

Article

A UWB Underground Mine Positioning Algorithm Based on Graph Neural Network

Zhongyang Yu ¹, Qinghua Liu ¹  and Yong Qian ^{2,*}

¹ Information and Communication College, Guilin University of Electronic and Technology, Guilin 541010, China; zyyu@mails.guet.edu.cn (Z.Y.)

² Guangxi Baise Industrial Investment Development Group Co., Ltd., Baise 533000, China

* Correspondence: 15078676173@163.com; Tel.: +86-15078676173

Abstract

Ultra-Wideband (UWB) positioning is a promising technique for underground mine localization, but its accuracy is strongly affected by complex tunnel topology, multipath propagation, and non-line-of-sight (NLOS) ranging errors. To address these challenges, this paper proposes a three-dimensional UWB positioning algorithm based on a Multi-head Attention Feature Fusion Graph Neural Network (MAFF-GNN). The localization problem is formulated as a graph-based node position regression task, where anchors and tags are represented as nodes and UWB ranging links are represented as edges. The proposed model integrates graph message passing, multi-head attention-based feature fusion, global skip connections, and geometry-constrained regularization to learn reliability-aware spatial representations from noisy ranging measurements. A physics-guided simulated underground mine environment is constructed by considering tunnel geometry, wall roughness, coal dust concentration, humidity attenuation, and controlled NLOS conditions. Three mine-like corridor topologies are generated, with 10,500 localization samples in total. Experimental results under controlled simulation conditions show that MAFF-GNN achieves an RMSE of 0.323 ± 0.093 m, an MAE of 0.274 ± 0.024 m, and a P90 error of 0.480 ± 0.038 m. Compared with weighted least squares and support vector regression, the proposed method reduces RMSE by 74.16% and 40.30%, respectively. Robustness tests under different simulated NLOS ratios further indicate that the proposed graph-attention framework maintains a gradual error-growth trend as NLOS severity increases. These results indicate the potential of attention-enhanced graph learning for UWB localization in a challenging mine-like environment.

Keywords: ultra-wideband technology; graph neural network; underground mine positioning; multi-head attention mechanism; non-line-of-sight error



Academic Editor: Jose María Alvarez Rodríguez

Received: 6 May 2026

Revised: 25 June 2026

Accepted: 28 June 2026

Published: 15 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The operational environment within underground mines is complex and hazardous; therefore, personnel positioning systems are crucial for safeguarding miners' lives and enhancing the efficiency of emergency rescue operations. Traditional mine positioning technologies, such as RFID, WiFi, and ZigBee, often fail to meet safety production requirements due to limitations such as low positioning accuracy and weak interference resistance. Ultra-Wideband (UWB) technology, characterized by centimeter-level ranging precision, strong penetration capabilities, and low power consumption, is considered an ideal solution for underground mine positioning. Compared with RFID-, WiFi-, and ZigBee-based systems,

UWB can provide higher ranging resolution and stronger multipath discriminability, which has motivated its use in underground localization and wireless sensing applications [1,2]. However, the underground mine environment differs significantly from conventional indoor settings. Mine tunnels are narrow and winding, containing numerous obstacles such as metal supports and mining equipment, resulting in an extremely complex signal propagation environment. In particular, ranging errors caused by Non-Line-of-Sight (NLOS) propagation represent the primary factor limiting UWB positioning accuracy. Therefore, achieving high-precision positioning in underground mine-like environments, especially under severe NLOS and multipath conditions, remains a challenging problem.

Recent learning-based studies have attempted to identify NLOS or multipath channels from CIR or related features and then compensate for ranging errors using machine-learning or deep-learning models [2–5]. These studies indicate the value of data-driven reliability estimation for UWB localization, but most of them focus on channel classification or ranging correction rather than directly modeling the anchor-tag topology as a graph. In recent years, significant progress has been made in ultra-wideband-based positioning methods. Su et al. [6] leveraged high-precision UWB Time-of-Arrival (TOA) measurements to establish theoretical precision limits using the Cramer-Rao Lower Bound (CRLB), extensively comparing its performance against machine-learning-enhanced Bluetooth RSSI across diverse obstructed environments and body postures to demonstrate superior detection confidence; nevertheless, their work was limited to one-dimensional proximity classification and did not provide continuous coordinates for multi-dimensional positioning. Addressing multi-agent spatial configurations, Alizad and Nobahari [7] introduced a master node equipped with a rotating UWB tag and integrated triangular geometric constraints with velocity magnitude laws into a decentralized estimator to simultaneously determine relative positions and accelerometer biases, yet their reliance on 2D planar motion and ideal communication assumptions limits applicability in complex 3D maneuvers. For harsh environments, Liang et al. [8] extracted high-dimensional Channel Impulse Response (CIR) features to train an AdaBoost ensemble classifier for channel identification, which was subsequently coupled with an Extended Kalman Filter (EKF) to dynamically correct range measurements in coal mine tunnels; however, a limitation of this study is that the validation relied on simulated data, which may not fully capture the stochasticity of real-world subterranean propagation. Finally, Liu and Li [9] developed a binary hypothesis test based on Signal-to-Noise Ratio (SNR) for initial NLOS detection, supplemented by a sliding window fusion strategy to recover missed detections before applying a Constrained Total Least Squares (CTLS) solver; however, this method relies on empirical thresholds, which risks data over-filtering and positioning outages if not precisely tuned. Moreover, these methods usually require manually designed thresholds, reference labels, or environment-specific parameters, which may limit their adaptability when NLOS conditions and tunnel geometry change.

In the specific context of underground mining environments, Hancke and Silva [1] highlighted that UWB's nanosecond pulses offer superior multipath resolution for tracking, yet their application is constrained by high infrastructure costs due to limited signal coverage. Addressing large stochastic errors, Cheng et al. [10] employed EKF coupled with a distance difference thresholding strategy to preprocess outlier measurements, effectively suppressing random errors and enhancing trajectory smoothness; however, this approach relies heavily on a preset empirical threshold for outlier rejection, which risks misclassifying valid high-dynamic maneuvers as anomalies, thereby leading to tracking lag or distortion. To balance accuracy and computational efficiency, Li and Li [11] leveraged reference labels as geometric benchmarks in the CBORF algorithm to statistically correct systematic offsets.

Even so, this approach is limited by the spatial error correlation between reference and target nodes, risking performance degradation in heterogeneous noise environments.

In recent years, graph neural networks (GNNs) have demonstrated unique advantages in processing non-Euclidean data. Messias et al. [12] pioneered the application of GNNs to positive-negative sample learning problems, effectively addressing classification tasks on graph-structured data through rewiring methods. Orfanoudakis et al. [13] proposed a large-scale coordinated reinforcement learning approach for electric vehicles based on GNNs, demonstrating the scalability of GNNs in handling massively networked systems. Komarovskiy et al. [14] applied spatio-temporal graph convolutional neural networks to traffic signal prediction, achieving high-precision forecasting by capturing spatial topology and temporal dependencies. Recent surveys further show that GNNs are well suited to IoT and wireless sensing systems because they can represent device topology, edge relations, and heterogeneous sensing attributes in a unified framework [15].

More recently, attention-based and graph-based UWB localization methods have started to incorporate geometric information and spatial-temporal relationships into learning-based positioning models [16,17]. Progress has been made in the application of graph neural networks for UWB positioning. He et al. [16] proposed a multi-tag UWB positioning method based on spatio-temporal graph neural networks, pioneering the construction of UWB ranging information into a graph structure. Nodes represent ranging data, while edges denote spatial and temporal relationships. This approach fully demonstrates the advantages of graph structures in capturing geometric constraints. However, as this method is primarily designed for indoor environments, its adaptability to the narrow, elongated tunnel structures and strong multipath effects characteristic of mining environments requires further validation. Nevertheless, existing GNN-based UWB localization studies are mainly designed for general indoor or multi-tag scenarios, and their ability to handle mine-specific propagation factors, such as tunnel roughness, coal dust, humidity, and high NLOS ratios, remains insufficiently investigated.

The central research question of this study is whether graph-structured UWB ranging information, when combined with attention-based reliability weighting and geometry-constrained regularization, can improve three-dimensional localization accuracy under severe simulated NLOS conditions in underground mine-like tunnel environments. We hypothesize that multi-head attention can adaptively assign lower weights to unreliable NLOS-distorted ranging links while preserving useful geometric constraints through graph message passing.

To answer this question, this study proposes a UWB underground mine positioning algorithm based on a Multi-head Attention Feature Fusion GNN (MAFF-GNN). The proposed algorithm incorporates a simulated mine environment model that considers propagation loss factors such as coal dust concentration, air humidity, tunnel roughness, and NLOS interference. The evaluation focuses on whether the proposed graph-attention framework can improve positioning accuracy and robustness under controlled simulated mine-like conditions.

The main contributions of this work are as follows: (1) the UWB localization problem is formulated as a graph-based node position regression task; (2) a multi-head attention feature fusion mechanism is introduced to distinguish reliable ranging links from NLOS-distorted links; (3) a simulated underground mine environment is constructed by considering tunnel geometry, rough-surface scattering, coal dust absorption, humidity attenuation, and controlled NLOS conditions; and (4) the proposed method is evaluated against classical positioning algorithms and representative neural-network baselines.

2. Materials and Methods

2.1. Principles of UWB Indoor Positioning and Ranging

The fundamental principle of UWB positioning systems relies on calculating the distance between nodes by measuring the signal propagation time. Traditional UWB positioning systems primarily employ TOA or Time Difference of Arrival (TDOA) measurement principles [18,19]. The TOA method calculates the distance d by measuring the one-way propagation time t from the transmitter to the receiver. This is expressed as $d = c \cdot t$, where c represents the speed of light. Three-dimensional positioning requires at least four non-coplanar anchor points to form intersecting distance spheres. TDOA constructs a hyperbolic positioning model based on the time difference in signal arrival between different anchors $\Delta d = c \cdot (t_j - t_i)$, determining the target's position through the intersection of multiple hyperboloids.

As illustrated in Figure 1, within a UWB positioning system, the position estimation of a target node is typically based on distance measurements acquired from multiple anchors. Let the tag position be denoted as $\mathbf{p} = [x, y, z]^T$, and the i -th anchor position as $\mathbf{a}_i = [x_i, y_i, z_i]^T$. By measuring the TOA or TDOA of UWB signals, distance observations d_i between the target and each anchor are obtained. Under ideal conditions, the distance measurement model is expressed in Equation (1):

$$d_i = \|\mathbf{p} - \mathbf{a}_i\|_2 = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2} \quad (1)$$

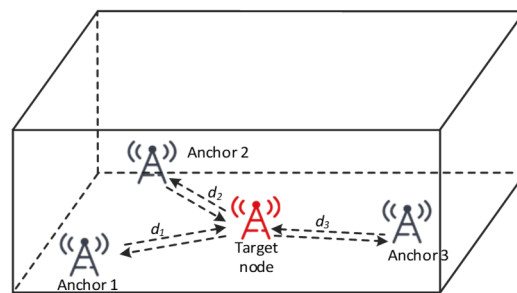


Figure 1. Schematic diagram of UWB trilateration ranging principle.

In practical scenarios, due to environmental noise and system errors, the measured distance is modeled as $\hat{d}_i = d_i + n_i$, where n_i denotes measurement noise. The position of the tag can then be estimated by minimizing the discrepancy between the measured distances and the geometric distances from the tag to all anchors, as shown in Equation (2):

$$\hat{\mathbf{p}} = \arg \min_{\mathbf{p}} \sum_{i=1}^N \left(\hat{d}_i - \|\mathbf{p} - \mathbf{a}_i\|_2 \right)^2 \quad (2)$$

where N denotes the number of anchors, $\hat{d}_i$ represents the measured distance between the tag and the i -th anchor, and $\hat{\mathbf{p}}$ is the estimated tag position. Due to the nonlinearity of Equation (2), direct optimization may incur high computational complexity and can be sensitive to initialization. A commonly used approach is to construct a linear least-squares system by subtracting distance-squared equations. First, the distance-squared equation for the i -th anchor is written as Equation (3):

$$(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2 = \hat{d}_i^2 \quad (3)$$

The first anchor is selected as the reference anchor. By subtracting the distance-squared equation of the i -th anchor from that of the first anchor, the quadratic terms x^2 , y^2 , and z^2

are eliminated. For each anchor with $i \geq 2$, the reference-subtracted equation is written as Equation (4):

$$(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 - [(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2] = \hat{d}_1^2 - \hat{d}_i^2 \quad (4)$$

After expanding and simplifying Equation (4), the linear equation in Equation (5) is obtained:

$$2(x_i - x_1)x + 2(y_i - y_1)y + 2(z_i - z_1)z = \hat{d}_1^2 - \hat{d}_i^2 + (x_1^2 + y_1^2 + z_1^2) - (x_i^2 + y_i^2 + z_i^2) \quad (5)$$

The resulting equations are assembled into the overdetermined linear system in Equation (6):

$$A \cdot p = b \quad (6)$$

where A denotes the design matrix of the overdetermined system:

$$A = \begin{bmatrix} 2(x_2 - x_1) & 2(y_2 - y_1) & 2(z_2 - z_1) \\ 2(x_3 - x_1) & 2(y_3 - y_1) & 2(z_3 - z_1) \\ \vdots & \vdots & \vdots \\ 2(x_N - x_1) & 2(y_N - y_1) & 2(z_N - z_1) \end{bmatrix} \quad (7)$$

and b is the corresponding right-hand-side vector:

$$b = \begin{bmatrix} \hat{d}_1^2 - \hat{d}_2^2 + (x_1^2 + y_1^2 + z_1^2) - (x_2^2 + y_2^2 + z_2^2) \\ \hat{d}_1^2 - \hat{d}_3^2 + (x_1^2 + y_1^2 + z_1^2) - (x_3^2 + y_3^2 + z_3^2) \\ \vdots \\ \hat{d}_1^2 - \hat{d}_N^2 + (x_1^2 + y_1^2 + z_1^2) - (x_N^2 + y_N^2 + z_N^2) \end{bmatrix} \quad (8)$$

Due to measurement noise, no exact solution generally exists for $Ax = b$. The least-squares estimate is given in Equation (9):

$$\hat{p} = (A^T A)^{-1} A^T b \quad (9)$$

In practical applications, TOA necessitates strict clock synchronization between the tag and anchors, as nanosecond-level timing discrepancies can result in meter-level positioning errors. While TDOA relaxes the synchronization requirements for the tag, it necessitates high-precision synchronization among the anchors. However, achieving such synchronization is challenging in complex underground mine environments.

2.2. Challenges of UWB Positioning in Mine Environments

The underground coal mine environment presents significant challenges to UWB signal transmission. These are primarily characterized by signal attenuation resulting from the complex structural features of mine tunnels and adverse environmental factors. Based on the electromagnetic wave propagation model for rectangular mine tunnels [14], this section analyzes the key loss mechanisms used to construct the subsequent simulation-based environmental model. This modeling step is used to parameterize the simulated underground mine environment and does not replace site-specific field calibration.

To facilitate analysis, the tunnel cross-section is modeled as a rectangle with width $2w$ and height $2h$. A coordinate system is established with the origin at the cross-section's center, where the x -, y -, and z -axes correspond to the tunnel width, height, and longitudinal axis, respectively. Figure 2 illustrates the cross-sectional structure of the rectangular tunnel.

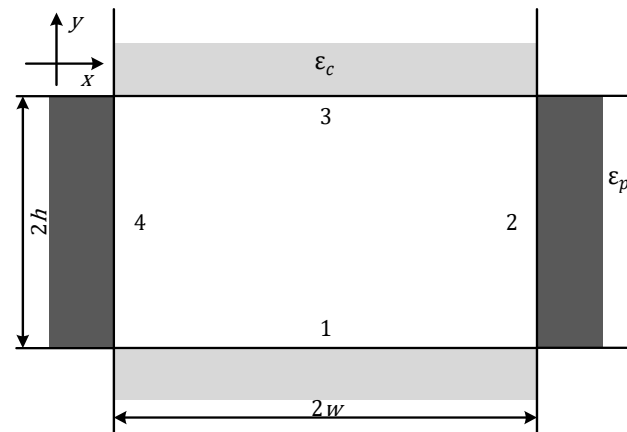


Figure 2. Cross-section of mine roadway electromagnetic propagation model.

As UWB signals undergo multiple reflections against tunnel walls, polarization-dependent reflection attenuation occurs, influenced by the dielectric properties of the coal walls. This attenuation is closely correlated with the signal wavelength λ , propagation distance z , and the relative permittivity of both the coal wall ϵ_p and the tunnel air ϵ_c . The total reflection attenuation model T_{ref} for horizontally and vertically polarised waves can be expressed as Equation (10):

$$T_{\text{ref}} = 4.343\lambda_z^2 \left(\frac{1 + \epsilon_p}{2w\sqrt{\epsilon_p - 1}} + \frac{1 + \epsilon_c}{2h\sqrt{\epsilon_c - 1}} \right) \quad (10)$$

Reflection-induced attenuation is significantly more pronounced in the low-frequency band (<1 GHz). Furthermore, vertically polarized waves exhibit greater loss compared to horizontally polarized waves.

In practical underground mine environments, tunnel walls often exhibit uneven surfaces due to geological formations, coal dust deposits, and equipment abrasion. When electromagnetic waves strike these rough surfaces, a portion of the energy is scattered from the original propagation path, resulting in multidirectional scattering. As illustrated in Figure 3, let $\sigma_{h,i}$ denote the standard deviation of roughness for the i -th wall surface, while $\theta_{a,b}$ represents the tunnel inclination angle [20]. $\sigma_{h,i}$ constitutes the primary determinant of scattering loss, exhibiting a positive correlation with the loss. The scattering loss model T_{sat} attributable to wall roughness is defined as Equation (11) [20]:

$$T_{\text{sat}} = \frac{16.214\pi}{\lambda^2} \left(\frac{1}{2w} (\sigma_{h,1}^2 + \sigma_{h,3}^2) + \frac{1}{2b} (\sigma_{h,2}^2 + \sigma_{h,4}^2) \right) \quad (11)$$

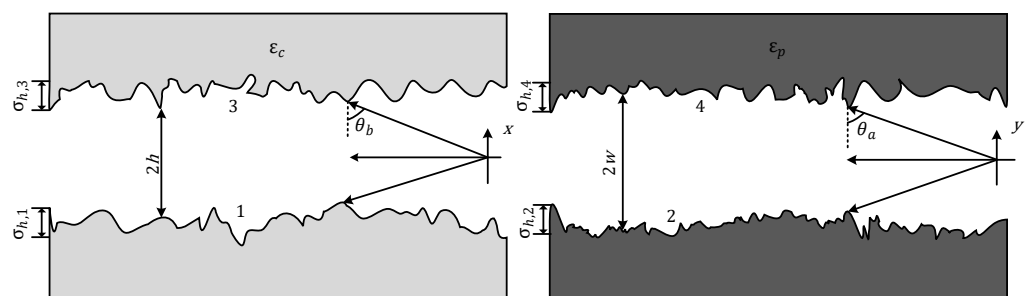


Figure 3. Longitudinal section of mine roadway electromagnetic propagation model.

Due to geological structures and mining-induced pressure, the roof and floor of underground tunnels often exhibit an inclination relative to the horizontal plane. During UWB signal propagation through inclined tunnels, electromagnetic waves undergo multiple asymmetric reflections between the tunnel walls. Each reflection results in energy loss, and the cumulative effect of these reflections induces what is termed inclination-induced attenuation, denoted as T_{tilt} in Equation (12) [21]:

$$T_{\text{tilt}} = \frac{5.791\pi\theta^2 z}{\lambda^2} (2m + 2n) \quad (12)$$

where m and n denote the half-wave mode numbers of the tunnel, respectively, and θ represents the dip angle.

Given that tunnel cross-sectional dimensions typically remain relatively stable, suspended coal dust particles represent a critical variable affecting electromagnetic wave propagation. The total loss attributed to the dust concentration P is quantified in Equation (13) [20,22]:

$$T_{\text{coa}} = 4.343PA \left(\frac{128\pi^5 (AL)^6 \left| \frac{(\varepsilon-1)}{(\varepsilon+2)} \right|^2}{(3\lambda^4)} + \frac{8\pi^2 (AL)^2 \left| \frac{(\varepsilon-1)}{(\varepsilon+2)} \right|}{\lambda} \right) \quad (13)$$

where A denotes the cross-sectional area, L represents the unit length, ε signifies the environmental dielectric constant, and P indicates the dust concentration. Notably, the absorption cross-section ψ_f of coal dust significantly exceeds its scattering cross-section ψ_s . This indicates that during UWB signal propagation, energy attenuation is primarily caused by dielectric absorption effects within coal dust particles, rather than scattering.

In mine tunnel environments, high humidity significantly alters the propagation characteristics of electromagnetic waves. Reported underground mine observations indicate that the relative humidity of underground air can exceed 70% [20]. This high-humidity environment induces attenuation in UWB signal propagation, and the corresponding attenuation model is described in Equation (14):

$$T_{\text{hum}} = 4.343 \times \frac{0.6\pi}{\lambda\rho} \ln \left(\frac{\varepsilon-1}{\varepsilon+1} \right) R \quad (14)$$

where R denotes the relative humidity of the air, and ρ denotes the density of water vapour within the tunnel.

In summary, the total signal loss within the roadway—comprising reflection attenuation, scattering loss, tilt loss, and attenuation due to coal dust and humidity—can be modeled as shown in Equation (15):

$$T_{\text{total}} = T_{\text{coa}} + T_{\text{hum}} + T_{\text{ref}} + T_{\text{sat}} + T_{\text{tilt}} \quad (15)$$

These terms provide the theoretical basis for constructing the simulated underground mine environment used in the subsequent experiments.

3. Experimental Design and the MAFF-GNN Framework

3.1. Multi-Topology Simulated Mine Scenarios and Graph Dataset Construction

The complex environment of underground mines presents significant challenges to the design and implementation of UWB positioning systems. In contrast to conventional indoor scenarios, mine tunnels are characterized by unique spatial structures and signal propagation properties. In this section, a simulation-based experimental environment

is developed to provide a controlled platform for subsequent graph-based localization model validation.

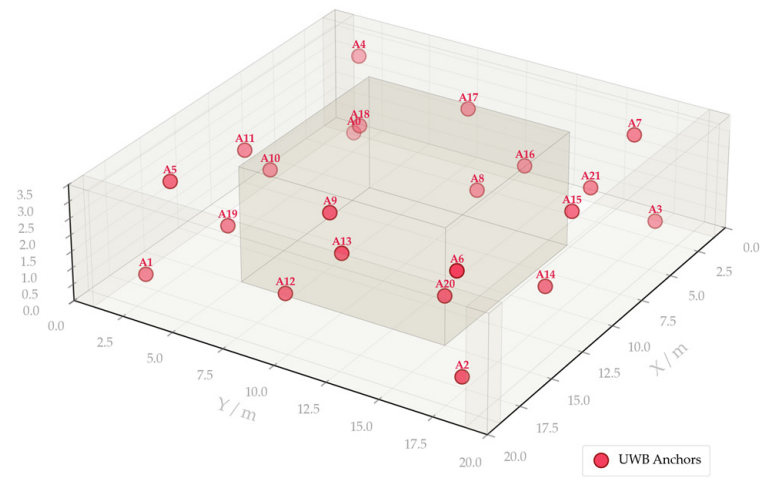
Based on typical roadway structures in underground mines, three simulated mine layouts with the same overall dimensions of $20\text{ m} \times 20\text{ m} \times 3.5\text{ m}$ were established. For clarity, the layouts are denoted as Scene A, Scene B, and Scene C. Scene A represents a ring-corridor layout with a major central obstacle; Scene B represents a cross-corridor layout formed by four obstacle blocks; and Scene C represents a dense multi-cell corridor layout with multiple internal obstacles. All these layouts comply with the Coal Mine Safety Regulation [23]. The three layouts use the same spatial scale and the same number of UWB anchors, but differ in obstacle topology, corridor connectivity, and anchor spatial distribution. Therefore, they provide controlled simulated conditions for evaluating localization performance under different mine-like corridor structures. The main characteristics of the three simulated layouts are summarized in Table 1.

Table 1. Structural characteristics and anchor deployment of the three simulated mine layouts.

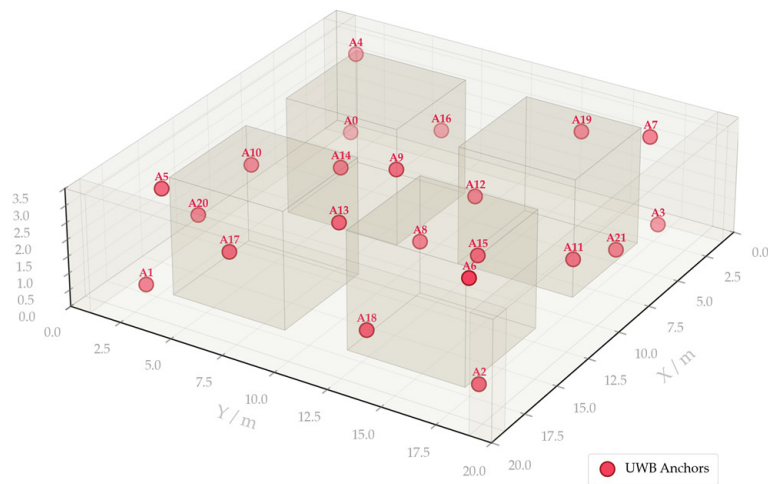
Scene	Layout Type	Simulation Space	Obstacle and Corridor Structure	Anchor Deployment
Scene A	Ring-corridor layout	$20\text{ m} \times 20\text{ m} \times 3.5\text{ m}$	A major central obstacle with a surrounding ring corridor	22 anchors; outer-corner anchors, central-observation anchors, and staggered mid-corridor anchors
Scene B	Cross-corridor layout	$20\text{ m} \times 20\text{ m} \times 3.5\text{ m}$	Four obstacle blocks forming a central cross corridor and peripheral passages	22 anchors; anchors distributed around corner regions, central crossings, and peripheral corridors
Scene C	Dense multi-cell corridor layout	$20\text{ m} \times 20\text{ m} \times 3.5\text{ m}$	Multiple internal obstacle blocks forming connected internal corridors	22 anchors; anchors deployed around outer corridors and internal corridor intersections

After the anchor deployments were fixed, 3500 valid tag sampling points were generated within the accessible corridor regions of each simulated scene. Each sampling point corresponds to one single-tag localization sample, and the tag was connected to the available communicable anchors to construct an anchor-tag ranging graph. Therefore, the complete graph dataset contains 10,500 localization samples from the three simulated mine topologies. Following a scene-stratified 10:2:2 split ratio, each scene contains 2500 training samples, 500 validation samples, and 500 test samples, resulting in 7500 training samples, 1500 validation samples, and 1500 test samples in total. The split was performed at the tag-location level within each scene to avoid overlap of identical tag coordinates across the training, validation, and test subsets.

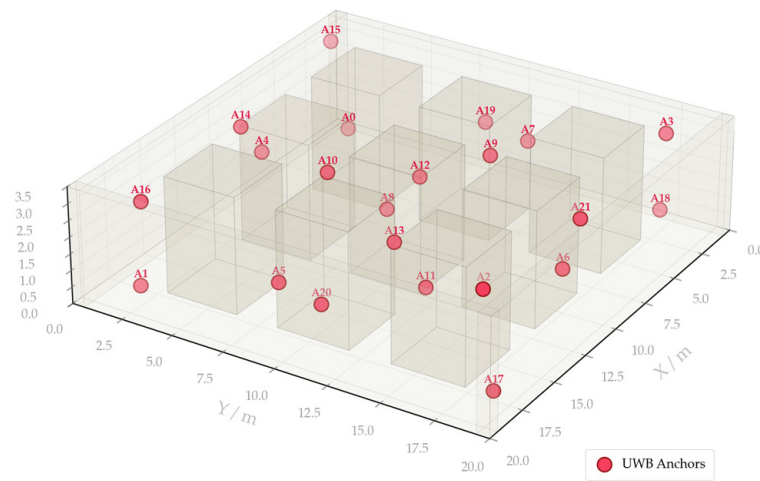
As illustrated in Figure 4, scenario-specific anchor deployments were designed according to the corridor topology and obstacle distribution of each layout. To ensure a fair comparison among different corridor structures, all three layouts employ the same number of UWB anchors. The anchor positions were distributed across corner regions, corridor intersections, and key passage areas with staggered height levels to enhance three-dimensional geometric diversity.



(a)



(b)



(c)

Figure 4. Schematic layouts of the three simulated mine scenarios and their UWB anchor deployments: (a) Scene A; (b) Scene B; and (c) Scene C.

Anchor deployment is designed according to three criteria: corridor coverage, geometric DOP (dilution of precision), and deployment redundancy. The objective function is expressed in Equation (16):

$$J = \alpha \cdot \text{Coverage} + \beta \cdot \text{DOP} + \gamma \cdot \text{Redundancy} \quad (16)$$

where Coverage represents the proportion of valid corridor regions covered by the anchor network, DOP reflects localization geometry quality, and Redundancy measures the availability of alternative ranging links under partial blockage.

After defining the layouts and anchor configurations, mine-specific environmental factors were incorporated into the ranging model. Based on the loss mechanisms described in Section 2.2, the model considers reflection attenuation, surface scattering, tunnel inclination loss, coal dust absorption, humidity attenuation, and shadowing effects. The total path loss and distance-dependent path loss are expressed in Equations (17) and (18), respectively.

$$PL_{total}(d) = PL(d) + T_{ref} + T_{sat} + T_{tilt} + T_{coa} + T_{hum} + X_{\sigma} \quad (17)$$

$$PL(d) = 20 \log_{10} \left(\frac{4\pi f}{c} \right) + 10n \log_{10}(d) \quad (18)$$

Simulation parameters were selected from reported empirical ranges [20,22,24]. The roadway-wall relative permittivity was set to 5.5–8.0, surface roughness to 0.05–0.15 m, tunnel inclination to 0–15°, coal dust concentration to 10–50 mg/m³, and relative humidity to 70–90%. By combining corridor topology, anchor deployment, and environmental attenuation factors, the simulated environments provide the basis for graph dataset construction, model training, and performance evaluation.

3.2. MAFF-GNN Architecture

In this study, Graph Neural Networks (GNNs) are employed for UWB positioning. The positioning problem is formulated as a tag-node position regression task on a graph structure, and the proposed model is termed MAFF-GNN. The core architecture integrates the spatial topology modeling capabilities of GNNs with the feature enhancement provided by multi-head attention mechanisms. Specifically, GNNs aggregate information along anchor-tag ranging links through message passing, while multi-head attention dynamically assigns weights to different ranging features to mitigate NLOS-induced errors in mine-like tunnels.

Following data pre-processing, each UWB positioning sample is converted into a graph $G = (V, E, W, F)$, where V is the set of nodes, E is the set of edges, W is the set of node features, and F is the set of edge features. The node index set $V = \{v_1, v_2, \dots, v_{N+M}\}$ comprises N anchor points and M tags, with each node v_i representing a physical entity within the positioning system. Corresponding to the node set V is the node feature matrix $W \in \mathbb{R}^{(N+M) \times d_v}$, where d_v denotes the node feature dimension, storing attribute information for each node. The feature information for node v_i is stored in the i -th row of matrix W , denoted as $w_i = [\text{type}_i, x_i, y_i, z_i, r_i]^T$, where $\text{type}_i \in \{0, 1\}$ indicates the node type (0 for an anchor, 1 for a tag); (x_i, y_i, z_i) represents the node's three-dimensional coordinates, which are known fixed values for anchors and parameters to be estimated for tags; r_i denotes the distance from the node to the coordinate origin. This design decouples the node's topological information V from its attribute information W , thereby facilitating graph-neural-network processing.

The edge set $E = \{(i, j) | \text{Distance measurement between node } v_i \text{ and node } v_j\}$ represents UWB ranging links. For each single-tag localization sample, edges are constructed between the tag node and all communicable anchor nodes. Corresponding to the edge set

E is the edge feature matrix $F \in R^{|E| \times d_e}$, which records the ranging-related information for each edge.

The k -th edge $e_k = (i, j)$ in the edge set corresponds to the k -th row of the edge feature matrix F . The edge feature vector contains only observable ranging-related information, including the measured UWB distance $\hat{d}_{ij}$, the NLOS-related indicator s_{ij} , and the signal-to-noise ratio snr_{ij} , i.e., $f_k = [\hat{d}_{ij}, s_{ij}, SNR_{ij}]^T$. The true Euclidean distance $d_{ij}^* = \|p_i - a_j\|_2$ is not used as an input feature and is only used for supervised loss calculation and evaluation.

To address NLOS error compensation, spatial relationship modeling under complex corridor topologies, and multipath distortion, MAFF-GNN transforms UWB localization into a tag-node position regression problem.

By leveraging alternating fusion of message passing and multi-head attention [25], this architecture presents a geometry-aware encoder–decoder framework built on MAFF-GNN.

As illustrated in Figure 5, the processing pipeline begins with a Feature Embedding layer, which projects heterogeneous raw node and edge attributes into a high-dimensional latent space to initialize the graph topology. The core Spatial Encoder then extracts deep topological features through an alternating mechanism of message passing and multi-head attention; this involves aggregating neighbor information, performing non-linear feature transformations, and utilizing multi-head attention to capture spatial dependencies from diverse perspectives. Crucially, to mitigate the over-smoothing problem inherent in deep networks, a Global Skip Connection is introduced to fuse the initial input embeddings directly with the high-level features, thereby preserving raw signal fidelity. These fused representations are subsequently decoded by a Hierarchical Decoder, comprising a non-linear MLP Projection Head for feature densification and a linear Regression Head that maps the latent features to the final 3D coordinate prediction.

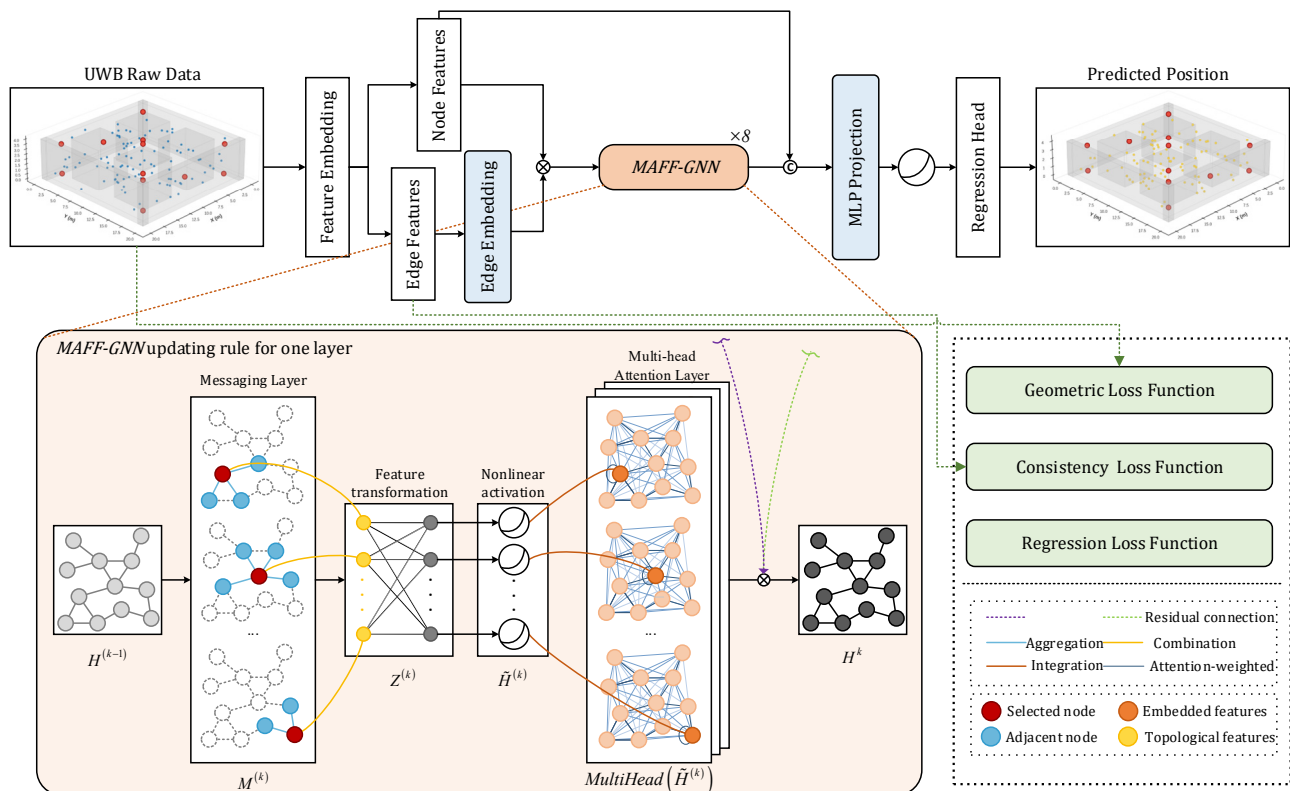


Figure 5. Overall architecture of the proposed MAFF-GNN framework: feature embedding; stacked message-passing and multi-head attention encoder; and hierarchical decoder for 3D coordinate regression with geometry-constrained regularization.

The geometry-constrained regularization term in the final objective is used to guide the predicted coordinates toward consistency with the observed ranging geometry. This mechanism ensures that feature extraction is governed by the physical geometry of the mining environment, thereby preventing the model from relying solely on statistical correlations.

To mitigate multipath effects and NLOS interference in underground mine environments, the multi-head attention feature-fusion module is designed as a reliability-aware spatial encoding unit. As shown in the expanded view of Figure 5, this unit functions as a physics-aware filter that transforms noisy ranging measurements into spatial embeddings. By integrating topological message passing with a multi-head attention mechanism, the unit can dynamically identify and weigh reliable LOS links while suppressing NLOS outliers, thereby ‘encoding’ the complex mine topology into high-fidelity feature representations.

The message passing mechanism transmits messages by updating node features. Let $H^{(0)} \in R^{|N+M| \times d_v}$ denote the input node feature matrix, N the number of anchors, M the number of tags, and d_v the node feature dimension. The initial feature vector $h_i^0 = [type_i, x_i, y_i, z_i, r_i]^T$ for each node i is then aggregated in layer k by first combining messages from all neighbouring nodes:

$$m_i^{(k)} = \sum_{j \in \mathcal{N}(i)} \alpha_{ij}^{(k)} \odot \text{MsgNet}^{(k)} \left([h_i^{(k-1)} \parallel h_j^{(k-1)} \parallel f_k] \right) \quad (19)$$

$\mathcal{N}(i)$ denotes the set of neighbours for node i , $\parallel$ represents vector concatenation, and $\odot$ denotes the Hadamard product. Edge weights $\alpha_{ij}^{(k)}$ are computed via the attention network:

$$\alpha_{ij}^{(k)} = \sigma \left(W_{edge}^{(k)} e_{ij} + b_{edge}^{(k)} \right) \quad (20)$$

where $\sigma(\cdot)$ denotes the sigmoid activation function, and $W_{edge}^{(k)}$ and $b_{edge}^{(k)}$ represent learnable parameters.

The node subsequently updates its state based on its own features, a process expressed as Equation (21).

$$h_i^{(k)} = \text{UPDATE}^{(k)} \left(h_i^{(k-1)}, m_i^{(k)} \right) \quad (21)$$

Here, UPDATE and AGGREGATE denote learnable functions. Figure 6 illustrates the schematic diagram of the GNN message-passing mechanism.

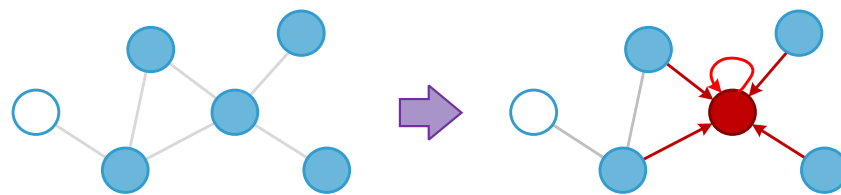


Figure 6. Schematic diagram of GNN message passing mechanism.

Upon completion of message transmission, the original node features are fused with the aggregated messages. Node features are updated through cascading and transformation:

$$Z^{(k)} = [H^{(k-1)} \parallel M^{(k)}] \quad (22)$$

$M^{(k)}$ denotes the aggregated message matrix. This concatenation strategy preserves the node’s historical information while integrating new information from the neighborhood, thereby forming a richer representation. The feature dimensions are doubled following concatenation, encompassing both local and global information.

Subsequently, a linear transformation and activation stage is applied, as formulated in Equation:

$$\tilde{H}^{(k)} = \text{ReLU}\left(\text{LN}\left(W_{\text{trans}}^{(k)} Z^{(k)} + b_{\text{trans}}^{(k)}\right)\right) \quad (23)$$

Here, the linear transformation matrix $W_{\text{trans}}^{(k)}$ maps the concatenated high-dimensional features back to the original dimension while learning the interactions between features. $\text{LN}(\cdot)$ denotes layer normalisation, which ensures the stability of feature distributions across different samples and aids in model training convergence. ReLU introduces non-linearity, enabling the model to learn complex feature representations. This step filters and reorganizes the fused information to extract valuable feature combinations.

The feature $\tilde{H}^{(k)}$ is subsequently fed into the multi-head attention layer. This mechanism captures feature relationships across distinct subspaces by computing multiple attention heads in parallel. Its core principle involves calculating and assigning attention weights to each position within the input sequence, thereby enabling the model to effectively capture critical information. Multi-head attention projects the input onto multiple distinct subspaces, computing several attention heads in parallel.

First, the feature matrix $\tilde{H}^{(k)}$ undergoes linear transformations to generate the Query, Key, and Value matrices: $Q = \tilde{H}^{(k)} W_Q$, $K = \tilde{H}^{(k)} W_K$, and $V = \tilde{H}^{(k)} W_V$. Here, $W_Q, W_K \in \mathbb{R}^{d_v \times d_k}$ and $W_V \in \mathbb{R}^{d_v \times d_v}$ denote learnable projection matrices, while d_k represents the dimensions of the keys and queries. These matrices are partitioned into multiple heads, each focusing on distinct subspaces within the feature space. Within each attention head, the Query matrix undergoes a dot product operation with the transpose of the Key matrix.

For the i -th attention head, the computation process is as follows:

$$\text{head}_i = \text{Attention}\left(QW_i^Q, KW_i^K, VW_i^V\right) \quad (24)$$

The attention function is defined as follows:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (25)$$

Here, $d_k = \frac{d_{\text{model}}}{h}$ denotes the dimension per head, and the scaling factor $\sqrt{d_k}$ prevents overly large dot-product values and stabilizes attention computation.

The outputs of h attention heads are obtained through linear transformation and concatenation:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O \quad (26)$$

$W^O \in \mathbb{R}^{hd_v \times d_{\text{model}}}$ represents the output projection matrix. The multi-head attention output is fused with the input feature $\tilde{H}^{(k)}$ via residual connections to yield the final output $H^{(k)}$.

$$H^{(k)} = \tilde{H}^{(k)} + \text{MultiHead}\left(\tilde{H}^{(k)}\right) \quad (27)$$

This design enables the model to selectively enhance crucial feature relationships while preserving the original information. The multi-head architecture enables the model to process data from multiple perspectives simultaneously, capturing diverse types of dependencies. Figure 7 illustrates the computational process of multi-head attention.

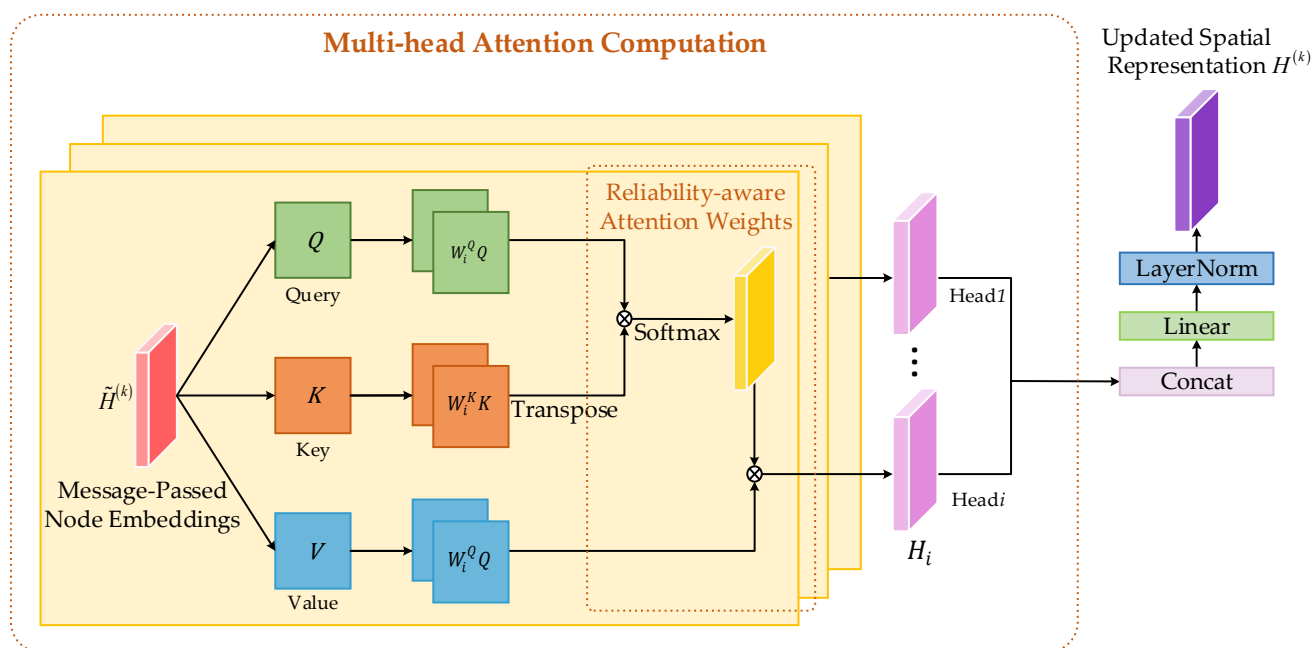


Figure 7. Multi-head attention-based spatial feature aggregation in the proposed MAFF-GNN.

To encourage the positioning predictions to remain consistent with ranging geometry rather than merely fitting data statistics, a geometry-constrained objective function is constructed. This function integrates supervised coordinate-regression error with ranging-consistency and geometric-regularization terms. The total loss $\mathcal{L}_{\text{total}}$ is formulated as a weighted multi-objective optimization problem:

$$L_{\text{total}} = \lambda_{\text{reg}} L_{\text{reg}} + \lambda_{\text{phy}} L_{\text{phy}} + \lambda_{\text{geo}} L_{\text{geo}} \quad (28)$$

where λ_{reg} , λ_{phy} , and λ_{geo} are hyperparameters balancing the contributions of coordinate regression, ranging consistency, and geometric topology, respectively.

The primary objective is to minimize the deviation between the predicted position $\hat{p}_i$ and the ground truth p_i . To mitigate the impact of outliers caused by severe NLOS bursts in mine tunnels, we employ the Huber Loss instead of the standard MSE. The Huber loss behaves quadratically for small errors and linearly for large errors:

$$\mathcal{L}_{\text{reg}} = \frac{1}{M} \sum_{i=1}^M \begin{cases} \frac{1}{2} \|\hat{p}_i - p_i\|^2, & \text{if } \|\hat{p}_i - p_i\| \leq \delta \\ \delta (\|\hat{p}_i - p_i\| - \frac{1}{2}\delta), & \text{otherwise} \end{cases} \quad (29)$$

The attention weights produced by the MAFF encoder are used as link-reliability indicators in the ranging-consistency term. Under ideal LOS ranging conditions, the Euclidean distance between the predicted tag position $\hat{p}_i$ and the anchor a_j should be consistent with the measured UWB range. However, blindly enforcing this constraint on NLOS edges introduces errors. Therefore, we propose an Attention-Guided Consistency Loss. We use the aggregated attention weights α_{ij} as reliability indicators. This allows the gradient descent to “trust” LOS edges while ignoring physical contradictions caused by NLOS edges:

$$\mathcal{L}_{\text{phy}} = \frac{1}{|\mathcal{E}|} \sum_{(i,j) \in \mathcal{E}} \alpha_{ij} \cdot \left| \|\hat{p}_i - a_j\| - \hat{d}_{ij} \right| \quad (30)$$

By integrating α_{ij} , this term acts as a soft ranging-consistency constraint, guiding the decoder to emphasize reliable ranging links.

To further regularize the solution space, we introduce a geometric constraint based on the subtended angles within the anchor topology. Unlike absolute distances, angular relationships are invariant to signal scaling errors. We define the geometric loss using the Cosine Theorem Constraint. For any triplet formed by the predicted tag $\hat{p}$ and two anchors a_j, a_k , the predicted cosine value $\cos(\hat{\theta}_{ijk})$ should align with the observed cosine value $\cos(\theta_{ijk})$ derived from raw measurements:

$$\mathcal{L}_{geo} = \frac{1}{C_N^3} \sum_{i < j < k} |\cos(\hat{\theta}_{ijk}) - \cos(\theta_{ijk})|^2 \quad (31)$$

where $\cos(\hat{\theta}_{j,k}) = \frac{(\hat{p}-a_j) \cdot (\hat{p}-a_k)}{\|\hat{p}-a_j\| \|\hat{p}-a_k\|}$. This term encourages topological consistency by constraining the predicted position to maintain geometrically valid angular relationships with the anchor layout.

For one forward pass, the computational complexity of MAFF-GNN can be approximated as $O(K(|E|d + hN^2d_h + Nd^2))$, where K is the number of encoder layers, $|E|$ is the number of ranging links, N is the number of graph nodes, d is the hidden feature dimension, h is the number of attention heads, and d_h is the feature dimension of each attention head. The first term corresponds to graph message passing, the second term corresponds to multi-head attention, and the third term corresponds to feature transformation and decoding. Since UWB localization graphs usually contain a limited number of anchors and tags, the graph size is much smaller compared to image-based or sequence-based networks. Therefore, the computational burden remains moderate, making the proposed method suitable for real-time localization under the simulated setting.

3.3. Baseline Models and Implementation Settings

To ensure a fair comparison, all methods were evaluated using the same simulated layouts, dataset splits, and UWB ranging inputs. Traditional methods were solved directly from ranging measurements and anchor coordinates, while learning-based methods were trained on the same samples. Graph-based models used the graph construction strategy described in Section 3.2.

Table 2 summarizes the compared methods. WLS and Taylor expansion were selected as traditional localization baselines, the Kalman filter as a filtering-based baseline, SVR as a conventional machine-learning method, MLP as a non-graph neural network, and GCN, GAT, and GraphSAGE as representative graph neural networks. The proposed MAFF-GNN was compared against these methods to evaluate the effectiveness of the proposed architecture.

Table 2. Implementation settings of the compared methods.

Method	Category	Training/Solving Protocol
WLS	Traditional positioning	No training; solved independently for each test sample
Taylor series expansion	Traditional positioning	No training; initialized using the centroid or WLS estimate
Kalman filter	Filtering-based positioning	No neural training; filtering applied to ordered samples or trajectories
SVR	Machine learning	Trained on the same training split
MLP	Neural network	Adam optimizer; same training split and early stopping
GCN	Graph neural network	Same graph dataset and training protocol
GAT	Graph neural network	Same graph dataset and training protocol
GraphSAGE	Graph neural network	Same graph dataset and training protocol
MAFF-GNN	Proposed	Same training split; geometry-constrained training objective

Table 3 lists the common training settings. All trainable models used the Adam optimizer with a learning rate of 0.001 and a batch size of 32. The maximum number of epochs was 500, with early stopping (patience = 200) applied to prevent overfitting. Each model was trained and tested five times, and the reported results are the mean and standard deviation.

Table 3. Training hyperparameters and computational settings.

Item	Setting
Framework	PyTorch 1.9.0
Python/CUDA	Python 3.8/CUDA 11.1
Hardware	NVIDIA TITAN RTX GPU with 24 GB memory
Data split	Training/validation/test = 2500/500/500 per scene
Optimizer	Adam
Learning rate	0.001
Batch size	32
Maximum epochs	500
Early stopping	200
Number of independent runs	5
Reported statistics	Mean $\pm$ standard deviation

4. Results and Discussion

4.1. Experimental Results

To evaluate whether the proposed MAFF-GNN remains stable under topology changes, the predicted tag positions were visualized in Scene A, Scene B, and Scene C. The three scenes correspond to the ring-corridor, cross-corridor, and dense multi-cell corridor layouts defined in Section 3.1, respectively.

Figure 8 visualizes the true tag positions, predicted tag positions, and UWB anchors in the three simulated mine topologies. The blue points denote the ground-truth tag positions, the orange points denote the predicted positions, and the red circles denote UWB anchors. Across the three scenes, most predicted positions are close to the corresponding true positions, whereas larger deviations mainly occur near obstacles, corridor intersections, and narrow passages where NLOS and multipath effects are more likely to appear.

Unless otherwise specified, all results in this section were evaluated on the scene-stratified test set containing 1500 samples, with 500 samples from each simulated mine topology.

To quantitatively evaluate positioning performance, Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are employed as key metrics. The RMSE, defined as

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \|p_i - \hat{p}_i\|_2^2}$$

serves as a measure of the overall positioning accuracy by penalizing large deviations. Complementarily, the MAE, defined as $MAE = \frac{1}{n} \sum_{i=1}^n \|p_i - \hat{p}_i\|_2$, provides an assessment of average accuracy that is more robust to outliers. P90 denotes the 90th percentile of the Euclidean localization error over the test samples.

RMSE and MAE are used as the primary accuracy metrics. To improve statistical reliability, the reported values should be presented as the mean $\pm$ standard deviation over five independent runs. Table 4 summarizes the overall positioning performance of the proposed MAFF-GNN and the compared baseline methods under the same simulated test setting.

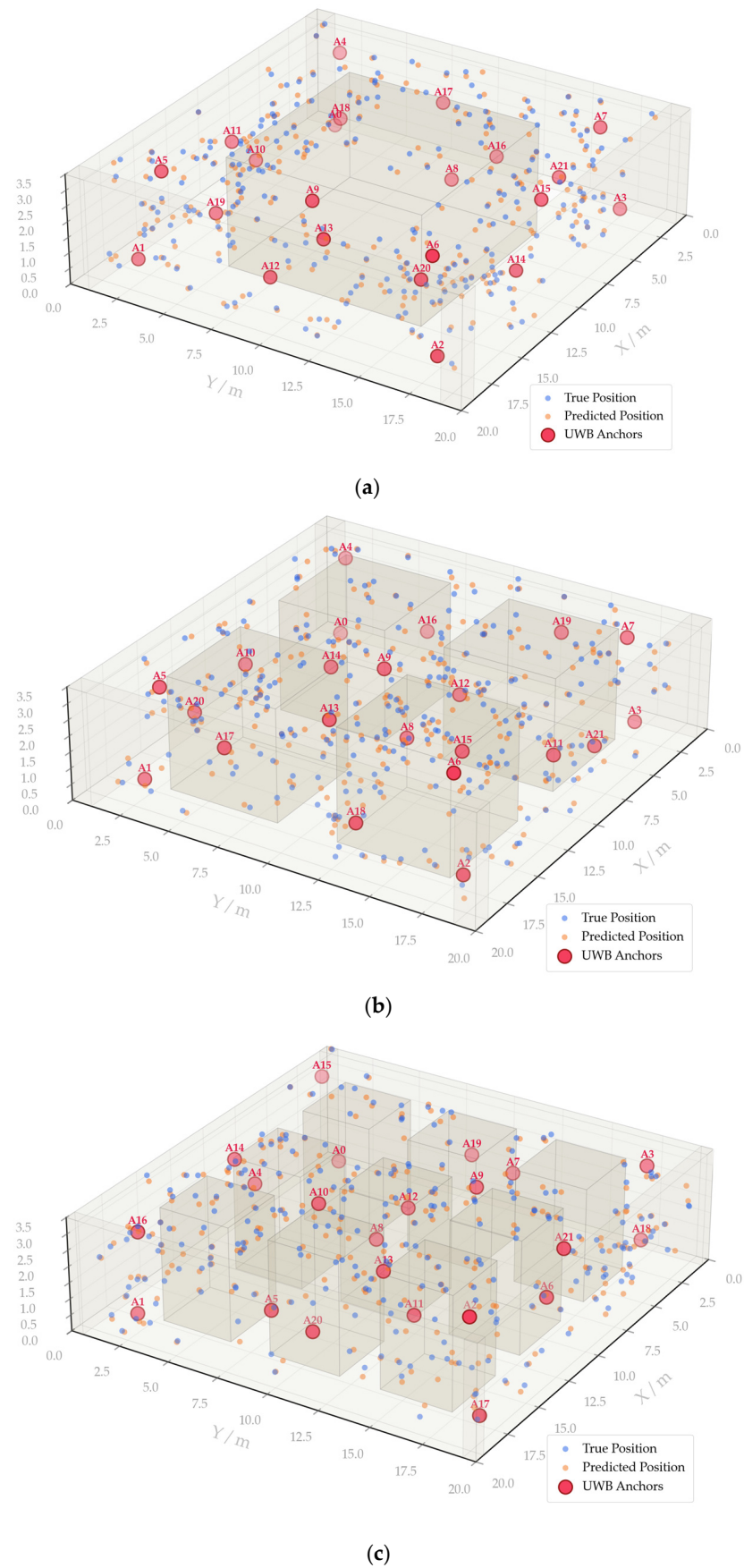


Figure 8. Positioning visualization in three simulated mine topologies: (a) Scene A; (b) Scene B; and (c) Scene C.

Table 4. Overall performance comparison of different positioning algorithms under simulated mine scenarios.

Method	RMSE (m, Mean $\pm$ SD)	MAE (m, Mean $\pm$ SD)	P90 Error (m, Mean $\pm$ SD)
WLS	1.25 $\pm$ 0.345	1.010 $\pm$ 0.306	1.719 $\pm$ 0.521
Taylor series expansion	0.966 $\pm$ 0.093	0.806 $\pm$ 0.071	1.701 $\pm$ 0.157
Kalman filter	1.76 $\pm$ 0.043	1.52 $\pm$ 0.047	2.281 $\pm$ 0.062
SVR	0.541 $\pm$ 0.046	0.493 $\pm$ 0.041	0.596 $\pm$ 0.059
MLP	0.845 $\pm$ 0.105	0.747 $\pm$ 0.076	1.256 $\pm$ 0.144
GCN	0.692 $\pm$ 0.031	0.625 $\pm$ 0.025	1.024 $\pm$ 0.040
GAT	0.504 $\pm$ 0.050	0.421 $\pm$ 0.043	0.762 $\pm$ 0.068
GraphSAGE	0.655 $\pm$ 0.033	0.593 $\pm$ 0.036	0.973 $\pm$ 0.023
MAFF-GNN	0.323 $\pm$ 0.093	0.274 $\pm$ 0.024	0.480 $\pm$ 0.038

The proposed MAFF-GNN achieves the best localization performance, with an RMSE of 0.323 ± 0.093 m, an MAE of 0.274 ± 0.024 m, and a P90 error of 0.480 ± 0.038 m. Compared with SVR, it reduces RMSE, MAE, and P90 error by 40.30%, 44.42%, and 19.46%, respectively. Relative to WLS, the RMSE reduction reaches 74.16%. These results demonstrate that MAFF-GNN consistently outperforms both traditional and learning-based baselines in the simulated mine environment.

Beyond average error values, the cumulative distribution function (CDF) curves provide a more complete view of localization reliability. Figure 9 compares the error distributions of WLS, MLP, GCN, GAT, GraphSAGE, and MAFF-GNN in the three simulated mine topologies.

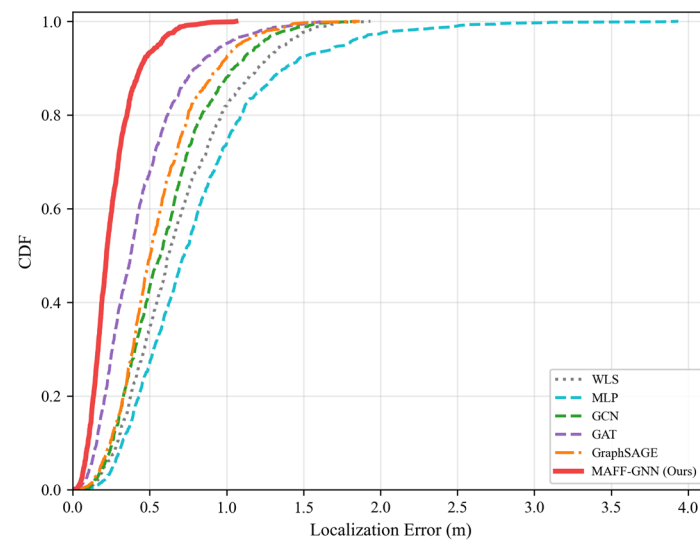
Across the three scenes, the MAFF-GNN curve is consistently shifted to the left of the baseline curves, indicating that a larger proportion of samples is localized within smaller error thresholds. This trend suggests that the proposed attention-based graph fusion mechanism is effective not only in average-error reduction, but also in suppressing large-error tails caused by NLOS and multipath propagation.

To clarify the influence of network depth, the number of MAFF-GNN encoder layers was varied from 2 to 10. Table 5 reports the average RMSE and convergence epochs over Scene A, Scene B, and Scene C. Performance Variation is defined as the RMSE difference relative to the best-performing depth configuration (Depth 8), and therefore, a smaller value indicates less degradation from the optimal configuration.

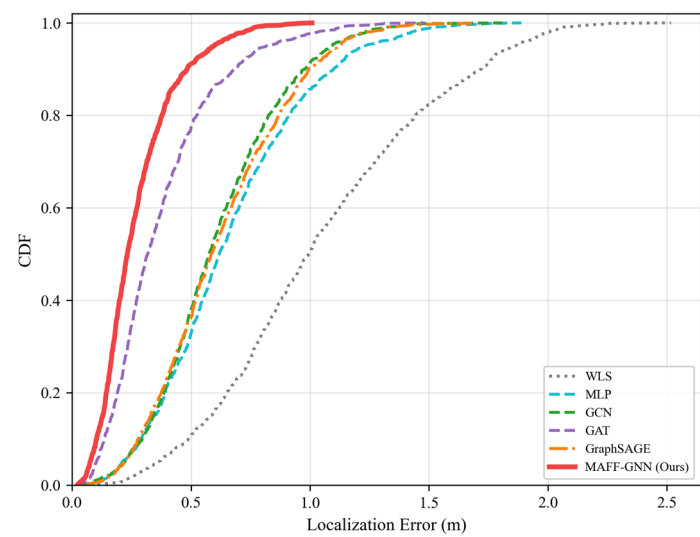
Table 5. Depth-sensitivity analysis of MAFF-GNN averaged over three simulated mine scenarios.

Depth	Scene A RMSE (m)	Scene B RMSE (m)	Scene C RMSE (m)	Average RMSE (m)
2	0.50	0.54	0.58	0.54
4	0.39	0.46	0.47	0.44
6	0.33	0.40	0.39	0.37
8	0.29	0.32	0.35	0.32
10	0.40	0.44	0.48	0.44

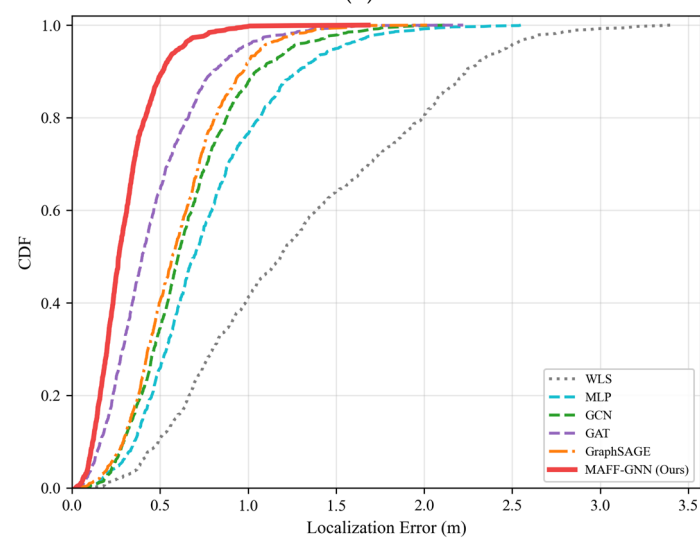
The results show that increasing the network depth from 2 to 8 layers gradually reduces the average RMSE from 0.54 m to 0.32 m, indicating that deeper graph feature extraction improves spatial representation across different corridor topologies. However, when the depth is further increased to 10 layers, the average RMSE rises to 0.44 m. This degradation may be attributed to over-smoothing or overfitting in the simulated graph dataset. Therefore, Depth 8 is selected as the default configuration in the subsequent experiments because it provides the best balance between localization accuracy and model depth.



(a)



(b)



(c)

Figure 9. CDF curves of localization errors in three simulated mine topologies: (a) Scene A; (b) Scene B; and (c) Scene C.

To evaluate the robustness of MAFF-GNN under different NLOS conditions, the simulated NLOS ratio in the test data was set to 30%, 50%, 60%, 70%, and 90%, respectively. During this test, the trained MAFF-GNN model was kept fixed, and only the NLOS ratio of the test samples was changed. The evaluation was conducted separately on Scene A, Scene B, and Scene C, and the average RMSE across the three scenarios was reported.

As shown in Figure 10, the localization error increases as the NLOS ratio becomes higher. In Scene B, the RMSE increases from 0.220 m at 30% NLOS to 0.440 m at 90% NLOS. Across the three simulated mine scenarios, the average RMSE increases from 0.227 m to 0.468 m. When the NLOS ratio is 60%, the average RMSE is 0.323 m. Among the three topologies, Scene A consistently obtains the lowest RMSE, whereas Scene C produces the highest RMSE because of its denser obstacle layout and more complex corridor connectivity.

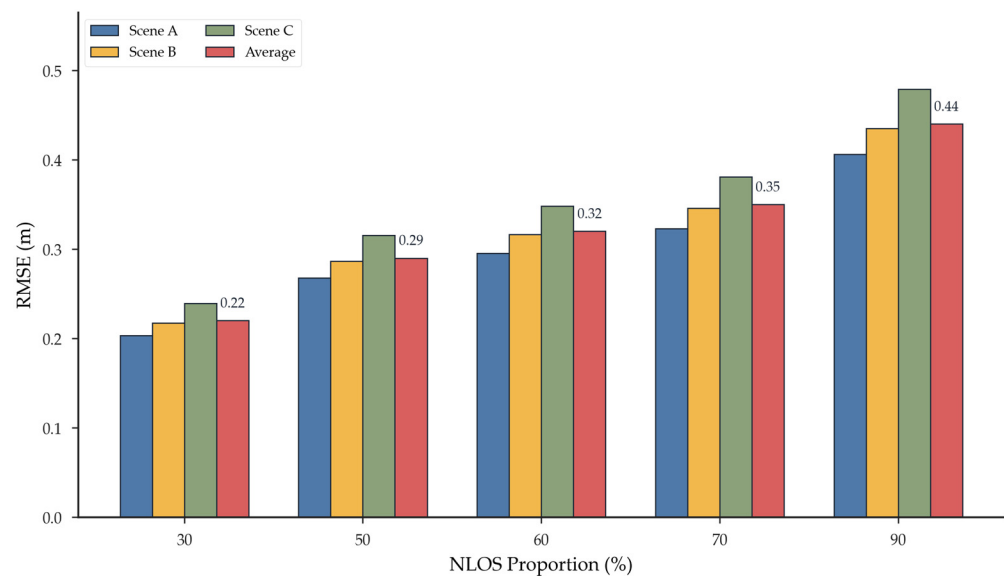


Figure 10. NLOS robustness test results.

These results indicate that MAFF-GNN maintains a gradual error-growth trend under increasing simulated NLOS ratios. However, the performance degradation under 90% NLOS also suggests that extremely severe obstruction remains challenging.

4.2. Limitations and Practical Implications

This study is based on physics-guided simulation rather than real underground UWB measurements. Therefore, the reported accuracy should be interpreted as evidence under controlled simulation conditions, not as deployment-validated performance. Real mine environments may involve additional uncertainties, including metallic equipment, dynamic human blockage, irregular tunnel deformation, time-varying humidity, and hardware synchronization errors.

The proposed method may also degrade under very sparse anchor deployment, rapid tag motion, or frequent LOS/NLOS transitions beyond the training distribution. Although three simulated mine topologies are used to reduce topology-specific overfitting, the generalization ability of MAFF-GNN still requires validation on more mine layouts and real field data.

Future work will focus on real-mine data collection, calibration between simulated and measured ranging errors, embedded-device inference tests, and online adaptation under changing NLOS conditions. These efforts are necessary to further assess the practical value of MAFF-GNN for underground personnel positioning and safety management.

5. Conclusions

This study investigated whether graph-structured UWB ranging information can be used to improve three-dimensional localization under severe simulated NLOS and multipath conditions in underground mine-like tunnel environments. The results suggest that UWB localization in such environments can be more effectively formulated as a relation-aware graph inference problem rather than as an independent range-fitting problem. By representing anchors and tags as nodes and ranging links as edges, the proposed MAFF-GNN framework exploits both spatial topology and measurement reliability, thereby providing a structured learning paradigm for NLOS-resilient mine localization.

The scientific implication of these findings is that attention-enhanced graph learning can serve as a data-driven mechanism for suppressing unreliable ranging links while preserving useful geometric constraints. Compared with conventional threshold-based or purely regression-based approaches, the proposed framework does not rely solely on manually defined NLOS rejection rules. Instead, it learns reliability-aware representations from noisy ranging relations, which is particularly relevant for elongated tunnel environments where multipath propagation, wall roughness, coal dust, humidity, and NLOS conditions jointly affect UWB measurements.

Under the controlled simulation setting, MAFF-GNN achieved an RMSE of 0.323 ± 0.093 m, an MAE of 0.274 ± 0.024 m, and a P90 error of 0.480 ± 0.038 m. These results indicate that combining graph message passing, multi-head attention, global skip connections, and geometry-constrained regularization can improve localization robustness across different simulated mine topologies and NLOS ratios. More broadly, the study provides evidence that topology-aware learning may be a promising direction for robust UWB positioning in underground intelligent mining systems.

Nevertheless, the present conclusions are based on physics-guided simulation rather than real underground mine measurements. Therefore, the reported performance should be interpreted as controlled simulation evidence rather than deployment-validated accuracy. Future work will focus on collecting real mine UWB datasets, calibrating the simulation model against measured ranging errors, evaluating cross-site generalization, and testing embedded deployment and online adaptation under dynamic LOS/NLOS transitions.

Author Contributions: Z.Y. made substantial contributions to the conception and design of the work and wrote the main manuscript text. Q.L. provided a critical revision of the work for important intellectual content. Y.Q. supervised the project and provided resources. All authors have read and agreed to the published version of the manuscript.

Funding: The research leading to these results received funding from the Guangxi Baise Scientific Research and Technology Development Plan Project under Grant Agreement No. Baik202432012 and the Innovation Project of Guangxi Graduate Education under Grant No. YCSW2026397 and Guangxi Innovation-driven Development Special Project under Project Number AA21077008.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on reasonable request from the corresponding author. The data are not publicly available due to privacy and ethical restrictions.

Conflicts of Interest: Author Yong Qian was employed by the company Guangxi Baise Industrial Investment Development Group Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CIR	Channel Impulse Response
CRLB	Cramer-Rao Lower Bound
CTLS	Constrained Total Least Squares
DOP	Dilution of Precision
EKF	Extended Kalman Filter
GAT	Graph Attention Network
GCN	Graph Convolutional Network
GNN	Graph Neural Network
IoT	Internet of Things
LOS	Line-of-Sight
MAE	Mean Absolute Error
MAFF-GNN	Multi-head Attention Feature Fusion Graph Neural Network
MLP	Multi-Layer Perceptron
NLOS	Non-Line-of-Sight
RMSE	Root Mean Square Error
P90	90th Percentile Error
RFID	Radio Frequency Identification
RSSI	Received Signal Strength Indicator
SNR	Signal-to-Noise Ratio
SVR	Support Vector Regression
TDOA	Time Difference of Arrival
TOA	Time-of-Arrival
UWB	Ultra-Wideband
WLS	Weighted Least Squares

References

1. Hancke, G.P.; Silva, B.J. Wireless Positioning in Underground Mines: Challenges and Recent Advances. *IEEE Ind. Electron. Mag.* **2021**, *15*, 39–48. [\[CrossRef\]](#)
2. Sang, C.L.; Steinhagen, B.; Homburg, J.D.; Adams, M.; Hesse, M.; Rückert, U. Identification of NLOS and Multi-Path Conditions in UWB Localization Using Machine Learning Methods. *Appl. Sci.* **2020**, *10*, 3980. [\[CrossRef\]](#)
3. Majeed, A.F.; Arsat, R.; Baharudin, M.A.; Latiff, N.M.A.; Albaidhani, A. Accurate Multiclass NLOS Channels Identification in UWB Indoor Positioning System-Based Deep Neural Network. *IEEE Access* **2024**, *12*, 179431–179448. [\[CrossRef\]](#)
4. Yang, H.; Wang, Y.; Seow, C.K.; Sun, M.; Coene, S.; Huang, L. Fuzzy Transformer Machine Learning for UWB NLOS Identification and Ranging Mitigation. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 8503817. [\[CrossRef\]](#)
5. He, X.; Mo, L.; Wang, Q. An Attention-Assisted UWB Ranging Error Compensation Algorithm. *IEEE Wirel. Commun. Lett.* **2023**, *12*, 421–425. [\[CrossRef\]](#)
6. Su, Z.; Pahlavan, K.; Agu, E.; Wei, H. Proximity Detection During Epidemics: Direct UWB TOA Versus Machine Learning Based RSSI. *Int. J. Wirel. Inf. Netw.* **2022**, *29*, 480–490. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Alizad, M.; Nobahari, H. Estimation of Accelerometers' Bias in Decentralized Relative Localization via the Rotating UWB Tag. *ISA Trans.* **2025**, *167*, 1006–1018. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Liang, B.; Zhu, X.; Liu, T.; Shan, G. Non-Line-of-Sight Error Compensation Method for Ultra-Wideband Positioning System. *Machines* **2025**, *13*, 1018. [\[CrossRef\]](#)
9. Liu, Z.; Li, Z. NLOS Mitigation and Indoor Positioning Method Based on UWB Ranging. *IEEE Sens. J.* **2025**, *25*, 43372–43381. [\[CrossRef\]](#)
10. Cheng, J.H.; Yu, P.P.; Huang, Y.R. Application of Improved Kalman Filter in Under-Ground Positioning System of Coal Mine. *IEEE Trans. Appl. Supercond.* **2021**, *31*, 0603904. [\[CrossRef\]](#)
11. Li, Z.; Li, Z. Reference Coordinate Based Chan Algorithm for UWB Personnel Localization in Underground Coal Mines. *Sci. Rep.* **2025**, *15*, 17922. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Messias, G.H.; Iasulaitis, S.; Valejo, A.D.B. Graph Neural Networks for Positive and Unlabeled Learning: A Rewiring Approach. *Knowl. Inf. Syst.* **2025**, *67*, 10727–10749. [\[CrossRef\]](#)

13. Orfanoudakis, S.; Robu, V.; Salazar, E.M.; Palensky, P.; Vergara, P.P. Scalable Reinforcement Learning for Large-Scale Coordination of Electric Vehicles Using Graph Neural Networks. *Commun. Eng.* **2025**, *4*, 118. [[CrossRef](#)] [[PubMed](#)]
14. Komarovskiy, S.; Haddad, J. Spatio-Temporal Graph Convolutional Neural Network for Traffic Signal Prediction in Large-Scale Urban Networks. *Transp. Res. Interdiscip. Perspect.* **2025**, *32*, 101482. [[CrossRef](#)]
15. Dong, G.; Tang, M.; Wang, Z.; Gao, J.; Guo, S.; Cai, L.; Gutierrez, R.; Campbell, B.; Barnes, L.E.; Boukhechba, M. Graph Neural Networks in IoT: A Survey. *ACM Trans. Sens. Netw.* **2023**, *19*, 1–50. [[CrossRef](#)]
16. He, S.; Yang, B.; Liu, T.; Zhang, H. Multi-Tag UWB Localization With Spatial-Temporal Attention Graph Neural Network. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 2531112. [[CrossRef](#)]
17. Tang, K.; Yang, B.; Ding, K. Deep Attention-Based Network Combing Geometric Information for UWB Localization in Complex Indoor Environments. *IEEE Access* **2024**, *12*, 31488–31497. [[CrossRef](#)]
18. Tu, C.; Zhang, J.; Quan, Z.; Ding, Y. UWB Indoor Localization Method Based on Neural Network Multi-Classification for NLOS Distance Correction. *Sens. Actuators A Phys.* **2024**, *379*, 115904. [[CrossRef](#)]
19. Jia, C.; Tao, C.; Yang, T.; Fu, M.; Zhou, X.; Huang, Z. UWB Indoor Localization Based on Artificial Rabbit Optimization Algorithm and BP Neural Network. *Biomimetics* **2025**, *10*, 367. [[CrossRef](#)] [[PubMed](#)]
20. Han, M.; Cao, B.; Wang, S.; Xu, M. Study on UWB Path Loss in Rectangular Mine Roadway. *Coal Technol.* **2024**, *43*, 218–222.
21. Cao, B.; Wang, S.; Lu, C. Autonomous Positioning Method of Shearer at the End of Coal Mining Face Based on UWB System. *Chin. J. Sci. Instrum.* **2022**, *43*, 108–117.
22. Li, D.; Sui, J.; Liu, G.; Zhao, Z. Technical Status and Development Direction of Coal Mine Dust Hazard Prevention and Control Technology in China. *Min. Saf. Environ. Prot.* **2019**, *46*, 1–7+13. (In Chinese) [[CrossRef](#)]
23. Ministry of Emergency Management of the People's Republic of China. *Coal Mine Safety Regulations; Promulgated by Order No. 87 of the State Administration of Work Safety on 25 February 2016 and Amended by Order No. 8 of the Ministry of Emergency Management on 6 January 2022*; Ministry of Emergency Management of the People's Republic of China: Beijing, China, 2022. Available online: <https://www.mem.gov.cn/> (accessed on 18 June 2026).
24. Tong, L.; Wang, X.; Xu, Y.; Zhang, C. Current Situation and Control Countermeasures of Respirable Dust Hazard in Underground Working Face. *China Coal* **2020**, *46*, 47–51. (In Chinese)
25. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 5998–6008. Available online: <https://arxiv.org/abs/1706.03762> (accessed on 1 May 2026).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.