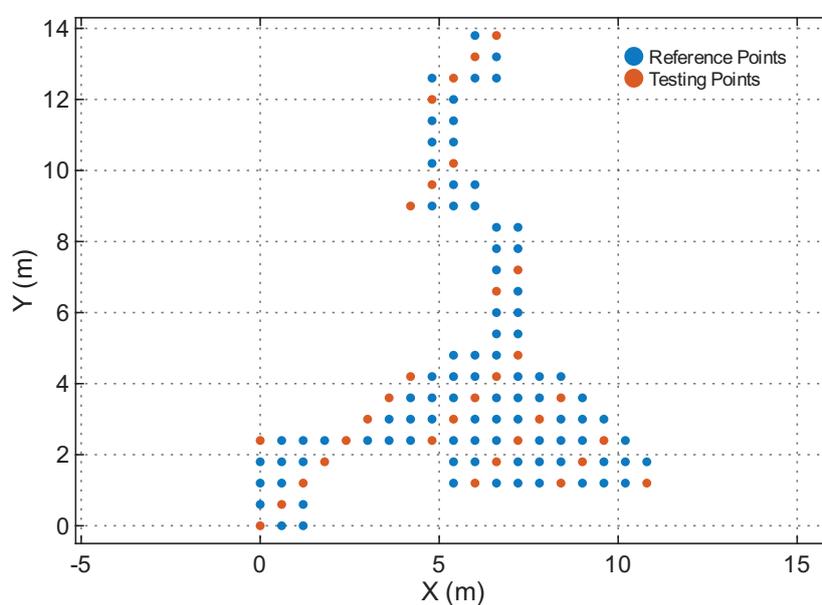| Environment	Area (m^2)	LOS Type	RPs/TPs (Train/Test)	Samples (Train/Test)
Lecture Theatre	15×14.5	LOS	88 / 32	5280 / 1920
Office	18×5.5	Mixed	81 / 27	4860 / 1620
Corridor	35×6	NLOS	85 / 29	5100 / 1740
Building Floor	92×15	Mixed	483 / 159	57960 / 19080

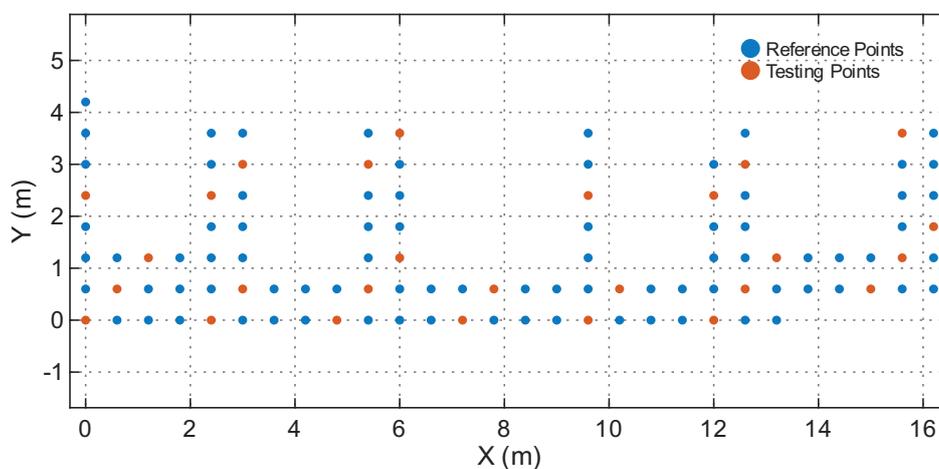**Figure 4.** Reference and testing points in the Lecture Theatre.

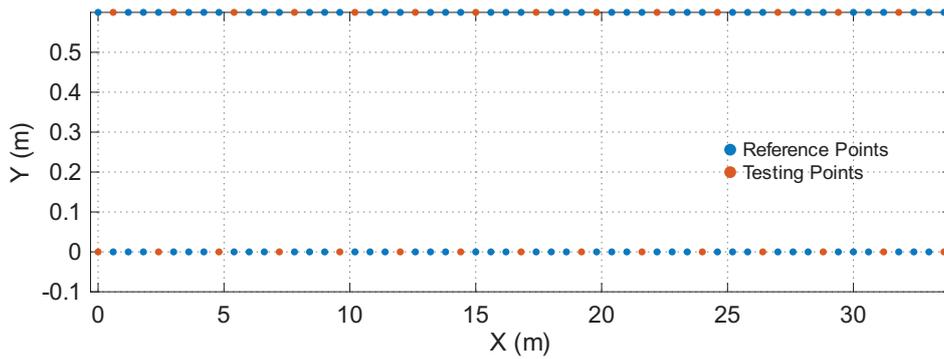**Figure 5.** Reference and testing points in the Office.

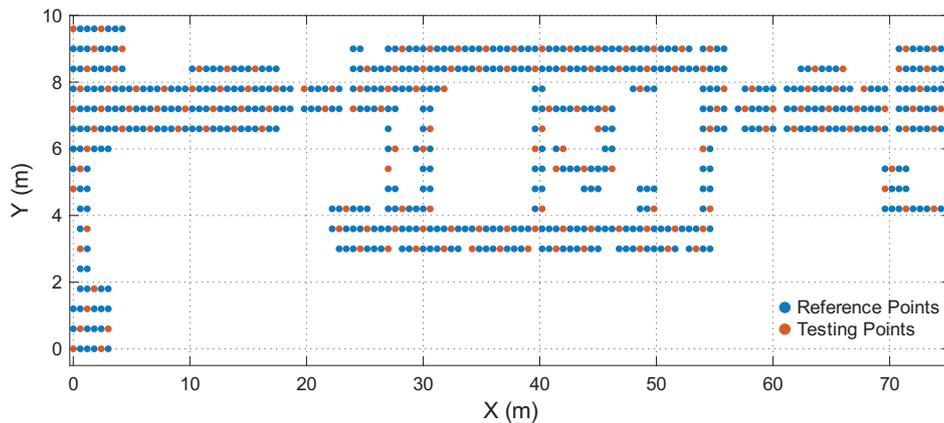**Figure 6.** Reference and testing points in the Corridor.

Figure 7. Reference and testing points in the Building floor.

The positioning performance is evaluated using the Root Mean Square Error (RMSE) of the 2D position estimates and the empirical Cumulative Distribution Function (CDF) of the localization error.

Let $\mathbf{y}_i = [x_i, y_i]^\top$ and $\hat{\mathbf{y}}_i = [\hat{x}_i, \hat{y}_i]^\top$ denote the true and estimated coordinates of the i-th test sample, respectively. The instantaneous Euclidean localization error is given by

$$d_i = \|\mathbf{y}_i - \hat{\mathbf{y}}_i\|_2. \tag{17}$$

The overall positioning accuracy is evaluated using the RMSE metric:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right]}, \tag{18}$$

where N is the total number of test samples. A lower RMSE indicates higher average positioning accuracy.

To provide a statistical view of localization performance, the empirical cumulative distribution function (CDF) is also reported. The CDF curve illustrates the probability that the positioning error does not exceed a given threshold, offering an interpretable measure of the model's reliability across test environments.

4.2. Model Parameters

The proposed GTCN model is implemented with the hyperparameter configurations summarized in Table 2. The architecture comprises three GCN layers for spatial feature extraction and three TCN blocks with causal, dilated convolutions for temporal modeling. This combination efficiently captures both inter-AP spatial relationships and time-varying signal dynamics while maintaining computational efficiency suitable for embedded deployment. A hidden feature dimension of 64 was selected as a balanced trade-off between model expressiveness and complexity, yielding stable training behavior

and avoiding overfitting on smaller datasets. The causal TCN blocks use dilation rates $(1, 2, 4)$ to exponentially expand the receptive field without increasing kernel size or depth. The model is trained using the AdamW optimizer with a learning rate of 1×10^{-3} and a weight decay of 1×10^{-4} to stabilize convergence and mitigate overfitting. Huber and Euclidean losses are combined with weights $(w_H, w_E) = (1.0, 0.5)$ to ensure robustness to outliers in RSS and RTT measurements while maintaining smooth gradient behavior.

Table 2. Hyperparameter settings for the proposed GTCN model.

Parameter	Value
Optimizer	AdamW
Learning rate	1×10^{-3}
Batch size	64
Training epochs	1500
GCN layers	3
TCN blocks (causal)	3
Convolution kernel size	3
Dilation rates	(1, 2, 4, 8)
Hidden dimension	64
Dropout rate	0.05
Weight decay	1×10^{-4}
Huber loss parameter β	1.0
Loss weights (w_H, w_E)	(1.0, 0.5)

Figure 8 illustrates the effect of training epochs on the positioning accuracy of the proposed GTCN model. The average distance error gradually decreases as the number of epochs increases, reflecting steady optimization of network parameters. Beyond approximately 1000 epochs, the model converges with a stable distance error of around 0.55 m, indicating that the proposed framework effectively learns spatial–temporal dependencies without overfitting. This convergence behavior demonstrates the reliability and robustness of the GTCN model across different indoor environments.
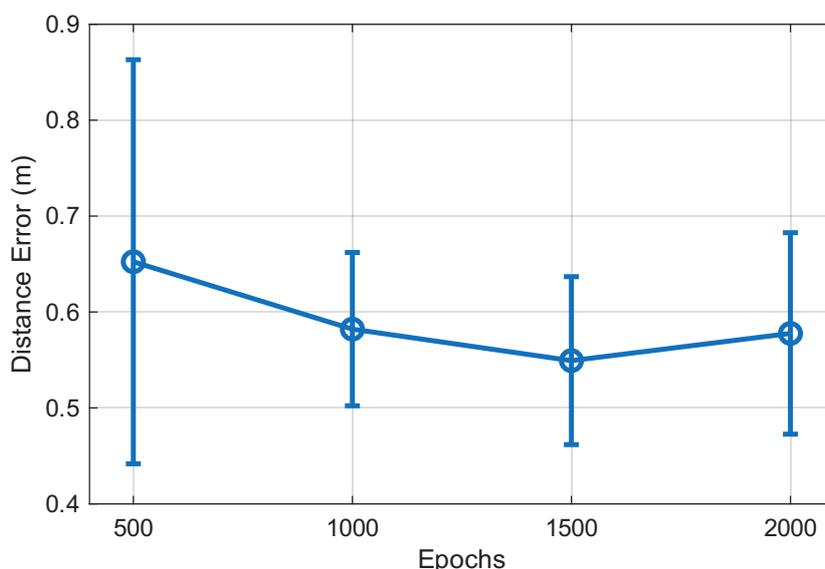

Figure 8. Effect of training epochs on the positioning error of the proposed GTCN model.

4.3. Positioning Performance

4.3.1. Comparison of RSS, RTT, and Hybrid RSS–RTT

We first evaluate the contributions of different input modalities: RSS-only, RTT-only, and combined RSS+RTT fingerprints which is denoted as both. Table 3 reports the positioning RMSE across four

environments and compares the proposed method with conventional fingerprinting and learning-based approaches. Figure 9 shows the corresponding CDFs of positioning error per environment, while left figure of Figure 11 illustrates the aggregated CDF across all testbeds.

Across all testbeds, RTT-only consistently outperforms RSS-only, confirming the advantage of geometric time-of-flight information under both LOS and NLOS conditions. The proposed causal GTCN further improves upon these baselines, particularly when fusing RSS and RTT. The hybrid configuration achieves sub-meter accuracy in the Office, Corridor, and Building environments and remains competitive in the Lecture Theatre. These results demonstrate that the propoaws GTCN successfully leverages complementary hybrid WiFi RSS and RTT characteristics and captures the underlying spatio–temporal structure of Wi-Fi fingerprints.

Table 3. Positioning RMSE (m) comparison across environments using different feature combinations.

| Method	Lecture Theatre	Office	Corridor	Building
JMT-SDAE [63]	0.716	0.857	0.705	1.032
RS-stacking [64]	0.724	0.824	0.672	0.967
NWEC [65]	0.663	0.781	0.599	0.965
RSS–RTT Fingerprinting [46]	0.612	0.729	0.612	0.989
RTT Fingerprinting [46]	0.559	0.718	0.704	0.988
RSS Fingerprinting [46]	2.356	1.423	1.315	1.730
Trilateration [46]	1.176	1.073	412.257*	7.503
Dynamic Model [46]	0.570	0.698	0.569	0.950
Proposed (RSS-only)	2.170	1.203	1.230	1.579
Proposed (RTT-only)	**0.449**	0.649	0.663	0.731
Proposed (Both)	0.501	**0.560**	**0.529**	**0.660**
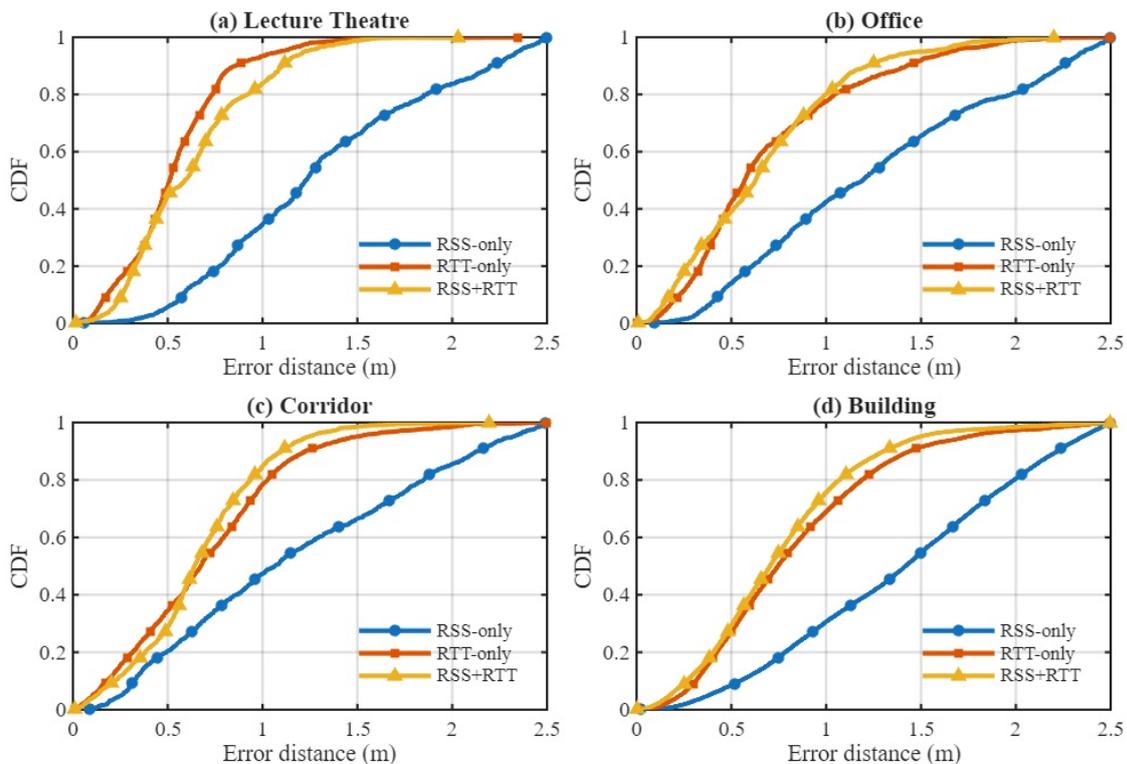
Figure 9. CDF of positioning error for RSS-only, RTT-only, and hybrid RSS+RTT across four environments using the proposed hybrid GTCN model.

4.3.2. Comparison of GCN, TCN, and Hybrid GTCN Model

Next, we compare three model architectures trained under identical conditions: a GCN, TCN, and the proposed GTCN. Figures 10 and right figure of Figure 11 depict the environment-wise and aggregated CDFs of positioning error, respectively. For consistency, all models in the Lecture Theatre were evaluated using their best RTT-based configurations.The hybrid GTCN consistently outperforms the standalone GCN and TCN across most test environments. In compact and predominantly LOS conditions such as the Lecture Theatre and Office, the GTCN exhibits a steeper CDF curve in the low-error region, reflecting enhanced sub-meter reliability and faster convergence of positioning accuracy. In the Corridor, where propagation is largely NLOS with strong multipath reflections, the GTCN surpasses the TCN by effectively exploiting graph-based spatial context among APs, confirming the benefit of incorporating AP topology into temporal modeling.
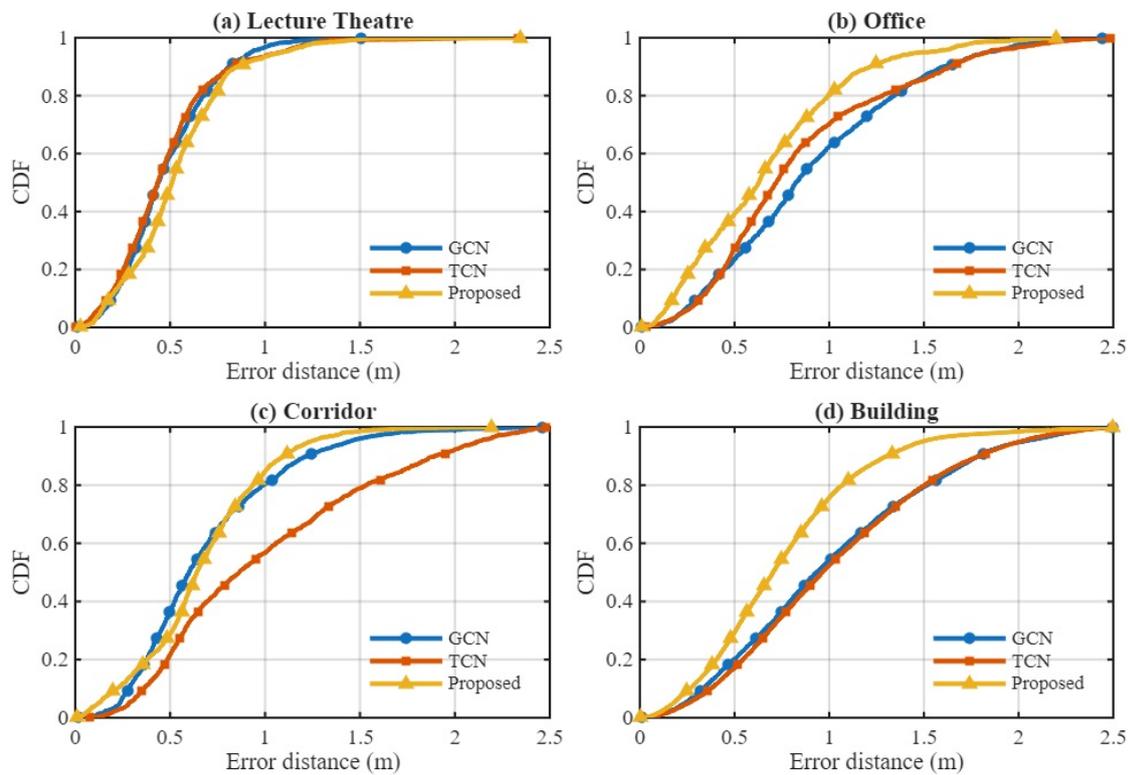

Figure 10. Per-environment CDF comparison of GCN, TCN, and the proposed hybrid GTCN model using RSS and RTT features.
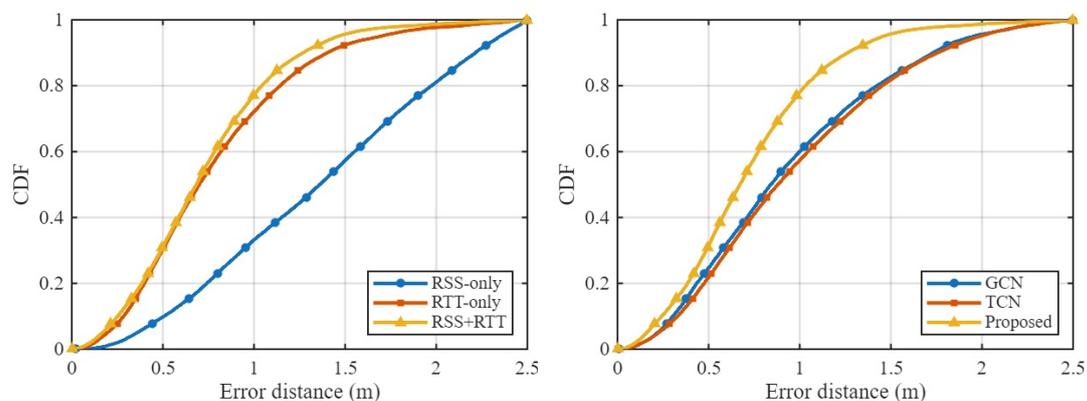

Figure 11. Aggregated CDF of positioning error based on input features and models across all four environments.

Table 4. Positioning performance comparison across environments and model variants.

Environment	Model	Feature Mode	RMSE (m)	Epochs
Lecture Theatre	GCN	RTT	0.426	1500
Lecture Theatre	TCN	RTT	**0.410**	1500
Lecture Theatre	Proposed	RTT	0.449	1500
Office	GCN	RTT	0.622	1500
Office	TCN	RTT	0.704	1500
Office	Proposed	Both	**0.560**	1500
Corridor	GCN	RTT	0.586	1500
Corridor	TCN	RTT	0.773	1500
Corridor	Proposed	Both	**0.529**	1500
Building	GCN	RTT	0.877	1500
Building	TCN	RTT	0.757	1500
Building	Proposed	Both	**0.660**	1500

The Building Floor testbed, the largest and most heterogeneous environment (92 m × 15 m, mixed LOS–NLOS), poses the most significant challenge due to its rich structural complexity and varying AP visibility. Here, the proposed GTCN model yields the highest probability of errors below 1 m, demonstrating strong scalability and generalization in large-scale deployments. Overall, the results show that RTT-based features dominate in pure LOS conditions(Lecture Theatre) because of their direct geometric range correlation, while the combined RSS and RTT features within the GTCN offers superior robustness in mixed and NLOS environments (Office, Corridor, Building). These findings confirm that jointly modeling spatial AP relationships and causal temporal dynamics is crucial for robust fingerprinting-based indoor positioning under diverse propagation conditions.

4.3.3. Impact of Number of APs

We further analyze the sensitivity of the proposed framework to the number of deployed APs. For each environment, the best-performing configuration (feature mode and model) is evaluated as the AP count varies, starting from a minimum of three. This lower bound is chosen because at least three APs are required to perform 2D indoor positioning via geometric trilateration, ensuring sufficient spatial diversity for reliable localization.

Tables 5–8 summarize the positioning RMSE, and Figure 12 presents the corresponding CDFs of positioning error. The results demonstrate that increasing AP density systematically enhances positioning accuracy, though the degree of improvement depends on the environmental characteristics. In the Lecture Theatre (compact LOS environment), the performance saturates beyond four APs, suggesting that a moderate deployment already provides sufficient geometric diversity for accurate time-of-flight estimation. Similarly, in the Office (mixed LOS–NLOS scenario), adding up to five APs significantly improves robustness, with the combination of RSS and RTT feature mode achieving 0.56 m RMSE and stable sub-meter performance thereafter.

Table 5. Effect of AP density on positioning performance in the Lecture Theatre.

# APs	RSS-only	RTT-only	RSS+RTT
3	2.337	1.929	1.364
4	2.542	**0.436**	0.480
5	2.170	0.449	**0.501**

Table 6. Effect of AP density on positioning performance in the Office.

# APs	RSS-only	RTT-only	RSS+RTT
3	1.612	0.701	0.829
4	1.417	0.713	0.687
5	1.203	0.649	**0.560**

Table 7. Effect of AP density on positioning performance in the Corridor.

# APs	RSS-only	RTT-only	RSS+RTT
3	1.749	**0.553**	0.631
4	1.230	0.663	**0.529**

In the Corridor (elongated NLOS environment), where multipath and reflection effects are dominant, the addition of even a single AP noticeably improves the probability of sub-meter accuracy by providing redundant spatial information to the GCN component. The most pronounced effect is observed in the Building Floor (large-scale mixed LOS–NLOS layout). As the number of APs increases from 3 to 13, the RTT-only RMSE decreases from over 9 m to below 1 m, and the combination of RSS and RTT configuration achieves 0.66 m RMSE. This strong monotonic gain highlights the scalability of the proposed framework in large and structurally complex indoor environments.

Overall, the proposed hybrid GTCN model effectively exploits additional APs when available, while maintaining robust performance under sparse deployments. These findings indicate that while dense AP coverage benefits large, heterogeneous spaces, compact environments can achieve comparable accuracy with only a few strategically placed APs when leveraging the joint spatial–temporal learning capability of the proposed model.
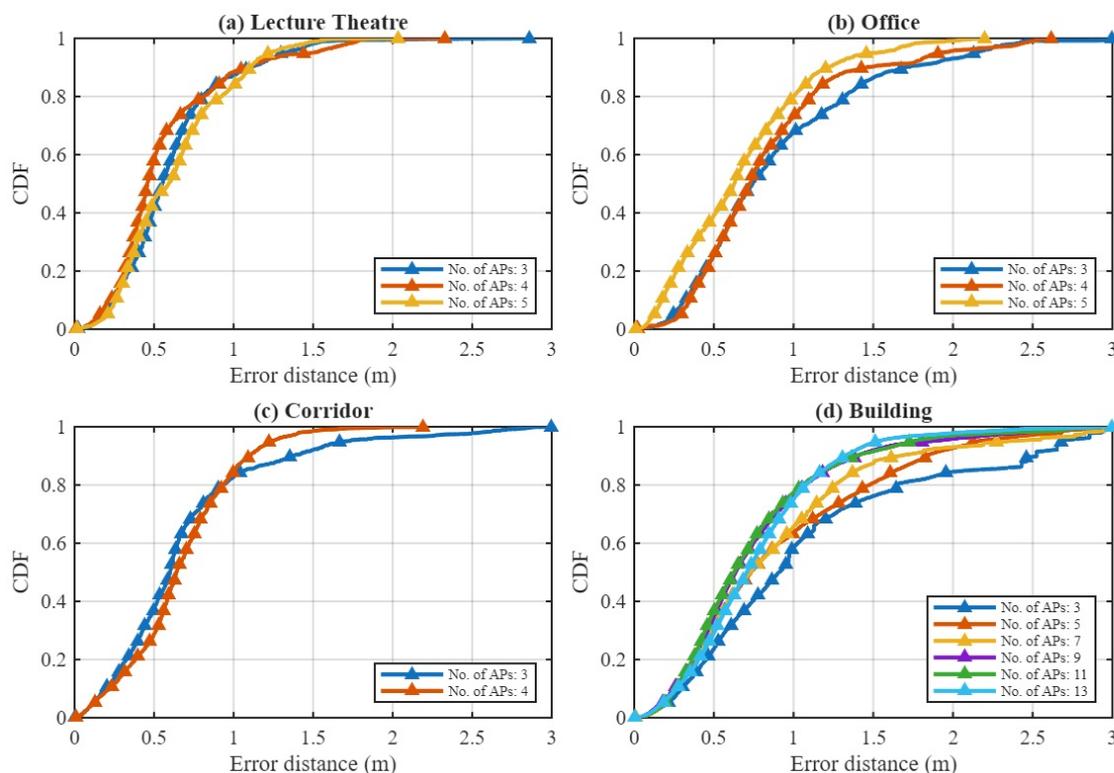

Figure 12. CDF of positioning error for varying numbers of APs across the four environments.

Table 8. Effect of AP density on positioning performance in the Building Floor.

# APs	RSS-only	RTT-only	RSS+RTT
3	9.693	9.687	9.716
5	3.082	2.393	2.376
7	2.258	1.644	1.586
9	1.689	0.951	0.803
11	1.538	0.825	0.725
13	1.579	**0.731**	**0.660**

4.3.4. Model Complexity

We assess the computational complexity of the GCN, TCN, and hybrid GTCN models under their best-performing configurations. Table 9 reports the total number of trainable parameters and floating-point operations (FLOPs) in the causal setting. Trainable parameters represent the total number of learnable weights and biases within a model. This value indicates the storage and memory footprint during both training and inference, directly reflecting the model size. FLOPs (Floating-Point Operations) quantify the total arithmetic operations (multiplications and additions) required to process one input sequence, providing an estimate of the computational workload and inference latency.

Table 9. Model complexity for the best configuration per environment and architecture (causal setting).

Environment	Model	Feature mode	#Params	FLOPs (M)
Building	GCN	RSS+RTT	21,102	53.18
Building	TCN	RTT-only	81,006	26.67
Building	Proposed GTCN	RSS+RTT	93,934	67.01
Corridor	GCN	RTT-only	19,221	7.72
Corridor	TCN	RTT-only	80,853	11.58
Corridor	Proposed GTCN	RSS+RTT	93,781	17.78
Lecture Theatre	GCN	RTT-only	19,230	9.52
Lecture Theatre	TCN	RTT-only	80,862	11.83
Lecture Theatre	Proposed GTCN	RTT-only	90,910	17.86
Office	GCN	RTT-only	19,230	9.52
Office	TCN	RTT-only	80,862	11.83
Office	Proposed GTCN	RSS+RTT	93,790	19.59

Among the compared architectures, the GCN exhibits the lowest complexity and is suitable for highly resource-constrained devices, though it lacks temporal modeling. The TCN effectively captures temporal dependencies but does not encode explicit AP topology. The proposed hybrid GTCN introduces only a modest computational overhead relative to TCN while consistently achieving the best positioning accuracy. With fewer than 10^5 trainable parameters and tens of MFLOPs per sequence, the GTCN offers an optimal balance between accuracy and efficiency, supporting real-time deployment on embedded and edge devices.

5. Conclusions

This study proposed a lightweight GTCN for hybrid Wi-Fi RSS and RTT-based indoor positioning. By jointly modeling spatial dependencies among APs using graph convolutions and temporal correlations across scans through causal dilated TCNs, the framework achieved sub-meter accuracy across diverse indoor environments, including LOS, NLOS, and mixed conditions. Experimental results showed up to a 25–30% reduction in RMSE compared to existing deep and ensemble-based methods, demonstrating the model's robustness and scalability with varying AP densities. The RTT component was found to be particularly beneficial in NLOS-dominant scenarios, effectively mitigating

multipath bias, while the hybrid RSS–RTT fusion provided more stable estimates in mixed conditions envioronments.

Moreover, increasing AP density systematically enhanced positioning accuracy, most notably in large, structurally complex layouts whereas compact LOS environments achieved comparable accuracy with only a few well-placed APs. With fewer than 10^5 trainable parameters and tens of MFLOPs per inference, the proposed GTCN model enables real-time, edge-compatible deployment. Future work will extend this framework to multi-floor and dynamic environments, including cross-environment positioning and adaptive graph updates to further generalize this approach for indoor positioning system.

References

1. Enge, P.K. The Global Positioning System: Signals, measurements, and performance. *International Journal of Wireless Information Networks* **1994**, *1*, 83–105. https://doi.org/10.1007/BF02106512.
2. Rusli, M.E.; Ali, M.; Jamil, N.; Din, M.M. An Improved Indoor Positioning Algorithm Based on RSSI-Trilateration Technique for Internet of Things (IOT). In Proceedings of the 2016 International Conference on Computer and Communication Engineering (ICCCE), 2016, pp. 72–77. https://doi.org/10.1109/ICCCE.2016.28.
3. Fedosova, I.; Pronina, O.; Piatykop, O. Exploring Wi-Fi-based Indoor Positioning Techniques for Mobile App. In Proceedings of the 2021 IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo), 2021, pp. 109–114. https://doi.org/10.1109/UkrMiCo52950.2021.9716608.
4. Tarrío, P.; Bernardos, A.M.; Casar, J.R. Weighted Least Squares Techniques for Improved Received Signal Strength Based Localization. *Sensors* **2011**, *11*, 8569–8592. https://doi.org/10.3390/s110908569.
5. Bahl, P.; Padmanabhan, V. RADAR: an in-building RF-based user location and tracking system. In Proceedings of the Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064), 2000, Vol. 2, pp. 775–784 vol.2. https://doi.org/10.1109/INFCOM.2000.832252.
6. Numan, P.E.; Park, H.; Laoudias, C.; Horsmanheimo, S.; Kim, S. DNN-based Indoor Fingerprinting Localization with WiFi FTM. In Proceedings of the 2022 23rd IEEE International Conference on Mobile Data Management (MDM), 2022, pp. 367–371. https://doi.org/10.1109/MDM55031.2022.00082.
7. Li, J.; Park, J.; Cui, S.; Dong, J.; Rana, L.; Hwang, J. An Indoor Positioning Method Using High Probability RSSI. In Proceedings of the 2024 9th International Conference on Computer and Communication Systems (ICCCS), 2024, pp. 483–488. https://doi.org/10.1109/ICCCS61882.2024.10603198.
8. Si, M.; Wang, Y.; Xu, S.; Sun, M.; Cao, H. A Wi-Fi FTM-Based Indoor Positioning Method with LOS/NLOS Identification. *Applied Sciences* **2020**, *10*. https://doi.org/10.3390/app10030956.
9. Hashem, O.; Youssef, M.; Harras, K.A. WiNar: RTT-based Sub-meter Indoor Localization using Commercial Devices. In Proceedings of the 2020 IEEE International Conference on Pervasive Computing and Communications (PerCom), 2020, pp. 1–10. https://doi.org/10.1109/PerCom45495.2020.9127363.
10. Yu, Y.; Chen, R.; Chen, L.; Xu, S.; Li, W.; Wu, Y.; Zhou, H. Precise 3-D Indoor Localization Based on Wi-Fi FTM and Built-In Sensors. *IEEE Internet of Things Journal* **2020**, *7*, 11753–11765. https://doi.org/10.1109/JIOT.2020.2999626.
11. Cao, H.; Wang, Y.; Bi, J.; Xu, S.; Si, M.; Qi, H. Indoor Positioning Method Using WiFi RTT Based on LOS Identification and Range Calibration. *ISPRS International Journal of Geo-Information* **2020**, *9*. https://doi.org/10.3390/ijgi9110627.

12. Si, M.; Wang, Y.; Seow, C.K.; Cao, H.; Liu, H.; Huang, L. An Adaptive Weighted Wi-Fi FTM-Based Positioning Method in an NLOS Environment. *IEEE Sensors Journal* **2022**, *22*, 472–480. https://doi.org/10.1109/JSEN.2021.3124275.

13. Cypriani, M.; Lassabe, F.; Canalda, P.; Spies, F. Open Wireless Positioning System: A Wi-Fi-Based Indoor Positioning System. In Proceedings of the 2009 IEEE 70th Vehicular Technology Conference Fall, 2009, pp. 1–5. https://doi.org/10.1109/VETECF.2009.5378966.

14. Liu, H.; Darabi, H.; Banerjee, P.; Liu, J. Survey of Wireless Indoor Positioning Techniques and Systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* **2007**, *37*, 1067–1080. https://doi.org/10.1109/TSMCC.2007.905750.

15. Zafari, F.; Gkelias, A.; Leung, K.K. A Survey of Indoor Localization Systems and Technologies. *IEEE Communications Surveys & Tutorials* **2019**, *21*, 2568–2599. https://doi.org/10.1109/COMST.2019.2911558.

16. Yeh, S.C.; Wang, C.H.; Hsieh, C.H.; Chiou, Y.S.; Cheng, T.P. Cost-Effective Fitting Model for Indoor Positioning Systems Based on Bluetooth Low Energy. *Sensors* **2022**, *22*. https://doi.org/10.3390/s22166007.

17. Khyam, M.O.; Noor-A-Rahim, M.; Li, X.; Ritz, C.; Guan, Y.L.; Ge, S.S. Design of Chirp Waveforms for Multiple-Access Ultrasonic Indoor Positioning. *IEEE Sensors Journal* **2018**, *18*, 6375–6390. https://doi.org/10.1109/JSEN.2018.2846481.

18. Zhang, D.; Yang, L.T.; Chen, M.; Zhao, S.; Guo, M.; Zhang, Y. Real-Time Locating Systems Using Active RFID for Internet of Things. *IEEE Systems Journal* **2016**, *10*, 1226–1235. https://doi.org/10.1109/JSYST.2014.2346625.

19. Werner, M.; Kessel, M.; Marouane, C. Indoor positioning using smartphone camera. In Proceedings of the 2011 International Conference on Indoor Positioning and Indoor Navigation, 2011, pp. 1–6. https://doi.org/10.1109/IPIN.2011.6071954.

20. Uradzinski, M.; Guo, H.; Liu, X.; Yu, M. Advanced Indoor Positioning Using Zigbee Wireless Technology. *Wirel. Pers. Commun.* **2017**, *97*, 6509–6518. https://doi.org/10.1007/s11277-017-4852-5.

21. Wang, J.; Park, J. A Novel Fingerprint Localization Algorithm Based on Modified Channel State Information Using Kalman Filter. *Journal of Electrical Engineering and Technology* **2020**, *15*, 1811–1819. Publisher Copyright: © 2020, The Korean Institute of Electrical Engineers., https://doi.org/10.1007/s42835-020-00441-6.

22. Brunato, M.; Battiti, R. Statistical learning theory for location fingerprinting in wireless LANs. *Computer Networks* **2005**, *47*, 825–845. https://doi.org/https://doi.org/10.1016/j.comnet.2004.09.004.

23. Li, J.L.; Dong, J.B.; Hwang, J.G.; Rana, L.; Park, J.G. Indoor Positioning Method Based on CGAN using Modified Fingerprint DB with IEEE 802.11 RSSI Measurements. In Proceedings of the Proceedings of the Symposium of the Korean Institute of Communications and Information Sciences (KICS), Jeju, South Korea, June 2024; pp. 131–132.

24. Cui, S.; Dong, J.; Hwang, J.G.; Rana, L.; Li, J.; Park, J.G. An enhanced WiFi indoor positioning method based on SNGAN. In Proceedings of the Proceedings of the 36th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2023), 2023, pp. 1684–1692.

25. Li, J.; Dong, J.; Hwang, J.G.; Rana, L.; Ryu, H.; Park, J.G. Deep Learning-Based Wi-Fi Signal Fingerprinting Indoor Positioning Technology. In Proceedings of the Proceedings of the 37th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2024), Baltimore, Maryland, USA, September 2024; pp. 1354–1362. https://doi.org/10.33012/2024.19868.

26. Maranò, S.; Gifford, W.M.; Wymeersch, H.; Win, M.Z. NLOS identification and mitigation for localization based on UWB experimental data. *IEEE Journal on Selected Areas in Communications* **2010**, *28*, 1026–1035. https://doi.org/10.1109/JSAC.2010.100907.

27. Rana, L.; Dong, J.; Cui, S.; Li, J.; Hwang, J.; Park, J. Indoor Positioning using DNN and RF Method Fingerprinting-based on Calibrated Wi-Fi RTT. In Proceedings of the 2023 13th International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2023, pp. 1–6. https://doi.org/10.1109/IPIN57070.2023.10332507.

28. He, S.; Chan, S.H.G. Wi-Fi Fingerprint-Based Indoor Positioning: Recent Advances and Comparisons. *IEEE Communications Surveys & Tutorials* **2016**, *18*, 466–490. https://doi.org/10.1109/COMST.2015.2464084.

29. IEEE Standard for Information Technology–Telecommunications and Information Exchange between Systems Local and Metropolitan Area Networks–Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 2: Enhanced Throughput for Operation in License-exempt Bands above 45 GHz. *IEEE Std 802.11ay-2021 (Amendment to IEEE Std 802.11-2020 as amendment by IEEE Std 802.11ax-2021)* **2021**, pp. 1–768. https://doi.org/10.1109/IEEESTD.2021.9502046.

30. Segev, J.; Berger, C. Next Generation Wi-Fi Positioning: An Overview of IEEE 802.11az. https://www.computer.org/csdl/video-library/video/1UXPYs1JCrS, 2025. Accessed on April 24, 2025.

31. Rana, L.; Park, J.G. An Enhanced Indoor Positioning Method Based on RTT and RSS Measurements Under LOS/NLOS Environment. *IEEE Sensors Journal* **2024**, *24*, 31417–31430. https://doi.org/10.1109/JSEN.2024.3441247.

32. IEEE Standard for Information Technology–Telecommunications and Information Exchange between Systems - Local and Metropolitan Area Networks–Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. *IEEE Std 802.11-2020 (Revision of IEEE Std 802.11-2016)* **2021**, pp. 1–4379. https://doi.org/10.1109/IEEESTD.2021.9363693.

33. Shoudha, S.N.; Helwa, S.; Van Marter, J.P.; Torlak, M.; Al-Dhahir, N. WiFi 5GHz CSI-Based Single-AP Localization With Centimeter-Level Median Error. *IEEE Access* **2023**, *11*, 112470–112482. https://doi.org/10.1109/ACCESS.2023.3323634.

34. Hwang, J.G.; Dong, J.; Rana, L.; Li, J.L.; Park, J.G. An Enhanced Access Points Placement Method for Indoor Positioning Using PPO. In Proceedings of the Proceedings of the Symposium of the Korean Institute of Communications and Information Sciences (KICS), Jeju, South Korea, June 2024.

35. Liu, X.; Wu, R.; Zhang, H.; Chen, Z.; Liu, Y.; Qiu, T. Graph Temporal Convolutional Network-Based WiFi Indoor Localization Using Fine-Grained CSI Fingerprint. *IEEE Sensors Journal* **2025**, *25*, 9019–9033. https://doi.org/10.1109/JSEN.2025.3525624.

36. Jurdi, R.; Chen, H.; Zhu, Y.; Loong Ng, B.; Dawar, N.; Zhang, C.; Han, J.K.H. WhereArtThou: A WiFi-RTT-Based Indoor Positioning System. *IEEE Access* **2024**, *12*, 41084–41101. https://doi.org/10.1109/ACCESS.2024.3377237.

37. Dong, J.; Rana, L.; Li, J.; Hwang, J.; Park, J. Indoor Positioning Using Wi-Fi RTT Based on Stacked Ensemble Model. In Proceedings of the 2024 9th International Conference on Computer and Communication Systems (ICCCS), 2024, pp. 1021–1026. https://doi.org/10.1109/ICCCS61882.2024.10602843.

38. Cao, H.; Wang, Y.; Bi, J.; Xu, S.; Si, M.; Qi, H. Indoor Positioning Method Using WiFi RTT Based on LOS Identification and Range Calibration. *ISPRS International Journal of Geo-Information* **2020**, *9*. https://doi.org/10.3390/ijgi9110627.

39. Cao, H.; Wang, Y.; Bi, J.; Xu, S.; Qi, H.; Si, M.; Yao, G. WiFi RTT Indoor Positioning Method Based on Gaussian Process Regression for Harsh Environments. *IEEE Access* **2020**, *8*, 215777–215786. https://doi.org/10.1109/ACCESS.2020.3041773.

40. Cao, H.; Wang, Y.; Bi, J.; Zhang, Y.; Yao, G.; Feng, Y.; Si, M. LOS compensation and trusted NLOS recognition assisted WiFi RTT indoor positioning algorithm. *Expert Systems with Applications* **2024**, *243*, 122867. https://doi.org/https://doi.org/10.1016/j.eswa.2023.122867.

41. Park, K.M.; Lee, B.h.; Lee, E.; Kim, S.C.; Choi, J. GPS-Aided Automatic Site Survey Method for WiFi RTT-Based Positioning. *IEEE Transactions on Vehicular Technology* **2023**, *72*, 13120–13129. https://doi.org/10.1109/TVT.2023.3275203.

42. Guo, G.; Chen, R.; Ye, F.; Peng, X.; Liu, Z.; Pan, Y. Indoor Smartphone Localization: A Hybrid WiFi RTT-RSS Ranging Approach. *IEEE Access* **2019**, *7*, 176767–176781. https://doi.org/10.1109/ACCESS.2019.2957753.

43. Dong, J.; Rana, L.; Cui, S.; Li, J.; Hwang, J.; Park, J. Investigation on Indoor Positioning by Improved RTT-RSS Fusion Ranging Method. In Proceedings of the 2023 IEEE 6th International Conference on Electronics and Communication Engineering (ICECE), 2023, pp. 49–53. https://doi.org/10.1109/ICECE59822.2023.10462313.

44. Gonzalez Díaz, N.; Zola, E.; Martin-Escalona, I. Assessing the Impact of Coupling RTT and RSSI Measurements in Fingerprinting Wi-Fi Indoor Positioning. In Proceedings of the Proceedings of the Int'l ACM Conference on Modeling Analysis and Simulation of Wireless and Mobile Systems, New York, NY, USA, 2023; MSWiM '23, p. 19–26. https://doi.org/10.1145/3616388.3617528.

45. Rizk, H.; Elmogy, A.; Yamaguchi, H. A Robust and Accurate Indoor Localization Using Learning-Based Fusion of Wi-Fi RTT and RSSI. *Sensors* **2022**, *22*. https://doi.org/10.3390/s22072700.

46. Feng, X.; Nguyen, K.A.; Luo, Z. A Wi-Fi RSS-RTT Indoor Positioning Model Based on Dynamic Model Switching Algorithm. *IEEE Journal of Indoor and Seamless Positioning and Navigation* **2024**, *2*, 151–165. https://doi.org/10.1109/JISPIN.2024.3385356.

47. Dong, Y.; Arslan, T.; Yang, Y. Real-Time NLOS/LOS Identification for Smartphone-Based Indoor Positioning Systems Using WiFi RTT and RSS. *IEEE Sensors Journal* **2022**, *22*, 5199–5209. https://doi.org/10.1109/JSEN.2021.3119234.

48. Guo, G.; Chen, R.; Ye, F.; Liu, Z.; Xu, S.; Huang, L.; Li, Z.; Qian, L. A Robust Integration Platform of Wi-Fi RTT, RSS Signal, and MEMS-IMU for Locating Commercial Smartphone Indoors. *IEEE Internet of Things Journal* **2022**, *9*, 16322–16331. https://doi.org/10.1109/JIOT.2022.3150958.

49. Guo, G.; Chen, R.; Niu, X.; Yan, K.; Xu, S.; Chen, L. Factor Graph Framework for Smartphone Indoor Localization: Integrating Data-Driven PDR and Wi-Fi RTT/RSS Ranging. *IEEE Sensors Journal* **2023**, *23*, 12346–12354. https://doi.org/10.1109/JSEN.2023.3267121.

50. Jung, H.; Ko, S.W.; Kim, S. Toward Cooperative Localization With Implicit Connectivity: Graph Neural Network Approach. *IEEE Wireless Communications Letters* **2025**, *14*, 3184–3188. https://doi.org/10.1109/LWC.2025.3588339.

51. Sun, M.; Wang, Y.; Wang, Q.; Chen, G.; Li, Z. Smartphone-Based WiFi RTT/RSS/PDR/Map Indoor Positioning System Using Particle Filter. *IEEE Transactions on Instrumentation and Measurement* **2025**, *74*, 1–15. https://doi.org/10.1109/TIM.2024.3509549.

52. Han, K.; Yu, S.M.; Kim, S.L.; Ko, S.W. WiFi Positioning with Mobility-Induced Graphs. In Proceedings of the 2024 IEEE 99th Vehicular Technology Conference (VTC2024-Spring), 2024, pp. 1–5. https://doi.org/10.1109/VTC2024-Spring62846.2024.10683593.

53. Yan, W.; Jin, D.; Lin, Z.; Yin, F. Graph Neural Network for Large-Scale Network Localization. In Proceedings of the ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2021, pp. 5250–5254. https://doi.org/10.1109/ICASSP39728.2021.9414520.

54. Liang, M.; Meyer, F. Neural Enhanced Belief Propagation for Cooperative Localization. In Proceedings of the 2021 IEEE Statistical Signal Processing Workshop (SSP), 2021, pp. 326–330. https://doi.org/10.1109/SSP49050.2021.9513853.

55. Bilge, A.; Ergen, E.; Soner, B.; Çöleri, S. Scalable Wi-Fi RSS-Based Indoor Localization via Automatic Vision-Assisted Calibration. 09 2025, pp. 1–6. https://doi.org/10.1109/ICAT66432.2025.11189271.

56. Al-qaness, M.; Dahou, A.; Trouba, N.; Elsayed Abd Elaziz, M.; Helmi, A. TCN-Inception: Temporal Convolutional Network and Inception modules for sensor-based Human Activity Recognition. *Future Generation Computer Systems* **2024**. https://doi.org/10.1016/j.future.2024.06.016.

57. Khor, C.; Ahmad, N. BLE-based Indoor Localization with Temporal Convolutional Network. *Journal of Engineering Research* **2025**. https://doi.org/10.1016/j.jer.2025.05.003.

58. Li, Z.; Jiang, T.; Yu, J.; Ding, X.; Zhong, Y.; Liu, Y. A Lightweight Mobile Temporal Convolution Network for Multi-Location Human Activity Recognition based on Wi-Fi. In Proceedings of the 2021 IEEE/CIC International Conference on Communications in China (ICCC Workshops), 2021, pp. 143–148. https://doi.org/10.1109/ICCCWorkshops52231.2021.9538870.

59. Zhu, A.; Tang, Z.; Wang, Z.; Zhou, Y.; Chen, S.; Hu, F.; Li, Y. Wi-ATCN: Attentional Temporal Convolutional Network for Human Action Prediction Using WiFi Channel State Information. *IEEE Journal of Selected Topics in Signal Processing* **2022**, *16*, 804–816. https://doi.org/10.1109/JSTSP.2022.3163858.

60. Lin, C.Y.; Lin, C.Y.; Liu, Y.T.; Chen, Y.W.; Shih, T.K. WiFi-TCN: Temporal Convolution for Human Interaction Recognition Based on WiFi Signal. *IEEE Access* **2024**, *12*, 126970–126982. https://doi.org/10.1109/ACCESS.2024.3428550.

61. Huber, P.J. Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics* **1964**, *35*, 73–101. https://doi.org/10.1214/aoms/1177703732.

62. Bai, S.; Kolter, J.Z.; Koltun, V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. *arXiv:1803.01271* **2018**.

63. Zhang, H.; Hu, B.; Xu, S.; Chen, B.; Li, M.; Jiang, B. Feature Fusion Using Stacked Denoising Auto-Encoder and GBDT for Wi-Fi Fingerprint-Based Indoor Positioning. *IEEE Access* **2020**, *8*, 114741–114751. https://doi.org/10.1109/ACCESS.2020.3004039.

64. JunLin, G.; Xin, Z.; HuaDeng, W.; Lan, Y. WiFi fingerprint positioning method based on fusion of autoencoder and stacking mode. In Proceedings of the 2020 International Conference on Culture-oriented Science & Technology (ICCST), 2020, pp. 356–361. https://doi.org/10.1109/ICCST50977.2020.00075.

65. Roy, P.; Chowdhury, C.; Kundu, M.; Ghosh, D.; Bandyopadhyay, S. Novel weighted ensemble classifier for smartphone based indoor localization. *Expert Systems with Applications* **2021**, *164*, 113758. https://doi.org/https://doi.org/10.1016/j.eswa.2020.113758.