Proactive Roaming Prediction using LSTM in Industrial Wi-Fi-based AGV Networks

Yonghan Kwon†, Jae Hong Shin*, and Jeongyeup Paek†

† Department of Computer Science and Engineering, Chung-Ang University, Seoul, Republic of Korea
* Manufacturing SW Platform R&D Team, Hyundai Motor Group, Uiwang-si, Gyeonggi-do, Republic of Korea
yh3155@cau.ac.kr, shjh@hyundai.com, jpaek@cau.ac.kr

***Abstract*—With the rise of smart factories employing Automated Guided Vehicles (AGVs) and Autonomous Mobile Robots (AMRs), ensuring reliable wireless communication has become a critical requirement for safe and continuous operation. However, in industrial Wi-Fi networks, frequent roaming often cause unexpected disconnections, undermining communication reliability and operational stability. This study aims to proactively predict roaming events in industrial Wi-Fi networks using large-scale data collected from AGVs and access points (APs) in real-world automobile factory. An Long Short-Term Memory(LSTM)-based model is employed to forecast each AGV's next AP, and roaming likelihood is further analyzed using the entropy of predicted probabilities. Our model detects more than 80% of roaming events in advance, contributing to improved communication stability in industrial environments.**

***Index Terms*—Automated Guided Vehicle (AGV), Software Defined Factory (SDF), Industrial Wireless, Industry 4.0**

I. Introduction

Factory automation is accelerating under Industry 4.0, leading to smart manufacturing environments where automated guided vehicles (AGVs) and autonomous mobile robots (AMRs) play key roles in material transport. To operate safely and efficiently, AGVs must continuously communicate with a central controller to exchange time-critical commands and sensor data. Thus, maintaining highly reliable and low-latency wireless links is essential for stable operation. If connectivity is interrupted, AGVs may stop or deviate from their assigned tasks, leading to line delays or productivity losses.

As wireless connectivity becomes mission-critical in smart manufacturing, traditional rule-based network management struggles with dynamic interference, multipath fading, and mobility-driven topology changes. Machine learning (ML) has emerged as a promising approach for network prediction [1], anomaly detection [2], and performance optimization. By leveraging time-series metrics such as RSSI, throughput, and latency, ML models–particularly sequence-learning architectures like long short-term memory (LSTM)–can capture temporal patterns that static threshold-based methods overlook, enabling proactive detection of link degradation or roaming behavior. Prior studies have also explored ML-based handover prediction in cellular and Wi-Fi networks [3]–[5], showing the feasibility of forecasting mobility events using sequence models.

This research was supported by the Hyundai Motor Group, also by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024-00359450), and also by the Institute of Information & Communications Technology Planning & Evaluation(IITP)-ITRC(Information Technology Research Center) grant funded by the Korea government(MSIT)(IITP-2025-RS-2022-00156353).

J. Paek is the corresponding author.

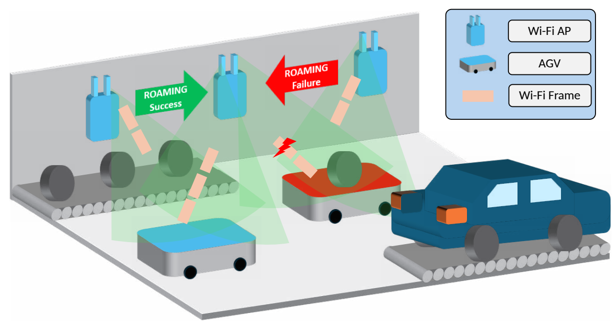


Fig. 1: Overview of AGV roaming scenario.

Despite these advances, predicting roaming in advance remains challenging because AGV mobility and channel conditions evolve quickly, and simple thresholding on RSSI or link quality fail to capture these dynamics. This motivates a data-driven sequential modeling approach tailored to industrial Wi-Fi.

In this paper, we propose a framework for proactive roaming prediction in factory Wi-Fi networks. Using real-world logs collected from AGVs and access points, we design an LSTM-based next-AP prediction model augmented with an entropy-based metric to estimate roaming likelihood and identify unstable link states.

The contributions of this paper are summarized as follows:

- We analyze real-world wireless traces from AGVs and Wi-Fi APs in an operational automobile factory to characterize roaming-related disconnect phenomena.
- We design an LSTM-based model for next-AP prediction and introduce an entropy-based metric to quantify roaming likelihood and link stability.
- We evaluate the proposed framework on field data and demonstrate its effectiveness in proactively identifying roaming events before they occur.

TABLE I: Total dataset analysis

	Total	Test	Ratio
Data size (timesteps)	93,083,287	16,436,677	17.65%
Roaming events (count)	101,333	22,167	21.87%
Roaming event ratio	0.108%	0.135%	-

II. Related Work

Manalastas et al. [3] developed a data-driven framework to predict and mitigate inter-frequency handover failures in 5G networks. Their focus lies in cellular environments and failure avoidance through radio-parameter adaptation, which differs from our objective of predicting proactive roaming behavior in Wi-Fi networks deployed in industrial facilities.

Lima et al. [4] treated handover prediction as a sequence learning problem using models such as LSTM and Gated Recurrent Unit (GRU). Although similar in methodology, their work targets conventional mobile networks with standardized RAN measurements, not dense industrial Wi-Fi systems with AGV-specific signal and mobility characteristics.

Dzaferagic et al. [5] demonstrated ML-based handover prediction in an O-RAN testbed using an LSTM-based xApp. While they addressed real-time roaming triggering, their goal was limited to determining whether a handover should occur, rather than identifying the next target cell or AP, which is central to proactive mobility management in industrial Wi-Fi.

In contrast to these studies, our work focuses on predicting the next AP in a factory Wi-Fi network using fine-grained AGV-side measurements and sequence-based learning, offering actionable insights for proactive roaming control in industrial environments.

III. Methodology

In this section, we present data analysis, labeling methodology, and the approaches evaluated.

A. Dataset

We utilize real-world operational data collected from a vehicle manufacturing plant in the United States. The dataset spans approximately four months, from November 20, 2024, to March 20, 2025. To ensure consistent production conditions, we exclude weekends and public holidays, resulting in 79 working days of valid logs. During this period, more than 80 AGVs operated concurrently across the plant, communicating through over 40 Wi-Fi[1] access points (APs).

For model training, we use AGV-side telemetry sampled every 2 seconds. Each record contains the currently connected AP, RSSI, latency, bitrate, noise floor, and connected time since association. Furthermore, to enable supervised learning for early *next-AP prediction*, we augment each record with the AP identifier observed 10 seconds later in time. If the AGV remained associated with the same AP, the future label is marked as unchanged; otherwise, it is labeled as the new AP that the AGV roamed to.

[1]This system uses IEEE 802.11ax Wi-Fi 6 for all devices.

Table I presents that the entire dataset consists of 93,083,287 samples, including 101,333 roaming instances. For evaluation, we reserve data from March 2025 as the test set, comprising 16,436,677 samples and 22,167 roaming events.

In summary, the input features and labels used in this study are as follows:

- **Input features:** datetime, latency, connected AP, bitrate, noise floor, RSSI, and connected time since association.
- **Target label:** next AP identifier observed 10 seconds after the current timestamp.

B. Motivation

When a disconnection occurred, we analyzed the dmesg logs of affected AGVs [6] to identify possible causes. Among all events with connectivity loss exceeding 6 seconds[2], a total of 331 cases contained valid dmesg traces. Remarkably, 297 cases (89.46%) were observed immediately before or after a roaming event. This strong correlation indicates that severe disconnections are most likely to occur near roaming transitions, emphasizing the importance of predicting roaming events in advance.

C. Approach

We design a data-driven framework for proactive roaming prediction based on sequence learning. The proposed model combines a Long Short-Term Memory (LSTM) network [7] and a Dense layer to predict the next AP that an AGV will connect to. The model takes as input a sequence of wireless features and outputs a probability distribution over all candidate APs. The AP with the highest probability is selected as the predicted next AP.

To evaluate the prediction confidence, we compute the entropy using probability distribution. In our design, high-entropy outputs often coincide with unstable wireless conditions or impending roaming transitions. Therefore, we use entropy as an indicator of potential roaming events: when entropy exceeds a defined threshold, the model flags a high likelihood of roaming occurrence. Based on this idea, when the entropy exceeds a threshold of 0.5, the model flags the current state as a high-likelihood roaming condition.

IV. Evaluation

In this section, we describe the experimental environment and present the results of our roaming prediction evaluation.

A. Evaluation Setup

All experiments are conducted on an Intel i5-14600KF CPU, an NVIDIA GeForce RTX 4070 SUPER (12 GB VRAM), and 48 GB RAM, running Ubuntu 20.04.6 LTS with Python 3.12, CUDA 12.6, and TensorFlow Keras 3.10.0 [8]. The input window size was set to 15 time steps. Numerical features (RSSI, latency, bitrate, noise floor, and connected time) are min–max scaled, while categorical features (current AP) were

[2]6 seconds is the threshold and a requirement set by the application to distinguish between a temporary disconnection and a long-term disconnection

Latency(ms)	Curr AP	Prev AP	Bitrate(Mbps)	RSSI(dBm)	Connect time(ms)	Target AP	Predicted AP	Entropy
2.746	AP08	AP24	68.8	-69	340	AP08	AP08, AP19	0.9919
3.695	AP08	AP24	68.8	-70	342	AP08	AP08, AP19	1.2271
10.659	AP08	AP24	74.8	-69	344	AP08	AP08, AP19, AP24	1.2984
3.781	AP08	AP24	74.8	-69	346	AP24	AP08, AP19, AP24	1.396
3.074	AP08	AP24	86	-70	348	AP24	AP08, AP19, AP24, AP25	1.4486
5.797	AP08	AP24	86	-71	350	AP24	AP08, AP19, AP24, AP25	1.4968
5.935	AP08	AP24	74.8	-71	352	AP24	AP08, AP19, AP24, AP25	1.5363
3.245	AP08	AP24	74.8	-72	354	AP24	AP08, AP19, AP24, AP25	1.6259
11.727	AP24	AP08	149.7	-59	0	AP24	AP24	0.2596
3.120	AP24	AP08	149.7	-57	2	AP24	AP24	0.0271

TABLE II: Example of test data and predict result

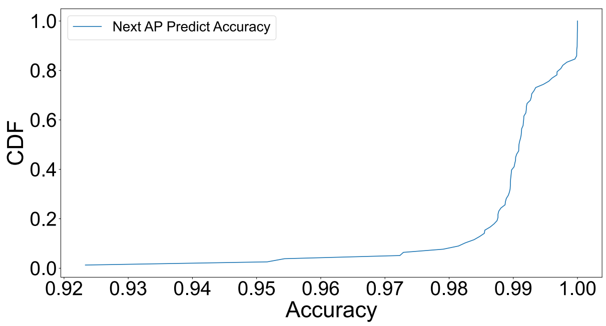


Fig. 2: CDF of next-AP prediction accuracy for all timesteps, including non-roaming cases.

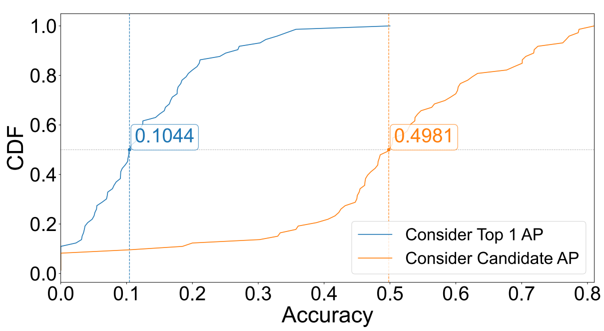


Fig. 3: CDF of roaming detection rate. The blue curve represents detection based on the predicted next AP differing from the current AP, while the orange curve shows detection based on prediction uncertainty (Entropy > 0.5).

one-hot encoded. The model consists of a single LSTM layer with 32 hidden units followed by a Dense layer with softmax over all AP classes. We optimize the network with Adam using sparse categorical cross-entropy, training for 100 epochs with a batch size of 128 and a validation split of 0.2.

B. Evaluation Results

We evaluate the proposed framework using AGV–AP traces collected from the operational automobile factory. Our evaluation focuses on three key dimensions: overall next-AP prediction performance, roaming-specific prediction accuracy, and entropy-based identification of imminent roaming.

Table II shows an example of input features, ground-truth labels, and model outputs for a roaming event. The first six columns (Latency through Connected Time) represent AGV-side input features. The *target AP* column denotes the true next AP observed 10 seconds later, while the *predicted AP* column lists APs with the highest predicted probabilities. The *Entropy* column indicates uncertainty in the model's output distribution. As illustrated, entropy rises before roaming and decreases once the AP association stabilizes, reflecting its relationship with prediction uncertainty.

Next-AP Prediction (Fig. 2) The model shows high accuracy in predicting the next AP association. However, this performance is somewhat inflated because roaming events are relatively rare compared with stable connections. In other words, the dataset is dominated by non-roaming samples, where the AGV continues to stay connected to the same AP, leading to naturally higher accuracy values.

Roaming Prediction Accuracy (Fig. 3) To better understand model behavior under roaming scenarios, we evaluate the prediction accuracy only during actual roaming transitions. The blue curve(top-1 accuracy) represents cases where the predicted next AP exactly matches the true AP, while the orange curve corresponds to candidate-based accuracy, where the true AP is included among candidates whose predicted probability is greater than 5%. The top-1 accuracy drops significantly in roaming conditions. In contrast, the candidate-based metric achieves considerably higher accuracy, showing that the model can still capture multiple plausible AP transitions. Approximately 10% of the samples exhibit 0% accuracy because they contain very few roaming occurrences during the observation window.

Roaming Detection Using Entropy (Fig. 4) Finally, we assess the ability to proactively detect roaming events. Roaming is considered detected when either (i) the predicted next AP differs from the current AP (blue curve) or (ii) the entropy of the predicted probability distribution exceeds 0.5 (orange curve). The entropy-based approach detects roaming earlier and more smoothly, indicating that prediction uncertainty serves as a reliable signal of imminent handover and link instability.

Temporal Visualization of Roaming Prediction (Fig. 5) Fig. 5(a) evaluates *where* the device will roam by comparing the ground-truth next AP with the model's predicted next AP at each time step. Vertical green and red markers indi-

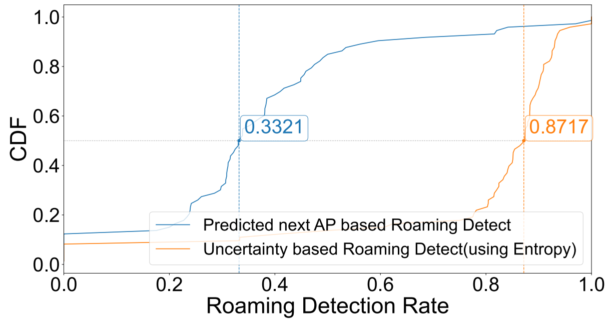
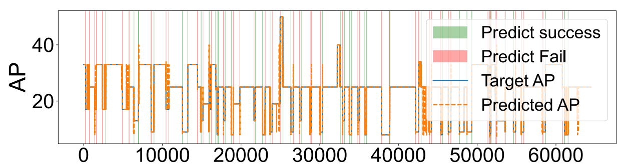
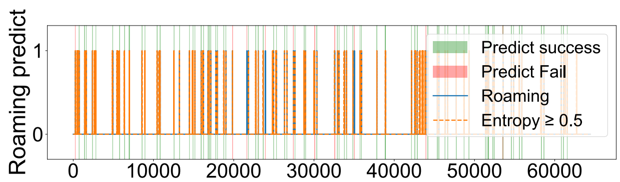


Fig. 4: CDF of next-AP prediction accuracy under roaming conditions. The blue curve indicates the accuracy of correctly predicting the next AP, whereas the orange curve represents cases where the true next AP appears among candidate APs with prediction probability above 5%.



(a) Comparison of predicted AP and target AP over time. Green and red bars denote correct and incorrect predictions.



(b) Time-series roaming prediction and entropy-based detection. The orange dashed lines indicate entropy ≥ 0.5.

Fig. 5: Qualitative visualization of roaming behavior along the AGV trajectory. The upper plot shows next-AP prediction performance at actual roaming transitions, while the lower plot illustrates entropy-based detection of roaming occurrence (roaming vs. non-roaming) in real time.

cate, respectively, correct and incorrect next-AP predictions at actual roaming instances. As shown in Fig. 5, the next-AP prediction accuracy during roaming remains around 50%, and the same trend is visible in the upper plot: the model correctly identifies roughly half of the actual roaming transitions, while the remaining cases exhibit deviations due to the inherent uncertainty of wireless conditions during handover.

Fig. 5(b) evaluates *whether* a roaming event will occur using entropy-based detection. Here, the orange dashed trace marks time instants where the entropy exceeds the roaming threshold ($H \geq 0.5$), and the blue markers denote ground-truth roaming occurrences. This binary signal does not predict the target AP; instead, it flags imminent roaming (roaming vs. non-roaming). The dense activations around transition points illustrate that entropy provides early warnings of link instability.

V. Conclusion

This paper presented a data-driven framework for proactive roaming prediction and connection stability estimation in industrial Wi-Fi networks. Using real-world datasets collected from AGVs and Wi-Fi access points in an automobile manufacturing plant, we analyzed disconnection phenomena and found that severe connectivity losses frequently occur near roaming transitions. However, anticipating roaming events is difficult for human operators due to rapidly changing wireless conditions, making automated prediction essential for reacting before link failure occurs. To address this, we proposed an LSTM-based next-AP prediction model that leverages time-series signal patterns and introduces an entropy-based metric to quantify roaming likelihood. Experimental results show that the proposed framework accurately predicts roaming events and unstable link states in advance. Entropy-based detection achieved up to 95% accuracy for AGVs with frequent roaming and a median accuracy of 87% across all AGVs, significantly improving industrial Wi-Fi reliability.

As future work, we plan to enable earlier diagnosis of disconnect after roaming and develop automated responses that can prevent disconnects before they impact AGV operation.

Acknowledgement

We thank the Hyundai Motor Group's Manufacturing SW Platform R&D Team within the Manufacturing Solution Division for providing us with real AGV–AP Wi-Fi communication logs from their car manufacturing factory in operation.

References

[1] M. Choi, J. H. Shin, and J. Paek, "Industrial Wi-Fi Traffic Forecasting Using Diverse Time-Series Models," in *ICTC Workshop on Big Data, CPS, and 5G&6G Communication Networks (IWBCN)*. IEEE, Oct 2025, pp. 670–673.

[2] S. Bang, J. H. Shin, and J. Paek, "Anomaly Detection of AGV Communication Failures in Factory Wi-Fi Network," in *ICTC Workshop on Big Data, CPS, and 5G&6G Communication Networks (IWBCN)*. IEEE, Oct 2025, pp. 666–669.

[3] M. Manalastas, M. U. B. Farooq, S. M. A. Zaidi, A. Abu-Dayya, and A. Imran, "A data-driven framework for inter-frequency handover failure prediction and mitigation," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 6, pp. 6158–6172, 2022.

[4] J. P. S. H. Lima, A. M. de Medeiros, E. P. de Aguiar, E. F. Silva, V. A. de Sousa, M. L. Nunes, and A. L. Reis, "Deep learning-based handover prediction for 5g and beyond networks," in *IEEE International Conference on Communications (ICC)*, 2023, pp. 3468–3473.

[5] M. Dzaferagic, B. Missi Xavier, D. Collins, V. D'Onofrio, M. Martinello, and M. Ruffini, "Ml-based handover prediction over a real o-ran deployment using ran intelligent controller," *IEEE Transactions on Network and Service Management*, vol. 22, no. 1, pp. 635–647, 2025.

[6] S. Bae, J. H. Shin, and J. Paek, "Diagnosing AGV Wi-Fi Disconnection via dmesg Logs: A Real-World Factory Case Study," in *ICTC Workshop on Big Data, CPS, and 5G&6G Communication Networks (IWBCN)*. IEEE, Oct 2025, pp. 863–866.

[7] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[8] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Józefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "Tensorflow: Large-scale machine learning on heterogeneous systems," https://www.tensorflow.org/, 2015.