SENSORCOMM 2025

The Nineteenth International Conference on Sensor Technologies and Applications

ISBN: 978-1-68558-304-0

October 26th - 30th, 2025

Barcelona, Spain

SENSORCOMM 2025 Editors

Jaime Lloret, Universitat Politècnica de València, Spain

SENSORCOMM 2025

Forward

The Nineteenth International Conference on Sensor Technologies and Applications (SENSORCOMM 2025), held between October 26th, 2025, and October 30th, 2025, in Barcelona, Spain, was a multi-track event covering related topics on theory and practice on wired and wireless sensors and sensor networks.

Sensors and sensor networks have become a highly active research area because of their potential of providing diverse services to broad range of applications, not only in science and engineering, but equally importantly on issues related to critical infrastructure protection and security, health care, the environment, energy, food safety, and the potential impact on the quality of all areas of life.

Sensor networks and sensor-based systems support many applications today on the ground. Underwater operations and applications are quite limited by comparison. Most applications refer to remotely controlled submersibles and wide-area data collection systems at a coarse granularity.

In wireless sensor and micro-sensor networks energy consumption is a key factor for the sensor lifetime and accuracy of information. Protocols and mechanisms have been proposed for energy optimization considering various communication factors and types of applications. Conserving energy and optimizing energy consumption are challenges in wireless sensor networks, requiring energy-adaptive protocols, self-organization, and balanced forwarding mechanisms.

We take the opportunity to warmly thank all the members of the SENSORCOMM 2025 technical program committee, as well as all the reviewers. The creation of such a high-quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and effort to contribute to SENSORCOMM 2025. We truly believe that, thanks to all these efforts, the final conference program consisted of top-quality contributions. We also thank the members of the SENSORCOMM 2025 organizing committee for their help in handling the logistics of this event.

We hope that SENSORCOMM 2025 was a successful international forum for the exchange of ideas and results between academia and industry for the promotion of progress in the field of sensor technologies and applications.

SENSORCOMM 2025 Chairs

SENSORCOMM 2025 Steering Committee
Anne-Lena Kampen, Western Norway University of Applied Sciences, Norge

SENSORCOMM 2025 Publicity Chairs
Lorena Parra Boronat, Universidad Politécnica de Madrid, Spain
Laura Garcia, Universidad Politécnica de Cartagena, Spain

SENSORCOMM 2025
Committee

SENSORCOMM 2025 Steering Committee

Anne-Lena Kampen, Western Norway University of Applied Sciences, Norge

SENSORCOMM 2025 Publicity Chairs

Lorena Parra Boronat, Universidad Politécnica de Madrid, Spain
Laura Garcia, Universidad Politécnica de Cartagena, Spain

SENSORCOMM 2025 Technical Program Committee

Majid Bayani Abbasy, National University of Costa Rica, Costa Rica
Younes Adriouch, Mohammed Five University of Rabat (UM5R), Agdal, Rabat, Morocco
Yassine Al-Amrani, Abdelmalek Essaadi University, Tetuan, Morocco
Amin Al-Habaibeh, Nottingham Trent University, UK
Jesús B. Alonso Hernández, Institute for Technological Development and Innovation in Communications (IDeTIC) | University of Las Palmas de Gran Canaria (ULPGC), Spain
Mário Alves, Politécnico do Porto - School of Engineering (ISEP), Portugal
Thierry Antoine-Santoni, University of Corsica, France
Kazutami Arimoto, Okayama Prefectural University, Japan
Anish Arora, The Ohio State University / The Samraksh Company, USA
Andy Augousti, Kingston University London, UK
Karim Baïna, Université Mohammed V de Rabat, Morocco
Michel Bakni, ESTIA Recherche, Bidart, France
Hind Bangui, Masaryk University, Czech Republic
Michail J. Beliatis, Aarhus University, Denmark
Boutaina Benhmimou, Mohammed Five University of Rabat (UM5R), Agdal, Rabat, Morocco
Reda Benkhouya, Ibn Tofail University, Morocco
Abdelmadjid Bouabdallah, University of Technology of Compiegne, France
Souheila Bouam, University of Batna 2, Algeria
An Braeken, Vrije Universiteit België, Belgium
Seddik Bri, Higher School of Technology - Moulay Ismail University, Meknes, Morocco
Alessandro Brighente, University of Padova, Italy
Juan-Vicente Capella-Hernández, Universitat Politècnica de València, Spain
Domenico Caputo, University of Rome "La Sapienza", Italy
Oscar Carrasco Quilis, Casa Systems Inc., Valencia, Spain
Vítor Carvalho**,** 2Ai Lab- School of Technology - IPCA / Algoritmi Research Center - Minho University, Portugal
Punyasha Chatterjee, Jadavpur University Salt Lake Campus, Kolkata, India
Mario Cifrek, University of Zagreb, Croatia
Luis J. de la Cruz Llopis, Universitat Politècnica de Catalunya, Spain
Baban P. Dhonge, Indira Gandhi Center for Atomic Research, Kalpakkam, India
Nabil El Akkad, ENSA | Sidi Mohamed Ben Abdellah University, Morocco

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

A Group-Based Access Coordination Scheme for Low-Latency and High-Throughput IoT over IEEE 802.11ah

Seung-ha Jee[1], Yu-ran Jeon[1], Il-Gu Lee*[1,2]
[1]Dept. of Convergence Security Engineering
Sungshin Women's University
[1,2]Dept. of Future Convergence Technology Engineering
Sungshin Women's University
Seoul, Korea
e-mail: {shjee2001, cseyrj}@gmail.com, iglee@sungshin.ac.kr
*Corresponding author

***Abstract*—As network infrastructure expands, the Internet of Things (IoT) demands extensive coverage, substantial throughput capacity, and real-time performance. The 802.11ah standard's raw mechanism was proposed to enhance efficiency in high-density environments. However, its implementation in latency-sensitive IoT network environments is constrained by inherent limitations. In this paper, we propose the Secure Restricted Access Window Based Group Coordination (SRAW-GC) technique, which prioritizes the processing of latency-sensitive traffic while aggregating high-throughput traffic for transmission, addressing both throughput and latency requirements. Experimental findings indicate that SRAW-GC improves performance metrics by 29.86% in throughput, 19.3% in latency, and 48.23% in energy consumption compared to conventional mechanisms. The proposed technique can ensure better availability in IoT network environments than the conventional RAW technique.**

Keywords- IoT (Internet of Things) Network; IEEE 802.11ah; Network Efficiency.

I. INTRODUCTION

The proliferation of Internet of Things (IoT) devices, along with the rapid increase in network traffic among these devices, has led to the development of technologies to address the diverse security and performance needs of IoT networks. Specifically, IoT networks require wide coverage, high throughput, and real-time performance. Therefore, developing wireless communication technologies that meet these performance criteria is crucial for advancing IoT technology [1]. The types of traffic exchanged between devices in IoT networks vary based on the purpose of transmission and network conditions. Each type of traffic has distinct requirements, such as latency, throughput, and reliability. To efficiently manage the vast amount of diverse traffic, it is essential to understand the characteristics of each traffic type and apply appropriate transmission and processing methods [2].

IoT communication technologies can be broadly divided into two categories: Wireless Personal Area Network (WPAN) technologies and Low-Power Wide Area Network (LPWAN) technologies. WPANs, exemplified by ZigBee and Bluetooth Low Energy (BLE), facilitate medium-level data transmission rates over short distances. In contrast, LPWANs, such as Long Range (LoRa), enable long-distance transmission at low data rates. Consequently, WPAN and LPWAN exhibit distinct advantages and disadvantages regarding throughput and coverage [3]. IEEE 802.11ah, also known as Wi-Fi HaLow, is a technique designed to overcome the limitations of conventional IoT communication technologies. It is gaining attention for its potential to enhance throughput, coverage, and power efficiency in IoT networks. IEEE 802.11ah provides a long-range, low-power, low-speed alternative to traditional Wi-Fi, supporting approximately 8,000 nodes per AP (Access Point) within a 1-2 km service radius. Unlike other IoT connectivity technologies, it does not require the implementation of separate controllers, hubs, or gateways, ensuring cost effectiveness and substantial scalability.

IEEE 802.11ah incorporates several pivotal features within the MAC layer, offering functionalities such as fast authentication and association, Restricted Access Window (RAW), Traffic Indication Map (TIM) segmentation, and Target Wake Time (TWT). Among these technologies, the RAW technique is noteworthy for providing a distributed channel access method that can enhance the efficiency of densely packed, energy-constrained Stations (STAs) and can be flexibly applied to varying network conditions. However, the conventional RAW technique groups STAs based on their required throughput levels and allocates time slots to ensure high network throughput, which limits its applicability in real-time IoT environments that require latency-sensitive traffic [4]. Consequently, this study proposes the Secure Restricted Access Window Based Group Coordination (SRAW-GC) technique, which considers both traffic throughput and latency requirements, prioritizes the processing of latency-sensitive traffic, and aggregates traffic requiring high throughput for transmission.

The contributions of this study are as follows:

- An SRAW-GC mechanism is proposed to address latency and throughput requirements. This mechanism utilizes a grouping approach to organize traffic and allocate time slots based on the characteristics of each group.
- The proposed SRAW-GC mechanism prioritizes time slots for latency-sensitive traffic while aggregating traffic that requires high throughput for transmission. Its efficacy is demonstrated by improvements in both network latency and throughput.

- An evaluation of the proposed SRAW-GC across various network environments shows a performance enhancement of 29.86% in throughput, 19.3% in latency, and 48.23% in energy consumption compared to the conventional RAW scheme.

The structure of this paper is as follows: Section 2 analyzes previous studies related to technologies introduced in 802.11ah, and Section 3 proposes the SRAW-GC technique to address the limitations of the RAW technique in 802.11ah. Section 4 evaluates and verifies the performance of the proposed SRAW-GC technique, while Section 5 concludes the paper.

II. Related Work

The IEEE 802.11ah standard introduces technologies such as TIM segmentation, TWT, and RAW to enable efficient channel access for resource-constrained STAs in dense IoT networks. TIM segmentation is a power-saving technique that divides TIM information in beacons into groups, allowing STAs to activate only during their corresponding groups, thereby improving energy efficiency. TWT is a reservation-based communication technique that facilitates the negotiation of activation timings between STAs and APs, enabling communication during designated time slots while preserving power-saving mode during other periods, thus significantly enhancing power efficiency. The RAW technique groups STAs to access the channel only during specified time slots, reducing network collisions and improving scalability. Table 1 summarizes existing studies relevant to TIM segmentation, TWT, and RAW technologies in 802.11ah.

TABLE I. Comparison of previous studies

Feature	Ref.	Contribution	Limitation
TIM segmentation	[5]	· The proposed network architecture enhances scalability by incorporating control loops and monitoring sensors into the network infrastructure.	· It has been demonstrated that if the beacon cycle is not optimized, there will be an increase in throughput and energy consumption.
	[6]	· The proposal entails the implementation of a link-layer mechanism, comprising downlink TIM and uplink RAW groups, to mitigate energy consumption.	· It is challenging to verify performance on real-time networks because link latency is not taken into account.
	[7]	· The proposed methodology involves implementing a multifaceted approach, integrating the utilization of RAW and TWT, to enhance network energy efficiency.	· It has been demonstrated that there is an increase in latency when using RAW and TWT in conjunction with one another.
TWT	[8]	· The proposal entails the implementation of a novel channel access methodology for STAs within a network environment characterized by the coexistence of RAW STA and TWT STA configurations.	· TWT transmission is contingent upon the availability of empty RAW slots, a factor that compromises energy efficiency and engenders augmented latency.
RAW	[9]	· The proposal entails implementing a RAW mechanism to identify concealed terminals and organize STAs into designated groups.	· It has been demonstrated that an increase in network latency is associated with a failure to consider traffic latency requirements.
	[10]	· The proposal entails the implementation of a machine learning-based mechanism within the RAW framework to facilitate the process of grouping and the subsequent control of channel access.	· The method of determining channel access depends on the number of collisions between STAs.

Seferagić et al. [5] propose a network that hosts a control loop to regulate its dynamic status. Additionally, the network includes monitoring sensors that periodically transmit measurement results. The purpose of this system is to enhance the scalability of IEEE 802.11ah. The proposed method utilizes a control loop to dynamically adjust the beacon

interval, ensuring compliance with latency requirements. However, it is important to note a limitation inherent to this approach: optimization of the beacon cycle is necessary since throughput and energy consumption increase with the beacon cycle. Bel et al. [6] propose a link-layer mechanism consisting of downlink TIM and uplink RAW groups to reduce energy consumption. The study demonstrated that energy efficiency can be enhanced to prolong the battery life of sensor nodes. However, it does not consider delays in uplink and downlink communications, complicating the verification of performance in real-time networks.

Santi et al. [7] call for research to enhance network energy efficiency by utilizing RAW and TWT technologies. Energy consumption calculations under various conditions demonstrate that the proposed method significantly improves energy efficiency. However, it is important to note the limitation of this method: a substantial increase in latency, which makes it difficult to apply in real-time networks. Santi et al. [8] analyze energy consumption rates in scenarios where RAW STAs and TWT STAs coexist, proposing a channel access method for STAs designed to enhance energy efficiency. The implementation of the IEEE 802.11ah TWT technique in an NS-3 environment has demonstrated the degradation of energy efficiency for TWT STAs caused by RAW STAs. Furthermore, a scheduling method has been proposed that allows TWT STAs to reserve empty RAW slots. However, this method raises concerns regarding energy consumption and latency when RAW slots are occupied, complicating TWT transmission. Similarly, the TIM technique has been observed to have increased beacon overhead. Additionally, TWT has the limitation of being challenged to apply in dynamically changing networks due to its reservation-based approach, which requires real-time performance.

Mondal et al. [9] proposed the HTAG (Hidden Terminal Aware Grouping) technique to address the hidden terminal problem that arises when employing the RAW technique in IEEE 802.11ah-based high-density IoT networks. The system detects hidden terminal devices through the Neighbor Detection Table (NDT) and engages in the grouping of these hidden nodes. However, this method has a limitation: it does not consider traffic delay requirements, which leads to increased network delays. Mahesh et al. [10] propose a machine learning mechanism to group STAs and control channel access for each STA group. After calculating the collision count for each RAW group using an unsupervised learning model, the AP adjusts the beacon interval based on this information and broadcasts it to the STAs. However, this method bases the channel access determination solely on the collision count between STAs, complicating the fulfillment of the network's real-time requirements, as ensuring smooth channel access for low-latency traffic is challenging. The conventional RAW technique groups STAs solely based on network throughput when managing traffic. However, low latency is essential in IoT environments with densely packed and interconnected sensors. This necessitates a mechanism that considers both latency and throughput when grouping traffic.

III. PROPOSED SCHEME

This section delineates the methodologies of SRAW-GC for facilitating efficient data transmission in dense network environments with IoT STAs. As illustrated in Figure 1, high-throughput STAs are grouped within the same Basic Service Set (BSS), and each STA transmits data to the AP.

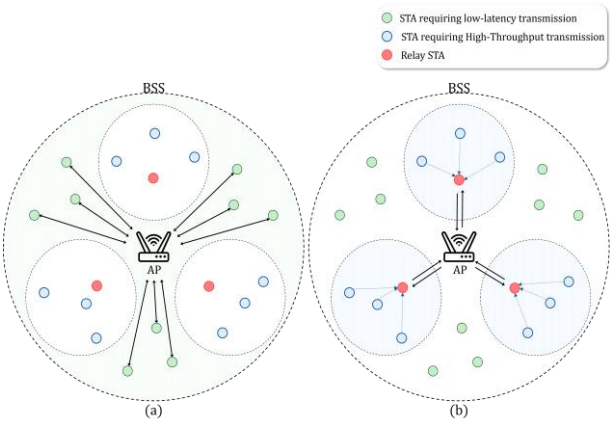


Figure 1. This is how STAs that require low latency and STAs that require high bandwidth transmit data to AP. (a) shows how low-latency STA communicates with AP, and (b) shows how high-bandwidth STA communicates with AP.

SRAW-GC has developed a methodology for classifying STAs at the application level. These STAs are divided into two distinct categories: those requiring low-latency transmission and those requiring high-throughput transmission. STAs that require low-latency transmit their data, known as MAC Protocol Data Unit (MPDU), to the AP individually. In contrast, STAs that require high throughput prioritize the transmission of MPDU to the Relay STA. The selection of relay STAs is determined by the Modulation and Coding Scheme (MCS) index, which provides a comprehensive assessment of the channel quality between the AP and the STAs. The selected Relay STA receives MPDUs from nearby STAs requiring high throughput and aggregates the data using the Aggregated MAC Protocol Data Unit (A-MPDU) method. Low-latency STAs generate data irregularly and require low-latency transmission rather than high throughput. Consequently, aggregating data and transmitting it in batches, as high-throughput STAs do, does not satisfy their low-latency transmission requirements. Instead, it is more efficient for them to transmit data immediately as the need arises. High-throughput transmission requires a throughput that exceeds the transmission delay rate. The A-MPDU method, as outlined in the extant 802.11n specification for high throughput, meets these requirements by allowing data to be transmitted in batches.

As illustrated in Figure 2, the operation methods for low-latency and high-throughput slots are sub-slots within the RAW slot.

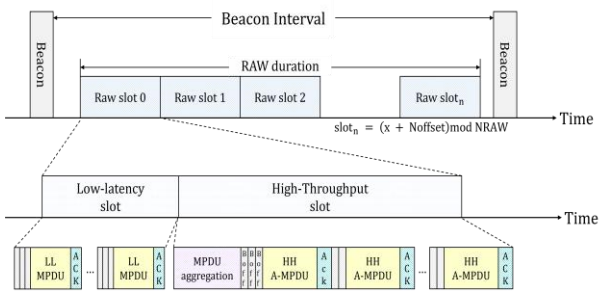


Figure 2. SRAW-GC technique showing a RAW slot divided into a low-latency sub-slot for individual transmissions and a high-throughput sub-slot for aggregated A-MPDU transmissions.

The SRAW-GC technique has been designed to be compatible with the 802.11ah RAW technique, and STAs are assigned to the n_{th} RAW slot according to (1) to access the channel:

$$slot_n = (x + N_{offset}) mod\ NRAW \quad (1)$$

In (1), $slot_n$ is the index number of the RAW slot allocated to the STA, and x represents either the AID position index or the AID of the STA. N_{offset} is the offset value expressed with the lower two bytes of the FCS field in the SIG beacon frame. $NRAW$ is the total number of slots in RAW.

The RAW slot is divided into two sub-slots: a low-latency sub-slot and a high-throughput sub-slot. In the low-latency sub-slot, STAs requiring low-latency transmission compete for channel access with the AP to transmit MPDU data. After this, the STA anticipates the designated back-off time and transmits its MPDU to the AP. It has been shown that since the low-latency sub-slot is prioritized within the RAW slot, STAs needing low-latency transmission can effectively address transmission delays caused by high-throughput STAs with long channel occupancy times. This situation is analogous to the existing RAW mechanism. Once the low-latency sub-slot concludes, a transition to the high-throughput sub-slot occurs. In this slot, the AP first selects relay STAs within the BSS. These designated relay STAs receive MPDUs from high-throughput STAs nearby and aggregate the data using the A-MPDU method. The aggregated data is then transmitted to the AP through channel competition among the relay STAs. The improved transmission efficiency observed in this scenario can be attributed to the superior channel quality and status maintained by the relay STAs with the AP, compared to situations where individual STAs transmit data directly to the AP. Furthermore, since only relay STAs are responsible for transmitting data to the AP, the probability of competition and subsequent collisions is significantly reduced, thereby enhancing the efficiency of high-throughput transmission.

IV. EVALUATION AND ANALYSIS

A. Evaluation Environment

This section delineates the experimental environment for evaluating the performance of the proposed SRAW-GC technique. The conventional model selected the 802.11ah RAW mechanism for performance comparison with that of SRAW-GC [9][10].

The experiment aimed to assess the performance of both the proposed and conventional models in a network environment based on the 802.11ah standard. The experiment was conducted in a Python 3.12 environment. The simulation environment was set up with one AP and 2,000 STAs within a single BSS. Performance evaluations were conducted for each scenario, considering the number of STAs, the ratio of low-latency STAs to high-throughput STAs, and the collision probability among STAs. The evaluation metrics used included throughput, latency, and energy consumption. The variables employed in the equations are detailed in Table 2, and throughput was calculated according to (2):

TABLE II. VARIABLES IN FORMULAS

Parameter	Meaning
D_i	Data successfully transmitted by i_{th} STA (byte)
T_{RAW}	Total RAW duration (ms)
P_{base}	0.1 W (basic transmission power)
P_{idle}	0.02 W (idle power)
$T_{active,n}$	Active transmission time of nth STA (ms)
$T_{idle,n}$	Idle time of nth STA (ms)
$E_{STA}(n)$	Energy consumption of node i (joules)
E_{total}	Sum of energy consumed by all N nodes plus energy consumed by the access point (joules)

$$Throughput(kbps) = \frac{\sum_{i=1}^{n} D_i(bytes) \times 8}{T_{RAW}(ms)} \quad (2)$$

To express throughput in kbps, the number of bytes successfully transmitted was multiplied by 8 to convert the unit to bits. This equation represents the successful transmission of data to the AP during the total RAW duration, T_{RAW}. Latency was determined via (3):

$$Latency(ms) = T_{Data\ transmission} + (T_{backoff} \times N_{backoff}) \quad (3)$$

Latency is calculated as the sum of data transmission time and back-off time. Subsequently, energy consumption could be predicted using (4):

$$E_{total}(J) = \sum_{n=1}^{N} E_{STA}(n) \quad (4)$$

The total energy consumption of BSS is defined as the sum of the energy consumption of all STAs and APs, as shown in Equation (4). The energy consumption of each STA can be calculated according to (5):

$$E_{STA}(J) = P_{base} \times \left(T_{active,n}/1000\right) + P_{idle} \times \left(T_{idle,n}/1000\right) \quad (5)$$

Within the same BSS, throughput, latency, and energy consumption were evaluated for each number of STAs, the ratio of low-latency STAs to high-throughput STAs, and the collision ratio. The simulation was repeated a total of 10,000 times.

B. Evaluation Results and Analysis

As illustrated in Figure 3, a comparative analysis was conducted to assess the throughput, latency, and energy consumption of SRAW-GC and conventional models in relation to the number of STAs.

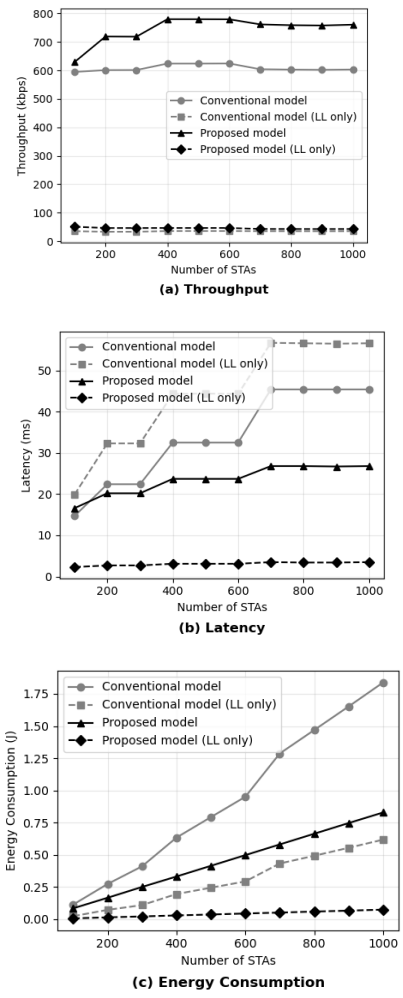


Figure 3. Performance evaluation results for (a) throughput, (b) latency, and (c) energy consumption according to the number of STAs.

As illustrated in Figure 3, the performance of the proposed model was compared to that of the conventional model in a network scenario where the ratio of STAs requiring low-latency transmission to those requiring high-throughput transmission is set at 1:1, and the collision probability is set at 0.3. A thorough analysis revealed that while throughput increased for both models as the number of STAs grew, they reached similar throughput levels starting from 400 STAs. This phenomenon can be attributed to the gradual increase in the number of STAs allocated to each RAW slot, which reduces the amount of data successfully transmitted within the slot. Consequently, both models achieved maximum throughput at 600 STAs, with the proposed model reaching 779.7 kbps and the conventional model achieving 624.4 kbps. When comparing the throughput of low-latency STAs alone, the proposed model achieved 46.3 kbps. In comparison, the conventional model attained 35.5 kbps at 600 STAs, indicating a 30.3% improvement in throughput for the proposed model. As the number of STAs increased, the average latency increased for both models. The conventional model exhibited an increase in latency from 14.7 ms (milliseconds) to 45.4 ms as the number of STAs increased from 100 to 1,000, while the proposed model demonstrated an increase from 16.6 ms to 26.8 ms. At 100 STAs, the proposed model had a latency that was 1.9 ms higher; however, as the latency of the conventional model increased sharply, the proposed model improved latency by up to 41.1% when the number of STAs reached 900. When examining the latency of low-latency STAs specifically, the proposed model showed a significant enhancement. This improvement results from the proposed model allocating sub-slots to prioritize the transmission of low-latency STAs within the RAW slot. In contrast, high-throughput STAs and low-latency STAs coexist within the conventional model, leading to competition within the same RAW slot. This competition increases latency for low-latency STAs due to the long channel occupancy times of high-throughput STAs.

As illustrated in Figure 4, an increase in the ratio of low-latency STAs led to a decline in throughput for both the proposed and conventional models. This phenomenon occurs because the data size transmitted by low-latency STAs is smaller than that of high-throughput STAs, resulting in an overall decrease in throughput. The proposed model demonstrated superior performance for both all STAs and low-latency STAs. In the initial phases, when the proportion of low-latency STAs was minimal, the proposed model exhibited higher latency compared to the conventional model. This can be attributed to the allocation of a minimum low-latency sub-slot within the RAW slot by the proposed model, which prioritizes the transmission of low-latency STAs. Consequently, even in the absence of low-latency STAs to transmit, high-throughput STAs must wait. However, as the proportion of low-latency STAs increased to an 8:1 ratio with high-throughput STAs, the latency of the proposed model exhibited a 62.2% improvement compared to the conventional model. This finding substantiates the efficacy of the proposed mechanism in reducing latency in network environments characterized by a high density of low-latency STAs.

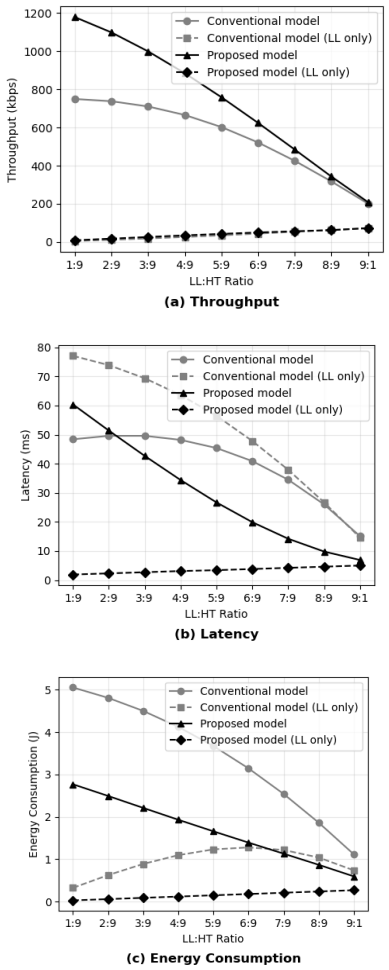


Figure 4. Performance evaluation results for (a) throughput, (b) latency, and (c) energy consumption based on the ratio of low-latency STAs to high-throughput STAs.

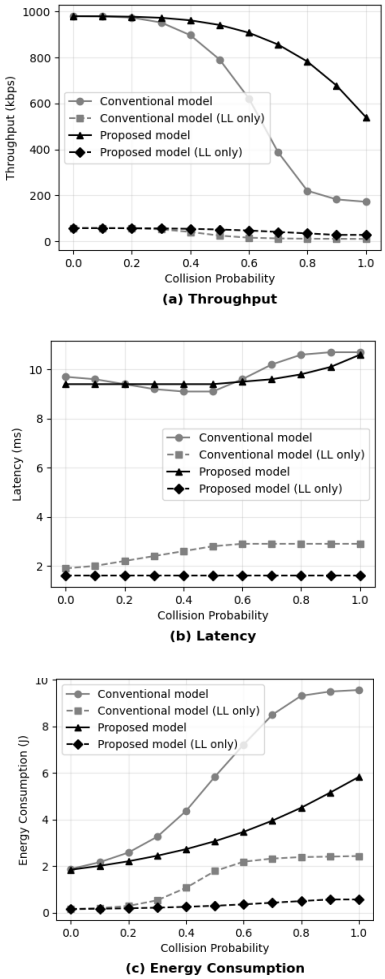


Figure 5. Performance evaluation results for (a) throughput, (b) latency, and (c) energy con-sumption based on collision probability.

A comparison of the latency exhibited by low-latency STAs reveals that the proposed model exhibited an average reduction of 93.4% in latency compared to the conventional model. As the proportion of low-latency STAs increased, both models exhibited a decline in energy consumption. It has been demonstrated that low-latency STAs maintain an active state for a shorter period than high-throughput STAs due to their shorter channel occupancy time, resulting in a reduced energy consumption rate. Nevertheless, the proposed model demonstrated a notable enhancement in energy efficiency, achieving an average savings of 51.2% compared to the conventional model. Figure 5 compares the throughput, latency, and energy consumption of SRAW-GC and the conventional model based on collision probability in an environment with 2,000 STAs. Figure 5 compares the throughput, latency, and energy consumption of SRAW-GC and the conventional model based on collision probability in an environment with 2,000 STAs.

As illustrated in Figure 5, an increase in collision probability led to a decline in throughput for both the proposed and conventional models.

The conventional model showed a sharp decrease in throughput, dropping from 980.1 kbps to 172.2 kbps (82.43%). In contrast, the proposed model exhibited a more modest reduction, from 980.1 kbps to 540.8 kbps (44.82%), demonstrating its effectiveness in maintaining throughput performance. The proposed technique employs the A-MPDU mechanism to transmit aggregated data from relay STAs with optimal channel conditions and transmission efficiency to the AP among high-throughput STAs. This allows the proposed model to sustain a higher throughput than the conventional technique. As the collision probability increased, the latency of both models also rose; however, the proposed model maintained a lower latency. This is due to the classification of traffic into low-latency and high-throughput slots by dividing the RWA slot, which mitigates collisions within the slot compared to the conventional RAW technique. Conversely, the conventional model competes for channel occupancy across all traffic within the same RAW slot, resulting in increased latency for STAs with low-latency requirements. A comparison of the proposed and conventional models in terms of latency for low-latency STAs shows that the conventional model suffers from

increased latency as the collision probability increases. In contrast, the proposed model maintained lower latency. This can be attributed to the fact that, even in scenarios with collisions, the competition among low-latency STAs mitigates the latency caused by high-throughput STAs, ensuring system availability and contrasting with the outcomes observed in the conventional method. As the collision probability increased, the retransmission mechanism was triggered, resulting in elevated energy consumption for both models. However, the proposed model demonstrated reduced energy consumption compared to the conventional model. Notably, when the collision probability was set at 0.8, the proposed model showed a significant reduction in energy consumption of 51.62%, underscoring its effectiveness in energy-efficient operations.

V. Conclusion and Future Work

As the proliferation of IoT devices continues to accelerate, the demand for efficient processing of the substantial volume of IoT traffic associated with wireless networks is increasing. While prior studies have focused on enhancing wide coverage and high throughput, it is crucial to recognize the need for advancements in real-time and low-latency data transmission within mission-critical IoT networks. Consequently, the RAW technique of 802.11ah has been proposed as a solution to improve latency. However, this technique groups STAs based on the required throughput level of the traffic and allocates time slots to ensure high network throughput, which limits its applicability to real-time IoT environments that require latency-sensitive traffic. This study proposes an A-MPDU grouping technique based on IEEE 802.11ah RAW to achieve low latency and high throughput. According to the experimental results, SRAW-GC enhances throughput by 29.86%, reduces latency by 19.3%, and decreases energy consumption by 48.23% compared to the conventional model. Furthermore, for STAs with low-latency requirements, the proposed SRAW-GC approach demonstrates improvements of 29.53% in throughput and 74.66% in latency compared to the conventional method. Consequently, the SRAW-GC technique can ensure better availability in IoT network environments than the conventional RAW technique. Future research will determine the optimal values for the low-latency sub-slot and the high-throughput sub-slot within the RAW framework.

Acknowledgment

This work was supported by the Ministry of Trade, Industry and Energy (MOTIE) under the Training Industrial Security Specialists for High-Tech Industry program (RS-2024-00415520), supervised by the Korea Institute for Advancement of Technology (KIAT); the Ministry of Science and ICT (MSIT) under the ICT Challenge and Advanced Network of HRD (ICAN) program (IITP-2022-RS-2022-00156310); and the Information Security Core Technology Development program (RS-2024-00437252), supervised by the Institute of Information & Communication Technology Planning & Evaluation (IITP).

References

[1] Q. V. Khanh, N. V. Hoai, L. D. Manh, A. N. Le, and G. Jeon, "Wireless communication technologies for IoT in 5G: Vision, applications, and challenges," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 3229294, 2022.

[2] L. H. Shen, C. H. Wu, W. C. Su, and K. T. Feng, "Analysis and implementation for traffic-aware channel assignment and contention scheme in LoRa-based IoT networks," *IEEE Internet of Things Journal*, vol. 8, no. 14, pp. 11368–11383, 2021.

[3] L. Tian, S. Santi, A. Seferagić, J. Lan, and J. Famaey, "Wi-Fi HaLow for the Internet of Things: An up-to-date survey on IEEE 802.11ah research," *Journal of Network and Computer Applications*, vol. 182, p. 103036, 2021.

[4] Wi-Fi Alliance, *"Wi-Fi CERTIFIED™ for WMM™—Support for multimedia applications with quality of service in Wi-Fi® networks,"* Austin, TX, USA: Wi-Fi Alliance, 2004.

[5] A. Seferagić, I. Moerman, E. De Poorter, and J. Hoebeke, "Evaluating the suitability of IEEE 802.11ah for low-latency time-critical control loops," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7839–7848, 2019.

[6] A. Bel, T. Adame, and B. Bellalta, "An energy consumption model for IEEE 802.11ah WLANs," *Ad Hoc Networks*, vol. 72, pp. 14–26, 2018.

[7] S. Santi, L. Tian, E. Khorov, and J. Famaey, "Accurate energy modeling and characterization of IEEE 802.11ah RAW and TWT," *Sensors*, vol. 19, no. 11, p. 2614, 2019.

[8] S. Santi, L. Tian, and J. Famaey, "Evaluation of the co-existence of RAW and TWT stations in IEEE 802.11ah using ns-3," in *Proc. 2019 Workshop on Next-Generation Wireless with ns-3 (WNS3)*, 2019.

[9] M. A. Mondal and M. I. Hussain, "Station grouping mechanism using machine learning approach for IEEE 802.11ah," *Ad Hoc Networks*, vol. 149, p. 103238, 2023.

[10] M. Mahesh and V. P. Harigovindan, "Hidden terminal aware grouping scheme for IEEE 802.11ah based dense IoT networks," *Computer Communications*, vol. 191, pp. 161–172, 2022

A Time based Sensor Data Analysis for Pre-Fall Prediction Using Machine and Deep Learning Approaches

Nazia Pathan
Queen Margaret University
Edinburgh, United Kingdom
npathan@qmu.ac.uk

Hongnian Yu
Edinburgh Napier University
Edinburgh, United Kingdom
H.Yu@napier.ac.uk

Michael Vassallo
Royal Bournemouth Hospital
Bournemouth, United Kingdom
michael.vassallo@uhd.nhs.uk

Pelagia Koufaki
Queen Margaret University
Edinburgh, United Kingdom
Pkoufaki@qmu.ac.uk

***Abstract*—Falls are a major cause of injury among older people, often leading to severe consequences, including death. To reduce this risk for both older and younger populations, Artificial Intelligence (AI) can play a critical role by predicting pre-fall states (conditions leading to a fall) and enabling timely intervention. Pre-fall prediction can be approached through various contexts, such as time-based, biological, and sensor data. This study focuses on predicting pre-falls through the time-based context by using the data from wearable sensors (accelerometer and gyroscope), while considering the time window feature of the dataset. The dataset used in this paper was collected using a MetaMotionR device and comprises two classes: "fall" and "no fall". A sliding time window approach of 5 seconds and 10 seconds was applied to prepare the dataset for pre-fall prediction. Notably, this type of dataset has not previously been utilised for pre-fall prediction. A variety of machine learning and Deep Learning algorithms were tested on this dataset. The machine learning models included Decision Tree (DT), Support Vector Machine (SVM), and Logistic Regression (LR), and Deep Learning models included Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN). Among machine learning algorithms, the DT demonstrated super performance, achieving accuracies of 95.99% and 95.75% for the 5-second and 10-second time windows, respectively. In the category of Deep Learning algorithms, Long Short Term Memory (LSTM) type of RNN models outperformed other approaches, with accuracies of 81.08% and 82.63% for the 5-sec and 10-sec windows, respectively.**

Keywords-Fall; Pre-Fall; Machine learning; Deep Learning.

I. INTRODUCTION

As the world's population ages more quickly, there is growing concern about the safety and health of the elderly. Unintentional falls are occurring frequently among older adults, which is associated to negative health outcomes. While falls can happen at any age, their impact is especially severe for the elderly, who often face longer recovery times and higher healthcare expenses; these factors can result in a reduced quality of life. The growing aging population underscores the urgency of addressing fall related risks. Therefore, fall prevention and early intervention are essential for maintaining well-being [1].

These challenges have made research on the detection and prevention of falls before they happen a priority, with recent developments in AI and wearable technology offering promising solutions [2][3]. By continuously tracking people's movements and predicting potential fall scenarios, AI systems can initiate timely interventions to prevent falls, ultimately saving lives and reducing injuries [4]. A key focus in this study is Pre-fall prediction, which can be approached from a sensor based perspective. In this approach, sensors collect data and timestamps to show early indicators of a possible fall. Each situation provides different perspectives on the elements that influence the risk of falling. This study adopts a sensor based approach, utilising gyroscope and accelerometer data to predict pre-fall instances. It highlights the importance of understanding the transitional period leading up to a fall, offering new insights into the factors that contribute to fall risk.

To facilitate this analysis, this study uses a publicly available dataset that includes sensor data collected during both fall and non-fall scenarios. To improve the understanding of Pre-fall (leading to fall) conditions, the dataset was segmented into fixed time windows of 5 and 10 seconds preceding each fall event. This segmentation captures the transitional phase before a fall and provides contextual data that enhance the predictive accuracy of the models. Both Machine Learning (ML) and Deep Learning (DL) algorithms were tested for their effectiveness in predicting pre-fall. The tested models included LR, DT Classifier, Support Vector Machine, Multi-Layer Perceptron, Gradient Boosting, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Long-Short-Term-Memory (LSTM).

Most existing studies focus on post fall detection and under-utilize temporal sensor data for pre-fall prediction. This study addresses these gaps by using gyroscope and accelerometer data within 5 and 10 second windows to enable early fall prediction and timely interventions.

The following are the main contributions of this study:

- Predicting pre-fall instances was achieved using wearable sensor data (gyroscope and accelerometer) segmented into 5 and 10-second time windows.
- A comparative analysis of ML and DL algorithms showed that DT performed the best among the machine learning models, while the LSTM model was the most effective Deep Learning model for pre-fall detection.

- The proposed framework for predicting fall risks in real time facilitates timely interventions, thereby reducing injuries caused by falls and providing overall safety through immediate fall risk assessment.

The remainder of this paper is organized as follows. Section II reviews related work and existing approaches. Section III presents the proposed methodology. Section IV reports experimental results, and Section V discusses the findings. Finally, Section VI concludes the paper and outlines future work.

II. Literature Review

With the aging population and associated risks, the need to address fall related challenges has become increasingly urgent. Researchers are now focusing on early detection systems and preventive measures to mitigate these risks [5][6][7]. The development of wearable sensor technologies, such as gyroscopes and accelerometers, has significantly transformed falls detection and prevention. These devices enable continuous, real-time motion tracking, making it possible to detect of unusual patterns associated with potential falls [8][9]. Due to their portability, non-invasive nature, and high data collection capacity, wearable sensors have proven to be extremely useful for developing ML and DL models [8][10]. While fall detection research has historically concentrated on post-fall identification, more recent studies emphasise pre-fall prediction to allow for prompt intervention. Pre-fall prediction identifies transitional movements indicating an elevated risk by analysing motion patterns during brief time windows before a fall [11][12].

Strong performance in classifying fall related data has been shown by machine learning techniques, such as Logistic Regression (LR), Decision Tree (DT), Support Vector Machine (SVM), and Gradient Boosting [5]. However, time series sensor data analysis is a perfect fit for Deep Learning models, especially LSTM networks, which have demonstrated exceptional performance in capturing temporal dependencies in sequential data [13][14].

When developing and accessing fall detection systems, datasets play an essential role. One example of such datasets are SisFall[15], which is gathered using an accelerometer and gyroscope. It contains classes for Activities of Daily Living (ADL) and falls, gathered from both younger and some older individuals. These activities were selected based on a literature survey [15]. UpFall includes the dataset of 17 individuals who performed 11 daily living activities, as well as falls [16]. The UMAFall dataset highlights the difference between the various approaches of machine learning to fall detection [17]. KFall is a comprehensive dataset for fall inertial sensors (acceleration, gyroscope and Euler angles) which are synchronised with video based fall labels [18]. These datasets were gathered from numerous sensors, both wearable and non-wearable, during everyday activities and simulated falls. Under controlled circumstances (Lab-based environment), these datasets allow researchers to train and validate ML and DL models. Recent advancements in fall prediction and detection are increasingly using wearable and vision based technologies. For instance, the system presented in the study [19] uses both wearable and vision-based sensors, giving a sensitivity of 96% using Hidden Markov Models (HMM) and a decision tree. Many other studies are focusing on real time applications, [20] employs the ConvLSTM network and techniques for real time fall detection and prediction, achieving a high accuracy rate of 98.3%. Similarly, KNN, GRU(Gated Recurrent Unit), and SVM algorithms, along with the wearable sensors, are used to predict falls with an accuracy of 93.5% [21]. Heterogeneous Hidden Markov Model (HHMM) is used for the effective recognition and prediction of falls by utilising the 3D Vision based body data with an accuracy of 81.5% [22]. Additionally, the Kinect System, along with Zero Moment Point (ZMP) and SVMs approaches, was used to reach an accuracy of 91.7% [23]. Deep Learning methods are commonly used in fall detection and prediction research, such as the use of CNNs with Class Activation Maps (CAM), which can detect the impact of a fall before it happens by utilising the wearable sensors. This approach has achieved an accuracy of 95.33% [24]. The PreFallKD system, which integrates CNNs and Vision Transformers with knowledge Distillation, demonstrates strong performance with a 92.66% F-1 score in real time fall prediction using wearable sensor data [25]. However, most existing fall detection systems focus on post fall identification, which limits the potential for prompt interventions. Additionally, temporal data for pre-fall prediction remains underutilised. This study addresses these gaps by leveraging accelerometer and gyroscope data within 5 and 10 second time windows to predict pre-fall conditions. We assess conventional machine learning models (e.g. SVM, DT, and more sophisticated DL architectures (RNN and CNN)), offering a framework for early intervention and fall risk mitigation.Table I shows the comparison of fall detection and prediction approches.

To the best of our knowledge, our study is among the first studies to use this specific MetaMonitor dataset with sliding time windows of 5 s and 10 s for pre-fall prediction, combining both ML and DL models to emphasize the role of temporal context in improving pre-fall prediction.

III. Methodology

The subsequent Figure 1 illustrates the methodological process in this study. The methodology includes various components, such as dataset sampling (data generation), dataset cleaning, preprocessing, modelling and evaluation.

A. Dataset

This study utilised a publicly accessible dataset [26] collected using the MetaMotionR sensor. Data was gathered using two wearable sensors (accelerometer and gyroscope) positioned at the user's waist. The dataset comprises recordings from 17 participants (4 females, 13 males) with an average age of 30 ± 8.02 years, height 174.18 ± 7.85 cm, and weight 74.35 ± 9.71 kg performing various Activities of Daily Living (ADLs) and simulated fall (lab based) events in controlled conditions. The ADLs included jumping, running and stopping, sitting on a chair, and pulling the sensor. The fall scenarios included

TABLE I. COMPARISON OF FALL DETECTION AND PREDICTION METHODS

No.	Refrence	Prediction	Detection	Real Time	Wearable	Vision	AI/ML	Accuracy / Perf.
1	[19]	Yes	Yes	No	Yes	Yes	HMM, DT	Sens: 96%
2	[20]	Yes	Yes	Yes	Yes	No	Conv, LSTM, Smoothing	Acc: 98.3%
3	[21]	Yes	Yes	Yes	Yes	No	KNN, GRU, SVM	Acc: 93.5%
4	[22]	Yes	Yes	Yes	No	Yes	HHMM	Acc: 81.5%
5	[23]	Yes	Yes	Yes	No	Yes	SVM, Mod. ZMP	Acc: 91.7%
6	[24]	Yes	Yes	Yes	Yes	No	CNN + CAM	Acc: 95.33%
7	[25]	Yes	Yes	Yes	Yes	No	CNN + ViT KD	F1: 92.66%

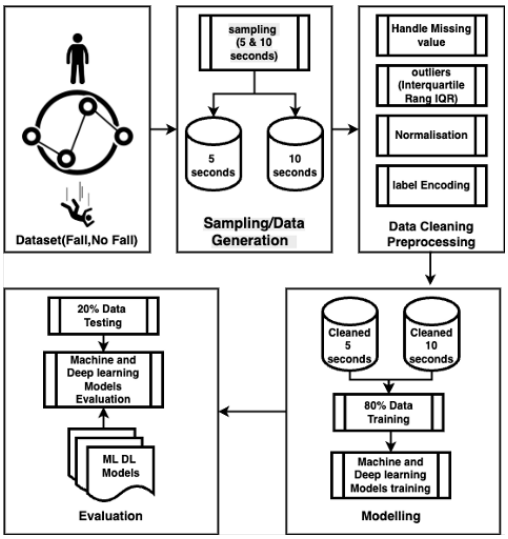


Figure 1. Methodology.

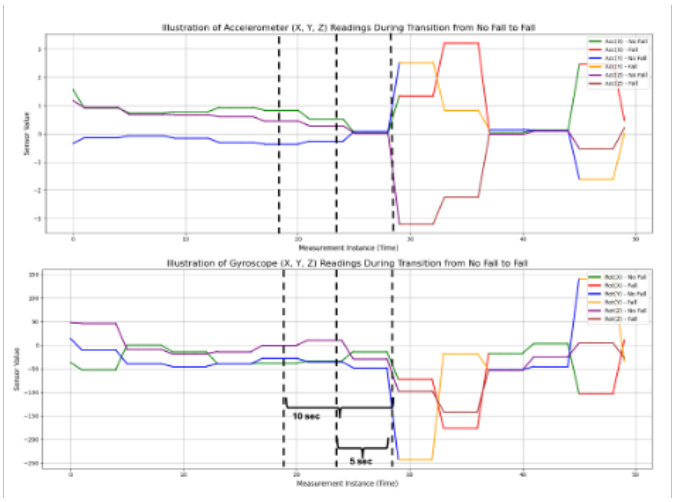


Figure 2. Illustration of all Instances (Fall, no fall) sub figure (a) shows the sensor values of accelerometer and sub figure (b) shows the sensor values of gyroscope.

forward falls, right-side falls, left-side falls, and backwards falls. MetaMotionR sensor records acceleration, rotation, and orientation. Falls were performed on a mat for safety, with a 1-second data window captured when acceleration exceeded 2.5 G. This dataset was chosen due to the nature of the sensor and time stamping for evaluating the performance of various ML and DL algorithms. In this study, we only considered two classes: fall and no fall. Figure 2 shows the values of features (x,y,z) from both the accelerometer and the gyroscope for instances of fall and no fall. It can be observed that both fall and no fall follow distinct patterns; the value of the accelerometer (Acc(X)) is lower in the fall instance and higher in the no fall instance. In case of Acc(Y), the values are higher for the fall event but lower when there is no fall. Acc(Z) shows lower values during fall and higher values for no fall. For the gyroscope readings, fall events are associated with higher Rot(X)and Rot(Y) values while Rot(Z) values are lower. These observed patterns highlight the potential of sensor based features in distinguishing between fall and non-fall events.

B. Sampling/ Data Generation

Data sampling was conducted to create the pre-fall dataset, capturing the time window preceding each fall event. The dataset comprises timestamps (e.g., 5 seconds, 10 seconds), sensor readings, and a binary fall indicator (e.g., 1 representing a fall). The timestamp denotes a fixed time window (e.g., 5 seconds) before each fall, facilitating the identification of conditions that lead to a fall event. Data was chronologically sorted by timestamp to generate the dataset, with fall events marked as 1 and Pre-Fall events as 0. This relationship can be expressed mathematically as follows:

$$prefall = t_{fall} - T_w \leq t_{event} \leq t_{fall} \quad (1)$$

Where t_fall is the timestamp of the fall event, T_w is the time window before fall, which is taken for prefill, which in this case is 5 and 10 seconds, and t_{event} is the timestamp of any row in the dataset. Figure 3 further illustrates the sensor (both accelerometer and gyroscope) reading from a typical no-fall to a fall transition. The pre-fall period is virtually highlighted for both 5 and 10-second windows preceding the fall, showing the temporal dynamics captured in the dataset.

C. Data Preprocessing /Cleaning

The data cleaning process contains several techniques. Firstly, the dataset was checked for missing values [27]. If any missing values are found, they were replaced by the mean of their respective columns. After addressing missing values, the next step was identifying and removing outliers. Outliers were removed using the interquartile range method to prevent them from destroying model training and accuracy [28]. Once the dataset is refined, normalisation is applied in standard scaling to ensure that all data points fall within a consistent range. While ML models often require normalization, feature selection, or handcrafted feature extraction, DL models can automatically

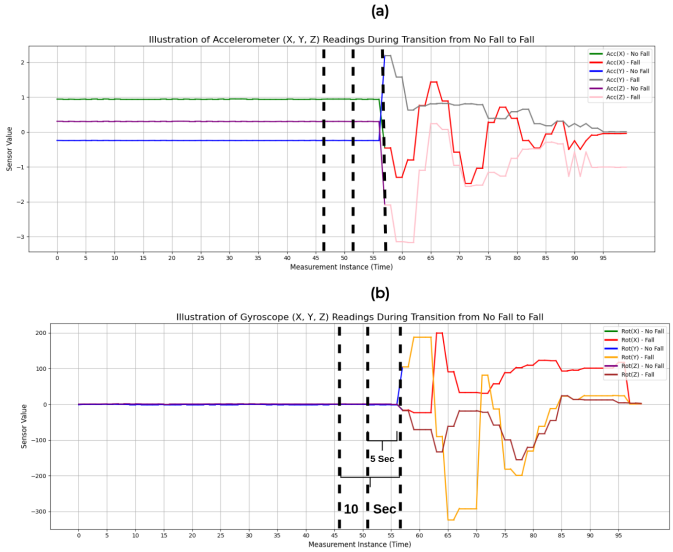


Figure 3. Illustration of sensor readings from Fall to no Fall sub figure (a) shows the sensor values of accelerometer and sub figure (b) shows the sensor values of gyroscope.

learn hierarchical features from raw sensor data, simplifying the overall workflow and potentially capturing more complex temporal patterns.

D. Data Modeling

After completing the preprocessing and data cleaning stages, 80% of the data was utilised to train ML and DL algorithms. For the ML, the data was trained using LR, DT, SVM, Multilayer perceptron (MLP), perceptron, and gradient boosting. The DL algorithms employed include LSTM, CNN, RNN, and DNN. Once the models were trained, the remaining 20% of the data was used to test the ML and DL algorithms. The model's performance was evaluated using Accuracy (the ratio of the number of correctly classified instances to the total number of instances predicted), Precision (the ratio of correctly predicted positive instances to all positively predicted instances) and Recall (the proportion of predicted positive instances to all actual positive instances) factors. These parameters provide comprehensive evaluation of model's performance to predict pre-fall.

IV. Results

The experiments for this study were conducted using Google Colab and Python, utilising 32 GB of RAM and 128 GB of storage. The DL models were trained over 20 epochs after which no significant performance gains were observed, and early stopping was applied to prevent overfitting. The results obtained from experiments using 5-second and 10-second window data by applying ML and DL algorithms, as proposed in the framework. The performance of ML and DL algorithms with the parameters Accuracy(Acc), Precision(Pre) and Recall(Rec) is summarised in Table II. Among all ML models, the DT classifier model has performed efficiently with an accuracy of 95.99% and 95.75% on 5-second and 10-second windows, respectively. For DL models, LSTM has performed efficiently with an accuracy of 81.08% and 82.63% on 5-second and 10-second windows, respectively. Figure 4 illustrates the precision-recall curve and ROC (Receiver operating characteristic) curve for the DT under 5 and 10-second windows. The curves demonstrate a high area under both metrics, indicating strong model accuracy. The 10-second window shows a slightly steeper curve, reflecting marginally improved performance.

TABLE II. ACCURACY, PRECISION AND RECALL FOR 5 AND 10-SECOND TIME WINDOW

5 Second Window			
Algorithms	**Acc**	**Pre**	**Rec**
Logistic Regression	78.75%	59.21%	46.86%
Decision Tree Classifier	95.99%	92.26%	91.58%
Support Vector Machine	82.02%	60.26%	81.41%
MLP Classifier	82.24%	62.02%	73.81%
Perceptron	69.72%	43.88%	77.93%
Gradient Boosting Classifier	84.58%	63.16%	91.18%
RNN	80.60%	56.84%	88.17%
CNN	78.53%	55.00%	70.26%
DNN	75.97%	54.86%	13.48%
LSTM	81.08%	57.42%	89.52%
10 Second Window			
Algorithms	**ACC**	**PRE**	**Rec**
Logistic Regression	78.98%	62.11%	53.78%
Decision Tree Classifier	95.75%	92.27%	91.72%
Support Vector Machine	83.49%	64.03%	86.57%
MLP Classifier	82.13%	65.51%	69.35%
Perceptron	73.18%	49.75%	81.00%
Gradient Boosting Classifier	85.08%	65.13%	94.55%
RNN	82.08%	60.79%	90.05%
CNN	78.74%	61.74%	50.74%
DNN	78.13%	57.89%	62.22%
LSTM	82.63%	61.69%	89.84%

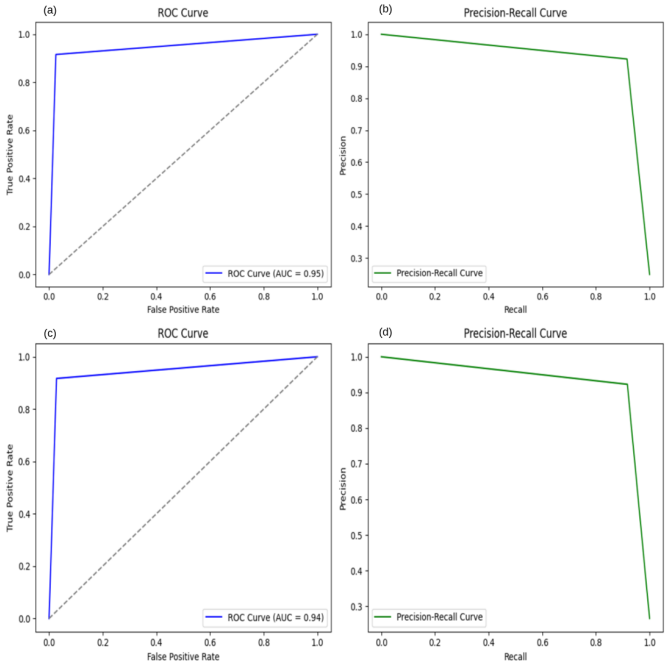


Figure 4. ROC Curve and Precision Recall Curve. (a) and (b) represent ROC curve and Precision recall curve for decision tree for 5 second window and (c) and (d) show ROC curve and Precision recall curve for decision tree for 10 second window.

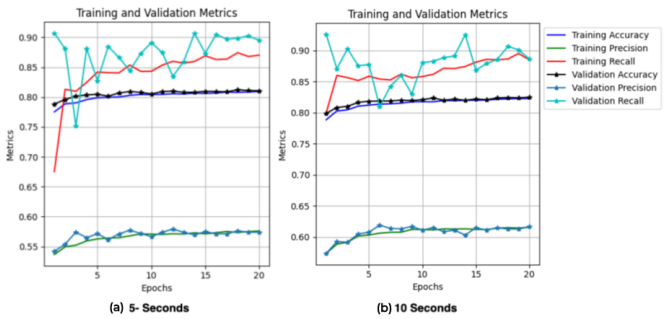


Figure 5. (a) Training, Validation and Testing Accuracy, precision and recall over 20 epochs for LSTM on 5 second window (b) Training, Validation and Testing Accuracy, precision and recall over 20 epochs for LSTM on 5 second window.

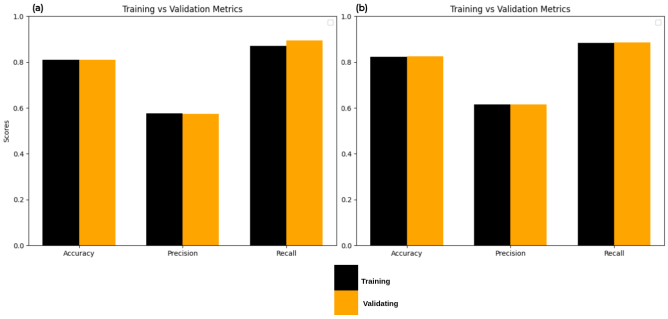


Figure 6. (a) Training and testing validation accuracy, precision and recall on 5 second window (b) Training and testing validation accuracy, precision and recall on 10 second window.

Figure 5 illustrates line graphs for training and validation accuracy, precision and recall over 20 epochs for the LSTM model. The graph highlights a steady improvement in these metrics as training progresses. A slight gap between training and validation metrics indicates that the model fits well and is generalised effectively.

Figure 6 compares training and validation precision and recall for the LSTM model across 5 and 10 second windows. The minimal difference between training and validation metrics suggests the model's robustness and adaptability to the use case.

V. Discussion

This study investigated pre-fall prediction using time stamped data collected from wearable sensors. The dataset included readings from the accelerometer and the gyroscope. The dataset consists of two classes, fall and no fall. To predict Pre-fall events, the dataset was transformed to 5 and 10-second time windows preceding fall occurrence. Both ML models, including Logistic Regression, Decision Tree Classifier, Support Vector Machine, MLP Classifier, Perceptron, Gradient Boosting Classifier and DL models, such as RNN, CNN, DNN, LSTM, were evaluated to identify their effectiveness for pre-fall prediction. The results of model tests indicate the that the DT Classifier is the best performing ML model, achieving an accuracy of 95% across both time windows. This means that the predictions made by the model for the pre-fall events were correct 95% of the time. The DT model was able to perform so well because of its ability to handle datasets with temporal features. In this study sensor readings were taken as temporal feature, which enhance the predative strength of DT model. Among DL models, LSTM performed well, achieving the accuracies of 81.08% and 82.63% for 5 and 10 second windows, respectively. Based on the comparative analysis, the results suggest that although LSTM is good with temporal features, traditional ML models, such as Decision Trees, are more suitable for this dataset due to their structure and features. The robustness of the proposed solution can be seen by the fact that models were tested across multiple time windows (5s and 10s) and using diverse ML and DL models. The consistent performance of DT in ML models and LSTM in DL models across both 5 s and 10 s windows demonstrates the framework's ability to generalize well under varying temporal conditions, which is crucial for reliable real world deployment. Since all of the experiments are performed on single dataset uniform sensor type i.e. accelerometer and gyroscope there exist the chance of data bias which only be studied and covered by including more dataset as discussed in future work.

VI. Conclusion and Future Work

In this study, a time-stamped dataset was used to predict pre-fall using machine learning and Deep Learning. The threshold windows set for pre-fall prediction were 5 seconds and 10 seconds. Based on these time frames, ML and DL algorithms are applied to this dataset. The results indicated that the best performing model is a decision tree with an accuracy of 95% for both 5 and 10-second windows. For DL, LSTM has been demonstrated to be the most suitable model. The nature of data favored traditional machine learning models such as decision trees. The main contribution of this study includes, to perform pre-fall prediction on time-stamped datasets and provide the evaluation scores for these techniques. Additionally, this study evaluated and compared the performance of ML and DL models for pre-fall prediction and established the baseline performance for future research. In the future, more advanced ML and DL approaches will be explored on real-time datasets to further enhance the accuracy and generalisation of pre-fall prediction systems.

References

[1] M. Montero-Odasso et al., "World guidelines for falls prevention and management for older adults: A global initiative", *Age and ageing*, vol. 51, no. 9, afac205, 2022.

[2] D. Mohan et al., "Artificial intelligence and iot in elderly fall prevention: A review", *IEEE Sensors Journal*, vol. 24, no. 4, pp. 4181–4198, 2024.

[3] A. Alarifi and A. Alwadain, "Killer heuristic optimized convolution neural network-based fall detection with wearable iot sensor devices", *Measurement*, vol. 167, p. 108 258, 2021.

[4] R. Jain and V. B. Semwal, "A novel feature extraction method for preimpact fall detection system using deep learning and wearable sensors", *IEEE Sensors Journal*, vol. 22, no. 23, pp. 22 943–22 951, 2022.

[5] S. Usmani, A. Saboor, M. Haris, M. A. Khan, and H. Park, "Latest research trends in fall detection and prevention using machine learning: A systematic review", *Sensors*, vol. 21, no. 15, p. 5134, 2021.

[6] S. K. Gharghan and H. A. Hashim, "A comprehensive review of elderly fall detection using wireless communication and artificial intelligence techniques", *Measurement*, vol. 226, p. 114 186, 2024.

[7] N. Kaur, S. Rani, and S. Kaur, "Real-time video surveillance based human fall detection system using hybrid haar cascade classifier", *Multimedia Tools and Applications*, vol. 83, no. 28, pp. 71 599–71 617, 2024.

[8] A. Choi et al., "Deep learning-based near-fall detection algorithm for fall risk monitoring system using a single inertial measurement unit", *IEEE transactions on neural systems and rehabilitation engineering*, vol. 30, pp. 2385–2394, 2022.

[9] S. Nooruddin, M. M. Islam, F. A. Sharna, H. Alhetari, and M. N. Kabir, "Sensor-based fall detection systems: A review", *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, no. 5, pp. 2735–2751, 2022.

[10] F. Kausar, M. Mesbah, W. Iqbal, A. Ahmad, and I. Sayyed, "Fall detection in the elderly using different machine learning algorithms with optimal window size", *Mobile Networks and Applications*, vol. 29, no. 2, pp. 413–423, 2024.

[11] G. Rescio, A. Leone, and P. Siciliano, "Supervised machine learning scheme for electromyography-based pre-fall detection system", *Expert Systems with Applications*, vol. 100, pp. 95–105, 2018.

[12] X. Hu and X. Qu, "Pre-impact fall detection", *Biomedical engineering online*, vol. 15, pp. 1–16, 2016.

[13] E. García et al., "Towards effective detection of elderly falls with cnn-lstm neural networks", *Neurocomputing*, vol. 500, pp. 231–240, 2022.

[14] J. Wu, J. Wang, A. Zhan, and C. Wu, "Fall detection with cnn-casual lstm network", *Information*, vol. 12, no. 10, p. 403, 2021.

[15] A. Sucerquia, J. D. López, and J. F. Vargas-Bonilla, "Sisfall: A fall and movement dataset", *Sensors*, vol. 17, no. 1, p. 198, 2017.

[16] L. Martínez-Villaseñor et al., "Up-fall detection dataset: A multimodal approach", *Sensors*, vol. 19, no. 9, p. 1988, 2019.

[17] E. Casilari, J. A. Santoyo-Ramón, and J. M. Cano-García, "Umafall: A multisensor dataset for the research on automatic fall detection", *Procedia Computer Science*, vol. 110, pp. 32–39, 2017.

[18] X. Yu, J. Jang, and S. Xiong, "A large-scale open motion dataset (kfall) and benchmark algorithms for detecting pre-impact fall of the elderly using wearable inertial sensors", *Frontiers in Aging Neuroscience*, vol. 13, p. 692 865, 2021.

[19] R. Rajagopalan, I. Litvan, and T.-P. Jung, "Fall prediction and prevention systems: Recent trends, challenges, and future research directions", *Sensors*, vol. 17, no. 11, p. 2509, 2017.

[20] M. A. Sarwar, B. Chea, M. Widjaja, and W. Saadeh, "An ai-based approach for accurate fall detection and prediction using wearable sensors", in *2024 IEEE 67th International Midwest Symposium on Circuits and Systems (MWSCAS)*, IEEE, 2024, pp. 118–121.

[21] M. Mahmoud et al., "Sudden fall detection and prediction using ai techniques", in *2024 21st Learning and Technology Conference (L&T)*, IEEE, 2024, pp. 308–312.

[22] O. Guendoul, H. A. Abdelali, Y. Tabii, R. O. H. Thami, and O. Bourja, "Enhanced fall detection and prediction using heterogeneous hidden markov models in indoor environnement", *IEEE Access*, 2024.

[23] M. Li, G. Xu, B. He, X. Ma, and J. Xie, "Pre-impact fall detection based on a modified zero moment point criterion using data from kinect sensors", *IEEE Sensors Journal*, vol. 18, no. 13, pp. 5522–5531, 2018.

[24] J. Shi, D. Chen, and M. Wang, "Pre-impact fall detection with cnn-based class activation mapping method", *Sensors*, vol. 20, no. 17, p. 4750, 2020.

[25] T.-H. Chi, K.-C. Liu, C.-Y. Hsieh, Y. Tsao, and C.-T. Chan, "Prefallkd: Pre-impact fall detection via cnn-vit knowledge distillation", in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2023, pp. 1–5.

[26] J. F.-B. Ruiz et al., "A low-cost and unobtrusive system for fall detection", *Procedia computer science*, vol. 192, pp. 2160–2169, 2021.

[27] R. J. Little and D. B. Rubin, *Statistical analysis with missing data*. John Wiley & Sons, 2019.

[28] C. C. Aggarwal and C. C. Aggarwal, *An introduction to outlier analysis*. Springer, 2017.

Machine Learning-Based Joint TX Power and RX Sensitivity Control for Overlapping Basic Service Set Interference Mitigation in Dense Internet of Things Wireless Networks

Jin-Min Lee, Hye-Yeon Shim, Il-Gu Lee
Department of Future Convergence Technology Engineering
Sungshin Women's University
Seongbuk-gu, Seoul 02844, Republic of Korea
email: csewa56579@gmail.com, 220237062@sungshin.ac.kr, iglee@sungshin.ac.kr

Abstract— **The recent increase in Internet of Things devices and wireless network equipment has led to frequent occurrences of overlapping basic service set environments, where multiple wireless networks share the same or adjacent channels within the same space. In these environments, network quality degrades owing to channel interference. Previous studies have attempted to avoid interference by blocking some links or using time-division methods; however, these methods have limitations in responding to real-time environmental changes and improving overall network throughput and spatial reuse rates. This study proposes a Machine Learning-based joint control technique for TX power and RX sensitivity. This technique is implemented in both centralized and distributed architectures. Each node recognizes the network state, predicts optimal parameters through a Machine Learning model, and applies them to minimize interference. Experimental results demonstrate that the proposed technique achieves up to 47.1% higher effective throughput and 29.6% better measured Signal-to-Interference-plus-Noise-Ratio compared with the conventional technique. The proposed distributed technique demonstrated approximately 46.4% higher effective throughput (21.43 Mbps) than the conventional central technique under low traffic load and maintained relatively high link quality even in environments with increased traffic load. While the proposed distributed method incurred higher control overhead owing to increased computational requirements compared with the conventional distributed method, the distributed architecture enables each Access Point to operate independently, allowing for parallel processing benefits in actual network deployments.**

Keywords- Overlapping Basic Service Set; Machine Learning; TX power and RX sensitivity Control; Internet of Things Wireless Networks.

I. INTRODUCTION

The recent rapid growth of Internet of Things (IoT) devices and wireless network equipment has led to the frequent occurrence of Overlapping Basic Service Set (OBSS) environments, where multiple wireless networks share the same or adjacent channels within the same space [1]. In these environments, the performance degradation due to channel interference increases significantly. Furthermore, attackers can intentionally generate interference signals or unnecessary traffic, resulting in jamming attacks that threaten the network availability and reliability [2]. Existing OBSS interference mitigation techniques primarily avoid interference issues by blocking certain links or applying time-division methods. However, these approaches have limitations: they degrade the overall network throughput and Spatial Reuse (SR) rates [3]. They often focus solely on TX power control (on sender side) or rely on predefined probability models, thereby failing to respond effectively to real-time changes in the network environment or dynamic traffic patterns. Furthermore, they do not consider controlling the RX sensitivity (on receiver side), which can also affect the interference. Therefore, this study views the OBSS environment as a resource to be managed efficiently, and not merely as a constraint to avoid. This study proposes a Machine Learning (ML)-based framework that jointly controls TX power and RX sensitivity. This study implemented and compared the performances of centralized and distributed architectures. The centralized approach utilizes network-wide information to enable global optimization, whereas the distributed approach allows each Access Point (AP) to perform predictions independently based solely on local information, ensuring scalability and practicality. We compared and analyzed the performance of the conventional technique and two proposed approaches. The main contributions of this study are as follows:

- Centralized and distributed ML architectures are proposed, demonstrating the trade-off between performance and control overhead in OBSS networks.
- TX power and RX sensitivity are optimized to support simultaneous connections for more devices.
- The trade-off between the Signal-to-Interference-plus-Noise-Ratio (SINR) of AP–Station (STA) communication pairs and the overall network connectivity is analyzed, and criteria for simultaneous connections are presented.

The structure of this paper is as follows: Section II reviews the research related to OBSS interference mitigation and ML-based network optimization. Section III describes the proposed technique, and Section IV presents the simulation model. Section V details the experimental environment and Section VI discusses the performance evaluation results. Finally, Section VII presents conclusions and directions for future research.

II. RELATED WORK

Jung et al. [4] proposed an OBSS packet detection SR technique based on an optimized TX power control to achieve high throughput in OBSS environments. The proposed technique derives the optimal TX power that maximizes the

communication success probability through probabilistic geometric analysis and adjusts the clear channel assessment threshold accordingly to reduce interference and increase channel access opportunities. However, this technique has limitations in that it calculates the optimal values based on predefined probability models, making it difficult to adapt flexibly to real-time changes in the network environment or dynamic traffic patterns. Zhu et al. [5] improved the performance of coordinated SR (CSR) in an IEEE 802.11be environment through TX power adjustment and distributed optimization using adaptive CSR and distributed CSR. However, this study did not address RX sensitivity control or adaptability to real-time environmental changes via ML, thereby limiting the comprehensive optimization of the transmit/receive parameters in dynamic traffic environments. In addition, Haxhibeqiri et al. [6] proposed a centralized CSR approach to centrally optimize transmit parameters to resolve OBSS interference issues and enhance network throughput. This approach aims to optimize TX power and Modulation and Coding Scheme (MCS) index to avoid interference at the main receiver. However, centralized structures have limited SR efficiency in dynamic environments owing to structural constraints, such as scalability, overhead, and single points of failure. It also has the limitation of focusing solely on TX power without simultaneously considering RX sensitivity joint control. Wojnar et al. [7] proposed a learning-based scheduling technique using multi-armed bandits (MABs) to optimize the TX power of multiple APs in an IEEE 802.11bn CSR environment. Specifically, they contributed to an 80% throughput improvement using hierarchical MAB (H-MAB) in a centralized manner. However, this study has limitations in terms of interference management, because it does not consider RX sensitivity control.

Previous studies have proposed various approaches to mitigate interference and enhance the SR efficiency in OBSS environments, such as TX power optimization and centralized or distributed parameter control. However, most of these approaches rely on predefined models, making them difficult to adapt flexibly to real-time changes in network environments and dynamic traffic patterns. Furthermore, comprehensive control strategies that simultaneously consider both TX power and RX sensitivity are still lacking, and fail to actively incorporate these dynamic factors through ML-based predictions.

III. OBSS Interference Management via ML-based Joint TX power and RX sensitivity Control

The proposed technique is illustrated in Figure 1. The left figure shows the problem of reduced overall network throughput due to OBSS interference when each AP and STA shares the same or adjacent channels in the existing OBSS environment. In contrast, the figure on the right shows the results of applying the proposed distributed control method. Each node dynamically adjusts its TX power and RX sensitivity through ML-based prediction, thereby minimizing interference and improving the overall network throughput. In the case of the proposed centralized control method, a single AP controls all STAs and APs to minimize interference.

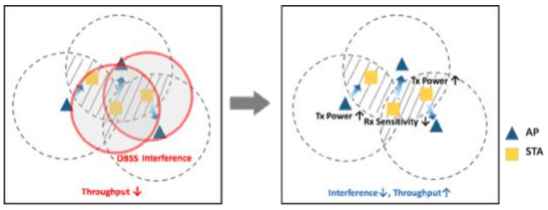


Figure 1. Distributed control for OBSS interference mitigation.

Figure 2 illustrates the overall operational flow of the TX power and RX sensitivity joint control framework proposed in this study.

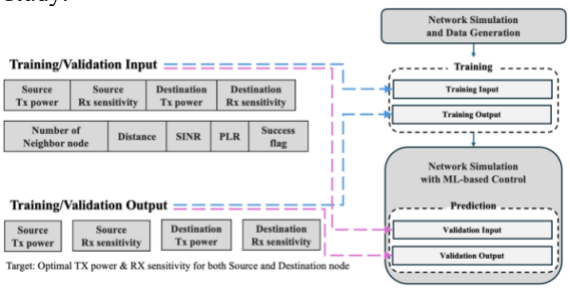


Figure 2. Overall flow for TX power and RX sensitivity control using ML.

First, simulations were repeatedly performed under various OBSS environments and network configurations to collect network environment data. This includes the transmit/receive parameters of each node, number of neighboring nodes, distance, SINR, Packet Loss Rate (PLR), and communication success. Subsequently, the ML model was trained on the collected dataset to predict the optimal TX power and RX sensitivity for each communication pair (source–destination). Detailed information regarding the real-time network environment is summarized in Table 1.

TABLE I. Network Environment Information

Type	Description	Scope
Source ID	Transmitting node	1–45
Destination ID	Receiving node	1–45
Source TX power	Transmitting node's TX power	15–23 dBm
Source RX sensitivity	Transmitting node's RX sensitivity	-90–-75 dBm
Destination TX power	Receiving node's TX power	15–23 dBm
Destination RX sensitivity	Receiving node's RX sensitivity	-90 – -75 dBm
Number of neighbors	Number of nodes within 50 m of the transmitting node	0–44
Distance	Distance between transmitter and receiver nodes	0–141.4 m
SINR	Estimated SINR at the receiving node	-10 – 40 dB
PLR	Packet loss rate	0–1
Success flag	Success of communication connection	0 or 1

The source ID is the identifier of the transmitting node (AP or STA). The destination ID is the identifier of the

receiving node. Source TX power and source RX sensitivity are the transmitting node's current TX power (range 15–23 dBm) and RX sensitivity (-90 – -75 dBm), respectively. Destination TX power and destination RX sensitivity refer to the TX power of the receiving node and the RX sensitivity, respectively. The number of neighbors is the number of surrounding nodes within 50 m of the transmitting node. Distance is the physical distance (m) between the transmitting and receiving nodes, which directly affects path loss. The SINR is the current SINR of the receiving node (dB), indicating the instantaneous link quality. Approximately -10 to 40 dB is an estimated value calculated during training data generation. PLR denotes the packet loss rate. The success flag indicates whether the communication connection was successful, represented by 0 or 1, and serves as the training label for supervised learning.

IV. SIMULATION MODEL

This section presents the simulation model considered in this study. It includes the high-density IEEE 802.11ax network configuration, the wireless channel assumptions, and the formulations of key variables such as TX power, RX sensitivity, and SINR. Figure 3 shows the network configuration used in the simulation.

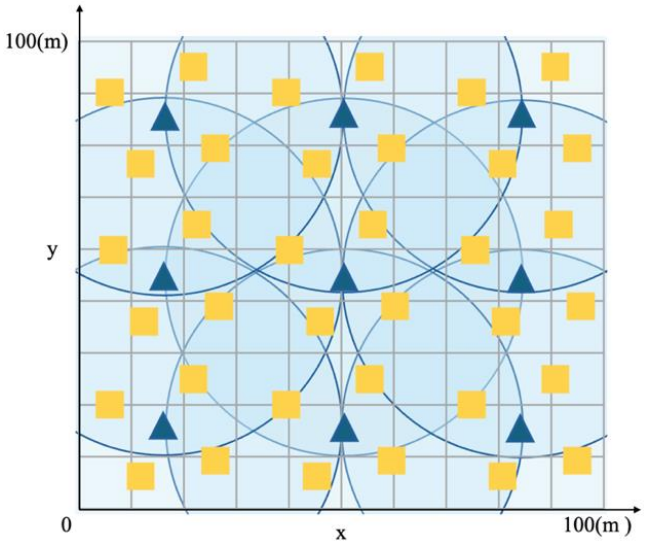


Figure 3. Network configuration.

This study assumes a high-density IEEE 802.11ax wireless network environment deployed within a 100 m × 100 m square area, operating only on a single 20 MHz channel in the 2.4 GHz band [8]. The network comprises nine APs arranged in a 3 × 3 grid pattern at 33.33 m intervals, with four STAs assigned per AP, randomly distributed across each area. This configuration creates an OBSS environment where multiple BSSs operate on the same channel, causing co-channel interference. The 33.33 m spacing between APs was specifically chosen to represent high-density deployment scenarios commonly assumed in smart building and industrial WLAN studies, where coverage overlap is unavoidable. This symmetric arrangement ensures that interference patterns are equally distributed from all directions, providing an unbiased testing environment for evaluating the proposed joint control algorithm's performance under realistic interference conditions.

The wireless channel is modeled using a log-distance path loss model that includes shadow fading, as shown in (1) [9][10].

$$PL(d) = PL_0 + 10\alpha log_{10}\left(\frac{d}{d_0}\right) + X_\sigma \quad (1)$$

Here, $PL_0 = 46.7dB$ is the path loss at the reference distance $d_0 = 1\,m$, and $\alpha = 3.5$ is the path loss exponent for indoor environments. d denotes the distance (m) between the transmitter and receiver, and $X_\sigma \sim N(0, \sigma^2)$ is the shadow fading component with $\sigma = 4dB$.
The noise power is calculated as in (2).

$$\begin{aligned} N &= N_0 + 10\log_{10}(BW) + NF \\ &= -174 + 10\log_{10}(20 \times 10^6) + 7 \\ &= -94\ dBm \end{aligned} \quad (2)$$

Here, $N_0 = -174\ dBm/Hz$ is the thermal noise power density at 290 K, $BW = 20\ MHz$ is the channel bandwidth, and $NF = 7\ dB$ is the receiver noise figure.

The SINR is a key indicator of link quality and achievable data transmission rates in wireless networks [11]. For each communication link, both the downlink and uplink SINR values are calculated. The SINR for the downlink transmission from AP_j to STA_i is calculated as shown in (3).

$$SINR_{DL(i,j)} = \frac{P_{rx(i,j)}}{N_i + I_i} \quad (3)$$

Here, $P_{rx(i,j)} = P_{tx(j)} - PL(i,j)$ represents the received signal power, where $P_{tx(j)}$ denotes the TX power of AP_j, and $PL(i,j)$ denotes the path loss between AP_j and STA_i. N_i is the noise power at STA_i, calculated as $N_i = kTB \cdot NF$, where k is the Boltzmann constant, T denotes the temperature, B denotes the bandwidth (20 MHz), and NF denotes the noise figure (7 dB). I_i is the interference from the other APs and active STAs, as shown in (4).

$$I_i = \sum_{k \neq j} P_{tx(k)} \cdot h_{ki} + \sum_{m \in STA_{active}} P_{tx(m)} \cdot h_{mi} \quad (4)$$

where h_{ki} and h_{mi} represent the channel gains from the interfering AP and STA, respectively, and STA_{active} denotes the set of STAs actively transmitting. The expression for the uplink transmission from STA_i to AP_j is given by (5).

$$SINR_{UL(i,j)} = \frac{P_{rx(j,i)}}{N_j + I_j} \quad (5)$$

The main difference in the uplink calculations is that STAs usually transmit at a lower power. This increases the probability of collisions, owing to the distributed properties of the CSMA/CA protocol.

V. EXPERIMENTAL ENVIRONMENT

In this study, experiments were conducted to analyze the impact of increasing traffic load on network performance by varying the data transmission rate to 3, 6, 12, 24, and 48 Mbps in a formula-based simulation using MATLAB 2021b [12]. Experiments lasting 10 s were repeated 1000 times for each traffic load level to measure the average performance. A ML model using XGBoost [13] was employed to predict and control TX power and RX sensitivity based on the network environment. The model was divided into two approaches: a centralized method, where a single AP handles data learning and prediction, and a distributed method, where nine APs perform data learning and prediction across a 40 m area. The experiments compared and analyzed the following four approaches:

TABLE II. CONTROL TECHNIQUES

Control techniques	Description
Conv (central)	Conventional method controlling Tx power in a centralized technique
Conv (dist)	Conventional method controlling Tx power in a distributed technique
Prop (central)	Proposed method controlling Tx power and Rx sensitivity in a centralized technique
Prop (dist)	Proposed method controlling Tx power and Rx sensitivity in a distributed technique

Table 2 summarizes the control techniques used as comparators in this experiment. Conv (central) is a conventional method that centrally controls TX power, corresponding to the approach by Wojnar et al. [7]. Conv (dist) is a conventional method for controlling TX power in a distributed manner. Prop (central) and prop (dist) are the proposed methods for controlling Tx power and Rx sensitivity in centralized and distributed manners, respectively.

The performance evaluation metrics used were the effective throughput, SINR, control overhead. To evaluate network performance, the achievable effective throughput of each STA-AP link was measured. Effective throughput follows the IEEE 802.11ac physical layer specification, with the transmission rate adaptively selected based on the channel quality [14]. The effective throughput T_i of each STA_i is calculated using (6).

$$T_i = R_{MCS}(SINR_i) \times (1 - PLR_i) \quad (6)$$

where $R_{MCS}()$ is the MCS selection function that maps the measured SINR to the corresponding data transmission rate. The IEEE 802.11ac standard defines 10 MCS levels (0–9), supporting rates from 6.5 Mbps (MCS 0, SINR ≥ 5 dB required) to 86.7 Mbps (MCS 9, SINR ≥ 33 dB required) on a 20 MHz channel. This function selects the highest MCS level that satisfies the minimum SINR requirement. PLR_i is the PLR that combines channel-induced errors and collision-induced losses, as shown in (7).

$$PLR_i = PLR_{channel}(SINR_i) + PLR_{collision(\rho)} - PLR_{channel} \times PLR_{collision} \quad (7)$$

where ρ represents network congestion. STAs with an SINR below 5 dB experience high packet loss (50–90%), whereas those with an SINR above 20 dB achieve low loss rates (below 5%). The control overhead represents the time required for parameter optimization, including ML prediction, file I/O, and SINR calculation. This is distinct from the data transmission period used in effective throughput measurements. While the distributed method theoretically allows nine APs to operate independently, our MATLAB implementation processes these operations sequentially, resulting in cumulative overhead.

VI. PERFORMANCE EVALUATION

In this section, the performance of the proposed ML-based joint control technique is evaluated. The centralized and distributed schemes are compared with the conventional methods, focusing on effective throughput, control overhead, and measured SINR under various traffic load conditions. Figure 4 shows the effective throughput of each method for different traffic loads.

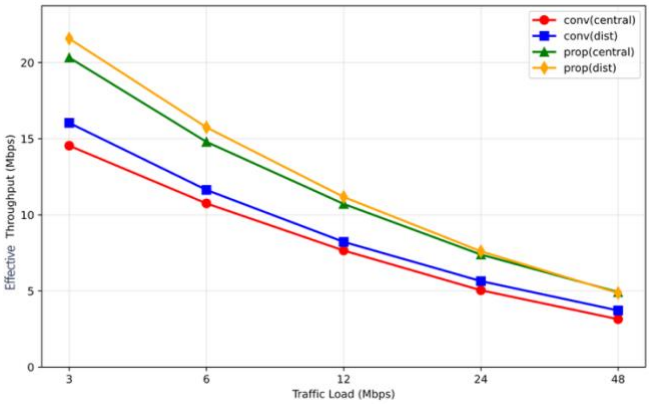


Figure 4. Effective throughput comparison of the four control schemes (conv (central), conv (dist), prop (central), and prop (dist)) with increasing traffic load.

As the traffic load increased, the effective throughput decreased across all methods. The prop methods (central, dist) demonstrated superior performance compared with the conv methods (central, dist). Specifically, prop (dist) achieved the highest effective throughput of 12.09 Mbps, showing an improvement of approximately 47.1% over conv (central). This improvement results from the effective interference control achieved by simultaneously optimizing TX power and RX sensitivity. The dist method exhibited higher effective throughput than the central method because each AP was optimized with values suitable for a 40 m radius area, enabling

finer control. At a low traffic load (3 Mbps), prop (dist) achieved the highest effective throughput at 21.43 Mbps, representing a performance improvement of approximately 46.4% compared with conv (central). However, as the traffic load increased to 48 Mbps, the interference caused a sharp decrease in effective throughput for all methods. Notably, prop (dist) exhibited an effective throughput of 4.71 Mbps, which was 0.44 Mbps lower than that of prop (central).

Figure 5 shows the results of the comparison of the control overhead for each method.

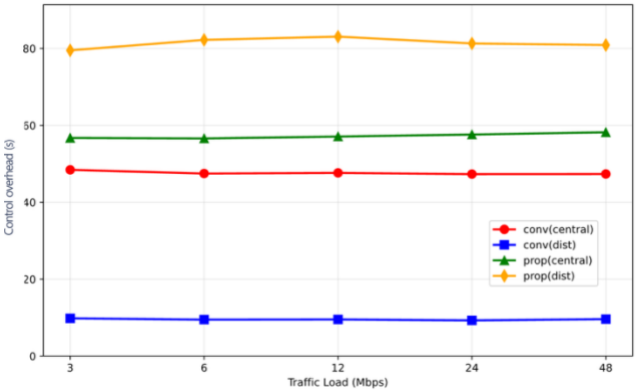


Figure 5. Control overhead comparison of the four control schemes (conv (central), conv (dist), prop (central), and prop (dist)) with increasing traffic load.

Conv (dist) exhibited the fastest control overhead, averaging 9.8 s, whereas prop (dist) required the longest control overhead, averaging 80.6 s. All the methods maintained consistent control overhead regardless of the traffic load, because the computational complexity of the algorithm was independent of the data transmission rate. Because of the complex model structure that simultaneously optimizes TX power and RX sensitivity, the control overhead for the prop methods (central, dist) increased compared with those of the conv methods (central, dist). Specifically, prop (dist) used four models, significantly increasing the overhead and requiring additional computation to predict both TX power and RX sensitivity based on network state information. However, in actual distributed systems, each AP operates independently; therefore, the benefits of parallel processing exist from the perspective of the entire network. Prop (central) increased by 18.1% compared with conv (central), averaging 57.5 s, whereas prop (dist) increased by 723.2% compared with conv (dist).

Figure 6 shows the SINR variation for each method under different traffic loads.

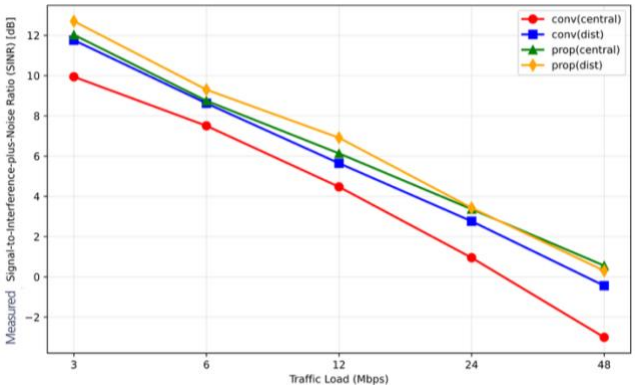


Figure 6. Measured SINR comparison of the four control schemes (conv (central), conv (dist), prop (central), and prop (dist)) with increasing traffic load.

The experimental results indicate that the prop method generally maintained a higher SINR than the conv methods (central, dist). Prop (dist) achieved the highest average SINR of 6.09 dB, representing a 29.6% improvement over conv (central). As the traffic load increased, all methods exhibited a decreasing trend in SINR. When traffic load increased from 3 to 48 Mbps, conv (central) decreased from 10.76 to -1.65 dB, and conv (dist) decreased from 11.61 to -0.73 dB. The proposed methods, prop (central) and prop (dist), also decreased from 11.66 to 0.28 dB and from 12.06 to -0.05 dB, respectively. However, even under high traffic load, the proposed methods maintained a relatively high SINR, providing better link quality. This demonstrates that the proposed methods can sustain a stable link quality even in high-traffic-load environments. The analysis indicates that the distributed approach maintains a higher SINR than the centralized approach because it can more accurately identify and control the interference characteristics within the local area.

VII. Conclusion and Future Work

The rapid increase in the number of devices utilizing wireless networks has exacerbated problems such as channel interference, degraded network quality, and jamming attacks in OBSS environments. Previous studies avoided interference by suspending communication on some links or applying time-division methods; however, these methods failed to reflect real-time changes in the network environment, limiting improvements in overall throughput and SR rates. To address these issues, this study proposes an ML-based simultaneous control technique for TX power and RX sensitivity. The proposed technique is implemented in both the centralized and distributed architectures. Each node recognizes the network state and then predicts and applies the optimal parameters through an ML model, effectively controlling the interference. Experimental results demonstrate that the proposed technique achieves up to 47.1% higher effective throughput and 29.6% improved measured SINR compared with conventional techniques. In particular, the proposed distributed approach achieved a 46.4% higher effective throughput than the proposed centralized approach under low traffic load

conditions while maintaining a relatively stable link quality even under high traffic loads. Although the control overhead increased significantly in the proposed distributed approach, the distributed structure enabled each AP to operate independently. This leverages the benefits of parallel processing, ensuring practical applicability in real-world environments. However, limitations were identified in the simulation environment. MATLAB is primarily designed for algorithm development and numerical computation, not for network simulation, and could not adequately reflect the parallel nature of distributed systems. The sequential processing of distributed operations in MATLAB resulted in higher control overhead, which prevented the observation of actual performance benefits that would occur when multiple APs operate independently in real networks. Future research will use ns-3 for more realistic distributed simulations and reinforcement learning for adaptive control, ultimately validating the framework in real WLAN scenarios.

ACKNOWLEDGMENT

This work is supported by the Ministry of Trade, Industry and Energy (MOTIE) under Training Industrial Security Specialist for High-Tech Industry [grant number RS-2024-00415520] supervised by the Korea Institute for Advancement of Technology (KIAT), the Ministry of Science and ICT (MSIT) under the ICAN (ICT Challenge and Advanced Network of HRD) program [grant number IITP-2022-RS-2022-00156310] and National Research Foundation of Korea (NRF) grant [RS-2025-00518150], and the Information Security Core Technology Development program [grant number RS-2024-00437252] supervised by the Institute of Information & Communication Technology Planning & Evaluation (IITP).

REFERENCES

[1] K. Shafique, B. A. Khawaja, F. Sabir, S. Qazi and M. Mustaqim, “Internet of things (IoT) for next-generation smart systems: A review of current challenges, future trends and prospects for emerging 5G-IoT scenarios,” IEEE Access, vol. 8, pp. 23022–23040, Feb. 2020, doi: 10.1109/ACCESS.2020.2970118.

[2] H. Pirayesh and H. Zeng, “Jamming attacks and anti-jamming strategies in wireless networks: A comprehensive survey,” IEEE Commun. Surv. Tutor., vol. 24, no. 2, pp. 767–809, 2nd Quart. 2022, doi: 10.1109/COMST.2022.3145244.

[3] L. Lanante and S. Roy, “Performance analysis of the IEEE 802.11ax OBSS_PD-based spatial reuse,” IEEE/ACM Trans. Netw., vol. 30, no. 2, pp. 616–628, Apr. 2021, doi: 10.1109/TNET.2021.3052594.

[4] J. Jung, J. Baik, Y. Kim, H.-S. Park and J.-M. Chung, “OTOP: Optimized transmission power controlled OBSS PD-based spatial reuse for high throughput in IEEE 802.11be WLANs,” IEEE Internet of Things J., vol. 10, no. 19, pp. 17110–17123, Oct. 2023, doi:10.1109/JIOT.2023.3275544.

[5] D. Zhu, L. Wang, G. Pan and S. Luan, “Two enhanced schemes for coordinated spatial reuse in IEEE 802.11be: Adaptive and distributed approaches,” Comput. Netw., vol. 258, Art. no. 111060, Jan. 2025, doi:10.1016/j.comnet.2025.111060.

[6] J. Haxhibeqiri, X. Jiao, X. Shen, C. Pan, X. Jiang, J. Hoebeke and I. Moerman, “Coordinated spatial reuse for WiFi networks: A centralized approach,” in Proc. IEEE Int. Workshop Factory Commun. Syst. (WFCS), 2024, pp. 1-8.

[7] M. Wojnar et al., “Coordinated spatial reuse scheduling with machine learning in IEEE 802.11 MAPC Networks,” preprint arXiv:2505.07278v2, 2025.

[8] IEEE standard for information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements—Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications, IEEE Std 802.11-2020, Dec. 2020.

[9] T. Adame, M. Carrascosa and B. Bellalta, “The TMB path loss model for 5 GHz indoor WiFi scenarios: On the empirical relationship between RSSI, MCS, and spatial streams,” in Proc. 2019 Wireless Days (WD), Apr. 2019, pp. 1–8, doi:10.1109/WD.2019.8734243.

[10] Z. Elkhaled, W. Ajib and H. Mcheick, “An accurate empirical path loss model for heterogeneous fixed wireless networks below 5.8 GHz frequencies,” IEEE Access, vol. 8, pp. 182755–182775, Sep. 2020, doi:10.1109/ACCESS.2020.3023141.

[11] 3GPP TR 38.901 v16.1.0, Study on channel model for frequencies from 0.5 to 100 GHz, Release 16, Nov. 2020, ETSI TR 138 901 V16.1.0.

[12] MathWorks, Inc., MATLAB Release 2021b. Natick, MA, USA, 2021. [Online]. Available: https://www.mathworks.com 2025.09.21.

[13] DMLC, “XGBoost.” [Online]. Available from: https://github.com/dmlc/xgboost 2025.09.21.

[14] W. Ciezobka, M. Wojnar, K. Rusek, K. Kosek-Szotta, S. Szott, A. Zubow and F. Dressler, “Using ranging for collision-immune IEEE 802.11 rate selection with statistical learning,” Comput. Commun., vol. 225, pp. 10–26, Sep. 2024, doi:10.1016/j.comcom.2024.07.001.

SCREEN2AIR: Exploiting Screen Savers for Covert Long-Distance Data Exfiltration and Defense

Ye-Rim Jeong
Department of Convergence Security Engineering
Sungshin Women's University
Seoul, Korea
Email: 220254016@sungshin.ac.kr

Yeon-Jin Kim
Department of Convergence Security Engineering
Sungshin Women's University
Seoul, Korea
Email: 220246046@sungshin.ac.kr

Chea-Yeon Park
Department of Convergence Security Engineering
Sungshin Women's University
Seoul, Korea
Email: 220254013@sungshin.ac.kr

Il-Gu Lee
Department of Convergence Security Engineering
Sungshin Women's University
Seoul, Korea
Email: iglee@sungshin.ac.kr

***Abstract*—An air-gapped network is used as a representative protection mechanism to strengthen cybersecurity by physically separating systems. However, to enhance the security of such environments practically, in-depth research on air-gap attack techniques should first be conducted. This study proposes SCREEN2AIR, a novel air-gap attack technique that utilizes screen savers and a high-dimensional modulation technique to encode large amounts of information. Screen savers generally do not cause user suspicion because they are automatically executed when the user is absent and exhibit excellent detection evasion. The experimental results demonstrated that a stable extraction success rate, up to 13 times higher than that of the conventional QR code-based method, can be maintained when the number of cells is small. In addition, we propose a technique to intentionally lower screen saver image quality to defend against decoding, and we experimentally demonstrate that the attack success rate can be reduced by up to 95% compared to using normal high-quality images.**

Keywords-Air-Gap; Data Leak; Screen Saver; Cybersecurity.

I. INTRODUCTION

In recent years, as our reliance on the Internet has grown, cyberattacks have manifested in diverse forms, such as viruses, worms, and ransomware [1]. These attacks can cause significant damage, such as data breaches and network paralysis, at both the individual and national levels. The establishment of air-gapped networks is recommended to mitigate the impact of such threats and protect critical national and industrial information [2]. An air-gapped network is a security system that physically and completely isolates an internal network from external networks, thereby minimizing the risk of external intrusion and internal data leakage [3]. Air gaps are used in environments that require high security, such as national infrastructures and military systems [4].

However, in recent years, security threats targeting air-gapped networks have become a reality; specifically, attacks that exfiltrate data from internal networks to the outside using electromagnetic signals, optical signals, and vibrations generated by the operation of computer components and Internet of Things (IoT) devices are being actively studied [5]. For instance, in 2010, the Stuxnet malware targeted an Iranian nuclear facility [6]. Additionally, in 2024, the Korea Hydro & Nuclear Power Research Institute was hacked, and in 2016, the Ministry of National Defense of the Republic of Korea also suffered a cyberattack [7]. These incidents demonstrate that cyberattacks against air-gapped systems have been reported in multiple countries. Such covert channel-based attacks are challenging to detect using traditional network-based security solutions, such as Intrusion Detection Systems (IDS), Intrusion Prevention Systems (IPS), and firewalls [8]. Therefore, a systematic analysis of potential attack vectors is essential to enhance the security of air-gapped environments.

Conventional air-gap attacks using various physical channels, including electromagnetic, optical, and vibration, have certain limitations. They are easy for users to recognize owing to their high visibility and narrow transmission range [9]. To address these limitations, this study proposes a novel information leakage technique based on a screen saver and a high-dimensional modulation method for encoding large volumes of data. The proposed technique divides the screen saver into cells of a fixed size and assigns binary data to each cell, generating a movement pattern that covertly transmits information that is recognizable only to the attacker. Because screen savers are standard system functions that are typically activated when the user is away, the attack does not raise suspicion and offers strong detection evasion. Furthermore, as the data transmission range scales with the screen size, this technique is advantageous for long-distance exfiltration.

The contributions of this paper are as follows:

- We propose a novel data leakage mechanism and a corresponding defense technique for air-gapped networks utilizing screen savers.
- We experimentally demonstrate the feasibility and effectiveness of the proposed attack with respect to the cell size and number of cells on the screen saver.

• We present a defense strategy tailored to screen-saver-based air-gap attacks and experimentally validate its practicality and performance.

The remainder of this paper is organized as follows. Section II reviews the related work. Section III introduces the proposed information leakage and defense techniques based on screen savers. Section IV presents the results of the performance evaluation. Finally, Section V concludes the study.

II. RELATED WORK

Morderchai Guri [10] proposed an air-gap attack that embeds a Quick Response (QR) code into a monitor display exploiting the limitations of the human visual system in rapidly perceiving blinking images and subtle grayscale patterns. In this study, 40 participants were tested for their ability to recognize Aurmented Reality (AR) codes visually, and the detection range was evaluated using both DSLR and smartphone cameras. However, the method achieved only a 75% success rate in data extraction at a distance of 1m when using a Digital Single-Lens Reflex (DSLR) camera with a 35mm lens, and the maximum recognition range using a smartphone camera was limited to 1.5 meters. Moreover, due to individual differences in visual sensitivity, there exists a risk that an embedded attack pattern may be noticeable to some users.

Anindya Maiti et al. [11] proposed an air-gap attack that leverages the infrared (IR) functionality of smart lighting systems. They encoded the binary data by dividing the brightness levels of smart lights into discrete steps and assigning bit values to each level. A TSOP48 IR sensor connected to an Arduino board was used to detect changes in infrared intensity. Additionally, an 80mm telescope with 45-255x magnification was employed to collect infrared signals and enhance the decoding performance. However, this technique is constrained by its reliance on smart lights equipped with infrared capabilities, which limits its applicability to specific environments.

Morderchai Guri [12] proposed an air gap attack technique that finely adjusts the brightness of the monitor screen, focusing on the fact that it is difficult for the human visual system to recognize the rapidly changing minute brightness difference. In this paper, the monitor screen was photographed using a security camera, webcam, and smartphone camera. OpenCV, an open-source library, was used to process the photographed image in real-time, and a C program that performs additional MATLAB processing by calculating frame brightness was developed and used for decoding. However, this technique has limitations in that the maximum transmission distance is only 1.5m when photographed with a smartphone camera, so the transmission range is limited, and the transmission speed is slow to 1 bit/s.

Previous studies have proposed attack scenarios that exfiltrate data using computer components or IoT devices located in air-gapped environments. While these studies introduced novel attack vectors that are challenging to detect using conventional network-based security solutions, they are limited by the restricted range of usable devices, short transmission distances, and high visibility of attack patterns, which hinder their practical applicability, To address these limitations, this paper proposes a new air-gap attack technique that exploits the screen saver of a computer monitor. The proposed method enables long-distance data exfiltration while remaining inconspicuous to users. Furthermore, we enhance transmission speed by applying a high-dimensional modulation technique and supporting large-volume data leakage.

III. INFORMATION LEAKAGE ATTACK USING SCREEN SAVER AND ITS DEFENSE TECHNIQUE

A. Information Leakage Attack Using Screen Saver

Figure 1 illustrates the operation of the proposed method, SCREEN2AIR. First, after collecting the information to infect and leak malware into a PC inside the airgap, it was converted into a binary number. Subsequently, the screen saver screen was divided into cells of a certain size, and binary bits of data were assigned to each area. When an attack begins, bubbles of a specific color among the bubbles of the screen saver are moved to the corresponding cell according to the information converted to a binary number and repeatedly stopped for a certain period. At this time, since the bubbles used in the attack operate mixed with ordinary bubbles, it poses a challenge for the user to discern this as an anomalous symptom. An attacker outside the airgap photographs the screen saver with a camera and then decrypts the information by analyzing the movement of the bubbles.

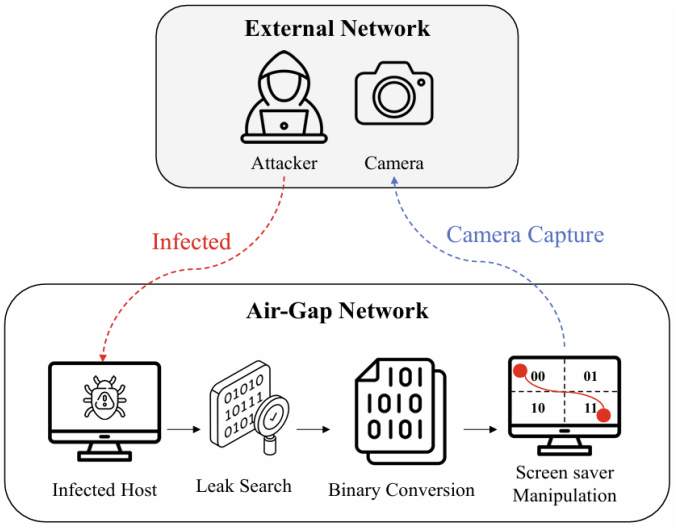


Figure 1. Method of attack using a screen saver

1) Data Encoder

In an environment where the screen saver is divided into 2^n cells, the pseudocode for transmitting information is presented in Algorithm 1. First, the information to be leaked outside the air-gap network is converted into binary data, and

n bit grouping is performed to map it to cells in the screen saver. The height and width of the monitor on which the screen saver was executed was assessed, and the color of the bubbles to be used for data leakage were selected. The screen was divided into 2^n cells based on its height and width and assigned n bits of binary data to each cell. According to the information converted into binary data, the selected bubbles moved to the corresponding coordinates within the cell and stopped at that position for a certain period while normal bubbles maintained their normal movement. For all n bit groups, the corresponding operation process was repeated, and the binary data were leaked to the outside through the position movement pattern of the bubbles.

Algorithm 1 Data Encoder Algorithm (Quadrant-based Transmitter)

```
function TransmitInformation(info)
    binaryData ← ConvertToBinary(info)
    bitPairs ← GroupIntoBitPairs(binaryData)
    screenSize ← GetScreenSize()
    width, height ← screenSize.width, screenSize.height
    targetBubbles ← SelectSpecialColoredBubbles(minCount = 2)
    for each bitPair in bitPairs do
        if bitPair = "00" then
            cx ← random value in range [width/2, width]
            cy ← random value in range [0, height/2)
        else if bitPair = "01" then
            cx ← random value in range [0, width/2)
            cy ← random value in range [0, height/2)
        else if bitPair = "10" then
            cx ← random value in range [0, width/2)
            cy ← random value in range [height/2, height]
        else                                          ▷ bitPair = "11"
            cx ← random value in range [width/2, width]
            cy ← random value in range [height/2, height]
        end if
        MoveBubblesTo(targetBubbles, (cx, cy))
        PauseAtQuadrant(targetBubbles, duration = FIXED_DURATION)
        UpdateNormalBubbles()
        Wait(TIME_INTERVAL)
    end for
end function
```

Algorithm 1. Data encoder pseudocode.

2) Data Decoder

The pseudocode for decrypting information in an environment where the screen saver is divided into 2^n cells is presented in Algorithm 2. First, the image captured by the screen saver is inserted into the decoder. The decoder analyzes the height and width of the image and converts it into an HSV color space. Subsequently, the image is analyzed to identify the location of the mark. Because the bubbles on the screen saver have translucent characteristics, two HSV (Hue, Saturation, Value) coordinates were set by adjusting the brightness and saturation according to the color of the mark. To distinguish colors accurately, the image was masked using Gaussian blur, and the color was recognized in the image according to the HSV coordinates. The area of each cell was classified according to the height and width of the analyzed image, and the binary data allocated according to the area where the recognized color was located was output.

Algorithm 2 Data Decoder Algorithm (Quadrant-based Reception)

```
Load the image from file
Convert image to RGB color space
Get image height and width
Convert image to HSV color space
Define two HSV ranges for detecting red
for each red HSV range do
    Create a mask for the current HSV range
    Print the number of non-zero pixels in the mask
    Apply Gaussian blur to the mask
    Find contours in the mask
    for each contour do
        if area of contour > 30 then
            Calculate the center point (cx, cy)
            if cx >= width/2 and cy < height/2 then
                bit ← "00"
            else if cx < width/2 and cy < height/2 then
                bit ← "01"
            else if cx < width/2 and cy >= height/2 then
                bit ← "10"
            else
                bit ← "11"
            end if
            Save bit value
            break inner loop
        end if
    end for
    if result is found then
        break outer loop
    end if
end for
if result is found then
    Print bit value
else
    Print "No red region detected"
end if
```

Algorithm 2. Data decoder pseudocode.

B. Defense Techniques for Information Leakage Attack Using Screen saver

Screen savers are automatically activated when a computer remains idle for a certain period, thereby preventing screen burn-in, preserving user privacy, and conserving power. Although they were originally developed to prevent burn-in in older display technologies such as Cathode-Ray Tube (CRT) and Plasma Display Panel (PDP), their necessity has diminished with the widespread adoption of Liquid Crystal Display (LCD) and Light-Emitting Diode (LED) displays. Nevertheless, screen savers are still used for screen protection and privacy enhancement.

Detecting data leakage attacks that exploit screen savers in air-gapped environments using traditional security solutions can be challenging because they utilize legitimate system programs as cover channels. In this study, we propose a defense technique that intentionally degrades the image quality of screen savers to reduce the stealthiness of such attacks and lower their success rate. The proposed method decreases the resolution of the screen saver to reduce the clarity of the visual information, thereby significantly diminishing the accuracy of data decryption when the screen is captured by an external camera.

IV. EVALUATION RESULTS AND ANALYSIS

A. *Evaluation of Air-Gap Attack Experiments Based on Screen Saver*

This experiment was conducted by photographing the screen saver screen of a Samsung Galaxy Book Pro laptop with the camera of a Galaxy S25 smartphone and then processing the image to extract information. Shooting was performed by gradually increasing the distance between the screen and camera, and the data leaked from the photographed image were decrypted using an automated script. The air-gap attack technique, which converts leaked data into a QR code and inserts it into an image with lower visibility to evaluate the performance of the proposed model and compare it with existing techniques, was implemented as a conventional model.

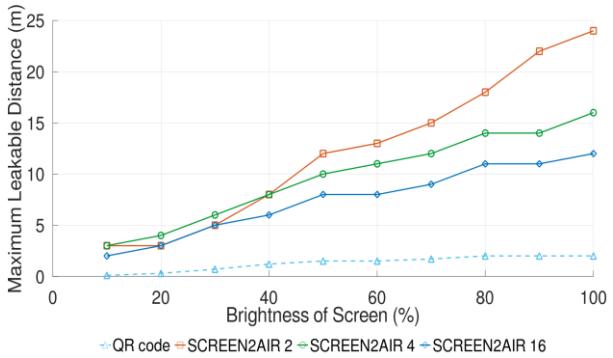


Figure 2. Maximum leakable distance depending on scree brightness.

Figure 2 illustrates the result of comparing the maximum outflow distance according to the screen's brightness. The proposed model (SCREEN2AIR 2) with two cells recorded the longest transmission distance because it can transmit information up to about 24m. On the other hand, the conventional study (QR code) was limited to a maximum of 2m. And as the number of cells increased, the transmission distance tended to decrease somewhat. Since the conventional technique relies on camera-based static image recognition, it was confirmed that the robustness against ambient illumination change is low, and due to this, there is a limit to the transmission distance.

Figure 3 illustrates the results of comparing the information leakage based on the transmission distance of the proposed model with 2, 4, and 16 cells, respectively. In this experiment, the word "hello" was attempted to be transmitted, and each character is expressed as 8 bits (ASCII code), so the maximum information leakage is 40 bits. In the proposed model with 16 cells, the outflow began to decrease from the 12m point and decreased to 3 bits at the 24m point. However, in the proposed model with 4 cells, the outflow decreased from 15m and recorded 13 bits at 24m. The proposed model with 2 cells exhibited the most stable performance, maintaining a 40-bit level even at 24m.

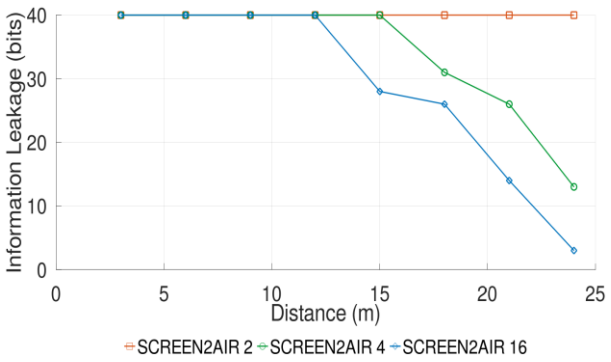


Figure 3. Information leakage by distance by cell count

This result is interpreted as follows: as the number of cells increases, the boundaries between the cells become closer to each other, and the cells where the bubbles are located are not distinguished, thereby increasing the probability of a decoding error. In particular, when the number of cells is 16, the gap between cells is very narrow; therefore, there is a high possibility of a signal hanging over the boundary or a recognition error occurring during long-distance transmission, resulting in a sharp decrease in information leakage. However, when the number of cells was as low as two, the size of the cells was large, and the boundaries were wide, enabling stable information leakage even over a long distance. However, the higher the number of cells, the higher the transmission efficiency because more binary data can be encoded in a single cell. In other words, it has the advantage of transmitting more information at once; however, reliability decreases in a long-distance environment because the gap between cells narrows. In addition, the results of this experiment show that as the screen size increases, the cell gap widens, so even if more cells are placed when using a large screen, a sufficient gap between cells is secured. Thus, a high outflow can be expected, even when long-distance transmissions are performed.

Additionally, increasing the resolution and quality of an image can potentially improve the success rate of information extraction. The pre-image optimization process must be performed since images with low or high resolution are likely to cause errors or missing information during the data extraction process. Therefore, in this paper, image quality processing was performed using "upscale.media [13]", and Artificial Intelligence (AI)-based image upscaling service. "upscale.media" uses an AI super-resolution algorithm to improve low-resolution images up to 8 times to preserve details and textures as much as possible. Rather than simply enlarging the image, the neural network improves the clarity without deteriorating the quality of the original image by analyzing the patterns and boundaries of the image and naturally restoring the missing pixels.

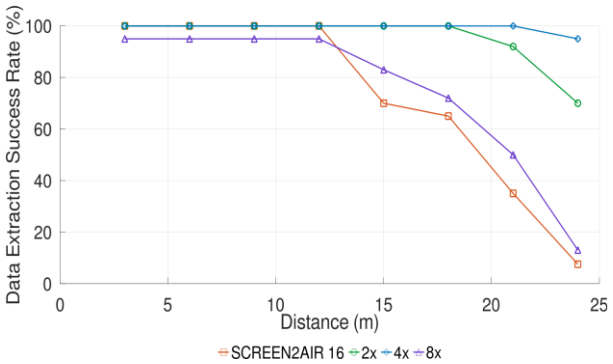


Figure 4. Success rate of information extraction according to distance by Scaling Factor when there are 16 cells

Figure 4 compares the performances of the same scaling factors when the number of cells is 16. The 4x scaling factor maintained a stable data extraction success rate of 100% up to 21m and slightly decreased to 95% at 24m. The 2x scaling factor maintained 100% performance up to 18m and then gradually decreased to 70% at 24m. The 8x scaling factor showed a 95% success rate up to 12m and then decreased to 13% at 24m. Consequently, when an appropriate scaling factor was applied, the information extraction success rate was improved in a long-distance transmission environment. In particular, when a 4x scaling factor was applied to a system with 16 cells, a high information extraction success rate of 95% was recorded at a distance of 24m; this shows that the scaling factor significantly impacts the information extraction performance when the scaling factor is not used, compared to 7.5%. These findings show that proper image scaling more clearly distinguishes inter-cell boundaries and amplifies the detailed features of the mark to effectively offset the effect of noise generated during long-distance transmission; this also suggests that excessive scaling can cause noise amplification or artifact generation, which can degrade leakage data-decoding performance.

Through experiments, it was confirmed that the optimization of the scaling factor plays an important role in increasing the effectiveness and reliability of information leakage. Therefore, it is expected that effective information extraction will be possible through the selection and optimization of the scaling factor in long-distance data leakage scenarios within an air-gap network in the future.

B. Experimental Evaluation of Defense Techniques for Attack Techniques Based on Screen Savers

Figure 5 shows the change in the information extraction success rate according to the attack distance. The information extraction performance was compared with that of normal screen savers with 2 and 4 cells, respectively, and screen savers with deteriorated image quality by applying a defense technique. Normal screen savers maintain a 100% data-extraction success rate up to a distance of 24m. However, in the case of screen savers with deteriorated image quality, the data extraction success rate decreased sharply when the distance was over 9m. As such, screen savers with deteriorated image quality decreased by 95% compared with normal screen savers at a distance of 24m and decreased by approximately 29.5% on average. This suggests that by lowering the image quality of the screen savers, the quality of the visual information decreases, which significantly decreases the decoding accuracy of the attack side. According to the experimental results, the deterioration in the image quality of screen savers can effectively reduce the reliability of an attack as the physical distance increases. This demonstrates that screen saver quality control can be used as a security enhancement technique in an air-gap environment.

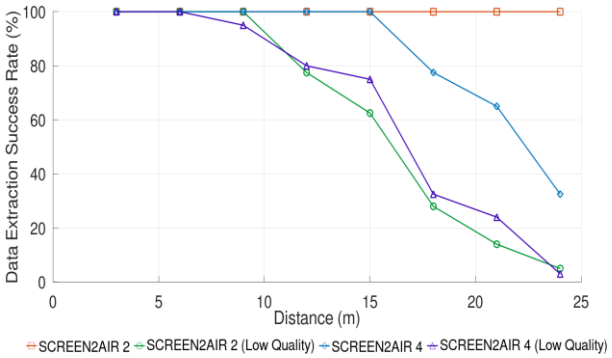


Figure 5. Success rate of information extraction according to distance by image quality

V. CONCLUSION

Building an air-gapped network is recommended to mitigate the impact of cyberattacks and protect critical national and industrial information. However, data leakage attacks that exploit various types of signals within air-gapped systems have been actively studied. Because covert channel-based attacks are difficult to detect using traditional network-based security solutions, proactive re-search on air-gap attack techniques is essential for enhancing the security of such environments.

Conventional air-gap attacks that utilize physical channels, such as electromagnetic and optical signals, are limited by their high visibility, which makes them easily detectable by users, and their short transmission range. To address these limitations, we propose a novel air-gap attack technique that leverages a screen saver combined with a high-dimensional modulation scheme to encode high-capacity information. The proposed method divides the screen saver into cells of fixed size and assigns binary data to each cell to generate a movement pattern for covert data transmission.

The experimental results revealed that a stable extraction success rate of up to 13 times higher can be achieved at long distances when the number of cells is small. Furthermore, applying an appropriate scaling factor improves the success

rate of information extraction by up to 7.6 times, even in long-distance scenarios, confirming the feasibility of effective data leakage over extended ranges.

In addition, we evaluate a defense technique that intentionally degrades the image quality of a screen saver to reduce the reliability of the attack. The experimental results showed that, while the data extraction success rate was maintained almost completely up to 24 m under normal image quality, it dropped sharply to a maximum of 5% at the same distance when the image quality was reduced. These findings experimentally demonstrate that lowering screen-saver image quality is an effective and practical defense technique.

In future work, we plan to compare and analyze additional defense strategies against information leakage attacks using screen savers. These strategies include inserting invisible watermarks into screen content and applying privacy-protection films on the monitor.

ACKNOWLEDGMENT

This study was supported by the Industrial Innovation Human Resources Growth Support Project (RS-2024-00415520), the Information Protection Core Source Technology Development Project (RS-2024-00437252) of the Ministry of Trade, Industry and Energy, and the Korea Institute of Industrial Technology Promotion in 2024.

REFERENCES

[1] Y. Li and Q. Liu, "A comprehensive review study of cyber-attacks and cyber security: Emerging trends and recent developments," Energy Reports, vol. 7, 2021, pp. 8176–8186..

[2] P. Vähäkainu, M. Lehto, and A. Kariluoto, "Cyberattacks against critical infrastructure facilities and corresponding countermeasures," in Cyber Security: Critical Infrastructure Protection, Springer, Cham, 2022, pp. 255–292.

[3] M. R. Na and K. B. Sundharakumar, "A study on air-gap networks," in *Proc. 2024 5th Int. Conf. Innovative Trends in Information Technology (ICITIIT)*, IEEE, 2024, pp. 1–6.

[4] J. Park *et al.*, "A survey on air-gap attacks: Fundamentals, transport means, attack scenarios and challenges," *Sensors*, vol. 23, no. 6, 2023, p. 3215.

[5] Y.-J. Kim, N.-E. Park, and I.-G. Lee, "Air-Fuzz: Feasibility analysis of fuzzing-based side-channel information leakage attack in air-gapped networks," in *Proc. Int. Conf. Information Security Applications*, Springer Nature, Singapore, 2024, pp. 231–242.

[6] S. Q. Abbas and H. Fatima, "Cyber security threats to Iran and its countermeasures: Defensive and offensive cyber strategies," *Journal of Research in Social Sciences*, vol. 12, no. 2, 2024, pp. 1–21.

[7] J. Lee *et al.*, "Optical air-gap attacks: Analysis and IoT threat implications," *IEEE Network*, vol. 38, no. 6, 2024, pp. 342–352.

[8] M. Guri, "Air-gap electromagnetic covert channel," *IEEE Transactions on Dependable and Secure Computing*, vol. 21, no. 4, 2023, pp. 2127–2144.

[9] M. T. Naz and A. M. Zeki, "A review of various attack methods on air-gapped systems," in *Proc. 2020 Int. Conf. Innovation and Intelligence for Informatics, Computing and Technologies (3ICT)*, IEEE, 2020, pp. 1–6.

[10] M. Guri, "Optical air-gap exfiltration attack via invisible images," *Journal of Information Security and Applications*, vol. 46, 2019, pp. 222–230.

[11] A. Maiti and M. Jadliwala, "Light ears: Information leakage via smart lights," *Proc. ACM Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 3, 2019, pp. 1–27.

[12] M. Guri, D. Bykhovsky, and Y. Elovici, "Brightness: Leaking sensitive data from air-gapped workstations via screen brightness," in *Proc. 2019 12th CMI Conf. Cybersecurity and Privacy (CMI)*, IEEE, 2019, pp. 1–6..

[13] upscale.media, https://www.upscale.media, [retrieved: June, 2025]

Invisible Watermarking for Image Data Protection in Sensor Network Environments

Seo-Yi Kim
Department of Convergence Security Engineering
Sungshin Women's University
Seoul, South Korea
email: sykim.cse@gmail.com

Na-Eun Park
Department of Future Convergence Technology Engineering
Sungshin Women's University
Seoul, South Korea
email: nepark.cse@gmail.com

Il-Gu Lee
Department of Convergence Security Engineering
Sungshin Women's University
Seoul, South Korea
email: iglee@sungshin.ac.kr

***Abstract*—Advancements in Artificial Intelligence (AI) have greatly increased the risk of digital-image tampering, underscoring the need to verify the integrity and authenticity of image data collected and transmitted within sensor networks and sensor-based systems. As visual threats, such as deepfakes and adversarial attacks proliferate, manipulated sensor images can trigger severe security incidents and false detections. This paper proposes a robust watermarking method that employs a three-level Discrete Wavelet Transform (DWT) and Singular Value Decomposition (SVD) to repeatedly embed a watermark into the singular values of both low- and selected high-frequency components. Designed to account for transmission noise and environmental distortions in multi-sensor settings, the proposed approach leverages redundancy across multiple frequency bands to enhance resistance to diverse signal-distortion attacks while keeping the watermark imperceptible. Experimental results show that the proposed method significantly surpasses conventional techniques in watermark extraction accuracy while preserving high image quality, establishing it as a reliable security solution for protecting image integrity and detecting tampering in sensor-based environments.**

Keywords-Sensor camera; Digital watermarking; Image protection.

I. INTRODUCTION

The increasing prevalence of malicious video-based attacks, such as deepfakes and adversarial attacks, has increased the need for technologies that can verify the integrity and authenticity of image data obtained from sensor cameras [1].

For instance, a notable 2019 incident in China involved bypassing a facial recognition access control system with deepfake technology. An attacker manipulated facial images from existing surveillance camera footage into real-time deepfake videos, which were then used to deceive the system and compromise physical security [2]. This incident underscores the vulnerability of image sensor-based systems, particularly those integral to public safety.

Similarly, a 2020 experiment in the United States targeting autonomous vehicles demonstrated the threat of adversarial patches. By placing specially crafted patterns on road signs, researchers deceived a vehicle's camera into misinterpreting a "STOP" sign as a "SPEED LIMIT 45" sign [3]. This attack exploited vulnerabilities in AI-based recognition systems, posing serious safety risks during road operations [4].

These examples illustrate the significant security threats that arise when malicious actors manipulate sensor-captured images. Consequently, verifying the integrity and authenticity of sensor-based image data has emerged as a critical security challenge [5].

Digital watermarking is a promising solution to this challenge. This technique embeds identifiable information into image data to detect unauthorized modifications or trace copyright ownership. To be effective, digital watermarking must satisfy two key requirements: robustness against external attacks and imperceptibility, which preserves the original image's visual quality. To meet these criteria, frequency-domain-based methods—particularly those using the Discrete Wavelet Transform (DWT)—are commonly employed. However, DWT-based methods can be vulnerable to certain attacks such as Joint Photographic Experts Group (JPEG) compression, Gaussian or salt-and-pepper noise, filtering (e.g., low-pass/median), and geometric transformations like rotation, scaling, and cropping [6].

To overcome these limitations, recent studies have combined DWT with Singular Value Decomposition (SVD). SVD facilitates watermark insertion by modifying an image's singular values, which represent its essential features, thereby avoiding noticeable distortion [7]. Embedding a watermark into the singular values of DWT-decomposed frequency components has been shown to enhance robustness against both noise and compression attacks [8].

In this study, we propose a method that applies a three-level DWT to decompose an image into its low- and selected high-frequency components. Subsequently, SVD is used to embed the watermark repeatedly into these components. Embedding the watermark in the low-frequency region, which contains the image's core structural information, helps ensure imperceptibility, as even minor modifications in this area can significantly impact the visual appearance. Simultaneously, embedding in high-frequency components enhances resistance to filtering and other frequency-based attacks.

During the extraction process, the correlation between the repeated watermark signals is leveraged to correct errors and accurately reconstruct the original watermark, even in the presence of distortion.

The main contributions of this study can be summarized as follows:

- We propose a novel invisible digital watermarking method that combines a three-level DWT with SVD for robust image integrity protection.
- The method demonstrates enhanced resilience against partial data loss and various signal distortion attacks, which is achieved by embedding the watermark with redundancy across multiple frequency components.
- We developed a comprehensive framework to systematically evaluate watermarking performance under diverse signal distortion conditions.
- Experimental validation confirms that the proposed method significantly outperforms conventional approaches in watermark extraction accuracy while maintaining high image quality.

The remainder of this paper is organized as follows: Section II discusses the conventional methods employed for image integrity protection, Section III details the proposed method, Section IV outlines the experimental setup and procedures, and Section V presents the performance evaluation results. Finally, Section VI concludes the paper.

II. BACKGROUND

Prior studies have employed various techniques to verify the integrity and authenticity of image data, including digital signatures, hashing, and digital watermarking. This section analyzes the conventional methods used for protecting image data.

A. Digital Signature

Albahadily et al. [9] proposed a hash-based digital signature scheme to verify the integrity and authenticity of digital documents. This method generates a unique hash value from the document and user information using the MD5 algorithm and embeds it as a signature. To detect tampering, the receiver extracts the hash value and compares it with a newly generated hash from the received content. This approach employs a lightweight hashing algorithm, enabling fast computation suitable for real-time processing, and is applicable to various data formats, including text and images. However, a key limitation is that the signature data must be stored separately from the image; therefore, the overall content integrity is compromised if the signature is lost or the image is partially modified.

B. Hashing

Khan et al. [10] proposed an ElGamal-based digital signature and encryption scheme to ensure both privacy and authentication for biometric image data. The method first randomizes the image's pixel positions using a 3D Arnold transform and then encrypts both the transform parameters and the image data with the ElGamal public-key cryptosystem. Integrity verification is subsequently achieved using an ElGamal digital signature. The scheme offers strong security by leveraging a public-key cryptosystem based on the discrete logarithm problem. Additionally, the integration of randomization and encryption enables both tamper detection and authentication while significantly reducing the risk of data leakage. However, the method's general applicability is limited, and its high computational overhead makes it unsuitable for lightweight or real-time environments such as Internet of Things (IoT) systems.

C. Digital Watermark

Zhan et al. [11] proposed a reversible fragile watermarking scheme that can verify the integrity of digital images and restore their original content. The method divides an image into blocks and generates two types of data for each: Verification Information (VI) and Recovery Information (RI). VI is embedded directly into its corresponding block to detect tampering, whereas RI, used for content restoration, is concealed in different block locations using the Arnold transform. This dual-verification approach achieves high detection accuracy and supports both tamper detection and content recovery. However, recovery accuracy decreases if the areas containing the watermarks are tampered with, and the complex decoding logic limits its use in real-time applications.

In a related study, Kusumaningrum et al. [12] proposed an image-watermarking technique combining a two-level DWT with SVD, where the watermark is embedded in the low-frequency (LL2) subband, and a non-blind extraction method is employed. The authors compared their method against approaches using only DWT or SVD, evaluating robustness under various attacks, including salt-and-pepper noise, Gaussian filtering, and JPEG compression. However, their evaluation was limited, as it did not consider varying attack intensities or a sufficiently broad range of attacks to comprehensively validate robustness. Although their method outperformed individual DWT and SVD models in watermark extraction, it exhibited poor performance under certain attacks.

Conventional methods demonstrate strengths in areas such as processing speed, security, and recoverability, but they typically involve trade-offs that make it challenging to satisfy all requirements simultaneously. Therefore, this paper presents a watermarking method that minimizes image quality degradation while maintaining robustness against external attacks and tampering during transmission.

III. IMAGE-WATERMARKING METHOD BASED ON DWT AND SVD

This study proposes an invisible watermarking scheme that is robust against signal distortion attacks. The proposed method applies a three-level DWT to decompose an image into multiple frequency subbands, followed by SVD on both the low-frequency and selected high-frequency components. The watermark is embedded repeatedly into the singular values, which enhances resistance to attacks that exploit signal distortions. During extraction, the watermarks embedded in

these multiple frequency regions are retrieved and integrated to successfully reconstruct the original watermark.

The design of the method leverages the different properties of an image's frequency components. High-frequency regions contain fine details such as edges and textures. Slight modifications to these regions are typically imperceptible to the human visual system, making them suitable for embedding invisible watermarks. However, these regions are vulnerable to noise attacks aimed at disrupting the watermark.

In contrast, an image's low-frequency components carry its global structure and essential information. Because modifications in this region can cause noticeable degradation in image quality and structure, embedding watermarks here requires minimal distortion to preserve visual fidelity. Watermarks in the low-frequency band are generally robust against JPEG compression, which primarily targets high-frequency content, and show lower sensitivity to attacks such as Gaussian noise and downsampling. As the low-frequency subband retains significant image information even after transformation, an embedded watermark can be reliably recovered unless the image undergoes severe degradation. However, this region has its vulnerabilities. High compression ratios can cause data loss in low-frequency components, and compression schemes like JPEG2000, which operate across the full frequency spectrum, can adversely affect the watermark. Moreover, global adjustments to image properties, such as brightness or contrast, can also impact the integrity of a watermark embedded in this region.

To address these respective challenges, the proposed method utilizes both low- and selected high-frequency components to implement a robust and invisible watermarking scheme.

A. Watermark Embedding Process

Although image-watermarking techniques that combine DWT and SVD typically follow a similar structure, specific procedures vary based on research objectives, such as enhancing robustness, imperceptibility, or efficiency. Typically, the process involves applying DWT to a host image to generate subbands (LL, LH, HL, HH), followed by performing SVD on a selected subband to embed a watermark by modifying its singular values.

The embedding process for the proposed method is illustrated in Figure 1. The size of the watermark image is fixed based on the host image's dimensions and the DWT level, as defined in (1):

$$W = \frac{N}{2^L} \quad (1)$$

where W denotes the side length of the watermark, N is the side length of the host image, and L represents the DWT level. In this study, a 512 × 512 host image and a three-level DWT were employed, necessitating a 64 × 64 watermark image.

When a three-level DWT is applied to the host image, the frequency domain is decomposed into four subbands: LL3, LH3, HL3, and HH3. SVD is then performed on the low-frequency (LL3) and selected high-frequency (LH3 and HL3) subbands to enable watermark embedding. The watermark is first embedded by modifying the singular values of these subbands, denoted as S_t. However, this modification can alter the host image's structural characteristics, which may degrade image quality or cause watermark extraction to fail if the new values do not align well with the original structure.

To address this potential issue, a second SVD is employed as a recalibration process to refine the modified singular values before reconstruction. This additional step helps integrate the modified singular values more naturally into the image's structural context, yielding new, updated singular values (S_w) that improve both the imperceptibility and robustness of the watermark. Using these updated values, the modified subbands (LL3t, LH3t, and HL3t) are reconstructed.

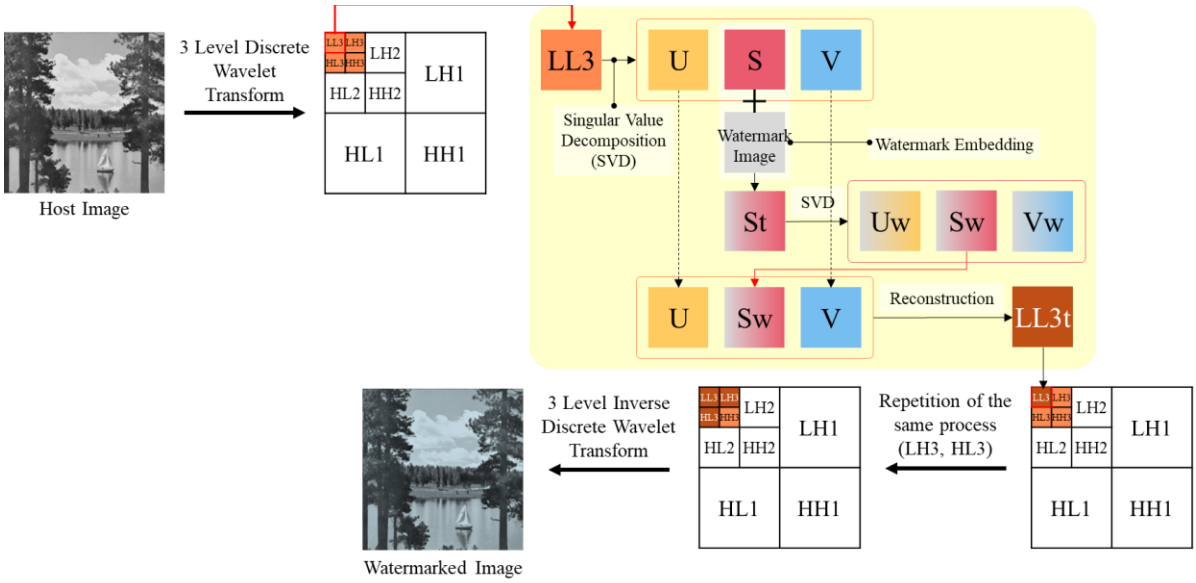


Figure 1. Proposed watermark-embedding process.

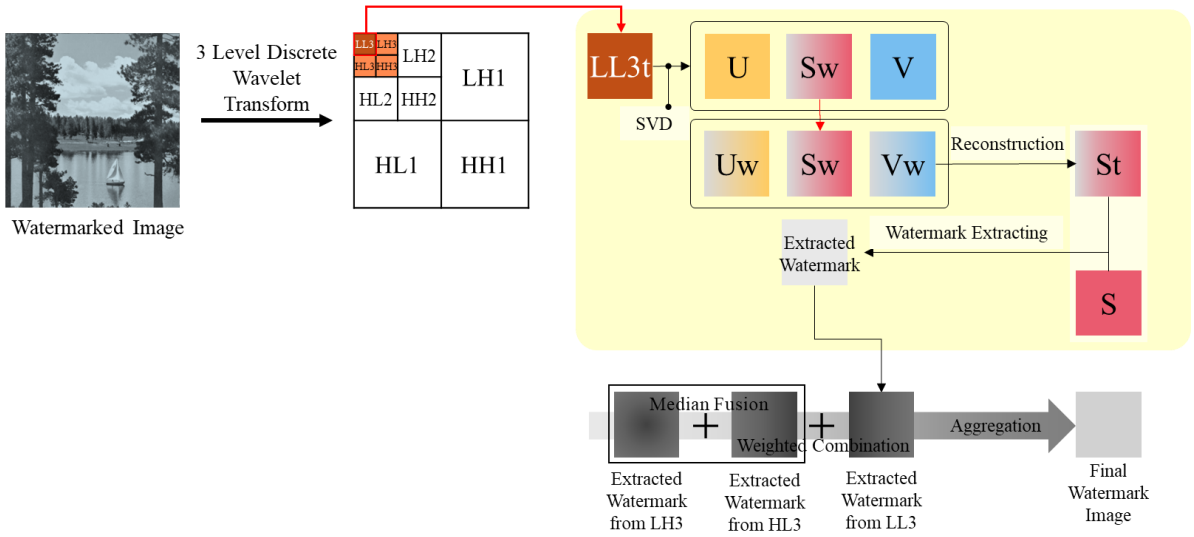


Figure 2. Proposed watermark-extraction process.

Finally, an Inverse DWT (IDWT) is performed to generate the watermarked image. This procedure results in the watermark being embedded thrice into different frequency subbands, creating a redundant watermark structure within the image.

B. *Watermark-Extraction Process*

The watermark extraction process, illustrated in Figure 2, follows a non-blind approach. First, a three-level DWT is applied to the watermarked image to decompose it into its constituent frequency subbands. SVD is then performed on the LL3, LH3, and HL3 subbands to extract the singular value matrices (S_w), where the watermark was embedded. Using these extracted matrices along with the corresponding original U_w and V_w matrices, the watermark images are reconstructed. Because the watermark is embedded separately into the LL3, LH3, and HL3 subbands, three distinct instances can be extracted for the final reconstruction.

The final watermark is reconstructed by fusing these three instances. Median fusion is first applied to the corresponding pixel values of the watermarks extracted from the high-frequency LH3 and HL3 subbands. This step integrates their information while reducing the influence of noise. The resulting intermediate watermark is then combined with the watermark from the LL3 subband using a weighted combination. Because the LL3 subband contains the most critical structural information and is least affected by distortions, its extracted watermark is assigned a higher weight. This ensures that the LL3 watermark plays a dominant role in the reconstruction, whereas the components from LH3 and HL3 serve as complementary sources of information.

IV. EVALUATION METHODOLOGY

This section details the methodology used to evaluate the performance of the proposed DWT-SVD image-watermarking method. IT describes the experimental setup, attack scenarios, evaluation metrics, and the procedure for embedding and extraction.

A. *Experiment Environments*

As shown in Figure 3, the experiments employed 512 × 512 pixel grayscale host images and a 64 × 64 pixel grayscale watermark image.

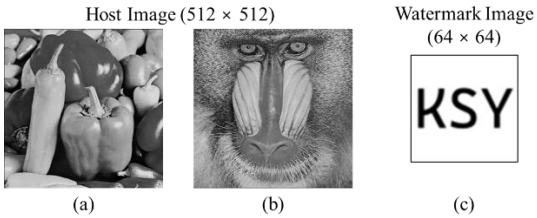


Figure 3. Host and watermark images used in the experiment: (a) Peppers, (b) Mandrill, and (c) watermark image.

To evaluate the robustness of the proposed watermarking scheme, seven distinct signal distortion attacks—encompassing noise, compression, and filtering—were applied to the watermarked images. Each attack was conducted at five intensity levels, from mild (Level 1) to severe (Level 5), to assess performance under varying conditions. The specific parameters controlling the intensity for each attack are summarized in TABLE I.

The intensity of each attack was controlled by specific parameters. For Gaussian noise, intensity was determined by the variance, where a higher value corresponds to stronger noise.

For salt-and-pepper noise, the density parameter represented the proportion of affected pixels; for instance, a density of 0.1 adds salt noise (white pixels, value = 255) to 5%

TABLE I. ATTACK PARAMETERS AND INTENSITIES.

Attack	Parameter	Attack intensity (level)				
		1	2	3	4	5
Gaussian noise	Variance	0.001	0.005	0.01	0.05	0.1
Salt-and-pepper	Density	0.01	0.03	0.05	0.1	0.2
Speckle noise	Probability	0.01	0.03	0.05	0.1	0.2
JPEG	Quality factor	90	70	50	30	10
JPEG2000		90	70	50	30	10
Blurring attack	Kernel size	3	5	7	9	11
Low-pass filtering		3	5	7	9	11

of the pixels and pepper noise (black pixels, value = 0) to another 5%, resulting in a total of 10% corrupted pixels.

Speckle intensity was controlled by a probability parameter, which defines the likelihood that any given pixel will be corrupted by noise. Here, higher probability results in noisier pixels.

For JPEG and JPEG2000 compression, the attack intensity was set by the quality factor, with lower factors indicating stronger compression and greater image quality loss.

Finally, for blurring and low-pass filtering, the kernel size determined the intensity. A larger kernel produces a stronger blur effect (greater information loss) or, in the case of low-pass filtering, removes more high-frequency components. For instance, a kernel size of 3 corresponds to a 3 × 3 filter. Each attack was applied at five intensity levels, from weak (Level 1) to very strong (Level 5), to evaluate the method's robustness under all scenarios.

B. *Experimental Procedure*

The experimental workflow is illustrated in Figure 4. The embedding process begins by applying a three-level DWT to the 512 × 512 host image, using the Daubechies 4 (db4) wavelet with periodization to decompose it into LL3, LH3, HL3, and HH3 subbands. SVD is then applied to the LL3, LH3, and HL3 subbands. The watermark is embedded into the singular value matrices using a scaling factor, α, followed by the second SVD recalibration step. The modified subbands (LL3t, LH3t, and HL3t) are then reconstructed and used in an inverse DWT (IDWT) to generate the final watermarked image.

For the robustness evaluation, each signal distortion attack was applied to the watermarked image. The watermark was then extracted from the attacked image by first applying a three-level DWT, followed by SVD on the LL3t, LH3t, and HL3t subbands. The same scaling factor α used during embedding is applied during extraction. The three extracted watermarks are then combined to reconstruct the final image. This is done by first applying median fusion to the watermark data from the LH3 and HL3 subbands to reduce noise and produce an intermediate watermark. This watermark is then combined with the LL3 watermark using a weighted combination, assigning a weight of 0.9 to the low-frequency data and 0.3 to the high-frequency data.

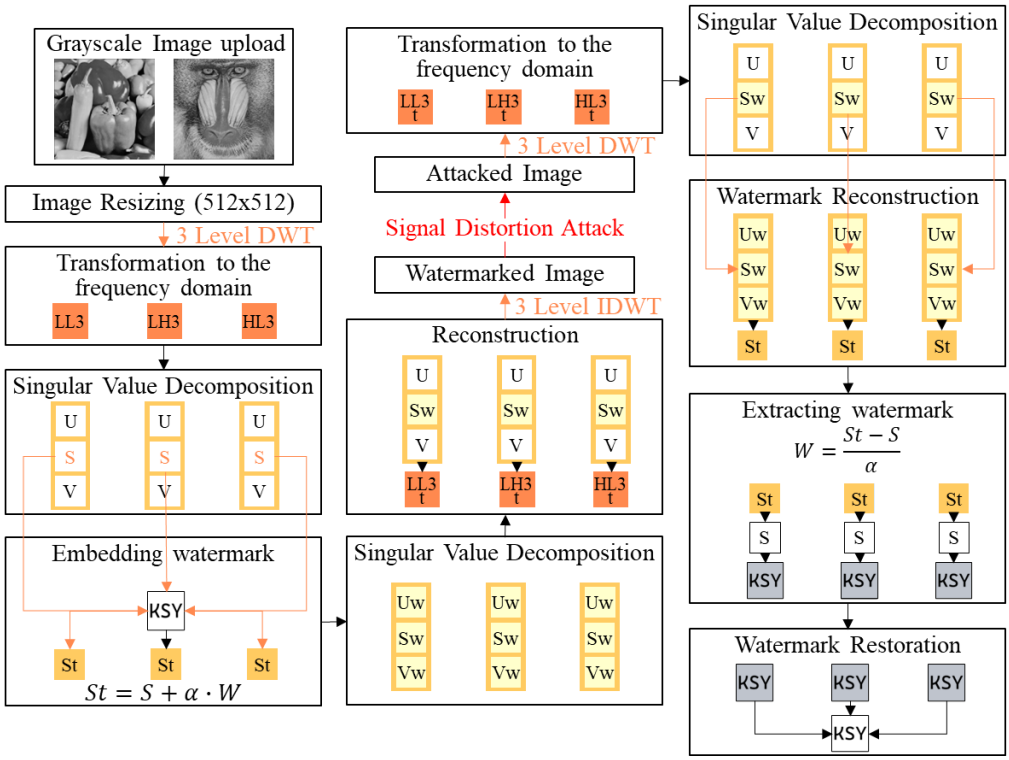


Figure 4. Flowchart of the experimental procedure.

C. Performance Evaluation Metrics

To assess watermark extraction accuracy and image quality, the following performance evaluation metrics were used:

Normalized Cross-Correlation (NCC) measures the similarity between two images, and in this study, it was used to compare the host image with the watermarked image and the original watermark with the extracted one [13].

Mean Squared Error (MSE) quantifies the pixel-wise numerical error between the original and altered images by averaging the squared differences between corresponding pixels, which evaluates the distortion caused by watermark embedding [14].

The Peak Signal-to-Noise Ratio (PSNR) is a widely used metric for assessing the quality of a distorted image compared to its original version; a higher PSNR value indicates better preservation of image quality after embedding [15].

The Structural Similarity Index Measure (SSIM) evaluates the structural similarity between two images by incorporating characteristics of the human visual system, such as luminance, contrast, and structure, making it a more perceptually relevant indicator than PSNR [16].

V. EXPERIMENTS

To validate the performance of the proposed method, a comparative analysis was conducted against a conventional method, which employs a two-level DWT and SVD, embedding the watermark only in the low-frequency (LL2) subband [12]. Both methods used the same watermark embedding strength (α), and robustness was evaluated by applying seven signal distortion attacks at five different intensity levels to assess performance under varying degrees of attack severity.

A. Image Quality Comparison

Figure 5 compares the image quality of the conventional and proposed methods using the *Peppers* and *Mandrill* images. The conventional method yielded slightly better visual quality because it only embeds the watermark in the low-frequency subband (LL2), preserving more of the original image content.

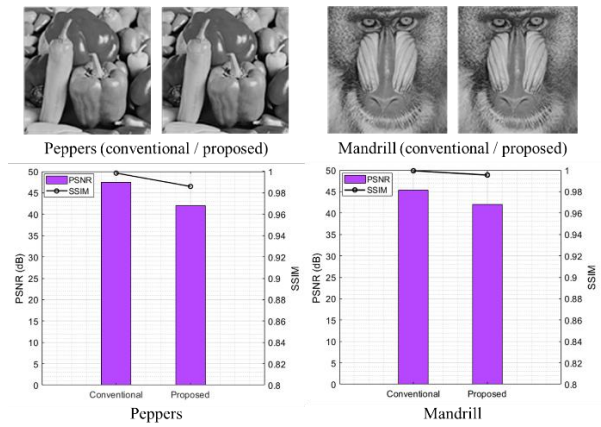


Figure 5. Image-quality comparison between the conventional and proposed methods: (a, c) Peppers and (b, d) Mandrill.

With the proposed method, the PSNR for the *Peppers* and *Mandrill* images decreased by 11.5% and 7.28%, respectively, although both values remained high, exceeding 40 dB. Similarly, the SSIM values showed only a marginal decline of 1.25% and 0.4%, respectively, with scores remaining above 0.98, indicating excellent perceptual similarity.

B. Watermark-Extraction Performance

To compare the watermark extraction performance of the conventional and proposed methods, the seven signal distortion attacks were applied to the watermarked images at five intensity levels.

The performance was then evaluated using the NCC and PSNR metrics.

As shown in Figure 6, the conventional method exhibited significant performance degradation in NCC for the *Peppers* and *Mandrill* images as the intensity of Gaussian noise, sparkle noise, and low-pass filtering attacks increased, with noticeable drops also observed for salt-and-pepper and blurring attacks. Specifically, as attack intensity rose from Level 1 to 5, image deteriorated by 75% (Gaussian noise), 89.99% (sparkle noise), and 82.55% (low-pass filtering). The *Mandrill* image showed similar degradation rates of 65.95%, 90.44%, and 92.34% for the same attacks.

By contrast, while the proposed method's performance also declined with increasing attack intensity, the degradation was significantly lower. For instance, under the most impactful low-pass filtering attack, the proposed method's performance dropped by only 14.26% for *Peppers* and 16.10% for *Mandrill*, demonstrating its superior robustness.

While there was no substantial performance difference for most compression attacks, the proposed method was superior under severe JPEG2000 (Level 5) compression, outperforming the conventional method by 31.47% for *Peppers* and 94.66% for *Mandrill*.

As presented in Figure 7, the conventional method showed a sharp decline in PSNR for nearly all attacks, failing to maintain stable performance even at weak, Level 1 intensities (except for JPEG compression). The most severe degradation occurred with the speckle noise attack; for the *Peppers* image, PSNR dropped from 12.60 dB (Level 1) to -12.83 dB (Level 5), a 201.86% decline. By contrast, the proposed method demonstrated consistently stable PSNR performance. Only minor degradation was observed for noise and low-pass filtering attacks between Levels 1 and 2, with values remaining relatively stable thereafter. Although compression attacks caused some degradation, the decline was considerably less severe than that with the conventional method, and the proposed method maintained higher extraction performance across all attack intensities.

VI. CONCLUSION AND FUTURE WORK

This paper presented a digital-image watermarking scheme that achieves both high robustness against signal distortion attacks and strong imperceptibility. The method combines a three-level DWT with SVD, repeatedly embedding a watermark into the singular values of the low-

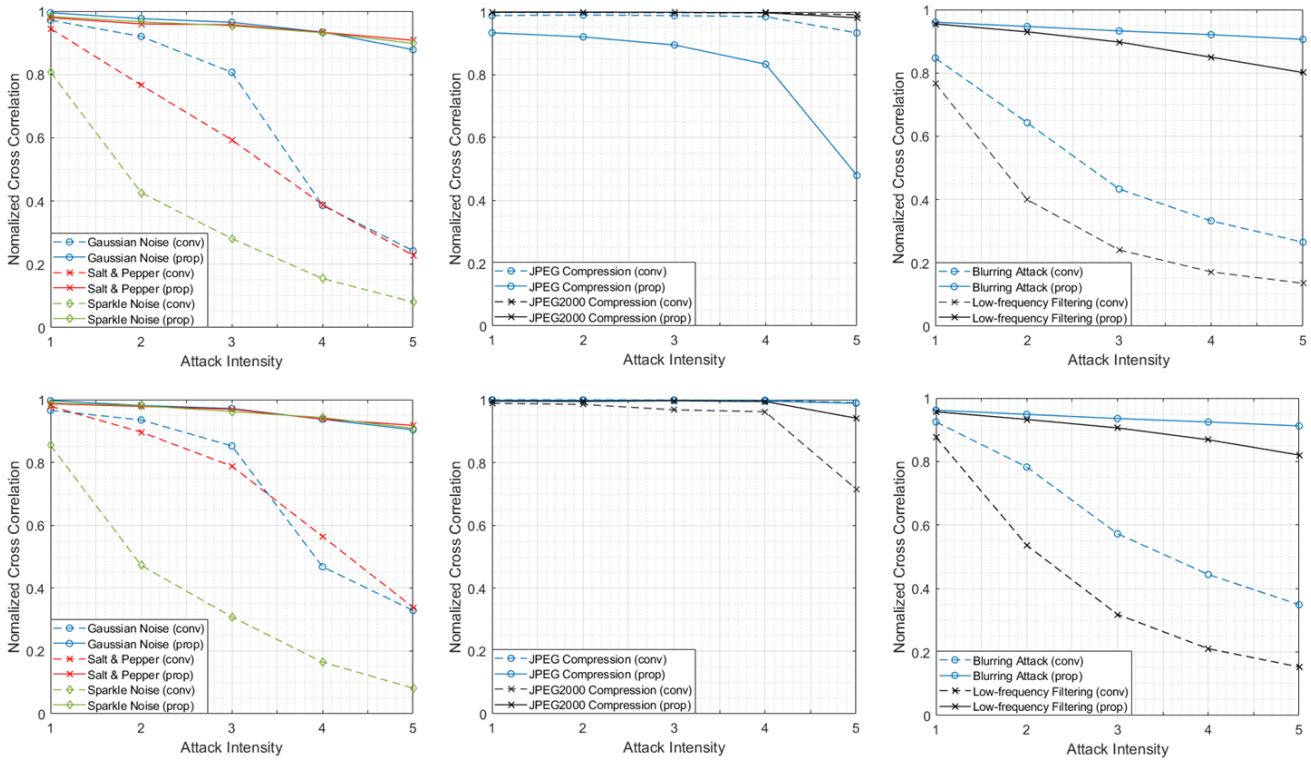


Figure 6. Extraction-performance comparison of conventional and proposed methods based on attack intensity—NCC (top: Peppers, bottom: Mandrill).

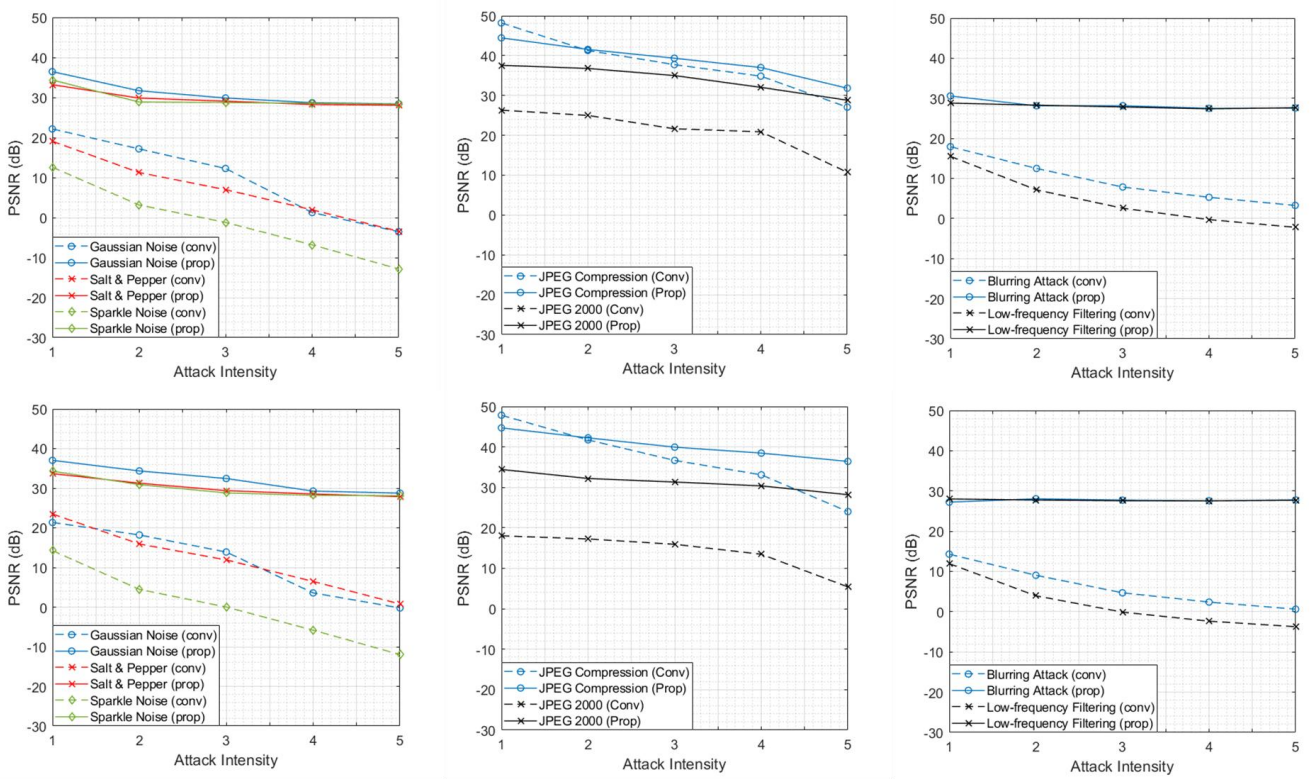


Figure 7. Extraction-performance comparison of conventional and proposed methods based on attack intensity—PSNR (top: Peppers, bottom: Mandrill).

frequency (LL3) and selected high-frequency (LH3, HL3) subbands. This redundant embedding enhances robustness against various attacks while allowing for the complementary recovery of damaged watermark data, effectively mitigating the typical trade-off between imperceptibility and robustness found in conventional methods.

Experimental results demonstrated that the scheme preserves excellent image quality, maintaining high PSNR and SSIM values after embedding. The redundancy led to significantly improved extraction performance; even when parts of the watermark were degraded, the copies enabled accurate reconstruction and reliable detection. Moreover, the method consistently showed strong performance under various levels of noise and compression attacks.

Therefore, the proposed method represents a practical solution for protecting image data in sensor network environments, offering an effective alternative for applications where high reliability and imperceptibility are essential.

ACKNOWLEDGMENT

This work was supported by the Ministry of Trade, Industry and Energy (MOTIE) under Training Industrial Security Specialist for High-Tech Industry (RS-2024-00415520), supervised by the Korea Institute for Advancement of Technology (KIAT), and the Ministry of Science and ICT (MSIT) under the ICAN (ICT Challenge and Advanced Network of HRD) program (No. IITP-2022-RS-2022-00156310), supervised by the Institute of Information & Communication Technology Planning & Evaluation (IITP).

REFERENCES

[1] P. Zhou, X. Han, V. I. Morariu, and L. S. Davis, "Two-stream neural networks for tampered face detection," in 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Jul. 2017, pp. 1831–1839. doi: 10.1109/CVPRW.2017.229.

[2] Y. Mirsky and W. Lee, "The creation and detection of deepfakes: a survey," ACM Comput Surv, vol. 54, no. 1, p. 7:1-7:41, Jan. 2021, doi: 10.1145/3425780.

[3] A. Zolfi, M. Kravchik, Y. Elovici, and A. Shabtai, "The translucent patch: a physical and universal attack on object detectors," in 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2021, pp. 15227–15236. doi: 10.1109/CVPR46437.2021.01498.

[4] N. Wang, Y. Luo, T. Sato, K. Xu, and Q. A. Chen, "Does physical adversarial example really matter to autonomous driving? Towards system-level effect of adversarial object evasion attack," in 2023 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE, 2023, pp. 4389–4400, doi: 10.14722/vehiclesec.2024.25014.

[5] S. Pavlitska, J. Robb, N. Polley, M. Yazgan, and J. M. Zöllner, "Fool the stoplight: realistic adversarial patch attacks on traffic light detectors," Jun. 05, 2025, arXiv: arXiv:2506.04823. doi: 10.48550/arXiv.2506.04823.

[6] D. Bhalke, C. Rupa, H. Dahiya, and V. Yadav, "Privacy protection of digital information using frequency domain watermarking technique," in *Proc. 2021 4th Int. Conf. Recent Trends Comput. Sci. Technol. (ICRTCST)*, Feb. 2022, pp. 202–206, doi: 10.1109/ICRTCST54752.2022.9781929.

[7] N. Zermi, A. Khaldi, M. R. Kafi, F. Kahlessenane, and S. Euschi, "Robust SVD-based schemes for medical image watermarking," Microprocess. Microsyst., vol. 84, p. 104134, Jul. 2021, doi: 10.1016/j.micpro.2021.104134.

[8] O. Evsutin and K. Dzhanashia, "Watermarking schemes for digital images: Robustness overview," Signal Process. Image Commun., vol. 100, p. 116523, Jan. 2022, doi: 10.1016/j.image.2021.116523.

[9] H. K. Albahadily, I. A. Jabbar, A. A. Altaay, and X. Ren, "Issuing digital signatures for integrity and authentication of digital documents," Al-Mustansiriyah J. Sci., vol. 34, no. 3, pp. 50–55, Sep. 2023, doi: 10.23851/mjs.v34i3.1278.

[10] Y. Qin and B. Zhang, "Privacy-preserving biometrics image encryption and digital signature technique using Arnold and ElGamal," Appl. Sci., vol. 13, no. 14, pp. 8117–8132, 2023, doi: 10.3390/app13148117.

[11] C. Zhan, L. Leng, C.-C. Chang, and J.-H. Horng, "Reversible image fragile watermarking with dual tampering detection," Electronics, vol. 13, no. 10, Art. no. 10, Jan. 2024, doi: 10.3390/electronics13101884.

[12] D. P. Kusumaningrum, E. H. Rachmawanto, C. A. Sari, and R. P. Pradana, "DWT–SVD combination method for copyrights protection," Sci J Inf., vol. 7, no. 1, p. 311, 2020, doi: 10.15294/sji.v7i1.21050.

[13] W. Wang, N. Zhang, S. Sun, and W. Wang, "Electronic certificate images forgery detection with electronic certificate images database based on NCC and SSIM algorithms," in International Conference on Image, Signal Processing, and Pattern Recognition (ISPP 2024), SPIE, Jun. 2024, pp. 924–931. doi: 10.1117/12.3033767.

[14] I. A. Sabilla, M. Meirisdiana, D. Sunaryono, and M. Husni, "Best ratio size of image in steganography using portable document format with evaluation RMSE, PSNR, and SSIM," in 2021 4th International Conference of Computer and Informatics Engineering (IC2IE), Sep. 2021, pp. 289–294. doi: 10.1109/IC2IE53219.2021.9649198.

[15] Y. Al Najjar, "Comparative analysis of image quality assessment metrics: MSE, PSNR, SSIM and FSIM," Int. J. Sci. Res. IJSR, vol. 13, no. 3, pp. 110–114, 2024, doi 10.21275/SR24302013533.

[16] Y.-J.-N. Gu, J. Zhang, Y. Piao, L.-J. Deng, and Q. Wang, "Integral imaging reconstruction system based on the human eye viewing mechanism," Opt. Express, vol. 31, no. 6, pp. 9981–9995, Mar. 2023, doi: 10.1364/OE.484176

Prioritized Self-Configuration for Self-Organized Sensor Networks

Rui Teng
Qilu Normal University
Shandong, China
e-mail: trqn.edu1@outlook.com

***Abstract*—Self-configuration of sequence order, including network addresses and data reporting times, is important in a sensor network. This process empowers sensors to dynamically assign short-length addresses, thereby enhancing energy efficiency in sensing-data reporting. In the absence of configuration servers, sensors self-organize into function roles of sever-client, enabling the dynamic formation of address servers to assign short-length addresses to sensors. This paper addresses the configuration distance problem with an aim of shortening the distance of configuration routes between the address server and clients. We propose a prioritized self-configuration method that employs spatial-temporal control of configuration according to the topological distance to the client server in each round of address configuration. Numerical evaluations are carried out to verify the performance of the proposed method. The evaluation results show that the proposed method enables a significant decrease of up to 30 percent in the configuration distance.**

Keywords-Self-organization; smallest-size address; self-configuration; high-priority zone.

I. INTRODUCTION

Self-configuration in a Wireless Sensor Network (WSN) is crucial for several key benefits. It facilitates the assignment of unique identifiers to each sensor node, enabling individual device identification and communication. Furthermore, self-configuration enables the utilization of short-length addresses, which significantly enhances energy efficiency by reducing the overhead associated with data packet transmission. Because sensing data can be only a few bytes long, address size becomes a critical factor that directly influences the energy consumption of each transmitted packet [1] [2] [3] [4].

Numerous methods of address configuration exist within Internet of Things (IoT) networks and ad-hoc networks [5] [6] [7]. Dynamic address assignment based configuration technique is a practical and straightforward approach to assigning unique addresses within an ad-hoc network [7] [8] [9]. For mobile ad-hoc networks, mobility and network partitions pose significant challenges in addressing and configuration. To address these issues, numerous configuration methods have been developed [10] [11].

On the other hand, sensor networks generally exhibit a static topology but demand the use of short-length addresses or sequence number for energy efficiency [3] [12] [13]. The self-organized server-client functionality plays a key role in the configuration of short-length addresses or sequence numbers [2] [14] [15]. The unique and short-length address or sequence number can be configured by a dynamic server-client structure, which are self-organized among sensors. Wireless mutlithop routing are employed for communication between the address server and clients. The topological distance between the server and an client is called the configuration distance. The configuration distance can be represented by the route length in terms of hops between the address client and server.

In this paper, we address the configuration distance problem in a sensor network. A large configuration distance significantly increases communication resource consumption and introduces delays. This problem becomes particularly impactive in large sensor networks.

To solve this problem, we propose a method that controls the configuration correlation between sensor nodes. The basic idea is that each client starts to configure an address with a high priority by using a probability function to control its access to the address server. The clients avoid initiating a configuration request when the topological distance between them is substantial, considering both spatial and temporal aspects of network connectivity.

Numerical evaluations are carried out to validate the proposed methods. We implelement the proposed scheme of prioritized configuration in a C++ based simulation, with comparasion to the basic method of dynamic server based configuration [2]. The evaluation results illustrate the significant effectiveness of the proposed method in terms of configuration distance and configuration overhead.

The rest of the paper is structured as follows. In Section II, we present system model and basic concept. In Section III, we introduce the proposed method of prioritized configuration with spatial-temporal control. In Section IV, we introduce the numerical evaluation and present evaluation results. Finally, we conclude the article in Section V.

II. SYSTEM MODEL AND BASIC CONCEPT

A. Network Model

A sensor network can be represented as a graph $G = (V, E)$, where V involves sensor nodes, and E is the collection of wireless links between sensors. Each sensor node has capabilities of sensing, computing, and wireless communication. The address of each sensor node is configurable. To save energy consumption, the address of a sensor and its size can be set up in an on-demand manner rather than being a predefined long-size address before networking. Since sensor node has a power constraint and short data size in transmission, the size of address is not ignorable. In a self-organized sensor network, sensor nodes are expected to cooperatively perform the role of network infrastructure, automatically configuring sensor nodes into network.

B. The Basic Concept of Dynamic Configuration Organization Method

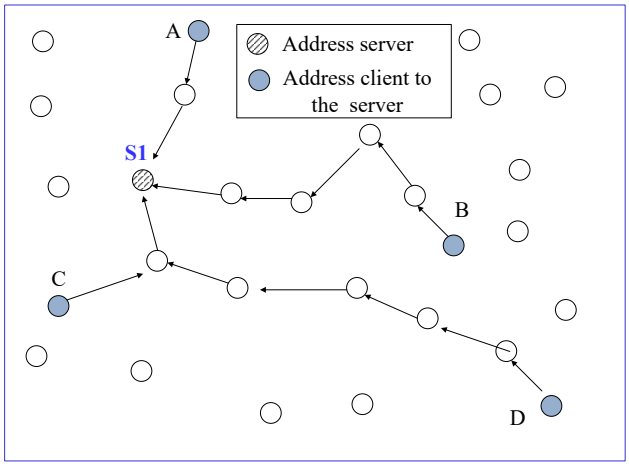
The dynamic configuration organization method attempts to configure a network-wide unique address for each sensor node [2]. In order to use a potential smallest address space, the self-configuration mechanism assigns address sequentially from low to high without the overuse of address space. Such a sequential assignment of node addresses desires a deterministic operation rather than an opportunistic operation so as to keep the consistency of address configuration in the self-organized sensor networks. Hence, a self-organized server-client structure is proposed. An address server is autonomously selected and serves the address configuration with a term limit. After a serving term expires, another sensor will be selected as the address server.

C. Problem

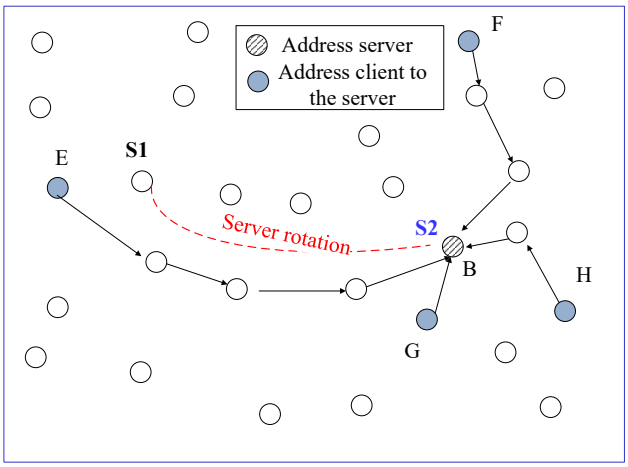
The problem addressed in this paper is the configuration distance problem in address configuration. The configuration distance is measured by the route length of address request and reply between address server and client. The configuration distance has a significant impact on energy consumption as well as the delay of address configuration. Meanwhile, the previous methods have not considered the efficient management of routes in the address configuration, leading to a scalability problem of the route length between the address server and the sensors that request an address.

Figure 1 shows an example of the configuration distance problem in dynamic self-configuration. In Figure 1(a), S1 is the address server. Node A, B, C, D are address clients that successfully issue address requests to S1. Node D needs to send the address request with a route length of 6, although there are nodes near to the server.

In Figure 1(b), S2 (node B) is the address server, which plays the server role after S1 transfers the server role to it. Node E, F, G, H are address clients that successfully issue address requests to S1. Node E needs to send requests with a route length of 4 although there are nodes near to the server.



(a) Address configuration with server 1



(b) Address configuration with server 2

Figure 1. Configuration distance problem in the address configuration of a large WSN.

III. Prioritized Configuration Method with Spatial-Temporal Control

We propose a spatial-temporal control based prioritized configuration method to the reduce the configuration distance. In a self-configuration procedure of a WSN, the server-client interaction is carried out using address request (AREQ), and address reply (AREP). For a sensor node, it needs to issue an AREQ to the address server with configuration contention among other sensor nodes. The proposed method differentiates the priority of configuration contention among sensors according to the topological distances between the address server and address clients.

As shown in Figure 4, there is a high priority zone around each address server. The priority zone contains a set of sensors that have a short topological distance to the address server. In the example shown in the figure, the radius of the high priority zone is two hops. The sensors in the high priority zone are assigned with a higher probability to obtain configured addresses from the address server than sensors outside the zone. In the example shown in Figure 4(a), the sensors A, B, C, D in the high-priority zone have the addresses configured by server 1. After the role of address server is shifted to server 2, sensors E, F, G, H, which are in the high priority zone have their addresses configured by the address server 2.

Algorithm 1 introduces the proposed algorithm for configuration control at sensor nodes of address clients. For each address server, address clients have different priorities to request the configuration response. Suppose that there is a time pool for a sensor node to issue an AREQ in a server term. In each time slot, a sensor node that has no address, attempts to request an address from the server. At time slot i in a server term, a node contents for a configuration seat by a probability control mechanism. The state of the configuration seat refers to whether the node can issue an AREQ in the time slot. There is a default

number pool in the range of $(0, Timepool1)$ for each node. A node randomly selects a control number from the configuration number pool. If the selected number matches with a predefined small number such as 1, the state configuration seat is then set to 1.

Therefore, to control the configuration probability based on the control number, we design a priority-based method based on the topological distance to the address server. According to the topological distance to the address server, the node zone is divided into two parts. The first part is the high priority zone, which is the set ZoneH that contains sensor[i] with the topological distance $TopoDist(i, ServerNow) < Threshold$. The second part is the low priority zone, which is the set ZoneL that contains the sensor[j] with the topological distance $TopoDist(j, ServerNow) > Threshold$. The $Threshold$ can be set to a certain value such as 3 hops, 5 hops, and so on.

In the high priority zone, the control number is generated as $ControlNum = Rand(0, Timepool2)$, where $Timepool2 = \frac{Timepool1}{k}$, where k is an integer such as 10. The hit probability of configuration state being 1 is $1/(Timepool1/k) = k/Timepool1$. In the low priority zone, the control number is generated as $ControlNum = Rand(0, Timepool1)$. The hit probability of configuration state being 1 is $1/Timepool1$.

Algorithm 2 shows the dynamic configuration method based on the server term control at address servers. Note that, in a Self-configuration of a WSN, the server-client interaction is carried out in address request and address reply. The server term control allows the dynamic generation of address servers among sensors, enabling the energy balancing for the configuration service at address servers. The locality AREQ and AREQ brings out merits of short route-length, low resource consumption, as well as low vulnerability to transmission failure and recovery cost.

IV. Numerical Evaluation

We carry numerical evaluation by C++ based simulation, in which the proposed method is implemented. The basic simulation setup is described in Table I. Two approaches are studied: the basic method of dynamic server based configuration [2], and the proposed scheme of prioritized configuration. The main metric employed in the simulation is the configuration distance in terms of hops, and the configuration overhead. The main evaluation target is to verify the efficiency of address or sequence auto-configuration in terms of configuration distance and configuration overhead. Figure 5 shows the setup of topoplogy of the network with 90 nodes in the evaluation. The second network scenario that employs 160 nodes has a topology of the same width of 100 m to the network with 90 nodes, but with a height of 160 m.

Figure 6 shows the average configuration distance (route length of AREQ/AREP) of sensor nodes in the scenario in which the network size is set to 90. For the conventional method of dynamic server based configuration, the average route length of configuration is about 6. In the proposed method of the prioritized configuration, the shortest configuration distance is achieved when the radius of the high-priority zone is set to 5 hops. With high-priority zone being set to 5 hops, the the results of configuration distance is 4.11 hops, which reduces 30 percent of configuration distance (route length) in the configuration. A very small priority-zone leads to the few nodes are enabled for prioritized configuration. A very large priority-zone weakens the impact of prioritized effect in configuration, leading to that the node faraway from the address server also have high probability to get an address being successfully configured

Algorithm 1 Prioritized configuration control at each sensor node.

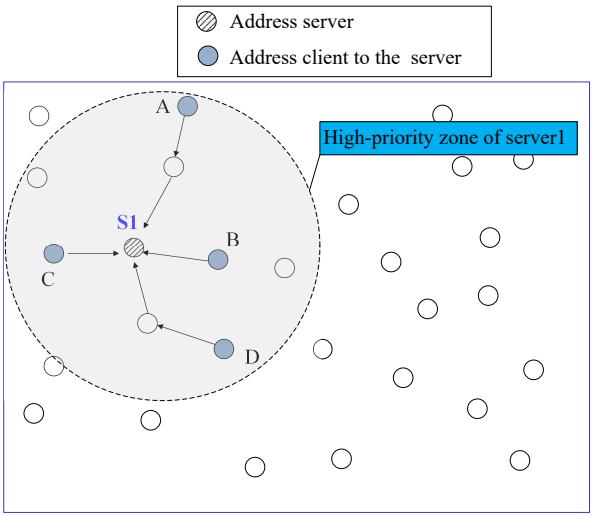
```
Input: network topology, Timepool1, TimePool2
Timepool2 ← Timepool1 / K
if ConfigurationState == 0 then
    if InHighPriorityState == 1 then
        ConfigSeat ← Random(0, Timepool2)
    end if
    if InHighPriorityState == 0 then
        ConfigSeat ← Random(0, Timepool1)
    end if
    if ConfigurationSeat == HitNum then
        Issue a request to the address server
    end if
    if
        Obtaining an address from the server then
        ConfigurationState ← 1
    end if
end if
```

Figure 2. Algorithm of prioritized configuration control at each sensor node.

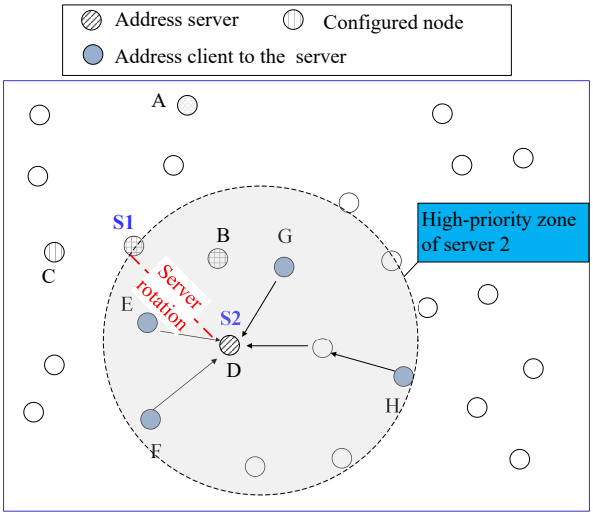
Algorithm 2 Serving term control.

```
Data Input: Serving term, network topology
for i = 1 → N do
    ConfigurationState[i] ← 0
end for
ServerID1 ← RandomlySelectedID
ConfigurationFinishState ← 0
while ConfigurationFinshState == 0 do
    if ServerRotationState == 0 then
        Configuration of the selected nodes
        if ServingCount == Term then
            ServerRotationState ← 1
            LastID ← LastConfiguredID
            NextServerID ← LastID
        end if
        Update ConfigurationFinshState
        Update server rotation state
    end if
end while
```

Figure 3. The algorithm of serving term control at the self-organized address servers.



(a) Address configuration with server 1



(b) Address configuration with server 2

Figure 4. Dynamic priority-zone based configuration control.

from the address server.

Figure 7 shows the average configuration distance (route length) of sensor nodes in the scenario in which the network size is set to 160. For the conventional method, the average

TABLE I. SIMULATION SETUP.

Basic Simulation Setup	
Parameters	Setup
Number of nodes in the network	90, 160
Server term	10
Communication range	10 m
Timepool1	200
Timepool2	20
Rounds of simulation	50

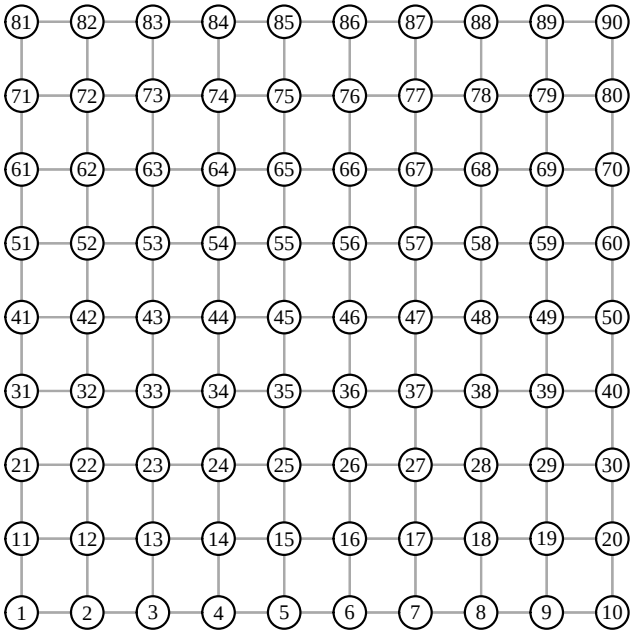


Figure 5. Evaluation scenario of 90 nodes.

route length of configuration is about 8.27. In the proposed method of the prioritized configuration, the shortest configuration distance is achieved when the radius of the high-priority zone is set to 5 and 6 hops. With the high-priority zones being set to 5 and 6 hops, the results of configuration distance is 5.47 hops, which reduces more than 33 percent of configuration distance in the configuration.

Figure 8 shows the average configuration overhead of sensor nodes in the the network with 90 nodes. For the conventional method of dynamic server based configuration, the average configuration overhead is about 5332. In the proposed method of the prioritized configuration, the minimum configuration overhead is achieved when the radius of the high-priority zone is set to 5 hops. With the setup of the optimal priority zone, the results of configuration overhead is 3803, which reduces 27 percent of configuration overhead in the configuration compared with the convention approach.

Figure 9 shows the average configuration overhead of sensor nodes in the the network with 160 nodes. For the conventional method of dynamic server based configuration, the average configuration overhead is about 16891. In the proposed method of the prioritized configuration, the minimum configuration overhead is achieved when the radius of the high-priority zone is set to 5 hops. With the setup of the optimal priority zone, the results of configuration overhead is 12210, which reduces 28 percent of configuration overhead in the configuration compared with the convention approach.

V. CONCLUSION AND FUTURE WORK

This paper addressed the configuration distance problem in a self-organized WSN. The large WSN desires a short configuration distance to avoid the large resources consumption

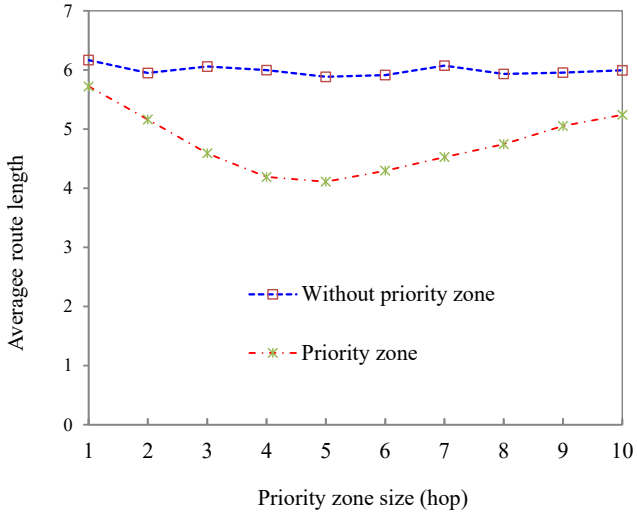


Figure 6. Average route length in network with 90 nodes.

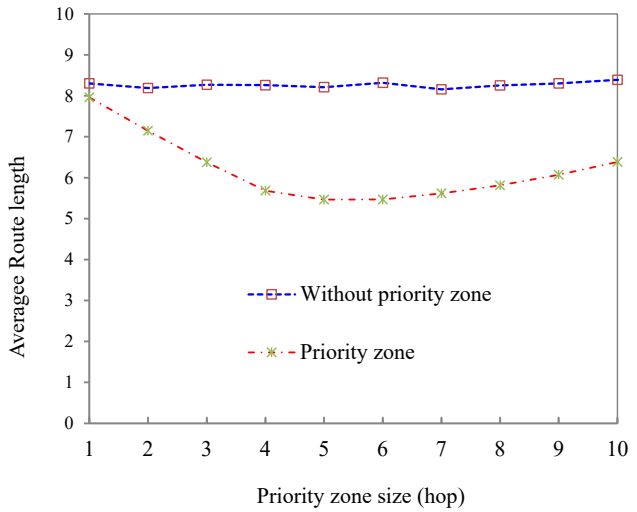


Figure 7. Average route length in network with 160 nodes.

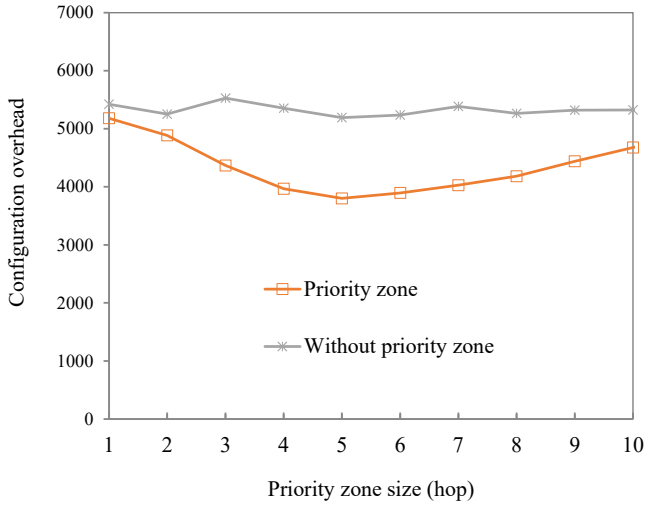


Figure 8. Configuration overhead in network with 90 nodes.

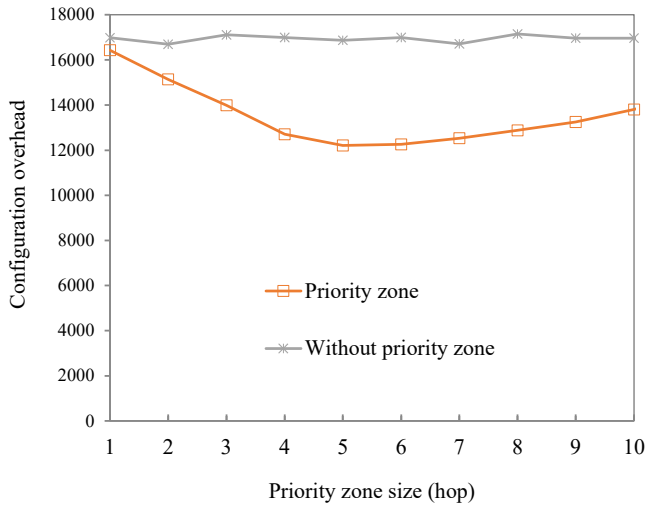


Figure 9. Configuration overhead in network with 160 nodes.

and link delay in auto-configuration of small-size addresses. We propose a prioritized configuration method with probability control based on spatial-temporal association between address clients and server. The evaluation results show the effectiveness of the proposed method in reducing the configuration distance. We find that there is an optimal setup of high-priority zone for the configuration to enable the shortest configuration distance. The short configuration distance is considered to have a significant impact on reducing the configuration delay and energy consumption at sensor nodes. Future work includes the study of adaptive setup of the priority zone for optimizing the configuration performance.

REFERENCES

[1] *Part 15.4: Wireless medium access control (mac) and physical layer (phy) specifications for low-rate wireless personal area networks (lrwpans)*, IEEE std. 802.15.4.

[2] R. Teng, "Sequence configuration with self-organized function structure for networking of smart sensors," *IEEE Systems Journal*, pp. 2594–2604, 2019.

[3] C. Intanagonwiwat, R. Govindan, D. Estfin, J. Heidemann, and F. Silva, "Directed diffusion for wireless sensor networking," *IEEE/ACM Transactions on networking*, vol. 11, 2–16, Feb. 2003.

[4] M. Chaudhary et al., "Underwater wireless sensor networks: Enabling technologies for node deployment and data collection challenges," *IEEE Internet of Things Journal*, pp. 3500–3524, Feb. 2023.

[5] G. K. et al., "Ipv6 addressing strategy with improved secure duplicate address detection to overcome denial of service and reconnaissance attacks," *Scientific Reports*, vol. 14, no. 25148, 2024.

[6] Pragya and B. Kumar, "A novel ipv6 address generation method for iot network," in *2024 7th International Conference on Contemporary Computing and Informatics (IC3I)*, 2024, pp. 1156–1161.

[7] X. Wang and H. Qian, "A distributed address configuration scheme for a manet," *Journal of Network and Systems Management*, vol. 22, pp. 559–582, Oct. 2014.

[8] S.-C. Kim, "Study on minimizing broadcast redundancy in strong dad," in *Proc. Seventh International Conference on Ubiquitous and Future Networks*, 2015.

[9] J. Jeong, C. S. Hong, and S. Park, "Dns configuration in ipv6: Approaches, analysis, and deployment scenarios," *IEEE Internet Computing*, vol. 17, pp. 48–56, Jul. 2013.

[10] K. Weniger, "Pacman: Passive autoconfiguration for mobile ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 3, pp. 507–519, Mar. 2005.

[11] S. H. Bouk and I. Sasase, "Ipv6 autoconfiguration for hierarchical manets with efficient leader election algorithm," *Journal of Communications and Networks*, vol. 11, 248–260, Jun. 2009.

[12] K. Awan et al., "Underwater wireless sensor networks: A review of recent issues and challenges," *Wireless Communications and Mobile Computing*, vol. 2019, 2019.

[13] R. Teng, "Property examination of double-zone presenting mechanism in the recovery of iot sensor failure," in *The 8th International Conference on Computer Science and Artificial Intelligence*, 2024.

[14] S. Camazine et al., *Self-organization in Biological Systems*. Princeton University Press, Oct. 2001.

[15] F. Dressler, *Self-Organization in Sensor and Actor Networks*. John Wiley & Sons, Ltd, 2007.

 ISBN: 978-1-68558-304-0