The Published Journal Article is available at https://doi.org/10.1016/j.comcom.2022.03.011.

On the Properties of Device-Free Multi-Point CSI Localization and Its Obfuscation

Marco Cominelli, Francesco Gringoli, Renato Lo Cigno

Department of Information Engineering (DII) — University of Brescia, Italy

Abstract

The use of Channel State Information (CSI) as a means of sensing the environment through Wi-Fi communications, and in particular to locate the position of unaware people, was proven feasible several years ago and now it is moving from feasibility studies to high precision applications, thus posing a serious threat to people's privacy in workplaces, at home, and maybe even outdoors. The work we present in this paper explores how the use of multiple localization receivers can enhance the precision and robustness of device-free CSI-based localization with a method based on a state-of-the-art Convolutional Neural Network. Furthermore, we explore the effect of the inter-antenna distance on localization, both with multiple receivers and with a single MIMO receiver. Next we discuss how a randomized pre-filtering at the transmitter can hide the information that the CSI carries on the location of one person indoor. We formalize the pre-filtering as a per-frame, per-subcarrier amplitude multiplication based on a Markovian stochastic process, and we discuss different signal clipping and smoothing methods highlighting the existence of a trade-off between communication performance and obfuscation efficiency. The methodology can in any case guarantee almost unhampered communications with very good localization obfuscation. Results are presented discussing two different ways of exploiting the multi-receiver or multi-antenna redundancy and how, in any case, properly randomized pre-distortion at the transmitter can prevent localization even if the attack is carried out with multiple localization devices (receivers controlled by the attacker) and not only with a multi-antenna (MIMO) receiver.

Keywords: CSI-based Wi-Fi Localization, Smart Spaces, Privacy Protection, Location Obfuscation, Markovian Modeling

1. Introduction and Motivation

Sensing as a side-service of Wi-Fi is becoming an industrial reality [1]. Channel State Information (CSI)-based localization in particular is attracting attention for device-free indoor positioning. This research field was opened about ten years ago by seminal works like [2, 3, 4, 5], then the academic community indulged in many variations of the topic, hinting at the possibility of identifying activities, poses, or gestures [6, 7, 8, 9], seeking health-care applications [10, 11] and many others. The most recent trend is exploiting Machine Learning (ML) and Artificial Intelligence (AI) to compensate for the difficulty of finding analytic models with the power of supervised learning techniques for classification purposes, often involving Deep Learning or Reinforcement Learning [12, 13, 14, 15, 16]. The recent survey [17] provides an excellent introduction to this topic for the interested reader.

Two topics received little attention instead:

1. If and how multi-point reception can improve the reliability of CSI-based localization; and
2. How much does localization impact the privacy of people.

Although the scarce attention to the first one may look surprising, we could find only a couple of works on the topic. The authors of [18] propose to use massive Multiple-Input Multiple-Output (MIMO) technologies to improve the quality of CSI-based localization with a neural network (NN). The work exploits up to 64 antennas but with a single logical measurement point; indeed, the work focuses on the learning technology and assumes that there is a service dedicated to localization, i.e., special frames are transmitted dedicated only to localization. Thus, such work should be compared mainly with technologies dedicated to localization (as those based on time-of-flight like [19]) rather than sensing as a side-effect of Wi-Fi communications. The research in [9, 20] has some similarities with our contribution, though those works focus on the localization of a device and not, as we do, on the localization of a person who does not carry any device. The work in [20] builds on the concept of channel charting [21], which lends to the possibility of semi-autonomous training because it uses differential positions and differential CSI, hence assuming a slowly changing channel with a CSI sampling that respects the Nyquist theorem, a condition that, for instance, cannot be assumed if Wi-Fi traffic is sporadic, or to detect the presence of a person in a room, a condition that implies a sort of discontinuity in the CSI.

The contributions of this paper, which extends the paper we presented at MedComNet 2021 [22], in the context of device-free localization and its obfuscation are two:

- First, we present the first experimental study that shows how, using multiple localization devices or multiple antennas connected to a single MIMO-like device, the precision of localization can be improved;
- Second, we show that also in these conditions, the privacy of users can be preserved with a refinement of the CSI randomization technique we first introduced in [23, 24], not

only protecting users from localization attacks but also preserving the communication performance. We also analyze the influence of different configurations of the randomization function at the transmitter.

Compared to the conference version [22], this paper extends the analysis to MIMO devices, explores the impact of different post-randomization filtering techniques, discusses the impact of antenna positioning, and investigates the training space of the Convolutional Neural Networks (CNNs) to improve localization efficiency. Besides, the literature review is extended, and the description of the entire system is improved and more detailed.

2. Related Work

In this section, we discuss the status of works and technologies dedicated to preventing Wi-Fi sensing and localization in particular. Special attention is given to fuzzing and obfuscation techniques as means to prevent attacks on users' privacy—which are directly related to our work—rather than as an attack to legitimate localization and sensing. The two perspectives are obviously related, but they differ in the attention posed in communication performance. For instance, an attacker using a Wi-Fi sensing system may not care about communication at all; on the other hand, preserving the users' privacy cannot hamper their legitimate right to communicate. We also address works tackling security as well as localization itself when they have influenced our work or may have an impact on it. Localization *in general* is instead not addressed here, as we are not proposing novel localization techniques, but rather means to prevent their illegitimate use.

Localization and sensing research seems to have completely disregarded privacy implications, which have been addressed only very recently by our group and others [22, 23, 24, 25, 26, 27, 28, 29, 30].

The first idea that comes to mind is using a reactive jamming device that selectively kills frames that belong to the localization attack, adapting for instance techniques like [25, 26]. To the best of our knowledge, this idea has never been explored in the literature, maybe because it kills traffic, thus if the frames used for localization also carry user data the communications will be heavily hampered. Additionally, such a technique requires knowing that an attack is underway and the ability to identify the frames used for localization, otherwise it would become simply a jamming denial-of-service. A system to counter Wi-Fi sensing was originally proposed in [27] to prevent gesture recognition. Similar to [29], this system is based on an independent device that relays frames with the goal of superimposing an additional "reflection" of the signal that *obfuscates* the information imprinted by the environments on frames, differently from the more common *jamming* that superimposes a different signal, sometimes just noise, with the goal of killing the frame reception.

The works that lay the foundations of this paper are [23, 24, 29]. In the first two, we proposed for the first time a technique to obfuscate, or hide, the information carried by the CSI that enables localization, focusing on passive attacks, i.e., those where the attacker controls only the localization devices. The core idea is to randomly distort the transmitted frames so that the CSI at the receiver looks like the one of a signal that has propagated through a different environment, unrelated to the one in which sensing is performed. In the third one, we tackled the problem of countering active attacks, i.e., those where the attacker controls both the transmitter and the localization device. In this case the countermeasures cannot be based on the pre-distortion of the transmitted frames, because transmitted frames are controlled by the attacker. The solution we proposed is based on a fast relay node, ideally an intelligent reflective surface like those discussed in [31], which introduces a time-varying additional reflection in the electromagnetic environment that prevents a localization device from pinpointing the position of a person.

Strictly related to the work we present is the demo in [30], which contributes to the community an open-source CSI fuzzer implemented in openwifi.[1] The fuzzer is based on a fixed-length, three-taps Finite Input Response (FIR) filter with random coefficients, that are limited to pre-defined values for the sake of easy implementation in the FPGA. The outcome is clearly similar to having two additional multipath reflections that are controlled by the FIR coefficients. The implementation is tested for good communication performance, while its efficiency in preserving privacy has not been tested.

Finally, exploiting techniques similar to those used in [24], the work in [28] proposes to manipulate the CSI with the goal of avoiding device radiometric fingerprinting and preventing impersonation attacks. The topic of the paper is not localization, but if a person holds a Wi-Fi device a double attack identifying the device and the location of the person is more than a possibility.

Moving to recently proposed localization techniques, [32] presents an ML technique where sophisticated pre-processing is applied to the CSI data before it is fed to the CNN, in particular the authors use wavelet transforms and principal component analysis to extract the features that the Deep Learning network analyzes and learns. The authors of [33] take instead a different approach, whereby the localizing device uses two directional antennas, one pointing directly to the transmitter and the other observing the area of interest. The localization is based on the differential analysis of the two signals, with declared similitude with radar analysis. The extremely recent work presented in [34] has many points of contact with ours. First of all the authors explore, like we do, the impact of multi-point sensing on localization; however, the goal of the paper is to improve localization precision. To achieve this goal the authors propose a data fusion technique where the CNN localization pipelines at each receiver do not output a position classification, but a probability map of the target position. All the probability maps are then fused together to obtain a more reliable position estimation. We

[1] openwifi (`https://github.com/open-sdr/openwifi`) is an open-source implementation of the 802.11 Medium Access Control (MAC) and PHY layers based on software-defined radios (SDRs) and widely used for research and innovation in Wi-Fi. In the GÈANT-funded project that partially supports this work we are implementing in openwifi the techniques proposed here and in previous papers.

note that this data fusion technique is different from both the methods we propose in Sect. 5 to exploit the additional information provided by multiple localization devices (or antennas). It would be interesting to compare also this methodology; however, the paper pre-print was published when our work was nearly concluded, and at the time of writing there is no open-source implementation of such technique available. Also, the CNN used in each pipeline is not available as an open-source project.

As a final remark of the work related to our contribution, we highlight that testing the obfuscation technique we propose against several localization techniques is of the utmost interest and importance, and we fully support whoever wishes to challenge our open-source implementation of the CSI obfuscator with other CSI-based localization techniques.

We think that localization techniques and countermeasures should be tested with real experiments, even if they are costly and time-consuming, because the complexity of the problem prevents easy shortcuts based on "standard data" or simulations. All the same, we plan to publish the dataset we are collecting to provide a useful tool for early-stage research, such as the one proposed in [35]. We could unfortunately not use this database for comparison, because it contains only clean CSI traces to test gesture recognition, and it is impossible for us to generate obfuscated data to compare with.

3. Attack Model and Scenario

This section presents the reference attack model that can be used to monitor the activities and the position of unaware people through the opportunistic reuse of the Wi-Fi signals pervading modern environments. Next, the setup we realized in our laboratory is described, while the details of single experiments are described in Sect. 7.

3.1. Attack Model

The attack model is shown in Fig. 1. We assume that the attacker (e.g., an employer whose goal is circumventing legislation on employees monitoring) can control multiple devices with the ability to extract CSI data from the received Wi-Fi frames. The large availability of extremely cheap and small platforms that can be converted into sensing nodes, like the Raspberry Pi [36], makes this feasible and cheap even on a large scale. A detailed analysis of the CSI structure at each single localization device, as well as a comparison between the CSI collected at different devices, enables the attacker to determine the precise position of a person in the room. The attack considered in this paper extends the techniques described in [24]. The attacker collects a set of CSI traces with the help of a collaborator standing in specific target positions; then, he trains a CNN with the collected data to use it later to determine the position of an unaware victim, e.g., an employee or a guest.

3.2. Experimental Facility

We carry out the experiments in a laboratory of the ANS[2] group at the University of Brescia, whose map is shown in Fig. 1.

[2]The Advanced Networking Systems (ANS) group is a research group in telecommunications at the University of Brescia, Italy.

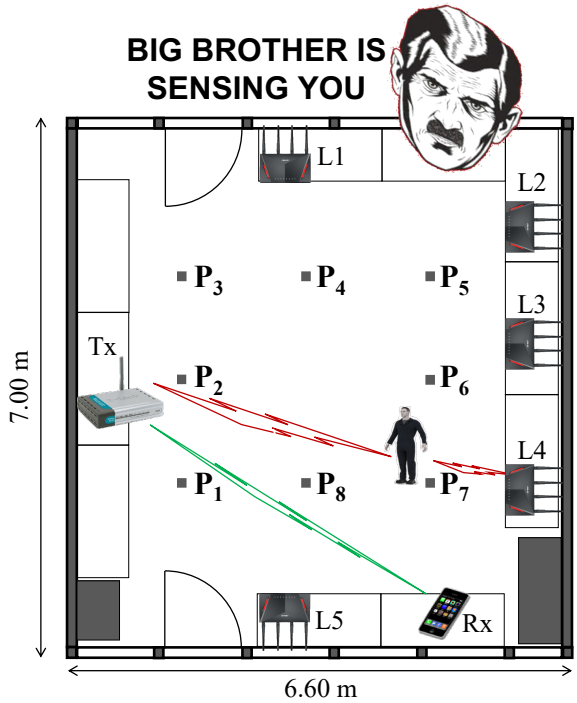


Figure 1: In the considered scenario the attacker can collect CSI data of the frames that Tx AP sends to the Rx smartphone simultaneously at five localization devices L1–L5 to locate a victim standing in one of 8 possible target locations (P1–P8).

Five localization devices (from L1 to L5) are positioned on desks aligned along three of the four sides of the lab, while the transmitter (Tx) lies on a desk on the fourth side and the legitimate receiver (Rx) is placed somewhere else in the room to measure the communication performance (its actual position is irrelevant for now). The attacker controls all the localization devices and knows their positions, while the transmitter, that can be one of the Access Points of a corporate network, is not under the control of the attacker, and its only requirement is to be in a fixed position. The goal of the attacker is to determine the correct location of the person in the room among 8 possible target positions (P1,. . . , P8). The configuration in Fig. 1 ensures that the person being tracked is always obstructing the line-of-sight (LoS) between the transmitter Tx and at least one of the localization devices (L1,. . . ,L5). In this way, the collected CSI should always be significantly affected at one or more devices, independently of the victim's position.

4. CSI-based Localization and its Obfuscation

The principle behind the localization technique is the interaction between the Wi-Fi signals and the human body. In fact, the presence in an environment of a human body that absorbs, scatters, and reflects electromagnetic waves induces peculiar variations in the spectrum of the received signal that depends on the body position and movements. These variations can be studied by analyzing the CSI evaluated by every Wi-Fi device upon receiving a frame. As we show in Fig. 2, the correct de-

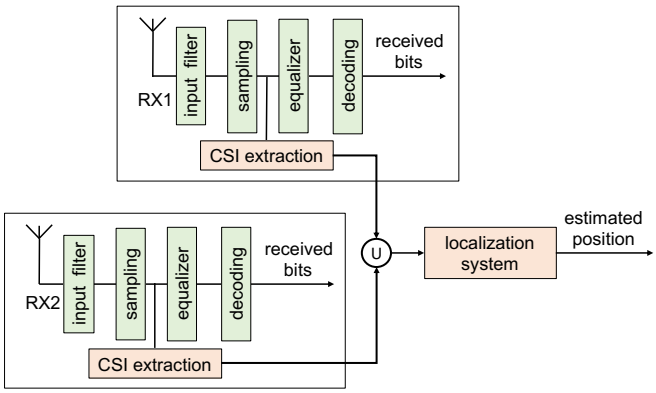


Figure 2: 802.11 modified receiver to infer people location: each receiver collects CSI and pushes everything to the localization system. Multiple CSI data originating from the same frame are jointly analyzed to improve the accuracy with respect to a single device configuration.

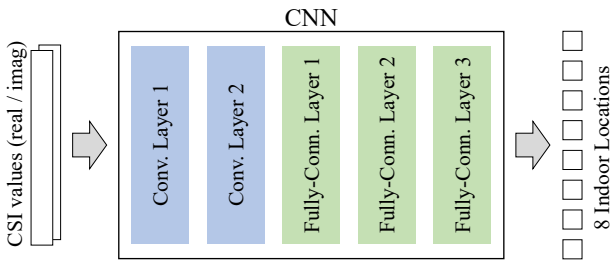


Figure 3: Architecture of the CNN used by our localization system.

coding of the frame's data requires equalization of the spectrum of the received signal to reduce the distortion introduced by the communication channel. Extracting the CSI data from the chipset internals, one can directly observe the modifications that depend on the position of the observed person, given that the rest of the environment, including the position of the transmitter and the localization device, remains relatively stable over time.

The localization frameworks that have received more attention are based on NNs trained with CSI obtained when someone is standing in a known position; then, during the attack, the victim's position is estimated by recognizing the same patterns in the CSI [37]. It is important to capture a large amount of data not only to speed up the training phase, but also to obtain multiple CSI snapshots corresponding to the same position of the victim, as this would allow to "average out" minor spectral variations. This is in general not a problem in modern workplaces, where Wi-Fi signals pervade the environment and an attacker can opportunistically use the signals transmitted by the Access Points (APs) of a corporate network, given that such Wi-Fi nodes are usually in well-known positions and generate the largest amount of traffic.

In this work, we build the multi-device localization system on top of what we introduced in [24], where we have developed an efficient implementation for a single device based on a CNN with good localization accuracy[3]. A high-level representation of the NN architecture is shown in Fig. 3. We consider 802.11ac frames transmitted on 80-MHz channels with a single spatial stream encoded using Orthogonal Frequency Division Multiplexing (OFDM); each CSI data point is an array of 256 complex values, one per OFDM subcarrier. After removing the unused subcarriers (eleven at the edges of the spectrum and three at the center) and splitting the real and imaginary parts of each value, we get as input for the CNN a 242×2 matrix. The first two convolutional layers of the CNN shown in Fig. 3 are used to extract complex features from the input data by exploiting the similarity of adjacent frequencies. In cascade to the convolutional layers, there are three fully-connected layers. The output of the last layer corresponds to a choice among one of the possible classes, i.e., positions that are decided during the training phase. All the layers but the last one (which uses a softmax function) use a common Rectified Linear Unit (ReLU) activation function. Finally, we use the Adaptive Momentum Estimation (ADAM) algorithm to adjust the weights of the CNN during the training phase.

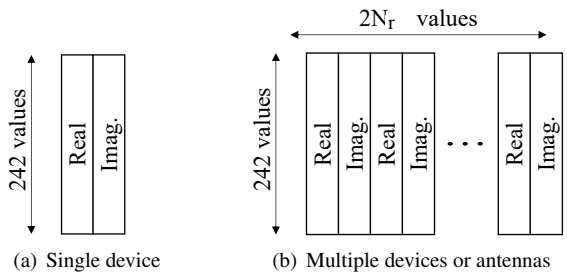
The considered CNN achieves good accuracy and this is clearly related to the unique and remarkably constant CSI data that are obtained for each target position of the person under tracking in the room. Simple reasoning suggests that a random pre-distortion of the transmitted signals should suffice to disrupt localization accuracy and obfuscate the person's position. In the feasibility study presented in [23, 24] we obtained excellent obfuscation results by selectively amplifying some OFDM subcarriers in the spectrum of transmitted frames, following a simple time correlation structure. As a consequence of this coarse manipulation, the presence of the obfuscator could have been detected (and eventually countered) with a careful analysis of the signal. Furthermore, the communication performance when the obfuscator was active was severely degraded, leading to highly-reduced throughput.

5. Improving Localization with Multiple or MIMO Devices

The literature on CSI-based localization is now starting to consider how combining CSI data collected at multiple points can improve the accuracy of wireless sensing systems. With enough devices it is possible, in fact, to always have at least one of them whose LoS to the transmitter intercepts the person under tracking. For instance, if the target person stands in position P_7 in Fig. 1, we expect minor interference effects on L1, but clear effects on L4 due to the obstructed LoS. Indeed, in principle, also the separate analysis of CSI data from different antennas of a MIMO system should yield a more accurate estimate, and hence a larger threat to users' privacy, even if the antennas are close to one another. In this work, we explore this possibility too. One of the goals of this work is to extend the localization system by combining the CSI captured at multiple devices or antennas as shown in Fig. 2; the second one is showing that also this powerful attack can be countered.

We present two methods for "combining" the localization data: we discuss here their pros and cons while we present the

[3]Further details on this line of research in our group, the software produced an so forth can be found at `https://ans.unibs.it/projects/csi-murder/` and `https://ans.unibs.it/projects/di-p2sl/`.



(a) Single device (b) Multiple devices or antennas

Figure 4: The network is trained with different types of input depending on the considered scenario; when we apply the CSI Data Fusion technique all the collected CSI are fed to a single NN.

experimental results in Sect. 8. Common to the two methods is that localization devices are positioned in the same indoor environment and they have the capability of matching the reception of the same packets (we use timestamps, but other techniques are just as fine).

5.1. Majority Vote

In the *Majority Vote* implementation, we combine the output of multiple stand-alone localization systems, each one associated to its own CNN. Given one CSI, each device, or *voter*, will produce an estimate for the victim's target position; the most-voted position will be the one estimated by the whole localization system. It is important to notice that a stand-alone localization system can be implemented for every antenna of the attacker's receivers. Thus, voters can either be single-antenna devices (one for each receiver in a different location L1,...,L5, see Fig. 1) or all the antennas belonging to a single MIMO receiver.

When the receivers in the room are located several wavelengths apart from each other, we can assume that the CSI vector at the input of each localization system, which we report in Fig. 4(a) as a 242×2 matrix, is independent from the others. Thus the CNNs actually learn and classify independent models, so that errors are independent, and we can assume that "summing" the results compensates for the random errors. Unfortunately, positions cannot be "summed" in an algebraic way, but we can decide based on the majority of decisions. Given that position errors are independent by construction, a majority vote also corresponds to a Maximum Likelihood Estimation (MLE) and should be optimal given the assumptions. The assumption of independence falls short when considering antennas in MIMO receivers, where antennas are spaced by roughly one wavelength only; however, even MIMO technologies are based on the assumption of an independence approximation. What is surely not satisfied is the LoS obstruction property discussed above, and the analysis of these differences is very interesting to understand the knowledge base carried by CSI.

Let N_r be the number of localization devices. Independently from N_r, it is always possible that the vote does not have a majority, e.g., 3 devices have classified 3 different locations, and no decision can be taken. We separate wrong classifications (a decision is taken, but it is for a wrong position), from undecidable situations when a majority is not reached.

5.2. CSI Data Fusion

More sophisticated use of the information at different devices is based on the fusion of the CSI vectors, assuming that an extended CNN can do a better job than simple majority voting. Note that majority voting is a MLE, but only under the assumptions of independent decisions. The data fusion changes the knowledge base of the estimator, thus we can hope in a more powerful technique.

In this second implementation we consider an extended CNN that processes the CSI vectors collected at the N_r devices or antennas as a "fused" and larger dataset, i.e., a $242 \times 2\,N_r$ matrix as we show in Fig. 4(b). The N_r CSI vector are always combined column-wise and "stacked" on different rows, so that the rows of the "fused" matrix always refer to the same subcarrier, and every pair of columns represents the real and imaginary parts of the CSI extracted by one device. Combining the data into a $242\,N_r \times 2$ matrix would put the highest subcarrier of one antenna/device near the lowest subcarrier of the next antenna/device, creating artificial "features" that may confound the learning process. The CNN can thus learn from a larger knowledge base, where features of the CSI related to the person position can be derived from the different antennas, possibly improving localization accuracy. Differently from the *Majority Vote* implementation, this method always outputs one single position estimation, as in the case of a single device, and does not have undecidable situations. We have not attempted to design a new architecture for the neural network for this task, so, whatever the results we obtain, we cannot exclude that a different learning method, based on CNNs or other network types, can obtain better results. We recall that the goal of this paper is not proposing a new localization technique, but exploring as much as possible the impact of our obfuscation techniques against multi-point localization attacks.

We highlight that the term 'fusion' here has a different meaning compared to its use in [34]. The technique proposed there merges probability maps derived from NNs, each one taking as input the CSI of a single antenna and not all the CSI as we do. In this sense, the methodology is a blend of merging the data at the analysis level as we propose here, and the majority vote we propose in Sect. 5.1, the method is very interesting and its impact on obfuscation techniques can be the subject of future research when it will be officially released by the authors.

6. Principles of CSI Randomization and Its Implementation

While any approach that can improve the localization accuracy represents a positive result, at the same time it can be considered an increasing hazard against the privacy of the tracked people. In addition, an improved localization technique can also have detrimental effects against simple obfuscation techniques like the one that we introduced in [24]. The goal of this section is hence to design an improved obfuscation technique that can be effective independently of the number of involved devices. Here we study how to defeat the localization mechanisms introduced in Sect. 5 to restore the privacy of the tracked person. A good obfuscation technique is as unobtrusive as possible, effective in

preventing localization, and also maintains good communication performance.

As discussed in [24], the frequency-dependent amplitude of the received signal is the feature that is mostly considered by the CNN in classifying CSI and mapping the person location. For this reason, we focus the randomization technique on the amplitude of the subcarriers that compose the spectrum of the transmitted signals. We consider a single transmission chain for the sake of simplicity: the extension to multiple transmission chains is left for future work. On the other hand we consider not only multiple devices with a single receiver chain, but also MIMO devices with 4 antennas and receiver chains.

Let N_{sc} be the number of carriers used by the Wi-Fi OFDM modulation; $f_i,\ i = 1, 2, \ldots, N_{sc}$ is the carrier number; $k = 1, 2, \ldots$ is the discrete-time index identifying the frame; and $\Delta_t(k)$ is the absolute (continuous) time between frame k, and frame $k-1$; $\Delta_t(1)$ is undefined, but it is irrelevant for our purposes. We only consider carriers that are not suppressed by the system, thus excluding the middle carrier and those in the guard bands. The magnitude of the spectrum at the receiver, or CSI, derived from the known initial symbols of the frame, represents the frequency and time-varying signal attenuation (or channel response) introduced by the channel $A_C(f_i, k)$. Notice that the entire Wi-Fi PHY layer is based on the assumption that the channel coherence is long enough to guarantee that the channel response is constant during a single frame, thus our modeling does not introduce significant approximations.

The goal of the obfuscator is to guarantee that the information in $A_R(f_i, k)$ does not allow an attacker to properly classify the position of a person in the room. Ideally, the obfuscator should guarantee that the mutual information between the CSI and the location of a person in the room is zero, but this theoretical analysis is out of the scope of this paper.

To achieve this goal we multiply the IQ samples of the digital signal before performing the Inverse Fast Fourier Transform (IFFT) conversion on the OFDM symbol, so that the actual CSI information at the receiver is:

$$A_R(f_i, k) = A_C(f_i, k) \times A_O(f_i, k) \tag{1}$$

where $\times$ is the standard algebraic multiplication applied separately carrier by carrier: $A_O(f_i, k)$ is a pre-distortion mask whose goal is adding random information to $A_C(f_i, k)$ so that $A_R(f_i, k)$ maintains the properties that allow demodulation and correct decoding of the frame given the CSI, but information on the real channel response is degraded to a point where localization is not better than a random guess.

The pre-distortion $A_O(f_i, k)$ must have the following characteristics:

1. It does not alter the frame power:

$$\sum_{i=1}^{N_{sc}} A_O(f_i, k) = K_A \tag{2}$$

   meaning that if some frequencies are amplified, then others must be attenuated, with K_A some appropriate constant.
2. It guarantees that the correlation in time is compatible with the standard movement of a person in a room;
3. It guarantees that the attenuation in frequency is compatible with the channel Doppler spread;
4. It cannot be inverted within a reasonable time, i.e., given the sequence $A_R(f_i, k); k = h, \ldots h + H$ it should not be possible to reconstruct the sequence $A_O(f_i, k)$, not even when using multiple devices; H is a design parameter whose impact on the system is left for future study;
5. It does not modify the communication performance of the system.

The formalization and (if possible) the proof that the five conditions above are feasible are beyond the scope of this work. In the following we present a heuristic Markovian methodology that we evaluate in Sect. 8.

Let $\mathbf{R}$ be a vector of independent, continuous random variables ρ_i of dimension N_{sc}, one for every OFDM subcarrier. Each random variable is drawn from the same distribution $f_R(\rho)$, and they are all independent one another. For reasons that will be clear shortly, we select $f_R(\rho)$ to be a uniform distribution with support $(\rho_{\min}, \rho_{\max})$. Consider now the multidimensional random process defined as:

$$\mathcal{R}(k) = e^{-\alpha\Delta_t(k)}\mathcal{R}(k-1) + \mathbf{R} \tag{3}$$

Eq. (3) defines a Uniform-Markov process, which, compared with the more popular Gauss-Markov process[4] exhibits uniform increments taken from $f_R(\rho)$ instead of Gaussian increments. The choice of uniform increments derives from the need of maximizing the causality of the choice while maintaining finite (and small) increments. As we discuss at the end of this section the random multiplier must be in any case clipped to guarantee that the amplitude is positive and does not saturate the transmitter amplifier, thus the choice of an infinite support (e.g, Gaussian) random variable as the process increment would be incongruous (or irrelevant).

The dimension of the process is N_{sc}. The process is discrete time, because the transmission of frames defines a discrete time index; however, the process memory depends on the absolute time $\Delta_t(k)$ in order to guarantee that frames transmitted far apart in time do not have excessive correlation. As it is well known from probability theory, this process exhibits a dependence in time that increases when α decreases, which allows tuning the obfuscation to the expected movements of people in the room.

The random process defined by Eq. (3) has no correlation in the frequency domain, which is in contrast with the desired characteristic 3) defined above. To overcome this unwanted feature, we use a simple convolutional filter (or weighted moving average) as follows:

$$A_O(k) = [1 + \mathcal{R}(k)] * \Theta_C \tag{4}$$

where $*$ is the standard convolutional sum and Θ_C is a symmetric filter with length C; C must be odd for symmetry and $3 \leq C \leq$

[4] For instance, the error of Global Navigation Satellite System positioning is normally modeled with a three-dimensional Gauss-Markov process.

N_{sc}. The dependency on f_i is implicit in the convolution, and appropriate leading and trailing zeros must be appended to $\mathcal{R}(k)$ to allow the convolution. The shape and characteristics of Θ_C can be studied to optimize the performance. In this work we do a simple moving average (all coefficients are 1) with $C = 3, 5, 7$, thus exploring the impact of the filter length.

Equations (3) and (4), together with the normalization (2) (to be run at every step) define, in our opinion, an appropriate randomized localization obfuscation that respects the characteristics illustrated above. $\rho_{\min}$, $\rho_{\max}$, α, C, and the values of Θ_C taps can be used for tuning the system, and can also be changed over time to make it harder for an attacker to invert the obfuscation. A proper sensitivity analysis on all these parameters is not feasible within a single paper, and we are more interested in fundamental properties than in finding the optimal setting, which may also depend on the considered scenario. A quick preliminary study was sufficient to select $\rho_{\min} = -0.3$, $\rho_{\max} = 0.3$, and $\alpha = 0.2$ for achieving acceptable performance. Notice that $\alpha = 0.2$ means that if $\Delta_t(k) \geq 15$ s, then $A_O(k)$ and $A_O(k\text{-}1)$ are almost completely uncorrelated (the correlation coefficient is below 5%, which is coherent with requirement 2) above.

Eqs. (3) and (4) cannot guarantee that the pre-distortion does not lead to amplitudes smaller than zero, which is obviously not implementable. Also zero or very small values are not desirable, as they imply that an entire subcarrier is suppressed, and this would introduce systematic transmission errors that should be avoided. To prevent this possibility we use a clipping function $[\cdot]_{\min}^{\max}$, that cuts the amplitude of each subcarrier between a min and a max value. Theoretically, only the min clip is necessary to avoid negative and too small values, but this asymmetry makes it difficult to guarantee that the average power of frames is not altered. Moreover, a very large amplification of a single carrier (recall that the amplification process at each carrier is independent from the others) makes all the others very small, because the analog part of the transmitter will saturate the largest subcarrier to the amplifier maximum power. This means that very large differences in subcarrier pre-distortion may lead to worst communication performance. We have shown in [22] that not clipping the maximum amplification does not result in any improvement in obfuscation, while instead penalizes more the communication performance. Hence, we do not report here results for max = ∞, concentrating instead on more interesting aspects of the experiments.

As an additional alternative, the clipping can be applied before or after the convolutional filter of Eq. (4). Predicting the consequences of clipping before or after the filter is very difficult, because clipping is highly non linear, thus we decided to explore the performance of both options implementing the equivalent of the following two equations:

$$A'_O(k) = [1 + \mathcal{R}(k)]_{\min}^{\max} * \Theta_C \tag{5}$$

$$A''_O(k) = [1 + \mathcal{R}(k) * \Theta_C]_{\min}^{\max} \tag{6}$$

As for the clipping values, we consider here only the symmetric case with min = 0.1 and max = 1.9. These value are arbitrary, but stem from the heuristic consideration that min significantly smaller than 0.1 would lead to systematic errors on the specific subcarrier, while larger values can only lead to less efficient randomization; max = 1.9 is a consequence of symmetry.

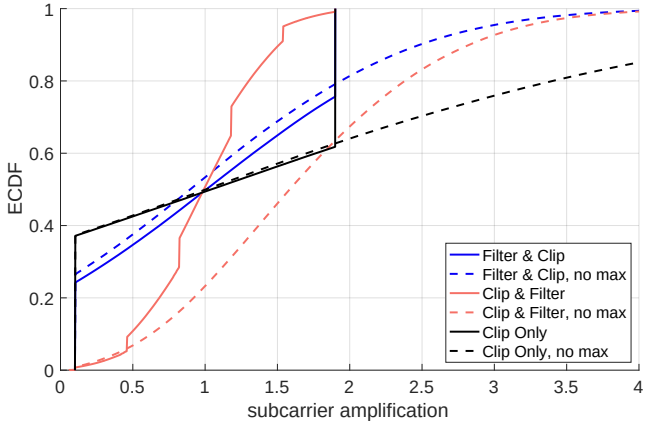


Figure 5: Experimental Cumulative Density Function (ECDF) of the amplification factors assigned to the OFDM subcarriers in Eqs. (5) and (6) with both symmetric and asymmetric clipping. Blue lines refer to Eq. (6) and red ones to Eq. (5) with a 5-tap filter (C=5); black lines are without filtering; 'no max' means that max = ∞.

6.1. Implementation

We describe here some implementation details, focusing on the constraints that may approximate the formulas introduced above. First of all, we note that $N_{sc} = 256$ as we work with 80 MHz 802.11ac frames. Second, we implemented the localization devices using Commercial Off-The-Shelf (COTS) Access Points from Asus: we chose the RT-AC86U model as it can extract CSI data from the transmitted frames. Third, we implemented the transmitter with an Ettus USRP N300 SDR whose bandwidth exceeds the 80 MHz requirement. We chose an SDR because we need precise control over the generation of each Wi-Fi frame in order to craft and apply the pre-distortion that modifies the CSI. We generate Wi-Fi frames with the MATLAB WLAN Toolbox running on a workstation equipped with an Intel Core i7 and 16 GB of RAM. We also prepare obfuscated frames directly in MATLAB before sending the corresponding sequence of IQ samples to the SDR. As the SDR does not run a MAC algorithm, there could be some uncontrolled collisions on the channel, even if we have selected a channel (157) that is not used in our University.

The MATLAB code implements the obfuscation techniques described by Eqs. (5) and (6) by applying the pre-distortion, if present, before the IFFT—that is, in the frequency domain. We avoid working in the time domain because it would require the usage of a circular convolution. Before sending the stream of IQ samples to the SDR, we normalize them to the highest value, i.e., we divide them by the one with the largest absolute module, so that we use the entire range of the SDR.

We highlight that the overall procedure cannot strictly guarantee that the frame power is not altered as required by Eq. (2): this entails evaluating the power spectrum of the entire frame, which we are unable to do; however, we deem that the pre-distortion, especially with symmetric clipping, does not change the power spectrum on the channel significantly.

Fig. 5 shows the ECDF of the marginal distribution of the amplitude of the processes described by Eq. (5) (Clip & Filter, in red) and Eq. (6) (Filter & Clip, in blue) and without filter at all (Clip Only, in black). The ECDF is computed over 10,000 frames, or equivalently, $256 \times 10,000 = 2,560,000$ samples. The solid lines refer to symmetric clipping: their median value is one, as expected. As the OFDM modulation is not constant envelope and its peak to average ratio depends on the payload, and can be as large as 12 dB, a pre-distortion that preserves the average amplitude should not significantly alter the power spectrum after the modulation, but this aspect requires further investigation. The dashed lines (no max) refer instead to the case when clipping is applied only to the min amplitude, and the median is obviously not one, but we do not know if this alters significantly the frame power spectrum. Filter and Clip (blue) and Clip Only (black) display a similar behavior with the same qualitative behavior, albeit without filtering the no max curve (black dashed) has much larger values; this is also reflected in the larger discrete part of the distribution when clipping. The behavior is expected as filtering concentrates the distribution around the average. The counter-intuitive behavior of the solid red curve (Clip & Filter) deserves a final comment. Indeed, by first clipping and then filtering, we introduce multiple discrete components (and not just two as in Filter & Clip) in the distribution: they are given by the combinations of clipped subcarriers in the 5-tap average filter, and are reflected in the ECDF. Changing the filter length will also change the position and weight of the multiple discrete components.

Fig. 6 reports qualitative results measured at one localization device. It visually shows why Wi-Fi sensing can locate people, and why the proposed obfuscation technique is a valid countermeasure. Without obfuscation, the two top plots, the channel response (amplitude of the CSI) is remarkably constant over time (x-axis) given a position, and the amplitude change with the position (left and right plots). The CNN can learn and classify the position of the person. Obfuscation, instead, keeps changing the CSI amplitude, thus preventing proper learning and classification as shown by the other three pairs of plots in the figure. It is clear that the filter position (and length, not shown in the figure for the sake of brevity) and the clipping strategy alters the channel response in different ways, but from these qualitative plots it is not possible to state what combination is more effective in obfuscating location, or which one will preserve or destroy communication performance. What can be seen from these plots is that the energy of the frames seems smaller when clipping is the last operation (blue dominates), while filtering as final operation preserves better the requirement to preserve frames power as there is a better distribution of blue and yellow. We would like to recall that the software we have used and the data we have collected are available on our website `https://ans.unibs.it`.

7. Experimental Setup

The experiments are all run in the laboratory quickly described in Sect. 3.2 and Fig. 1. Experiments have been run on several different days between spring and autumn 2021 to verify that the performance is not influenced by random effects, and it is reasonably time invariant.

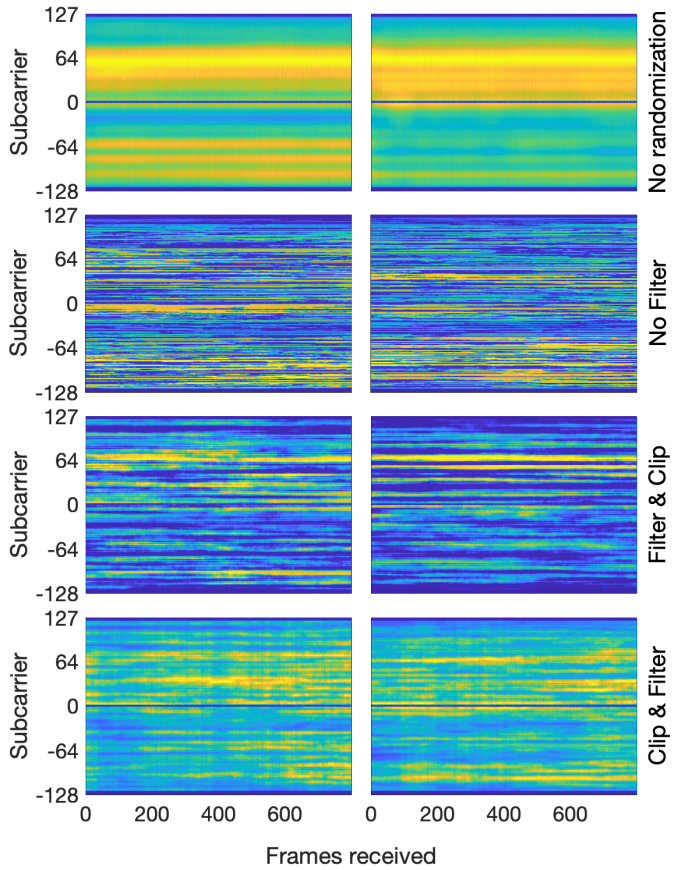


Figure 6: Magnitude of the CSI collected from 800 frames with a person standing in positions P1 and P2 (brighter yellow-ish colors mean larger magnitude). The two plots at the top (first row) refer to clean transmissions: the horizontal bands show that CSI are constant over time if the person does not move, enabling localization by ML. The other six plots refer to the same positions when different obfuscation techniques are in place: the second row corresponds to Clip Only, the third row to Filter & Clip, and the last one to Clip & Filter, both with C=5.

Each experiment consists of two different and separate phases.

Phase 1: Training. During this phase a person stands in one of the 8 positions P1 . . . P8, reported in Fig. 1 and well known to the localization system. The transmitter (the AP, or the SDR in our case) keeps sending a continuous flow of packets. For each position, each receiver controlled by the attacker decodes and collects 1200 Wi-Fi frames, extracting the corresponding CSI. Based on this dataset, all the CNNs that are part of the localization system (one per antenna/device) are trained ten times with ten different initialization vectors. Training is performed on $8 \cdot 1000$ CSI samples, while the remaining $8 \cdot 200$ samples are used to validate the model and to stop the training as soon as the accuracy on the validation set converges.

Phase 2: Attack. In ML terminology this is the testing phase of the CNN. After several minutes (sometimes hours) from

Clean	No obfuscation technique is implemented
NF	No Filter, only the clipping is applied
CF3	Clip & Filter with 3 taps
CF5	Clip & Filter with 5 taps
CF7	Clip & Filter with 7 taps
FC3	Filter & Clip with 3 taps
FC5	Filter & Clip with 5 taps
FC7	Filter & Clip with 7 taps

Table 1: Acronyms of the obfuscation techniques experimented.

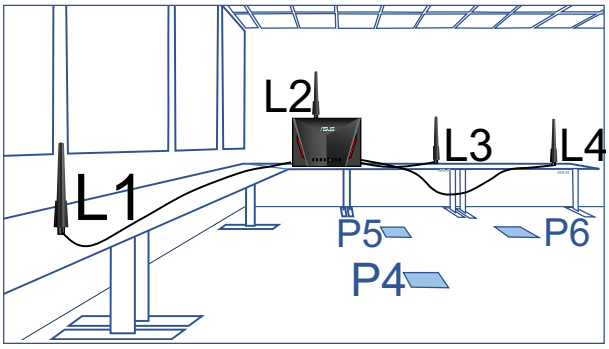


Figure 7: Octopus configuration derived from a photo: antennas of the MIMO localization system are placed far from the L2 device, approximately in the positions occupied by L1, L3, and L4.

Phase 1, a person—not necessarily the same that did the training—stands in the same 8 positions, and another 1000 packets are collected by every device. The localization techniques discussed in 5 are applied to all the 10 CNN configurations that are the outcome of the 10 training with different initialization. All results are kept allowing to compute both the average performance and the standard deviation. Performance results distribution are clustered enough that the standard deviation is a good indication of dispersion, since there are no outliers.

Training is always performed separately for each of the 7 obfuscation designs discussed in Sect. 4 plus the one without obfuscation summarized with the acronym used in commenting results in Tab. 1; clipping is always symmetric with min = 0.1, max = 1.9. Obviously, independent training is done also for each of the 4 experimental configurations discussed below: base, mimo, Octopus, Frankenstein.

base: This is the basic configuration represented in Fig. 1. One transmitter (Tx) with a single antenna and transmission chain, sends frames to a generic receiver within the same room (Rx), whose position is completely irrelevant for the localization, and that can indeed be changing. The receiver is not involved in localization and decodes the frames (used to estimate the Packet Delivery Rate (PDR)) with its standard hardware characteristics. Each localization device (Ln) extracts the CSI from a single, external antenna and uses it to estimate the position of the person in the room as described in Sect. 4 and 5. The cooperation between localization devices is necessarily off-line moving the information collected to a computer that operates the multi-point analysis.

Figure 8: Picture of the Frankenstein configuration: 4 devices are placed very close to one another in position L2, only 1 antenna per device is used for localization.

mimo: In this configuration the localization devices, in the same positions as in the base configuration work exploiting all the 4 antennas of the MIMO device. The multi-point localization is indeed local to each localization device, and the 5 different devices and positions are used to gain confidence in the results themselves.

Octopus: This setup is meant to understand together with Frankenstein, what is the impact of antenna positions in the multi-point analysis when a MIMO device is used. Its name, well explained by Fig. 7, came to us observing the long 'tentacles' going to the antennas: i.e., the cables that allows placing the external antennas of the MIMO device in the positions L1, L3, and L4, while the localization device itself (and the last antenna) is in position L2.

Frankenstein: Opposite to Octopus, this configuration puts 1 antenna of 4 different devices in a single location (L2) as shown in Fig. 8. The name refers to an (intelligent?) body built with different and spare parts: A sort of MIMO system built with different devices aimed at improving the localisation performance. The goal is to understand if the localization performance is affected primarily by antenna location or, conversely, by the use of different devices once the antennas are separated by at least a wavelength, making the channel transfer functions from the transmitter roughly orthogonal.

8. Results and Discussion

Both localization and its obfuscation depend in non-trivial ways on many parameters. This section is split in three parts, each one devoted to one or more configurations we have tested

	L1	L2	L3	L4	L5
Clean	78 (6)	89 (4)	93 (3)	78 (6)	70 (5)
NF	13 (2)	23 (3)	15 (2)	34 (3)	25 (3)
CF3	10 (2)	20 (3)	11 (2)	17 (2)	21 (5)
CF5	16 (1)	30 (3)	17 (2)	28 (2)	24 (4)
CF7	17 (3)	33 (5)	14 (3)	29 (4)	26 (2)
FC3	12 (2)	25 (5)	16 (3)	20 (2)	18 (5)
FC5	14 (2)	27 (5)	14 (2)	19 (3)	22 (2)
FC7	15 (2)	29 (4)	14 (3)	26 (3)	22 (3)

Table 2: Accuracy and standard deviation for the localization at each single location (L1–L5) with all the obfuscation techniques; both accuracy and (standard deviation) are in percentage [%].

in our experiments. For each configuration we give an interpretation of the results obtained, laying the foundations for better CSI-localization theory as well as how to prevent its unauthorized use.

First, Sect. 8.1 analyzes the performance of localization and its obfuscation in the BASE configuration, to gain insight on the advantages of multi-point localization and the performance of the CSI randomization at the transmitter as a means of obfuscation. Second, Sect. 8.2 discusses the impact of CSI randomization on communication performance, allowing us to analyze and select the randomization scheme that offers the best trade off between position obfuscation and data throughput. Next, Sect. 8.3 tackles the problem of extending the techniques to a single, MIMO device rather than multiple single-antenna devices. Finally, Sect. 8.4 explores the role of antenna disposition and device diversification, both from the attacker perspective, i.e., what is the most efficient setup to infer people's position, and from the privacy protection point of view, i.e., how efficient obfuscation remains in face of the different attack configurations.

8.1. BASE: *Multi-Devices Experiments*

During every experiment, we use $5 \times 8 \times 1000 = 40,000$ frames to train the system ($8,000$ more are used for validation), and other $40,000$ frames to test it. The data collected during an experiment are stored as raw CSI, allowing post-processing and reducing the number of experiments to be performed. This allows applying the two different multi-point localization techniques (see Sect. 5.1 and 5.2) exactly on the same measured data, thus avoiding that the difference observed in the two techniques is related to differences in the experimental environment and not an intrinsic property. Since we consider 8 possible positions, a random guess would lead to an average accuracy of 12.5%, and this is our reference for evaluating the quality of the obfuscation.

Tab. 2 reports the accuracy of the localization attack performed *separately* in each location L1–L5. For each experiment, we train the model ten times with ten random initialization vectors, and we report the average percentage of correctly classified positions as well as the standard deviations. Fig. 9 reports instead the average localization accuracy over all the five locations (multi-point attack) for all the obfuscation designs, together with the standard deviations (vertical bars at the top of the bars)

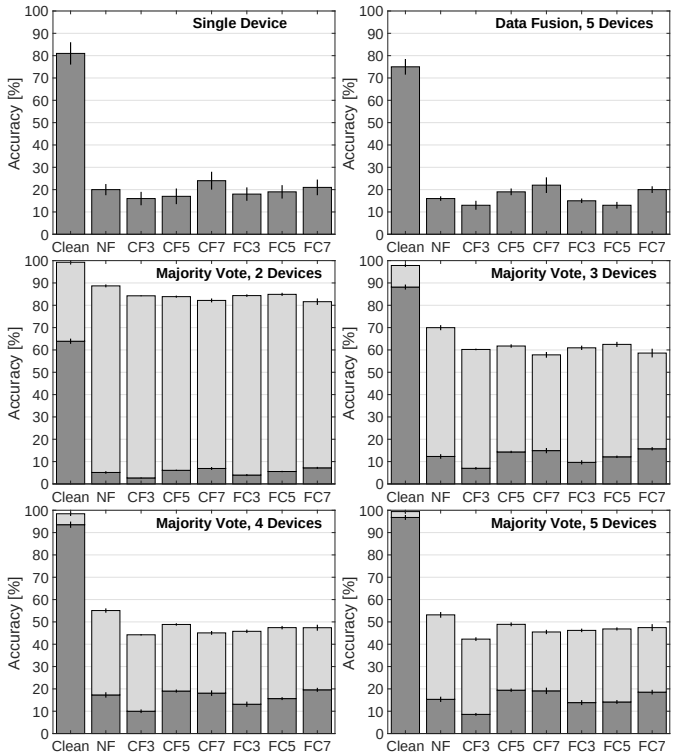


Figure 9: Localization precision for all the multi-point localization analyzed, light gray areas in majority vote results refer to undecidable situations; vertical bars represent the standard deviation of the measures.

measured again from the accuracy of each trained model at all locations.

A careful analysis of the Clean data in Tab. 2 highlights, as we already observed in [29], that the position of the localization device has a strong influence on the sensing performance independently from the position of the person: L5 performs much worse than the others, while L3 significantly better. We also notice that the localization accuracy decreases drastically with every obfuscation technique, regardless of the position of the localization device. The actual performance seems to depend more on the location of the device than on the obfuscation technique, but it never exceeds 33% (L2, CF7), which is an accuracy hardly usable for any meaningful attack. In general the accuracy remains above the 12.5% of a random guess, indicating that there is still some information embedded in the CSI. We believe that further research can lead both to better localization techniques and to better obfuscation ones.

Coming to the results achieved with the multi-device attack, the first thing that emerges is that a simple and traditional Majority Voting outperforms the apparently more sophisticated Data Fusion approach, which does not seem to improve the performance even in the absence of obfuscation as it clearly emerges from the Clean bars in all the plots of Fig. 9. As we already commented in Sect. 5, this does not exclude that some other data fusion technique may outperform Majority Voting, but just that improving CNN-based localization simply adding information may be more difficult than expected. A possible explanation lies in the fact that the Data Fusion approach does not make the

localization errors of different devices independent one another, thus the strong correlation in all the data induced by the CNN makes some errors dominant.

The six plots of Fig. 9 convey a lot of information and summarize the 8 experiments, for a total of more than 600,000 frames collected and analyzed. Each of the bars reports the average of all the experiments with the relative standard deviation. For instance, the top left plot, which summarizes the results for single devices, reports the overall average of the measures reported in Tab. 2. Clearly, multi-point results are obtained from merging the measurements from the five locations L1–L5.

The four plots referring to Majority Vote in Fig. 9 report the localization accuracy in dark grey and the undecidable situation (i.e., a majority decision has not been reached, see Sect. 5.1) in light grey, with the relative standard deviations. The plots for $N_r = 2$, 3, and 4 devices are obtained with all the possible combinations out of the 5 locations, clearly for 5 devices only a single combination is possible.

As expected, undecidable situations are particularly high when we consider only two devices, while increasing the number of receivers increases the performance and decreases the number of undecidable cases, until localization without obfuscation becomes nearly perfect for $N_r = 5$ devices, but we can consider the localization almost deterministic also with 4 devices, where the accuracy is already above 90%. Notice that given the amount of data we collected, these results are extremely reliable and significant, making Wi-Fi-based localization a true threat for location privacy. For $N_r \leq 5$ the results reported indicate that multi-point localization also makes the technique less sensitive to the device position.

Considering the obfuscation results, it is immediately clear that the technique we propose is robust also against a sophisticated attack brought with five devices strategically placed in the surveyed room. Independently from the specific manipulation all obfuscation techniques are efficient and make localization useless for any practical purpose. It is clear that there exists some residual information on the position embedded in the CSI at the receivers, as the accuracy is always above the 12.5% of a random guess (apart from the 2 devices Majority Vote, where most of the situations are undecidable); however, this information seems difficult to exploit as neither the Majority Vote nor the Data Fusion technique seems able to extract this information to make a guess that reliably goes above 20%, which is barely more than the random guess.

It is interesting that when the obfuscation is active, for Majority Vote, there is a high percentage of wrong decisions, which were instead completely absent without obfuscation. This is a strong indication that the randomized pre-distortion implemented by the obfuscator successfully deceives the localizer, so that it does not learn anything really meaningful during the training phase. Since the training phase is done with the obfuscator active, the classifier (recall that the CNN learning is supervised, so training should be effective in any case if there is information to exploit) indeed learns random patterns and it is not able to single-out the channel characteristics, resulting in classification errors and not only undecidable situations.

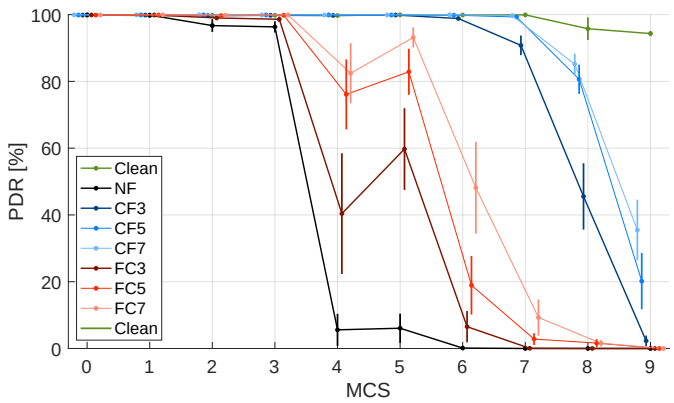


Figure 10: PDR as a function of the MCS with different types of randomization applied to the transmitted signal. Each point represent the average result, while the vertical bars identify the 90% confidence interval. Points are slightly offset from the integers they refer to for the sake of readability.

8.2. Communication Performance

Sect. 8.1 has clarified that a multi-point localization attack can be really effective, but also that proper position obfuscation is feasible, and many possible obfuscation configurations are possible and roughly equivalent from the localization performance point of view. What we have not yet discussed is the impact on the communication performance, i.e., if the CSI manipulation affects the throughput (or equivalently the PDR) at the legitimate receiver (Rx in 1).

In Wi-Fi, the achievable throughput depends on the chosen Modulation and Coding Scheme (MCS), while localization is solely based on the CSI, which is independent from the MCS. The 802.11ac standard defines ten MCS over 80 MHz and 800 ns guard period with corresponding throughput increasing from 29.3 Mbit/s (MCS 0) to 390 Mbit/s (MCS 9). A higher MCS enables higher throughput, but it is more sensitive to distortion, noise and interference because of more advanced modulation techniques and less robust correction codes.

Fig. 10 shows the impact of different randomization techniques on the PDR, i.e., the percentage of Wi-Fi frames correctly received. We transmit 1000 frames for every possible MCS and for every randomization technique we are considering, and we count how many frames we correctly decode at each receiver. First of all we highlight that without obfuscation ('clean' results) the performance is in line with throughput documented in the literature, with 100% PDR but for MCS= $8, 9$, where some frames are nearly always lost due to the high modulation index (256-QAM), and reduced Forward Error Correction (FEC) codes capabilities. On the other hand, pure randomization of the frames subcarriers ('nofilter' results), penalizes performance as soon as the modulation index of each subcarrier is not constant-envelope (MCS= 3 and larger use QAM modulations), with the exception of MCS= 3 where the robustness of the rate $\frac{1}{2}$ FEC compensates even for systematic errors.

The other results show the influence of different smoothing filter lengths and relative positions of clipping and filtering on the communication performance. There is a profound difference

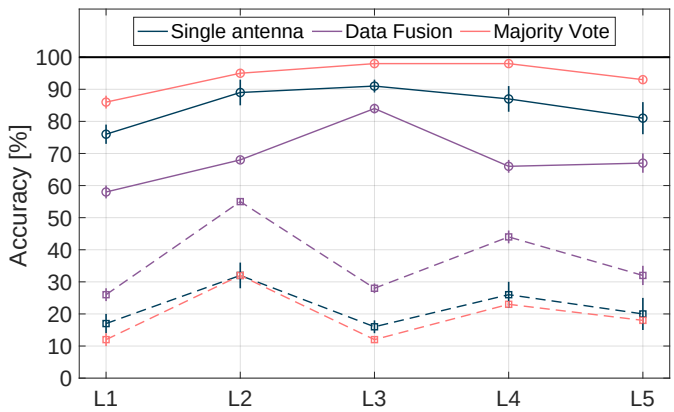


Figure 11: Percentage of correctly classified positions in the MIMO case for different configurations: results for the Clean case are reported with solid lines and circles; results for the CF5 case are identified by dashed lines and squares.

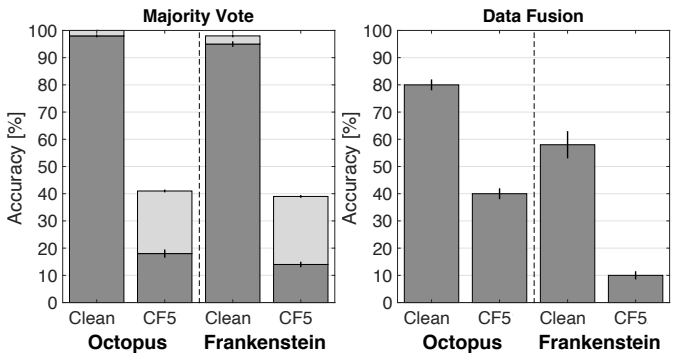


Figure 12: Localization accuracy of the OCTOPUS and FRANKENSTEIN systems in different conditions and with different techniques.

in the performance achieved by Clip & Filter versus Filter & Clip. It is evident that filtering the randomization mask before applying clipping has a destructive effect on the communication performance, and this is particularly evident for MCS ≥ 3, when QAM modulations are used, which are more sensitive to amplitude distortion.

While it is difficult to provide a theoretical justification for this evidence, we can try to explain this by noticing that, before clipping, the values assumed by the random process may be well-above or below the clipping threshold. The clipping operation forces the mask into the admissible range of values, but also introduces some high-frequency components in the signal spectrum, which may interfere with other subcarriers, introducing systematic errors. In the transition phase between good and bad performance (MCS 3, 4, and 5), the behavior is particularly chaotic due to the interplay between the systematic errors and the convolutional codes that change the rate with different MCSs. On the other hand, filtering the mask when the clipping has been already applied has the effect of producing a "smoother" mask with less high-frequency components, helping the receiver's equalizer to compensate the distortions. The filter length, on the other hand, does not have a major impact on the PDR, although longer filters improve the PDR slightly. It is clear that we are still unable to perform perfect obfuscation without hampering communications at all, but CF5 and CF7 provide a good trade-off with similar performance both in obfuscation and in communications

For the reasons discussed above, in the next two subsections we present results only for the Clip & Filter methodology with 5 taps C&F5. We highlight that experiments have been run for all the configurations and this choice is done for the sake of briefness and clarity, as the other results confirm what we present and discuss without adding significant insight.

8.3. MIMO: *Localizing with MIMO Devices*

The multi-point attack considered in Sect. 8.1 is powerful, but costly and eventually difficult to mount. A legitimate question concerns the possibility to use a single position with MIMO technology and whether obfuscation is effective also in this case. To this end, we collected the CSI from all 4 antennas of the localization devices in positions L1–L5, and we trained the localization algorithm with this data, with the same approach of the BASE scenario.

Fig. 11 reports the localization accuracy and standard deviation for the 5 location for the Clean and CF5 cases with single antennas (the average of the 4 *separate* decisions), data fusion, and Majority Vote (4 antennas, there are a few cases of undecidable situations, but they are marginal and not reported for the sake of clarity). Considering the Clean case, once again the Majority Vote yields very good results, with a slight degradation for L1 and L5, confirming that the position of the attacking device does have an impact on results. The data fusion method has results worse than the single antenna ones, indicating that the CNN we have adopted is not suited to analize multiple CSI at the same time, and a re-design of the localization attack is needed also for MIMO technologies if a data fusion method is desired.

Interestingly, when obfuscation is active, the performance of the methodologies is inverted: Majority Vote performs as single antenna does, while data fusion roughly doubles the localization accuracy, even exceeding 50% for L2. This results is qualitatively confirmed by all the other obfuscation configurations: NF, CF3/7 and FC3/5/7. Even if localization accuracy with obfuscation remains well below any usability threshold, this behavior suggests that an accurate study of the electromagnetic environment may lead to better localization performance with MIMO technologies. This study is not only beyond the scope if this paper, but it requires an extensive collection of specific data from many different locations and scenarios.

8.4. OCTOPUS *and* FRANKENSTEIN: *Impact of Devices and Antenna Position*

The results for MIMO suggest to explore the role of the specific antenna position and MIMO technology on the localization performance, which is the goal of the OCTOPUS and FRANKENSTEIN configurations (Figs. 7 and 8). Fig. 12 summarizes the insight we get from these configurations for the Clean and CF5 cases. Again all other obfuscation cases confirm these results.

The key findings are two. First, looking at the Clean results for Majority Vote (4 antennas) it is clear that diversifying the

position of the CSI collection improves the localization performance. This is an indication that the typical MIMO antenna separation, not being designed for this task, might be too small for robust sensing. Second, collecting the CSI from the same device improves the performance of data fusion when obfuscation is active compared to its collection on different devices. Once again we do not have a theoretical explanation for this behavior, but it is possible that hardware imperfections introduce some fake fingerprints in the CSI collected from different devices that deceits the CNN. An indication that this may be the case comes from studies on device fingerprinting and its obfuscation (see [28]).

The final remark regards the main goal of this work: privacy protection. All the experiments we did and the results we present here indicate that CSI distortion is a viable solution to prevent unauthorized localization of unaware victims.

9. Conclusions

Recent works have shown that:

1. It is feasible to exploit communication signals opportunistically to sense indoor environments and localize unaware victims, paving the road for diffuse and illicit surveillance;
2. CSI randomization represents an effective countermeasure against such types of passive (CSI-based) localization attacks and, if properly implemented, it does not hamper communication performance and may find its way in future standards.

The work we presented in this paper contributes a significant step in the state of the art. For the first time, CSI obfuscation against a full-scale multi-point localization attack is considered, both with multiple devices and with the combination of MIMO technologies. Our results show that such an attack can improve the localization accuracy to a level where peoples privacy and security as well are at risk. On the other hand, the experimental results show that proper CSI randomization techniques can still disrupt localization attacks carried out with more than one device: Even when five devices are used, a case that make localization without obfuscation practically perfect, proper randomization completely destroys the possibility of localizing a victim. This is achieved while preserving most communications, although without further improvements it is still impossible to reduce the frame loss rate to zero for the highest MCSs.

This work also lays the foundation for more theoretically-sound randomization techniques that are virtually identical to real channel responses instead of altering the signal with prominent features that may eventually be detected (and overcome) by a powerful attacker. We emphasize that the requirements for the localization prevention system that we discussed in this paper are general and sound, meaning that they apply to any such technique. At the same time they suggest that finding a tractable formalization of the randomization process is a tough task, because of the frequency-time correlation structure of the process itself, and the complex interaction with the propagation environment.

We think that the topic discussed in this paper from a purely experimental point of view contributes a significant insight in the area of Wi-Fi sensing and people's localization (or even activity and gesture recognition). Its formalization, the design of protocols to ensure legitimate and useful sensing applications while preventing illicit ones, the optimization of techniques once a formal model is available are all very interesting topics we hope the research community will tackle and help to solve.

Acknowledgement

This work has been partially funded at the University of Brescia by GÉANT Educational Activities and Services Agreement ref. SER-21-142, Project "Design and Implementation of an 802.11 Privacy Preserving Sub-Layer (DI-P^{2}SL)"

References

[1] F. Restuccia, IEEE 802.11bf: Toward Ubiquitous Wi-Fi Sensing (3 2021). arXiv:2103.14918.

[2] K. Chetty, G. Smith, K. Woodbridge, Through-the-Wall Sensing of Personnel Using Passive Bistatic WiFi Radar at Standoff Distances, IEEE Trans. on Geoscience and Remote Sensing 50 (4) (2012) 1218–1226.

[3] F. Adib, D. Katabi, See through walls with WiFi!, in: Conf. of the Special Interest Group on Data Communication (SIGCOMM), ACM, Hong Kong, Aug. 2013, 2013, pp. 75–86.

[4] Z. Yang, Z. Zhou, Y. Liu, From RSSI to CSI: Indoor Localization via Channel Response, ACM Comput. Surv. 46 (2) (2013) 25:1–25:32.

[5] K. Wu, J. Xiao, Y. Yi, D. Chen, X. Luo, L. Ni, CSI-Based Indoor Localization, Trans. Parallel Distrib. Syst. 24 (7) (2013) 1300–1309.

[6] H. Abdelnasser, M. Youssef, K. A. Harras, WiGest: A ubiquitous WiFi-based gesture recognition system, in: IEEE Conference on Computer Communications (INFOCOM), 2015, pp. 1472–1480.

[7] F. Zhang, C. Chen, B. Wang, K. J. R. Liu, WiSpeed: A Statistical Electromagnetic Approach for Device-Free Indoor Speed Estimation, IEEE Internet of Things Jou. 5 (3) (2018) 2163–2177.

[8] J. Ding, Y. Wang, WiFi CSI-Based Human Activity Recognition Using Deep Recurrent Neural Network, IEEE Access 7 (2019) 174257–174269.

[9] L. Guo, Z. Lu, X. Wen, S. Zhou, Z. Han, From Signal to Image: Capturing Fine-Grained Human Poses With Commodity Wi-Fi, IEEE Communications Letters 24 (4) (2020) 802–806.

[10] Y. Wang, K. Wu, L. M. Ni, WiFall: Device-Free Fall Detection by Wireless Networks, IEEE Trans. on Mobile Computing 16 (2) (2017) 581–594.

[11] N. Damodaran, E. Haruni, M. Kokhkharova, J. Schäfer, Device free human activity and fall recognition using WiFi channel state information (CSI), CCF Trans. on Pervasive Computing and Interaction 2 (2020) 1–17.

[12] X. Wang, L. Gao, S. Mao, S. Pandey, CSI-based Fingerprinting for Indoor Localization: A Deep Learning Approach, Trans. Veh. Technol. 66 (1) (2017) 763–776.

[13] G.-S. Wu, P.-H. Tseng, A Deep Neural Network-Based Indoor Positioning Method using Channel State Information, in: Int. Conf. on Computing, Networking and Communications (ICNC), IEEE, Maui, HI, USA, Mar. 2018, 2018, pp. 290–294.

[14] E. Schmidt, D. Inupakutika, R. Mundlamuri, D. Akopian, Sdr-fi: Deep-learning-based indoor positioning via software-defined radio, IEEE Access 7 (2019) 145784–145797.

[15] M. Abbas, M. Elhamshary, H. Rizk, M. Torki, M. Youssef, WiDeep: WiFi-based Accurate and Robust Indoor Localization System using Deep Learning, in: IEEE Int. Conf. on Pervasive Computing and Communications (PerCom), 2019, pp. 1–10.

[16] A. Foliadis, M. H. Castañeda Garcia, R. A. Stirling-Gallacher, R. S. Thomä, CSI-Based Localization with CNNs Exploiting Phase Information, in: IEEE Wireless Communications and Networking Conf. (WCNC), 2021, pp. 1–6.

[17] Ma, Y. and Zhou, G. and S. Wang, S., WiFi sensing with channel state information: A survey, ACM Computing Surveys 52 (3) (2019) pp. 46:1–36.

[18] M. Widmaier, M. Arnold, S. Dorner, S. Cammerer, S. ten Brink, Towards Practical Indoor Positioning Based on Massive MIMO Systems, in: 90th IEEE Vehicular Technology Conf. (VTC2019-Fall), 2019, pp. 1–6.

[19] Ricciato, Fabio and Sciancalepore, Savio and Gringoli, Francesco and Facchi, Nicolò and Boggia, Gennaro, Position and Velocity Estimation of a Non-Cooperative Source From Asynchronous Packet Arrival Time Measurement, IEEE Trans. on Mobile Computing 17 (8) (2018) 2166–2179.

[20] E. Gönültaş, E. Lei, J. Langerman, H. Huang, C. Studer, CSI-Based Multi-Antenna and Multi-Point Indoor Positioning Using Probability Fusion (2020). arXiv:2009.02798.

[21] C. Studer, S. Medjkouh, E. Gönültaş, T. Goldstein, O. Tirkkonen, Channel Charting: Locating Users Within the Radio Environment Using Channel State Information, IEEE Access 6 (2018) 47682–47698.

[22] M. Cominelli, F. Gringoli, R. Lo Cigno, Passive Device-Free Multi-Point CSI Localization and Its Obfuscation with Randomized Filtering, in: 19th IEEE Mediterranean Communication and Computer Networking Conference (MedComNet), 2021, pp. 1–8.

[23] M. Cominelli, F. Kosterhon, F. Gringoli, R. Lo Cigno, A. Asadi, An Experimental Study of CSI Management to Preserve Location Privacy, in: 14th ACM Workshop on Wireless Network Testbeds, Experimental evaluation & CHaracterization (WiNTECH), London, UK, 2020, pp. 1–8.

[24] M. Cominelli, F. Kosterhon, F. Gringoli, R. Lo Cigno, A. Asadi, IEEE 802.11 CSI randomization to preserve location privacy: An empirical evaluation in different scenarios, Elsevier Computer Networks 191 (22) (2021) 107970.

[25] D. Nguyen, C. Sahin, B. Shishkin, N. Kandasamy, K. R. Dandekar, A Real-Time and Protocol-Aware Reactive Jamming Framework Built on Software-Defined Radios, in: ACM Workshop on Software Radio Implementation Forum, 2014, p. 15–22.

[26] M. Schulz, F. Gringoli, D. Steinmetzer, M. Koch, M. Hollick, Massive Reactive Smartphone-Based Jamming Using Arbitrary Waveforms and Adaptive Power Control, in: 10th ACM Conf. on Security and Privacy in Wireless and Mobile Networks (WiSec), 2017, p. 111–121.

[27] Y. Qiao, O. Zhang, W. Zhou, K. Srinivasan, A. Arora, PhyCloak: Obfuscating Sensing from Communication Signals, in: 13th USENIX Conf. on Networked Systems Design and Implementation (NSDI'16), Santa Clara, CA, USA, 2016, p. 685–699.

[28] Abanto-Leon, Luis F. and Bäuml, Andreas and Sim, Gek Hong (Allyson) and Hollick, Matthias and Asadi, Arash, Stay Connected, Leave no Trace: Enhancing Security and Privacy in WiFi via Obfuscating Radiometric Fingerprints, Proc. ACM Meas. Anal. Comput. Syst. 4 (2020) 44:1–44:31.

[29] M. Cominelli, F. Gringoli, R. Lo Cigno, Non Intrusive Wi-Fi CSI Obfuscation Against Active Localization Attacks, in: 16th IFIP/IEEE Conf. on Wireless On demand Network Systems and Services (WONS), 2021, pp. 87–94.

[30] X. Jiao, M. Mehari, W. Liu, M. Aslam, I. Moerman, Openwifi CSI Fuzzer for Authorized Sensing and Covert Channels, in: 14th ACM Conf. on Security and Privacy in Wireless and Mobile Networks, 2021, p. 377–379.

[31] M. Di Renzo, M. Debbah, D. Phan-Huy, et al., Smart radio environments empowered by reconfigurable AI meta-surfaces: an idea whose time has come, J Wireless Com Network 129 (2019).

[32] Y. Zhang, C. Qu, Y. Wang, An Indoor Positioning Method Based on CSI by Using Features Optimization Mechanism With LSTM, IEEE Sensors Journal 20 (9) (2020) 4868–4878.

[33] W. Li, R. J. Piechocki, K. Woodbridge, C. Tang, K. Chetty, Passive WiFi Radar for Human Sensing Using a Stand-Alone Access Point, IEEE Trans. on Geoscience and Remote Sensing 59 (3) (2021) 1986–1998.

[34] E. Gönültaş, E. Lei, J. Langerman, H. Huang, C. Studer, CSI-Based Multi-Antenna and Multi-Point Indoor Positioning Using Probability Fusion, IEEE Trans. on Wireless CommunicationsAvailable online Sept. 10 2021 (2021).

[35] L. Guo, L. Wang, C. Lin, J. Liu, B. Lu, J. Fang, Z. Liu, Z. Shan, J. Yang, S. Guo, Wiar: A Public Dataset for Wifi-Based Activity Recognition, IEEE Access 7 (2019) 154935–154945.

[36] F. Gringoli, M. Schulz, J. Link, M. Hollick, Free your CSI: A channel state information extraction platform for modern Wi-Fi chipsets, in: 13th Int. Workshop on Wireless Network Testbeds, Experimental Evaluation and Characterization (WiNTECH '19), ACM, Los Cabos, Mexico, Oct. 2019, 2019, pp. 21–28.

[37] C. Cai, L. Deng, M. Zheng, S. Li, PILC: Passive Indoor Localization Based on Convolutional Neural Networks, in: Ubiquitous Positioning, Indoor Navigation and Location-Based Services (UPINLBS), Wuhan, China, 2018, pp. 1–6.