Practical identity recognition using WiFi's Channel State Information

Cristian Turetta*, Florenc Demrozi*, Philipp H. Kindt†, Alejandro Masrur†, and Graziano Pravadelli*
*Department of Computer Science, University of Verona, Italy, Email: name.surname@univr.it
†Department of Computer Science, TU Chemnitz, Germany, Email: name.surname@informatik.tu-chemnitz.de

***Abstract*— Identity recognition is increasingly used to control access to sensitive data, restricted areas in industrial, healthcare, and defense settings, as well as in consumer electronics. To this end, existing approaches are typically based on collecting and analyzing biometric data and imply severe privacy concerns. Particularly when cameras are involved, users might even reject or dismiss an identity recognition system. Furthermore, iris or fingerprint scanners, cameras, microphones, etc., imply installation and maintenance costs and require the user's active participation in the recognition procedure. This paper proposes a non-intrusive identity recognition system based on analyzing WiFi's Channel State Information (CSI). We show that CSI data attenuated by a person's body and typical movements allows for a reliable identification – even in a sitting posture. We further propose a lightweight deep learning algorithm trained using CSI data, which we implemented and evaluated on an embedded platform (i.e., a Raspberry Pi 4B). Our results obtained using real-world experiments suggest a high accuracy in recognizing people's identity, with a specificity of 98% and a sensitivity of 99%, while requiring a low training effort and negligible cost.**

***Index Terms*—Identity recognition, WiFi, Channel State Information, Deep learning, Convolutional Neural Networks**

I. Introduction

Identity recognition, viz., automatically identifying (human) users, has become essential in different industrial, healthcare, and defense settings, as well as in consumer devices. To this end, the available methods are based on collecting and analyzing biometric data, which ranges from physical traits like the iris, fingerprints, hand shape, etc., to more behavioral traits, such as gait, voice and hand writing [1]. Also cameras, e.g., for facial recognition [2], [3], are becoming more widespread.

However, all of these approaches have a number of shortcomings. First, dedicated sensors incur additional costs that make devices based on them more expensive. Second, they typically require close proximity and/or a physical contact, e.g., touching a fingerprint sensor. This increases the chances of transmitting infectious diseases [4], especially in public device installations. Finally, they require an active participation of the user, e.g., by staring into a camera without moving for a certain amount of time. Lately, a number of works have demonstrated the capabilities of radio signals, in particular, WiFi, to detect human activity [5]–[7] or even to perform identity recognition [8]–[10]. These methods have the potential to overcome the limitations of conventional approaches described above. For example, since WiFi is already ubiquitous, existing networks can be re-used and hence, no additional cost is induced. WiFi-based methods are intrinsically privacy-preserving, since no explicit biometric information (e.g., face images, fingerprints, etc.) need to be used/stored.

This research work has been partially supported by the project Dipartimenti di Eccellenza 2018-2022 funded by the Italian Ministry of Education, Universities and Research (MIUR).

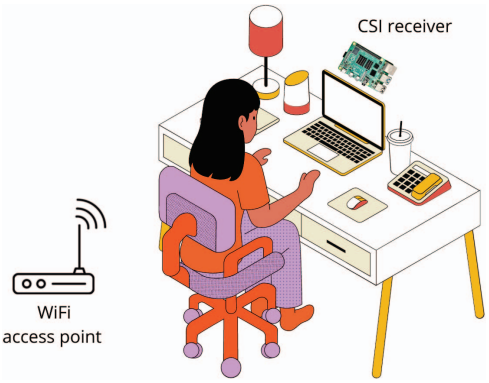

Fig. 1. Overview of our setup.

On the other hand, even though identifying persons using WiFi has been studied previously, the existing approaches are not yet practical. The vast majority, e.g., [8], [9], [11], [12], have considered human gait as the main feature for identity recognition. This requires a subject to walk, typically, on a pre-defined path. In addition, such approaches achieve relatively low accuracies, which makes them unsuitable in many scenarios of practical relevance. For example, *WiFi-ID* was shown to have an identification accuracy between 93% to 77% on a set of 2 to 6 subjects [11], whereas *WiWho* achieves an identification accuracy of 92% on a set of 20 subjects [12].

Contributions. This paper addresses the above shortcomings and proposes a practical, highly accurate system for WiFi-based identity recognition (Figure 1). To this end, we propose a specialized deep learning model, more specifically, a Convolutional Neural Network (CNN), which allows for a high identification accuracy while only requiring a low amount of time for being trained. The model is lightweight with a size of only 1.5 MB. This makes it suitable to be deployed on resource-constrained embedded devices, as we demonstrate by an implementation on a Raspberry Pi 4B.

Our experiments suggest a specificity (i.e., true negative rate) of 99% and a sensitivity (i.e, true positive rate) of 98%, thereby outperforming the previous works we are aware of. Our approach captures the CSI data emitted by only one standard WiFi AP, in contrast to many of the cited approaches

from the literature. Such an AP broadcasts approximately 10 *beacon frames* per second to advertise its presence, which we show to be sufficient for identity recognition. To the best of our knowledge, our approach significantly outperforms all existing ones and makes identity recognition practical for the first time. Particularly, our results suggest that users can be identified in a reliable manner and with minimum costs by using a commercial off-the-shelf WiFi SoC (System on a Chip), e.g., the one integrated by the Raspberry Pi 4B. Among others, we envision the following applications of our proposed system.

Applications. Identifying a sitting user is beneficial in a multitude of different applications. For example, our proposed technique can unlock a computer/notebook when a person sits in front of it. Likewise, a vehicle can automatically detect the driver and adjust the position of the seat or steering wheel accordingly. Furthermore, a user's identity can also be recognized in a standing position without requiring the person to walk along a predefined path. E.g., the identity of a user standing in front of an ATM/cash dispenser can be detected to prevent fraud with credit cards. Furthermore, persons arriving at their homes can be identified using WiFi, e.g., unlocking the door. Moreover, in industrial scenarios, a practical identification of employees/workers is necessary to allow them to access restricted areas or use equipment that requires special training.

Structure of the paper. We have organized this paper as follows: Section II presents the proposed identification methodology, including the architecture of the system and our setup. A description on the gathered data and experimental results are then discussed in Section III, while Section IV concludes the paper.

II. Identity recognition

With reference to Figure 2, the proposed approach consists of the following 3 steps:

- CSI data collection,
- CSI data preprocessing, and
- training of the pattern recognition model.

Finally, the trained model is used for identity recognition. Details are reported in the next sections.

A. CSI data collection

We collected CSI data using the following setup. A standard AP is positioned at a distance of around two-meters from the Raspberry Pi model 4B. The AP is placed in front of the sitting person, more specifically, behind the monitor/notebook, as depicted in Figure 1. Whereas the AP emits WiFi signals, the Raspberry Pi acts as a passive observer collecting the necessary CSI data. Here, the observer features a particular WiFi SoC, i.e., Broadcom BCM43455c0. However, WiFi SoCs usually do not make CSI data available to the host computer/board. This issue is resolved by patching the BCM43455c0 firmware using the Nexmon project [13], [14] to "*unlock*" CSI perception. The extracted CSI data consists of an amplitude and phase for each WiFi subcarrier and frame.

The changes in phase and amplitude of each subcarrier are quantified by the aforementioned CSI, which we denote by H. When a WiFi sender emits a certain signal X, the signal Y seen by the receiver is given by

$$Y = X \cdot H + N. \tag{1}$$

Since the signal propagation is frequency dependent, X, Y, H and N are vectors of length n, with N being related to noise and n being the number of subcarriers. Further, the term $X \cdot N$ in Equation 1 represents an element-wise multiplication. Clearly, H accounts for the difference between the received and the transmitted signal. Each value in H is a complex number, which can be translated into a signal attenuation and a phase shift. Our approach makes use of multiple successively received WiFi frames. Hence, we obtain a sequence of CSI vectors $H_1, H_2, H_3, ..., H_x$, where x stands for the number of frames received.

B. CSI preprocessing

Besides noise removal, there are a number of preprocessing steps that need to be conducted.

- **Amplitude and phase separation:** Since the CSI phase is prone to hardware distortions, we only use the CSI amplitude. We hence split every complex value of H into an amplitude and phase. To perform identity recognition, we only use the amplitude information.
- **Pilot carrier nulling:** Some values of H are related to pilot subcarriers that do not contain any valid CSI data. We hence remove them to avoid that they negatively impact the detection accuracy.
- **AGC compensation:** The receiving WiFi SoC performs AGC to bring the received signal power into a desired range for decoding. This also affects the absolute value of every CSI sample in H. We adopt the method proposed in [15] to cancel the effect of AGC using RSSI. Before this, we additionally denoise the RSSI signal.

C. Training the pattern recognition model

This section is on creating a pattern recognition model based on the collected CSI data, which consists of the following steps.

1) Segmentation: We use a constant sampling frequency f and consider a certain time window w. This way, a segment has a size of $m \times n$, where $m = f \cdot w$ and n corresponds to the number of subcarriers, as described before. In this paper, we consider a WiFi network in the 2.4 GHz band and a channel of 20 MHz bandwidth, which implies 64 subcarriers . Standard APs, as the one used in this paper, emit around 10 beacon frames per second, which corresponds to a sampling frequency of $f = 10Hz$. We consider a time window of $w = 1s$. Due to clock offsets, the number of beacons captured in this time window varies typically between 9 and 11 beacons. To avoid such variations, we only consider the first 9 beacon frames captured in a window and discard those that arrive later.

As a result, in a time window of $w = 1s$, we always have segments of size 9×64 CSI samples. Clearly, different window sizes result in segments of different size. Moreover,

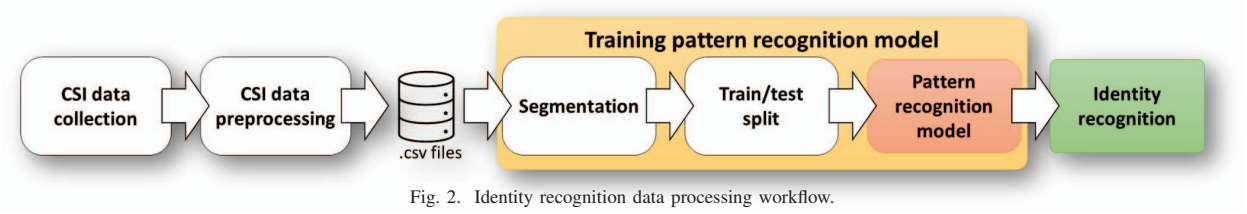

Fig. 2. Identity recognition data processing workflow.

each segment has a label called *Subject Id*, which describes the person that is actually present.

2) Training/test data partitioning: Once multiple CSI data segments have been collected, we need to define training and test datasets for our pattern recognition model. In particular, we use 75% of the obtained CSI data for training, while the remaining 25% are reserved for testing purposes.

D. Pattern recognition model

Our CNN model takes the CSI data segments as its input. The input data goes through two convolutional layers operating on the vectors $v_s \in \mathbb{R}^a$, i.e., a portion of a CSI data segment or sub-segment, and $v_f \in \mathbb{R}^b$, i.e., a convolutional filter, to return a vector $r \in \mathbb{R}^{a-b+1}$. Each element r_i in this vector r is given by $r_i = v_f^T \times v_s[i : i+b-1]$. This means that r_i is computed as a scalar product between the convolutional filter v_f and the sub-segment v_s. This process empowers our CNN model to extract features from CSI data. Between the two convolutional layers, and after the second one, there are two *max pooling* layers. These have the task of reducing the dimensions of the extracted features, by acting *de facto* as a feature selector. Basically, the max pooling layer divides the output from the preceding layer into slices of size σ, which are known as pools, selecting the maximum value in each pool to reduce dimensions. In our implementation, we use a pool size of $\sigma = 2$. After passing the last max pooling layer, the data, which is in matrix form, goes through a flatten layer that reshape the data into a one-dimensional vector. The output of the previous layer pass through the two dense fully connected layers, these are responsible for the classification and have 256 and 128 neurons, respectively. The classification result is returned by a soft-max layer, which estimates the probability of a given person being identified correctly.

III. Experimental results

This section presents and discusses the results of our experiments. In addition, we present a comparison of the proposed CNN with other state-of-the-art machine learning techniques.

A. Our setup

As depicted in Figure 1, our experimental setup consists of the following components:

- a Raspberry Pi model 4B featuring a Broadcom BCM43455c0 WiFi SoC. It is placed in front of the user, but behind a monitor/notebook.
- a Fritz Box 7590 as WiFi AP positioned behind the user at around $2m$ from the observer.

Our proposed CNN model needs to be trained for a specific arrangement or environment.

The Raspberry Pi in our setup receives all beacon frames emitted by the AP and extracts the CSI data, in particular, the CSI amplitude. Figure 3 presents an example of collected CSI data. The x-axis shows the time line, the y-axis represents subcarriers, and the z-axis (i.e., color intensity) represents the CSI amplitude.

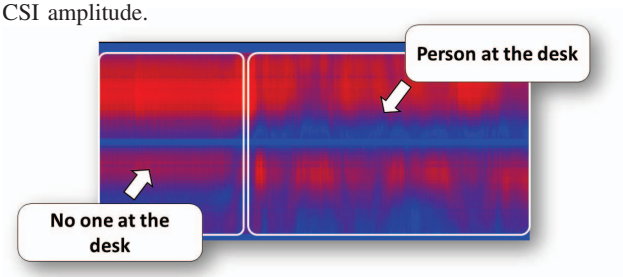

Fig. 3. Real-time visualization of CSI amplitude.

B. Collected dataset

Nine different subjects (i.e., four females and five males) were recruited for data collection. Each subject was asked to remain seated for 15 consecutive minutes, as shown in Figure 1. During this time, they could perform any action (e.g., working on the PC, reading an article, using the smartphone, talking, etc.), except standing up and leaving the workplace. In addition to collecting data related to the different subjects, we also recorded data for the case of an empty desk (i.e., no one sitting on the chair for 15 minutes).

The weights and heights of the subjects range from 49 to 107 kg and 168 to 195 cm, with a standard deviation of 16.6 kg and 10.5 cm. This data was emitted by only one AP and amounts to a total of 600 MB.

C. Identification performance

The CNN model described in Section II-D has been trained and tested using the collected dataset by following the training procedure described in Section II-C. Table I presents the obtained results in terms of sensitivity, specificity, precision and F1-score. Our CNN model is compared with four state-of-the-art machine learning (ML) models (i.e., k-nearest Neighbor (k-NN), Decision Tree (DT), Random Forest (RF) and Linear Discriminant Analysis (LDA)) using the BHAR framework [16]. To this end, we consider four different time windows (i.e., 1, 2, 3, and 5 seconds).

From Table I, we highlight that the ML models present a slight degradation of the results as the time window size increases. This is related to the fact that the subject's movements play a more significant role in a larger time window

than in shorter time windows. Unlike our CNN model that adapts its extracted features to the size of the time windows, thus, performing better, the features by the other ML models do not adapt well. This is because they are tied to specific data characteristics, such as dominant frequency, the sum of the CSI amplitudes, the standard deviation of the CSI amplitudes, etc., which do not adapt well to larger windows.

Model	Sensitivity	Specificity	Precision	F1-score
CNN	98-96-97-98	99-99-99-99	98-96-97-98	98-96-97-98
k-NN	94-93-93-91	99-99-99-99	94-93-93-91	94-93-93-91
DT	85-82-80-80	98-98-97-97	85-82-80-81	86-82-80-81
RF	94-92-92-93	99-99-99-99	94-92-93-93	94-92-93-93
LDA	91-90-85-80	99-98-98-97	91-90-86-80	91-90-86-80

Results for time windows of 1, 2, 3 and 5 seconds

TABLE I
AVERAGE RESULTS OF THE PROPOSED CNN AND ML MODELS.

In summary, these results suggest that our CNN model was able to correctly recognize the subject at the desk in approximately 98.2% of the cases.

Figure 4 presents the boxplot of the CSI amplitude grouped by the subject Id to investigate the obtained results further. As can be seen, there is a clear difference between each subject, essentially due to her/his physique and sitting behavior (i.e., posture, typical movements, etc.). This evidences that the studied context and the implemented system provide representative data and that the proposed network can classify it correctly.

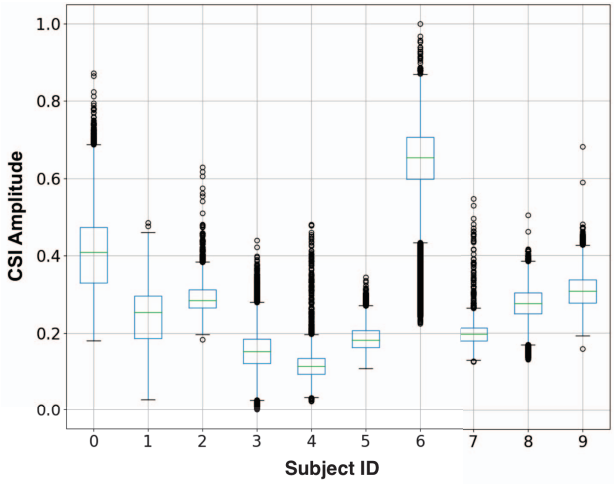

Fig. 4. Boxplot view of the normalized CSI amplitude (y-axis) grouped by test subjects (x-axis) e.g., Subject ID = 6 represents the empty environment scenario.

IV. CONCLUSION AND FUTURE WORK

In this paper, we proposed a practical identity recognition method based on WiFi's CSI data. In contrast to known approaches from the literature, our method allows for an identification accuracy of more than 98% and can be deployed at a very low cost. In particular, we only require a standard WiFi AP that broadcasts beacon frames (e.g., a typical frequency of around 10 Hz already suffices). We focused on a scenario involving sitting persons, which is highly relevant in an office context or within a vehicle. However, the same technique can also be used to identify an individual's upper body in a standing position, e.g., at an ATM or in an industrial context to operate sensitive equipment, etc. Our experiments evidence that collecting only 15 minutes of training data per subject/person suffices for a reliable identity recognition.

As future work, we plan to evaluate and further extend our approach towards various other setups. In the future, we will study the usage of multiple APs for identity recognition, whether and how different WiFi packet rates affect the classification results, and evaluate different frequency bands. Moreover, since CSI-based methodologies suffer from the high variability of the measured CSI data, which is related to changes in the environment (e.g., a change in the furniture's position, temperature changes, humidity, etc.), we will concentrate on exploring methodologies to mitigate these issues.

REFERENCES

[1] N. Duta, "A survey of biometric technology based on hand shape," *Pattern recognition*, vol. 42, no. 11, pp. 2797–2806, 2009.

[2] P. Zhou, X. Han, V. I. Morariu, and L. S. Davis, "Two-stream neural networks for tampered face detection," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2017.

[3] P. Vadakkepat, P. Lim, L. C. De Silva, L. Jing, and L. L. Ling, "Multimodal approach to human-face detection and tracking," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 3, pp. 1385–1393, 2008.

[4] A. K. Pitol and T. R. Julian, "Community transmission of SARS-CoV-2 by surfaces: risks and risk reduction strategies," *Environmental Science & Technology Letters*, vol. 8, no. 3, pp. 263–269, 2021.

[5] F. Demrozi, C. Turetta, F. Chiarani, P. H. Kindt, and G. Pravadelli, "Estimating indoor occupancy through low-cost BLE devices," *IEEE Sensors Journal*, 2021.

[6] F. Demrozi, F. Chiarani, and G. Pravadelli, "A low-cost BLE-based distance estimation, occupancy detection and counting system," in *2021 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2021.

[7] L. Sun, S. Sen, D. Koutsonikolas, and K.-H. Kim, "Widraw: Enabling hands-free drawing in the air on commodity WiFi devices," in *Annual International Conference on Mobile Computing and Networking (Mobicom)*, 2015.

[8] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using WiFi signals," in *International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*, 2016.

[9] T. Xin, B. Guo, Z. Wang, M. Li, Z. Yu, and X. Zhou, "Freesense: Indoor human identification with wi-fi signals," in *IEEE Global Communications Conference (GLOBECOM)*, 2016.

[10] C. Shi, J. Liu, H. Liu, and Y. Chen, "Smart user authentication through actuation of daily activities leveraging WiFi-enabled IoT," in *International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 2017.

[11] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "Wifi-id: Human identification using wifi signal," in *International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2016.

[12] Y. Zeng, P. H. Pathak, and P. Mohapatra, "WiWho: WiFi-based person identification in smart spaces," in *International Conference on Information Processing in Sensor Networks (IPSN)*, 2016.

[13] M. Schulz, D. Wegemer, and M. Hollick. (2017) Nexmon: The C-based firmware patching framework. [Online]. Available: https://nexmon.org

[14] F. Gringoli, M. Schulz, J. Link, and M. Hollick, "Free your CSI: A channel state information extraction platform for modern Wi-Fi chipsets," in *International Workshop on Wireless Network Testbeds, Experimental Evaluation& Characterization (WiNTECH)*, 2019.

[15] Z. Gao, Y. Gao, S. Wang, D. Li, and Y. Xu, "CRISLoc: Reconstructable CSI fingerprinting for indoor smartphone localization," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3422–3437, 2020.

[16] F. Demrozi, C. Turetta, and G. Pravadelli, "B-har: an open-source baseline framework for in depth study of human activity recognition datasets and workflows," *arXiv preprint arXiv:2101.10870*, 2021.