Toward privacy in IoT mobile devices for activity recognition

Théo Jourdan, Antoine Boutet, Carole Frindel

Toward privacy in IoT mobile devices for activity recognition

Théo Jourdan
Univ Lyon, INSA Lyon, CNRS, Inserm, CREATIS UMR 5220, U1206, F-69621 VILLEURBANNE, France
theo.jourdan@creatis.insa-lyon.fr

Antoine Boutet
Univ Lyon, INSA Lyon, Inria, CITI, F-69621 VILLEURBANNE, France
antoine.boutet@insa-lyon.fr

Carole Frindel
Univ Lyon, INSA Lyon, CNRS, Inserm, CREATIS UMR 5220, U1206, F-69621 VILLEURBANNE, France
carole.frindel@creatis.insa-lyon.fr

ABSTRACT

Recent advances in wireless sensors for personal healthcare allow to recognise human real-time activities with mobile devices. While the analysis of those datastream can have many benefits from a health point of view, it can also lead to privacy threats by exposing highly sensitive information. In this paper, we propose a privacy-preserving framework for activity recognition. This framework relies on a machine learning technique to efficiently recognise the user activity pattern, useful for personal healthcare monitoring, while limiting the risk of re-identification of users from biometric patterns that characterizes each individual. To achieve that, we first deeply analysed different features extraction schemes in both temporal and frequency domain. We show that features in temporal domain are useful to discriminate user activity while features in frequency domain lead to distinguish the user identity. On the basis of this observation, we second design a novel protection mechanism that processes the raw signal on the user's smartphone and transfers to the application server only the relevant features unlinked to the identity of users. In addition, a generalisation-based approach is also applied on features in frequency domain before to be transmitted to the server in order to limit the risk of re-identification. We extensively evaluate our framework with a reference dataset: results show an accurate activity recognition (87%) while limiting the re-identifation rate (33%). This represents a slightly decrease of utility (9%) against a large privacy improvement (53%) compared to state-of-the-art baselines.

KEYWORDS

Activity Recognition, Privacy, IoT Healthcare

1 INTRODUCTION

The emergence of medical Internet of Things (IoT) devices have paved the way for personal healthcare monitoring at home or in hospital environments. These devices record electronic health measurements from a variety of sensors (most commonly an accelerometer, a gyroscope and a magnetometer) and send these patient data to an application server to be processed and analysed. These processing and analysis include for instance advanced signal processing and machine learning algorithms to provide a variety of services such as (1) motion tracking: number of steps, burned calories, traveled distance and sleep monitoring and (2) vital signs measurement: heart rate, skin temperature, electrocardiogram (ECG) and electroencephalogram (EEG) [18].

Due to their nature, collected data from medical IoT devices are highly sensitive. Advances in wireless communication and web technologies facilitate the remote real-time monitoring of such systems [39]. However, the complex workflow of collected medical data multiplies the security and privacy risks all along the life-cycle of the data including the data collection and transmission [3, 38], as well as the processing and the storage [31]. When such medical data can be accessed by an adversary, risks of privacy threats like leakages of sensitive information or user re-identification are very high (e.g., the re-identification of Governor William Weld's medical information [22]).

In the context of activity recognition through mobile devices, the challenge is to identify data that can preserve the privacy of individuals while still being relevant enough for machine learning tasks [33]. This challenge raises two important questions: 1) Is the collected data protected enough so that no one can misuse it to infer sensitive information or to re-identify the owner? 2) How to assess whether the protected data are still accurate enough for researchers in the health domain? Achieving this balance between data utility and data privacy is an important objective to send secure and reliable data through mobile devices and to strengthen end-user confidence and adoption.

In this paper, we propose a privacy-preserving framework for activity recognition from mobile devices. This framework relies on a machine learning technique to efficiently recognise the user activity pattern, useful for personal healthcare monitoring, while limiting the risk of re-identification of users from biometric patterns that characterizes each individual. To achieve that, firstly we extracted multiple features from raw signal and deeply analysed their impact on both the activity recognition and the user re-identification. We show that features in temporal domain are useful to discriminate user activity while features in frequency domain lead to discriminate the user identity.

Based on this observation, we design a novel privacy-preserving framework. In this framework, data records are processed locally on the user device and only relevant features are extracted. Additionally, features in the frequency domain (i.e., features leading to discriminate users) are normalized. This normalization can be viewed as a generalization-based approach. However compared to other generalization-based approaches based on k-anonymity that are well known to drastically reduce the utility of the protected data [15], our solution keeps a high utility (i.e. activity recognition) while providing a good privacy (i.e. small user re-identification).

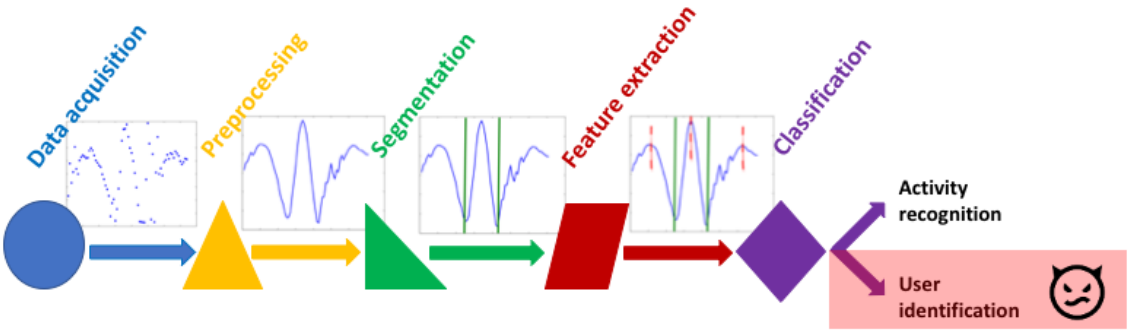


Figure 1: Traditional IoT healthcare workflow for activity recognition, an adversary can misuse the classifier to re-identify users.

Once normalized, this information are periodically upload to the application server. Each batch of features is stored independently on the server (i.e., with a different pseudonym) to avoid to link both batches to individuals and batches together. Moreover, to avoid centralizing both the data and the associated identity of their owners on the same node, the mapping between the pseudonyms and the user identities is only retained by the hospital practitioners.

We exhaustively evaluated our framework with the use of a reference dataset. Results show an accurate activity recognition of 87% in average while limiting the user re-identification rate up to 33%. We also compared our solutions against different baselines. Our solution provides a better privacy-utility trade-off with a slightly decrease of utility (9%) against a large increase of privacy (53%).

Our contributions can be summarized as follow:

- We quantify both the risk assessment associated to the re-identification of users (90% in average) and the capacity to detect the user activity (97% in average) from signal from mobile devices.
- We deeply analysed the impact of multiple features on both the activity recognition and the user re-identification. We show that features in the temporal domain tend to discriminate the user activity while features from the frequency domain tend to discriminate users.
- We propose an efficient workflow and machine learning technique to recognise user activity with high utility while limiting the risk of user re-identification. Our solution provides a better privacy-utility trade-off with a slightly decrease of utility (9%) against a large increase of privacy (53%) compared to state-of-the-art baselines.

In this paper, we present background on IoT healthcare workflow in Section 2 before to define the adversary model in Section 3. We then quantify and analyze the capacity of both recognizing the activity of user and their identity in Section 4. Section 5 details our privacy-preserving framework and Section 6 presents its evaluation. Finally, related work is reviewed in Section 7 before to conclude in Section 8.

2 BACKGROUND ON IOT HEALTHCARE WORKFLOW

This section explains the methodology we followed for activity recognition and user re-identification using IoT mobile devices. Although this description is specific to our methodology, it is typical and provides background on IoT healthcare workflow. Figure 1 depicts the whole workflow including data acquisition (Section 2.1), signal preprocessing (Section 2.2), segmentation (Section 2.3), feature extraction (Section 2.4), and classification (Section 2.5).

2.1 Data acquisition

Data acquisition relies on sensors that are present in IoT devices, such as smartphones, smartwatches, smart wristbands, tablets and medical sensors. There exist a variety of sensors that allow the acquisition of various types of data, which can then be used for different types of tasks. For the recognition of physical activities, the authors in [29] propose to use of inertial sensors, i.e., accelerometers and gyroscopes, complemented with orientation measurement using magnetic sensors, e.g., a compass and a magnetometer, and location measurement using location sensors, e.g., a global positioning system (GPS).

The data acquisition process is accomplished by a specific module in the mobile device and consists of the measurement and conversion of the electrical signals received by each sensor into a readable format [32]. Several challenges are associated with the data acquisition process when recognizing physical activities, including the positioning of the mobile device, the data sampling rate and the number of sensors to be used and hence managed [7]. All these factors directly influence the correct extraction of meaningful features. As the sensors are embedded in the mobile device, they cannot be located separately in different parts of the body; rather, the mobile device needs to be situated in a usual and comfortable position. Another issue related to mobile devices is the power consumption of the data acquisition tasks. Multitasking execution patterns differ among mobile devices, because these

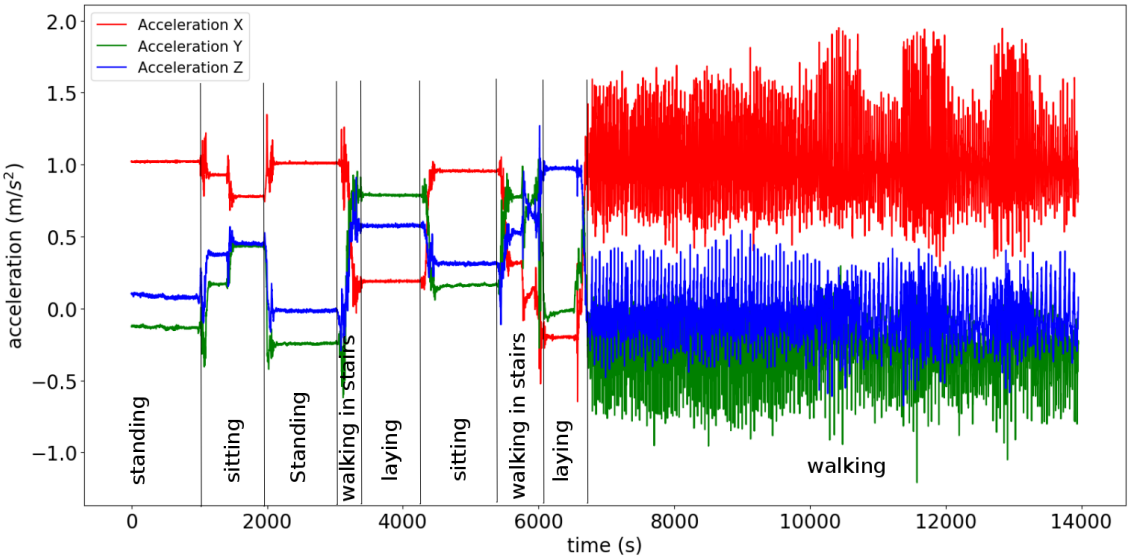


Figure 2: Visualization of accelerometer signals in x, y and z dimensions and associated activities.

depend on their processing ability, memory and power capabilities and on the operating system and on the number and type of mobile applications currently installed and/or running. The selection of the best data acquisition methods depends on the purpose of use, the type of data acquired and their environment [12, 28].

2.2 Signal preprocessing

Sensor signals are typically preprocessed by the application of a series of filters. First, noise was reduced with a median filter and a third order low-pass Butterworth filter with a cutoff frequency of 20 Hz. This frequency threshold was selected from the work presented in [21] which states that the energy spectrum of the human body motion is below 15 Hz. The resulting signals were further filtered to break them down into channels that make sense from a physical point of view as displayed in Figure 3. For example, linear acceleration signal was decomposed in two principal channels: gravitational and body motion components. This step was performed using another low-pass filter and assuming that the gravitational component mainly refer to the lowest frequencies [2]. Subsequently, body motion acceleration and gyration signals were derived in time to obtain jerk that reflect the temporal variations of the signals. Finally, signals were decomposed according to their acquisition axes (x, y, z, respectively) in order to observe them in a specific direction (vertical, lateral or longitudinal) as depicted Figure 2. The magnitude of associated signals has also been calculated to produce an average signal less sensitive to how the device is fixed on the person. This filtering step allowed to reach 20 channels in total.

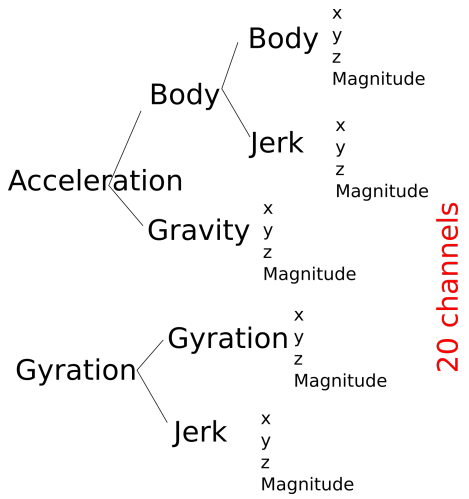


Figure 3: Channels considered for feature extraction.

2.3 Segmentation

Channel signals are typically segmented using a fixed sliding window technique. Windows with a span of 2.5 seconds and an overlap of 50% were captured. An overlap degree of 50% means that the window is shifted by half of its size, in other words 50% of the previous data are included in the next window. The choice of the window size is not trivial especially for an activity recognition algorithm. A small window size could split an activity signal while large window size could contain multiple activity signals. We decided to calibrate our window size on the most complex activity: walking. Hence, the window size has been chosen to take into account at least a full walking cycle of two steps: the cadence range of an

	Function	Description	Formulation
Time domain	mean $(\boldsymbol{s})$	Arithmetic mean	$\bar{s} = \frac{1}{N}\sum_{i=1}^{N} s_i$
	std $(\boldsymbol{s})$	Standard deviation	$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(s_i - \bar{s})^2}$
	mad $(\boldsymbol{s})$	Median absolute deviation	$\text{median}_i\left(\lvert s_i - \text{median}_j(s_j)\rvert\right)$
	max $(\boldsymbol{s})$	Largest values in array	$\max_i(s_i)$
	min $(\boldsymbol{s})$	Smallest value in array	$\min_i(s_i)$
	sma $(\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3)$	Signal magnitude area	$\frac{1}{3}\sum_{i=1}^{3}\sum_{j=1}^{N}\lvert s_{i,j}\rvert$
	iqr $(\boldsymbol{s})$	Interquartile range	$\text{Q3}(\boldsymbol{s}) - \text{Q1}(\boldsymbol{s})$
	autoregression $(\boldsymbol{s})$	4th order Burg Autoregression coefficients	$\boldsymbol{a} = \text{arburg}(\boldsymbol{s}, 4), \boldsymbol{a} \in \mathbb{R}^4$
	correlation $(\boldsymbol{s_1}, \boldsymbol{s_2})$	Pearson Correlation coefficient	$C_{1,2}/\sqrt{C_{1,1}C_{2,2}}, C = \text{cov}(\boldsymbol{s}_1, \boldsymbol{s}_2)$
	angle $(\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, v)$	Angle between triaxial signal mean and vector	$\tan^{-1}\left(\lVert[\bar{s}_1, \bar{s}_2, \bar{s}_3] \times \boldsymbol{v}\rVert, [\bar{s}_1, \bar{s}_2, \bar{s}_3] \cdot \boldsymbol{v}\right)$
Frequency domain	skewness $(\boldsymbol{s})$	Frequency signal Skewness	$\text{E}\left[\left(\frac{\boldsymbol{s}-\bar{s}}{\sigma}\right)^3\right]$
	kurtosis $(\boldsymbol{s})$	Frequency signal Kurtosis	$\frac{\text{E}\left[(\boldsymbol{s}-\bar{s})^4\right]}{\text{E}\left[(\boldsymbol{s}-\bar{s})^2\right]^2}$
	maxFreqInd $(\boldsymbol{s})$	Largest frequency component	$\arg\max_i(s_i)$
	energy $(\boldsymbol{s})$	Average sum of the squares	$\frac{1}{N}\sum_{i=1}^{N} s_i^2$
	entropy $(\boldsymbol{s})$	Signal Entropy	$\sum_{i=1}^{N}(c_i \log(c_i)),\ c_i = s_i / \sum_{j=1}^{N} s_j$
	meanFreq $(\boldsymbol{s})$	Frequency signal weighted average	$\sum_{i=1}^{N}(is_i) / \sum_{j=1}^{N} s_j$
	energyBand $(\boldsymbol{s}, a, b)$	Spectral energy of a frequency band $[a, b]$	$\frac{1}{a-b+1}\sum_{i=a}^{b} s_i^2$

Figure 4: List of measures for computing feature vectors. N: signal vector length, Q: quartile.

Xacc_body_iqr	Xacc_body_max	Xacc_body_mean	Xacc_body_med	Xacc_body_min	Xacc_body_ropy	Xacc_body_std	...	pers	act
0.77666792327	1.01481659060	0.32071585656	0.34988767835	-0.49054102707	4.7489565343	0.4194744014	...	10	2
0.66693512370	1.43263647481	0.26841672908	0.43411692212	-1.41238613704	4.7722314297	0.6610481443	...	10	3
1.02907915173	1.43263647481	-0.10075092775	0.08232560553	-1.42686548654	4.7899910551	0.6334978541	...	10	3
0.23557396729	0.74911155782	0.33652443467	0.26582976888	0.10360618631	4.8056637254	0.1568877474	...	10	4
0.35584093169	0.78654658654	0.21654485464	0.27026656791	-0.72443435405	4.7194139887	0.3592030590	...	10	4

Figure 5: A sample dataset with features and labels, input of the classification step.

average person walking corresponds to minimum speed of 1.5 steps by second according to [6].

2.4 Feature Extraction

From each window of each channel signal, a feature vector was extracted which contained 17 measures estimated in the time and frequency domains respectively. The Discrete Fourier Transform (DFT) was used to extract the descriptors of each window in the frequency domain. The choice of these descriptors was made on the basis of an earlier review on effective descriptors for gait recognition [34] : e.g. for time domain mean, standard deviation (STD), signal magnitude area (SMA) and signal-pair correlation (Corr); and for frequency domain energy and entropy. The selected measures to obtain the feature vector are depicted in Figure 7. A feature vector was calculated from each experiment window sample and labeled according to the user and activity it belongs. Figure 5 shows an example of the dataset format, where lines correspond to window samples and columns to features (except the two last ones which correspond to the labels). Such dataset is used as an input for the classification task. A total of 340 features (20 channels x 17 measures) are extracted. The notation for naming a descriptor in the rest of this article is the following $\{orientation\}\_\{channel\}\_\{descriptor\}$.

2.5 Classification

2.5.1 Machine learning algorithm. Random Forest (RF for short) was chosen for the multi-class classification tasks, respectively classes referring to activity recognition and classes associated to user identities in case of an adversary willing to misuse the classifier to re-identify users. In general, the RF algorithm is a supervised classifier having fast training time and very high performance without fine tuning [25]. RF operates by building a large ensemble of decision trees, where

each tree is built on a bootstrapped sample of the original data [10]. The classification trees are built based on recursive binary splits: for each split, a randomly-chosen subset of input variables is used to find the optimal binary split that corresponds to a condition on a feature. The optimal splits are determined using the Gini impurity index [20]. The function "RandomForestClassifier" in the Python Scikit Learn package [26] was used for constructing the RF classifier. In this work, according to the instances and features of our classification problem, 700 is chosen as the number of trees in the forest, $\sqrt{n}$ random features are considered in building each tree and 10 is set as the maximum depth of each tree.

2.5.2 Utility and privacy measures. To measure the classification quality based on the proposed features with RF, we computed the accuracy from the confusion matrix [19]:

$$Accuracy = \frac{|TP| + |TN|}{|TP| + |TN| + |FP| + |FN|},$$

where $|TP|$ (True Positive): is the number of correct predictions for a specific event value, $|TN|$ (True Negative): is the number of correct predictions for non-event values, $|FP|$ (False Positive): is the number of incorrect of predictions for a specific event value, and $|FN|$ (False Negative): is the number of incorrect predictions for non-event values.

Accuracy reflects the number of correct predictions made by the model over all kinds predictions made. Accuracy is comprised in $[1 : 0]$ where a value of 1 corresponds to a perfect prediction. We use this metric to compute the quality of our classification to predict both the activity of the user and the user identity. We called *Accuracy(activity)* the result when it is applied to the activity recognition, and we call *Accuracy(re-identification)* the result when it is applied to the user identity. We prefer the accuracy rather than the f-score because the variable classes in the data are nearly balanced.

Algorithm 1: Feature selection

Input : List of features sorted by importance f and associated initial accuracy a; $threshC = 0.7$; $threshA = 0.03$
Output : List of selected features

1 **for** *each feature* $f_i \in f$ **do**
2 Compute the Pearson correlation values C for each feature in $\{f - f_i\} : fcorre$
3 **for** *each feature* $f_j \in fcorre$ **do**
4 **if** $| C(f_j) | > threshC$ **then**
5 Compute accuracy $newA$ of classification for $\{f - f_j\} : newa$
6 **if** $a - newA < threshA$ **then**
7 Erase feature f_j from f
8 **end**
9 **end**
10 **end**
11 **end**

2.5.3 Feature ranking and selection. The RF algorithm can be used to rank features according to their importance in the classification. When training a tree, it can be computed how much each feature decreases the Gini impurity index [20] in a tree. For a forest, the impurity decrease from each feature can be averaged and the features are ranked according to this measure.

The RF algorithm can also be used for feature selection [10]. This is done via measuring the mean decrease of accuracy when a particular feature is removed from the set of features in the trees. If the accuracy deterioration after feature exclusion is negligible, the feature is less important and vice versa. The importance scores of the features in the RF classifier [10, 16] can therefore be evaluated and used as a feature selection criteria. For more details, see the Algorithm 1: It consists of two nested loops, one corresponding to features ranked by importance (line 1) and one corresponding to features correlated to each of the features of the first loop (line 3). The correlation is calculated using the Pearson coefficient (line 2). If the correlation between two features is greater than a certain threshold (line 4), then the accuracy of the random forest algorithm is recalculated after removal of the correlated feature (line 5) and if the corresponding decrease in accuracy is below a certain threshold (line 6) this feature is eliminated for good (line 7).

3 ADVERSARY MODEL

Before presenting our privacy-preserving framework in Section 5, we describe our assumptions and the adversary model against which our solution is designed. The framework presented in this paper involves three premises: the client running on the smartphone of users, the application server storing the features and performing the classification, and the hospital practitioner monitoring the patient activity. First, we assume that the client application and the smartphone on which it is run are trusted. This means that the data acquisition, the preprocessing, the segmentation, the feature extraction, and the normalisation cannot deviate from a correct behaviour. Moreover, we do not consider limitation on the sampling rate of the data acquisition as in [35].

Second, we assume that the application server runs on public cloud platforms. We consider that this cloud platform is honest but curious [14]. This means that the application server behaves correctly when it comes to processing data received from clients. More precisely, this means that the data is stored correctly in the database, that no forged information can be injected in the database, and that the classifier model cannot be maliciously tampered. However, we assume that the adversary is able to collect part or the entire information stored in the database. Each information corresponds to independent batches of data unlinked to users (i.e., with a different random pseudonym for each batch). Additionally, we assume that the adversary is able to collect data relative to the gestures of each user from a malicious IoT device for instance. This prior knowledge on each user is used by the adversary to build a classifier model. This classifier exploits the same preprocessing, segmentation, and features than our classifier but with the objective to predict the identity of the user for each batch of data stored in the database.

Activity	Accuracy(activity)
Walking	0.97
Walking upstairs	0.95
Walking downstairs	0.94
Sitting	0.97
Standing	0.98
Laying	0.99

Table 1: User activities can be recognised with a high success rate (recognition using the methodology presented Section 2).

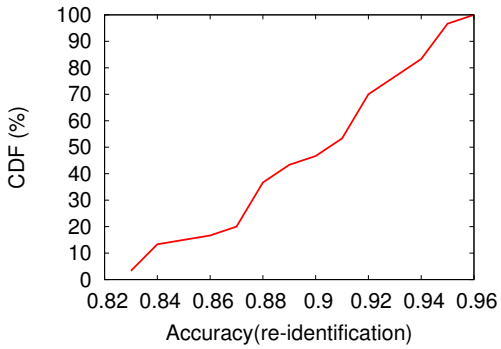


Figure 6: Cumulative distribution of the accuracy for the user re-identification task: users can be easily re-identified from their data.

Third, we assume that the server used by the hospital practitioner is trusted. This server is used to store the mapping between the batches of data sent to the application server and the identity of the users.

Lastly, all communications between nodes (i.e., clients, the application server, and server of the hospital practitioner) are secured. We assume that no information can be inferred from these secured communications.

4 QUANTIFYING ACTIVITY RECOGNITION AND USER RE-IDENTIFICATION

We carried out an extensive evaluation of the capacity to recognise the activity of users and to re-identify them. We show that following the methodology described in Section 2, we are able to predict the activity of the user with a very high rate of success. In addition, we show that without any protection scheme, data from mobile devices act as a personal fingerprint and lead to re-identify users. We first describe the dataset used in this evaluation in Section 4.1 before to quantify the activity recognition and the user re-identification in Section 4.2 and Section 4.3, respectively. Finally, we analyse the impact of extracted features in Section 4.4.

4.1 Dataset

The dataset used in this work is available online for public use as the "Human Activity Recognition using Smartphones" dataset in the UCI Machine Learning Repository [2]. It is composed of the 3-axial raw data from accelerometer and gyroscope sensors read at a constant frequency of 50 Hz. A group of 30 volunteers were selected to follow a protocol of activities while wearing a smartphone on waist. The experiment was planned in order to contain six basic activities: three static postures (standing, sitting, lying-down) and three ambulation activities (walking, walking-downstairs and walking-upstairs). Figure 2 displays accelerometer signal of one of the experiments and the associated activities. The protocol of activities is detailed in [30]. The duration of an entire experiment was around 15 minutes and was repeated ten times. All the experiments were recorded on video to have a ground truth to annotate the performed activities on acceleration and gyration signals.

4.2 Activity Recognition

Table 3 summarizes the accuracy for the recognition of the different activities. Results show that our machine learning framework is able to highly recognise activities with an average accuracy of 0.97. As the table indicates, the accuracy is lower for ambulatory activities in stairs. A possible explanation for this is that these activities correspond to the smallest acquisition times (Figure 2).

4.3 User Re-Identification

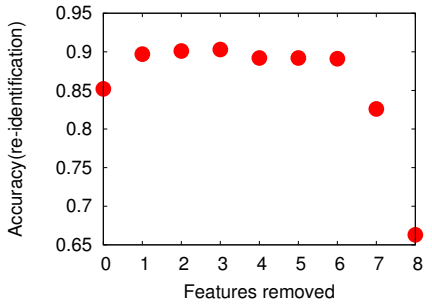
Figure 6 depicts the cumulative distribution of the accuracy for the user re-identification task. Accuracy ranges from 0.82 to 0.96 among the 30 users with an average of 0.90. These results indicate that the data collected from the gesture of users characterizes each individual and can lead to re-identify them with a high success rate. However, the task of re-identification is slightly more difficult than that of recognizing activities with lower accuracy.

4.4 Impact of Features

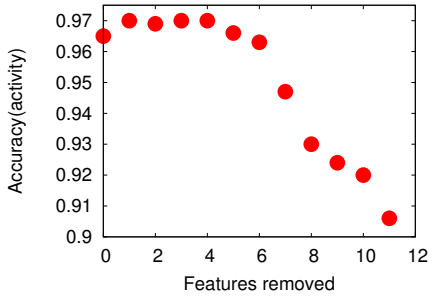
Features	Importance
Y_grav_std	0.175
Z_grav_med	0.163
Z_grav_energy	0.137
X_grav_max	0.128
Magn_grav_max	0.123
Y_gyro_mean	0.107
Y_gyro_irq	0.088
Y_body_zcross	0.079

Table 2: Most important features for user re-identification (features in the frequency domain are in grey).

The previous experiments are also used to rank features (from the 340) according to their importance. Eight and eleven features were respectively selected for the activity recognition and user re-identification tasks given the correlation and accuracy analysis (see Algorithm 1 for methodology and Tables 2 and 3 for results). Indeed, many features are



(a) Privacy



(b) Utility

Figure 7: Impact of the number of features (depicted in Table 2 and Table 3) retained in the RF learning process on user's privacy and utility metric (features were sorted by increasing order of importance).

Features	Importance
X_grav_max	0.144
X_grav_min	0.127
Magn_grav_max	0.109
X_gyro_min	0.104
X_body_var	0.098
Magn_body_var	0.085
X_gyro_max	0.082
Y_gyro_irq	0.078
X_gyro_mean	0.077
Magn_gyro_mean	0.074
Y_body_entropy	0.020

Table 3: Most important features for activity classification (frequency-based features are in Grey).

alike and contain similar information on the original sensor data. Compared to using all 340 features, using only these 19 relevant features lowers only slightly ($< 4\%$) the two classification tasks performance (97% vs 96% for activity classification and 90% vs 86% for user re-identification). This can be observed more precisely in the Figures 7a and 7b, where the importance of each selected feature is independently tested for the task of interest: there is a strong correlation between the importance of a specific feature and the performance of the RF algorithm after removing it.

Based on these ranking results, it is interesting to note that the task of activities recognition (i.e., utility) is almost exclusively (9 of the 11 selected features) operated in the time domain whereas the task of user identification (i.e., privacy) is based (5 of the 8 selected features) on features in the frequency domain. These results can be explained by the fact that the activities are mainly distinguished from each other by their level of amplitude in acceleration and gyration (Figure 2) and therefore their associated statistics. Conversely, the user identification is more related to the pace or cadence at which this person performs the activity and is strongly related to biomechanics (e.g., age, size, weight).

5 PRIVACY-PRESERVING ACTIVITY RECOGNITION FRAMEWORK

To ensure privacy, our framework relies on both an architecture limiting the exposure of sensitive information and a data normalisation applied on features leading to re-identify user (Section 4.4). These normalisations act as a form of generalisation-based obfuscation. In this section, we first present the architecture of our framework (Section 5.1) before to describe the normalisation of each sensitive feature (Section 5.2).

5.1 Architecture

The design of our privacy-preserving framework comprises three main elements: a client application running on the user smartphone communicating with its IoT environment, the application server, and the hospital practitioner. To limit the exposition of sensitive information, the application server does not store identified data but only batches of features where each batch is randomly pseudoanonymized. Only the hospital practitioner retained the mapping between the user identities and pseudonyms, and requests the application server to monitor the activity of users.

The architecture of our privacy-preserving activity recognition framework is depicted Figure 8. Firstly, IoT devices (e.g. smartwatch) or directly the smartphones perform the data acquisition (❶). In both cases, these raw data are stored locally on the smartphone. The client application then performs the preprocessing, the segmentation and the features extraction following the methodology described in Section 2. On the basis of our analysis on the importance of features, this feature extraction only concerns the 19 features identified as important (Section 4.4). Moreover, the client conducts the normalisation of the features identified as leading to the re-identification of users. All these normalisations are described in the following sub-section. As all the aforesaid actions performed on the smartphone only concern the associated user on one batch of data (i.e., one day for instance), the resulting computational cost is cheap. On a commodity computer, these operations applied on all the data of one user spend 2.5 seconds in our experiments. Secondly, the client application associates a random pseudonym to each

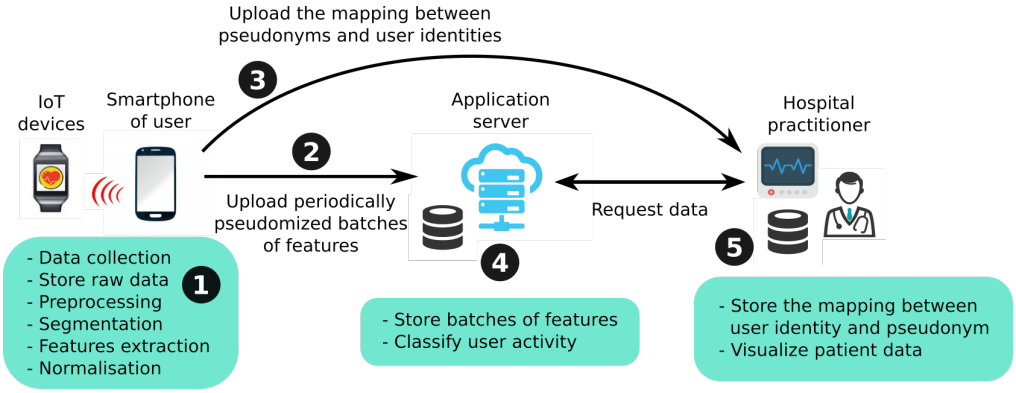


Figure 8: Architecture of our framework: the user smartphone is leveraged to extact relevant features and only these features are uploaded periodically to the application server.

timestamped batch of features before to periodically upload them to the application server (❷). The client application then sends to the hospital practitioner the list of pseudonyms associated to its identity (❸).

When a batch of features is received by the application server, it stores this information in a database (❹). Consequently, each batch in this database does not contain the identity of the user but a random pseudonym. The application server then periodically performs the classification to detect the activity associated to each batch of features.

Finally, when the hospital practitioner wants to monitor the activity of a specific users, firstly it retrieves locally all the pseudonyms associated to the specified user and then requests the application server to have the activity history of the specified pseudonyms (❺).

5.2 Normalisation

In order to limit the re-identification of users, we propose a normalisation scheme which generalises the effect of the different descriptors identified as important for the task of user re-identification. In other words, we try to mitigate their characteristics allowing the re-identification of the user without removing them completely because they also have an impact on the recognition of activities. Given the data from the sensors noted S and of size n, applying the normalisation approach on S will output the so-called "normalised data" noted S^*. In this work, we distinguished five normalisations, each of them referring to the features in the frequency domain listed in Table 2. Regarding the temporal features, we simply delete them.

5.2.1 Normalisation by mean. (Y_gyro_mean)

$$S_i^* = S_i - \mu + \mu^*, \quad i \in [0, n], \tag{1}$$

with μ and μ^* being respectively the data means before and after normalization.

5.2.2 Normalisation by interquantile range. (Y_gyro_irq) The interquantile range (IQR) is a measure of statistical dispersion, being equal to the difference between 75th and 25th percentiles.

$$S_i^* = \frac{S_i}{IQR} IQR^*, \quad i \in [0, n], \tag{2}$$

with IQR and IQR^* being respectively the data interquartile ranges before and after normalisation.

5.2.3 Normalisation by standard deviation. (Y_grav_std)

$$S_i^* = \frac{S_i}{\sigma} \sigma^*, \quad i \in [0, n], \tag{3}$$

with σ and σ^* being the data standard deviations before and after normalisation.

5.2.4 Normalisation by root mean square. (Z_grav_energy)

$$S_i^* = \frac{S_i}{\sqrt{\frac{1}{n}\sum_{j=1}^{n} S_j^2}}, \quad i \in [0, n]. \tag{4}$$

5.2.5 Normalisation by maximum and minimum. (X_grav_max)

$$S_i^* = (S_i - Min)\frac{newMax - newMin}{Max - Min} + newMin, \quad i \in [0, n], \tag{5}$$

with Max and Min being respectively the maximum and minimum of the original data and $newMax$ and $newMin$ the maximum and minimum of the normalised data.

The reference values after normalisation for mean, IQR, standard deviation and newMin and newMax were chosen by taking the average of the values before normalisation.

6 EVALUATION OF OUR FRAMEWORK

We carried out an extensive evaluation of our framework. In this section, we start with a description of the comparison baselines (Section 6.1) before evaluating the performance of our approach in term of utility-privacy trade-off (Section 6.2).

6.1 Comparison Baselines

To highlight the benefits of our approach, we compare the performance of our framework with that of two alternatives. The first alternative follows a perturbation scheme. Similarly

to [1] that applies a perturbation scheme in the frequency domain of aggregated time series in the context of location privacy, this alternative (called *perturbation*) adds a Gaussian noise in the signal in frequency domain before the extraction of features. The second alternative is based on simply the removing of features identified as leading to the user re-identification (Section 4.4). The incentive behind this alternative (called *suppression*) is that without these features, the re-identification is harder.

6.2 Privacy Improvement

Figure 9 reports for our solution and the baseline approaches the trade-off between the utility captured by the accuracy to recognise the activity and the privacy captured by the accuracy to re-identify users. For the baseline based on the suppression of features, each point of the curves corresponds to the deletion of a feature (from the 8 selected ones for the re-identification task). For the baseline based on perturbation, in turn, each point refers to the addition of an increasing fixed amount of noise (noise is centered on zero and its standard deviation is, for each point, increased by 2). Finally, in our framework, each point corresponds to the normalisation of a growing number of features (in order of increasing importance).

Results show that the suppression approach (slope: 0.12) seems the most advantageous in terms of compromise between utility and privacy. However it is very quickly limited by the number of selected features and therefore in privacy and utility metrics; for instance the best obtained performance are respectively 0.66 and 0.93. The perturbation approach (slope: 0.34) is very effective in loss of identification however at the cost of a very important loss of utility too, with for best performance in privacy and utility metrics respectively 0.51 and 0.84. Our approach is between the two (slope: 0.21) and provides the best utility and privacy trade-off (respectively 0.87 and 0.33). Our approach based on normalisation gives a better control on the weight of each feature in the protection, unlike the suppression approach for which limits their impact to consideration or not.

Lastly, we also considered an adversary that trains a classifier only with features leading to the re-identification (Table 2), in this case the accuracy in term of re-identification is less efficient than with our framework (0.17).

7 RELATED WORK

With the technological advances of recent years, the medical domain is changing fast raising important privacy issues. For instance, new high throughput DNA sequencing technologies have drastically reduced the price and democratized DNA analysis. Due to the highly sensitive nature of this data, an important research area has emerged to address the quantification of the risk associated to this information and to protect it [5, 36]. The widespread adoption of medical IoT has also introduced new security and privacy questions and concerns. These security and privacy concerns emerge at multiple stages in the life-cycle of the data [31]. In the data transmission for

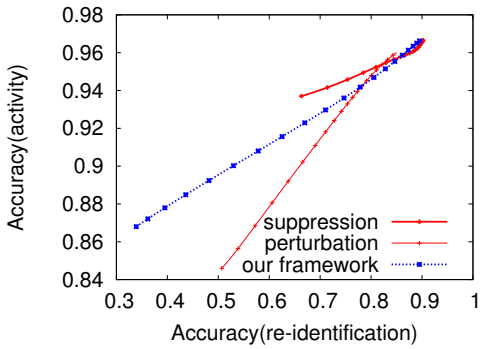


Figure 9: Our framework provides a better utility and privacy trade-off than baseline approaches.

example, [38] proposed a method to capture network traffic from medical IoT devices and automatically detect clear-text information that may reveal sensitive medical conditions and behaviors. [3], in turn, presented PDI, a framework which aims to prevent an adversary from inferring certain sensitive information about subjects using the encrypted data that they disclosed during communication with an intended recipient. Other approach such as the NeuroSENS architecture [13] tries to improve the security and the privacy of neurogical gait monitoring at several levels (data storage, mobile and web apps and data transmission). Although gesture recognition attracts many attention currently [37], to the best of our knowledge, our work is the first one that addresses the protection of data dedicated to activity recognition through wearable devices in the medical domain. The identification of relevant features for both the activity recognition and the user re-identification is also novel.

Several well known reported user re-identifications have shown that hiding explicit identity information through pseudonymity is not enough to guarantee the anonymity of users [22]. Indeed, many criteria lead to uniquely identifying users. Previous researches have shown that individuals can be identified from their mobility [8, 23], their touch-based gestures on touch-screen devices [24], or their Web browsers [11] to name a few. Following these studies, we also demonstrate in this paper that an user can be easily identified from its gestures collected by sensors.

Compared to other approaches that obfuscate independently every record [4], only features leading to the re-identification of users are obfuscated. In addition, although this obfuscation based on a normalization does not provide the same privacy guaranty as other generalization-based approaches ensuring k-anonymity, the utility (i.e., activity recognition) remains high while providing a good privacy (i.e., a small re-identification rate).

Lastly, splitting sensitive information (i.e., both the identity of users and their data) on different nodes have already showed its benefits in terms of privacy [17, 27]. In addition, by processing the signals at the edge of the network on the smartphone of users, our framework inherently reduces the

operational costs of the application [9] and strengthens the control of users on their data.

8 CONCLUSION

We present a privacy-preserving IoT framework in the context of activity recognition for healthcare monitoring with wearable devices. Our framework processes the signal and extracts relevant features locally on the user smartphone. In addition, accordingly to the observation that the frequency domain prevails in the user identification task, a normalization is performed on the frequency-based features to obfuscate the re-identification of users. Finally, only a set of features unlinked to the identity of its owner is uploaded to the application server which is then able to recognise the activity of the users with a high accuracy while reducing the risk of user re-identification. An extensive validation of our framework has been performed on reference data sets yielding good results in terms of privacy-utility trade-off: a high activity recognition with few user re-identification.

REFERENCES

[1] G. Acs and C. Castelluccia. 2014. A case study: Privacy preserving release of spatio-temporal density in paris. In *KDD*. 1679–1688.

[2] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz. 2013. A Public Domain Dataset for Human Activity Recognition using Smartphones.. In *ESANN*.

[3] D. Aranki and R. Bajcsy. 2015. Private Disclosure of Information in Health Tele-monitoring. *CoRR* abs/1504.07313 (2015).

[4] R. Assam, M. Hassani, and T. Seidl. 2013. Differential Private Trajectory Obfuscation. In *MOBIQUITOUS*. 139–151.

[5] E. Ayday and M. Humbert. 2017. Inference Attacks against Kin Genomic Privacy. *S&P* 15, 5 (2017), 29–37.

[6] C. BenAbdelkader, R. Cutler, and L. Davis. 2002. Stride and cadence as a biometric in automatic person identification and verification. In *FG*. 372–377.

[7] S. D. Bersch, D. Azzi, R. Khusainov, I. E. Achumba, and J. Ries. 2014. Sensor data acquisition and processing parameters for human activity classification. *Sensors* 14, 3 (2014), 4239–4270.

[8] A. Boutet, S. Ben Mokhtar, and V. Primault. 2016. *Uniqueness Assessment of Human Mobility on Multi-Sensor Datasets*. Research Report. LIRIS UMR CNRS 5205. https://hal.archives-ouvertes.fr/hal-01381986

[9] A. Boutet, D. Frey, R. Guerraoui, A.-M. Kermarrec, and R. Patra. 2014. HyRec: Leveraging Browsers for Scalable Recommenders. In *Middleware*. 85–96.

[10] L. Breiman. 2001. Random forests. *Machine learning* 45, 1 (2001), 5–32.

[11] P. Eckersley. 2010. How unique is your web browser?. In *PETS'10*. 1–18.

[12] C. Frindel and D. Rousseau. 2017. How Accurate Are Smartphone Accelerometers to Identify Intermittent Claudication?. In *HealthyIoT*. 19–25.

[13] P. Gard, L. Lalanne, A. Ambourg, D. Rousseau, F. Lesueur, and C. Frindel. 2018. A Secured Smartphone-Based Architecture for Prolonged Monitoring of Neurological Gait. In *HealthyIoT*. 3–9.

[14] O. Goldreich. 2003. Cryptography and Cryptographic Protocols. *Distrib. Comput.* 16, 2-3 (2003), 177–199.

[15] M. Gramaglia and M. Fiore. 2015. Hiding mobile traffic fingerprints with GLOVE. In *CoNEXT*. 26:1–26:13.

[16] B. Gregorutti, B. Michel, and P. Saint-Pierre. 2017. Correlation and variable importance in random forests. *Statistics and Computing* 27, 3 (2017), 659–678.

[17] S. Guha, M. Jain, and V. N. Padmanabhan. [n. d.]. Koi: A Location-Privacy Platform for Smartphone Apps. In *NSDI*. 183–196.

[18] M. Haghi, K. Thurow, and R. Stoll. 2017. Wearable devices in medical internet of things: scientific research and commercially available devices. *HIR* 23, 1 (2017), 4–15.

[19] J. Han, J. Pei, and M. Kamber. 2011. *Data mining: concepts and techniques*. Elsevier.

[20] G. James, D. Witten, T. Hastie, and R. Tibshirani. 2013. *An introduction to statistical learning*. Vol. 112. Springer.

[21] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler. 2006. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *TITB* 10, 1 (2006), 156–167.

[22] Lamberg L. 2001. Confidentiality and privacy of electronic medical records. *JAMA* 285, 24 (2001), 3075–3076.

[23] D. Manousakas, C. Mascolo, A. R. Beresford, D. Chan, and N. Sharma. 2018. Quantifying Privacy Loss of Human Mobility Graph Topology. *PETS* 2018, 3 (2018), 5–21.

[24] R. Masood, B. Zi Hao Zhao, H. J. Asghar, and M. A. Kâafar. 2018. Touch and You're Trapp(ck)ed: Quantifying the Uniqueness of Touch Gestures for Tracking. *PoPETs* 2018, 2 (2018), 122–142.

[25] S Mehrang, J. Pietilä, and I. Korhonen. 2018. An Activity Recognition Framework Deploying the Random Forest Classifier and A Single Optical Heart Rate Monitoring and Triaxial Accelerometer Wrist-Band. *Sensors* 18, 2 (2018), 613.

[26] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. 2011. Scikit-learn: Machine learning in Python. *Journal of machine learning research* 12, Oct (2011), 2825–2830.

[27] A. PETIT, T. Cerqueus, S. Ben Mokhtar, L. Brunie, and H. Kosch. 2015. PEAS: Private, Efficient and Accurate Web Search. In *TrustCom*.

[28] I. M. Pires, N. M. Garcia, N. Pombo, and F. Flórez-Revuelta. 2016. From data acquisition to data fusion: a comprehensive review and a roadmap for the identification of activities of daily living using mobile devices. *Sensors* 16, 2 (2016), 184.

[29] S. J. Preece, J. Y. Goulermas, L. PJ Kenney, D. Howard, K. Meijer, and R. Crompton. 2009. Activity identification using body-mounted sensorsa review of classification techniques. *Physiological measurement* 30, 4 (2009), R1.

[30] J. L. Reyes-Ortiz. 2015. *Smartphone-based human activity recognition*. Springer.

[31] M. Rushanan, A. D. Rubin, D. F. Kune, and C. M. Swanson. 2014. SoK: Security and Privacy in Implantable Medical Devices and Body Area Networks. In *S&P*. 524–539.

[32] S Scalvini, D. Baratti, G. Assoni, M. Zanardini, L. Comini, and P. Bernocchi. 2014. Information and communication technology in chronic diseases: a patients opportunity.

[33] B. Seref and E. Bostanci. 2016. Opportunities, threats and future directions in big data for medical wearables. In *BDAW*. 15:1–15:5.

[34] S. Sprager and M. B. Juric. 2015. Inertial sensor-based gait recognition: a review. *Sensors* 15, 9 (2015), 22089–22127.

[35] Y. Tang and C. Ono. 2016. Detecting Activities of Daily Living from Low Frequency Power Consumption Data. In *MOBIQUITOUS*. 38–46.

[36] F. Tramèr, Z. Huang, J.-P. Hubaux, and E. Ayday. 2015. Differential Privacy with Bounded Priors: Reconciling Utility and Privacy in Genome-Wide Association Studies. In *CCS*. 1286–1297.

[37] H. Watanabe, T. Terada, and M. Tsukamoto. 2016. Gesture Recognition Method Based on Ultrasound Propagation in Body. In *MOBIQUITOUS*. 288–289.

[38] D. Wood, N. Apthorpe, and N. Feamster. 2017. Cleartext Data Transmissions in Consumer IoT Medical Devices. In *IoT S&P*. 7–12.

[39] Z.-K. Zhang, M. C. Y. Cho, C.-W. Wang, Ch.-W. Hsu, C.-K. Chen, and S. Shieh. 2014. IoT security: ongoing challenges and research opportunities. In *SOCA*. 230–234.