BFId: Identity Inference Attacks Utilizing Beamforming Feedback Information

Julian Todt
KASTEL Security Research Labs
Karlsruhe Institute of Technology
Karlsruhe, Germany
julian.todt@kit.edu

Felix Morsbach
KASTEL Security Research Labs
Karlsruhe Institute of Technology
Karlsruhe, Germany
felix.morsbach@kit.edu

Thorsten Strufe
KASTEL Security Research Labs
Karlsruhe Institute of Technology
Karlsruhe, Germany
thorsten.strufe@kit.edu

Abstract

Beamforming, as introduced in WiFi 5, requires clients to broadcast observations of their channel characteristics. This introduces a new information source for WiFi sensing with privacy threats that have not been explored, so far. With WiFi networks being ubiquitous in our everyday lives, the impact of unknown privacy threats is likely severe.

To investigate this concern, we introduce **BFId**, the first identity inference attack using BFI-based sensing and evaluate its efficacy on a novel dataset containing WiFi recordings of 197 individuals. We show that we can infer the identity of individuals with very high accuracy, across different walking styles and perspectives, even with large sample sizes.

CCS Concepts

• **Security and privacy → Pseudonymity, anonymity and untraceability**; **Privacy protections**.

Keywords

privacy, identification, wireless sensing, beamforming

ACM Reference Format:
Julian Todt, Felix Morsbach, and Thorsten Strufe. 2025. BFId: Identity Inference Attacks Utilizing Beamforming Feedback Information. In *Proceedings of the 2025 ACM SIGSAC Conference on Computer and Communications Security (CCS '25), October 13–17, 2025, Taipei, Taiwan.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3719027.3765062

1 Introduction

Wireless networks are ubiquitous – at home, at work, in a *smart* city. With it comes WiFi sensing, the inference of information about the networks environment from its signal propagation characteristics. As signals propagate through matter, they interfere with it – they are either transmitted, reflected, absorbed, polarized, diffracted, scattered, or refracted [1]. By comparing an expected signal with a received signal, the interference can be estimated and used for error correction of the received data.

However, this estimation inherently contains information about the environment that the signal traveled through. For example, a human in the signal's path will lead to more interference than a clear path. By carefully analyzing the signal's interference with the environment, even with previous generations of WiFi standards, certain aspects of the environment can be inferred. For example, whether humans are present [3], what activities they are doing [14, 25, 32, 61], or who they are [5, 18, 26, 31, 34, 42, 46, 48, 49, 51–53, 56, 64, 66–68, 71–74, 76, 77, 81]. While much of the extant literature focuses on the utility of WiFi sensing, the privacy threats appear clear: with WiFi networks in the most private places and inferences such as activity recognition shown to have high accuracy, adversaries may be able to infer very sensitive information. In the following, we will focus on identity inference as one archetypical privacy threat.

Identity inference based on WiFi can be done by analyzing different sources, but most prominent in recent years has been the analysis of *Channel State Information (CSI)*, a built-in part of the physical layer of WiFi. CSI, however, while being a rich source of information for many different sensing applications, comes with a significant drawback: As it was not envisioned to be used outside of error correction, it is part of the physical layer of WiFi and thus accessing this information requires modified firmware which is only readily available for specific hardware [38]. On the one hand, this makes utilizing WiFi for genuine sensing tasks harder, but on the other hand, it also decreases the privacy risk associated with the sensing capabilities of WiFi.

To enable higher bandwidths, WiFi 5 (802.11ac) introduced *beamforming*. Beamforming utilizes similar information on the physical environment as CSI, but on the sender instead of the receiver side. In a typical WiFi scenario, clients send *Beamforming Feedback Information (BFI)* back to the access point, a compressed representation of the current signal characteristics. The key difference to CSI-based sensing is that BFI is broadcast back to the access point, unencrypted. Thus, this information can be easily accessed with common-off-the-shelf hardware, reducing the barrier for developing and deploying genuine WiFi sensing applications, but at the same time making it also easier for malicious actors to build and execute attacks. Furthermore, while CSI-based systems can only access the perspective from access point to malicious node, for BFI a single malicious node can record every perspective between access point and legitimate clients as long as it is anywhere within the broadcasting range. With the IEEE consortium planning to standardize WiFi sensing for broad applications in 802.11bf – without any privacy protections – we feel it necessary to further investigate and highlight the privacy risks associated with WiFi sensing, particularly those of BFI.

The general ability of CSI-based sensing systems to infer identities has been documented, and to some extent investigated by previous work [33, 38]. However, BFI-based sensing has not been investigated to date. As the transmitted beamforming feedback

CCS '25, Taipei, Taiwan

ACM ISBN 979-8-4007-1525-9/2025/10
https://doi.org/10.1145/3719027.3765062

information is a compressed version of CSI at a lower time resolution, one might argue that the sensing capabilities of BFI based systems are in principle inferior than those of CSI based systems. However, the compression also inherently represents some form of pre-processing that filters out noise from the raw signal, possibly even enhancing the sensing potential. Whether BFI is still a rich enough source of information for identity inference remains unknown. At the same time, BFI and CSI have different attack surfaces with BFI's broadcasting facilitating trivial collection from various perspectives (i.e., multiple devices at different positions relative to the subjects). The impact of this on sensing accuracy remains unclear. Finally, previous studies of WiFi sensing based identification struggle with unrealistically low numbers of participants (see Table 1). It is therefore unclear how WiFi sensing for identification scales and to which extent it may be used in real-world organizational or smart city contexts. In summary, it is unclear to which extent BFI-based sensing is able to infer the identity of humans. Thus, and due to its significantly weaker adversary model and thereby potentially higher privacy risk, we investigate this privacy threat.

To address these open problems, we introduce *BFId*, the first identity inference attack using beamforming feedback information. As our goal is not to improve the absolute efficacy of WiFi sensing based identification, but to investigate and demonstrate the potential identification threat of BFI, the attack processes raw WiFi artifacts (BFI or CSI) by utilizing recurrent neural works. This makes it easier to deploy compared to previous attacks. It also more accurately approximates the lower bound of the associated privacy risk. To evaluate the attack's efficacy, we conducted an extensive user study in which we recorded WiFi artifacts of 197 participants, including multiple perspectives and walking styles. These recordings include BFI and CSI information simultaneously, which allows us to directly compare the efficacy of both sensing approaches. We show that individuals can be recognized with very high accuracy (99.5% ± 0.38) with our BFI-based attack. Furthermore, BFId is not only able to infer the identity of individuals, our experiments also demonstrate that in a direct comparison it is able to do so better than CSI-based attacks for large populations. This also holds for identifying individuals across walking styles, from multiple different perspectives, and at reduced sample rates. We find that BFI-based attacks both assume a weaker adversary model—they do not require specialized hardware and custom firmware—and achieve higher accuracy compared to CSI-based attacks.

To summarize, our main contributions in this paper are:

- We propose BFId, the first identity inference attack using beamforming feedback information
- We conduct an extensive user study with WiFi recordings (BFI & CSI) of 197 participants and make this dataset available to interested researchers
- We demonstrate the efficacy of our attack with very high accuracy and its superior robustness to CSI-based attacks

We thereby highlight the identification threat of beamforming and advance the understanding of the inherent privacy threats of WiFi sensing.

The remainder of this paper is organized as follows: We start by introducing the relevant background knowledge in Section 2 and highlight key works in the literature in Section 3. We introduce our attack in Section 4, describe our dataset for evaluation and its collection in Section 5, and evaluate the attack with the dataset in Section 6. We discuss the results in Section 7 and conclude the paper in Section 8.

2 Background

While WiFi is most commonly known for its data transmission capabilities, a variety of implicit or explicit properties of these transmissions can also be used to make inferences about the environment. A signal emitted by a transmitter is subject to interference based on the surrounding environment, which results in a slightly different signal received by a receiver compared to the signal the transmitter emitted. The difference between the emitted and received signal is the result of the physical properties of the electromagnetic waves that make up this signal and the matter through that they travel and interact with. For example, a concrete wall within the signal's path will result in parts of the signal being reflected and absorbed, which means the received signal will be less strong. Similarly, a human walking through a signal's path will cause different reflection, refraction and transmission properties. By analyzing these signal alterations, it is inherently possible to infer information about the environment. This information can and is then used to, for example monitor service quality or improve the transmission via error correction. However, the same information can also be used for sensing applications.

The most straightforward property of WiFi transmissions that can be used for sensing is the received signal strength (RSS). It can be accessed from the application layer of a device and allows simple sensing tasks, such as human detection [3] and localization [70, 79]. At the same time, it's an aggregation of a variety of information and is not robust which means it cannot be used for fine-grained sensing tasks [4].

2.1 Channel State Information (CSI)

To increase transmission rates, WiFi uses multiple subcarriers (i.e. different frequencies) and divides the data to send over them using orthogonal frequency-division multiplexing (OFDM). Every subcarrier interacts with the environment slightly differently, similar to how different wavelengths of light interact differently with a prism. Further, due to the reflection of the radio wave, the signal will arrive at the receiver through multiple paths, each causing a unique attenuation and phase shift of the signal [4].

The magnitude and phase shifts of each subcarrier at the receiver is commonly referred to as channel state information (CSI). To measure CSI, pre-defined symbols, so called long training fields (LTFs), are sent in the packet preamble of each WiFi transmission. The receiver can then compare the expected signal (i.e. the predefined LTFs) with the actual received signal and use this information for error correction of the transmission's remaining payload. This results in a three-dimensional complex-valued CSI matrix $H \in \mathbb{C}^{N \times M \times K}$, where the N and M give the number of receiving and transmitting antennas and K the number of used subcarriers [38]. Recording and processing a timeseries of CSI-matrices can then be used for sensing applications.

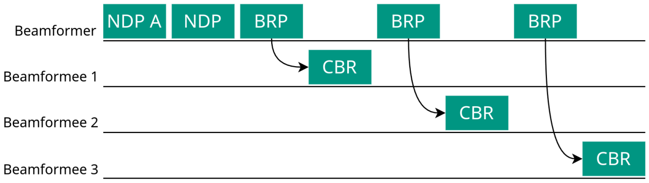

Figure 1: Beamforming channel sounding procedure with a null data packet (NDP), its announcement frame (NDPA), beamforming report polls (BRP), and the compressed beamforming reports (CBR)

A multitude of sensing applications have been realized using CSI, for example activity recognition [14, 25, 32, 61], gesture recognition [75, 80], object recognition [58, 82], human localization and tracking [35, 43, 44, 54, 78], respiratory rate estimation [37, 55], and human counting [11, 20, 60]. In this paper, we focus on human identification. We discuss a variety of these approaches in our related work section.

2.2 Beamforming Feedback Information (BFI)

To further increase transmission rates, particularly at medium range, the 802.11ac standard introduced the option for *beamforming*. With beamforming, the multiple antennas of WiFi devices send transmissions no longer unidirectional, but rather steer them towards the receiver. For this, the sender (then called: beamformer) requires information about the environment towards the receiver (beamformee) in order to steer the transmissions. For this, the standard defines a channel sounding procedure (shown in Figure 1) which is initiated by the access point (beamformer) regularly through a null data packet (NDP) announcement frame. Beamformees will reply to this announcement. The actual NDP is then sent by the access point which contains one VHT-LTF (very high throughput long training field) per spatial stream used in the transmission. Beamformees will then use the CSI of these VHT-LTFs to calculate so-called feedback matrices for each subcarrier. The feedback matrix is compressed into beamforming angles which are sent back to the beamformer. The beamformer can then calculate a steering matrix which can be used to direct the transmission towards the beamformee [6, 16].

The compressed beamforming report (CBR) contains the beamforming angles which are referred to as BFI. Their size depends on the number of beamformees (single user vs. multi user) and the number of subcarriers. As they are based on CSI, similar sensing applications as with CSI are possible and as they transmitted over the air unencrypted, they can be more easily processed compared to CSI. We discuss some approaches utilizing BFI for sensing in the following section.

3 Related Work

Many works discuss WiFi based identification and sensing in general. There have been multiple comprehensive surveys [33, 38] on this topic. Thus, in this chapter, we focus on identification as the sensing task and the latest state-of-the-art studies on CSI-based systems and a general overview on BFI-based systems. To the best of our knowledge, we are the first to investigate BFI-based identification.

CSI-based identification. Table 1 shows an overview over CSI-based identification systems. There are a multitude of studies which, while having minor differences in their approaches, all reach the same conclusion: the inference of identities via CSI-based WiFi sensing is reliably possible. The vast majority of approaches use recordings of gait sequences for identification, but there are exceptions, e.g. using lip-motions [42], keystrokes [18] or no moving at all, but the individual just standing [53] or sitting [51]. When using gait, most approaches record individuals while walking orthogonally to the line-of-sight (LOS) between sender and receiver. Two early approaches had participants walk parallel to the LOS [26, 71] and some approaches have opted to have participants walk freely, but only one approach considered multiple perspectives [76]. This shows that the influence of perspectives does still appear to be an underinvestigated topic while gait is the most robust biometric trait for identification.

Similarly, the majority of approaches use the 5 GHz band for WiFi sensing, though some approaches do use the 2.4 GHz band. The difference between those bands has been explored further in [19]. There, the authors find only small differences between the 2.4 and 5 GHz band, largely resulting from the larger number of subcarriers that are used in 5 GHz. WiFi 6E also introduced the 6 GHz band with even greater bandwidths. It has not yet been evaluated for use in WiFi sensing, although the difference between the 6 GHz and 5 GHz is even smaller than that between 5 GHz and 2.4 GHz which means that the impact of this choice of band should be minimal.

Another point of interest in comparing CSI-based identification approaches is the choice of WiFi network interface card (NIC). As the extraction of CSI requires modified firmware, only very few NICs are compatible with CSI-based sensing. Most commonly, the Intel 5300 is being used, a NIC released in 2008, even in studies published as late as 2023. We will discuss this further in Section 4 when comparing adversary models.

Pre-processing has been done with a variety of complex combinations of filtering, (inverse) Fourier transformations, and PCAs (see Table 1). Only few approaches limit their pre-processing and rely on their deep neural networks to identify relevant features. For model architectures, there appears to be a trend towards convolutional neural networks (CNNs), this however limits their ability to consider the time dimension of CSI. Only WiHF uses a recurrent neural network (RNN) [34]. An overview and benchmark over model architectures can also be found in SenseFi [65]. In summary, there is no consensus on which pre-processing steps and model architectures are the most beneficial to WiFi sensing, which suggests that the problem is not yet well understood. There is no comprehensive comparison of pre-processing options. Even with in-depth domain knowledge, it remains unclear how to implement the best processing chain.

The biggest limitation of related work we find is the number of identities that are being considered. The number of identities is an important factor, as the larger the population, the harder the problem becomes. Only a single approach attempts to distinguish more than 50 individuals [72]. Considering that many approaches claim they can be used in smart spaces or smart cities, this is clearly unrealistic. To make the study at least a little more realistic, and the identification more challenging than a simple classification task, we

Table 1: Comparison of CSI-based identification approaches (ordered by year of publication)

Paper	Identities	Accuracy	Pre-Processing	Model Arch.	Perspective	NIC	Channel
FreeSense [64]	6	88.9	PCA, DWT, DTW	kNN	orthogonal	Intel 5300	2.4 GHz
WifiU [56]	50	79.3	Butterworth Filter, PCA	SVM	orthogonal	Intel 5300	5 GHz
WFID [26]	9	93.9	Subcarrier Amplitude Frequency	SVM	parallel	Intel 5300	unknown
WiFi-ID [73]	6	77	Butterworth Filter, CWT	SAC	orthogonal	Intel 5300	5 GHz
WiWho [71]	6	80	FFT, Butterworth Filter	Dec. Tree	parallel	Intel 5300	unknown
SLL+17 [49]	11	94	Bandpass Filter, Subcarrier Select.	SVM	orthogonal	Intel 5300	unknown
WiID [48]	15	93.4	PCA, STFT, Freq. Time-Series	SVM	gestures	Intel 5300	2.4 GHz
CrossSense [72]	100	80	Various different tested	MLP	orthogonal	Intel 5300	5 GHz
NeuralWave [42]	24	87.8	Phase Calib., Wavelet Denoising	CNN	orthogonal	Intel 5300	5 GHz
BioID [77]	5	90	Butterworth Filter, PCA, DWT	kNN	lip-motions	Intel 5300	2.4 GHz
AutoID [81]	20	91	Shapelet Coefficient Matrix	DNN[a]	free	TP-Link N750	5 GHz
WiPIN [53]	32	92	Butterworth Filter, IFFT, FFT	SVM	standing	Intel 5300	5 GHz
XModal-ID [31]	6	80	STFT, Hermite Spectogram	MLP	orthogonal	Intel 5300	5 GHz
WiHF [34]	6	96.7	Bandpass Filter, STFT	RNN	gestures	Intel 5300	5 GHz
LW-WiID [5]	50	99.7	Frequency Energy Graph	CNN	orthogonal	Intel 5300	5 GHz
Gate-ID [74]	20	75.7	Butterworth Filter, PCA	DNN[a]	orthogonal	Intel 5300	5 GHz
WiOne [18]	16	94.7	Rician Fading	CNN	keystrokes	Intel 5300	unknown
CAUTION [52]	15	88.9	None	CNN	free	TP-Link N750	5 GHz
EfficientFi [67]	15	98.6	CNN-based Compression	CNN	orthogonal	TP-Link N750	5 GHz
GaitSense [76]	11	76	Gait Body-Coord. Velocity Profile	CNN	6 different[b]	Intel 5300	5 GHz
TDK+22 [51]	9	98.2	AGC compensation	CNN	sitting	Raspberry Pi	2.4 GHz
AutoFi [66]	14	83.3	None	CNN-MLP	orthogonal	Intel 5300	5 GHz
3D-ID [46]	28	85.3	2D AoA image, SMPL	DNN[a]	free	Intel 5300	5 GHz
SecureSense [68]	12	95.8	None	CNN	orthogonal	TP-Link N750	5 GHz
BFId (ours)[c]	170	84.3	None	LSTM	4 different	Intel AX210	6 GHz

[a] Custom complex DNN model that combines various building blocks [b] 4 different angles across LOS + walking in square and circle [c] applied to our CSI data, not BFI

suggest that datasets of at least over a hundred different individuals are required to produce insights on convincing identification risks in real environments.

BFI-based sensing. Table 2 shows an overview over existing BFI-based sensing approaches. There already has been a few but valuable insights into the sensing capabilities of BFI, from human detection [40, 41, 62] over localization [15, 50] to activity recognition [23, 24]. However, for many conceivable sensing applications, investigations are still missing, especially compared to CSI-based sensing.

Similar to the CSI-based approaches, we find a variety of pre-processing steps and model architectures. Many pre-processing approaches evolve around attempting to restore the CSI that was used to calculate the transmitted BFI. This allows these systems to then use the same further processing steps as they do for the related CSI sensing task. At the same time, we do see more approaches that do not use pre-processing anymore, but rather rely on their deep learning models to extract relevant features from the raw data. This shows that also the understanding of suitable processing pipelines for BFI is limited.

Due to the wider availability of devices supporting BFI extraction compared to CSI, we see a variety of different setups. While many approaches take advantage of wider channels in newer WiFi standards, this is not ubiquitous. Also, we see most access points using four antennas and nodes using either one or two. Only two approaches also use APs as their beamformees which results in 4x4 antenna setups. This however, is clearly unrealistic, as in most circumstances, access points will communicate with clients and not with each other. As such, a realistic setup would use four or even more antennas for the beamformer, as this is now standard for new devices, and two antennas for the beamformee.

Mitigation approaches. There have a been a limited number of studies on the privacy problems of WiFi sensing and approaches to mitigate them for CSI-based sensing, in particular human localization. All of them are based on adding randomized noise to the training fields of WiFi transmissions which results in a significant loss of accuracy of the sensing approaches. The noise can be added directly by the transmitter which relies on a passive adversary which does not compromise the transmitter [8, 10]. It can also be added by an intelligent reflective surface (IRS) which acts as a repeater of the transmitted data, but with added noise [9, 45]. This approach requires additional hardware. Either approaches have currently only been prototyped using software defined radios and not COTS hardware. Due to the noise added, utility degradations (decreased bandwidth) must also be expected. In essence, no promising approaches for a real-world mitigation of WiFi sensing privacy risks have been proposed.

Table 2: Comparison of BFI-based WiFi sensing approaches (ordered by year of publication)

Paper	Inference	Pre-Processing	Model Arch.	Bandwidth	Antennas	Sampling Rate
MMI+18 [41]	human detection	Decompression	n/a	20 MHz	4x1	unknown
MIF+19 [40]	human detection	Low-pass filter, Normal.	MLP	20 MHz	4x1	100 Hz
FMA+19 [15]	human localization	None	kNN, SVM, RandF	unknown	4x1	100 Hz
TIF+19 [50]	human localization	Low-pass filter, PCA	MLP	unknown	4x1	100 Hz
CSI2Image [29]	image reconstruction	None	GAN	80 MHz	2x2	unknown
KSA+22 [28]	respiratory rate	PCA, DFT, Bandpass filter	n/a	80 MHz	4x4	5 Hz
KIY+22 [30]	human localization	None	RandF, SVM, LightGBM	20 MHz	4x4	10 Hz
DeepCSI [39]	device identification	Decompression	CNN	80 MHz	3x1/2	unknown
Wi-BFI [23]	activity recognition	None	CNN	80 MHz	3x1	unknown
BeamSense[a] [24]	activity recognition	None	CNN	80 MHz	3x1	unknown
BeamSense[a] [59]	device localization, object tracking, sign language recognition	CSI reconstruction	CNN	80 MHz	Multiple	30 Hz
BeamCount [7]	crowd counting	Time Series Prediction	CNN	unknown	4x2	unknown
Wi2DMeasure [57]	object size	CSI ratio, Singularity extract.	n/a	80 MHz	4x1	unknown
BFMSense [69]	respiratory rate, walking speed	BFM ratio calculation	n/a	80 MHz	4x1	10 Hz
LeakyBeam [62]	occupancy detection	Phase norm., subcarrier fusion	n/a	20/40/80 MHz	Multiple	10-17 Hz
BFId (ours)	identity	None	LSTM	160 MHz	4x2	10 Hz

[a] Both approaches have been named ’BeamSense’ but have no relation to each other.

4 BFId

In this section, we introduce *BFId*, our WiFi sensing identity inference attack. We would like to highlight that, even though related work may consider their identification systems to be "privacy-preserving" [12, 18, 51, 64–67, 76, 81], "less privacy intrusive" [72], "not cause any privacy concerns" [52] or "avoid[ing] [..] personal privacy invasion" [5], we consider this to be an attack on privacy. While legitimate use-cases may exist, considering the pervasiveness of WiFi, especially in public areas, the ability to record through walls, and the simplicity to do so without consent, we see severe potential threats to everyone’s privacy. We therefore start by explicitly and elaborately defining the threat and adversary model that we consider in this paper and highlight the differences between CSI and BFI-based sensing. Finally, we describe the technical details on the used pre-processing and machine learning models.

4.1 Threat model

Our threat model primarily addresses identity disclosure. WiFi sensing identifies an individual by linking independent recordings of individuals that walk through areas covered by WiFi networks—breaking anonymous presence in public—and thereby causes privacy harm. Thus, we define identification as identity disclosure by linking together recordings of the same individuals from different points in time: Once during known presence and once during potential presence of that individual. As recordings, we solely consider WiFi signal observations which capture the individuals in a WiFi field. Linking together such recordings facilitates identification and thus breaks anonymity, even if no direct mapping to a real-world identity (i.e., name or social security number) is initially available.

Put colloquially, the adversary aims to recognize an individual who they have observed before, similar to how one might recognize the same person on the bus to work everyday, but without necessarily knowing the real name of this person. This inherently requires the adversary to only have access to previous recordings of the target individual but not to have any further auxiliary information on them. At the same time, further linking the recordings to a real-world identity is generally possible down the line via auxiliary information and would most likely amplify the privacy harm.

This identity disclosure can cause harm in multiple ways, even if there is no direct link to a real-world identity. For instance, if the adversary can link an individual’s recording in a malicious or compromising situation to a recording of the same individual in a benign situation, they could act upon this link in a harmful way by confronting the individual in a benign situation. For example, imagine an oppressive state records individuals on their way to protest through a coffee shop’s WiFi along the way. Later, the state’s militia could await these individuals while they are on a benign walk along this coffee shop. This makes this threat especially problematic in low-regulation and high-surveillance contexts, such as authoritarian regimes, as government actors may use this technique to track dissidents or suppress protests.

Compared to traditional surveillance mechanisms, such as CCTV, WiFi-based surveillance is even more problematic in two ways: Stealthiness and availability.

First, WiFi infrastructure is ubiquitous and primarily associated with benign services. Cameras, in comparison, are less common, their usage might be regulated, and their sole purpose is surveillance. WiFi sensing allows an adversary to employ surveillance without raising suspicion, effectively creating an "inverse panopticon", where individuals behave as though they are unobserved, while being silently tracked. Thus, an adversary might choose WiFi-based attacks to give individuals a false sense of security. In the protester example, individuals might purposely avoid an area with video cameras, but will ignore seemingly harmless WiFi APs.

Second, modern environments host a wide range of devices that are capable of sending and especially receiving BFI reports. As a result, an adversary may be able to leverage existing infrastructure, including remote devices that they have not deployed themselves, but can take control of. Particularly since many WiFi-enabled devices (for instance IoT devices) are insecure in practice and poorly maintained, vulnerable WiFi-enabled coffee makers for instance. This lowers the barrier for conducting large-scale and long-term tracking.

4.2 Adversary model

We consider a passive adversary who aims to compromise the anonymity of individuals that walk through a WiFi field. They do this by observing WiFi signals and creating recordings of them. These recordings are then used to create an identification system that links recordings from different times to the same individual.

To do this, the adversary requires labeled training data. Our adversary model does not assume that such training data exists and thus needs to be created by the adversary. To record WiFi sensing data — for training and inference — the adversary must have access to a WiFi device within the broadcasting range of the target individual and network. However, the adversary does not need to have access to the WiFi access point or its network (e.g. the network password is not needed). Any recording can be done completely passively, it relies solely on the unencrypted parts of legitimate traffic within the WiFi network. This also hides the presence of the adversary, as they only monitor natural traffic.

Our scenario assumes the adversary to have continuous access to a device in the network's range, but only sporadic access to the ground truth label, obtained on an arbitrary side-channel. The labeled training examples would typically correspond to benign situations, whereas unlabeled examples for inference may be from compromising situations.

For example, one might consider the individual's smartphone's MAC address as their label. Then, the adversary can link recordings where the individual was carrying their smartphone with ones where they were not. This matches our protester scenario where individuals carry their phone in benign situations but intentionally do not in compromising situations, a common mitigation strategy often recommended by civil rights organizations. This enables attacks both in retrospect and in near-real time. Overall, this adversary model aligns well with the (mostly implicit) assumptions of previous works.

Differences between BFI- and CSI-based sensing. While large parts of the adversary model are the same for both CSI and BFI, there are differences regarding hardware requirements and recording perspectives. For CSI, the perspective being recorded is always between access point and the malicious node. Any traffic that the access point broadcasts can be received by the malicious node, its CSI calculated and thereby the channel between access point and malicious node estimated. For this, it is irrelevant where and who the legitimate recipient of the traffic is, as the access point broadcasts the signal omnidirectional. This also means that in order to record multiple perspectives, multiple malicious nodes are required.

For beamforming on the other hand, BFI reports are sent by legitimate nodes whenever prompted by the access point. During this channel sounding procedure, all legitimate clients will broadcast information about the channel between access point and themselves. Due to the fact that this communication back to the access point happens unencrypted, a single malicious node can passively record all these beamforming reports (even without being part of the network). This means that for BFI, the malicious node can be placed anywhere within range of the network and will receive information about all channels between the access point and the legitimate devices. At the same time, for CSI, the malicious node must be placed in the exact location for which the channel should be estimated and it will receive information about only this perspective.

CSI and BFI are processed by the network card on different layers of the protocol stack which has implications for the adversary model. CSI is an implicit part of WiFi's physical layer which is primarily used for error correction. Access to this information is therefore generally not possible from applications because CSI is processed directly by the network card. To access CSI, one needs modified firmware for the WiFi adapter which only exists for few NICs, primarily the Intel 5300 (via the Linux 802.11n CSI tool [21]), certain Atheros 802.11n PCI/PCI-E chips (for example used in TP-Link N750 APs; via Atheros CSI tool [63]), certain broadcom chips (for example used in Raspberry Pis; via nexmon CSI extractor [17, 47], and, most recently, the Intel AX210 via PicoScenes [27]. BFI on the other hand is easily accessible and does not require specific hardware because clients broadcast it back to the access point on the MAC layer. Any WiFi NIC can be set to monitor mode and the relevant packets can then be easily recorded. For context, Wu et al. [59] find that less than 6% of real-world deployed WiFi devices support the extraction of CSI while even in 2023 over half of them supported BFI, and this number most likely will only increase in the future as it is part of the WiFi standard.

After recording, many related WiFi sensing approaches employ significant pre-processing of their data to extract useful features for classification, for example bandpass-filtering or Fourier transformations (see Table 1). This indirectly assumes that the adversary possesses the required domain knowledge to properly choose, implement and tune such pre-processing steps. However, assuming such domain knowledge as requirement to mount such attacks underestimates the privacy risk, if it turns out that such domain knowledge is not actually required. While such pre-processing steps might enhance the attack's efficacy, it requires a stronger adversary. But as our goal in this paper is to demonstrate and investigate the lower bound privacy threat of BFI-based identification, we deem such assumptions unnecessarily strict and thus do not employ any pre-processing steps besides feature standardization.

A summary and comparison of the adversary models can be found in Table 3. It can be concluded that the BFI adversary model is weaker. It has less requirements on hardware and positioning of the malicious node while enabling multiple perspectives.

Table 3: Comparison of adversary models

	CSI	BFI
Network access req.	×	×
Fully passive	✓	✓
No. of perspectives	1	Any
Perspective from AP to	malicious node	legitimate nodes
Specific hardware req.	✓	×

4.3 Our attack

All the identification attacks that we discussed in Section 3 rely on a fixed length of each sample. For example, Pokkunuru et al. [42] make every sample exactly four seconds long and assume a consistent sample rate of 2000 Hz of CSI data points. This is clearly unrealistic. While one person might take four seconds to walk through the WiFi field, another one might make the same trip in just three seconds. Fixed length samples ignore that walking speed is a simple and straightforward, yet important biometric feature [22]. Furthermore, fixing the sample length assumes a constant sample rate, which is unrealistic, especially for CSI, as it is dependent on legitimate traffic to generate data points. Rather, samples most likely are of different length and the sample rate might be inconsistent, such that the time difference between points in the timeseries is variable. We therefore design our attack to be able to process variable length inputs by deploying simple and straightforward recurrent neural networks with a softmax classification as the final layer. We also add an additional feature to the timeseries data that contains the time elapsed since the last observed data point.

As we want to establish a lower bound for the privacy threat of BFI, we purposely keep our processing pipeline simple. We do not employ any pre-processing besides feature standardization of our WiFi recordings, as in our adversary model we do not assume any specific domain knowledge. Rather, we propose that an ordinary machine learning model in combination with sufficient training data should be able to learn relevant features and their dependencies within the data itself. As such model, we utilize a standard LSTM followed by two fully-connected layers, each with batch normalization and ReLU activation functions, and ending in a softmax layer. We utilized this architecture for both CSI and BFI and trained the models with the ADAM optimizer until convergence. The only difference being the input dimension which is given by the number of features in the CSI and BFI sequences. Additional learning hyperparameters, such as the initial learning rate, the batch size, weight decay, but also further architectural hyperparameters such as the sizes of the LSTM and fully-connected layer have been optimized individually for CSI and BFI using the Tree-structured Parzen Estimator of the Optuna library [2].

In our evaluations, we split the samples between training and test 80:20. If not stated otherwise, we repeated each experiment five times with independent splits and measure identification accuracy. That is the number of correct classifications over the total number of identifications. As the maximum, a value of 1 means that all samples where correctly identified, while as a lower bound 1 divided by the number of identities is the chance level of a correct identification.

5 Data Collection

Existing data sets for WiFi sensing based identification are severely limited in the number of participants that they recorded, see Table 1. They also do not record multiple perspectives simultaneously and, most importantly for our attack, do not record beamforming feedback reports. Therefore, we conducted our own user study in which 197 participants walked through a WiFi field, exercising five different walking styles, while being recorded with CSI and BFI simultaneously from four different perspectives. In this section, we describe the technical recording setup, the study protocol and the recording post-processing. We also consider ethical aspects of our recording.

5.1 Setup

5.1.1 *Physical Layout.*

We record both CSI and BFI from four different perspectives each. The perspectives and walking path layout are shown in Figure 2.

Both the access points and all antennas, except the antennas for perspective 2, are mounted at (or slightly above) hip height, as done in previous work, as this is expected to result in the most gait information being captured. Perspective 3 is the most common in related work, in which the walking path is crossing the line-of-sight orthogonally. Some related works also recorded from a perspective that is parallel to the walking path, which corresponds to our perspective 1. Perspective 2 is mounted in the same location as perspective 3, but higher at a height of 210cm. This allows us to investigate the impact of the antenna height. Finally, perspective 4 is positioned such that the participant is not walking through the line-of-sight, enabling us to test non-LOS scenarios. This setup allows us to investigate the impact of different perspectives on identification accuracy which is a limitation of the existing literature.

5.1.2 *Technical Setup.*

Due to CSI and BFI requiring different traffic patterns, two separate WiFi networks on different channels are needed and therefore two access points are required. We use two TP-Link Archer BE800, as they fullfill all requirements for our experiment setup. As channels, we use channels 37 and 85, the lowest two non-overlapping 160MHz-wide channels in the 6GHz range that were introduced in WiFi 6E. This enables us to minimize inference with other WiFi devices in the vicinity. It also allows us

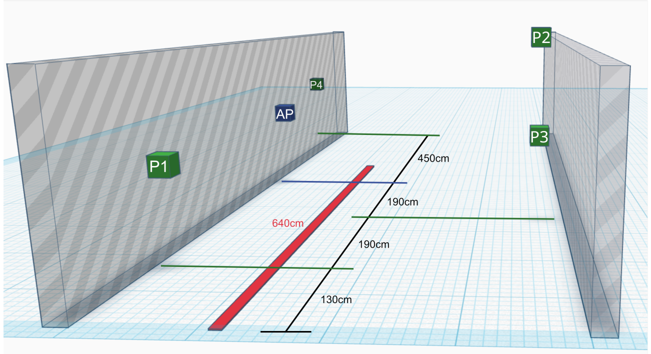

Figure 2: Positions of access points, perspectives and the walking path in our setup.

to investigate the suitability of the 6 GHz band for WiFi sensing, something that has not previously been done in related work, see Section 3.

For each perspective, a set of two antennas, one for each BFI and CSI, are connected to Intel AX210 WiFi NICs. These NICs were chosen as the single option to support both 6 GHz and CSI extraction. For each NIC, the two antennas are mounted orthogonally to each other, one pointing up and the other orthogonally to the line-of-sight to the access point.

For CSI, the access point sends traffic to a separate traffic generation node. The traffic is generated via iperf3. While the traffic generator is located beneath perspective 1 in our setup, it should be noted that its location is irrelevant, as it simply requests the traffic, but the packets that are recorded by the four perspectives are broadcasted by the access point. Each perspective passively monitors the frequencies of this access point and records the CSI data of all packets send using PicoScenes [27].

For BFI, all perspectives are connected to the WiFi network established by the access point and receive traffic from it. This traffic triggers the channel sounding procedure regularly. All beamforming reports are recorded centrally by an additional beamforming recording node which is not part of the network. Again, the exact location of this additional node is not relevant as long as they are within range since beamforming reports are broadcast by the beamformees. The traffic is once again generated using iperf3.

The exact parameters for the traffic pattern for both CSI and BFI were determined empirically by performing a parameter optimization with the goal of maximizing the sample rate. The optimization determined a bitrate of 200 Mb/s of TCP traffic for BFI and 30 Kb/s of UDP traffic for CSI[1], for each perspective. This results in average samples rates of ≈ 10 Hz for BFI and ≈ 285 Hz for CSI.

5.2 Study Protocol

In order to consider recognition in a variety of smart city scenarios, we ask participants to walk within the recording area in multiple walking styles: normal, with a backpack, carrying a bottle crate, through a turnstile and at a faster speed. The normal walking style is repeated 20 times (back and forth), the others 10 times.

To allow for an unobstructed gait recording, participants were instructed not to wear any baggy clothes, skirts, dresses or heeled shoes. We recorded 197 different individuals. Of these participants, average age was 23.2 (standard deviation 3.3) and 58.9% identified as male, 39.6% as female and the rest either answered with a custom response or refused to answer. Due to technical unreliabiltities, not all recordings resulted in usable data. For our experiments, we use 170 and 161 participants for CSI and BFI, respectively. To summarize, an overview over the parameters of our data collection can be found in Table 4.

5.3 Ethical & Open Science Considerations

The user study data collection was approved by the ethics commission of the Karlsruhe Institute of Technology (research project "Smart City Privacy") and was conducted in accordance with the Declaration of Helsinki. All data was collected in November 2024.

Table 4: Parameters of our data set

Parameter	Options
Participants	197 (CSI 170; BFI: 161)
WiFi artifacts	Beamforming Feedback Information (BFI) & Channel State Information (CSI)
Perspectives	Parallel to LOS (1), Across LOS High (2) and Low (3), & Non LOS (4)
Walking styles	normal, with backpack, carrying crate, through a turnstile, & fast

Participants were recruited from a local student panel and reminded that they could refuse to answer any question and withdraw from the study at any time. The study's information sheet explicitly stated that biometric data, which is part of the special categories of personal data, as defined by Art. 9 of the GDPR, would be recorded of them. Participation took up to one hour and participants were paid 15€.

In addition to the recording, participants could optionally agree to the sharing of their data with other scientists for research purposes, which 181 of 197 did. While an unconditional publication of all recorded data may be desirable from an Open Science perspective, considering the vastness of personal information in the data and the possibility of further, currently still unknown, inferences from the data, we consider this approach to be an advisable compromise. Researchers can request access to our dataset at **https://ps.kastel.kit.edu/bfid-dataset**.

5.4 Post-Processing

The raw CSI recordings follow a custom data format specific to the PicoScenes recording framework. However, processing these files is straightforward due to their provided Python toolbox. Our CSI recordings contain three distinct sets of CSI reports, differentiated by the number of subcarriers used for a transmission. While the bandwidth of a WiFi network defines the maximum number of subcarriers that are available for communication, the number of subcarriers actually used for a communication burst is not constant and varies. In our recordings, the vast majority (≈ 80%) of CSI reports include 53 subcarriers and ≈ 16% include 2025 subcarriers. In order to match the extant work's sampling rate for CSI, we extracted and utilized only the 53 subcarrier reports. The BFI recordings are simple traffic dumps (i.e., pcaps) which we parsed according to IEEE 802.11 standards and extract the for us relevant compressed beamforming reports.

After processing the raw CSI and BFI recordings, each atomic sample is a timeseries of variable length due to different walking speeds. BFI recordings have 740 features, as our setup yields 10 (quantized) angles for 74 channels each. CSI recordings have 212 features, that is, both phase and magnitude for 53 subcarriers for each of the two antennas. Additionally, each timeseries contains one additional features which is the time difference between two consecutive data points. In total, our dataset contains 6798 normal walking CSI sequences with an average sequence length of 1241.2 across 170 individuals and 6312 BFI normal walking sequences with an average length of 40 across 161 individuals. For the other walking styles (with backpack, carrying a crate, walking fast or

[1] Higher bandwidths are possible and result in higher sample rates, but cause CSI recording to be unstable with PicoScenes.

through a turnstile) the number of sequences is halved. As expected, the sample rate for CSI is significantly higher than for BFI (leading to longer sequences), as CSI information is calculated for each transmission whereas BFI is only requested periodically by the AP.

6 Evaluation

Using the data obtained in the user study that we described in the previous section, we now evaluate our attack. For this, we define a set of hypotheses based on our theoretical knowledge that we want to test, design matching experiments, report their results and analyze them.

6.1 BFI is identifying

We start by evaluating the main hypothesis of this paper: **H1**: Beamforming feedback information can be used to infer the identity of individuals. To be precise, we hypothesize that we can link multiple recordings of the same individual together. For this we apply the attack described in Section 4 to the BFI data that we collected as described in Section 5. We split all normal walking sequences (40 per individual) into training and test sets and trained our model on the training dataset. Afterwards, our attack is able to identify the samples in our test set with an accuracy of 99.5% ± 0.38.

For our study setup we attempted to maximize the sample rate of BFI (and CSI equivalently). To this end, we tuned the traffic pattern to maximize the frequency with which the channel sounding procedure is triggered by the access point. During this preliminary experimentation, we observed that the frequency of channel sounding procedures heavily depends on the traffic pattern. In an everyday scenario, this frequency might therefore be lower, which could decrease our attack's efficacy.

To investigate how the number of beamforming reports per sequence (and thus the traffic pattern) influences the efficacy of our attack, we continuously and uniformly remove entries from each sequence (reducing the temporal resolution) and retrain the classification model. We hypothesize that **H2**: while time-coarsening will reduce identification accuracy, a significant number of reports need to be removed for this to have a severe impact.

The results of this experiment can be found in Figure 3. While the accuracy slightly decreases as we reduce the number of beamforming reports, the effect is very small. Reducing the frequency with which channel sounding procedures are conducted could also be seen as a straightforward mitigation strategy for the threat of identification. However, as the results indicate that BFI-based identification is extremely robust to lower sample rates, a simple time-coarsening based mitigation appears to be ineffective.

6.2 Comparison of CSI and BFI

After comparing the adversary models of CSI- and BFI-based sensing in Section 4, we now compare the attack efficacy given these two information sources. For **H3**, we expect that CSI achieves slightly higher accuracies than BFI for all scenarios, for the reasons described below. In essence, beamforming reports are lossy compressed CSI at a lower temporal resolution. This means any information present in BFI is also present in CSI while the higher temporal resolution of CSI means that additional information is available which we expect to increase identification accuracy.

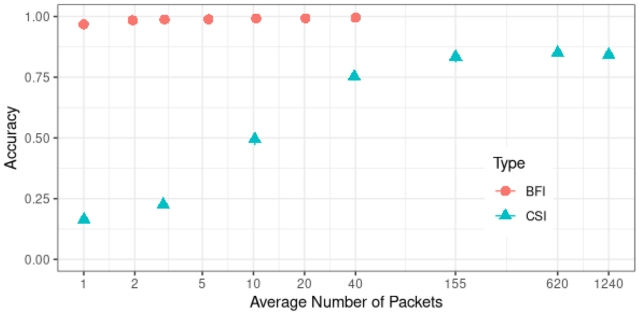

Figure 3: Test accuracy for identification when decreasing the sample rate.

We perform the same experiment for CSI as we did for BFI in the previous subsection. However, it should be noted that reducing the temporal resolution of CSI directly translates to reduced bandwidth capacity, and to the fact that less legitimate traffic can be sent: As CSI is calculated from every transmission from the access point, lowering the sample rate of CSI is only possible by lowering the transmission rate of the access point, thereby decreasing its throughput. As a mitigation strategy this would directly cause a loss of utility for the WiFi network.

The results for CSI can also be found in Figure 3. We find that we can identify individuals based on their normal walking style using CSI with high accuracy, here 82.4% ± 0.62. Related work, as seen in Table 1, often achieves higher accuracies, up to 99.7% claimed by LW-WiID. These methods are, however, generally tested on significantly smaller datasets (none that exceed 90% accuracy have more than 50 individuals) and have a CSI-specific design and pre-processing. Our approach, in contrast, is tested on the largest available CSI dataset (170 individuals) and assumes no domain knowledge by the adversary, therefore utilizing a purposely straightforward model without pre-processing. Both of these factors will decrease the measured identification accuracy. We validate this assumption by comparing our approach to the state-of-the-art in Section 6.6.2.

We see the expected trend that a lower number of CSI reports results in a decreased accuracy though the effect is more significant than for BFI. Interestingly, decreasing the sample rate slightly to an average of 620 packets actually increases the identification accuracy slightly. We consider this to be an artifact of our machine learning architecture in which the LSTM does eventually forget earlier inputs and therefore samples that are too long might be harder to identify.

Finally, regarding our hypothesis that CSI surpasses BFI, we find ourselves contradicted. In fact, BFI-based identification even performs significantly better than CSI. Possible explanations for this include that the compression of the feedback matrix into the angles that we use for our attack in BFI work as a form of pre-processing and noise removal. Also, the channel sounding procedure with one VHT-LTFs per spatial stream could allow for a higher spatial resolution in BFI than the regular LTFs that are used for CSI, which relates well to the larger amount of features (740) per point in time for BFI compared to the amount of features for CSI (212). This would imply that spatial resolution is more important than temporal resolution.

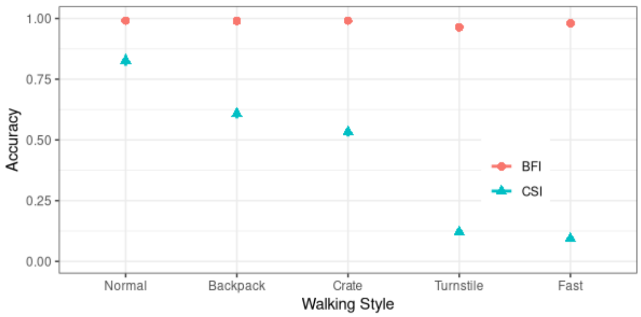

Figure 4: Test accuracy for training on the normal walking style and testing on the other walking styles.

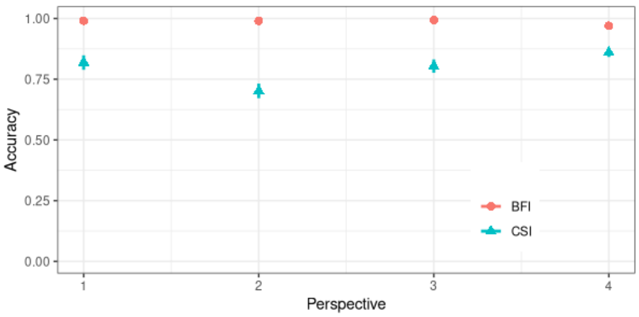

Figure 5: Test accuracy for identification for all four different perspectives.

6.3 Across walking styles

Individuals do not always exhibit the same walking style in every situation. For example, one might be in a hurry to catch a bus, wear a backpack, or carry home a crate full of beverages. This significantly alters the movement patterns and possibly the capabilities of recognition systems to identify individuals. Therefore, we also investigate the possibility to identify them across walking styles. We hypothesize for **H4** that we can still identify individuals when they walk with a backpack, a crate, or fast with both CSI and BFI, although accuracy might be decreased. This is because we expect our machine learning model to learn the gait of individuals based on the normal walking style. The other walking styles are less similar to the training data and therefore it becomes harder for our model to extract the gait information, which could lead to some misclassifications, so slightly reduced accuracy can be expected. In our experiment, we use the same models which we previously trained on the normal walking style only and measure the identification accuracy of the other walking styles. We want to emphasize that we do not retrain our models—but rather use the models already trained on normal walking to now infer identities based on recordings of other walking styles, thereby testing whether the capabilities of the system transfer across these styles.

Figure 4 reports the corresponding results. We find that for BFI, we can still reliably identify individuals for all walking styles. For CSI, we find accuracies between 50 and 60% for the modalities backpack and crate, while they are around 10% for fast and turnstile. This is insofar expected as backpack and crate are a smaller change from normal, as the movement of the legs (most relevant for gait [22]) stays the same. At the same time, fast and turnstile do also significantly change the movement of the legs and therefore a lower accuracy might be expected. We do see these same trends also for BFI where accuracies for fast and turnstile are lower than the rest, but this difference is much smaller. This shows that BFI is much more robust in its identification potential.

6.4 Different perspectives

As highlighted before, there are multiple ways how individuals can walk through the WiFi field and related work considers multiple different options, but there are few comparative insights on which perspectives are the most effective. Thus, in this subsection, we consider the different perspectives from which we recorded information and compare them against each other.

We hypothesize that **H5**: perspective's accuracy decreases as the walking path intersects less with the path between AP and perspective. The identification potential of WiFi sensing is based on the amount of biometric information—particularly gait—that is captured by the system. Signals between the AP and receiver will contain more biometric information when the individual is walking through the line-of-sight path between AP and receiver. Therefore, attack accuracy is highest when the walking path intersects directly with it and lowest when it does not. This assumption is also implicit in related work where most studies regard individuals walking orthogonally across the line-of-sight, see Table 1.

It should be noted that considering their adversary models, using multiple perspectives in CSI requires multiple malicious nodes, while multiple perspectives in BFI still only require a single malicious node. In the experiment, we always test and train with data from the same perspective and measure the identification accuracy.

The results for this experiments are shown in Figure 5. We find that all perspectives allow the reliable identification of participants. Also, we find the trend of BFI outperforming CSI form the previous experiment repeated for all perspectives. Generally, differences between perspectives are very small for BFI. The only noticeable difference to be found is slightly lower accuracy for perspective 4, which is the non line-of-sight perspective in the neighboring room. This matches our hypothesis. However, we would also expect perspective 1 to have a lower accuracy because the walking path does not intersect its line-of-sight. An explanation could be that while the impact of an individual on perspective 1 is smaller, because only some signal paths will interfere with them, this impact is along the entire walking path. In comparison, the impact on perspective 3 might be bigger, but only in one part of the walking path.

With the results for CSI, we find all perspectives with clients at hip height (1, 3, and 4) to achieve very similar accuracies. Only perspective 2 which is mounted a bigger height achieves a lower accuracy. This indicates that the height of clients is relevant and that positioning the antennas and access points such that gait can be captured well is important. Therefore, CSI also matches our hypothesis, with the same exception for perspective 1 as BFI.

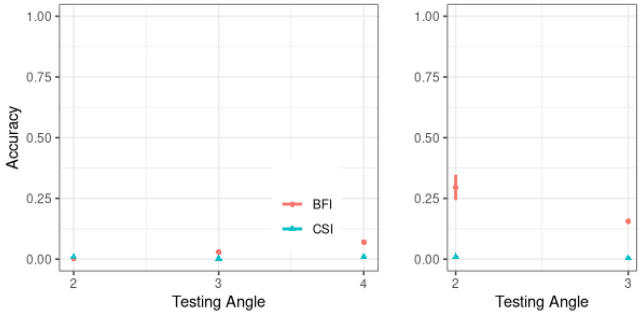

Figure 6: Test accuracy for identification for mismatched perspectives. Left, training with perspective 1, right training with perspective 3 and 2 respectively.

6.5 Across perspectives

As the adversary can easily record beamforming reports for different perspectives, and it may be that not all clients always participate in the channel sounding procedure, they may want to attempt to match recordings from one perspective with recordings from another perspective. When considering the identification across different perspectives (training the recognition system with samples from one perspective and testing with samples from another), we hypothesize **H6**: that this will not be successful. This is due to the identification being based on how individuals interfere with a constant signal from access point to a client. An individual causing a specific interference with the signal of one client does not necessarily match the interference they may have with the signal of a different client. For example, if the gait of an individual causes a specific frequency to be refracted which results in a different path of the signal between access point and client and this is used as an identifying feature, then there's no reason to assume that the same frequency will also be refracted in the same way from a different perspective, as the incoming angle will be different, and the walls that may be part of the different path will be different. As such, this identifying feature will not work across perspectives.

In our experiment, we train our model using the normal walks from perspective 1 because this perspective achieved the highest average single perspective accuracy. We then test using the recordings from the other perspectives. Additionally, because perspectives 2 and 3 are the closest, we also train on one of them and test on the other.

Our results can be found in Figure 6. We find that, as expected, identification accuracies are very low. For those cases where we trained with perspective 1, all accuracies are below 5%. This is also true for CSI when mismatching perspectives 2 and 3, but BFI is slightly higher, even surpassing 25% for the case of training with perspective 3 and testing on perspective 2. This is not surprising because 2 and 3 are the most similar and BFI surpasses CSI in all of our experiments.

6.6 Validation

We validate our results by conducting multiple further experiments: We confirm that we cannot identify empty rooms and compare our CSI-based results with related work in this subsection. We also investigate the generalizability of our model by testing alternative data splits in Appendix B.

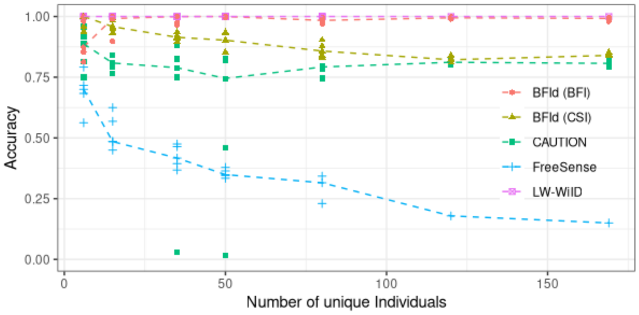

Figure 7: Test accuracy for identification for BFId (CSI and BFI), and alternative CSI-based recognition methods over differently sized subsets of our dataset. Trendlines are based on median performance over five independent subsets.

6.6.1 *BFI in empty rooms.*

As our experiments demonstrate an unexpectedly high accuracy for identifying individuals with BFI – especially compared to the accuracies for CSI – we set out to validate our results with a set of further experiments.

Before recording participants in our study, we also recorded the empty room as a baseline for 45s. To test whether our trained recognition system actually identifies the biometric information contained in the BFI recordings, we let the trained model predict participants for these empty room recordings. If the model was able to link a recording of the empty room just before the participant entered the room to the respective participant, the model would not have learned to recognize the biometric trait but some other session signal. Thus, for **H7**, we hypothesize that we cannot match these empty room recordings to the participants that were recorded before or after. This would support our claim that identification is achieved based on the biometric data that is contained in the signal propagation characteristics of WiFi and is not based on some unexpected session information. For the empty room identification, we use our models that have been trained on the normal walking style and test with the empty room recordings. We specifically use top-2-accuracy here, as most empty room recordings have participants both before and after. A high top-2-accuracy would imply that our model reliably matches empty rooms to either the participant before or after them which could be interpreted as evidence that we identify sessions rather than people.

We find that the BFI-model only reaches a top-2-accuracy of 2.34%(±0.64). This supports our hypothesis and indicates that the source of information relevant to the identification is the biometric data of the individuals walking through our experiment setup.

6.6.2 *CSI baseline.*

Because our CSI-based attack only achieves an accuracy of 82.4% ± 0.62 in comparison to related work that regularly achieves over 90%, we conduct additional experiments to validate these results. We hypothesize for **H8**, that the slightly lower accuracy for CSI is a result of our purposefully straightforward machine learning model and our larger dataset that contains more identities than any CSI dataset before. More identities trivially makes identification harder and increases the chance of misclassifications thereby decreasing accuracy. Our adversary model assumes no domain knowledge which results in a simple machine learning model without pre-processing to establish a lower bound on the privacy risk. This is uncommon for related work which generally uses purpose-built custom pre-processing and models, designed to achieve the highest possible accuracy.

To test this hypothesis, we vary our baseline CSI experiment (only normal walking, parallel to line-of-sight perspective 1) in two factors: the number of identities and the recognition model. First, we randomly sample subsets of identities from our full dataset, with the subsets being the sizes often used by related work (i.e., 6, 15, 35, 50, 80, 120). We repeat this five times for each subset size. Second, we select three CSI-based identification methods from our related work (see Table 1) and train them on the subsets: LW-WiID (state-of-the-art CNN with some pre-processing and the highest accuracy on a larger dataset) [5], CAUTION (only encoding learning CNN, no pre-processing) [52] and FreeSense (one of the first approaches with extensive pre-processing and no deep learning) [64]. More details on their implementation can be found in Appendix A.

Our results can be found in Figure 7. We find our implementations of the extant CSI-based identification methods attain their claimed accuracies (LW-WiID: 50 / 99.7 / 100.0 ± 0.0; CAUTION 15 / 88.9 / 80.3 ± 2.74; FreeSense 6 / 88.9 / 69.1 ± 8.28 (identities/claimed accuracy/our result)). Most noticeably, FreeSense, whose claims were based on a small dataset, performs significantly worse for larger datasets.[2] At the same time, LW-WiID, which claimed the highest accuracy on 50 individuals still performs very well on our full dataset, even surpassing our BFId model used with CSI data, which we attribute to its CSI-specific designed pre-processing and model. CAUTION, which also has a CSI-specific design, performs similarly to our BFId model on CSI data on the full dataset, but worse on medium sized datasets, showing particularly high variations between different subsets. We deduce that it requires more data to train its model which makes it sensitive to the exact participants that are chosen.

In conclusion, we find that our hypothesis to be supported. Even with its simplistic design, BFId on CSI matches or significantly surpasses some methods that were only tested on significantly smaller datasets. It is only surpassed by a method with CSI-specific design on the full dataset, showing the impact of its purposely simplistic design that models an adversary without domain knowledge. BFId on BFI can match even the highest accuracies of LW-WiID in defiance of its weaker adversary model and simple design.

7 Discussion

In our evaluation, we found BFI to be a highly identifying source of information that even achieves higher identification accuracy than CSI. The identification is robust across perspectives and walking styles, even with large sample sizes. We will discuss these results in this section, particularly the implications for everyone's privacy.

7.1 Implications

As outlined in the related work section, numerous studies have investigated the capabilities of WiFi sensing—primarily focusing on its utility in sensing environments, while largely overlooking the privacy risks it exposes. In fact, many of them actually claim to be privacy-preserving [12, 18, 51, 64–67, 76, 81]. While there may be legitimate use-cases, we explicitly consider identity inference via WiFi sensing a privacy attack. This view reflects the serious risks associated with the ubiquity of WiFi networks, their ability to sense through walls and in non-line-of-sight scenarios, and the fact that this would likely happen without explicit consent.

In this paper, we demonstrated the privacy threats of WiFi sensing, particularly those linked to BFI, but also the ability of CSI-based attacks to scale to large populations. We argue that future work on wireless sensing must account for the threats that their approaches introduce, and actively pursue mitigation strategies. Existing mitigation approaches for CSI-based sensing are inapplicable in real-world scenarios and are unable to protect against BFI-based sensing. Additionally, our identity inference attack also amplifies the risk and harm that existing WiFi sensing applications pose, because it allows adversaries to attach the inferred activities and other attributes to specific individuals.

Furthermore, not just other researchers should be aware of the privacy threats that BFI-based sensing poses, but also the public. We have shown robust identity inference with common-of-the-shelf hardware which is already in widespread adoption in many homes and public areas. Without possible mitigation strategies, with standardization of WiFi sensing in work and the integration of similar joint-communication-and-sensing (JCAS) approaches planned for 6G and beyond [13], we feel it necessary to communicate the associated privacy threats to the public. Particularly, the planned standardization of WiFi sensing in 802.11bf should strongly consider adding effective privacy protection, or abandon beamforming entirely.

7.2 Limitations

We acknowledge some limitations of our investigation in this paper. First, we use a straightforward machine learning model and no data pre-processing. While this allows us to set a lower bound for identification accuracy, it should be noted that further fine-tuning of the model architecture, hyperparameters, and additional pre-processing could improve the attack's efficacy. Note, that the recognition rate is already high—especially considering the size of our dataset—despite this limitation. Furthermore, as to align with prior work for comparability, we use a softmax classifier. Therefore we cannot test to which extent our model generalizes to individuals that are not part of the training data.

Second, while the number of individuals in our data set is significantly larger than any of those used in previous work, its size does not match expected sample sizes in large organizational or smart city contexts. To better investigate such scenarios, data sets with thousands of identities would be desirable, like those which are commonly used for evaluating face recognition systems. However, it is beyond common academic capabilities to collect datasets of the necessary size, and participation of industry would be required. We believe that our study, given the sample size being an order

[2] Please note that due to FreeSense's unfavorable runtime scaling with datapoints, it was not possible to run all five repetitions for larger (120+) population sizes.

of magnitude larger than what is common in the community, already highlights noteworthy threats for identity disclosure in public spaces and small to medium sized organizations.

Third, since BFI identification capabilities were still unknown, we also purposely collected the dataset under controlled conditions, for instance regarding the clothing of individuals. As these restrictions were meant to ensure that the gait of subjects could be recorded by preventing clothing from obstructing it, we may be overestimating the identification potential of BFI. At the same time, these restrictions also lead to less diversity in recordings (as clothing across participants was more similar), which makes identification more difficult. Therefore, we suggest that future work considers a data collection scenario closer to a real-world setting to investigate this.

Last, it remains unclear how exactly human gait influences beamforming reports, as we are missing a deeper understanding of the semantics of them. While we can use them to infer information using machine learning, domain knowledge based explanations of individual features within BFI are still missing. This information could potentially be leveraged to further increase recognition rates. However, we would like to point out that the recognition accuracy, even without a deeper understanding of the semantics and with our simple machine learning approach is already alarmingly high.

7.3 Future Work

We consider the most crucial area for future research to be mitigation strategies for WiFi sensing based privacy threats. The utility of WiFi in modern everyday life is immense, and therefore an effective mitigation that can reliably prevent the privacy threats discussed in this paper while not sacrificing its utility is crucial. Considering the severe privacy threat that we demonstrated in this paper and the ubiquitousness of devices that could implement this attack, a defense against it seems imperative.

Existing countermeasures against CSI-based sensing (cf. Section 3) either rely on modified firmware, or require additional hardware, and introduce performance trade-offs. Their effectiveness against BFI-based sensing remains unclear. Since BFI is derived from CSI (specifically, from VHT-LTFs), injecting noise into CSI could degrade BFI-based sensing. However, the BFI computation pipeline involves compression and other transformations that may attenuate such noise. Inaccurate BFI, in turn, risks beamforming misalignment, leading to packet loss and degraded network performance. As shown in our evaluation, reducing the time resolution of beamforming reports by modifying access points to initiate the channel sounding procedure less often, also only has very limited effect. Approaches could consider encrypting the transmitted beamforming information, such that at least the credentials to the network would have to be known to the adversary, but this would require modification of the WiFi standard and would potentially result in devices being incompatible. As such, there does not seem to be a straightforward path to mitigate the threats shown by BFId, which highlights this as an important open problem that requires community attention.

Further work should also deepen our understanding of BFI's identification capabilities through additional experiments. Open questions include the effects of dynamic or transient objects in the environment, long-term robustness, and interference from other WiFi networks.

Finally, since current insights into BFI privacy risks are largely empirical, a theoretical framework for analyzing and bounding such risks remains an interesting open problem.

8 Conclusion

In this paper, we have introduced BFId, the first identity inference attack using beamforming feedback information. We have shown it to be a robust method of identification, even across walking styles and perspectives and with large sample sizes. For this, we have created a novel WiFi sensing dataset containing BFI and CSI recordings of 197 individuals, which is available to interested researchers. We have also shown the BFI identification accuracy to surpasses that of CSI in a direct comparison, even though it uses a weaker adversary model. As BFI is transmitted unencrypted over the air, no specialized hardware with custom firmware is necessary to record it and it is easier to record multiple perspectives. This highlights the privacy threats associated with BFI-based sensing. With this hardware making its way into millions of homes, the privacy concerns are severe.

Acknowledgments

We thank Lutz Thies for piloting parts of this study and Fritz Windisch for implementing the BFI processing. This work was funded by the Topic Engineering Secure Systems of the Helmholtz Association (HGF) and supported by the KASTEL Security Research Labs, the Cluster of Excellence "Centre for Tactile Internet with Human-in-the-Loop" (EXC 2050/1 CeTI – DFG Project ID 390696704), and the Research Project SynthiClick (FKZ 16KISA108K). The user studies were conducted via the Karlsruhe Decision&Design Lab (KD^{2}Lab), an experimental lab funded by the DFG and the Karlsruhe Institute of Technology (INST_12138411-1_FUGG).

References

[1] National Aeronautics and Science Mission Directorate Space Administration. 2010. *Wave behaviors.* http://science.nasa.gov/ems/03_behaviors

[2] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. 2019. Optuna: A Next-generation Hyperparameter Optimization Framework. In *Proc. of KDD.*

[3] Apidet Booranawong, Nattha Jindapetch, and Hiroshi Saito. 2018. A System for Detection and Tracking of Human Movements Using RSSI Signals. *IEEE Sensors Journal* (2018).

[4] Jiannong Cao and Yanni Yang. 2022. *Wireless Sensing: Principles, Techniques and Applications.* Springer Intl. Publishing, Cham.

[5] Yangjie Cao, Zhiyi Zhou, Chenxi Zhu, Pengsong Duan, Xianfu Chen, and Jie Li. 2021. A Lightweight Deep Learning Algorithm for WiFi-Based Identity Recognition. *IEEE Internet of Things Journal* (2021).

[6] Cheng Chen, Hao Song, Qinghua Li, Francesca Meneghello, Francesco Restuccia, and Carlos Cordeiro. 2023. Wi-Fi Sensing Based on IEEE 802.11bf. *IEEE Communications Magazine* (2023).

[7] Siyu Chen, Hongbo Jiang, Jie Xiong, Hu Jingyang, Penghao Wang, et al. 2024. BeamCount: Indoor Crowd Counting Using Wi-Fi Beamforming Feedback Information. In *Proceedings of the Twenty-fifth Intl. Symp. on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing.*

[8] Renato Lo Cigno, Francesco Gringoli, Marco Cominelli, and Lorenzo Ghiro. 2022. Integrating CSI Sensing in Wireless Networks: Challenges to Privacy and Countermeasures. *IEEE Network* (2022).

[9] Marco Cominelli, Francesco Gringoli, and Renato Lo Cigno. 2022. AntiSense: Standard-compliant CSI obfuscation against unauthorized Wi-Fi sensing. *Computer Communications* (2022).

[10] Marco Cominelli, Felix Kosterhon, Francesco Gringoli, Renato Lo Cigno, and Arash Asadi. 2020. An Experimental Study of CSI Management to Preserve

Location Privacy. In *Proceedings of the 14th Intl. Workshop on Wireless Network Testbeds, Experimental evaluation and Characterization*.

[11] Simone Di Domenico, Mauro De Sanctis, Ernestina Cianca, and Giuseppe Bianchi. 2016. A Trained-once Crowd Counting Method Using Differential WiFi Channel State Information. In *Proceedings of the 3rd Intl. on Workshop on Physical Analytics*.

[12] Lijie Fan, Tianhong Li, Rongyao Fang, Rumen Hristov, Yuan Yuan, and Dina Katabi. 2020. Learning Longterm Representations for Person Re-Identification Using Radio Signals. In *Proc. of CVPR*.

[13] Xinran Fang, Wei Feng, Yunfei Chen, Ning Ge, and Yan Zhang. 2023. Joint Communication and Sensing Toward 6G: Models and Potential of Using MIMO. *IEEE Internet of Things Journal* (2023).

[14] Parisa Fard Moshiri, Reza Shahbazian, Mohammad Nabati, and Seyed Ali Ghorashi. 2021. A CSI-Based Human Activity Recognition Using Deep Learning. *Sensors* (2021).

[15] Takeru Fukushima, Tomoki Murakami, Hirantha Abeysekera, Shunsuke Saruwatari, and Takashi Watanabe. 2019. Evaluating Indoor Localization Performance on an IEEE 802.11ac Explicit-Feedback-Based CSI Learning System. In *2019 IEEE 89th Vehicular Technology Conf. (VTC2019-Spring)*.

[16] Matthew Gast. 2013. *802.11ac: a survival guide*. O'Reilly, Beijing.

[17] Francesco Gringoli, Matthias Schulz, Jakob Link, and Matthias Hollick. 2019. Free Your CSI: A Channel State Information Extraction Platform For Modern Wi-Fi Chipsets. In *Proceedings of the 13th Intl. Workshop on Wireless Network Testbeds, Experimental Evaluation and Characterization*.

[18] Yu Gu, Huan Yan, Mianxiong Dong, Meng Wang, Xiang Zhang, Zhi Liu, and Fuji Ren. 2021. WiONE: One-Shot Learning for Environment-Robust Device-Free User Authentication via Commodity Wi-Fi in Man–Machine System. *IEEE Transactions on Computational Social Systems* (2021).

[19] Haobin Guan, Aryan Sharma, Deepak Mishra, and Aruna Seneviratne. 2023. Experimental Accuracy Comparison for 2.4GHz and 5GHz WiFi Sensing Systems. In *ICC 2023 - IEEE Intl. Conf. on Communications*.

[20] Xiaonan Guo, Bo Liu, Cong Shi, Hongbo Liu, Yingying Chen, and Mooi Choo Chuah. 2017. WiFi-Enabled Smart Human Dynamics Monitoring. In *Proceedings of the 15th ACM Conf. on Embedded Network Sensor Systems*.

[21] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2011. Tool release: gathering 802.11n traces with channel state information. *SIGCOMM Comput. Commun. Rev.* (2011).

[22] Simon Hanisch, Evelyn Muschter, Admantini Hatzipanayioti, Shu-Chen Li, and Thorsten Strufe. 2023. Understanding Person Identification Through Gait. *Proceedings on Privacy Enhancing Technologies* (2023).

[23] Khandaker Foysal Haque, Francesca Meneghello, and Francesco Restuccia. 2023. Wi-BFI: Extracting the IEEE 802.11 Beamforming Feedback Information from Commercial Wi-Fi Devices. In *Proceedings of the 17th ACM Workshop on Wireless Network Testbeds, Experimental evaluation and Characterization*.

[24] Khandaker Foysal Haque, Milin Zhang, Francesca Meneghello, and Francesco Restuccia. 2025. BeamSense: Rethinking Wireless Sensing with MU-MIMO Wi-Fi Beamforming Feedback. *Computer Networks* (2025).

[25] Khandaker Foysal Haque, Milin Zhang, and Francesco Restuccia. 2023. SiMWiSense: Simultaneous Multi-Subject Activity Classification Through Wi-Fi Signals. In *WoWMoM*.

[26] Feng Hong, Xiang Wang, Yanni Yang, Yuan Zong, Yuliang Zhang, and Zhongwen Guo. 2016. WFID: Passive Device-free Human Identification Using WiFi Signal. In *Proceedings of the 13th Intl. Conf. on Mobile and Ubiquitous Systems: Computing, Networking and Services*.

[27] Zhiping Jiang, Tom H. Luan, Xincheng Ren, Dongtao Lv, Han Hao, Jing Wang, Kun Zhao, Wei Xi, Yueshen Xu, and Rui Li. 2022. Eliminating the Barriers: Demystifying Wi-Fi Baseband Design and Introducing the PicoScenes Wi-Fi Sensing Platform. *IEEE Internet of Things Journal* (2022).

[28] Takamochi Kanda, Takashi Sato, Hiromitsu Awano, Sota Kondo, and Koji Yamamoto. 2022. Respiratory Rate Estimation Based on WiFi Frame Capture. In *2022 IEEE 19th Annual Consumer Communications and Networking Conf. (CCNC)*.

[29] Sorachi Kato, Takeru Fukushima, Tomoki Murakami, Hirantha Abeysekera, Yusuke Iwasaki, et al. 2021. CSI2Image: Image Reconstruction From Channel State Information Using Generative Adversarial Networks. *IEEE Access* (2021).

[30] Sota Kondo, Sohei Itahara, Kota Yamashita, Koji Yamamoto, Yusuke Koda, Takayuki Nishio, and Akihito Taya. 2022. Bi-Directional Beamforming Feedback-Based Firmware-Agnostic WiFi Sensing: An Empirical Study. *IEEE Access* (2022).

[31] Belal Korany, Chitra R. Karanam, Hong Cai, and Yasamin Mostofi. 2019. XModal-ID: Using WiFi for Through-Wall Person Identification from Candidate Video Footage. In *The 25th Annual Intl. Conf. on Mobile Computing and Networking*.

[32] Bing Li, Wei Cui, Wei Wang, Le Zhang, et al. 2021. Two-Stream Convolution Augmented Transformer for Human Activity Recognition. *Proc. of AAAI* (2021).

[33] Chenning Li, Zhichao Cao, and Yunhao Liu. 2022. Deep AI Enabled Ubiquitous Wireless Sensing: A Survey. *Comput. Surveys* (2022).

[34] Chenning Li, Manni Liu, and Zhichao Cao. 2020. WiHF: Enable User Identified Gesture Recognition with WiFi. In *IEEE INFOCOM*.

[35] Xiang Li, Daqing Zhang, Qin Lv, Jie Xiong, Shengjie Li, Yue Zhang, and Hong Mei. 2017. IndoTrack: Device-Free Indoor Human Tracking with Commodity Wi-Fi. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* (2017).

[36] Chi Lin, Jiaye Hu, Yu Sun, Fenglong Ma, Lei Wang, and Guowei Wu. 2018. WiAU: An Accurate Device-Free Authentication System with ResNet. In *IEEE SECON*.

[37] Xuefeng Liu, Jiannong Cao, Shaojie Tang, Jiaqi Wen, and Peng Guo. 2016. Contactless Respiration Monitoring Via Off-the-Shelf WiFi Devices. *IEEE Transactions on Mobile Computing* (2016).

[38] Yongsen Ma, Gang Zhou, and Shuangquan Wang. 2020. WiFi Sensing with Channel State Information: A Survey. *Comput. Surveys* (2020).

[39] Francesca Meneghello, Michele Rossi, and Francesco Restuccia. 2022. DeepCSI: Rethinking Wi-Fi Radio Fingerprinting Through MU-MIMO CSI Feedback Deep Learning. In *2022 IEEE 42nd Intl. Conf. on Distributed Computing Systems (ICDCS)*.

[40] Masahiko Miyazaki, Shigemi Ishida, Akira Fukuda, Tomoki Murakami, and Shinya Otsuki. 2019. Initial Attempt on Outdoor Human Detection Using IEEE 802.11ac WLAN Signal. In *2019 IEEE Sensors Applications Symp. (SAS)*.

[41] Tomoki Murakami, Masahiko Miyazaki, Shigemi Ishida, and Akira Fukuda. 2018. Wireless LAN-Based CSI Monitoring System for Object Detection. *Electronics* (2018).

[42] Akarsh Pokkunuru, Kalvik Jakkala, Arupjyoti Bhuyan, Pu Wang, and Zhi Sun. 2018. NeuralWave: Gait-Based User Identification Through Commodity WiFi and Deep Learning. In *IECON*.

[43] Kun Qian, Chenshu Wu, Zheng Yang, Yunhao Liu, and Kyle Jamieson. 2017. Widar: Decimeter-Level Passive Tracking via Velocity Monitoring with Commodity Wi-Fi. In *Proceedings of the 18th ACM Intl. Symp. on Mobile Ad Hoc Networking and Computing*.

[44] Kun Qian, Chenshu Wu, Zheng Yang, Chaofan Yang, and Yunhao Liu. 2016. Decimeter Level Passive Tracking with Wifi. In *Proceedings of the 3rd Workshop on Hot Topics in Wireless*.

[45] Yue Qiao, Ouyang Zhang, Wenjie Zhou, Kannan Srinivasan, and Anish Arora. 2016. PhyCloak: Obfuscating Sensing from Communication Signals. In *Proceedings of the 13th Usenix Conf. on Networked Systems Design and Implementation*.

[46] Yili Ren, Yichao Wang, Sheng Tan, Yingying Chen, and Jie Yang. 2023. Person Re-identification in 3D Space: A WiFi Vision-based Approach. In *Proc. of USENIX Security*.

[47] Matthias Schulz, Daniel Wegemer, and Matthias Hollick. 2017. Nexmon: The C-based Firmware Patching Framework.

[48] Muhammad Shahzad and Shaohu Zhang. 2018. Augmenting User Identification with WiFi Based Gesture Recognition. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* (2018).

[49] Cong Shi, Jian Liu, Hongbo Liu, and Yingying Chen. 2017. Smart User Authentication through Actuation of Daily Activities Leveraging WiFi-enabled IoT. In *Proceedings of the 18th ACM Intl. Symp. on Mobile Ad Hoc Networking and Computing*.

[50] Ryo Takahashi, Shigemi Ishida, Akira Fukuda, Tomoki Murakami, and Shinya Otsuki. 2019. DNN-based Outdoor NLOS Human Detection Using IEEE 802.11ac WLAN Signal. In *2019 IEEE SENSORS*.

[51] Cristian Turetta, Florenc Demrozi, Philipp H. Kindt, Alejandro Masrur, and Graziano Pravadelli. 2022. Practical identity recognition using WiFi's Channel State Information. In *DATE*.

[52] Dazhuo Wang, Jianfei Yang, Wei Cui, Lihua Xie, and Sumei Sun. 2022. CAUTION: A Robust WiFi-Based Human Authentication System via Few-Shot Open-Set Recognition. *IEEE Internet of Things Journal* (2022).

[53] Fei Wang, Jinsong Han, Feng Lin, and Kui Ren. 2019. WiPIN: Operation-Free Passive Person Identification Using Wi-Fi Signals. In *IEEE GLOBECOM*.

[54] Ju Wang, Hongbo Jiang, Jie Xiong, Kyle Jamieson, Xiaojiang Chen, Dingyi Fang, and Binbin Xie. 2016. LiFS: Low Human-Effort, Device-Free Localization with Fine-Grained Subcarrier Information. In *Proceedings of the 22nd Annual Intl. Conf. on Mobile Computing and Networking*.

[55] Pei Wang, Bin Guo, Tong Xin, Zhu Wang, and Zhiwen Yu. 2017. TinySense: Multi-user Respiration Detection Using Wi-Fi CSI Signals. In *2017 IEEE 19th Intl. Conf. on E-Health Networking, Applications and Services (Healthcom)*.

[56] Wei Wang, Alex X. Liu, and Muhammad Shahzad. 2016. Gait recognition using wifi signals. In *Proceedings of the 2016 ACM Intl. Joint Conf. on Pervasive and Ubiquitous Computing*.

[57] Xuanzhi Wang, Junzhe Wang, Kai Niu, Jie Xiong, Fusang Zhang, et al. 2024. Wi2DMeasure: WiFi-based 2D Object Size Measurement. In *Proceedings of the 22nd ACM Conf. on Embedded Networked Sensor Systems*.

[58] Myounggyu Won, Shaohu Zhang, and Sang H. Son. 2017. WiTraffic: Low-Cost and Non-Intrusive Traffic Monitoring System Using WiFi. In *ICCCN*.

[59] Chenhao Wu, Xuan Huang, Jun Huang, and Guoliang Xing. 2023. Enabling Ubiquitous WiFi Sensing with Beamforming Reports. In *ACM SIGCOMM*.

[60] Wei Xi, Jizhong Zhao, Xiang-Yang Li, Kun Zhao, Shaojie Tang, et al. 2014. Electronic Frog Eye: Counting Crowd Using WiFi. In *IEEE INFOCOM*.

[61] Chunjing Xiao, Daojun Han, Yongsen Ma, and Zhiguang Qin. 2019. CsiGAN: Robust Channel State Information-Based Activity Recognition With GANs. *IEEE Internet of Things Journal* (2019).

[62] Rui Xiao, Xiankai Chen, Yinghui He, Jun Han, and Jinsong Han. 2025. Lend Me Your Beam: Privacy Implications of Plaintext Beamforming Feedback in WiFi. In *Proc. of NDSS*.

[63] Yaxiong Xie, Zhenjiang Li, and Mo Li. 2015. Precise Power Delay Profiling with Commodity WiFi. In *Proceedings of the 21st Annual Intl. Conf. on Mobile Computing and Networking*.

[64] Tong Xin, Bin Guo, Zhu Wang, Mingyang Li, Zhiwen Yu, and Xingshe Zhou. 2016. FreeSense: Indoor Human Identification with Wi-Fi Signals. In *2016 IEEE Global Communications Conf. (GLOBECOM)*.

[65] Jianfei Yang, Xinyan Chen, Han Zou, Chris Xiaoxuan Lu, Dazhuo Wang, Sumei Sun, and Lihua Xie. 2023. SenseFi: A library and benchmark on deep-learning-empowered WiFi human sensing. *Patterns* (2023).

[66] Jianfei Yang, Xinyan Chen, Han Zou, Dazhuo Wang, and Lihua Xie. 2023. AutoFi: Toward Automatic Wi-Fi Human Sensing via Geometric Self-Supervised Learning. *IEEE Internet of Things Journal* (2023).

[67] Jianfei Yang, Xinyan Chen, Han Zou, Dazhuo Wang, Qianwen Xu, and Lihua Xie. 2022. EfficientFi: Toward Large-Scale Lightweight WiFi Sensing via CSI Compression. *IEEE Internet of Things Journal* (2022).

[68] Jianfei Yang, Han Zou, and Lihua Xie. 2024. SecureSense: Defending Adversarial Attack for Secure Device-Free Human Activity Recognition. *IEEE Transactions on Mobile Computing* (2024).

[69] Enze Yi, Dan Wu, Jie Xiong, Fusang Zhang, Kai Niu, Wenwei Li, and Daqing Zhang. 2024. BFMSense: WiFi Sensing Using Beamforming Feedback Matrix. In *21st USENIX Symp. on Networked Systems Design and Implementation (NSDI 24)*.

[70] Moustafa Youssef and Ashok Agrawala. 2005. The Horus WLAN Location Determination System. In *Proceedings of the 3rd Intl. Conf. on Mobile Systems, Applications, and Services*.

[71] Yunze Zeng, Parth H. Pathak, and Prasant Mohapatra. 2016. WiWho: WiFi-Based Person Identification in Smart Spaces. In *2016 15th ACM/IEEE Intl. Conf. on Information Processing in Sensor Networks (IPSN)*.

[72] Jie Zhang, Zhanyong Tang, Meng Li, Dingyi Fang, Petteri Nurmi, and Zheng Wang. 2018. CrossSense: Towards Cross-Site and Large-Scale WiFi Sensing. In *Proceedings of the 24th Annual Intl. Conf. on Mobile Computing and Networking*.

[73] Jin Zhang, Bo Wei, Wen Hu, and Salil S. Kanhere. 2016. WiFi-ID: Human Identification Using WiFi Signal. In *2016 Intl. Conf. on Distributed Computing in Sensor Systems (DCOSS)*.

[74] Jin Zhang, Bo Wei, Fuxiang Wu, Limeng Dong, Wen Hu, Salil S. Kanhere, Chengwen Luo, Shui Yu, and Jun Cheng. 2021. Gate-ID: WiFi-Based Human Identification Irrespective of Walking Directions in Smart Home. *IEEE Internet of Things Journal* (2021).

[75] Yi Zhang, Yue Zheng, Kun Qian, Guidong Zhang, Yunhao Liu, Chenshu Wu, and Zheng Yang. 2021. Widar3.0: Zero-Effort Cross-Domain Gesture Recognition with Wi-Fi. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021).

[76] Yi Zhang, Yue Zheng, Guidong Zhang, Kun Qian, Chen Qian, and Zheng Yang. 2022. GaitSense: Towards Ubiquitous Gait-Based Human Identification with Wi-Fi. *ACM Transactions on Sensor Networks* (2022).

[77] Zhiwei Zhao, Zifei Zhao, Geyong Min, Chang Shu, Zhe Wang, and Hancong Duan. 2018. Non-Intrusive Biometric Identification for Personalized Computing Using Wireless Big Data. In *2018 IEEE SmartWorld*.

[78] Rui Zhou, Xiang Lu, Pengbiao Zhao, and Jiesong Chen. 2017. Device-Free Presence Detection and Localization With SVM and CSI Fingerprinting. *IEEE Sensors Journal* (2017).

[79] Xiuyan Zhu and Yuan Feng. 2013. RSSI-based Algorithm for Indoor Localization. *Communications and Network* (2013).

[80] Han Zou, Jianfei Yang, Yuxun Zhou, Lihua Xie, and Costas J. Spanos. 2018. Robust WiFi-Enabled Device-Free Gesture Recognition via Unsupervised Adversarial Domain Adaptation. In *2018 27th Intl. Conf. on Computer Communication and Networks (ICCCN)*.

[81] Han Zou, Yuxun Zhou, Jianfei Yang, Weixi Gu, Lihua Xie, and Costas Spanos. 2018. WiFi-Based Human Identification via Convex Tensor Shapelet Learning. *Proc. of AAAI* (2018).

[82] Yongpan Zou, Yuxi Wang, Shufeng Ye, Kaishun Wu, and Lionel M. Ni. 2017. TagFree: Passive Object Differentiation via Physical Layer Radiometric Signatures. In *2017 IEEE Intl. Conf. on Pervasive Computing and Communications (PerCom)*.

A CSI baseline implementation details

Some of the implemented CSI-based recognition approaches required minor modifications in order to be usable for our experiments, which we want to acknowledge here. Generally, we performed a hyperparameter optimization for any options of the approaches, as they may be specific to our dataset.

For LW-WiID, the size of the sliding window and the step size during pre-processing were also determined using hyperparameter optimization. As the Baloon mechanism used within the model design primarily has the goal of decreasing the size of the model while preserving accuracy, we did not include it in our implementation. Instead, we used the CNN architecture from WiAU [36]. LW-WiID compares its architecture to it during their evaluation and they find insignificant accuracy changes, but a larger model – which is not relevant for our experiments.

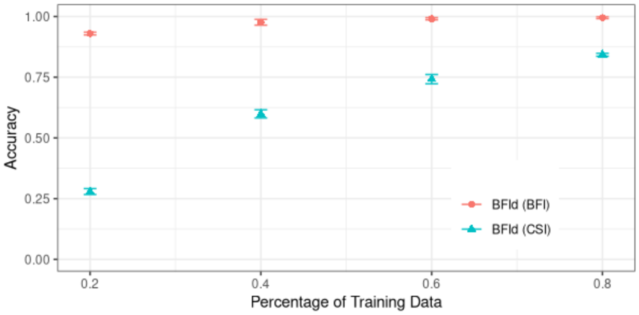

Figure 8: Test accuracy for identification for BFId (CSI and BFI) with varying data splits.

As CAUTION does not support variable length samples, but requires every CSI recording to have the same number of frames, we cut or pad them zeros. The fixed length is determined through hyperparameter optimization.

While FreeSense generally supports variable length samples (since DTW does), we found accuracy significantly improved when cutting our samples to a common size, similar to CAUTION. Due to FreeSense being significantly less efficient than the other deep learning approaches due to using a k-NN with custom (complex) distance function, we had to perform its hyperparameter optimization on a random, but fixed subset of our dataset with 10 individuals.

B Impact of training data split

To further investigate how well our model generalizes, we test the impact of the split between training and test data. We hypothesize that less training data decreases the model's accuracy though a identification remains generally possible. This is because less training data means that it becomes more difficult to learn the representative features from the data that can be used for identification. At the same time, we have already seen BFI and CSI to be very distinctive, so even if the extracted features are less optimal, identification will often still be possible.

To test this hypothesis, we run our baseline experiment for both CSI and BFI (normal walking, perspective 1) with varying data splits. In our other experiments, we use the common 80:20 split between training and test data. Here, we test multiples of 20 as values for this split. As before, we repeat each experiment five times.

The results can be found in Figure 8. As expected, the identification decreases with lower amounts of training data. While it remains over 90% for BFI even with only 20% of the data for training, it decreases to as low as 28% for CSI. This difference between WiFi artifacts could be attributed to CSI being significantly higher dimensional. This means the data is more complex and it is more difficult to extract the relevant features for identification. This also matches with the theory that BFI's inherent compression already extracts relevant features compared to raw CSI. Therefore, more data is needed to train a system for CSI than for BFI.