Journal of Network and Computer Applications 242 (2025) 104267

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

Research paper

SafeAcc: A lightweight and accurate user identification scheme using location-identity learning for indoor Internet access

Mohsen Ali Alawami [a], Sang-Hoon Choi [b], Ki-Woong Park [c,*]

[a] Division of Computer Engineering, Hankuk University of Foreign Studies, Yongin-si, 17035, South Korea
[b] SysCore Lab, Sejong University, Seoul, 05006, South Korea
[c] Department of Information Security and Convergence Engineering for Intelligent Drone, Sejong University, Seoul, 05006, South Korea

ARTICLE INFO

Keywords:
Location information
Wireless network
User identification
Smartphones
Indoor environments

ABSTRACT

Wireless networks, especially Wi-Fi access points (APs), have been an indispensable part of our daily life. For example, cafes and restaurants provide free Wi-Fi access with/without password settings only for customers (legitimate users). However, undesired users who reside outside but still within the coverage can easily use the Internet for free, misuse the Wi-Fi network, or apply harmful threats to the connected users and devices. In case administrators protect Wi-Fi networks using the cryptographic keys in WPA2 and WPA3, they have two limitations: (1) secure Wi-Fi APs increase the burden on customers not only because of manually inputting passwords but also finding the location of Wi-Fi information (*i.e.,* SSID and password), and (2) the security of Wi-Fi APs is only as good as their passwords. To enhance the security and ensure the usability of Internet access in indoor environments, we present a lightweight and accurate identification scheme named `SafeAcc` that uses the smartphone's physical fingerprints sensed at the user's location inside the area-of-interest (AoI) and meanwhile hard for an attacker to mimic. In this work, we explore the synergistic cooperation of the user's identity and location information to adapt the networks to only grant Internet access to the legitimate users who reside within the AoI. Our idea is to use Wi-Fi signals and the light intensity readings of smartphones as physical fingerprints to identify users' locations via a location-identity learning protocol and grant or disable Internet access accordingly. To evaluate the feasibility of `SafeAcc` in real-life scenarios, we developed an android-collector application and collected a real-world dataset containing fingerprints (Wi-Fi and light scans) from 40 locations distributed in two large and adjacent neighbor areas in a building. The data collection process is repeated for ten rounds that extend for a period of a month (each round is conducted during three periods a day: Morning, Afternoon, and Evening) and using two smartphones. The experimental results show that the identification accuracy using Wi-Fi fingerprints is always higher than those of light fingerprints — F1 scores ranging from 92.4% to 99.2% for Galaxy Note5 across the ten rounds and the three periods (Morning, Afternoon, and Evening). Also, the results show a slight accuracy difference when changing the device to Galaxy S8 providing F1 scores ranging from 86.5% to 99.5%. In addition, results showed reliable performance when `SafeAcc` was evaluated against unseen fingerprints (*i.e.,* drift concept) collected after two weeks (achieved F1 scores ranging from 95% to 98.3%) and after five weeks (achieved F1 scores ranging from 97.5% to 98.7%) respectively. We also measured the identification time required for training and testing models to ensure the usability of the `SafeAcc` in real-world usage.

1. Introduction

With the proliferation of Wi-Fi routers and the development of mobile devices, wireless communication using Wi-Fi access points becomes the most popular and successful for ubiquitous wireless network access in indoor environments. This dramatically increases the adaptation of access connectivity to the Internet across various devices, including networking products, tablets, smartphones, and IoT devices. Furthermore, the wide and fast increase in Internet use in indoor environments (*e.g.,* restaurants, cafes, bookshops, schools, universities, shopping malls, *etc.*.) motivates the need for safe and usable network access connectivity solutions, where only legitimate devices can connect to the Internet

* Corresponding author.
E-mail addresses: mohsencomm@hufs.ac.kr (M.A. Alawami), csh0052@gmail.com (S.-H. Choi), woongbak@sejong.ac.kr (K.-W. Park).
URL: http://syscore.sejong.ac.kr/ (K.-W. Park).

https://doi.org/10.1016/j.jnca.2025.104267
Received 27 July 2023; Received in revised form 23 May 2025; Accepted 3 July 2025
Available online 17 July 2025

via Wi-Fi home networks.

Wireless networks can be discovered by anyone nearby. Specifically, connecting to the Wi-Fi access points is easy for users who physically reside within the network coverage. However, not only legitimate users can use the Wi-Fi traffic data, but nearby undesired users (*e.g.,* attackers) can also get access to the Wi-Fi network and use the Internet service once they are within the communication range (but outside the target physical area). The fundamental reason is that the actual communication range is much larger than the valid usage range. This situation creates security and privacy challenges such as slowing down the traffic speed, consuming the limited bandwidth of the Internet, and bringing threats to legitimate users and connected devices.

Therefore, security protocols and schemes are essential to prevent unauthorized access to wireless networks and provide secure communication ways for connected devices. For protecting Wi-Fi networks, WPA2 and WPA3 are the most widely used. However, the cryptographic keys used WPA2 and WPA3 are mainly derived from a user's password alone — in consequence, their security is only as good as the user's password. Furthermore, some examples of advanced authentication schemes are particularly relevant to scenarios like captive portals. First, the "EAP (Extensible Authentication Protocol)" scheme has various methods for network authentication such as EAP-TLS (Transport Layer Security) and EAP-PEAP (Protected EAP). The former enhances the security of password-based methods by using client-side certificates for authentication, while the latter secures username/password by encapsulating EAP protocol within the TLS tunnel. Another authentication scheme is "Passpoint (Hotspot 2.0) with WPA3-Enterprise" which reduces the reliance on web-based captive portals by using pre-configured credentials that enable automatic login to Wi-Fi networks. Also, the "MACsec (Media Access Control Security)" authentication scheme can be integrated with IEEE 802.1X for dynamic key management to ensure secure access control for wired networks. Lastly, the "OAuth 2.0 for Network Authentication" network scheme uses social accounts such as Google and Microsoft to use OAuth-based login instead of traditional username/password login.

On the other side, many vulnerabilities in wireless Wi-Fi networks are as follows. Moreover, administrators (especially for public indoor environments such as cafes, restaurants, bookshops, schools, and shopping malls) often choose to set too simple and easy-to-guess passwords or leave Wi-Fi APs available without passwords for any user. In the case of secured Wi-Fi APs, administrators usually display WiFi information (*i.e.,* SSID and password) in places where only customers can see it and update it regularly. Two drawbacks are that (1) customers need to enter the password manually which increases the burden on them, (2) sometimes they do not know the location of the information. In the context, several studies (Eian et al., 2020; Fikriyadi et al., 2020; Ee et al., 2020) conducted experiments to perform attacks and penetration tests against wireless networks demonstrating that Wi-Fi networks are susceptible to both passive (e.g., eavesdropping) and active attacks (e.g., denial of service).

Previous studies using audio steganography and acoustic signals techniques presented alternative solutions to imperceptibly broadcast Wi-Fi information through a speaker that continuously plays music at the place (Tan et al., 2019; Eichelberger et al., 2019a; Bai et al., 2020b,a; Cai et al., 2022). The intuition behind this is that acoustic signals attenuate rapidly through walls and only customers inside the target area can receive Wi-Fi information. Although audio steganography has its strengths as hiding information technology, it has weaknesses that prevent its applicability in the real-world application such as: (1) it requires continuously running music to keep hiding Wi-Fi information which is noisy in indoor environments, (2) it requires adding encoding and decoding OFDM circuits to the system for hiding and broadcasting Wi-Fi information which increases the complexity of the system's usability. Other works suggested Internet access control using visible light communication (VLC) (Pathak et al., 2015; Rehman et al., 2019; Blinowski, 2019; Kumar and Singh, 2019; Oyewobi et al., 2022) techniques which use light bulbs connected at room ceil as transmitters and users' phones as receivers. These works require additional and specialized modulation (installed at source light bulbs) circuits and demodulation modules (mounted on smartphones or laptops) for transmitting and receiving the data correctly. In other words, transmitting information via VLC requires modulating the visible light spectrum (400–700 nm) of LED lights at home, offices, and public spaces that are used for illumination and network access control. The modulation process requires adding special-designed electronic circuits to the LEDs for data transmission through the unlicensed visible light spectrum. We know that the spectrum of the lights is more concentrated as the line of sight and provides reliable transmission for VLC networks consisting of nodes with different Field of View (FOV). However, the lights cannot cross walls and should be restricted within the same area (e.g., room, office, lab). Therefore, the data transmission is valid as long as the receiver demodulation circuit (e.g., photodiodes) is located within the lighting area and near enough to the LED source bulbs.

Recently, location-based services (LBSs) have become attractive attributes and experienced a surge such that users can use wireless location-dependant information (e.g., Wi-Fi, Cellular, or visible lights) for various applications such as localization, tracking, navigation, identification, and authentication (Alawami and Kim, 2020; Ali et al., 2018; Chen et al., 2019; Zafari et al., 2019; Alawami et al., 2022, 2019, 2020). However, location-based network access techniques have not designed yet to operate alongside existing wireless security protocols in order to strengthen protection mechanisms of data transfer for Wi-Fi devices. Particularly, the access connectivity to the network can be restricted to a single indoor location (or room) by exploiting the location fingerprint information of users who are located within the targeted area. This will ultimately provide location-based access to the Wi-Fi access points that require to share of sensitive and confidential data. In addition, there is no system defined in terms of the size of the coverage area that the users are allowed to connect to the network. This means an attacker or undesired user, as long as he resides within the coverage area of the Wi-Fi AP, still can get to access a specific Wi-Fi node and use the network.

The objective of this work is to mitigate this ongoing problem by developing an accurate and usable user identification system to continuously associate the identity of legitimate users with valid area-of-the-interest and distinguish them from those nearby undesired users who should be disconnected from the network. Our motivation emerges from the fact that legitimate and undesired users have different physical conditions such as location areas that they are staying in for the most of time (e.g., inside/outside the cafe and inside/outside the restaurant), leading to a different history of the location fingerprints profiles. By assuming that undesired users are able to crack the Wi-Fi passwords and get access to the network, our method using location information still can predict their identities and remove the devices from the network as long as they are keeping connecting to the network from outside of the target area. Regarding the applicability, our method can be adopted in restaurants, cafes, secure offices, research labs, classrooms, and homes which would limit/restrict the connectivity access coverage of Wi-Fi networks to the users who reside only within these trusted regions (room or sub-room). Once the model correctly identifies a user's location fingerprints inside the area of interest (AoI), his/her smartphone stays connected to the network. Therefore, we basically convert the application of user identification inside target areas (*e.g.,* cafes and restaurants) into a classification problem to determine user identity as a legitimate or undesired user based on their location fingerprints. The success rates of identifications are as good as the classification accuracy produced by system models.

In this paper, we specifically aim to advantage of the concept of using location information to develop a usable and secure system that enhances the connection's security of the Wi-Fi network in a target

area-of-interest (AoI). Here, location information based on RSSI and light features is used for identification purposes. Once a user is identified as legitimate inside the target AoI, the Wi-Fi keeps connected and provides safe Internet access. We present SafeAcc aims to develop a protocol, that only uses the location information (i.e., RSSI and light fingerprints) to identify the user's identity (i.e., legitimate or undesired) to validate his/her connection to the Internet on a specific Wi-Fi AP inside target indoor environment. To ensure the lightweight of the SafeAcc, we rely only on commercial off-the-shelf (COST) Wi-Fi routers and smartphones without any modification on the AP side or requiring any additional devices.

Indeed, the classification of locations at legitimate and undesired areas that are adjacent and neighboring is a challenging task because of three reasons: ① Inside the legitimate areas, some locations may show different fingerprints patterns of Wi-Fi and light from other locations, and in the meanwhile, they are closely similar to those outside locations (*i.e.,* attackers' location), which leads to confusing the models and raising false classifications (false positives). ② In contrast, inside the undesired areas, some of the attackers' locations may show fingerprint patterns similar to those inside the target area (especially locations at the border that expect attackers to reside), which leads to the wrong detection (false accept) by the model and hence granting Internet access. ③ We found that fingerprints of Wi-Fi and light change over time even if they are sensed at the same locations. This reduces matching scores between the fresh scans (*i.e.,* online) and the stored scans of location fingerprints in the database (*i.e.,* trained models), which negatively affects the classification accuracy. Specifically, the change in the network properties due to network usage affects the quality of Wi-Fi fingerprints in terms of the amount of sensed network nodes and their RSSI values as well as changes in daily weather conditions affect light readings. ④ The availability of many smartphone types leads to variations in Wi-Fi and light values due to different manufacturing model specifications.

We summarize our contributions as follows.

1. We present SafeAcc, a user identification scheme that enhances Internet access control of the Wi-Fi network in a target area of interest (AoI) using location information.
2. We develop a protocol that only uses the location information (*i.e.,* Wi-Fi RSSI and light fingerprints) to identify user identity (i.e., legitimate or undesired) and provide safe access to the Internet without requiring any additional hardware or user intervention.
3. To evaluate the performance of SafeAcc, we developed a data-collector application on smartphones and collected a real-world dataset of location Wi-Fi and light fingerprints from 40 locations of two neighbors and adjacent indoor areas.
4. To investigate the impacts of changing devices and weather conditions over time, we emphasized collecting the dataset using two smartphones over the course of ten rounds (*i.e.,* each round was conducted on a separate day) and through three periods per each round (*i.e.,* Morning, Afternoon, Evening).
5. We conducted extensive experiments and our evaluation results under various scenarios show that the identification accuracy using Wi-Fi fingerprints is always higher than those of light fingerprints — F1 scores ranging from 92.4% to 99.2% for Galaxy Note5 across the ten rounds and the three periods (Morning, Afternoon, Evening), while the results show a slight accuracy difference when changing the device to Galaxy S8 providing F1 scores ranging from 86.5% to 99.5%. In addition, to evaluate the SafeAcc against unseen fingerprints, we got average F1 scores ranging from 95% to 98.3% for data collected after two weeks while the data after five weeks achieved average F1 scores ranging from 97.5% to 98.7%. We also measured the identification time required for training and testing models to ensure the usability of the SafeAcc for real-world usage.

The remainder of this paper is structured as follows. In Section 2, we review related works. We present the design overview of SafeAcc in Section 3. Section 4 demonstrates the methodology details of SafeAcc. Section 5 shows experiment settings and results from various evaluation approaches. We finally provide our conclusion in Section 6.

2. Related work

In this section, we explain the related works proposed for identifying customers in side indoor environments (*e.g.,* cafes) for getting Wi-Fi access. We categorize them into three techniques: (1) Audio Stenography-based techniques, (2) Acoustic Signals-based techniques, and (3) Visible light communication (VLC) based techniques.

The study of audio steganography technology has been used in information hiding for broad applications and prospects. One of the most audio steganography methods is "LBS encoding" in which the least significant part of each digital audio sample is replaced with the equivalent secret message in the binary sequence (Tan et al., 2019). Zhang et al. (2019) provided a solution named "No more free riders" that aims to secretly share Wi-Fi information with only customers using acoustic signals inside cafes. They implemented a prototype of Wi-Fi providers encoding Wi-Fi information with acoustic signals and broadcast through a speaker to two customers' smartphones that receive the signals and decode them to connect to the Wi-Fi. Besides that this method needs the speaker to be running all time with high volume (dB) to keep broadcasting, their BER rates are very high reaching 20% within a maximum distance of one meter only. Eichelberger et al. (2019b) used the psychoacoustic masking effect that used OFDM subcarriers next to frequencies of high amplitude to transmit data within music played from loudspeakers. The system achieved a transmitting rate of 900 bits per second without any degradation in the sound quality of the music and the BER can be kept at 10% within a distance of 15 m. Similarly, Wang et al. (2021) proposed "ChirpMu" system that encodes information into chirp symbols and mixes them with audible music to share Wi-Fi secrets or coupons in shopping malls. The work was tested with a speaker and three smartphones and achieved BER up to 20% in a distance ranging from 2 to 10 m.

With advancements in acoustic-based sensing, recent studies have demonstrated the feasibility and effectiveness of using acoustic signals using audio infrastructures integrated into mobile and IoT devices for a broad of applications including localization, security monitoring, tracking, and human activity recognition (Bai et al., 2020b; Cai et al., 2022). Range finding and device-to-device data transfer applications using acoustic signals use near-ultrasound frequencies 18–24 kHz — this prevents acoustic applications from benefiting high-frequency ultrasounds (> 24 kHz) due to hardware limitations on commercial-off-the-shelf (COTS) mobile devices. Therefore, chen (Chen et al., 2018) presented the iChemo system that enables COTS mobile devices to sense high-frequency ultrasound signals — he customized the coprime sampling algorithm on COTS devices to detect the power spectral density (PSD) of high-frequency ultrasound signals. Nadeem and Uddin (2022) proposed "Acoustic-WiFi" framework that contains a Power Saving Mode (PSM) scheme (A2PSM) and a smart contention resolution scheme (Harmony) to develop more efficient Wi-Fi networks. A2PSM scheme addressed the inefficiency of the existing power-saving schemes in smart devices and leveraged acoustic signals to reduce the wakeup events of the Wi-Fi interface when it is in power-saving mode. The Harmony scheme was proposed to reduce the overall collisions that occur from the overhead of the Wi-Fi backoff among devices. From their experiments, Acoustic-WiFi saved up to more than 25% more power and gained 40% in throughput over the 802.11 Wi-Fi backoff scheme.

Inaudible Acoustic communication become a popular alternative to short-range communications and peer-to-peer manner such as QR and NFC — but suffers from limited throughput leading to limited employment on applications. Bai et al. (2020a) redesigned the mechanism of acoustic communication by remodeling the non-linearity of

the microphone of mobile devices using OFDM over multiple orthogonal channels with an ultrasound frequency carrier and increased the throughput rates by 12×on mobile devices. Nandakumar et al. (2013) proposed "Dhwani", an acoustics-based NFC system that eliminates the need for any specialized NFC hardware and only uses the microphone and speakers on mobile phones. Zhou et al. (2018) proposed a dual-channel communication system named "Dolphin" using real-time acoustic signals by leveraging the masking effects of human audio signals. Dolphin enables data communication between the speaker and microphone in a real-time and unobtrusive manner and supports an average data rate of 240 bps at 2 m. Zhu et al. (2019) proposed a method named "HyperEar" to find a small object in indoor environments such as a house or an office by using time-difference-of-arrival (TDoA) measurements over acoustic signals issued from the object (*e.g.,* speaker) and received by smartphone's microphones. Zhang et al. (2014) addressed the data confidentiality problem in the short-range communication technology on smartphones when used for contactless payment and device pairing. The author proposed "PriWhisper", a keyless and secure software-based using acoustic communication without a key-exchange phase by leveraging jamming techniques. Zhou et al. (2019) presented a practical attack for unlocking a smartphone's pattern lock using acoustic signals when the victim is drawing the pattern. The system utilized fingertip information contained in the acoustic signals sent from the speaker and received by the microphone of the smartphone that can be leveraged to infer the pattern.

In summary, although acoustic-based sensing is widely used for various applications, there are many practical drawbacks as follows. First, acoustic-based sensing requires the speaker that continuously plays music to keep hiding and sending secret information which causes a noisy environment for customers inside cafes. Second, it requires a few meters of distance between the speaker and microphone to keep low BER rates of communication. Third, several propagation properties of acoustic signals (*e.g.,* reflection, diffraction, and scattering) as well as the noise of ambient music (or sounds) affect the communication quality. Fourth, all acoustic-based sensing requires complex encoding with OFDM and decoding techniques that require additional circuits and increase the overhead of systems on mobile devices.

Visible light communication (VLC) is another paradigm where information is transmitted by modulating the visible light with high-speed data communication, unregulated and unlicensed spectrum, low latency, high capacity, and high security (Oyewobi et al., 2022; Rehman et al., 2019; Pathak et al., 2015). Recently, VLC technology has been considered an attractive alternative solution for solving challenges for various applications in 5G and IoT systems. The work in Suduwella et al. (2018) proposed a protocol for location-based Internet access using visible light communications technology (VLC) such that the Wi-Fi AP restricts transferring confidential and sensitive data to the users who are only located in the coverage of specific locations. This is done when the users are physically present under the VLC-based light bulbs installed in the ceils, point their devices (e.g., Laptops) that are capable to demodulate the secrete keys received from the VLC transmitter circuits, and eventually are granted access to the network whenever the location successfully authenticated. The work in Bakar and Dahnil (2017) proposed implementing a location-based authentication protocol using visible light communication to collect location information of the user within a small area (e.g., the room). They used the existing LED light-bulbs infrastructure with control circuits (transmitters) to transmit the shared secret keys to the smartphones (receivers). They only explained their work theoretically to explain how the user's location information can be used for authentication. The work in Zhao et al. (2017) proposed an authentication framework that uses visible light communications (VLC) as credentials to distinguish users for indoor Wi-Fi network connectivity. However, VLC-based works have drawbacks as follows: First, the protocols are designed based on VLC-enabled bulbs that require installing of specially-designed electronic circuits with the bulbs to perform the modulation/demodulation processes and share unique ID/secret keys to the receivers (smartphones). Although such electronic circuits are mandatory for VLC-based systems, it is expensive to be equipped in the real world with each light bulb (transmitter) as well as each smartphone (receiver). This limits the deployment of the system in real life. Second, the receiver devices require careful alignment with the transmitters (VLC-based bulbs) because of the narrow Field Of View (FOV) -i- this limits the mobility of users inside the area since the user is exactly required to be under the bulb. Third, to provide synchronized transmission of the shared ID/keys and achieve good connectivity between transmitters and receivers, the VLC bulbs required a specific arrangement design, which adds complexity to the communication protocol.

In contrast, our work overcomes the drawback of acoustic-based and VLC-based works and proposed a novel location-bade identification scheme "`SafeAcc`" that does not require playing music, has no limitation in transmission distance, does not require any additional hardware or circuits for encoding/decoding implementation, and has no limit on the user's mobility inside the target area. Our system is a classification-based and software-only approach that depends on sensing Wi-Fi and light fingerprints at any location inside the target area for identification.

3. `SafeAcc` system design

In this section, we introduce an overview of `SafeAcc`'s workflow, developed Android application, design goals, communication protocol, and threat model.

3.1. Overview of `SafeAcc`'s workflow

First, we demonstrate the concept of legitimate (*e.g.,* customers) and undesired (attackers) users that we considered in this paper. Without loss of generality, we consider the area-of-interest (AoI) inside a cafe, restaurant, company, office, or bookshop as a legitimate area (target). Also, all users who reside, work, or use these legitimate areas are considered legitimate users who provide valid location fingerprints inside the legitimate AoI. On the other hand, we consider adjacent areas such as open neighborhoods and nearby places are undesired areas. Both legitimate and undesired areas have many location points such that each location point represents a user seat or work cubicle, for example. In this work, we set unique IDs for each location point, as well as the area of interest to identify the collected location fingerprints and then determine of user's identity accordingly. We demonstrate the scenario of the user's identity determination and the system workflow of `SafeAcc` in Fig. 1. The process starts when a user located at an area-of-interest (AoI) connects to the network AP using his/her personal smartphone, simultaneously `SafeAcc` starts to collect the user's location fingerprints (Wireless Wi-Fi characteristics and light intensity readings) using an Android application installed on the smartphone — the collection process is implicitly (*i.e.,* background service) and shortly (*i.e.,* within a few seconds). Then, `SafeAcc` automatically forwards collected fingerprints to the third-party server (the server is under our control) for creating models of location-identity learning for user identification purposes. After that, `SafeAcc` decides which area is the location of the user's device belongs to. If it is a legitimate area, the `SafeAcc` keeps the device to stay connected to the network. In contrast, `SafeAcc` removes the user's device from the network when the server re-identifies (for more than one time) that the fingerprints of the location belonging to the undesired AoI where the malicious users are potentially located in. This process of identification can be repeated periodically every period of time (*e.g.,* couple of minutes) to track the user's location changes such as going in or going out of the target area. For example, an undesired user who already entered the cafe, identified correctly inside the cafe area, and connected to the network; can leave the cafe after some time to the neighboring areas to use-for-free the cafe's network. Since we assumed that the connection

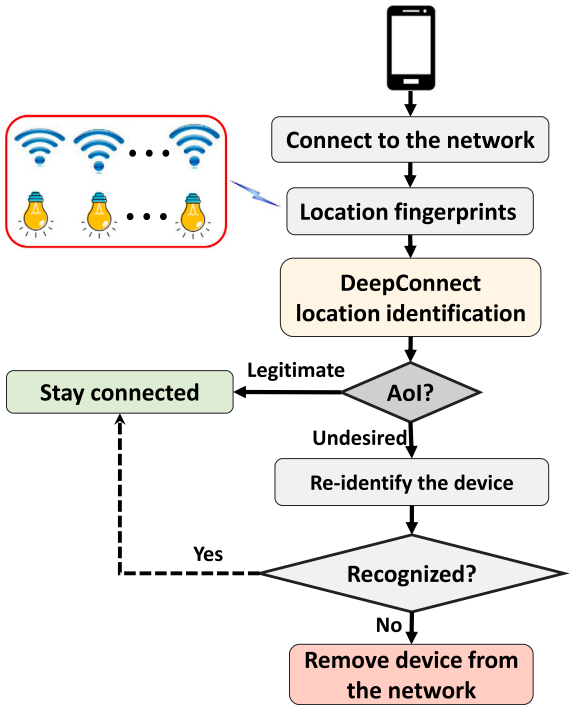
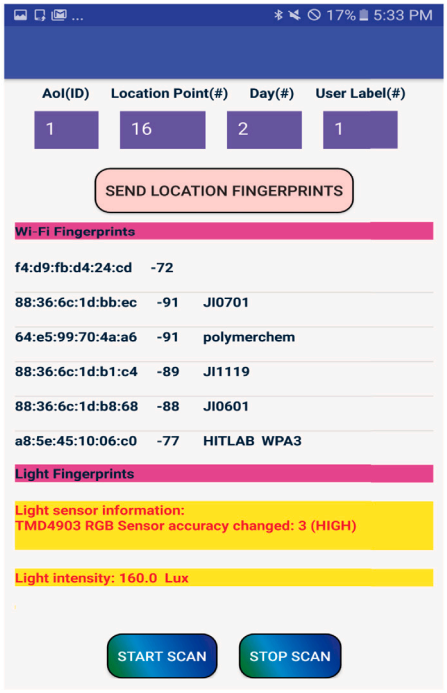
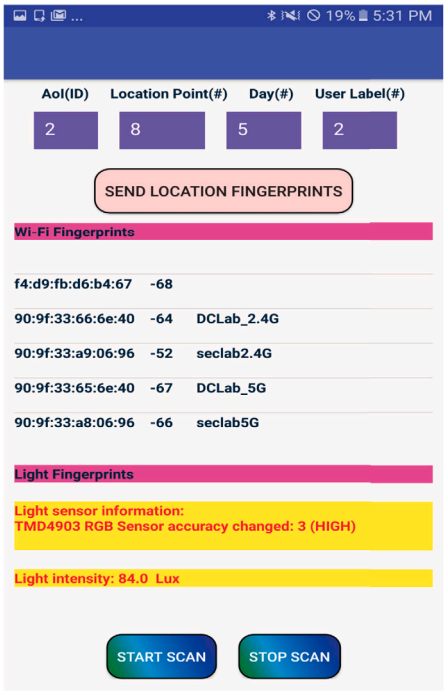


Fig. 1. A high-level overview of SafeAcc workflow process.



(a)



(b)

Fig. 2. Screenshots of developed SafeAcc android application for collecting location Wi-Fi and light fingerprints: (a) Legitimate area (AoI_1). (b) Undesired area (AoI_2).

to the cafe's network is through an application installed on the user's smartphone, the undesired user should have the application installed on his smartphone from the prior connection. This means that the SafeAcc periodically runs in the background and collects his location fingerprints and sends them to the server for periodic identification. If a user stops the application, the connection to the cafe's network will be lost. Therefore, our perception guarantees periodic identification and grants Internet access for authorized users as long as their location fingerprints belong to the legitimate AoI.

During the work, we found that some factors such as changing the AoI for the portion of time, device diversity influence, and instability of fingerprints readings over time are substantial reasons why SafeAcc needs to re-identify user's device periodically (shown in Fig. 1) before removing it from the network. For example, in situations when a user's device is identified as an undesired device due to either the instability of location fingerprints or the user's transitions among AoIs for long periods, SafeAcc re-collect location fingerprints to reconnect the network.

To collect wireless fingerprints observed at each location, we develop an Android application using Java language on the Android Studio platform and hence we install the app on the user's smartphone. Without loss of generality, SafeAcc aimed to be employed for realistic and ubiquitous scenarios such that it works at any smart places, with no requirement for prior get of the coordinates of surrounding network nodes or light sources, and no need for extra hardware. Fig. 2 shows the screenshot of the SafeAcc app that we used during data collection experiments which describes two types of fingerprints recorded at each location within each area-of-interest (AoI): Wi-Fi characteristics of surrounding wireless network nodes (e.g., BSSID, SSID, and RSSI) and light readings (Intensity values in Lux). Additionally, to arrange the datasets, we set four "Edit Text" settings to enter the ID of the interested area, the

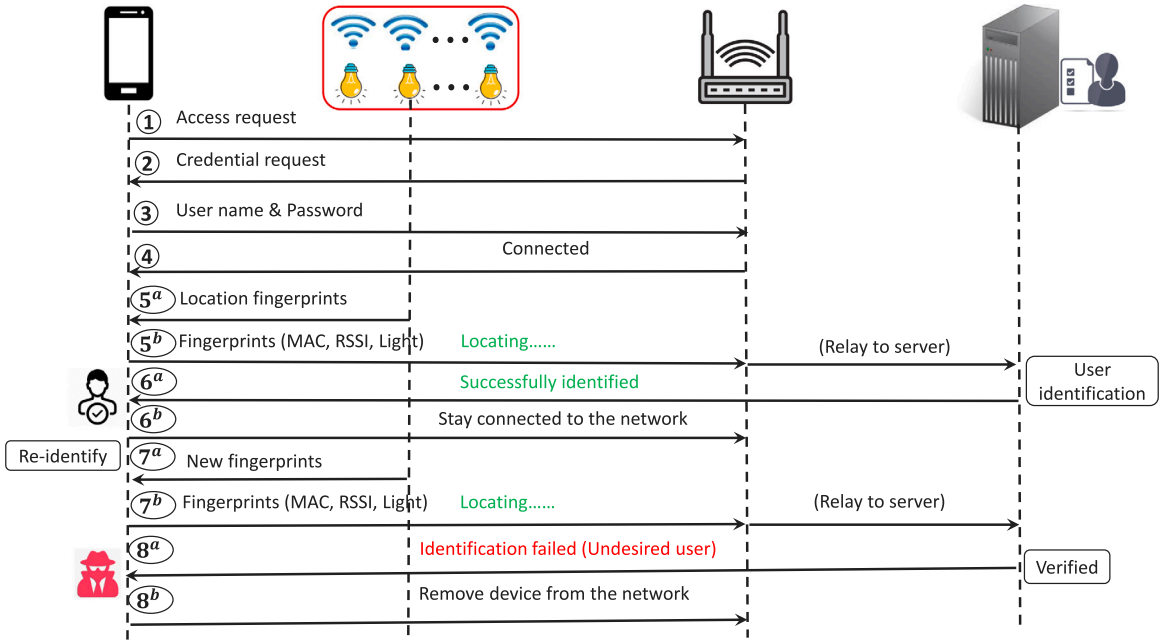


Fig. 3. An illustration of SafeAcc communication protocol.

ID of the location inside the area, the number of collection days, and the user's label (Legitimate or Undesired). Eventually, the app shows start/stop scan buttons to control the data collection process. All these recorded fingerprints of locations are automatically saved in (.csv) files and sent to our server for implementing SafeAcc learning.

3.2. Design goals and communication protocol

1. SafeAcc should not pose modifications to the Wi-Fi access point side nor require any additional hardware devices.
2. No need for passwords when connecting to the Internet, instead a user's smartphone can automatically connect when located inside the cafe's area.
3. SafeAcc uses physical and distinct features (i.e., location fingerprints) implicitly collected using the user's smartphone, which is difficult to mimic unless the undesired user is physically located inside the targeted area.
4. The system is costless (only relies on installed Wi-Fi APs and smartphones) and applicable everywhere in indoor environments.
5. The system is performed on the user side such that he/she recognizes the eligibility of the current location to gain access to the network. This operation will happen only once whenever a user visits the AoI.
6. To prevent the scenario of a non-customer user quickly entering and exiting the target area to authenticate his location and connect to the Internet, SafeAcc periodically (*e.g.,* every a few minutes) and implicitly collects new location fingerprints for identification.

We design a communication protocol of the SafeAcc system (shown in Fig. 3) to provide network access connectivity constraints based on an association between user identity and location information fingerprints. We explain the detailed sequence of the communication protocol as follows.

1. User/the smartphone requests network access from the Access Point (AP).
2. Network Access Point (AP) replies with a credential request to the user.
3. The device will send an authentication password and wait for connection.
4. Once connected, the smartphone implicitly (background app) collects current location information (Wi-Fi and Light) and sends it to the identification server.
5. Server uses location data to run pre-trained models for performing location-identity association to accept or reject the user's identity:
6. When the user is successfully identified, the server simply sends the identification result to the application which is responsible for keeping the user connected to the AP which means the user is permitted to access Wi-Fi by the server's response.
7. The above steps are repeated periodically (every a few minutes) and implicitly on the user's smartphone by collecting updated location fingerprints to re-identify the user's area without disturbing the user. For example, when the user temporarily leaves the cafe, then the SafeAcc immediately identify her as an undesired user, disconnects the device from Cafe's Wi-Fi for a while, and re-identify the user again after a period of a few minutes.
8. In contrast, when the identification fails for consecutive times and the result is detected as an undesired user, the server sends the response to the application and the connection with AP should be removed which means that the access is denied by the server's response.

It is a worthy note that the first three steps are for secured Wi-Fi AP with a password — in case of an open network, SafeAcc starts the protocol once the device is connected (*i.e.,* from step 3).

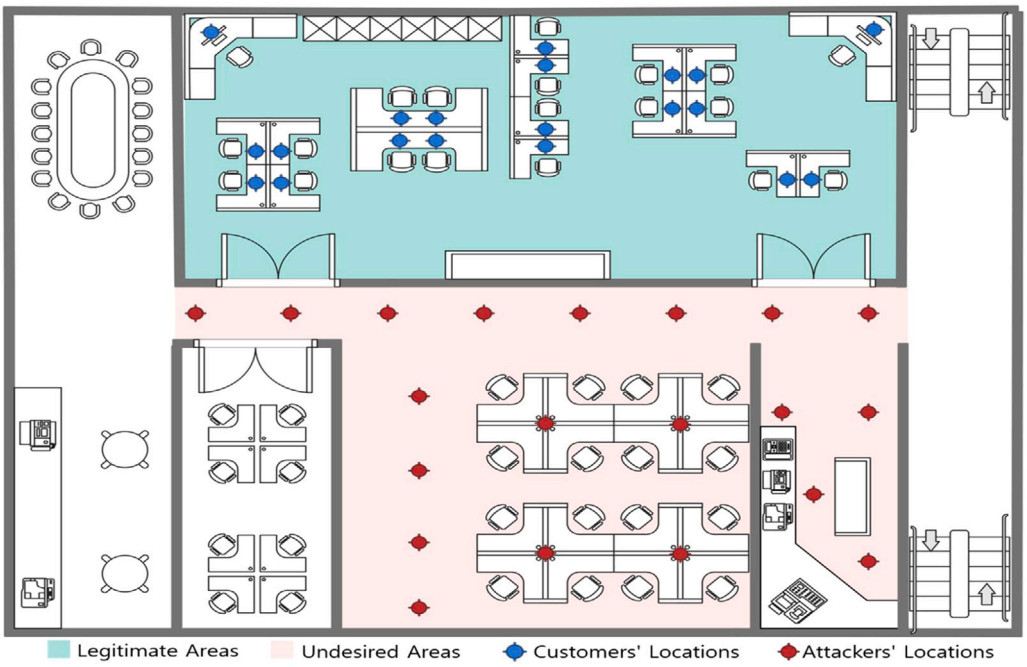


Fig. 4. Top view illustration of the tested environment layout inside a building.

3.3. Threat model

Some of the serious problems caused by nearby undesired users can potentially happen such as slowing the network, surfing the Internet for free, monitoring network traffic, or conducting harmful threats on the network. For example, in the US, the FBI's Internet Crime Complaint Center has released its annual report — it reported around 791,790 complaints of suspected Internet crime and losses exceeding $ 4.2 billion in 2020. In 2021, the FBI's Internet Crime Complaint Center (IC3) received 847,376 cybercrime complaints and US consumers lost $ 6.9 billion to internet crime, with an increase of 7 percent from the year before (Blog, 2022) (Blog, 2021). Additionally, some privacy and security-related issues may even result from harmful actions of nearby undesired users such as injecting/installing malware files to the network, changing/capturing MAC addresses of the connected legitimate users, collecting users-related network information, or shutting down the network service.

To address these threats, we propose an additional security method using location-related fingerprints as a second identification factor of the user to enhance the security level along with the existing network security protocols. Specifically, we expect that the location fingerprints of the undesired users would exhibit different characteristics and patterns that can be classified and distinguished from the location fingerprints of legitimate users. Thus, our `SafeAcc` model converts the identification process into a location-identity association problem using learning models in order to limit the Internet access of the Wi-Fi access point to only customers within the targeted area. As a result, any suspicious access connection to the network from undesired users should be rejected by the model as long as they are showing different location fingerprints that are detected from outside the target area.

4. Methodology of `SafeAcc`

In the section, we describe in detail the methodology of `SafeAcc`, including dataset collection, pre-processing, some of the motivating examples, and location-identity learning to create machine learning models.

4.1. Location dataset collection

For conducting data collection experiments, we choose two neighboring and adjacent areas: The first area (AoI_1) is a big hall that has 20 distributed locations to simulate customers inside a cafe or restaurant. The second area (AoI_2) is the surrounding space such as a corridor and open lounges that have 20 locations to simulate the adjacent locations of attackers who hear the Wi-Fi network of AoI_1. Fig. 4 shows the layout of the tested environment that we used to implement the data collection experiments for this work, highlighting AoI_1 with light-blue color and AoI_2 with light-red color. The blue circles mean the exact location we put the smartphone for each customer's location and the red circles mean the exact location of the attacker's location. To collect the dataset, we developed a prototype application for Android devices using the Android studio platform with Java. Specifically, our implementation of the app uses *BroadcastReceiver()* method from WifiManager for sensing Wi-Fi characteristics from all observed network nodes and *onSensorChanged* method from *SensorManager* for sensing light intensity readings (in Lux values) from the built-in light sensor on smartphones.

About the size of the dataset, we collected ten rounds distributed over five consecutive weeks. Each round was conducted on a different day, containing data from the 40 locations, repeated though three time periods [Morning: (9:30–12:00), Afternoon: (14:00–18:00), Evening: (20:30–23:00)] per day, and using two devices [(Galaxy S8 with Android 9, API 28) and (Galaxy Note 5 with Android 6, API 23)]. In detail, locations' fingerprints were collected in a cyclical movement starting from location 1 to location 40, (1 to 20 are labeled for AoI_1 and 21 to 40 are labeled for AoI_2). During each location, we conducted m scans ($m = 12$) for the length of one-minute time intervals (*i.e.,* one scan every 5 s) and then moved to the next location and so on. We emphasize that

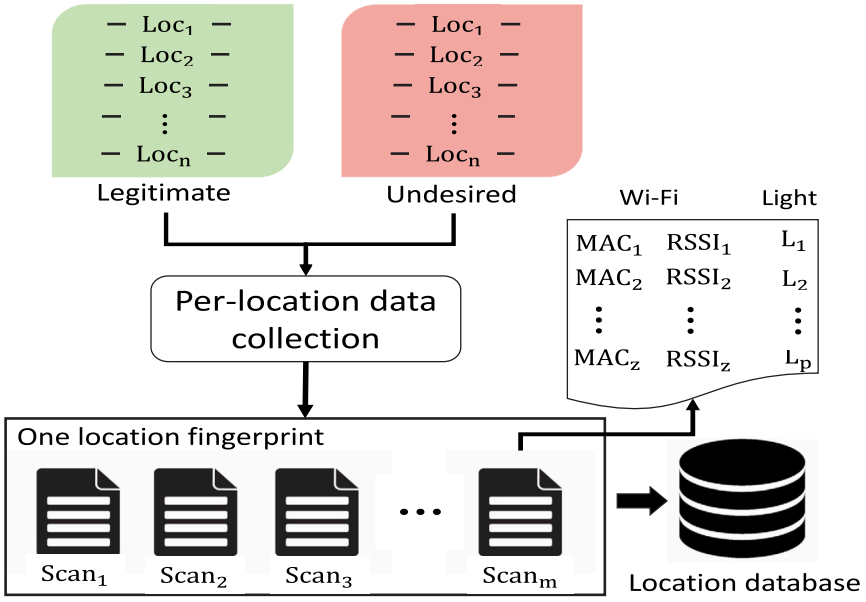


Fig. 5. Data collection stage of SafeAcc System.

the decision to collect locations' fingerprints in three periods a day (not hourly intervals) has been taken personally and only because of time limitations during the experiment. Specifically, we found it was hard with us to conduct experiments of collecting 40 locations' data in an hour. That is because side challenges consume more time so the hourly interval will not be enough and hard be hard to implement. These challenges were (for example) regarding arranging the rooms where the experiments were conducted, a large number of locations, moving smartphones around each location, and verifying that the dataset was received correctly by the server. Therefore, to conduct the experiments conveniently, we decided to use the three-period profile a day instead of hourly. Also, regarding weather conditions, as we explained the dataset was deliberately collected for ten rounds (i.e., ten days) distributed over consecutive five weeks (more than a month) to include data from different weather conditions such as cloud/sunny/rainy days.

As shown in Fig. 5, each scan captures two columns (MAC addresses and RSSI readings) of hundreds of network nodes observed from the surroundings and a column of light intensity readings in Lux values. After finishing m scans, we dump fingerprints into two .csv files (Wi-Fi.csv and light.csv) and save them into our server. In total, the whole dataset contains 4800 (.csv) files because we have 2 fingerprint types (Wi-Fi and light) × 40 locations × 3 periods per day × 10 rounds × 2 smartphones.[1]

Note: the experiments and data collection process of the work were conducted at two different but adjacent indoor areas (AoI_1, AoI_2) each of which has 20 physical locations as shown in Fig. 4. It is worth mentioning that we could not collect datasets from other indoor areas (or other buildings) because of difficulty regarding allocating rooms to run the experiments and collect the dataset for a month. However, we expect that if we conduct experiments in another building, the same process will be conducted. In other words, the models should be first trained on legitimate locations inside the trustworthy area, then should be tested for both legitimate and undesired locations (at neighboring locations). Thus, we do not think that testing models in another indoor environment (e.g., another building) would be a technical problem or affect the performance.

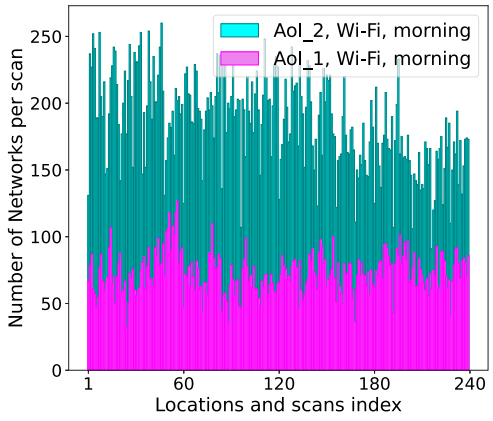
4.2. Preprecoessing of datasets

After collecting the whole dataset, we apply preprocessing steps on the data files to prepare data before inputting it into the learning models. First, we conduct an alignment step such that all Wi-Fi and light scans collected at a specific location during one iteration will be filtered and gathered separately. For example, for each iteration, we put smartphones in the required location and collect $m = 12$ scans of Wi-Fi RSSI and light readings. Thus, each location produces a matrix (dataframe) of $n \times m$ of RSSI values and a matrix (dataframe) of $n \times m$ light readings (n represents hundreds of rows and it is different from one scan to another even in the same location). After that, to sum all locations' data that are within the same area, we concatenate all the above dataframes, label them ("0" for the legitimate area and "1" for the undesired area), and save them into the server's database. We repeat the above steps for the dataset of the ten rounds and the two smartphones. So, for each round, we finally grouped the dataset into 12 larger (.csv) files named with three different variables [(AoI_1, AoI_2), (Wi-Fi, Light), (Morning, Afternoon, Evening)]. For example, one file is named "AoI_1-Wi-Fi-Morning" and another is named "AoI_2-Light-Evening". So, the total amount of data files after preprocessing is 12 files per each round × 10 rounds × 2 smartphones = 240 (.csv files). Each file includes data of one type of fingerprint (Wi-Fi or Light) that was captured from all 20 locations in a specific area (AoI_1 or AoI_2), on a specific day, during a specific period (Morning, Afternoon, or Evening) using a specific smartphone (Galaxy Note5 or Galaxy S8).

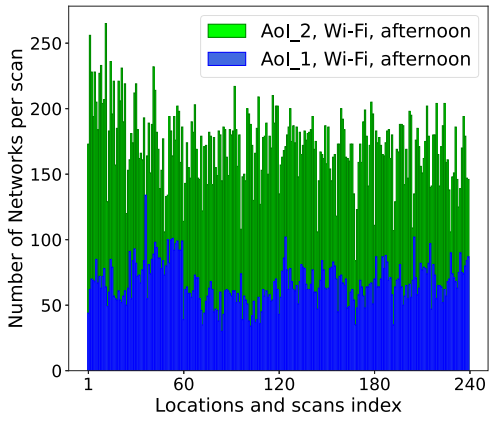
4.3. Motivating examples

To visualize the patterns of fingerprints captured inside each of the two selected areas, we plot an illustration of how many network nodes our data collector can detect. Specifically, we computed the number of

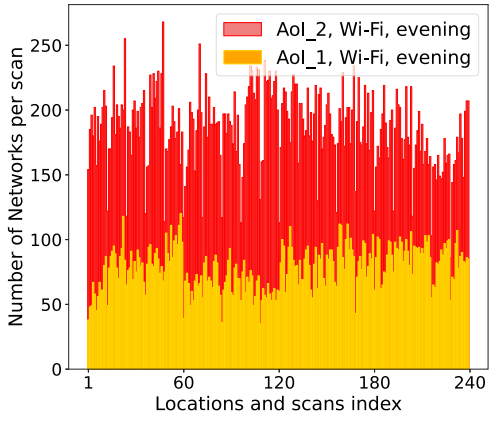
[1] The collected location-based dataset will be available upon request.



(a)



(b)



(c)

Fig. 6. An illustration of motivating examples that show differences of Wi-Fi fingerprints (number of sensed network nodes) collected at 40 locations in the two legitimate and undesired areas during three periods per day.

network nodes of $m = 12$ scans for 20 locations inside AoI_1 and AoI_2, and during the day's three periods (Morning, Afternoon, and Evening). Fig. 6 shows that the differences between the two areas (AoI_1 and AoI_2) are clearly observed in the number of detected network nodes over all 240 scans (12 scans per location × 20 locations per area). For example, in Fig. 6(a), we can see that during the morning time of a collection day, all scans in AoI_1 (magenta color) range up to 120 network nodes per scan. In contrast, we can see that the neighboring area AoI_2 (cyan color) which is outside, shows a high number of detected network nodes reaching up to 250 per scan. Note that, we count the number of network nodes based on their MAC addresses recorded during the scan process — they are unique and appear frequently. We repeated data collection during the afternoon period on the same day's round and extracted the number of detected network nodes for all 20 locations and in the two neighboring areas. Fig. 6(b) shows the distribution of the signal signatures in which the majority of scans that were captured inside AoI_1 (blue color) provide a number of network nodes less than 100 (most range between 70 ~ 80) and the minority of scans cross over 100. In contrast, we can see that the neighboring area AoI_2 (green color) shows the majority of scans can detect the number of network nodes between 150 and 200 and the minority of scans cross over 200 and even reach 250.

Lastly, to be more sure, we repeated the experiments during the evening with data collected from 40 locations (240 scans) as shown in Fig. 6(c). We found the same signal signatures are kept with the majority of AoI_1 (gold color) providing a number of network nodes less than 100 (most of them range between 50 ~ 70). In contrast, the neighboring area AoI_2 (red color) shows the majority of scans can reach network nodes up to 200 and some scans cross over 200 and even reach 250. We interpret the reason behind the phenomenon of detecting a larger number of network nodes in AoI_2 than AoI_1 because of the building geometry and structure. Specifically, the target area (*i.e.,* AoI_1) usually has a closed structure design surrounded by walls in which the smartphone can detect a small number of network nodes that are observed inside the area. In contrast, the undesired area (*i.e.,* AoI_2) usually has an open structure design in which the smartphone can detect a large number of network nodes that are observed from neighbor areas and buildings. From these motivating experiments, we can see that there is a feasibility of distinguishing AoI_1 from AoI_2 based on signatures of the network nodes sensed from surrounding environments using smartphones.

Similarly, we visualize the pattern of light intensity readings captured at the same 40 locations in the areas (AoI_1 and AoI_2). Smartphones can capture light data as discrete and integer values every time event (*e.g.,* seconds-level). For our experiments, we collected light-intensity scans at every location during the three-day periods (Morning, afternoon, and evening), each of which was represented in a separate column of data. Then, we used a boxplot to visually show the distribution of the numerical light data and skewness by displaying the data quartiles and averages for each location. Fig. 7 shows a boxplot for each column providing the median, Q1, Q3, IQR, minimum, and maximum values during the evening period in the two areas (AoI_1 colored with red and AoI_2 colored with gold). Circles indicate the outlier values that are distributed far away from the majority values detected at the same location. As we can see all locations of AoI_1 (indexed from 1 to 20) mostly provide average distribution light intensity values above 125 up to 225 Lux. In contrast, most of the other 20 locations (indexed from 21 to 40) locations located outside (*i.e.,* in the undesired area AoI_2) provide average distribution light intensity values lower than 125 except for four locations (locations 21, 22, 26, and 30). Additionally, we repeated the same experiments during morning and afternoon periods — the boxplots are provided in Appendix. We can see that during the afternoon all 20 locations of AoI_1 still show the average distribution of light intensity values above 150 Lux and all 20 locations of AoI_2 clearly show light intensity values below 150 Lux except three locations which are 21, 22, and 32. However, during

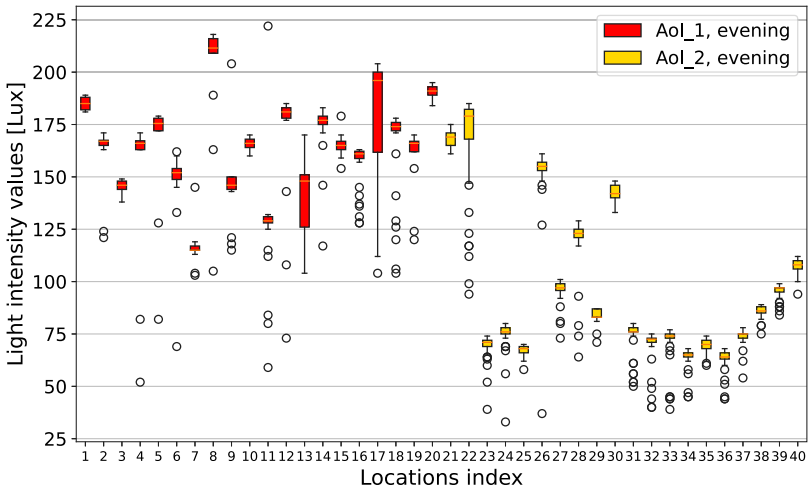


Fig. 7. An illustration of a motivating example shows differences of light fingerprints (intensity readings in Lux) collected at 40 locations in the legitimate and undesired areas during three periods per day (morning and afternoon plots are in Appendix, Fig. 12).

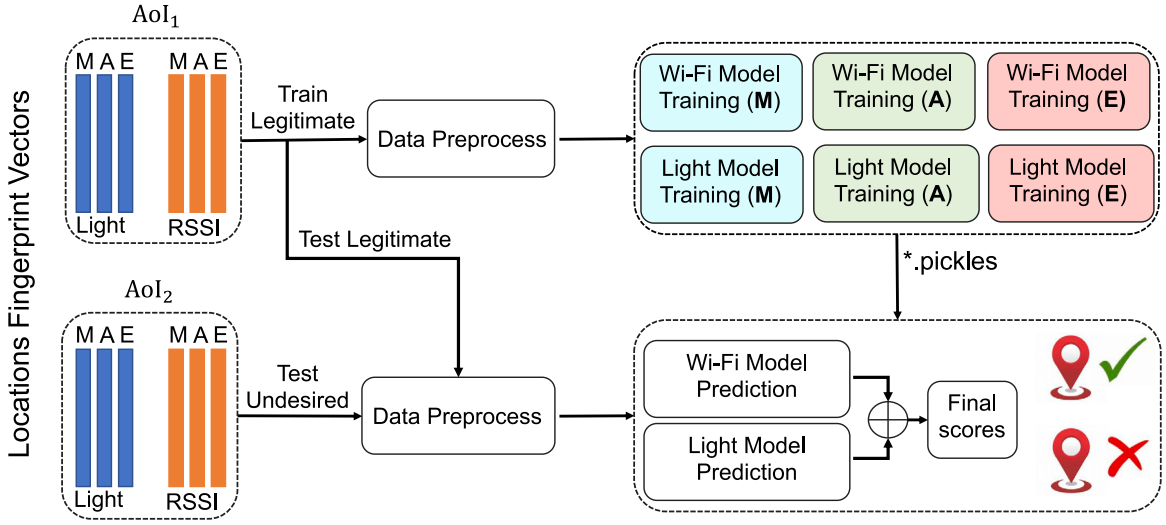


Fig. 8. The pipeline of location-identity learning including data preparation, model training, and location prediction.

the morning period, the distribution is not that clear for classification since many locations of both areas AoI_1 and AoI_2 show light intensity values above 150.

We are motivated by the findings of these preliminary experiments (especially Wi-Fi fingerprints) to build models and go forward with these light and Wi-Fi features for a location-identity identification system that can be used in indoor environments (*e.g.*, cafes).

4.4. Location-identity learning

To build practical location-identity models, we need first to build templates of customers' locations using their location Wi-Fi and light fingerprints inside the legitimate area (*e.g.*, cafes) to train models. In contrast, the location fingerprints of attackers who are located outside the legitimate area are unknown during the training phase. Since all neighbor areas can be potential undesired areas that attackers can use, it is hard to determine them and expect the exact attackers' locations.

So, collecting fingerprints from attackers cannot be applied in real-world scenarios because there are massive prospective undesired areas. Therefore, the binary classification approach used in the literature is no longer valid for building machine learning models for this work. The only available dataset for developing practical models during the training stage is the customers' fingerprints. Therefore, one-class detectors a.k.a. anomaly detection is the only possible solution to build a practical model with the customers' data to detect attackers whose fingerprints are deviated. In this work, we decided to select the most common one-class detector named One-class SVMs (OCSVM). One-class SVM is an unsupervised algorithm that learns a decision function for novelty detection by classifying new data as similar or different to the training set. OCSVM is one of the most convenient methods to approach one-class problem statements, including anomaly detection. It works on the basic idea of minimizing the hypersphere of the single class of examples in training data and considers all the other samples outside the hypersphere to be outliers or out of training data distribution. We implemented OCSVM through SK-learn from *libsvm*. The most hyperparameters of one-class SVM algorithm are *kernel* and *gamma*.

Fig. 8 explains the pipeline steps of location-identity learning used in the work to evaluate the performance of SafeAcc system in identifying and classifying fingerprints of legitimate locations from those of attackers. First, we prepared each location's Wi-Fi RSSI and light fingerprints in the form of six columns — three columns for Morning (M), Afternoon (A), and Evening (E) for the light and RSSI features. After that, we applied preprocessing steps such as building train dataframes containing columns of all locations inside the legitimate area (AoI_1). Then, we decided to create six models by separately training them on dataframes that have two properties: (1) Feature type (*i.e.,* Wi-Fi or Light), (2) Period type (Morning, Afternoon, or Evening). This is because we need to deeply investigate the performance impact of each case and which case is the most suitable for real-life usage. Note that, since we used one-class classification, we only used legitimate data of customers' locations to train models — no attackers' data are required for training which gives more practicality to our system. Lastly, we test trained models (pickles files) using the dataset collected from both areas (AoI_1 and AoI_2) for the purpose of performance evaluation and getting final prediction scores of location identification.

5. Performance and evaluation results

In this section, we first demonstrate the performance metrics and evaluation approaches that we used in this work. Then, we go into detailed explanations of experimental results that show the effectiveness of our proposed method.

5.1. Evaluation metrics and approaches

To evaluate the effectiveness of SafeAcc, we use the following metrics. *True negative (TN)*: The location fingerprints from users located inside the legitimate area (*e.g.,* cafes) are correctly identified by the system as customers' data and hence the SafeAcc system grants Internet access. *True positive (TP)*: The location fingerprints from attackers located outside the legitimate area are correctly identified as untrusted data and hence the SafeAcc system block disconnect/block the Internet access. *False positive (FP)*: The location fingerprints from users located inside the legitimate area are incorrectly rejected by the system. *False negative (FN)*: The location fingerprints from users located outside the legitimate area are incorrectly identified as trusted data and hence granted by the system to get Internet access. *Precision* is defined as in Eq. (1), which measures the portion of true positives divided by the summation of true positives and false positives. *Recall* is defined as in Eq. (2), which measures the portion of true positives divided by the summation of true positives and false negatives. The formula for the F1 score is shown in Eq. (3), which is a weighted average metric emphasizing the model's performance regarding false positives and false negatives.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \tag{3}$$

Our evaluation approaches to assess the performance of SafeAcc was based on several settings as follows. First, since we collected 10 rounds of datasets from the same 40 locations of the two areas (AoI_1 and AoI_2), we evaluated each round separately using six different data and period cases: (1) Light(M), (2) Light(A), (3) Light(E), (4) Wi-Fi(M), (5) Wi-Fi(A), (6) Wi-Fi(E). Second, to evaluate the robustness of the trained models against unseen data, we deliberately tested them using data collected at other rounds (different and far away from the training rounds). Lastly, to ensure the usability and applicability of the SafeAcc system, we computed the training and identification times of the models for the same six cases. Note that, we repeated all the above evaluation settings for the two tested devices (Galaxy Note5 and Galaxy S8).

5.2. Approach 1: Performance results

In this section, we investigate whether the SafeAcc system is feasible or not for area identification by computing the classification performance of location fingerprints. In other words, we conducted evaluations of the 10 rounds in which models are trained and tested using the dataset collected during the same round. Additionally, we considered six evaluation cases (*i.e.,* six OCSVM models): (1) Light(M), (2) Light(A), (3) Light(E), (4) Wi-Fi(M), (5) Wi-Fi(A), (6) Wi-Fi(E). For example, for the case of Wi-Fi(M), we use part of the round 1 Wi-Fi fingerprints that were collected from all 20 locations in the AoI_1 (one-class detection) during the morning period and tested them using Wi-Fi locations datasets collected from the two areas (AoI_1, AoI_2) on the same round and period (Round1, morning). All data samples of the legitimate locations are labeled as "0" and all data samples of attackers' locations are labeled as "1". We calculated F1, precision (Pr), and recall (Re) scores for each case and round independently. This process is repeated for ten rounds.

Table 1 shows the evaluation results of the 10 rounds dataset containing Wi-Fi and light fingerprints collected at 40 locations in two neighbor areas. The six evaluation cases (*i.e.,* six OCSVM models) are shown in the "Data (period)" row in the table. For example, in round 1, the OCSVM model of case 4 (*i.e.,* Wi-Fi(M)) has been trained using 50% of the RSSI readings collected at only the 20 locations inside AoI_1 (one-class). Then tested using the remaining 50% data size of the AoI_1 and the same amount of the data collected from the 20 locations located at the attacker's area AoI_2. The model predicts the labels of the tested locations' data samples and calculates the performance scores of an F1 score of 94.9%, precision of 90%, and recall of 99.1%. Similarly, in the same round, the OCSVM model of case 1 (*i.e.,* Light(E)) was trained on the 50% of the light intensity values that were collected in the morning from all 20 locations of AoI_1, and then tested using light samples from the 40 locations of the two areas (AoI_1, AoI_2). However, the model of Light(E) predicts a lower performance of an F1 score of 83.3%, precision of 93.7%, and recall of 75%. We repeated evaluations using the same scenario for all six cases and over the course of the ten rounds as shown in Table 1. Note that, the data samples used for testing models are different from those used for the training. In other words, we trained the models using data collected in the 20 locations in AoI_1, then we used other data (different scans but at the same locations) from both areas for the testing. This is to simulate the real application where the administrator needs to take scans for all cafe locations just one-time during the training phase and save them in the

Table 1
Evaluation results of the 10 rounds location datasets using Galaxy Note5 smartphone.

Data (period)	Light (M)			Light (A)			Light (E)			WiFi (M)			WiFi (A)			WiFi (E)		
Metrics	F1	Pr	Re	F1	Pr	Re	F1	Pr	Re	F1	Pr	Re	F1	Pr	Re	F1	Pr	Re
Round 1: Train (Area_1)																		
Test (Area_1, Area_2)	65.6	50	95	**89.5**	94.4	85	83.3	93.7	75	97.9	98.3	97.5	97.9	99.1	96.7	95.9	92.9	99.2
Round 2: Train (Area_1)																		
Test (Area_1, Area_2)	80	93.3	70	76.9	78.9	75	89.47	94.4	85	92.4	99	86.7	95.2	91.5	99.2	95.7	98.2	93.3
Round 3: Train (Area_1)																		
Test (Area_1, Area_2)	83.3	93.7	75	70	70	70	86.4	94.1	80	97.4	99.1	95.8	96.2	95.8	96.6	98.7	99.2	98.3
Round 4: Train (Area_1)																		
Test (Area_1, Area_2)	86.4	94.1	80	85.1	76.9	95.2	**92.3**	94.7	90	97.1	97.5	96.7	96.6	99.1	94.2	**99.2**	99.2	99.2
Round 5: Train (Area_1)																		
Test (Area_1, Area_2)	**89.5**	94.4	85	82.6	73.1	95	86.5	94.2	80	97.9	97.5	98.4	**98.8**	98.4	99.2	95.7	99.1	92.5
Round 6: Train (Area_1)																		
Test (Area_1, Area_2)	85.7	81.8	90	77.8	87.5	70	81.1	88.2	75	92.4	99.1	86.7	98.3	97.5	99.7	94.9	90.9	99.2
Round 7: Train (Area_1)																		
Test (Area_1, Area_2)	78.9	83.3	75	85.7	81.8	90	84.2	88.8	80	96.7	95.2	98.3	96.6	99.2	94.2	93.4	88.2	99.2
Round 8: Train (Area_1)																		
Test (Area_1, Area_2)	80	93.3	70	87.8	85.7	90	90	90	90	96.6	99.1	94.2	92.5	87.9	97.5	98.3	98.3	98.3
Round 9: Train (Area_1)																		
Test (Area_1, Area_2)	84.2	88.9	80	86.5	94.2	80	82.9	80.9	85	**98.3**	99.2	97.5	95.7	99.1	92.5	97.4	99.7	95.8
Round 10: Train (Area_1)																		
Test (Area_1, Area_2)	84.2	88.8	80	87.2	89.5	85	90	90	90	**98.3**	97.5	99.2	97.2	95.9	98.3	96.6	99.1	94.2

Table 2
Evaluation results of the 10 rounds location datasets using Galaxy S8 smartphone.

Data (period)	Light (M)			Light (A)			Light (E)			WiFi (M)			WiFi (A)			WiFi (E)		
Metrics	F1	Pr	Re	F1	Pr	Re	F1	Pr	Re	F1	Pr	Re	F1	Pr	Re	F1	Pr	Re
Round 1: Train (Area_1)																		
Test (Area_1, Area_2)	61.9	59.1	65	83.3	93.7	75	79.2	67.8	95	94.9	90.9	99.1	97.4	99.1	95.8	88.5	80	99.2
Round 2: Train (Area_1)																		
Test (Area_1, Area_2)	80	93.3	70	80.9	77.3	85	87.2	89.5	85	86.5	76.8	99.2	**99.2**	99.1	99.1	89.1	89.8	88.3
Round 3: Train (Area_1)																		
Test (Area_1, Area_2)	82.1	84.2	80	80	80	80	88.4	82.6	95	96.3	95.1	97.5	96.3	95.1	97.5	97.4	99.1	95.8
Round 4: Train (Area_1)																		
Test (Area_1, Area_2)	83.3	93.75	75	86.9	80	95.2	85	85	85	97.1	95.2	99.1	97.2	95.2	99.2	97.1	99.1	95
Round 5: Train (Area_1)																		
Test (Area_1, Area_2)	81.1	88.2	75	88.4	82.6	95	80	72	90	**99.2**	99.2	99.2	97.5	96	99.2	92.2	90.4	94.2
Round 6: Train (Area_1)																		
Test (Area_1, Area_2)	**84.2**	88.9	80	78.9	83.3	75	84.2	88.9	80	96.4	93.7	99.2	97.2	95.2	99.2	96.3	95.2	97.5
Round 7: Train (Area_1)																		
Test (Area_1, Area_2)	76.2	72.7	80	**88.9**	83.3	95.2	83.7	78.3	90	97.1	99.1	95	94.4	90.7	98.3	**99.5**	99.2	99.8
Round 8: Train (Area_1)																		
Test (Area_1, Area_2)	81.1	88.2	75	86.4	79.2	95	83.7	78.3	90	88.1	89.7	86.7	96.7	95.2	98.3	95	95	95
Round 9: Train (Area_1)																		
Test (Area_1, Area_2)	80	80	80	80	68.9	95.2	84.2	88.9	80	97.2	95.2	99.2	94.7	92.8	96.7	92.6	86.9	99.2
Round 10: Train (Area_1)																		
Test (Area_1, Area_2)	83.7	78.3	90	86.4	79.2	95	**89.5**	94.4	85	93.1	90.5	95.8	92.7	86.9	99.2	95.5	92.9	98.3

server to create the models. In the testing phase, whenever a customer wants to get Internet access, the `SafeAcc` is required to take a new scan at his location to test the model on the server and predict the identity for granting/refusing the connection.

In general, by looking at the identification results shown in Table 1, we can see that the accuracy slightly changes from one round to another, among the three-day periods (M, A, E), and from light to Wi-Fi fingerprints. In detail, the best performance of Light(M) models (colored with cyan) is provided at round 5 of an F1 score equal to 89.5% (Bold-Black color), while the worst performance is given at round 1 of an F1 score equal to 65.6% (Bold-Red color). The performance of the light feature enhanced during the afternoon period (*i.e.,* Light(A) colored with the light-green) where the best performance is improved to the F1 = 89.5% at round 1 and the worst performance is given at round 3 with an F1 score of 70%. Lastly, by evaluating case 3 during the evening period (*i.e.,* Light(E) colored with the light-red), we found that the performance of identification using only light intensity measurements improved more and provided the best F1 score of 92.3% at round 4 while the worst performance is given at round 6 of an F1 score equal to 81.1%.

Similarly, for the other three Wi-Fi cases (*i.e.,* Wi-Fi (M), Wi-Fi (A), Wi-Fi (E)), we can see better overall performance than the light-based fingerprint. In detail, the best performance of the Wi-Fi(M) model (colored with cyan) is provided at round 9 and round 10 of an F1 score equal to 98.3% while the worst performance is given at round 2 of an F1 score equal to 92.4%. The performance is slightly enhanced during the afternoon period (*i.e.,* Wi-Fi(A) colored with the light-green) achieving the best F1 score of 89.8% at round 5, while round 8 gives the worst F1 value equal to 92.5%. Lastly, we evaluate the Wi-Fi fingerprint during the evening (*i.e.,* Wi-Fi (E) colored with the light-red) and achieved the best performance ever of an F1 score equal to 99.2%, while the worst performance decreased to only an F1 score of 93.4%. Note that the F1 scores represent identification results of fingerprints' labels of both legitimate and attacker locations, which means that 99.2% (for example) indicates that almost all labels data samples of the 40 tested locations have been classified correctly to their areas with less than 1% false positives and false negatives. This ensures the robustness of our proposed features and system models. We conclude that Wi-Fi-based results shown in cases Wi-Fi(M), Wi-Fi(A), and Wi-Fi(E) in Table 1 are much higher than those results of light-based as well as more stable for all ten rounds and day periods.

The above results were evaluated on the ten rounds dataset that was collected using only one smartphone (*i.e.,* Galaxy Note5). However, to provide practical evaluations, customers who visit the cafe could have different types of devices and smartphones. Therefore, we need to investigate whether different models of smartphones could affect the performance of our system or not. Typically, smartphones from different manufacturers are equipped with different models of wireless antennas on the boards as well as different light sensor specifications. These differences in the hardware of smartphones could lead to variations in the measurements of collected location fingerprints of Wi-Fi and light and hence may affect the identification results. Since we already collected datasets using two smartphones (as explained in Section 4.1), we repeated the above evaluation process but on the dataset of the ten rounds collected from the other smartphone (*i.e.,* Galaxy S8). Table 2 shows the evaluation results of the 10 rounds dataset containing Wi-Fi and light fingerprints collected at 40 locations in two neighbor areas using a Galaxy S8 device. By looking at the best F1 scores computed for each of the six cases and comparing them with those results in Table 1 provided by Galaxy Note5, we can observe that by subtracting the best F1 scores of Galaxy Note5 from the F1 scores of Galaxy S8 for the six OCSVM models, there are no big differences as follows. Light(M) = +5.3%, Light(A) = +0.6%, Light(E) = +2.8%, Wi-Fi(M) = −0.9%, Wi-Fi(A) = −0.4%, Wi-Fi(E) = −0.3%, where "+" sign indicates that the best F1 score of Galaxy Note5 is higher than the best F1 score of Galaxy S8 and the "–" sign means the opposite. Thus,

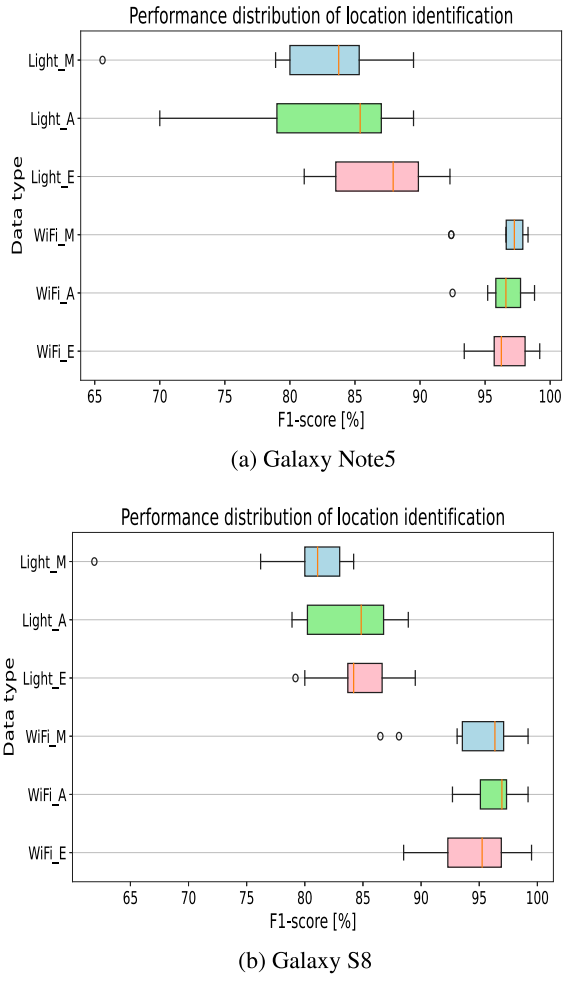


Fig. 9. Performance distribution of 10 rounds of locations datasets collected using two devices.

we conclude that the Galaxy note5 provides performance outperforms (up to around 5%) Galaxy S8 using the three light models. However, it is the opposite when using three Wi-Fi models where Galaxy S8 provides performance slightly (less than 1%) outperforms Galaxy Note5. In general, both smartphones show accurate results for identification especially when using Wi-Fi fingerprints (from 98.3% to 99.5%).

In the above explanation of results, we just focused on the best and worst F1 scores provided by the six OCSVM models over the ten rounds and using two smartphones. However, we need to describe the distribution of identification through the ten rounds to see to what extent the F1 score varies from one round to another. Therefore, we used box plots to graphically display five-number of parameters and statistics which are the minimum, first quartile, median, third quartile, and maximum. In a box plot, we draw a box from the first quartile to the third quartile. A vertical line goes through the box at the median. Fig. 9 shows box-plot graphs of the performance distribution of F1 scores over the ten rounds for the six OCSVM models. In Fig. 9(a), we draw the box-plots that represent the performance distribution of the six OCSVM models for Wi-Fi and light data using Galaxy Note5. In detail, we can see that Wi-Fi-based models over the ten rounds have less distribution (narrow range of Q3-Q1), which indicates the stability

of the system performance over the ten rounds that were collected on different days. Also, it is shown that all medians (vertical lines) are above 96%, and maximum values reach 99%. We noticed two outliers during the morning and afternoon periods down to 94%. In contrast, the performance of the light-based models shows a wide distribution of box-plot (large range of Q3–Q1) for the three OCSVM models in the morning, afternoon, and evening. This indicates that the performance changes from one round to another, which lead to less stability. Although light-based box-plots show an outlier and minimum values of F1 scores decrease to range between 65% and 70% respectively, the median values still keep performance between 85% to 90%.

Similarly, Fig. 9(b) shows the box-plots of the performance distribution of the six OCSVM models for Wi-Fi and light data using Galaxy S8. Here, it is the opposite. In detail, we can see that light-based models are more stable than those in the Galaxy Note5 because models provide a narrow range of Q3-Q1 through the ten rounds. But, Wi-Fi-based models are less stable because their box-plots seem a little wider ranges of Q3-Q1 than those in the Galaxy note5 for the ten rounds. The medians of the three light-based models range from 82% to 85%, while the medians of WiFi-based models range from above 95% to 97%. From approach 1 of evaluation, we conclude that our experimental results show the robustness and validity of our system even when evaluated under several impacts such as various ten rounds, three periods a day, and different devices.

5.3. Approach 2: Identification delay

In this evaluation approach, we investigate the effectiveness of the `SafeAcc` system under another impact, which is the required time for identification. In addition to evaluating the performance of the `SafeAcc` in detecting location fingerprints (as explained in the previous approach), we need to estimate how much time the models take for identification to ensure their usability in real-world applications. Achieving both high accuracy (security) and less time consumption (usability) are two important factors in designing models for the deployment and practical identification of locations. To do this, we computed the time consumed by models for the training and testing during the ten rounds. Fig. 10 shows the average train and test timing distributions of the ten rounds using the two tested devices. In detail, for each round, we calculated train time and test time during each day period of the light-based models (*i.e.,* morning, afternoon, and evening) and then average them to represent it as one point in the "Train_Light" and "Test_Light" curves. Similarly, we did the same thing for the three Wi-Fi-based models to create "Train_Wi-Fi" and "Test_Wi-Fi" curves. All times are recorded in milliseconds (ms).

By looking at Fig. 10(a) of Galaxy Note5, we can generally see that the average train and test times required for light-based models through the ten rounds are lower than those of Wi-Fi-based models. In detail, all ten rounds of datasets consumed training times below 4 ms and testing times below 2 ms for light-based models. In contrast, all ten rounds of datasets consumed training times up to 6 ms (except round 1) and testing times near 4 ms for Wi-Fi-based models. We repeated calculations for the ten rounds of datasets collected using Galaxy S8 smartphone as shown in Fig. 10(b), where training and testing times for both light and Wi-Fi models are slightly changed. In detail, for light-based models, most of the ten rounds of datasets consumed training times decrease to below 3.5 ms (except rounds 5 and 10) and testing times decrease to below 2 ms. Also, for Wi-Fi-based models, most of the ten rounds of datasets consumed training times decrease to below 4.5 ms (except rounds 3, 4, and 6), and testing times decrease to below 4 ms. From the approach 2 experiments, we conclude two observations: (1) Wi-Fi-based models always consume training and testing times higher than those consumed by light-based models for the ten rounds and using two devices. (2) There are trade-offs between identification accuracy (explained in approach 1) and identification delay (explained in approach 2). In other words, Wi-Fi-based models provide identification accuracy much higher than the accuracy provided by light-based models, while they consumed a little more training and testing times (in milliseconds) compared to the delays of light-based models.

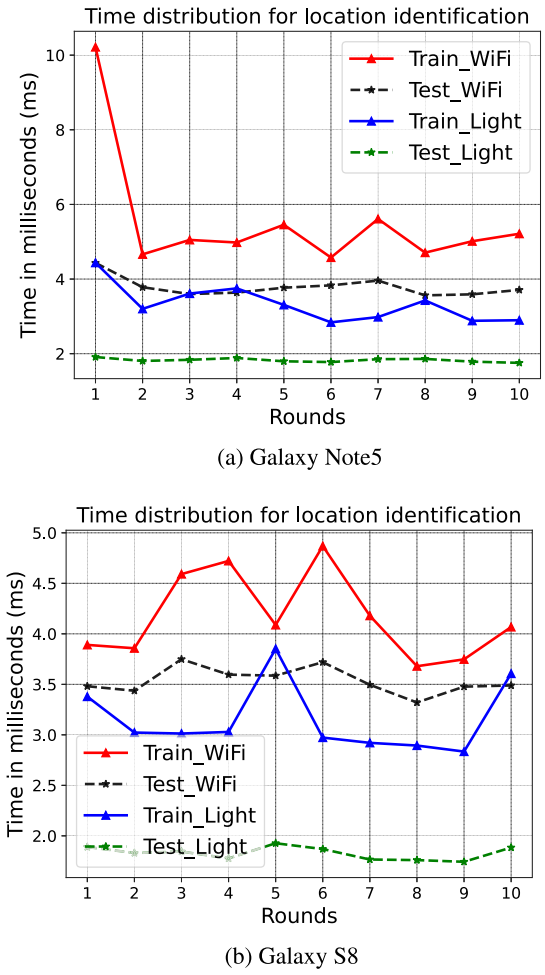
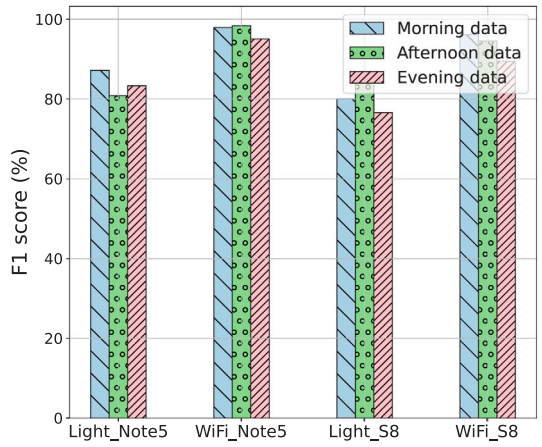


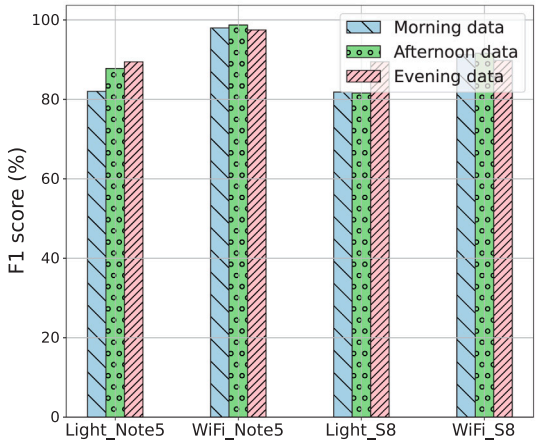
Fig. 10. Train and test time distributions of 10 rounds of locations datasets collected using two devices.

5.4. Approach 3: Unseen identification

In the experiments of above approaches 1 and 2, we considered training and testing models for each round independently because we focused on studying the feasibility and validity of the system. However, in reality, it is difficult for the cafe's administrator to collect Wi-Fi and light fingerprints from all cafe locations every day to train and build models on a daily basis. In approach 3, we addressed the *drift concept* in which we evaluated the performance of `SafeAcc` system on unseen fingerprints that were collected after several weeks of fingerprints used for training and creating the models. This is more realistic in that the cafe's administrator needs to collect location fingerprints just one time a month (for example), build models on the server, and customers' phones download them implicitly using `SafeAcc` application for online identification. After that, whenever the identification accuracy is decreasing over time (by increasing false positives and negatives),



(a) Train: R1, Test: R5



(b) Train: R1, Test: R10

Fig. 11. Performance Evaluation of the system when tested using unseen and time-gaped location fingerprints. (a) Train: R1 and test: R5. (b) Train: R1 and test: R10 for the two devices.

the Cafe's administrator needs to re-collect new location fingerprints and retain models for improving identification accuracy. The time gap between retraining the models should be for at least several weeks to avoid increasing the burden on administrators.

Since we have collected ten rounds of location fingerprint datasets separated over the period of a month, we focused on approach 3 to train models using the dataset collected in round 1 to simulate administrators who collect the cafe's locations just for one time (on the first day). After the models are trained using round1 fingerprints and saved into the server, we deliberately tested them using unseen fingerprints collected in round 5 (*i.e.,* after two weeks) and round 10 (*i.e.,* after five weeks) for the two smartphones to implement the case when trained models are downloaded into customers' phones at the first day and then used for identification using new fingerprints during a period of a month as shown in Fig. 11. By looking at Fig. 11(a) we can see that when testing the models using unseen fingerprints collected at round 5 (*i.e.,* after two weeks), Wi-Fi models still provide high F1 scores ranging from 95% to 98.3% using Galaxy Note5 and from 89.4% to 96.1% using Galaxy S8. However, the performance decreases for the light-based models to the range of 80.85% to 87.17% using Galaxy Note5 and from 76.6% to 85.71% using Galaxy S8.

Similarly, Fig. 11(b) shows the evaluation results when testing the models using unseen fingerprints collected at round 10 (*i.e.,* after five weeks). In detail, Wi-Fi models provide F1 scores ranging from 97.5% to 98.7% using Galaxy Note5 and from 89.8% to 91.6% using Galaxy S8. However, light-based models provide F1 scores ranging from 82% to 89.47% using Galaxy Note5 and from 81.6% to 89.47% using Galaxy S8. From experiments of approach 3, we conclude that even when evaluating the performance of our system using unseen datasets that were collected after two and five weeks from the collection time of the training data, `SafeAcc` still provides promising results that are reliable against the drift concept and then valid for practical identification scenarios.

6. Conclusion

In this paper, we tackle the problem of free riders — users who illegally use Internet access provided by Wi-Fi access points in indoor environments (*e.g.,* cafes or restaurants). Our goal is to only grant free and implicit (*i.e.,* without burdening customers to manually input Wi-Fi passwords) Internet access to customer users who reside inside the target area and disable connection from all undesired users (attackers) located outside. To do this, we proposed a location-based identification system named "`SafeAcc`" on mobile devices in which sensed fingerprints (Wi-Fi and light) at all target area locations should be classified correctly and hence locations are identified as legitimate. To evaluate the validity of our method, we developed an Android application and collected a real-world and large-scale dataset from 40 locations in two neighboring areas (one legitimate and one underside) for ten rounds distributed over a month using two smartphones. Our experimental results show the robustness of our identification models under various evaluation approaches including per-round performance, identification delay, and unseen identification.

CRediT authorship contribution statement

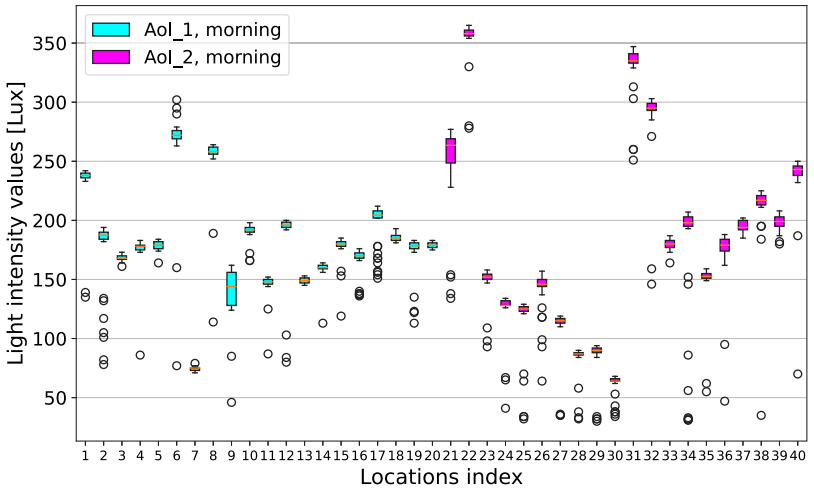
Mohsen Ali Alawami: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Sang-Hoon Choi:** Writing – original draft, Software. **Ki-Woong Park:** Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

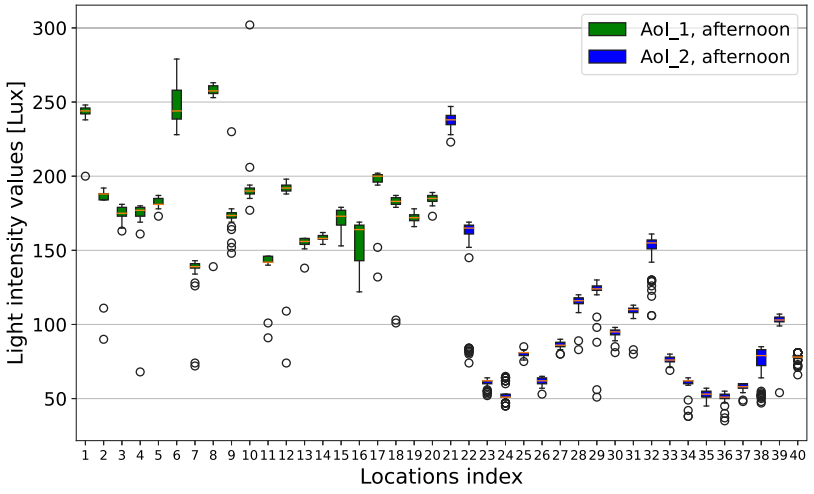
The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Ki-woong Park reports was provided by Sejong University. Ki-woong Park reports a relationship with Sejong University that includes: employment.

Acknowledgments

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) (Project No. RS-2023-00228996, 30%; RS-2024-00438551, 20%; IITP-2025-RS-2021-II211816, 10%), the Culture, Sports and Tourism R&D Program through the Korea Creative Content Agency grant funded by the Ministry of Culture, Sports and Tourism in 2025 (Project Name: Training Global Talent for Copyright Protection and Management of On-Device AI Models, Project No. RS-2025-02221620, Contribution Rate: 20%), and the National Research Foundation of Korea (NRF) grant funded by the Korean Government (Project No. RS-2023-00208460, 20%).



(a)



(b)

Fig. 12. An illustration of a motivating example shows differences of light fingerprints (intensity readings in Lux) collected at 40 locations in the two legitimate and undesired areas during (a) Morning, and (b) Afternoon.

Appendix

See Fig. 12.

Data availability

Data will be made available on request.

References

Alawami, Mohsen A., Aiken, William, Kim, Hyoungshick, 2019. The light will be with you. Always – a novel continuous mobile authentication with the light sensor (poster). In: Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services. MobiSys '19, Association for Computing Machinery, New York, NY, USA, pp. 560–561.

Alawami, Mohsen Ali, Aiken, William, Kim, Hyoungshick, 2020. LightLock: User identification system using light intensity readings on smartphones. IEEE Sensors J. 20 (5), 2710–2721.

Alawami, Mohsen A., Kim, Hyoungshick, 2020. LocAuth: A fine-grained indoor location-based authentication system using wireless networks characteristics. Comput. Secur. 89, 101683.

Alawami, Mohsen Ali, Vinay, Aishwarya Ram, Kim, Hyoungshick, 2022. LocID: A secure and usable location-based smartphone unlocking scheme using wi-fi signals and light intensity. IEEE Internet Things J. 9 (23), 24357–24372.

Ali, M., ElBatt, T., Youssef, M., 2018. SenseIO: Realistic ubiquitous indoor outdoor detection system using smartphones. IEEE Sensors J. 18 (9), 3684–3693.

Bai, Yang, Liu, Jian, Lu, Li, Yang, Yilin, Chen, Yingying, Yu, Jiadi, 2020a. BatComm: enabling inaudible acoustic communication with high-throughput for mobile devices. In: Proceedings of the 18th Conference on Embedded Networked Sensor Systems. pp. 205–217.

Bai, Yang, Lu, Li, Cheng, Jerry, Liu, Jian, Chen, Yingying, Yu, Jiadi, 2020b. Acoustic-based sensing and applications: A survey. Comput. Netw. 181, 107447.

Bakar, Khairul Azmi Abu, Dahnil, Dahlila Putri, 2017. Design of location based authentication system using visible light communication. J. Theor. Appl. Inf. Technol. 95 (1), 147–154.

Blinowski, Grzegorz, 2019. Security of visible light communication systems—a survey. Phys. Commun. 34, 246–260.

Blog, ALTA, 2021. FBI Releases the Internet Crime Complaint Center 2020 Internet Crime Report. https://www.fbi.gov/news/press-releases/fbi-releases-the-internet-crime-complaint-center-2020-internet-crime-report-including-covid-19-scam-statistics, Online; (Accessed June 26, 2023).

Blog, AL.T.A., 2022. Cyber Losses Hit $6.9B in 2021. https://blog.alta.org/2022/03/cyber-losses-hit-69b-in-2021.html, Online; (Accessed June 26, 2023).

Cai, Chao, Zheng, Rong, Luo, Jun, 2022. Ubiquitous acoustic sensing on commodity iot devices: A survey. IEEE Commun. Surv. Tutor. 24 (1), 432–454.

Chen, Yuchi, Gong, Wei, Liu, Jiangchuan, Cui, Yong, 2018. I can hear more: Pushing the limit of ultrasound sensing on off-the-shelf mobile devices. In: IEEE INFOCOM 2018-IEEE Conference on Computer Communications. IEEE, pp. 2015–2023.

Chen, Zhenghua, Zou, Han, Yang, JianFei, Jiang, Hao, Xie, Lihua, 2019. WiFi Fingerprinting Indoor Localization Using Local Feature-Based Deep LSTM. IEEE Syst. J. PP, 1–10.

Ee, Sun Jun, Tien Ming, Jeshua Woon, Yap, Jia Suan, Lee, Scott Chuen Yuen, et al., 2020. Active and passive security attacks in wireless networks and prevention techniques. TechRxiv.

Eian, Isaac Chin, Lim, Ka Yong, Yeap, Majesty Xiao Li, Yeo, Hui Qi, Fatima, Z, 2020. Wireless networks: Active and passive attack vulnerabilities and privacy challenges. Preprints.

Eichelberger, Manuel, Tanner, Simon, Voirol, Gabriel, Wattenhofer, Roger, 2019a. Imperceptible audio communication. In: ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing. ICASSP, IEEE, pp. 680–684.

Eichelberger, Manuel, Tanner, Simon, Voirol, Gabriel, Wattenhofer, Roger, 2019b. Receiving data hidden in music. In: Proceedings of the 20th International Workshop on Mobile Computing Systems and Applications. pp. 33–38.

Fikriyadi, Fikriyadi, Ritzkal, Ritzkal, Prakosa, Bayu Adhi, 2020. Security analysis of Wireless Local Area network (WLAN) network with the penetration testing method. J. Mantik 4 (3, Nov), 1658–1662.

Kumar, Sanjeev, Singh, Preeti, 2019. A comprehensive survey of visible light communication: potential and challenges. Wirel. Pers. Commun. 109, 1357–1375.

Nadeem, Tamer, Uddin, Mostafa, 2022. Acoustic-WiFi: Acoustic support for Wi-Fi networks in smart devices. IEEE Trans. Commun. 70 (6), 3977–3994.

Nandakumar, Rajalakshmi, Chintalapudi, Krishna Kant, Padmanabhan, Venkat, Venkatesan, Ramarathnam, 2013. Dhwani: secure peer-to-peer acoustic NFC. ACM SIGCOMM Comput. Commun. Rev. 43 (4), 63–74.

Oyewobi, Stephen S., Djouani, Karim, Kurien, Anish Matthew, 2022. Visible light communications for internet of things: Prospects and approaches, challenges, solutions and future directions. Technologies 10 (1), 28.

Pathak, Parth H., Feng, Xiaotao, Hu, Pengfei, Mohapatra, Prasant, 2015. Visible light communication, networking, and sensing: A survey, potential and challenges. IEEE Commun. Surv. Tutor. 17 (4), 2047–2077.

Rehman, Saeed Ur, Ullah, Shakir, Chong, Peter Han Joo, Yongchareon, Sira, Komosny, Dan, 2019. Visible light communication: a system perspective—overview and challenges. Sensors 19 (5), 1153.

Suduwella, Chathura P., Ranasinghe, Yohani S., De Zoysa, Kasun, 2018. Visible light communication based authentication protocol designed for location based network connectivity. In: 17th International Conference on Advances in ICT for Emerging Regions, ICTer 2017 Proceedings, 2018-Janua, pp. 158–165.

Tan, Dingwei, Lu, Yuliang, Yan, Xuehu, Wang, Xiaoping, 2019. A simple review of audio steganography. In: 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference. ITNEC, IEEE, pp. 1409–1413.

Wang, Yu, Zhu, Xiaojun, Han, Hao, 2021. Chirpmu: Chirp based imperceptible information broadcasting with music. In: 2021 IEEE/ACM 29th International Symposium on Quality of Service. IWQOS, IEEE, pp. 1–9.

Zafari, Faheem, Gkelias, Athanasios, Leung, Kin K., 2019. A Survey of Indoor Localization Systems and Technologies. IEEE Commun. Surv. Tutor. 21 (3), 2568–2599, arXiv:1709.01015.

Zhang, Bingsheng, Zhan, Qin, Chen, Si, Li, Muyuan, Ren, Kui, Wang, Cong, Ma, Di, 2014. PriWhisper: Enabling keyless secure acoustic communication for smartphones. IEEE Internet Things J. 1 (1), 33–45.

Zhang, Li, Zhu, Xiaojun, Wu, Xiaobing, 2019. No more free riders: Sharing wifi secrets with acoustic signals. In: 2019 28th International Conference on Computer Communication and Networks. ICCCN, IEEE, pp. 1–8.

Zhao, Zhiwei, Chen, Feiyu, Pang, Yaoyao, Min, Geyong, Dong, Wei, 2017. Improving user identification for WiFi networks using visible light based authentication.

Zhou, Man, Wang, Qian, Ren, Kui, Koutsonikolas, Dimitrios, Su, Lu, Chen, Yanjiao, 2018. Dolphin: Real-time hidden acoustic signal capture with smartphones. IEEE Trans. Mob. Comput. 18 (3), 560–573.

Zhou, Man, Wang, Qian, Yang, Jingxiao, Li, Qi, Jiang, Peipei, Chen, Yanjiao, Wang, Zhibo, 2019. Stealing your android patterns via acoustic signals. IEEE Trans. Mob. Comput. 20 (4), 1656–1671.

Zhu, Hongzi, Zhang, Yuxiao, Liu, Zifan, Chang, Shan, Chen, Yingying, 2019. Hyperear: Indoor remote object finding with a single phone. In: 2019 IEEE 39th International Conference on Distributed Computing Systems. ICDCS, IEEE, pp. 678–687.