Poison to Cure: Privacy-preserving Wi-Fi Multi-User Sensing via Data Poisoning

Jingzhi Hu* Xin Li* Jin Gan Jun Luo
College of Computing and Data Science, Nanyang Technological University (NTU), Singapore
Email: jingzhi.hu518@gmail.com, {l.xin, jin010, junluo}@ntu.edu.sg

ABSTRACT

Wi-Fi human sensing, boosted by latest progress in both system innovation and deep analytics, has demonstrated ever-increasing resolution of users' activities. Nonetheless, it may become a spy on users' private activities such as password entry or intimate social interactions. Existing countermeasures include signal obfuscation and adversarial perturbations to hamper and confuse Wi-Fi sensing, yet they both require substantial changes in Wi-Fi hardware/firmware, and they at most stay at *user level* in protection granularity. This paper presents *Poison2Cure*, the first *semantic-level* privacy-preserving framework for Wi-Fi human sensing systems, with full compatibility to any underlying hardware. The innovation behind Poison2Cure lies in feeding poisoned training data from (privacy-sensitive) users to the neural model for Wi-Fi sensing, degrading only the sensing for *private* activities while retaining that for *regular* ones. Moreover, we tackle the harsh conditions where the neural model is kept confidential and/or preceded by data cleansing. Our extensive evaluations demonstrate that Poison2Cure reduces over 76% of the accuracy for the private activities while keeping the accuracy for regular activities largely intact.

CCS CONCEPTS

• **Security and privacy → Privacy protections**; *Mobile and wireless security*; • **Computing methodologies → Machine learning**.

KEYWORDS

Wi-Fi multi-user sensing, human activity recognition, privacy protection, poisoning attack, domain adaptation.

ACM Reference Format:

J. Hu, X. Li, J. Gan, and J. Luo. 2025. Poison to Cure: Privacy-preserving Wi-Fi Multi-User Sensing via Data Poisoning. In *The 31st Annual International Conference on Mobile Computing and Networking (ACM MobiCom'25), November 4–8, 2025, Hong Kong, China.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3680207.3723459

* Both authors contributed equally to this research.

ACM MobiCom'25, November 4–8, 2025, Hong Kong, China

ACM ISBN 979-8-4007-1129-9/2025/11.
https://doi.org/10.1145/3680207.3723459

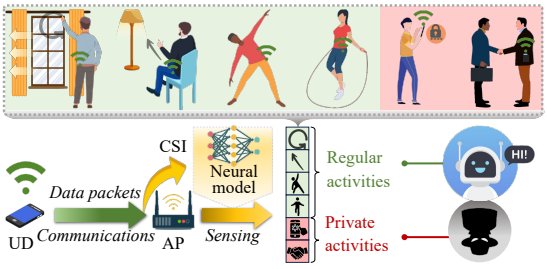


Figure 1: Wi-Fi human sensing services, albeit being desirable, may cause user privacy breaches.

1 INTRODUCTION

Wi-Fi-based human sensing techniques have been actively studied for more than a decade in diverse applications [64], including vital sign monitoring [44, 79], gesture detection [66, 88], and activity recognition [16, 33]. In distinct contrast to sensing techniques based on cameras [60, 85], radars [2, 4], lidars [47, 59], acoustic sensors [6, 56], and wearable add-ons [21, 58], Wi-Fi human sensing possesses the unique advantage of piggybacking on Wi-Fi infrastructure that is already extensively deployed. Recent progress in system innovation has fully enabled *multi-user sensing* leveraging only the channel state information (CSI) carried in normal data traffic from user devices (UD) to an access point (AP) [29, 42]. With this inherent multi-user sensing capability, manufacturers can readily endow their APs with *context-aware edge intelligence* that supports applications such as cognitive communications [8, 26] and augmented/virtual reality [18, 35], thus making UDs both lighter and more efficient.

While being a fascinating prospect, transforming every AP into an intelligent sensor may disturb today's privacy-conscious users. Although Wi-Fi human sensing records no portraits or voices of the users, it can still cause serious privacy breaches due to the ubiquity of Wi-Fi signals. As illustrated in Figure 1, with a trained neural model processing a user's CSI, an AP is able to recognize sensitive activities of the user, such as its key information entry [40, 74] and intimate social interaction [27, 78]. In this regard, Wi-Fi human

sensing is controversial: users desire edge intelligence that responds to gestures or assists health monitoring, yet they are concerned about their sensitive activities being spied on. This concern is extremely hard to eliminate due to the ubiquity of wireless signals, which allows Wi-Fi human sensing to bypass common obstructions and extend beyond the range limits of cameras and microphones.

To prevent Wi-Fi sensing from compromising users' privacy, existing methods involve either wireless transceivers obfuscating CSI to hide activity-related information [10, 12, 45, 57, 63] or encrypting CSI to prevent unauthorized access [11, 25, 48, 52]. Such methods can be deemed as a switch to turn off the sensing potentials of CSI *completely* for a user. Nevertheless, to simultaneously allow for regular sensing services while disabling sensitive ones, privacy-preserving Wi-Fi sensing should be realized at a *semantic level*, i.e., users should be able to determine which activities can or cannot be sensed; we refer to the former as *regular* activities and the latter as *private* ones. One feasible approach is injecting *adversarial perturbations* to the CSI at Wi-Fi transmitters [30] or receivers [84, 89], which particularly prevents neural models from recognizing private activities. However, perturbing CSI at the physical layer requires substantial changes to Wi-Fi hardware/firmware, making them incompatible with prevalent infrastructure.

Aiming to close the research gap, this paper presents *Poison2Cure*, the first semantic-level privacy-preserving framework for Wi-Fi human sensing with full compatibility to any Wi-Fi hardware. The innovation behind Poison2Cure lies in users feeding *poisoned* training data to poison the neural model during its calibration (also known as *fine-tuning*) process. Such a calibration process is often necessary because of the much lower spatial resolution of Wi-Fi sensing compared to that of camera-driven computer vision techniques [53], which necessitates user-provided training data for neural models to learn user-specific mapping between activities and CSI variations. As the poisoning targets at this calibration process, Poison2Cure has the desirable characteristic that it requires no real-time manipulations of physical CSI and only needs to be performed once, yet achieving a lasting effect. This characteristic makes it fully compatible with all prevalent Wi-Fi standards and hardware.

To undermine the sensing accuracy for only the private activities, the poisons injected into the training data need to be carefully crafted. Nevertheless, how a user can craft the poisons poses a major challenge for Poison2Cure, because the relationship between the poisoned training data and the accuracy of the calibrated neural model is highly intricate and hard to model, especially by users with generally limited computational resources. Furthermore, poisoning the training data may encounter harsher conditions. For example, the details of the AP's neural model may be held confidential to the users, and/or data cleansing methods may be used to remove abnormal CSI data and filter out noise and jittering. Such conditions exacerbate the complications for users to craft effective poisons.

To tackle the above challenges, we design an efficient CSI poisoning method for realizing Poison2Cure in practice. Instead of modeling the neural model's post-calibration performance, we convert the problem into aligning the gradient of the neural model's parameters towards minimizing the user's semantic-level privacy preservation loss, which entails a much more efficient solution. We then extend the efficacy of our method to harsher conditions, handling a confidential neural model by substituting it with an ensemble of surrogates, while using random dropout and stochastic model batching techniques to enhance the generalizability of crafted poisons. Besides, we treat the data cleansing as power and doppler frequency constraints and enforce their compliance via proximal projection and spectral cutoff. Our key contributions are summarized as follows:

- We propose Poison2Cure, the first privacy-preserving framework for Wi-Fi human sensing, fully compatible with any underlying hardware; it preserves privacy at the semantic level by reducing sensing accuracy only for private activities.
- We propose an efficient CSI poisoning method to craft poisoned training data; then we extend it to countering confidential neural models and CSI data cleansing.
- We evaluate Poison2Cure with extensive experiments on multiple users and environments, confirming that it reduces the accuracy for private activities by over 76% while maintaining high accuracy for regular activities.

The rest of the paper is organized as follows. Section 2 introduces the preliminary of multi-user Wi-Fi sensing systems and experimentally motivates us to leverage the calibration process. Section 3 presents the design of Poison2Cure, including the efficient CSI poisoning method and its extensions to harsher conditions. Section 4 specifies the implementation of Poison2Cure and its evaluation setup. Section 5 reports the experimental results. Section 6 discusses the practicability and susceptibility of Poison2Cure. Related works are briefly captured in Section 7. Finally, Section 8 concludes this paper.

2 PRELIMINARY AND MOTIVATION

We first introduce the preliminary of multi-user Wi-Fi sensing and demonstrate its *user*-specific nature. We then describe potential methods for enabling cross-user sensing and experimentally compare their performance, demonstrating the necessity of user calibration. In addition, we examine existing privacy-preserving approaches and demonstrate their inefficiency and incompatibility.

2.1 Multi-User Wi-Fi Sensing

We start with a general scenario shown in Figure 1, where N users, each with an UD near him/her, connect to a Wi-Fi network held by an AP. Receiving data packets from the UD of a user, the AP captures the Long Training Sequence in the packets' preambles [1] and obtains CSI for the wireless channel between it and the UD. Then, with the edge intelligence enabled by a neural model, the AP infers the gestures and activities of the user by processing the CSI.

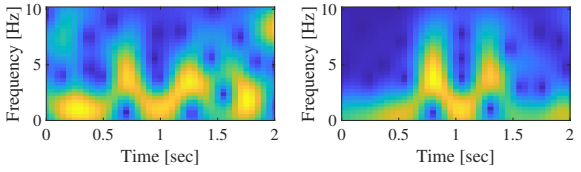
For the n-th UD in the proximity of the n-th user, the CSI of a certain frequency at time t can be expressed as:

$$h(t) = h_{\mathrm{S}} + h_{\mathrm{D}}(t) + h_n(t) + \textstyle\sum_{n' \neq n}^{N} h_{n'}(t), \tag{1}$$

where h_{S} represents the collective channel gain of static environment reflection, scattering, and direct line-of-sight paths, $h_{\mathrm{D}}(t)$ is the dynamic channel gain due to surrounding movements and hardware fluctuations, and $h_n(t)$ and $h_{n'}(t)$ respectively denote the channel gains corresponding to the paths via the n-th and n'-th users.

Based on Eqn. (1), motions of the n-th user cause variations of $h_n(t)$, which are the "driving force" behind sensing his/her gestures and activities with $h(t)$. Although the variations of $h_{\mathrm{D}}(t)$ also change $h(t)$, the randomness and generally low power of $h_{\mathrm{D}}(t)$ barely affect the perception of $h_n(t)$ [19]. Moreover, the variations of $h_{n'}(t)$ is much smaller than that of $h_n(t)$ owing to the n-th UD being much closer to the n-th user than to the others [29]. As a result, the variations of $h_n(t)$ dominate the variations of $h(t)$, indicating that the motions of each user can be sensed separately.

Moreover, the domination of $h_n(t)$ also implies that **multi-user Wi-Fi sensing is not susceptible to environmental interference but highly sensitive to respective users.** Figure 2 experimentally demonstrates this, showing the CSI variations for drawing circle of different users measured in two environments. As expected, environmental differences



(a) Meeting room, User A. (b) Lecture room, User A.

(c) Meeting room, User B. (d) Lecture room, User B.

Figure 2: Spectrograms of CSI variations for drawing circle of different users in two environments.

do not significantly alter the pattern of CSI variations, yet user differences have a more prominent impact, which may be owing to the user specifics, e.g., their different body contours and habitual behaviors. Consequently, multi-user Wi-Fi sensing has a *user-specific* nature, which demands the neural model to have *cross-user* sensing capability.

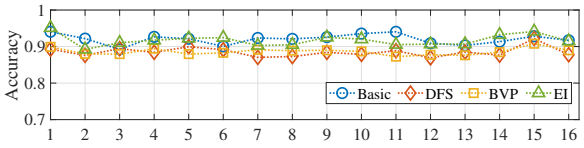
2.2 Methods for Cross-User Sensing

Wi-Fi sensing based on commodity devices is generally deemed highly domain-specific due to the lack of spatial resolution [7]. Our analysis in Section 2.1 indicates that a *domain* can be reduced from a complete surrounding environment to a *user*; yet, can a neural model be endowed with the cross-*user* sensing capability? To answer it, we start with analyzing existing approaches for cross-domain Wi-Fi sensing, which are divided into two categories below.

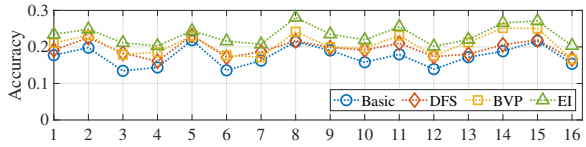
2.2.1 Domain Independent Representation (DIR). The first category of approaches does not require calibration for a target domain, as they leverage DIR of CSI as the input of the neural model to allow cross-domain sensing. The most commonly adopted DIR is the spectrogram of Doppler frequency shift (DFS) of CSI variations [19, 29, 37, 51, 88], which decomposes the CSI variations into time-frequency components to disentangle the influence from environment and user's activities. In addition, for gesture recognition, the body velocity profile (BVP) [88] and other similar methods [19, 54] are proposed to estimate the velocities and directions of body movements, which are derived from the DFS at multiple Wi-Fi receivers. Moreover, with the help of adversarial learning techniques, the environment-independent (EI) feature extraction in [32] and other similar studies [41, 46, 67] propose to learn a DIR extractor automatically.

However, our practical experiment results below show that these DIR approaches fail to enable multi-user Wi-Fi sensing for cross-user sensing. In the experiment, we collect a dataset comprising 20 hours of labeled sequences of CSI for ten gesture and body activities of 16 users (see Section 4.2 for details). Leveraging the dataset and the neural model specified in Section 4.1, we test four representation methods of CSI, including the basic method (using amplitudes and phases of channel gains) and three DIR methods, including DFS spectrogram, BVP, and EI.

In particular, the CSI variations represented by the basic and DFS methods are respectively handled by the GRU-based and CNN-based neural models in Section 4.1, while for BVP and EI, they are handled by the models implemented according to [88] and [32], respectively. During training the neural model, the dataset is split at a 9:1 ratio for training and validation, and the test results are evaluated on the validation set. Additionally, we adopt the "leave-one-out" strategy [69] to study their cross-user capability: an test user is selected



(a) Accuracy for training users in different cases of left-out test users.



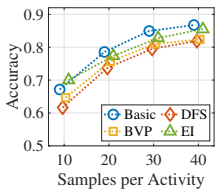
(b) Accuracy for different left-out test users.

Figure 3: Comparison of the DIR methods' sensing accuracy for training users and test users.

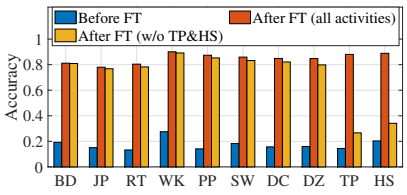
while the rest of the users are for the training. In evaluation, we focus on *accuracy*, i.e., the proportion of CSI samples whose corresponding activities are correctly classified relative to the total number of samples. Figures 3(a) and 3(b) show that, though all the methods achieve high accuracy for training users, their accuracy for test users is much lower, showing the inefficacy of DIR methods.

The reason for DIRs' inefficacy is twofold. On one hand, the proximity between a user and its UD makes the CSI variations of activities highly dependent on the nuances of the user's body figure and behaviors. Such dependency is challenging for the DIR methods since they assume a similar relationship between activities and CSI variations across domains. On the other hand, the CSI to sense a user is obtained from a single AP-UD link with normal packet rates, as opposed to multiple links with near-saturated traffic in [32, 88]. These distinctions in physical complexity and CSI sufficiency significantly hinder the effectiveness of DIRs.

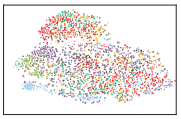
2.2.2 Fine-Tuning (FT). In contrast to DIR-based approaches, the FT-based approach [81, 83] use a small set of labeled data collected in the target domain, referred to as an *FT dataset*, to calibrate the neural model for a few epochs with a small learning rate, adapting the neural model to the specifics of the target domain. By conducting the above calibration process for each user, the FT approach effectively enables cross-user sensing capabilities. To demonstrate their efficacy, we let each test user provide an FT dataset of 10 to 40 samples per activity and continue with the experiments in Section 2.2.1. Figure 4(a) shows that the calibration in FT substantially improves the accuracy for test users of all the methods: With 30 samples per activity, the average accuracy raises from 0.20 to 0.82. Nevertheless, the FT approach has a disconcerting potential to raise the accuracy even for the activities whose samples are removed from the FT dataset due to being deemed private. This is demonstrated in Figure 4(b), where using the FT dataset without TP and HS samples still leads to an increase in their average accuracy, from 0.17 to 0.30.[1]



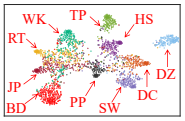
(a) Impact of FT datasets.



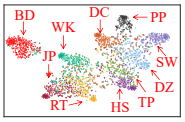
(b) Accuracy vs. activity.



(c) tSNE w/o FT.



(d) FT (all activities).



(e) FT (w/o TP&HS)

Figure 4: Efficacy of FT. (b) to (e) are obtained in the case of basic method and 30 samples per activity.

This phenomenon can be analyzed via the t-distributed Stochastic Neighbor Embedding (tSNE) [50] of extracted features shown in Figures 4(c)–4(e). Here, the extracted features are the latent embeddings at the layer preceding the final classification layer of the neural model trained in the basic method case. Comparing Figures 4(c) with 4(d), one can observe that the FT enhances the boundary clarity of feature clusters of the activities. When no TP or HS samples are used, the FT on other samples still allows the neural model to learn the influence of user specifics on the CSI-activity relationship. Owing to the generalizability of such influence, learning it helps the neural model recognize TP and HS, leading to more concentrated clusters of TP and HS in Figure 4(e) than those in Figure 4(c). Meanwhile, the fact that the clusters of the other activities have clearer boundaries also contributes to more accurate recognition of TP and HS.

Although the average accuracy of 0.30 for TP and HS is relatively low, it raises a non-neglectable privacy risk to users. In particular, given that an AP is deployed for the long-term, a relatively-low accuracy still enables the AP to gradually accumulate CSI data of users' private activities, from which sensitive information of users can be derived. Thus, to eliminate the privacy risk, users have strong motivation to pursue that the AP has *zero* sensing accuracy for private activities.

Remarks: With the above results, we acquire two key motivations: i) An FT process of user calibration is necessary for enabling cross-user capability in multi-user Wi-Fi sensing; ii) Removing private activity samples from the FT dataset is inadequate for preserving user privacy as regular activity samples help neural model recognize private activities.

[1] In Figure 4, the acronyms are bending (BD), jumping (JP), rotating (RT), walking (WK), push&pull (PP), sweeping (SW), drawing circle (DC), drawing zigzag (DZ), typing-on-phone (TP), and hand-shaking (HS).

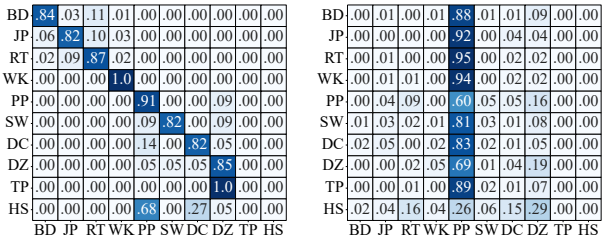
Additionally, we remark that the FT dataset needs to be provided by the user. Although the AP can also measure CSI given channel reciprocity, it cannot establish a labeled CSI dataset without the user specifying the start and end time of activities and providing accurate activity labels.

2.3 Existing Privacy Preserving Approaches

To achieve cross-user sensing while preserving user privacy, existing studies resort to *adversarial perturbations* [36]. Specifically, they consider modifying either Wi-Fi transmitters [30] or receivers [72, 84, 89] to apply adversarial perturbations to input CSI of neural models. Nevertheless, they face the following difficulties in practice.

Firstly, applying adversarial perturbations to real-time CSI is incompatible with prevalent Wi-Fi infrastructure. Although accessing CSI from the physical layer is widely supported [23, 34], modifying it requires a systematic revamp to the communication protocol and underlying firmware, which is prohibitively expensive. Injecting adversarial perturbations at the application layer rather than the physical layer may seem to be a plausible option, yet it fails to thwart malicious AP manufacturers. Because AP manufacturers can leave a backdoor to bypass the perturbations at the application layer by directly obtaining clean CSI from the physical layer. A few adversarial schemes leveraging third-party devices to attack Wi-Fi sensing may be exploited for privacy preservation, which adversarially perturb CSI through strategical spectrum jamming [43], multipath manipulation [90], and medium access competition [31]. However, such schemes inevitably interfere with normal communications and may lead to security issues owing to their attack nature.

Secondly, due to the user-specific nature of multi-user Wi-Fi sensing, adversarial perturbations face challenges in achieving cross-user efficacy as well. Thus, applying the perturbations designed for training users may be ineffective for test users, even resulting in complete destruction of sensing abilities. To demonstrate this, continuing with the experiment in Section 2.2.2, we arbitrarily choose 15 users as the training group, design adversarial perturbations for them using [36], and test the perturbations on the remaining user. Figure 5(a) verifies that, for the training group, the adversarial perturbations work as expected, causing the neural model to misclassify only the private activities. However, upon applying the adversarial perturbations to the CSI data of the test user, the neural model misclassifies all the activities as PP, as shown in Figure 5(b). This anomalous outcome is probably because PP is the default class to which the neural model categorizes all unrecognized activities.



(a) For a user in the training group. (b) For the test user.

Figure 5: Confusion matrices for activities of (a) an user in the training group and (b) the test user, given CSI being adversarially perturbed.

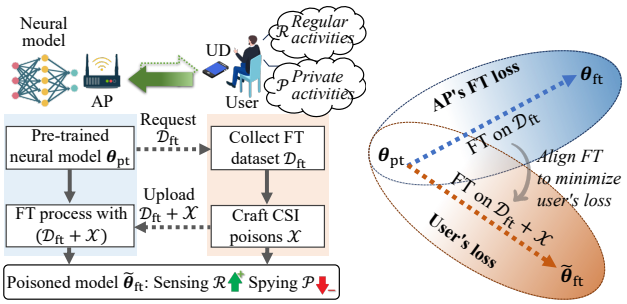
3 DESIGN OF POISON2CURE

Aiming to preserve users' private activities in multi-user Wi-Fi sensing systems, we hereby propose Poison2Cure, the first framework that achieves semantic-level privacy preservation with full compatibility to any underlying Wi-Fi hardware. Leveraging the calibration process necessary for cross-user sensing, Poison2Cure allows users to poison their FT datasets, causing neural models to misclassify user-defined private activities. Below, we start with a theoretical formulation of poisoning the FT process as an optimization problem. Then, we design an efficient CSI poisoning method to handle the problem and extend it to harsher conditions.

3.1 Poisoning the Fine-Tuning

Following the scenario in Section 2.1, Poison2Cure is based on a Wi-Fi network set up by an AP serving multiple users. Due to the symmetry statuses of the users, we focus on an arbitrary user in the following. We assume the user intends to keep a set $\mathcal{P}$ of P activities private while requiring Wi-Fi sensing services for another set $\mathcal{R}$ of R regular activities. The AP, however, does not cooperate in preserving private activities and uses a neural model to recognize all the activities based on CSI data. Without loss of generality, the neural model is represented by $\mathbf{g}(\cdot|\boldsymbol{\theta})$ with parameters $\boldsymbol{\theta}$, mapping a CSI data $\boldsymbol{C} = (\boldsymbol{c}_1, \cdots, \boldsymbol{c}_L) \in \mathbb{R}^{L\times T}$ to an activity $a \in \mathcal{R} \cup \mathcal{P}$. Here, $\boldsymbol{c}_\ell \in \mathbb{R}^T$ $(\ell = 1, \cdots, L)$ denotes the sequence of the ℓ-th CSI feature at time $t_1, \cdots, t_T$, which are associated with the random packet arrivals and thus may be non-equally spaced in time. In addition, the neural model has been pre-trained by the AP manufacturer, resulting in parameters $\boldsymbol{\theta}_{\text{pt}}$.

To achieve cross-user sensing, the AP calibrates its neural model by FT, requiring the user to provide a small FT dataset of M labeled CSI data. Using this FT dataset, the AP trains the neural model and calibrate its parameters to $\boldsymbol{\theta}_{\text{ft}}$. Due to the user's privacy concern, it only provides labeled CSI data for activities in $\mathcal{R}$. The FT dataset can be expressed as $\mathcal{D}_{\text{ft}} = \{(\boldsymbol{C}_k, a_k)|a_k \in \mathcal{R}\}_{k=1}^{M}$, where a_k is the corresponding activity label of $\boldsymbol{C}_k$. However, as demonstrated in Section 2.2.2, FT on $\mathcal{D}_{\text{ft}}$ enhances the sensing accuracy for the private activities unintendedly. For the purpose of privacy



(a) Workflow of Poison2Cure. (b) Effect of poisoning.

Figure 6: Illustrations of (a) Poison2Cure's workflow and (b) the effect of poisoning the FT dataset.

preservation, Poison2Cure enables the user to inject *poisons* into the FT dataset, resulting in a poisoned FT dataset:

$$\tilde{\mathcal{D}}_{\text{ft}} = \{(\boldsymbol{C}_k + \boldsymbol{X}_k, a_k) | (\boldsymbol{C}_k, a_k) \in \mathcal{D}_{\text{ft}}\}, \tag{2}$$

where $\boldsymbol{X}_k$ denotes the *CSI poison* added to CSI data $\boldsymbol{C}_k$. The key of Poison2Cure is to craft the set of CSI poisons $\mathcal{X} = \{\boldsymbol{X}_k\}$ so that FT on $\tilde{\mathcal{D}}_{\text{ft}} = \mathcal{D}_{\text{ft}} + \mathcal{X}$ will shift the parameters to $\tilde{\boldsymbol{\theta}}_{\text{ft}}$ instead of $\boldsymbol{\theta}_{\text{ft}}$. The parameters $\tilde{\boldsymbol{\theta}}_{\text{ft}}$ should preserve the user's privacy at a semantic level, solely diminishing the neural model's ability in recognizing activities in $\mathcal{P}$. The workflow of Poison2Cure is summarized as Figure 6(a).

Accordingly, the poison crafting of Poison2Cure can be formulated as the optimization problem below:

$$\text{(P1)}: \min_{\mathcal{X}} L_{\text{u}}(\tilde{\boldsymbol{\theta}}_{\text{ft}}) = \mathop{\mathbb{E}}_{(\boldsymbol{C},a)\sim\Gamma}\Big[L\big(\mathbf{g}(\boldsymbol{C}|\tilde{\boldsymbol{\theta}}_{\text{ft}}), a\big)\cdot\big(I_{\mathcal{R}}(a) - I_{\mathcal{P}}(a)\big)\Big],$$

where the expectation is taken over the joint distribution Γ of CSI data and activities of the user, $L(\cdot)$ denotes the cross-entropy loss function, and $I_{\mathcal{R}}(a)$, $I_{\mathcal{P}}(a)$ denote the indicator functions, which equal one if a in set $\mathcal{R}$ and $\mathcal{P}$, respectively; otherwise, zero. We refer to the loss function $L_{\text{u}}(\tilde{\boldsymbol{\theta}}_{\text{ft}})$ in (P1) as the *user's loss*.

Nevertheless, solving (P1) faces the following challenges: *Firstly*, due to the massive number of parameters and their iterative update during the FT, the relationship between $L_{\text{u}}(\tilde{\boldsymbol{\theta}}_{\text{ft}})$ and $\mathcal{X}$ is extremely complex. *Secondly*, considering the limited computational resources of the user, it is also essential to ensure efficiency in crafting the poisons, resulting in even more difficulties in solving (P1). Moreover, harsher conditions such as a confidential neural model and data cleansing of the AP may further hinder the poison crafting.

3.2 Efficient CSI Poisoning

We first design an efficient CSI poisoning method handling the first two challenges, given that the parameters of the neural model are known by the user. Consider the iterative parameter update in the FT, which can be expressed as:

$$\tilde{\boldsymbol{\theta}}^{(j+1)} = \tilde{\boldsymbol{\theta}}^{(j)} - \eta\nabla_{\boldsymbol{\theta}} L_{\text{ft}}(\tilde{\boldsymbol{\theta}}^{(j)}). \tag{3}$$

Here, $j = 1, \cdots, J$ is the index of iteration with $\tilde{\boldsymbol{\theta}}^{(0)} = \boldsymbol{\theta}_{\text{pt}}$ and $\tilde{\boldsymbol{\theta}}_{\text{ft}} = \tilde{\boldsymbol{\theta}}^{(J)}$, η is the learning rate in the FT, and $L_{\text{ft}}(\tilde{\boldsymbol{\theta}}^{(j)})$ represents the AP's *FT loss* given parameters $\tilde{\boldsymbol{\theta}}^{(j)}$, which can be expressed as

$$L_{\text{ft}}(\tilde{\boldsymbol{\theta}}^{(j)}) = \sum_{(\boldsymbol{C}+\boldsymbol{X},a)\in\tilde{\mathcal{D}}_{\text{ft}}} L\big(\mathbf{g}(\boldsymbol{C}+\boldsymbol{X}|\tilde{\boldsymbol{\theta}}^{(j)}), a\big). \tag{4}$$

We note that the AP's FT loss in Eqn. (4) differs from the user's loss defined in (P1) since the AP aims to reduce the average recognition loss while ignoring the user's privacy concerns.

Even though the number of iterations J is typically on the order of tens (to avoid overfitting), it can still lead to an extremely complex relationship between $\mathcal{X}$ and $\tilde{\boldsymbol{\theta}}_{\text{ft}}$ due to the iterative gradient calculations. Therefore, to convert (P1) into a tractable form, it is imperative to approximate $\tilde{\boldsymbol{\theta}}_{\text{ft}}$. Intuitively, we could approximate $\tilde{\boldsymbol{\theta}}_{\text{ft}}$ by $\tilde{\boldsymbol{\theta}}^{(1)}$, resulting in the objective function for (P1) to be $L_{\text{u}}(\tilde{\boldsymbol{\theta}}^{(1)})$. However, crafting $\mathcal{X}$ by $\mathcal{X} \leftarrow \mathcal{X} - \alpha\nabla_{\mathcal{X}} L_{\text{u}}(\tilde{\boldsymbol{\theta}}^{(1)})$, with α being the step size for poison crafting, still incurs heavy computational burdens owing to the calculation of $\nabla_{\mathcal{X}} L_{\text{u}}(\tilde{\boldsymbol{\theta}}^{(1)})$, which can be expressed as below based on the *chain rule*:

$$\begin{aligned}\nabla_{\mathcal{X}} L_{\text{u}}(\tilde{\boldsymbol{\theta}}^{(1)}) &= \nabla_{\tilde{\boldsymbol{\theta}}^{(1)}} L_{\text{u}}(\tilde{\boldsymbol{\theta}}^{(1)})\nabla_{\mathcal{X}}\tilde{\boldsymbol{\theta}}^{(1)}, \\ &= \eta^2\nabla_{\boldsymbol{\theta}} L_{\text{u}}(\tilde{\boldsymbol{\theta}}^{(1)})\nabla^2_{\boldsymbol{\theta}} L_{\text{ft}}(\boldsymbol{\theta}_{\text{pt}})\nabla_{\mathcal{X}}\big(\nabla_{\boldsymbol{\theta}} L_{\text{ft}}(\boldsymbol{\theta}_{\text{pt}})\big).\end{aligned} \tag{5}$$

The second order derivatives $\nabla^2_{\boldsymbol{\theta}} L_{\text{ft}}(\boldsymbol{\theta}_{\text{pt}})$ and $\nabla_{\mathcal{X}}\big(\nabla_{\boldsymbol{\theta}} L_{\text{ft}}(\boldsymbol{\theta}_{\text{pt}})\big)$ require second-order back-propagation (BP) processes, which results in the heavy computational burdens.

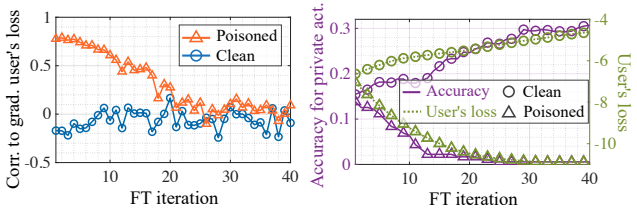
In Eqn. (5), $\nabla_{\mathcal{X}}\tilde{\boldsymbol{\theta}}^{(1)}$ is necessary for evaluating the influence of the CSI poisons on the parameter update and can hardly be further simplified. Thus, we focus on alleviating the burden of computing $\nabla_{\tilde{\boldsymbol{\theta}}^{(1)}} L_{\text{u}}(\tilde{\boldsymbol{\theta}}^{(1)})$. Since $\tilde{\boldsymbol{\theta}}^{(1)} \approx \boldsymbol{\theta}_{\text{pt}}$, we approximate $\nabla_{\tilde{\boldsymbol{\theta}}^{(1)}} L_{\text{u}}(\tilde{\boldsymbol{\theta}}^{(1)})$ with $\nabla_{\boldsymbol{\theta}} L_{\text{u}}(\boldsymbol{\theta}_{\text{pt}})$. Then, Eqn. (5) can be approximated by

$$\nabla_{\mathcal{X}} L_{\text{u}}(\tilde{\boldsymbol{\theta}}^{(1)}) \approx \nabla_{\boldsymbol{\theta}} L_{\text{u}}(\boldsymbol{\theta}_{\text{pt}})\nabla_{\mathcal{X}}\tilde{\boldsymbol{\theta}}^{(1)} \tag{6}$$

$$= \eta\nabla_{\mathcal{X}}\Big(-\nabla_{\boldsymbol{\theta}} L_{\text{u}}(\boldsymbol{\theta}_{\text{pt}})\cdot\nabla_{\boldsymbol{\theta}} L_{\text{ft}}(\boldsymbol{\theta}_{\text{pt}})\Big). \tag{7}$$

We note that through the approximation applied in Eqn. (6), the original $\nabla_{\tilde{\boldsymbol{\theta}}^{(1)}} L_{\text{u}}(\tilde{\boldsymbol{\theta}}^{(1)})$ that requires a second-order BP is reduced to a constant gradient vector. By this means, the computational expense is at least halved since only one second-order BP process is needed now. Therefore, with Eqn. (7), CSI poisons can be crafted efficiently by a resource-limited user. We note that the user can calculate $\nabla_{\boldsymbol{\theta}} L_{\text{u}}(\boldsymbol{\theta}_{\text{pt}})$ by jointly using its prepared FT dataset and several additionally-collected samples per private activity.

More specifically, to craft CSI poisons $\mathcal{X}$, the user first calculates the gradient term $\nabla_{\boldsymbol{\theta}} L_{\text{u}}(\boldsymbol{\theta}_{\text{pt}})$ in (7), which is independent from $\mathcal{X}$ and thus can be considered as a constant vector when optimizing $\mathcal{X}$. Then, with $\mathcal{X}$ added to the FT



(a) Gradient correlation. (b) Accuracy and user's loss.

Figure 7: Influence of CSI poisons on (a) correlation between gradient of the FT and that of user's loss and (b) accuracy for private activities and the user's loss.

dataset, the user evaluates $\nabla_{\boldsymbol{\theta}} L_{\text{ft}}(\boldsymbol{\theta}_{\text{pt}})$. Since the inner product $-\nabla_{\boldsymbol{\theta}} L_{\text{u}}(\boldsymbol{\theta}_{\text{pt}}) \cdot \nabla_{\boldsymbol{\theta}} L_{\text{ft}}(\boldsymbol{\theta}_{\text{pt}})$ is a scalar and depends on $\mathcal{X}$, it can be treated as a loss function of $\mathcal{X}$. Its gradient with respect to $\mathcal{X}$ can be calculated by the BP, which derives $\nabla_{\mathcal{X}} L_{\text{u}}(\tilde{\boldsymbol{\theta}}^{(1)})$ in Eqn. (6).

One can readily observe in Eqn. (7) that the poison crafting is equivalent to aligning $\nabla_{\boldsymbol{\theta}} L_{\text{u}}(\boldsymbol{\theta}_{\text{pt}})$ and $\nabla_{\boldsymbol{\theta}} L_{\text{ft}}(\boldsymbol{\theta}_{\text{pt}})$ by maximizing their inner product correlation. Notably, this accords with the intuition that the CSI poisons should **align the neural model's parameter update in the FT toward minimizing the user's loss**, which is illustrated in Figure 6(b). Furthermore, although Eqn. (7) focuses on a single iteration step, the crafted CSI poisons can continuously align the gradients in the FT toward reducing the user's loss.

To demonstrate this, continuing from Section 2.2.2, we poison the FT dataset in the 30-sample case using poisons crafted with Eqn. (7). The user additionally uses 30 samples per private activity (TP and HS) in calculating $\nabla_{\boldsymbol{\theta}} L_{\text{u}}(\boldsymbol{\theta}_{\text{pt}})$; while the complete dataset for the test user is adopted to evaluate the user's loss after the FT. Figure 7(a) shows that the gradient in the poisoned FT and that of the user's loss retain a positive correlation for around twenty iterations, proving that the efficacy of the poisons extends well beyond the first FT iteration. In stark contrast, for a clean FT dataset, the correlation is mostly negative, since the clean FT raises the accuracy for private activities, which contradicts the user's goal. Figure 7(b) demonstrates that with this positive correlation achieved by poisoning, the user's loss and the accuracy for private activities are significantly reduced, while FT on the clean dataset leads to an increase of both.

3.3 Extending to Harsher Conditions

We further extend the CSI poisoning method to handle the following conditions: i) the neural model is confidential, and ii) filtering methods are employed to cleanse CSI data.

3.3.1 Poisoning Confidential Neural Models. While open-source neural models are recognized as the key to the popularization and trustworthiness of artificial intelligent services [61], an AP manufacturer might also hold details of its neural model confidential to thwart piracy by its commercial rivals. For instance, the AP may deny access to the specific components and parameters of the neural model, only compromising to disclose its functionality and coarse architecture to gain users' trust. Under this condition, the poison crafting is impeded as the user cannot calculate the gradient of the neural parameters required in Eqn. (7).

To tackle this issue, a common approach is to leverage an ensemble of surrogate models—neural models that share the same functionality and general architecture as the confidential one, but with full access to details—then craft the poisons based on these surrogates instead. The intuition is that the crafted poisons exhibit *generalizability* to models with similar architectures [22], the principle of which has been systematically analyzed in [15]. Specifically, to select the surrogate models, we integrate three techniques below, explicitly addressing the uncertainty of the confidential model in terms of its architecture, initial point, and pre-training:

- *Diversified Architecture*: The ensemble includes multiple neural models based on the coarse architecture, each with a different number of neural layers and feature dimensions.
- *Randomized Initialization*: The surrogate models in the ensemble are initialized with different random parameters.
- *Staggered Epochs*: The surrogate models are pre-trained for staggered numbers of epochs within an empirical range.[2]

However, the limited resources of the user restrict the ensemble size, resulting in challenges to craft generalizable CSI poisons. To enhance the generalizability of the crafted poisons, during each step of poison crafting, we adopt the dropout technique [62] to randomly shortcut neurons of the surrogate models, allowing them to represent a larger variety of architectures.

Additionally, at each poison crafting step, we randomly select a batch of surrogate models from the ensemble and calculate the gradients of the average loss of all the selected models. We iteratively update a single set of CSI poisons using the sign instead of the exact values of the calculated gradients. The gradient signs represent an update direction for CSI poisons aligned with a broader range of potential surrogate models, rather than being specific to one of them.

3.3.2 Poisoning Against CSI Data Cleansing. In the pursuit of higher training data quality, the AP may employ some data cleansing methods [49], which potentially neutralize the CSI poisons. Common cleansing methods for CSI data include: i) *Low-pass filter*, which eliminates high-frequency temporal CSI variations due to hardware jittering [70]; ii) *Outlier removal*, which eliminates CSI data with abnormally

[2] If the pre-training dataset is unavailable, the user can resort to similar public datasets of Wi-Fi sensing instead. We show in later Figure 13(a) that pre-training the surrogate models with datasets collected in different environments does not affect the efficacy of Poison2Cure.

high or low power owing to strong interference or temporary blockage [39].

To ensure that the poisoned CSI data withstand these cleansing methods, we establish frequency and power constraints for CSI poisons, which are represented by Doppler frequency upper bound f_{ub} and power upper bound P_{ub}. The power upper bound can be enforced by the *projected gradient descent* (PGD) [55, 68], projecting the CSI poisons after each crafting step back to the Euclidean ball with radius $\sqrt{P_{\text{ub}}}$:

$$\boldsymbol{X} \leftarrow \sqrt{P_{\text{ub}}}(\boldsymbol{X} + \Delta\boldsymbol{X})/\big(\max(\|\boldsymbol{X} + \Delta\boldsymbol{X}\|_2, \sqrt{P_{\text{ub}}})\big). \quad (8)$$

Here, $\Delta\boldsymbol{X}$ denotes the original change to CSI poison $\boldsymbol{X}$, which is calculated based on Eqn. (7), and $\|\cdot\|_2$ denotes the Euclidean norm of the argument.

Compared with the power upper bound, the frequency upper bound is more difficult to enforce. In multi-user Wi-Fi sensing, the difficulty is exacerbated since CSI poison $\boldsymbol{X}$ is non-equally spaced in time as a result of CSI data $\boldsymbol{C}$ spanning across irregular time intervals. To handle this difficulty, rather than filtering the CSI poisons in the time domain, Poison2Cure directly crafts them in the Doppler frequency domain with a cutoff at f_{ub}: for CSI data $\boldsymbol{C} \in \mathbb{R}^{L\times T}$ (defined in Section 3.1), a spectral matrix $\boldsymbol{S} = (\boldsymbol{s}_1, \cdots, \boldsymbol{s}_L) \in \mathbb{C}^{L\times F}$ is used to generate the CSI poison $\boldsymbol{X}$ by:

$$\boldsymbol{X} = \boldsymbol{S}\boldsymbol{E} = \boldsymbol{S}\,\mathbf{exp}\big(-2\pi\mathrm{i}\cdot \boldsymbol{f}\boldsymbol{t}^{\top}\big), \quad (9)$$

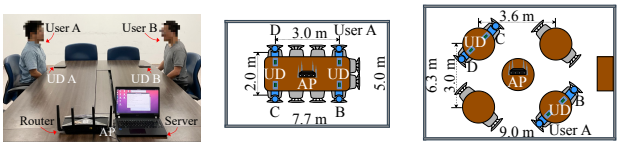
where $\mathbf{exp}(\cdot)$ is the element-wise exponential function, i is the imaginary unit, $\boldsymbol{t} = (t_1, \cdots, t_T)$ is the sequence of sampling time of the CSI data, $\boldsymbol{f} = (0, f_{\text{ub}}/(F-1), \cdots, f_{\text{ub}})$ comprises F frequency components below f_{ub}, and $(\cdot)^{\top}$ denotes the transposition. It can be observed from Eqn. (9) that $\boldsymbol{X}$ should have no frequency components higher than f_{ub}.

4 PROTOTYPING & EXPERIMENT SETUP

We describe the implementation of Poison2Cure in a commodity Wi-Fi network and then specify its experiment setup.

4.1 Implementation of Poison2Cure

Our prototype of Poison2Cure is built upon a multi-user Wi-Fi human sensing system shown in Figure 8(a). The system consists of an AP and four UDs, each placed around 20 cm from its user's chest. The UDs include a Google Pixel 6A, a Xiaomi 14 Ultra, an iPhone XR, and an iPhone 13, with CSI poison crafting uniformly conducted on an Alienware laptop to streamline prototype development. The AP is a Netgear Nighthawk X10 router connected to a server emulated by an Acer TravelMate laptop, providing activity recognition services based on the CSI of data traffic. The Wi-Fi network setup by the AP follows IEEE 802.11ac standard [1] and operates at 5.28 GHz with a bandwidth of 40 MHz. The server uses PicoScenes [34] to extract raw CSI samples from QoS Data packets. Each raw CSI sample is a complex matrix of size 117×2, corresponding to the channel gains of 117 subcarriers across two channels between the router and a UD.



(a) Prototype. (b) Meeting room. (c) Lecture room.

Figure 8: The prototype of Poison2Cure (a) and its two test environments shown in (b) and (c).

To pre-process a raw CSI sample into CSI data, we use the basic method in Section 2.2.1 because it is computationally efficient and preserves all information, while we also use the DFS method in the benchmark comparison. For both methods, the conjunction multiplication [86] is applied to the channel gains of the two channels to cancel random phase errors [7]. For the basic method, we concatenate the CSI amplitudes and the cosine and sine values of the CSI phases to form a 1D feature vector of length 351. For the DFS method, we apply the short-time Fourier transform (STFT) [65] on the channel gain sequences of all the subcarriers after resampling by 500 points per second. A one-second time window with a half-second overlapping is used in STFT, and the resulting spectra of all the subcarriers are arranged as a 2D feature map with two channels for the real and imaginary parts, which is finally resized to be a $128 \times 128 \times 2$ array. These pre-processing steps are common for CSI-based Wi-Fi sensing, extracting primitive features from CSI.

For the CSI data in the user-provided FT dataset, we assume it comprises the primitive features of CSI since using primitive features reduces the overhead in uploading the FT dataset. As the relationship between the primitive features and the raw CSI is definitive and well-established, poisons for the primitive features can be readily mapped to corresponding poisons for the raw CSI. Therefore, regardless of whether the user provides the primitive features or raw CSI, Poison2Cure remains effective.

The neural model of the AP is designed following two common architectures: i) the gated recurrent unit (GRU)-based one [88], which handles sequences of feature vectors obtained by the basic method, and ii) the convolutional neural network (CNN)-based one [32], which tackles the 2D feature map obtained by the DFS method. The GRU-based model comprises a two-layered multi-layer perceptron (MLP), two GRU layers, and a two-layered linear classifier. The CNN-based model comprises three 2D CNN layers with a kernel size of 3 and a stride of 2, each followed by an average pooling layer with a factor of 4, and then a one-layered MLP and a two-layered linear classifier. In both models, the output hidden features of MLPs have 64 dimensions by default.

4.2 Experiment Setup

4.2.1 Datasets. We collect a large dataset comprising 20 hours of raw CSI samples for 10 activities of 16 users in two environments, including a meeting room (MR) and a lecture room (LR), as illustrated in Figures 8(b) and 8(c), respectively. The 16 users are invited volunteers of 11 males and 5 females between the ages of 21 and 53. The 10 activities include 8 regular ones with low privacy concerns, comprising 4 body activities: bending (BD), jumping (JP), rotating (RT), and walking (WK); and 4 gesture activities: push&pull (PP), sweeping (SW), drawing circle (DC), and drawing zigzag (DZ). The remaining 2 activities are deemed as more privacy-sensitive, which are typing-on-phone (TP) and hand-shaking (HS) with one another.[3]

During the data collection for an activity, four users at the positions indicated in Figures 8(b) or 8(c) perform the activity simultaneously,[4] with each sample consisting of two seconds of activity and one second of rest. To generate normal data traffic, the UDs engage in an online Zoom meeting. After segmenting, pre-processing, and labeling the CSI samples, the dataset is formed. We select an arbitrary user as the test user and use the data from the rest of the users to pre-train the neural model. The test user uses 30 samples per regular activity to form its FT dataset and additionally uses 30 samples per private activity during crafting CSI poisons. The rest of data of the test user is used to test the objective of (P1). These experiments strictly follow our IRB requirements.

4.2.2 Hyper-parameters. In the pre-training, the learning rate is 10^{-3}, and the number of epochs is 100, which ensures convergence. As for the FT process, a smaller learning rate of 10^{-4} is adopted to avoid over-fitting, and the number of FT epochs is 40. When crafting CSI poisons, by default, we begin with a zero initial point and update the CSI poisons iteratively for 100 steps based on Eqn. (7) by the sign stochastic gradient descent [5] with a step size of 0.006.

When the neural model is confidential, we employ an ensemble of 32 surrogate models, whose architectures, initial points, and numbers of pre-training epochs are different from the actual model, with values sampled between 0.5 and 1.5 times those of the actual ones. In addition, the size of the stochastic model batch in each step of poison crafting is 1, and the dropout rate is 0.1. When crafting CSI poisons against cleansing, the power upper bound ensures a 30dB power ratio between the CSI data and the CSI poison; and the frequency upper bound is 32Hz, where the spectrum is discretized into 256 components. The influence of these hyper-parameters is further evaluated in Sections 5.1.2 and 5.1.3.

[3]These activities raise privacy concerns because TP may allow password inference [28], and HS can expose social interaction [78].

[4]When the four users perform HS, they form two pairs at diagonal positions.

5 EVALUATION RESULTS

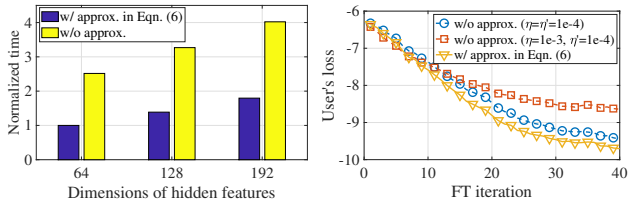
We evaluate the overall performance through micro-benchmark studies, benchmark comparison, and analyses of factors impacting poisoning.

5.1 Micro-benchmark Studies

5.1.1 Efficiency of Crafting CSI Poisons. In Figure 9(a), we compare the computational complexity of the proposed efficient CSI poisoning method with the approximation in Eqn. (6) and the poisoning method without it, i.e., poison crafting based on Eqn. (5). We measure the crafting time of CSI poisons for neural models of different sizes, determined by the dimensions of hidden features. One can observe that with the approximation in our method, the computational time is reduced by about 58%, which is in accordance with our analysis for Eqn. (6).

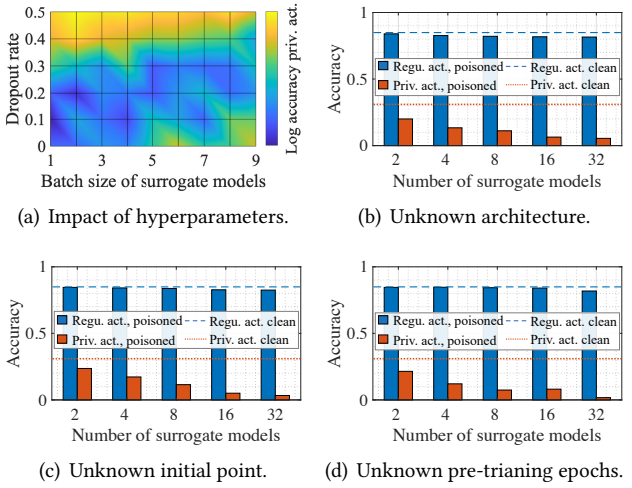
Besides, Figure 9(b) shows that the proposed method crafts more effective CSI poisons for minimizing the user's loss compared to the methods without the approximation. This is probably because maximizing the gradient correlation in Eqn. (7) utilizes the information in all the parameters of the neural model, yielding higher generalizability to multiple FT iterations. Moreover, one can observe in Figure 9(b) that the method without the approximation is susceptible to the difference between the FT learning rate η assumed in poison crafting and the actual one η'. In distinct contrast, the proposed CSI poisoning method is unaffected by such difference because it is not dependent on predicting the neural model's parameters after an FT iteration, which is required for the other method to calculate $\nabla_{\boldsymbol{\theta}} L_{\mathrm{u}}(\tilde{\boldsymbol{\theta}}^{(1)})$ in Eqn. (5).

5.1.2 Efficacy of Poisoning Confidential Model. Secondly, we evaluate the efficacy of poisoning a confidential model with the method proposed in Section 3.3.1. The actual model is a GRU-based model with two GRU layers and 64 dimensions of hidden features, which is pre-trained for 100 epochs. The ensemble comprises the surrogate models, which are randomly sampled from models with 1 to 3 GRU layers, 32 to 96 dimensions of hidden features, and pre-trained for 50 to 150 epochs from different random initial points. Figure 10(a) illustrates the accuracy for the test user's private activities



(a) Computational complexity. (b) User's loss.

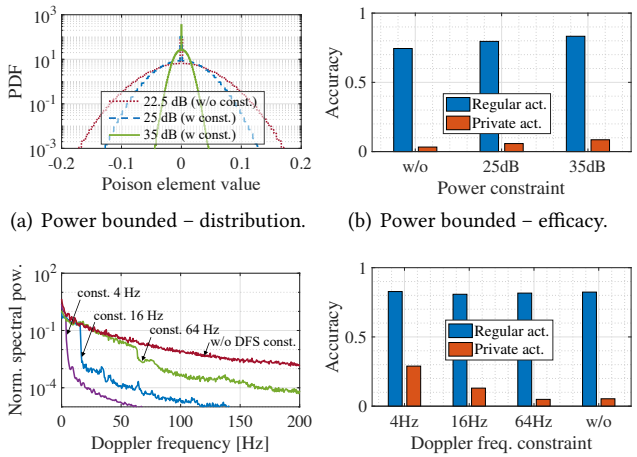
Figure 9: Comparison between the CSI poisoning methods with and without the proposed approximation.



(a) Impact of hyperparameters. (b) Unknown architecture.

(c) Unknown initial point. (d) Unknown pre-trianing epochs.

Figure 10: Poisoning the confidential neural model (a) given different hyperparameters and (b)-(d) in different cases of confidentiality. Interpolation is applied to (a) to fill in the regions between the evaluation points.



(a) Power bounded – distribution. (b) Power bounded – efficacy.

(c) Frequency bounded – distribution. (d) Frequency bounded – efficacy.

Figure 11: Comparison between the CSI poisons given different power and Doppler frequency upper bounds.

given different dropout rates and model batch sizes during the poison crafting. It can be observed that using a small dropout rate and a small model batch size improves the generalizability of CSI poisons to poison the confidential model. Thus, we adopt the best hyper-parameters, i.e., the dropout rate of 0.1 and the model batch size of 1, in our experiments.

In Figures 10(b)–10(d), we evaluate the individual cases of confidential architecture, initial point, and number of pre-training epochs. It is evident that they show similar trends: as the number of surrogate models increases, the crafted CSI poisons become more effective for the confidential model in terms of preserving private activities. These results prove the capability of Poison2Cure to achieve semantic-level privacy preservation for confidential neural models.

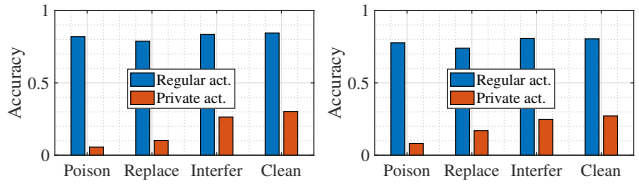
5.1.3 Efficacy of Poisoning under Constraints. Thirdly, we show that the proposed method in Section 3.3.2 can craft effective CSI poisons under different power and Doppler frequency upper bounds. Figure 11(a) demonstrates the value distributions of the crafted poisons without constraint (which results in an average 22.5 dB power ratio between CSI data and CSI poisons) and those bounded by 25 dB and 35 dB, and Figure 11(b) compares their respective performance in terms of the average accuracy for regular and private activities after the FT. It is evident in Figure 11(a) that with our method, the power of the crafted CSI poisons is effectively reduced, while it maintains effective semantic-level privacy preservation, as proven by Figure 11(b). In addition, Figure 11(b) shows that constraining the CSI poisons in power mitigates the poisons' impact on the accuracy for regular activities.

On the other hand, Figure 11(c) shows the spectral magnitudes of the crafted poisons without a Doppler frequency bound and those bounded by 4Hz, 16Hz, and 64Hz, while Figure 11(d) illustrates their respective performance. It is evident that the proposed method effectively suppresses the frequency components of the CSI poisons that exceed the frequency upper bound, despite a slight degree of leakage. In Figure 11(d), two interesting observations can be drawn: *Firstly*, the crafted CSI poisons can remain quite effective even under a stringent upper bound of 16Hz. *Secondly*, the 64Hz frequency upper bound slightly enhances the efficacy of CSI poisons for preserving private activities. In addition, we note that the efficacy decrease given the bounds of 4Hz and 16Hz is mainly because the datasets for both pre-training and FT are not filtered accordingly. When the neural model is trained with data filtered by similar bounds, the CSI poisons bounded by an even lower Doppler frequency (e.g., 2Hz) can remain highly effective (see Section 5.3.3).

5.2 Benchmark Comparison

Besides the proposed method (*Poison*) and the baseline with no privacy preservation (*Clean*), we compare Poison2Cure with two benchmark methods below in terms of the resulting accuracy for regular and private activities, using both the GRU- and CNN-based neural models.

- *Data Replacing (Replace)*: In the FT dataset, randomly select CSI data of regular activities and replace them with CSI data of private activities, while keeping their labels unchanged to make the neural model misclassify private activities. This emulates the label-flipping attack. For a fair comparison, the replacement ensures all regular and private activities have the same number of CSI data.



(a) GRU-based neural model. (b) CNN-based neural model.

Figure 12: Benchmark comparison under different cases of neural models.

- *Channel Interfering (Interfer)*: Without changing the FT dataset, the user injects noise into the CSI by interfering with the Wi-Fi channel when performing private activities. For covertness and a fair comparison, the noise power is set to be 30 dB below the CSI power.

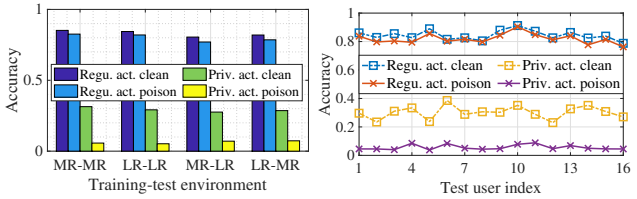
Figures 12(a) and 12(b) show that Poison2Cure outperforms the two benchmarks significantly. Averaged across the cases for GRU-based and CNN-based models, *Poison* reduces the accuracy for private activities by 76%, compared to 52% and 11% for *Replace* and *Interfer*, respectively. Moreover, the second-best method, *Replace*, sacrifices 7.4% of original accuracy for regular activities, which is 2.3 times that of *Poison*.

The reason that *Replace* is less effective than *Poison* is twofold: i) *Poison* adjusts the complete FT dataset, while *Replace* can only change part of it to maintain the accuracy for regular activities; ii) *Replace* makes the neural model *minimize* the loss for some intentionally erroneous classifications, which is generally less efficient than directly *maximizing* the classification loss as achieved in *Poison*. Moreover, *Replace* achieves lower accuracy for regular activities since introducing incorrectly labeled CSI data reduces the already small amount of training data for regular activities. As for *Interfer*, it does not achieve satisfactory privacy preservation under the same power constraint as *Poison*, owing to the general robustness of neural models to random noises.

5.3 Factors Impacting Poisoning

Below, we evaluate the impact of six practical factors on Poison2Cure. We focus on the GRU-based model case, since the results for the CNN-based model are similar.

5.3.1 Environments and Users. To evaluate the impact of environments, we consider four cases in Figure 13(a). In the name of each case, the left word indicates the environment whose CSI data is used for the pre-training, and the word on the right indicates the environment where the test user obtains and poisons its FT dataset. Figure 13(a) shows that environments have no substantial impacts on the sensing performance, regardless of whether the FT dataset is poisoned or not. The results of the "MR-LR" and "LR-MR" cases also



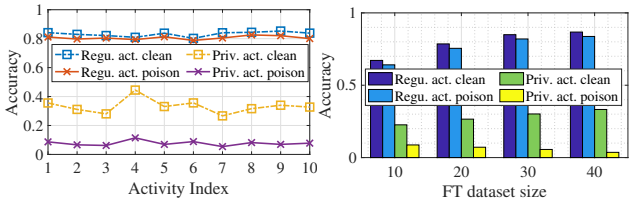
(a) Impact of environments. (b) Impact of test users.

Figure 13: Impact of environments and test users.

imply that FT with a dataset collected in a different environment than the pre-training one has no significant impact on Poison2Cure. On the other hand, Figure 13(b) shows that the sensing accuracy after FT varies across different test users. This is because the user-specific nature of multi-user Wi-Fi sensing makes the sensing accuracy of the neural model dependent on test users even after the FT. Nevertheless, across different test users, Poison2Cure demonstrates significant semantic privacy preservation consistently, reducing accuracy for private activities by 81% on average.

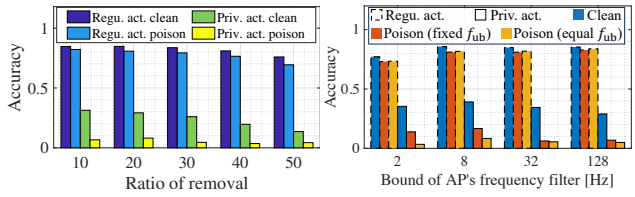
5.3.2 Private Activities and FT Datasets. We evaluate the impact when the test user defines different activities as private while treating the others as regular. From Figure 14(a), one can observe that after a clean FT process, the accuracy for private activity varies depending on which activity is deemed private; however, Poison2Cure consistently reduces it to less than 0.08. Besides, we evaluate the influence of the size of FT dataset. Figure 14(b) verifies that when the FT dataset is larger, the FT process enhances the sensing accuracy for both regular and private activities to a larger extent. More importantly, the larger FT dataset also benefits the semantic-level privacy preservation since using more CSI data also yields a higher design freedom for CSI poisons.

5.3.3 Outlier Removal and Low-Pass Filter. Finally, we evaluate the impact of AP's data cleansing. We consider the cases where the outlier removal of the AP removes 10% to 50% of the CSI data with either the lowest or highest power in the FT dataset. As we restrict the power of CSI poisons to be 30 dB lower than the CSI data, poisoning the FT dataset hardly affects the removal. Moreover, Figure 15(a) demonstrates



(a) Impact of private activity. (b) Impact of FT dataset size.

Figure 14: Impact of private activity and size of FT dataset.



(a) Impact of outlier removal. (b) Impact of frequency filter.

Figure 15: Impact of the outlier removal and low-pass filter, while power and Doppler frequency bounds are employed to counter the CSI data cleansing.

that removing as much as 50% of the poisoned FT data does not undermine the semantic-level privacy preservation of Poison2Cure, despite the general sensing accuracy declining due to the reduced size of the FT dataset.

As for the lower-pass filter, we let the AP filter out high Doppler frequencies in both the pre-training and the FT datasets, with the cutoff frequency ranging from 2 Hz to 128 Hz. Two cases are evaluated: i) the Doppler frequency bound set by the user *equals* that of the AP's filter, and ii) the Doppler frequency bound of the user is *fixed* to 16 Hz. As shown in Figure 15(b), when the datasets of both pre-training and FT are filtered accordingly, even a Doppler frequency bound of 2 Hz does not compromise the efficacy of Poison2Cure. Besides, Figure 15(b) confirms that the efficacy of Poison2Cure is resilient to the discrepancy between the Doppler frequency bounds adopted by the user and the AP, regardless of whether the user's bound is lower or higher.

6 DISCUSSION

We discuss two important aspects of Poison2Cure in its pursuit of becoming a standard paradigm for privacy-preserving multi-user Wi-Fi sensing: its practicability and susceptibility.

Practicability. Though we do not assume the AP to be fully cooperative (otherwise, there is no need for privacy preservation), a certain level of cooperation between the AP and the user is needed by Poison2Cure. Firstly, the AP manufacturer needs to disclose the functionality and rough architecture of its neural model and allow users to upload their prepared FT datasets. Despite the seemingly demanding appearance of these conditions, the AP manufacturer has the incentive to make such compromises because it is business-oriented, i.e., its primary goal is to popularize its products and sensing services, while collecting (private) information from users is its secondary goal.

For the primary goal of the AP, open-source is recognized as a key factor in the popularization and trustworthiness of AI services. Many successful AI services, such as Google's Gemma [13], Meta's Llama [3], and DeepSeek-AI's DeepSeek LLM [14], have gained users' trust and expanded market share by open-sourcing their models. In the field of Wi-Fi sensing, projects such as Widar3.0 [88], MUSE-Fi [29], and SenseFi [75], etc., have also open-sourced their neural models, contributing to transparency and collaboration. While commercial deployment of Wi-Fi sensing is still in its early stage, the above examples suggest the AP manufacturers' strong incentives to disclose partial information about their neural models when entering the market, thereby encouraging users to adopt their services. In this regard, certain agreements and legislation could be established to enforce these conditions.

Secondly, although the users' motivation is clearly driven by their desire for sensing services, collecting an FT dataset could be laborious as tens of samples are needed for each activity. Fortunately, this labor could be alleviated by enhancing the FT process with advanced few-shot learning [76] and one-shot learning [71] techniques, which can significantly reduce the number of required samples per activity to just a few or only one. In these cases, the FT becomes more sensitive to the small number of samples, potentially making it more convenient to craft CSI poisons for privacy preservation.

Susceptibility. Poison2Cure would be susceptible if the AP were able to detect and remove CSI poisons or could restore the original clean CSI data. In Figure 16, we compare a pair of clean and poisoned CSI data through their spectrograms obtained by continuous wavelet transformation (CWT) [82] and STFT [65]. It is evident that, in terms of their temporal-spectral features, the poisoned CSI data is virtually indistinguishable from the clean one and thus can be hardly detected.

This advantage, however, may turn out to be a new susceptibility. Specifically, given channel reciprocity, the AP can record non-poisoned CSI data of the user and match it with those in the user-uploaded FT dataset to acquire their activity labels. However, since the actual timing for the user's CSI collection is unpredictable by the AP, the AP has to continuously record reciprocal CSI to prepare for the matching, leading to prohibitive memory and computational costs due to its limited hardware resources.

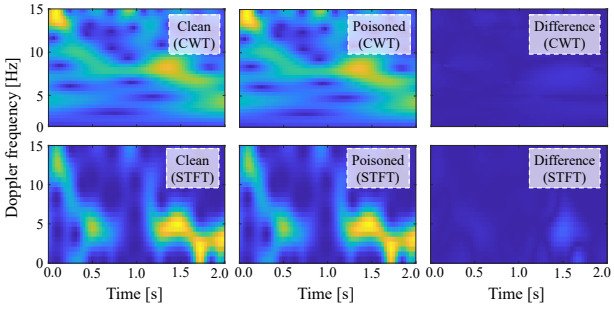


Figure 16: CWT and STFT spectrograms of the clean and the poisoned CSI data. The color intensity of their difference are magnified 10x to facilitate observation.

In summary, the poisoned CSI data are undetectable by the AP unless at an unacceptable cost. Furthermore, the susceptibility to matching can be completely prevented if data condensation techniques [77, 87] are employed to condense the poisoned FT dataset into a distinct and much smaller one with equivalent training efficacy. In this way, not only are the transmission overheads reduced, but also the matching between labels and non-poisoned CSI data is prevented.

7 RELATED WORKS

Poison2Cure is closely related to cross-domain Wi-Fi sensing and privacy preservation techniques for wireless sensing.

Cross-Domain Wi-Fi sensing. Due to the multipath effect, Wi-Fi sensing techniques are generally domain-sensitive. In literature, several cross-domain techniques are proposed, which Poison2Cure is highly related to as a special cross-domain paradigm. One of the main branches of work [29, 37, 51] adopts the DFS of CSI for cross-domain sensing, leveraging its time-frequency decomposition of the CSI variations. Widar 3.0 [86] processes the DFS at multiple receivers to obtain the BVP for cross-domain gesture recognition. Gao *et al.* [20] use multiple pairs of transceivers and derive cross-position features for motion direction changes, which is extended to DFS in [54]. Cross-domain Wi-Fi sensing techniques have also been actively studied from the perspective of machine learning. EI [32] uses adversarial learning to extract domain-invariant features from CSI automatically, which is enhanced with data augmentation in [67]. Additionally, with powerful generative adversarial networks, Li *et al.* [41] propose to further enhance the generalizability of sensing by using synthesized CSI data of diverse domains.

However, in multi-user Wi-Fi sensing scenarios, the impact of the user's nuances becomes significant and harder to predict. In this case, real samples from the target domain are necessary to fine-tune the neural model. Using collected real samples, CrossSense [81] and SIDA [80] learn the mapping between the data in the target domain and that in the training domains. To reduce the number of required samples, Fewsense [76] utilizes few-shot learning, while RF-Net [16] and Wi-Learner [17] resort to one-shot learning. Yet, these approaches have no privacy-preserving mechanisms and may compromise users' sensitive activities.

Privacy Preservation in Wireless Sensing. Given the pervasive presence of wireless signals in our living environments, users' privacy is inevitably exposed to wireless sensing. PhyCloak [57] first uses a relay to distort the physical information in the signals received by unauthorized devices. Similarly, Cominelli *et al.* [11, 12] focus on obfuscating the location-relevant information carried by the CSI. Besides, IRShield [63] leverages an intelligent reflecting surface as a novel approach for CSI obfuscation. MIMOCrypt [48] and Secur-Fi [52] use codebooks shared between authorized transceivers to encrypt CSI and prevent eavesdropping. Nevertheless, the above approaches cannot preserve privacy at a semantic level. To thwart the sensing of only the private activities, existing works resort to adversarial attacks [45, 72, 84, 89] and add specially designed perturbations to CSI at the receiver side, inducing neural models to misclassify private activities. Furthermore, Huang *et al.* [30] propose a transmitter-side on-device perturbation method against only wireless positioning. However, similar to CSI obfuscation, adversarial perturbations of CSI are also domain-specific, and their reliable injection requires manipulation of real-time CSI, rendering them hardly compatible with prevalent Wi-Fi hardware.

8 CONCLUSION

Aiming at preserving users' privacy from ubiquitous Wi-Fi sensing, Poison2Cure has pioneered a data poisoning framework for semantic-level privacy preservation. Leveraging the necessary cross-user calibration of the AP, where a user provides the AP's neural model with labeled CSI data for FT, Poison2Cure has successfully enabled the user to poison the CSI data, neutralizing the AP's sensing ability for private activities without undermining that for regular ones. This success is attributed to our efficient CSI poisoning method; we have achieved a significant complexity reduction to fit it for resource-limited users and extended its efficacy over confidential neural models and CSI data cleansing. Our extensive evaluations have evidently confirmed that Poison2Cure can work under diverse conditions, including different preprocessing of CSI data and distinct neural model architectures. Furthermore, it has also been verified that Poison2Cure is robust to changes in environments, users, and private activities. Owing to its proven efficacy and full compatibility with any underlying Wi-Fi hardware and firmware, we believe the paradigm-shifting Poison2Cure can be widely popular in future practice. Meanwhile, we are also studying security of Wi-Fi sensing under ISAC framework [9, 24] and also its co-existence with other co-channel communication systems [38, 73] as part of our future work.

ACKNOWLEDGEMENT

This research is supported by the National Research Foundation Singapore and DSO National Laboratories under the AI Singapore Programme (AISG Award No: AISG2-GC-2023-006), the National Research Foundation Singapore and Infocomm Media Development Authority under its Future Communications Research & Development Programme, and MOE Tier 1 grant RG16/22.

REFERENCES

[1] 2021. IEEE Standard for Information Technology–Telecommunications and Information Exchange between Systems - Local and Metropolitan Area Networks–Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications - Redline. *IEEE Std 802.11-2020 (Revision of IEEE Std 802.11-2016) - Redline* (2021), 1–7524.

[2] Karan Ahuja, Yue Jiang, Mayank Goel, and Chris Harrison. 2021. Vid2Doppler: Synthesizing Doppler Radar Data from Videos for Training Privacy-Preserving Activity Recognition. In *Proc. of the 39th ACM CHI*. 1–10.

[3] Meta AI. 2024. Llama Models. Available: https://github.com/meta-llama/llama-models. Access: Dec. 24, 2024.

[4] Mohammed Alloulah, Anton Isopoussu, and Fahim Kawsar. 2018. On Indoor Human Sensing Using Commodity Radar. In *Proc of the 20th ACM UbiComp*. 1331–1336.

[5] Jeremy Bernstein, Yu-Xiang Wang, Kamyar Azizzadenesheli, and Animashree Anandkumar. 2018. signSGD: Compressed Optimisation for Non-convex Problems. In *Proc. of the 35th ICML*. 560–569.

[6] Yetong Cao, Chao Cai, Anbo Yu, Fan Li, and Jun Luo. 2023. EarAcE: Empowering Versatile Acoustic Sensing via Earable Active Noise Cancellation Platform. *Proc. of the ACM IMWUT* 7, 2 (2023), 1–23.

[7] Chen Chen, Gang Zhou, and Youfang Lin. 2023. Cross-Domain WiFi Sensing with Channel State Information: A Survey. *Comput. Surveys* 55, 11 (2023), 1–37.

[8] Yong Chen, Yueming Cai, Guoru Ding, Baoquan Yu, and Chenglong Xu. 2023. Age of Information for Short-packet Relay Communications in Cognitive-radio-based Internet of Things with Outdated Channel State Information. *IEEE Transactions on Cognitive Communications and Networking* 9, 3 (2023), 722–737.

[9] Zhe Chen, Tianyue Zheng, Chao Hu, Hangcheng Cao, Yanbing Yang, Hongbo Jiang, and Jun Luo. 2023. ISACoT: Integrating Sensing with Data Traffic for Ubiquitous IoT Devices. *IEEE Communications Magazine* 61, 5 (2023), 98–104.

[10] Renato Lo Cigno, Francesco Gringoli, Marco Cominelli, and Lorenzo Ghiro. 2022. Integrating CSI Sensing in Wireless Networks: Challenges to Privacy and Countermeasures. *IEEE Network* 36, 4 (2022), 174–180.

[11] Marco Cominelli, Francesco Gringoli, and Renato Lo Cigno. 2022. AntiSense: Standard-compliant CSI Obfuscation Against Unauthorized Wi-Fi Sensing. *Computer Communications* 185 (2022), 92–103.

[12] Marco Cominelli, Felix Kosterhon, Francesco Gringoli, Renato Lo Cigno, and Arash Asadi. 2021. IEEE 802.11 CSI Randomization to Preserve Location Privacy: An Empirical Evaluation in Different Scenarios. *Computer Networks* 191 (2021), 107970.

[13] Google DeepMind. 2024. Gemma: Lightweight Open Models. Available: https://github.com/google-deepmind/gemma. Access: Dec. 24, 2024.

[14] DeepSeek-AI. 2024. DeepSeek-LLM. Available: https://github.com/deepseek-ai/deepseek-LLM. Access: Dec. 24, 2024.

[15] Ambra Demontis, Marco Melis, Maura Pintor, Matthew Jagielski, Battista Biggio, Alina Oprea, Cristina Nita-Rotaru, and Fabio Roli. 2019. Why Do Adversarial Attacks Transfer? Explaining Transferability of Evasion and Poisoning Attacks. In *Proc. of the 28th USENIX Security*. 321–338.

[16] Shuya Ding, Zhe Chen, Tianyue Zheng, and Jun Luo. 2020. RF-Net: A Unified Meta-learning Framework for RF-enabled One-shot Human Activity Recognition. In *Proc. of the 18th ACM SenSys*. 517–530.

[17] Chao Feng, Nan Wang, Yicheng Jiang, Xia Zheng, Kang Li, Zheng Wang, and Xiaojiang Chen. 2022. Wi-learner: Towards One-shot Learning for Cross-domain Wi-Fi based Gesture Recognition. *Proc. of the ACM IMWUT* 6, 3 (2022), 1–27.

[18] Zhangjie Fu, Jiashuang Xu, Zhuangdi Zhu, Alex X Liu, and Xingming Sun. 2018. Writing in the Air with WiFi Signals for Virtual Reality Devices. *IEEE Transactions on Mobile Computing* 18, 2 (2018), 473–484.

[19] Ruiyang Gao, Wenwei Li, Yaxiong Xie, Enze Yi, Leye Wang, Dan Wu, and Daqing Zhang. 2022. Towards Robust Gesture Recognition by Characterizing the Sensing Quality of WiFi Signals. *Proc. of the ACM IMWUT* 6, 1 (2022), 1–26.

[20] Ruiyang Gao, Mi Zhang, Jie Zhang, Yang Li, Enze Yi, Dan Wu, Leye Wang, and Daqing Zhang. 2021. Towards Position-independent Sensing for Gesture Recognition with Wi-Fi. *Proc. of the ACM IMWUT* 5, 2 (2021), 1–28.

[21] Yan Gao, Yang Long, Yu Guan, Anna Basu, Jessica Baggaley, and Thomas Ploetz. 2019. Towards Reliable, Automated General Movement Assessment for Perinatal Stroke Screening in Infants Using Wearable Accelerometers. *Proc. of the ACM IMWUT* 3, 1 (2019), 1–22.

[22] Micah Goldblum, Dimitris Tsipras, Chulin Xie, Xinyun Chen, Avi Schwarzschild, Dawn Song, Aleksander Mądry, Bo Li, and Tom Goldstein. 2023. Dataset Security for Machine Learning: Data Poisoning, Backdoor Attacks, and Defenses. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 2 (2023), 1563–1580.

[23] 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 *Proc. of the 13th ACM WiNTECH*. 21–28.

[24] Yinghui He, Jianwei Liu, Mo Li, Guanding Yu, Jinsong Han, and Kui Ren. 2023. SenCom: Integrated Sensing and Communication with Practical WiFi. In *Proc. of the 29th ACM MobiCom*. 1–16.

[25] Steven M Hernandez and Eyuphan Bulut. 2023. Scheduled Spatial Sensing Against Adversarial WiFi Sensing. In *Proc. of 21st IEEE PerCom*. 91–100.

[26] Chung Duc Ho, Toan-Van Nguyen, Thien Huynh-The, Tien-Tung Nguyen, Daniel Benevides da Costa, and Beongku An. 2021. Short-packet Communications in Wireless-powered Cognitive IoT Networks: Performance Analysis and Deep Learning Evaluation. *IEEE Transactions on Vehicular Technology* 70, 3 (2021), 2894–2899.

[27] Hande Hong, Chengwen Luo, and Mun Choon Chan. 2016. SocialProbe: Understanding Social Interaction Through Passive WiFi Monitoring. In *Proc. of the 13th EAI Mobiquitous*. 94–103.

[28] Jingyang Hu, Hongbo Wang, Tianyue Zheng, Jingzhi Hu, Zhe Chen, Hongbo Jiang, and Jun Luo. 2023. Password-Stealing Without Hacking: Wi-Fi Enabled Practical Keystroke Eavesdropping. In *Proc. of 30th ACM CCS*. 239–252.

[29] Jingzhi Hu, Tianyue Zheng, Zhe Chen, Hongbo Wang, and Jun Luo. 2023. MUSE-Fi: Contactless MUti-person SEnsing Exploiting Near-field Wi-Fi Channel Variation. In *Proc. of the 29th ACM MobiCom*. 1135–1149.

[30] Pengzhi Huang, Emre Gönültaş, Maximilian Arnold, K Pavan Srinath, Jakob Hoydis, and Christoph Studer. 2024. Attacking and Defending Deep-learning-based Off-device Wireless Positioning Systems. *IEEE Transactions on Wireless Communications* 23, 8 (2024), 8883–8895.

[31] Pei Huang, Xiaonan Zhang, Sihan Yu, and Linke Guo. 2021. IS-WARS: Intelligent and Stealthy Adversarial Attack to Wi-Fi-based Human Activity Recognition Systems. *IEEE Transactions on Dependable and Secure Computing* 19, 6 (2021), 3899–3912.

[32] Wenjun Jiang, Chenglin Miao, Fenglong Ma, Shuochao Yao, Yaqing Wang, Ye Yuan, Hongfei Xue, Chen Song, Xin Ma, Dimitrios Koutsonikolas, et al. 2018. Towards Environment Independent Device Free Human Activity Recognition. In *Proc. of the 24th ACM MobiCom*. 289–304.

[33] Wenjun Jiang, Hongfei Xue, Chenglin Miao, Wang Shiyang, Lin Sen, Chong Tian, Srinivasan Murali, Haochen Hu, Zhi Sun, and Lu Su. 2020. Towards 3D Human Pose Construction Using WiFi. In *Proc. of the 26th ACM MobiCom*. 23:1–14.

[34] Zhiping Jiang, Tom H. Luan, Xincheng Ren, Dongtao Lv, Han Hao, Jing Wang, Kun Zhao, Wei Xi, Yueshen Xu, and Rui Li. 2021. Eliminating

the Barriers: Demystifying Wi-Fi Baseband Design and Introducing the PicoScenes Wi-Fi Sensing Platform. *IEEE Internet of Things Journal* (2021), 1–21.

[35] Manikanta Kotaru and Sachin Katti. 2017. Position Tracking for Virtual Reality using Commodity WiFi. In *Proc. of the 30th IEEE/CVF CVPR*. 68–78.

[36] Alexey Kurakin, Ian J Goodfellow, and Samy Bengio. 2018. Adversarial Examples in the Physical World. In *Artificial Intelligence Safety and Security*. Chapman and Hall/CRC, 99–112.

[37] Chenning Li, Manni Liu, and Zhichao Cao. 2020. WiHF: Enable user identified gesture recognition with WiFi. In *Proc. of the 39th IEEE INFOCOM*. 586–595.

[38] Feng Li, Jun Luo, Gaotao Shi, and Ying He. 2017. ART: Adaptive fRequency-Temporal Co-Existing of ZigBee and WiFi. *IEEE Trans. on Mobile Computing* 16, 3 (2017), 662–674.

[39] Hong Li, Wei Yang, Jianxin Wang, Yang Xu, and Liusheng Huang. 2016. WiFinger: Talk to Your Smart Devices with Finger-grained Gesture. In *Proc of the 18th ACM UbiComp*. 250–261.

[40] Mengyuan Li, Yan Meng, Junyi Liu, Haojin Zhu, Xiaohui Liang, Yao Liu, and Na Ruan. 2016. When CSI Meets Public WiFi: Inferring Your Mobile Phone Password via WiFi Signals. In *Proc. of 23rd ACM CCS*. 1068–1079.

[41] Xinyi Li, Liqiong Chang, Fangfang Song, Ju Wang, Xiaojiang Chen, Zhanyong Tang, and Zheng Wang. 2021. CrossGR: Accurate and Low-cost Cross-target Gesture Recognition Using Wi-Fi. *Proc. of the ACM IMWUT* 5, 1 (2021), 1–23.

[42] Xin Li, Hongbo Wang, Zhe Chen, Zhiping Jiang, and Jun Luo. 2024. UWB-Fi: Pushing Wi-Fi towards Ultra-wideband for Fine-Granularity Sensing. In *Proc. of the 22nd ACM MobiSys*. 42–55.

[43] Jianwei Liu, Yinghui He, Chaowei Xiao, Jinsong Han, and Kui Ren. 2024. Time to Think the Security of WiFi-Based Behavior Recognition Systems. *IEEE Transactions on Dependable and Secure Computing* 21, 1 (2024), 449–462.

[44] Jian Liu, Yan Wang, Yingying Chen, Jie Yang, Xu Chen, and Jerry Cheng. 2015. Tracking Vital Signs During Sleep Leveraging Off-the-Shelf WiFi. In *Proc. of the 16th ACM MobiHoc*. 267–276.

[45] Jianwei Liu, Chaowei Xiao, Kaiyan Cui, Jinsong Han, Xian Xu, and Kui Ren. 2023. Behavior Privacy Preserving in RF Sensing. *IEEE Transactions on Dependable and Secure Computing* 20, 1 (2023), 784–796.

[46] Shijia Liu, Zhenghua Chen, Min Wu, Chang Liu, and Liangyin Chen. 2023. WiSR: Wireless Domain Generalization Based on Style Randomization. *IEEE Transactions on Mobile Computing* (2023), 1–13.

[47] Fei Luo, Stefan Poslad, and Eliane Bodanese. 2020. Temporal Convolutional Networks for Multiperson Activity Recognition Using a 2-D LIDAR. *IEEE Internet of Things Journal* 7, 8 (2020), 7432–7442.

[48] Jun Luo, Hangcheng Cao, Hongbo Jiang, Yanbing Yang, and Zhe Chen. 2024. MIMOCrypt: Multi-User Privacy-Preserving Wi-Fi Sensing via MIMO Encryption. In *Proc. of the 45th IEEE S&P*. 1–19.

[49] Yongsen Ma, Gang Zhou, and Shuangquan Wang. 2019. WiFi Sensing with Channel State Information: A Survey. *Comput. Surveys* 52, 3 (2019), 1–36.

[50] Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing Data Using t-SNE. *Journal of Machine Learning Research* 9 (2008), 2579–2605.

[51] Francesca Meneghello, Domenico Garlisi, Nicolò Dal Fabbro, Ilenia Tinnirello, and Michele Rossi. 2022. SHARP: Environment and Person Independent Activity Recognition with Commodity IEEE 802.11 Access Points. *IEEE Transactions on Mobile Computing* 22, 10 (2022), 6160–6175.

[52] Xuanqi Meng, Jiarun Zhou, Xiulong Liu, Xinyu Tong, Wenyu Qu, and Jianrong Wang. 2023. Secur-Fi: A Secure Wireless Sensing System Based on Commercial Wi-Fi Devices. In *Proc. of the 42nd IEEE INFOCOM*. 1–10.

[53] Xuan Son Nguyen, Luc Brun, Olivier Lézoray, and Sébastien Bougleux. 2019. A Neural Network based on SPD Manifold Learning for Skeleton-based Hand Gesture Recognition. In *Proc. of the 32nd IEEE/CVF CVPR*. 12036–12045.

[54] Kai Niu, Fusang Zhang, Xuanzhi Wang, Qin Lv, Haitong Luo, and Daqing Zhang. 2021. Understanding WiFi Signal Frequency Features for Position-independent Gesture Sensing. *IEEE Transactions on Mobile Computing* 21, 11 (2021), 4156–4171.

[55] Neal Parikh, Stephen Boyd, et al. 2014. Proximal algorithms. *Foundations and Trends in Optimization* 1, 3 (2014), 127–239.

[56] Prasoon Patidar, Mayank Goel, and Yuvraj Agarwal. 2023. VAX: Using Existing Video and Audio-based Activity Recognition Models to Bootstrap Privacy-Sensitive Sensors. *Proc. of the ACM IMWUT* 7, 3 (2023), 1–24.

[57] Yue Qiao, Ouyang Zhang, Wenjie Zhou, Kannan Srinivasan, and Anish Arora. 2016. PhyCloak: Obfuscating Sensing from Communication Signals. In *Proc. of the 13rd USENIX NSDI*. 685–699.

[58] Vitor Fortes Rey, Peter Hevesi, Onorina Kovalenko, and Paul Lukowicz. 2019. Let There be IMU Data: Generating Training Data for Wearable, Motion Sensor based Activity Recognition from Monocular RGB Videos. In *Proc of the 21st ACM UbiComp*. 1331–1336.

[59] Hamada Rizk, Yuma Okochi, and Hirozumi Yamaguchi. 2022. Demonstrating Hitonavi-μ: A Novel Wearable LiDAR for Human Activity Recognition. In *Proc. of the 28th ACM MobiCom*. 756–757.

[60] M. S. Ryoo. 2011. Human Activity Prediction: Early Recognition of Ongoing Activities from Streaming Videos. In *Proc. of the 13th IEEE ICCV*. 1036–1043.

[61] Yash Raj Shrestha, Georg Von Krogh, and Stefan Feuerriegel. 2023. Building open-source AI. *Nature Computational Science* 3, 11 (2023), 908–911.

[62] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *The Journal of Machine Learning Research* 15, 1 (2014), 1929–1958.

[63] Paul Staat, Simon Mulzer, Stefan Roth, Veelasha Moonsamy, Markus Heinrichs, Rainer Kronberger, Aydin Sezgin, and Christof Paar. 2022. IRShield: A Countermeasure Against Adversarial Physical-layer Wireless Sensing. In *Proc. of the 43rd IEEE S&P*. 1705–1721.

[64] Sheng Tan, Yili Ren, Jie Yang, and Yingying Chen. 2022. Commodity WiFi Sensing in Ten Years: Status, Challenges, and Opportunities. *IEEE Internet of Things Journal* 9, 18 (2022), 17832–17843.

[65] Raghav H Venkatnarayan, Griffin Page, and Muhammad Shahzad. 2018. Multi-user Gesture Recognition using WiFi. In *Proc of the 16th ACM MobiSys*. 401–413.

[66] Aditya Virmani and Muhammad Shahzad. 2017. Position and Orientation Agnostic Gesture Recognition Using WiFi. In *Proc. of the 15th ACM MobiSys*. 252–264.

[67] Dazhuo Wang, Jianfei Yang, Wei Cui, Lihua Xie, and Sumei Sun. 2022. AirFi: Empowering WiFi-based Passive Human Gesture Recognition to Unseen Environment via Domain Generalization. *IEEE Transactions on Mobile Computing* 23, 2 (2022), 1156–1168.

[68] Yunjuan Wang, Poorya Mianjy, and Raman Arora. 2021. Robust Learning for Data Poisoning Attacks. In *Proc. of the 38th ICML*. 10859–10869.

[69] Tzu-Tsung Wong. 2015. Performance Evaluation of Classification Algorithms by k-fold and Leave-one-out Cross Validation. *Pattern Recognition* 48, 9 (2015), 2839–2846.

[70] Fu Xiao, Jing Chen, Xiaohui Xie, Linqing Gui, Lijuan Sun, and Ruchuan Wang. 2018. SEARE: A System for Exercise Activity Recognition and Quality Evaluation based on Green Sensing. *IEEE Transactions on Emerging Topics in Computing* 8, 3 (2018), 752–761.

[71] Rui Xiao, Jianwei Liu, Jinsong Han, and Kui Ren. 2021. OneFi: One-Shot Recognition for Unseen Gesture via COTS WiFi. In *Proc of the 19th ACM SenSys*. 206–219.

[72] Leiyang Xu, Xiaolong Zheng, Xiangyuan Li, Yucheng Zhang, Liang Liu, and Huadong Ma. 2022. WiCAM: Imperceptible Adversarial Attack on Deep Learning based WiFi Sensing. In *Proc. of the 19th IEEE SECON*. 10–18.

[73] Ruitao Xu, Gaotao Shi, Jun Luo, Zenghua Zhao, and Yantai Shu. 2011. MuZi: Multi-Channel ZigBee Networks for Avoiding WiFi Interference. In *Proc. of the 4th IEEE/ACM CPSCOM*. 323–329.

[74] Edwin Yang, Qiuye He, and Song Fang. 2022. WINK: Wireless Inference of Numerical Keystrokes via Zero-Training Spatiotemporal Analysis. In *Proc. of the 29th ACM CCS*. 3033–3047.

[75] Jianfei Yang, Xinyan Chen, Han Zou, Chris Xiaoxuan Lu, Dazhou Wang, Sumei Sun, and Lihua Xie. 2023. SenseFi: A library and benchmark on deep-learning empowered WiFi human sensing. *Patterns* 4, 3 (2023), 10073.

[76] Guolin Yin, Junqing Zhang, Guanxiong Shen, and Yingying Chen. 2022. Fewsense, Towards a Scalable and Cross-domain Wi-Fi Sensing System Using Few-shot Learning. *IEEE Transactions on Mobile Computing* 23, 1 (2022), 453–468.

[77] Ruonan Yu, Songhua Liu, and Xinchao Wang. 2024. Dataset Distillation: A Comprehensive Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 46, 1 (2024), 150–170.

[78] Camellia Zakaria, Youngki Lee, and Rajesh Balan. 2021. Detection of Social Identification in Workgroups From a Passively-sensed WiFi Infrastructure. *Proc. of the ACM HCI* 5, 71 (2021), 1–19.

[79] Youwei Zeng, Dan Wu, Jie Xiong, Jinyi Liu, Zhaopeng Liu, and Daqing Zhang. 2020. MultiSense: Enabling Multi-Person Respiration Sensing with Commodity WiFi. In *Proc. of the 22nd UbiComp*. 102:1–29.

[80] Jin Zhang, Yuyan Dai, Jie Chen, Chengwen Luo, Bo Wei, Victor CM Leung, and Jianqiang Li. 2023. SIDA: Self-Supervised Imbalanced Domain Adaptation for Sound Enhancement and Cross-Domain WiFi Sensing. *Proc. of the ACM IMWUT* 7, 3 (2023), 1–24.

[81] Jie Zhang, Zhanyong Tang, Meng Li, Dingyi Fang, Petteri Nurmi, and Zheng Wang. 2018. CrossSense: Towards Cross-site and Large-scale WiFi Sensing. In *Proc. of the 24th MobiCom*. 305–320.

[82] Jin Zhang, Bo Wei, Wen Hu, and Salil S Kanhere. 2016. WiFi-ID: Human Identification using WiFi Signal. In *Proc of the 12nd IEEE DCOSS*. 75–82.

[83] Lei Zhang, Zhirui Wang, and Liu Yang. 2019. Commercial Wi-Fi based Fall Detection with Environment Influence Mitigation. In *Proc. of the 16th IEEE SECON*. 1–9.

[84] Wei Zhang, Siwang Zhou, Dan Peng, Liang Yang, Fangmin Li, and Hui Yin. 2020. Understanding and Modeling of WiFi Signal-based Indoor Privacy Protection. *IEEE Internet of Things Journal* 8, 3 (2020), 2000–2010.

[85] Yunhua Zhang, Hazel Doughty, Ling Shao, and Cees G. M. Snoek. 2022. Audio-Adaptive Activity Recognition Across Video Domains. In *Proc. of the 35th IEEE/CVF CVPR*. 13791–13800.

[86] 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* 44, 11 (2021), 8671–8688.

[87] Bo Zhao, Konda Reddy Mopuri, and Hakan Bilen. 2021. Dataset Condensation with Gradient Matching. In *Proc. of the 9th ICLR*. 1–20.

[88] Yue Zheng, Yi Zhang, Kun Qian, Guidong Zhang, Yunhao Liu, Chenshu Wu, and Zheng Yang. 2019. Zero-Effort Cross-Domain Gesture Recognition with Wi-Fi. In *Proc. of the 17th ACM MobiSys*. 313–325.

[89] Siwang Zhou, Wei Zhang, Dan Peng, Yonghe Liu, Xingwei Liao, and Hongbo Jiang. 2020. Adversarial WiFi Sensing for Privacy Preservation of Human Behaviors. *IEEE Communications Letters* 24, 2 (2020), 259–263.

[90] Yuxuan Zhou, Huangxun Chen, Chenyu Huang, and Qian Zhang. 2022. WiAdv: Practical and Robust Adversarial Attack Against WiFi-based Gesture Recognition System. *Proc. of the ACM IMWUT* 6, 2 (2022), 1–25.