Intelligent Campus Network Monitoring and Optimization

IoT and Machine Learning

Review 2

Hemesh Yeturu	(CB.SC.U4AIE23166)
Joel John	(CB.SC.U4AIE23131)
Abhishek Sankaramani	(CB.SC.U4AIE23107)
Adarsh P	(CB.SC.U4AIE23109)

Table of Contents

Background
Context

Introduction: Real-time 3D WiFi Signal Strength Mapper with Device Tracking
Importance: Addresses the growing demand for seamless wireless connectivity by analyzing WiFi coverage, detecting signal dead zones, and tracking device movements.
Technology Stack: Integrates IoT and Machine Learning (ML) to enhance real-time monitoring and analysis. .
Impact: Helps in optimizing network performance and ensuring seamless connectivity..

- **Key technologies used:** IoT: Real-time data acquisition from wireless access points and devices. Machine Learning (ML): Signal strength classification and dead zone detection using clustering algorithms (DBSCAN, KMeans). Data Visualization: Interactive 3D mapping of WiFi coverage and device locations using Dash and Plotly.

Project Objectives

- **First Goal:** Collect real-time WiFi signal strength data across various locations. Generate a 3D visualization of WiFi coverage to identify weak signal areas. Dead Zone Detection
- **Second Goal:** Detect areas with poor or no WiFi coverage using clustering algorithms and dynamic thresholding. Provide insights for WiFi network optimization. Device Tracking and Distance Estimation
- **Third Goal:** Store WiFi signal data and device locations in structured CSV files. Ensure efficient data handling for further analysis and predictive modeling. User-friendly Visualization and Dashboard
- **Fourth Goal:** Develop an interactive dashboard using Dash to monitor signal strength and device movement. Provide real-time insights for users and network administrators.

Significance of the project

- **Challenge:** Unreliable Wi-Fi coverage due to environmental factors, interference, and improper router placements.
- **Goal:** Develop an automated, intelligent system to map network performance, predict connectivity issues, and suggest optimizations.
- **ML Relevance:** Machine learning is utilized to distinguish between normal signal variations and genuine dead zones. Supervised learning models such as decision trees or regression models can classify weak signal zones efficiently.

Justification

- **Relevance:** Ensuring strong and reliable campus-wide connectivity supports research, communication, and digital learning.
- **Impact:** Improves resource allocation, enhances IoT-based applications, and provides a data-driven approach to network management.
- **Why ML?:** Traditional network planning relies on static heuristics, whereas ML enables dynamic adaptation by learning from real-time data patterns, enhancing accuracy in dead zone detection.

Project Scope

- Signal strength monitoring using ESP32 and laptops.
- Real-time data collection and preprocessing.
- Map and visualization of dead zones.
- Basic ML-based predictions for network performance.

Mathematical Modelling

- Radio Wave Propagation: Friis Transmission Equation, Path Loss Models.
- Multipath Fading: Rayleigh and Rician fading models.
- Triangulation: Distance-based localization using multilateration.

Excluded:

- Advanced AI-based self-healing networks.
- Full-scale deployment beyond campus limits.

Inclusions and Exclusions

Inclusions:

- IoT sensor-based data collection.
- ML for anomaly detection and predictions.
- Real-time visualization and dashboard.
- Integration with existing databases (SQLite, PostgreSQL, MySQL, InfluxDB).

Exclusions:

- Hardware development beyond sensor deployment.
- Enterprise-level networking solutions.(e.g., Cisco proprietary systems).
- User behavior analysis beyond network performance metrics.

Project Phases

Phase	Deliverables
1	Initial data collection and dead zone identification.
2	Integration of triangulation techniques for object localization.
3	ML-based predictive modeling for network performance.

Algorithm Choice: Triangulation techniques such as centroid localization and multilateration will be used to estimate object positions. These techniques leverage signal strength variations to infer relative distances from multiple access points.

Flowcharts Infographics :

ML algorithm decision processes.

IoT device network interactions.

Signal preprocessing flowcharts.

Comparison with Existing Research

1.1 WiFi Signal Strength Estimation

Existing Research:

- Several studies focus on WiFi signal strength estimation using path loss models and empirical formulas.
- Traditional methods include log-distance models and free-space path loss calculations.
- Some papers explore machine learning models for signal strength prediction but do not integrate them with real-time systems.

How Our Project Differs:

- We use an adaptive filtering mechanism combined with machine learning models to improve signal strength estimation accuracy.
- Unlike static models, our approach continuously refines signal predictions using Kalman filtering and dynamic thresholding.
- Our system applies statistical noise modeling and supervised learning to enhance the precision of RSSI-based signal predictions.

Methodology - 2

Comparison with Existing Research

1.2 WiFi-Based Localization

Existing Research:

- WiFi-based localization is widely studied using trilateration, fingerprinting, and RSSI-based methods.
- Most approaches rely on pre-recorded signal maps that are not adaptable to real-time changes.

How Our Project Differs:

- We implement real-time device tracking using RSSI-based distance estimation, making it more dynamic.
- Our project incorporates Gaussian Process Regression and ML models to adapt to environmental conditions automatically.
- We integrate Dash-based visualization for interactive and real-time tracking, which is missing in most studies.

Comparison with Existing Research

1.3 Dead Zone Detection

Existing Research:

- Many studies use static thresholding (e.g., classifying dead zones if RSSI ¡ -80 dBm).
- Some ML-based approaches exist but primarily focus on static classification of weak signal areas.

How Our Project Differs:

- We use DBSCAN clustering to dynamically detect and classify dead zones based on real-time data.
- Our system incorporates disconnection data, considering areas where devices frequently drop connections as dead zones.
- Unlike conventional studies, our approach uses Gaussian Process Regression to refine dead zone classification, making it more adaptable.

Comparison with Existing Research

1.4 IoT and ML for WiFi Analysis

Existing Research:

- Many papers discuss either IoT-based WiFi monitoring or ML-based signal prediction, but they are rarely combined.
- Most IoT-based studies collect data but lack real-time interactive analytics.

How Our Project Differs:

- Our system seamlessly integrates IoT and ML, using real-time WiFi data to train and refine models dynamically.
- We offer a web-based interactive dashboard (Dash + Plotly) for real-time monitoring, an aspect missing in previous studies.
- Instead of just predicting signal strength, our system actively detects changes and optimizes network performance dynamically.

Comparison with Existing Research

1.5 WiFi Network Optimization

Existing Research:

- Most studies focus on static optimization strategies for router placement.
- Some research uses simulation-based approaches for optimizing coverage.

How Our Project Differs:

- Our system provides continuous, real-time tracking of network performance, allowing for dynamic adjustments.
- Instead of using predefined signal maps, our system adapts to real-world changes in WiFi conditions, improving optimization over time.

Overview of the Approach

Structured Methodology:

- **Data Collection:** Real-time acquisition of WiFi signal strength and device tracking data.
- **Preprocessing Feature Engineering:** Cleaning and structuring data for analysis.
- **Dead Zone Detection:** Using Random Forest Classification to identify low-signal areas.
- **Distance Estimation Tracking:** Applying Gaussian Process Regression (GPR) for accurate device positioning.
- **Visualization Real-time Monitoring:** Dash-based interactive dashboards for live analysis.

Overview of the Approach

Structured Methodology: System Workflow Stages:

- **Step 1: Data Collection using IoT Network Scanning**
- WiFi Signal Scanner: Captures SSID, BSSID, Signal Strength (RSSI), and Noise Levels.
- Device Tracker: Monitors devices and estimates distances.
- Dead Zone Logger: Logs weak signal areas and frequent disconnections.
- Data is stored in structured CSV files for further processing.

Dead Zone Detection, Distance Estimation Visualization

Step 2: Data Preprocessing Feature Engineering

Cleaning Handling Missing Values: Removes inconsistencies from collected data.

Feature Extraction: WiFi signal strength variations, device movement patterns, and environment classification.

Outlier Detection: Adaptive filtering (Kalman Filter) to refine noisy signal readings.

Step 3: Dead Zone Detection using Random Forest

- **Model Selection:** Random Forest Classifier used for dead zone detection.
- **Training Data:**
  - Features: Signal Strength, Noise, Device Dropout Patterns.
  - Labels: "Dead Zone" or "Normal Coverage".
- **Process:**
  - Model classifies weak signal areas as dead zones.
  - Classifier is periodically retrained with real-time data.

Dead Zone Detection, Distance Estimation Visualization

Step 4: Distance Estimation using Gaussian Process Regression (GPR)

- **Why GPR?**
  - Traditional log-distance path loss models are inaccurate in dynamic environments.
  - GPR provides uncertainty estimation for fluctuating WiFi conditions.
- **Process:**
  - Learns the relationship between RSSI and distance.
  - Predicts device distances with confidence intervals.

Step 5: Visualization Real-time Monitoring

- **Dashboard Implementation (Dash + Plotly)**
  - Displays WiFi signal strength in 3D mapping.
  - Tracks device movement and estimated distances.
  - Highlights dead zones using heatmaps and clustering results.

  Interactivity Alerts
  - Users can filter and analyze WiFi signal variations over time.
  - Real-time alerts for newly detected dead zones.
  - Live tracking of devices with dynamic distance updates.

Tools, Techniques, and Frameworks Used

Hardware IoT Components

- WiFi Adapters: Captures real-time RSSI values.
- Raspberry Pi / IoT Edge Devices: Runs scanning and preprocessing tasks.

Software Frameworks

- Python Pandas: For data processing and manipulation.
- Dash Plotly: For interactive visualization.
- Scikit-learn:
  - Random Forest Classifier for dead zone detection.
  - Gaussian Process Regression (GPR) for distance estimation.
- Scipy NumPy: For signal filtering and mathematical modeling.
- Geopy Scapy: For network scanning and geolocation processing.

Technical Depth Enhancements and Mathematical Formulations

Technical Depth Enhancements:

ESP32 Signal Processing: Signal filtering, Kalman filter implementation.

Machine Learning Model Architecture: Feature selection, model choice

Cloud Integration: Strategies for real-time network data in cloud storage.

Edge Computing Considerations: Processing data at IoT nodes for real-time decision-making.

Signal Strength Prediction (Kalman Filter Model):

$$X_t = AX_{t-1} + BU_t + W_t \tag{1}$$

- X_t: Estimated state of the network.
- A, B: Transition matrices.
- U_t: Active devices.
- W_t: Noise.

1. Completed Milestones

- Implemented WiFi signal strength scanning and device tracking.
- Integrated real-time data acquisition from WiFi adapters.
- Developed a structured CSV-based data storage system.

1.2 Data Preprocessing Feature Engineering Completed

- Implemented data cleaning, handling missing values, and feature extraction.
- Applied adaptive filtering (Kalman Filter) to smooth noisy RSSI values.
- Structured data for use in Random Forest (dead zone detection) and GPR (distance estimation) models.

1.3 Dead Zone Detection using Random Forest Completed

- Trained an initial Random Forest Classifier on collected signal strength data.
- Classified areas into dead zones vs. normal zones.
- Integrated DBSCAN clustering to refine dead zone detection.
- Outputting results to the dashboard visualization.

1.4 Distance Estimation using Gaussian Process Regression (GPR) Completed

- Implemented GPR for device distance estimation using RSSI.
- Improved model accuracy with environmental classification (Indoor/Outdoor).
- Integrated distance estimation into real-time tracking system.

1.5 Visualization Dashboard Implementation Partially Completed

- Developed interactive dashboards using Dash Plotly.
- 3D WiFi signal strength visualization is operational.
- Device tracking map is implemented, but UI refinements performance optimizations are in progress.
- Need to finalize real-time updates for dead zone visualization.

2. Work in Progress

2.1 Model Optimization Retraining

- Fine-tuning Random Forest and GPR models for better real-world accuracy.
- Enhancing training dataset with real-time network fluctuations.

2.2 Real-time Data Processing Improvements

- Optimizing data flow between modules to reduce processing latency.
- Implementing parallel processing for live device tracking updates.

2.3 Final Dashboard Enhancements

- Adding real-time alert system for dead zone detection.
- Improving responsiveness and UI elements for better usability.

3. Upcoming Tasks Planned

3.1 Integration Final Testing

- Perform end-to-end testing to validate all interconnected modules.
- Optimize data pipelines for efficiency.
- Finalize dashboard deployment usability testing.

3.2 Documentation Research Paper Finalization

- Completing technical documentation of methodologies.
- Writing comparative analysis and results discussion.

4. Summary of Progress

- Core Data Collection, Preprocessing, and Model Implementation (Dead Zone Distance Estimation) are completed.
- Dashboard Visualization is functional but requires optimization.
- Work remains on refining real-time data processing and final UI improvements.
- Final integration, testing, and research documentation are upcoming steps.

Challenges Risks

Technical: Accurate noise modeling for real-world conditions.Hardware limitations of ESP32 in signal measurement.

Multipath Fading Challenge:

- When signals reflect off surfaces, they can interfere with direct signals, affecting accuracy.
- ML-based signal processing can mitigate this by identifying anomalies in RSSI data.

Operational:

- Ensuring continuous real-time data collection.
- Managing data storage and processing at scale.

Performance Metrics

- **Signal Strength Measurement Protocols:** Standardized procedures for data collection.
- **Quantitative Benchmarks:** Evaluating localization accuracy, signal stability.
- **Error Rate Analysis:** Localization error vs. signal quality.
- **Comparative Performance Graphs:** Visualizing improvements over baseline models.

Conclusion

- **Summary:** The CNIS project aims to leverage IoT and ML to enhance network intelligence by identifying dead zones, optimizing network coverage, and enabling object localization.
- **Takeaway:** The integration of real-time data analytics, noise modeling, and ML-based predictions can revolutionize wireless network optimization for smart campuses.
- **Future Enhancements:** The project can be extended to adaptive learning techniques where the ML models continuously refine based on feedback, paving the way for autonomous network management systems.

Literature Review

No.	Title	Author(s)	Outcome Implemented in Code
1	Utilizing Multiple Condition RSSI Distance Conversion on WiFi Localization Utilizing Multiple Condition RSSI Dist	Ming Yen Tsaia, Fu Ching Tsai	Implemented logarithmic regression and triangulation to improve WiFi-based device localization, reducing errors in weak signal conditions.
2	Real-Time Experimentation and Analysis of Wifi Spectrum Utilization in Microwave Oven Noisy Environment	Yakubu S. Baguda	Used Wireshark and Spectrum Analyzer to measure WiFi signal degradation and throughput loss due to microwave interference.
3	Improving Signal Strength estimation in IOT using WI-FI Netowrk Performance Data	Apostol Todorov, Vanya Stoykova, Zlatin Zlatev	Developed a second-order polynomial regression model using RReliefF and PCA to accurately predict WiFi signal strength, achieving 94% accuracy in IoT network performance estimation

Literature Review - Part 2

No.	Title	Author(s)	Outcome Implemented in Code
4	Wi-Fi Received Signal Strength Indicator (RSSI) Factors in Influencing Indoor Signal: A Review	Mohd Shah Shafie, Akhyari Nasir	Supported the development of strategies to mitigate interference effects impacting signal strength.
5	Estimating the Physical Distance between Two Locations with Wi-Fi Received Signal Strength Information Using Obstacle-aware Approach	Tomoya Nakatani, Takuya Maekawa, Masumi Shirakawa, Takahiro Hara	Provided foundational insights for modeling noise and interference, influencing the implementation of noise analysis.
6	A Comprehensive Survey of WiFi Analyzer Tools	Jivthesh M. R.	Inspired the implementation of the Kalman filter to smooth fluctuating RSSI data in the measurement pipeline.
7	SDN-Based Multipath Data Offloading Scheme using Link Quality Prediction for LTE and WiFi Networks	Santhosh Kamath, Aravinda Raman	Motivated the moving average filtering approach for preprocessing noisy signal data prior to analysis.

Literature Review - Part 3

No.	Title	Author(s)	Outcome Implemented in Code
8	An Enhanced Indoor Positioning Algorithm Based on Fingerprint Using Fine-Grained CSI and RSSI Measurements	Jingjing Wang, Joongoo Park	Influenced the triangulation-based object localization method used for identifying and eliminating dead zones.
9	Free Device Location Independent WiFi-based Localisation Using Received Signal Strength Indicator and Channel State Information	Mohammed Al-Andoli	Informed the real-time data collection and sensor integration strategy using ESP32 modules and laptops.
10	Indoor WiFi-Beacon Dataset Construction Using Autonomous Low-Cost Robot for 3D Location Estimation	Suleiman Abu Kharmeh, Emad Natsheh	Provided guidance for applying machine learning (e.g., regression models) to predict network performance and dead zones.

Literature Review - Part 4

No.	Title	Author(s)	Outcome Implemented in Code
11	Integrating GPS and WiFi Signal Strength Analysis for Enhanced Building Entrance Localization Using Fuzzy Logic	Ahmad Abadleh	Implemented a fuzzy logic system that classifies GPS and WiFi signals into weak, medium, and strong categories to improve building entrance detection for seamless indoor-outdoor localization.
12	An RSSI Classification and Tracing Algorithm to Improve Trilateration-Based Positioning	Yong Shi, Wenzhong Shi, Xintao Liu, Xianjian Xiao	Developed a tri-partition RSSI filtering algorithm that reduces signal variance and improves trilateration-based positioning stability by 4.45 times, with an accuracy improvement of 20.5%.
13	Supplementary Open Dataset for WiFi Indoor Localization Based on Received Signal Strength	Jingxue Bi, Yunjia Wang, Baoguo Yu, Hongji Cao, Tongguang Shi, Lu Huang	Created an open dataset (SODIndoorLoc) covering WiFi RSS-based localization in multi-floor buildings, optimized for machine learning models in indoor positioning with 2.3m accuracy.

Literature Review - Part 5

No.	Title	Author(s)	Outcome Implemented in Code
14	Wireless Communications	Andrea Goldsmith	Introduced multipath fading models (Rayleigh/Rician) and path loss exponent models for accurate wireless signal propagation predictions, which are used in network planning and optimization.
15	WiFi-Based Human Sensing with Deep Learning: Recent Advances, Challenges, and Opportunities	Iftikhar Ahmad, Arif Ullah, Wooyeol Choi	Applied deep learning models (CNN, LSTM) to WiFi CSI-based human activity recognition for applications in healthcare, security, and smart environments
16	WiFi Signal Propagation at 2.4 GHz	Muzaiyanah Hidayab, Abdul Halim Ali, Khairul Bariah Abas Azmi	Modeled WiFi signal propagation using the Log Distance Path Loss Model, deriving path loss coefficients for different environments (indoor, outdoor, and between buildings).

Literature Review - Part 6

No.	Title	Author(s)	Outcome Implemented in Code
17	WiFi Access Points Line-of-Sight Detection for Indoor Positioning Using the Signal Round Trip Time	Xu Feng, Khuong An Nguyen, Zhiyuan Luo	Developed a machine learning-based feature selection algorithm that uses WiFi RTT and RSSI to detect line-of-sight conditions, improving indoor positioning accuracy up to 98%.
18	Using Wi-Fi Signal Strength to Localize in Wireless Sensor Networks	Eddie C.L. Chan, George Baciu, S.C. Mak	Implemented a Kalman Filter-enhanced WiFi fingerprinting method for real-time localization in wireless sensor networks, reducing tracking errors in non-line-of-sight environments.

Thank You

- In this work, we reviewed several key papers that contribute to our understanding of wireless communication, machine learning, and sensor networks.
- The implementation of these concepts in our system helped refine the signal strength measurement, noise analysis, and data collection techniques.
- By combining adaptive algorithms, real-time signal analysis, and advanced machine learning methods, we have made significant strides towards accurate network performance prediction.
- The lessons learned from the literature have also shaped the design of our project, especially in overcoming challenges related to environmental factors, interference, and signal strength fluctuations.
- Future work will focus on expanding the system to incorporate more complex models and enhance the real-time performance of the solution.