Vers la protection de la vie privée dans les objets connectés pour la reconnaissance d'activité en santé

Théo Jourdan, Antoine Boutet, Carole Frindel

Vers la protection de la vie privée dans les objets connectés pour la reconnaissance d'activité en santé

Théo Jourdan [1,2], **Antoine Boutet** [2], **Carole Frindel** [1]

1. *Univ Lyon, INSA Lyon, CNRS, Inserm, CREATIS UMR 5220, U1206, F-69621 VILLEURBANNE, France*
   carole.frindel@creatis.insa-lyon.fr
2. *Univ Lyon, INSA Lyon, Inria, CITI, F-69621 VILLEURBANNE, France*
   antoine.boutet@insa-lyon.fr

RÉSUMÉ. *Les progrès récents en matière de capteurs sans fil pour les soins de santé personnels permettent de reconnaître les activités humaines en temps réel. L'analyse de ces flux de données peut présenter de nombreux avantages du point de vue de la santé, mais elle peut également conduire à des menaces concernant la vie privée en exposant des informations extrêmement sensibles. Dans cet article, nous proposons un cadre préservant la vie privée pour la reconnaissance d'activité. Ce cadre repose sur une technique d'apprentissage automatique permettant de reconnaître efficacement le modèle d'activité de l'utilisateur, utile pour la surveillance des soins de santé personnels, tout en limitant le risque de ré-identification des utilisateurs à partir de modèles biométriques caractérisant chaque individu. Pour y parvenir, nous avons d'abord analysé en profondeur différents schémas d'extraction de descripteurs dans les domaines temporel et fréquentiel. Nous montrons que les descripteurs du domaine temporel sont utiles pour discriminer l'activité de l'utilisateur, tandis que les descripteurs du domaine de fréquence permettent de distinguer l'identité de l'utilisateur. Sur la base de cette observation, nous avons ensuite conçu un nouveau mécanisme de protection qui traite le signal brut sur le smartphone de l'utilisateur et transfère au serveur d'application uniquement les descripteurs pertinents non liées à l'identité des utilisateurs. De plus, une approche basée sur la généralisation est également appliquée sur les descripteurs du domaine fréquentiel avant leur transmission au serveur afin de limiter les risques de ré-identification. Nous évaluons de manière approfondie notre cadre avec un ensemble de données de référence: les résultats montrent une reconnaissance précise de l'activité (87%) tout en limitant le taux de réidentification (33%). Cela représente une légère diminution de l'utilité (9%) par rapport à une amélioration importante de la confidentialité (53%) par rapport à l'état de l'art, tout en réduisant le coût en temps de calcul sur le serveur applicatif. Enfin, nous validons notre approche en l'appliquant sur une autre base de données contenant des signaux plus perturbés par du bruit.*

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, while reducing the computational cost on the application server. Finally, we also validate our framework with a dataset containing signals more perturbed by noise.

MOTS-CLÉS : *Reconnaissance d'activité, Protection de la vie privée, Objets connectés, Santé*

KEYWORDS: *Activity Recognition, Privacy, IoT, Healthcare.*

DOI:10.3166/RIA.28.1-27

1. Introduction

L'émergence de l'Internet des Objets (IoT) en matière de santé a ouvert la voie à la surveillance individuelle et permanente à domicile ou en milieu hospitalier. En effet, ces appareils enregistrent des paramètres de santé électroniques à partir de divers capteurs (le plus souvent un accéléromètre, un gyroscope et un magnétomètre) et envoient ces données du patient à un serveur d'application pour traitement et analyse. Ces traitements et analyses incluent, par exemple, des algorithmes avancés de traitement du signal et d'apprentissage automatique pour fournir une variété de services tels que (1) le suivi du mouvement: nombre de pas, calories brûlées, surveillance de la distance parcourue et du sommeil et (2) mesure des paramètres vitaux: fréquence cardiaque , température de la peau, électrocardiogramme (ECG) et électroencéphalogramme (EEG) (Haghi *et al*., 2017). En raison de leur nature, les données collectées à partir d'objets connectés sont extrêmement sensibles. D'autant plus, que le flux de travail des données médicales ainsi collectées multiplie les risques pour la sécurité et la confidentialité tout au long du cycle de vie des données: au moment de la collecte et de la transmission (Aranki, Bajcsy, 2015 ; Wood *et al*., 2017), ainsi que durant le traitement et le stockage (Rushanan *et al*., 2014). Lorsque l'adversaire a accès à ces données médicales, les risques d'atteinte à la vie privée – tels que la fuite d'informations sensibles ou la réidentification d'un utilisateur – sont très élevés (par exemple, la ré-identification des informations médicales du gouverneur William Weld (Lamberg, 2001)). Dans le contexte de la reconnaissance d'activité via des objets connectés, le défi consiste à identifier les données qui peuvent préserver la vie privée des individus tout en restant suffisamment pertinentes pour les tâches d'apprentissage automatique (Sprager, Juric, 2015). Ce défi soulève deux questions importantes: 1) Les données collectées sont-elles suffisamment protégées pour que personne ne puisse les détourner pour en déduire des informations sensibles? ou pour ré-identifier le propriétaire? 2) Comment déterminer si les données protégées sont encore suffisamment précises pour les chercheurs dans le domaine de la santé? Atteindre cet équilibre entre l'utilité et la confidentialité des données est un objectif important pour l'envoi de données sécurisées et fiables via des objets connectés et pour renforcer la confiance et l'adoption de l'utilisateur final.

Dans cet article, nous proposons un cadre préservant la confidentialité des données pour la reconnaissance d'activité à partir d'appareils mobiles. Ce cadre s'appuie sur une technique d'apprentissage automatique pour reconnaître efficacement le motif d'activité des utilisateurs – utile pour la surveillance individualisée des soins – tout en limitant le risque de ré-identification des utilisateurs à partir des motifs biométriques caractérisant chaque individu. Pour y parvenir, nous avons d'abord extrait plusieurs descripteurs du signal brut et analysé en profondeur leur impact à la fois sur la reconnaissance d'activité et la ré-identification de l'utilisateur. Nous montrons que les descripteurs du domaine temporel sont utiles pour reconnaître l'activité de l'utilisateur alors que les descripteurs du domaine fréquentiel permettent d'identifier l'utilisateur. Sur la base de cette observation, nous proposons un nouvelle approche permettant la protection de la vie privée. Dans ce contexte, les données collectées sont traitées lo-

calement sur l'appareil de utilisateur et uniquement les descripteurs pertinents sont extraits. De plus, les descripteurs du domaine fréquentiel (i.e. ceux permettant d'identifier les utilisateurs) sont normalisés. Cette normalisation peut être considérée comme une approche basée sur la généralisation. Cependant, comparé à d'autres méthodes de ce type comme le k-anonymat – bien connu pour réduire considérablement l'utilité des données protégées (Gramaglia, Fiore, 2015) – notre solution conserve une grande utilité (i.e. la reconnaissance d'activité) tout en assurant une bonne confidentialité (i.e. faible taux d'identification de l'utilisateur). Une fois normalisées, ces descripteurs sont périodiquement téléchargées sur le serveur d'application. Chaque lot de descripteurs est stocké indépendamment sur le serveur (i.e. avec un pseudonyme différent) pour éviter de lier les lots à des individus et à lier des lots ensemble. De plus, pour éviter de centraliser à la fois les données et l'identité associée sur le même noeud, la correspondance entre les pseudonymes et les identités des utilisateurs est uniquement connue par les praticiens de l'hôpital.

Nous avons évalué de manière exhaustive notre approche grâce à un jeu de données de référence. Les résultats témoignent d'une reconnaissance d'activité de 87% en moyenne tout en limitant la ré-identification de l'utilisateur à 33%. Nous avons également comparé notre approche à différentes approche de l'état de l'art. Notre approche fournit un meilleure compromis entre confidentialité et utilité avec une légère diminution de l'utilité (9%) contre une forte augmentation de la confidentialité (53%). Nos contributions peuvent être résumées comme suit:

– Nous quantifions à la fois le risque associé à la ré-identification des utilisateurs (90% en moyenne) et la capacité à détecter l'activité des utilisateurs (97% en moyenne) à partir du signal brut d'appareils mobiles.

– Nous avons analysé l'impact de multiples descripteurs à la fois sur la reconnaissance d'activité et sur la ré-identification des utilisateurs. Nous montrons que les descripteurs du domaine temporel sont utiles pour reconnaître l'activité de l'utilisateur alors que les descripteurs du domaine fréquentiel permettent d'identifier l'utilisateur.

– Nous proposons une approche efficace – basée sur de l'apprentissage automatique – pour reconnaître l'activité des utilisateurs avec une grande utilité tout en limitant les risques de ré-identification de l'utilisateur. Notre solution offre un meilleur compromis entre protection de la vie privée et utilité au regard de l'état de l'art actuel: i.e. une légère diminution de l'utilité (9%) et une forte augmentation de la confidentialité (53%), tout en réduisant le coût en temps de calcul sur le serveur applicatif (une réduction de temps de 81%).

– Nous évaluons et validons notre solution avec une autre base de données contenant des signaux plus bruités. Nous montrons que l'impact sur l'utilité est réduit, et que cet impact peut être supprimé en adaptant l'étape de pré-traitement (filtrage) selon les signaux considérés.

Cet article est structuré de la manière suivante. Le pipeline de traitements des données est introduit dans la section 2 avant de définir le modèle d'adversaire dans la section 3. Nous quantifions et analysons ensuite la capacité de reconnaissance de l'activité et de l'identité dans la section 4. La section 5 détaille notre nouvelle approche de

préservation de la vie privée et la section 6 présente son évaluation. Enfin, les travaux connexes sont passés en revue dans la section 7 avant de conclure dans la section 8.

2. Pipeline de traitements des données

Cette section explique la méthodologie que nous avons utilisée pour la reconnaissance d'activités et la ré-identification de l'utilisateur à l'aide d'un dispositif mobile. Bien que cette description soit spécifique à notre méthodologie, elle peut fournir des informations de base au regard de l'utilisation de dispositifs mobiles pour la santé. La Figure 1 décrit l'ensemble du pipeline, y compris l'acquisition des données (section 2.1), le prétraitement du signal (section 2.2), la segmentation (section 2.3), l'extraction des descripteurs (section 2.4), et la classification (section 2.5).

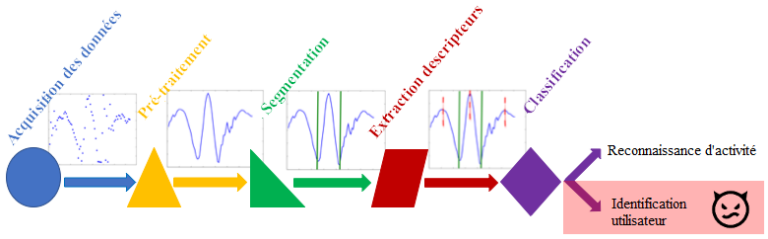


Figure 1. Pipeline de traitement de données IoT pour la reconnaissance d'activité. Un adversaire peut également utiliser un classificateur pour faire de la ré-identification de l'utilisateur.

2.1. Acquisition des données

L'acquisition des données repose sur des capteurs présents dans les appareils mobiles tels que les smartphones, montres connectées, bracelets intelligents et capteurs médicaux. Il existe une grande variété de capteurs permettant l'acquisition de différents types de données, qui peuvent ensuite être utilisées pour différents types de tâches. Pour la reconnaissance des activités physiques, les auteurs dans (Preece *et al.*, 2009) proposent l'utilisation de capteurs inertiels, c'est-à-dire d'accéléromètres et de gyroscopes, complétés par une mesure de l'orientation (e.g. magnétomètre) et une mesure de la localisation (e.g. GPS). Le processus d'acquisition des données est réalisé par un module propre dans l'appareil mobile et consiste en la mesure et conversion des signaux électriques reçus par chaque capteur dans un format lisible (Scalvini *et al.*, 2014). Plusieurs défis sont associés au processus d'acquisition de données dans le contexte de la reconnaissance d'activités, y compris le positionnement du dispositif mobile, la taux d'échantillonnage des signaux et le nombre de capteurs à utiliser et donc à gérer (Bersch *et al.*, 2014). Tous ces facteurs influencent directement l'extraction de descripteurs corrects. Comme les capteurs sont tous intégrés au niveau de l'appareil mobile, ils ne peuvent pas être placés séparément sur différentes parties

du corps; l'idée est plutôt de placer l'appareil mobile dans une position habituelle et confortable. Un autre problème lié aux appareils mobiles est la consommation d'énergie des tâches d'acquisition de données. L'exécution multitâche diffère d'un modèle d'appareil mobile à l'autre, car elle dépend de leur capacité de traitement, de leur mémoire, de leur système d'exploitation et du nombre et type d'applications mobiles installées et/ou en fonctionnement. La sélection des meilleures méthodes d'acquisition de données dépend donc de la finalité de l'utilisation, du type de données acquises et leur environnement (Frindel, Rousseau, 2017 ; Pires *et al.*, 2016).

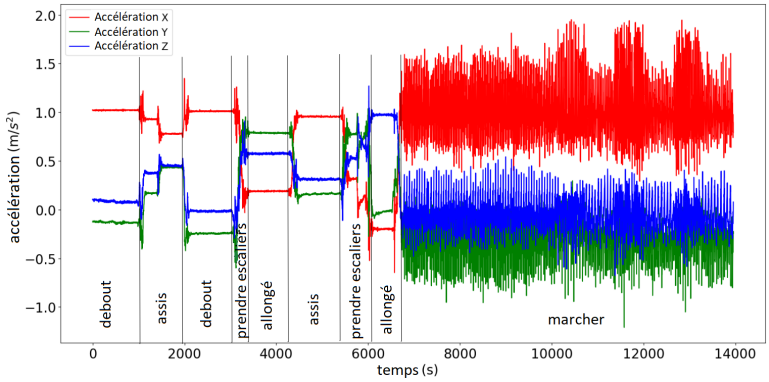


Figure 2. Visualisation d'un signal d'accéléromètre dans les dimensions x, y et z et les activités associées.

2.2. Prétraitement des signaux

Les signaux issus des capteurs sont prétraités par une série de filtres. Tout d'abord, le bruit a été réduit grâce à un filtre médian et un filtre de Butterworth passe-bas du troisième ordre (fréquence de coupure de 20 Hz). Ce seuil de fréquence a été sélectionné en fonction de travaux antérieurs (Karantonis *et al.*, 2006) qui indiquent que le spectre d'énergie du mouvement du corps humain est inférieur 15 Hz. Les signaux résultants ont ensuite été filtrés pour les séparer en canaux qui ont un sens physique, comme le montre la Figure 3. Par exemple, le signal d'accélération linéaire a été décomposé en deux canaux principaux: les composantes de la gravitation et du mouvement du corps. Cette étape a été réalisée en utilisant un autre filtre passe-bas et en supposant que la composante gravitationnelle se réfère principalement aux fréquences les plus basses (Anguita *et al.*, 2013). Par la suite, les signaux d'accélération du mouvement corporel et de giration ont été dérivés par rapport au temps afin d'obtenir la secousse (jerk) reflétant les variations temporelles des signaux. Enfin, les signaux ont été décomposés en fonction de leurs axes d'acquisition (x, y, z respectivement) afin de les observer dans une direction spécifique (verticale, latérale ou longitudinale), comme illustré sur la Figure 2. La magnitude associée aux signaux a également été calculée pour produire un signal moyen moins sensible à la manière dont l'appareil a

été fixé sur la personne. Au total, cette étape de filtrage a permis d'obtenir 20 canaux différents.

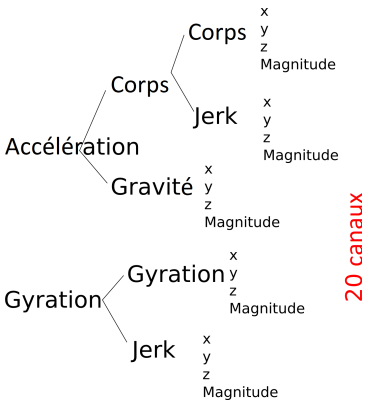


Figure 3. Canaux résultant de l'étape de pré-traitement.

2.3. Segmentation

Les signaux des canaux ont été segmentés en utilisant une technique de fenêtre glissante fixe. Les fenêtres avec une durée de 2,5 secondes et un chevauchement de 50% ont été capturées. Un degré de chevauchement de 50% signifie que la fenêtre est décalée de la moitié de sa taille. En d'autres termes, 50% des données précédentes sont incluses dans la fenêtre suivante. Le choix de la taille de la fenêtre n'est pas anodin, en particulier pour un algorithme de reconnaissance d'activité: une taille de fenêtre trop petite peut diviser un signal d'activité alors qu'une taille de fenêtre trop grande peut contenir plusieurs signaux d'activité. Nous avons donc décidé de calibrer la taille de notre fenêtre sur l'activité la plus complexe: la marche. Par conséquent, la taille de la fenêtre a été choisie pour prendre en compte au moins un cycle de marche complet (deux pas): la plage de cadence d'une marche moyenne correspond à une vitesse minimale de 1,5 pas par seconde, conformément à (BenAbdelkader *et al.*, 2002).

2.4. Extraction des descripteurs

De chaque fenêtre du signal de chaque canal, un vecteur de descripteurs contenant 17 mesures estimées dans les domaines temporels et fréquentiels a été extrait. La transformée de Fourier discrète (DFT) a été utilisée pour extraire les descripteurs du domaine fréquentiel. Le choix de ces descripteurs a été effectué sur la base d'une étude antérieure sur les descripteurs effectifs pour la reconnaissance de la marche (Sprager, Juric, 2015): par exemple, pour le signal dans le domaine temporel, la moyenne (mean), l'écart type (std), la zone de magnitude du signal (sma) et la corrélation (corr); et pour

le signal dans le domaine fréquentiel, l'énergie et l'entropie. Les mesures sélectionnées pour obtenir le vecteur de descripteurs sont illustrées à la Figure 4 . Un vecteur de descripteurs a été calculé à partir de chaque fenêtre des signaux et étiqueté en fonction de l'activité et de l'utilisateur auxquels il appartient. La Figure 5 montre un exemple de jeu de données, où les lignes correspondent aux échantillons de fenêtres et les colonnes aux descripteurs (à l'exception des deux dernières qui correspondent aux étiquettes). Un tel jeu de données est utilisé comme entrée pour la tâche de classification qui suit. Un total de 340 descripteurs (20 canaux x 17 mesures) est extrait. La notation pour nommer un descripteur dans la suite de cet article est la suivante {orientation}_{canal}_{descripteur}.

	Fonction	Description	Formule
Domaine temporel	mean $(\boldsymbol{s})$	Moyenne arithmétique	$\bar{s} = \frac{1}{N}\sum_{i=1}^{N} s_i$
	std $(\boldsymbol{s})$	Ecart-type	$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(s_i - \bar{s})^2}$
	mad $(\boldsymbol{s})$	Ecart médian absolue	$\text{median}_i\,(\,\lvert s_i - \text{median}_j(s_j)\rvert\,)$
	max $(\boldsymbol{s})$	Maximum	$\max_i(s_i)$
	min $(\boldsymbol{s})$	Minimum	$\min_i(s_i)$
	sma $(\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3)$	Magnitude	$\frac{1}{3}\sum_{i=1}^{3}\sum_{j=1}^{N}\lvert s_{i,j}\rvert$
	iqr $(\boldsymbol{s})$	Ecart interquartile	$\text{Q3}(\boldsymbol{s}) - \text{Q1}(\boldsymbol{s})$
	autoregression $(\boldsymbol{s})$	Coefficients d'autoregression (4ème ordre, méthode de Burg)	$\boldsymbol{a} = \text{arburg}(\boldsymbol{s}, 4), \boldsymbol{a} \in \mathbb{R}^4$
	correlation $(\boldsymbol{s}_1, \boldsymbol{s}_2)$	Coefficient de corrélation de Pearson	$C_{1,2}/\sqrt{C_{1,1}C_{2,2}}, C = \text{cov}(\boldsymbol{s}_1, \boldsymbol{s}_2)$
	angle $(\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, v)$	Angle entre la moyenne triaxiale et le vecteur	$\tan^{-1}(\lVert[\bar{s}_1, \bar{s}_2, \bar{s}_3] \times \boldsymbol{v}\rVert, [\bar{s}_1, \bar{s}_2, \bar{s}_3] \cdot \boldsymbol{v})$
Domaine fréquentiel	skewness $(\boldsymbol{s})$	Asymétrie du signal fréquentiel	$\text{E}\left[\left(\frac{\boldsymbol{s}-\bar{s}}{\sigma}\right)^3\right]$
	kurtosis $(\boldsymbol{s})$	Coefficient d'aplatissement	$\frac{\text{E}\left[(\boldsymbol{s}-\bar{s})^4\right]}{\text{E}\left[(\boldsymbol{s}-\bar{s})^2\right]^2}$
	maxFreqInd $(\boldsymbol{s})$	Argument du maximum	$\arg\max_i(s_i)$
	energy $(\boldsymbol{s})$	Moyenne des carrés	$\frac{1}{N}\sum_{i=1}^{N} s_i^2$
	entropy $(\boldsymbol{s})$	Entropie	$\sum_{i=1}^{N}(c_i \log(c_i)),\ c_i = s_i / \sum_{j=1}^{N} s_j$
	meanFreq $(\boldsymbol{s})$	Moyenne pondérée du signal fréquentiel	$\sum_{i=1}^{N}(is_i) / \sum_{j=1}^{N} s_j$
	energyBand $(\boldsymbol{s}, a, b)$	Energie spectrale sur une bande fréquentielle [a,b]	$\frac{1}{a-b+1}\sum_{i=a}^{b} s_i^2$

Figure 4. Mesures du domaine temporel et fréquentiel sélectionnées pour calculer les vecteurs de descripteurs. N: taille en nombre d'échantillons du signal issu d'un canal, 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. Un échantillon de données avec les descripteurs et étiquettes associés: entrée de l'étape de classification.

2.5. *Classification*

2.5.1. *Algorithme d'apprentissage automatique*

Les forêts aléatoires (RF) ont été choisies pour les tâches de classification multi-classes, respectivement les classes faisant référence à la reconnaissance d'activité et

les classes associées aux identités des utilisateurs (dans le cas d'un adversaire cherchant à utiliser le classifieur de manière abusive pour réidentifier les utilisateurs). En général, l'algorithme RF est un classifieur supervisé ayant un temps d'apprentissage rapide et de très bonnes performances sans réglage précis (Mehrang *et al*., 2018). Les RF fonctionnent en construisant un grand ensemble d'arbres de décision, chaque arbre étant construit sur un échantillon bootstrap des données d'origine (Breiman, 2001). Les arbres de classification sont construits sur la base de divisions binaires récursives: pour chaque division, un sous-ensemble de variables d'entrée choisi de manière aléatoire est utilisé pour trouver la division binaire optimale correspondant à une condition sur un descripteur. Les divisions optimales sont déterminées à l'aide de l'indice d'impureté de Gini (James *et al*., 2013). La fonction "RandomForestClassifier" du package Scikit Learn de Python (Pedregosa *et al*., 2011) a été utilisée pour construire le classifieur RF. Dans ce travail, selon les instances et les descripteurs de notre problème de classification, 700 est choisi comme le nombre d'arbres dans la forêt, $\sqrt{n}$ descripteurs aléatoires sont considérés dans la construction de chaque arbre et 10 est défini comme la profondeur maximale de chaque arbre.

2.5.2. Mesures d'utilité et de confidentialité

Pour mesurer la qualité de la classification sur la base des descripteurs proposés et de l'algorithme de RF, nous avons calculé la précision à partir de la matrice de confusion (Han *et al*., 2011):

$$\frac{\mid TP \mid + \mid TN \mid}{\mid TP \mid + \mid TN \mid + \mid FP \mid + \mid FN \mid}$$

où $\mid TP \mid$ (vrais positifs) est le nombre de prédictions correctes pour une valeur d'événement spécifique, $\mid TN \mid$ (vrais négatifs) est le nombre de prédictions correctes pour les valeurs non-événement, $\mid FP \mid$ (faux positifs) correspond au nombre de prédictions incorrectes pour une valeur d'événement spécifique, et $\mid FN \mid$ (faux négatifs) correspond au nombre de prédictions incorrectes pour des valeurs non-événement. La précision reflète le nombre de prédictions correctes effectuées par le modèle (numérateur) sur tous les types de prédictions (dénominateur). La précision est comprise dans [0:1] où 1 correspond à une prédiction parfaite. Nous utilisons cette métrique afin de calculer la qualité de notre classification pour prédire à la fois l'activité et l'identité de l'utilisateur. Nous avons appelé *Précision(activité)* le résultat quand il est appliqué à la reconnaissance d'activité, et *Précision(ré-identification)* lorsqu'il est appliqué à l'identité de l'utilisateur. Nous préférons la précision au F-Score, car, dans notre cas, les classes de variables dans les données sont presque équilibrées.

2.5.3. Classement et sélection des descripteurs

L'algorithme RF peut être utilisé pour classer les descripteurs en fonction de leur importance dans le processus de classification. Lors de l'apprentissage d'un arbre, il est possible de calculer de combien chaque descripteur diminue l'indice d'impureté de Gini (James *et al*., 2013) dans l'arbre. Pour une forêt, la diminution d'impureté de

chaque descripteur peuvent être moyennée et les descripteurs sont classés en fonction de cette mesure.

Algorithme 1 : Sélection des descripteurs

Input : Liste des descripteurs triés par ordre d'importance f et la précision initiale associée a; $threshC = 0.7$; $threshA = 0.03$
Output : Liste des descripteurs sélectionnés

```
1  for chaque descripteur f_i ∈ f do
2      Calcul du coefficient de corrélation de Pearson C pour chaque descripteur dans {f − f_i} : fcorre
3      for chaque descripteur f_j ∈ fcorre do
4          if | C(f_j) | > threshC then
5              Calcul de la précision newA de la classification pour {f − f_j} : newa
6              if a − newA < threshA then
7                  Enlève le descripteur f_j de f
8              end
9          end
10     end
11 end
```

L'algorithme RF peut également être utilisé pour la sélection des descripeurs (Breiman, 2001). Cela se fait en mesurant la diminution moyenne de la précision quand un descripteur particulier est retiré de l'ensemble des descripteurs dans les arbres. Si la diminution de la précision après exclusion du descripteur est négligeable, le descripteur est peu important et inversement. Les scores d'importance des descripteurs du classifieur RF (Breiman, 2001 ; Gregorutti *et al*., 2017) peuvent donc être évalués et utilisés comme critères de sélection des descripteurs. Pour plus de détails, voir l'algorithme 1: Il se compose de deux boucles imbriquées, une correspondant aux descripteurs classés par importance (ligne 1) et une correspondant aux descripteurs corrélées à chacune des descripteurs de la première boucle (ligne 3). La corrélation est calculée en utilisant le coefficient de Pearson (ligne 2). Si la corrélation entre deux descripteurs est supérieure à un certain seuil (ligne 4), la précision de l'algorithme RF est recalculée après suppression du descripteur corrélé (ligne 5) et si la diminution correspondante de la précision est inférieure à un certain seuil (ligne 6), ce descripteur est définitivement supprimé (ligne 7).

3. Le modèle d'adversaire

Avant de présenter notre approche préservant la vie privée à la section 5, nous décrivons nos hypothèses et le modèle d'adversaire pour lequel notre solution est conçue. Le cadre présenté dans cet article repose sur trois entités: le client s'exécutant sur le smartphone des utilisateurs, le serveur d'application stockant les descripteurs et effectuant la classification, et le praticien hospitalier surveillant l'activité du patient. Tout d'abord, nous supposons que l'application client et le smartphone sur lequel elle est exécutée sont fiables. Cela signifie que l'acquisition des données, le prétraitement, la segmentation, l'extraction des descripteurs et la normalisation ne peuvent

pas s'écarter d'un comportement correct. De plus, nous ne considérons pas la limitation du taux d'échantillonnage de l'acquisition de données comme dans (Tang, Ono, 2016). Deuxièmement, nous supposons que le serveur d'application s'exécute sur des plates-formes de cloud public. Nous considérons que cette plateforme de cloud est honnête mais curieuse (Goldreich, 2003). Cela signifie que le serveur d'application se comporte correctement en ce qui concerne le traitement des données reçues des clients. Plus précisément, cela signifie que les données sont correctement stockées dans la base de données, qu'aucune information falsifiée ne peut être injectée dans la base de données et que le modèle de classifieur ne peut pas être manipulé de manière malveillante. Cependant, nous supposons que l'adversaire est capable de collecter une partie ou la totalité des informations stockées dans la base de données. Chaque information correspond à des lots indépendants de données non liées aux utilisateurs (c'est-à-dire avec un pseudonyme aléatoire différent pour chaque lot). De plus, nous supposons que l'adversaire est capable de collecter des données relatives aux gestes de chaque utilisateur à partir d'un dispositif IoT malveillant, par exemple. L'adversaire utilise cette connaissance préalable de chaque utilisateur pour créer un modèle de classificateur. Ce classificateur exploite les mêmes prétraitements, segmentations et descripteurs que notre classifieur, mais avec l'objectif de prédire l'identité de l'utilisateur pour chaque lot de données stockées dans la base de données. Troisièmement, nous supposons que le serveur utilisé par le praticien hospitalier est fiable. Ce serveur est utilisé pour stocker le lien entre les lots de données envoyés au serveur d'application et l'identité des utilisateurs. Enfin, toutes les communications entre les noeuds (clients, serveur d'application et serveur de l'hôpital) sont sécurisées. Nous supposons qu'aucune information ne peut être déduite de ces communications sécurisées.

4. Quantification de la reconnaissance d'activité et de la ré-identification de l'utilisateur

Nous avons effectué une évaluation approfondie de la capacité de notre classifieur à reconnaître l'activité des utilisateurs et à les réidentifier. Nous montrons qu'en suivant la méthodologie décrite dans la section 2, nous sommes en mesure de prédire l'activité de l'utilisateur avec un taux de réussite très élevé. De plus, nous montrons que sans système de protection, les données des appareils mobiles agissent comme une empreinte digitale personnelle et conduisent à une identification aisée des utilisateurs. Nous décrivons d'abord le jeu de données utilisé dans cette évaluation à la section 4.1 avant de quantifier la reconnaissance de l'activité et la ré-identification de l'utilisateur aux sections 4.2 et 4.3, respectivement. Enfin, nous analysons l'impact des descripteurs extraits à la section 4.4.

4.1. Jeu de données

L'ensemble des données utilisé dans ce travail est disponible en ligne pour un usage public en tant que jeu de données «Human Activity Recognition using Smartphones» dans le dépôt UCI Machine Learning (Anguita *et al.*, 2013). Il est composé

des données brutes triaxiales des accéléromètres et gyroscopes lus à une fréquence constante de 50 Hz. Un groupe de 30 volontaires ont été sélectionnés pour suivre un protocole d'activités tout en portant un smartphone à la taille. L'expérience a été planifiée pour contenir six activités de base: trois postures statiques (debout, assis, couché) et trois activités dynamiques (marcher, descendre et monter les escaliers). La Figure 2 montre le signal de l'accéléromètre d'une des expériences ainsi que les activités associées. Le protocole des activités est détaillé dans (Reyes-Ortiz, 2015). La durée d'une expérience entière était d'environ 15 minutes et a été répétée dix fois. Toutes les expériences ont été enregistrées sur vidéo afin d'avoir une vérité terrain pour annoter les activités effectuées.

4.2. Reconnaissance d'activités

Le tableau 1 résume la précision de la reconnaissance des différentes activités. Les résultats montrent que notre approche d'apprentissage automatique est capable de reconnaître les activités avec une précision moyenne de 0,97. Comme le tableau 1 l'indique, la précision est moindre pour les activités ambulatoires dans les escaliers. Une explication possible à cela est que ces activités correspondent aux temps d'acquisition les plus faibles (voir Figure 2).

Activité	**Précision(activité)**
Walking	0.97
Walking upstairs	0.95
Walking downstairs	0.94
Sitting	0.97
Standing	0.98
Laying	0.99

TABLE *1. Les activités des utilisateurs peuvent être reconnues avec un fort taux de précision (reconnaissance utilisant la méthodologie présenté dans la Section 2).*

4.3. Ré-identification de l'utilisateur

La Figure 6 illustre la distribution cumulative de la précision pour la tâche de ré-identification de l'utilisateur. La précision varie de 0,82 à 0,96 parmi les 30 utilisateurs avec une moyenne de 0,90. Ces résultats indiquent que les données collectées sur le mouvement des utilisateurs caractérisent chaque individu et peuvent conduire à les réidentifier avec un taux de réussite élevé. Cependant, la tâche de ré-identification est légèrement plus difficile que celle de reconnaissance d'activités, donc avec une précision moindre.

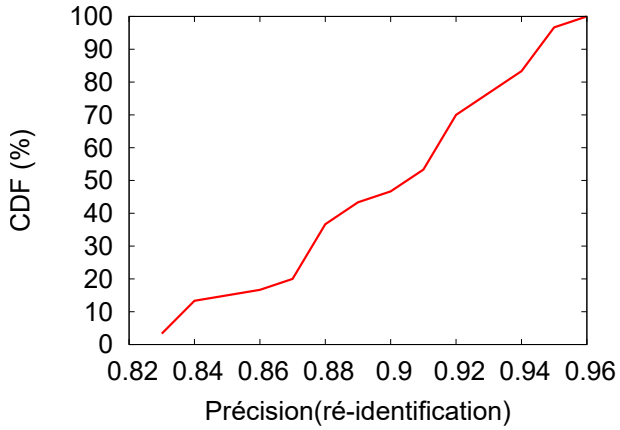


Figure 6. Distribution cumulative de la précision pour la tâche de ré-identification: les utilisateurs peuvent être facilement ré-identifiés à partir de leurs données.

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. *Descripteurs les plus importants pour la ré-identification de l'utilisateur (les descripteurs dans le domaine fréquentiel sont en gris).*

4.4. Impact des descripteurs

Les expériences précédentes servent également à classer les descripteurs (parmi les 340) en fonction de leur importance. Huit et onze descripteurs ont été respectivement sélectionnés pour les tâches de reconnaissance d'activités et de ré-identification de l'utilisateur, compte tenu de l'analyse de corrélation et de précision (voir l'algorithme 1 pour la méthodologie et les tableaux 2 et 3 pour les résultats). En effet, de nombreux descripteurs se ressemblent et contiennent des informations similaires. Par rapport à l'utilisation des 340 descripteurs, l'utilisation de seulement 19 descripteurs pertinemment choisis réduit légèrement (<4%) les performances des deux tâches de classification (97% contre 96% pour la classification par activité et 90% vs 86% pour la ré-identification de l'utilisateur). Cela peut être observé plus précisément sur les Figures 7a et 7b, où l'importance de chaque descripteur sélectionné est testée indé-

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. Descripteurs les plus importants pour la reconnaissance d'activités (les descripteurs dans le domaine fréquentiel sont en gris).*

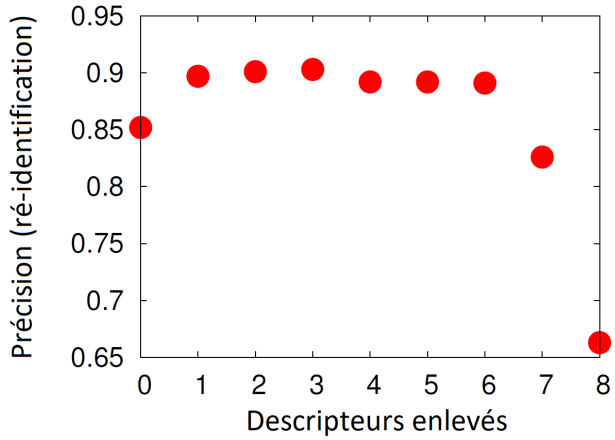
pendamment pour la tâche à étudier: il existe une forte corrélation entre l'importance d'un descripteur spécifique et les performances de l'algorithme RF après l'avoir supprimé. Sur la base de ces résultats de classement, il est intéressant de noter que la tâche de reconnaissance des activités (utilité) est presque exclusivement (9 des 11 descripteurs sélectionnés) opérée dans le domaine temporel, tandis que la tâche d'identification de l'utilisateur (confidentialité) est basée (sur 5 des 8 descripteurs sélectionnés) sur des descripteurs du domaine fréquentiel. Ces résultats peuvent s'expliquer par le fait que les activités se distinguent principalement les unes des autres par leur niveau d'amplitude d'accélération et de giration (voir Figure 2) et donc leurs statistiques associées. Inversement, l'identification de l'utilisateur est davantage liée au rythme ou à la cadence à laquelle cette personne effectue l'activité et est fortement liée à la biomécanique (par exemple, l'âge, la taille, le poids).

5. Approche de reconnaissance d'activité avec protection de la vie privée

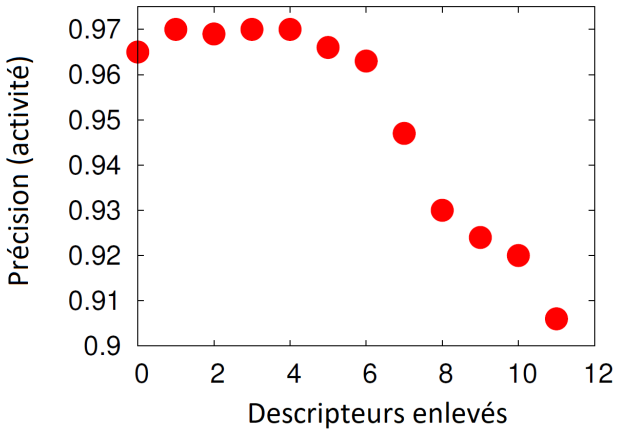
Pour assurer la confidentialité, notre approche repose à la fois sur une architecture limitant l'exposition d'informations sensibles et sur une normalisation appliquée aux descripteurs conduisant à la ré-identification de l'utilisateur (section 4.4). Ces normalisations agissent comme une forme d'obfuscation fondée sur la généralisation. Dans cette section, nous présentons d'abord l'architecture de notre approche (Section 5.1) avant de décrire la normalisation de chaque descripteur sensible (Section 5.2).

5.1. Architecture

La conception de notre approche de préservation de la vie privée comprend trois éléments principaux: une application client s'exécutant sur le smartphone de l'utilisateur communiquant avec son environnement IoT, le serveur d'application et le



(a) Confidentialité



(b) Précision

Figure 7. Impact du nombre de descripteurs retenus (décrits dans Table 1 et Table 2) dans le processus de sélection (les descripteurs sont triés par ordre croissant d'importance).

praticien de l'hôpital. Pour limiter l'exposition d'informations sensibles, le serveur d'application ne stocke pas les données identifiées mais uniquement des lots de descripteurs où chaque lot est pseudo-anonymisé aléatoirement. Seul le praticien de l'hôpital connaît le lien entre les identités et les pseudonymes des lots, et demande au serveur d'application de surveiller l'activité des utilisateurs. L'architecture de notre système est illustrée dans la Figure 8. Tout d'abord, des dispositifs IoT (par exemple, une montre intelligente) ou directement les smartphones effectuent l'acquisition de données (❶). Dans les deux cas, ces données brutes sont stockées localement sur le smartphone. L'application client effectue ensuite le prétraitement, la segmenta-

tion et l'extraction des descripteurs en suivant la méthodologie décrite dans la section 2. Sur la base de notre analyse de l'importance des descripteurs, cette extraction ne concerne que les 19 descripteurs considérés comme importants (section 4.4). De plus, le client procède à la normalisation des descripteurs identifiés comme conduisant à la ré-identification des utilisateurs. Toutes ces normalisations sont décrites dans la sous-section suivante. Étant donné que toutes les actions susmentionnées sont effectuées sur le smartphone et ne concernent que l'utilisateur associé à un seul lot de données (par exemple, pour quelques heures dans la journée), le coût de calcul qui en résulte est bas. Sur un ordinateur standard, ces opérations appliquées à toutes les données d'un utilisateur durent 2,5 secondes dans nos expériences. Deuxièmement, l'application client associe un pseudonyme aléatoire à chaque lot daté de descripteurs avant de les transférer périodiquement sur le serveur d'application (❷). L'application client envoie alors au praticien hospitalier la liste des pseudonymes associé à son identité (❸). Lorsqu'un lot de descripteurs est reçu par le serveur d'application, celui-ci stocke ces informations dans une base de données (❹). Par conséquent, chaque lot de cette base de données ne contient pas l'identité de l'utilisateur mais un pseudonyme aléatoire. Le serveur d'application effectue ensuite périodiquement la classification pour détecter l'activité associée à chaque lot de descripteurs. Enfin, lorsque le praticien hospitalier souhaite surveiller l'activité d'un utilisateur spécifique, il récupère d'abord localement tous les pseudonymes associés à l'utilisateur spécifié, puis demande au serveur d'application de disposer de l'historique des activités des pseudonymes spécifiés (❺).

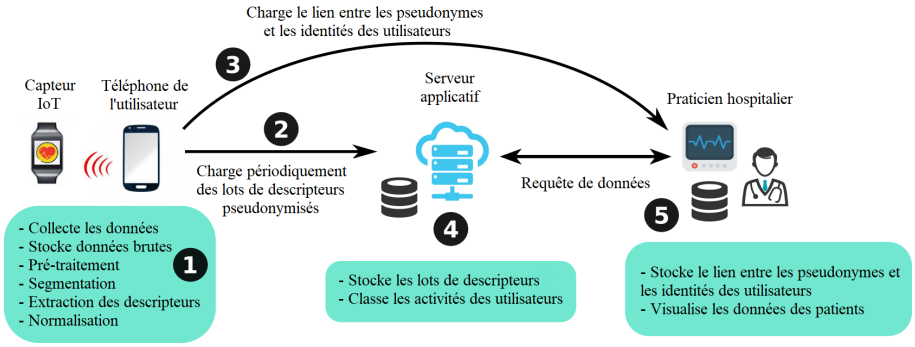


Figure 8. Architecture de notre approche: le smartphone de l'utilisateur extrait les descripteurs pertinents et juste ces descripteurs sont envoyés périodiquement au serveur d'application.

5.2. Normalisation

Afin de limiter la ré-identification des utilisateurs, nous proposons un schéma de normalisation qui généralise l'effet des différents descripteurs identifiés comme importants pour la tâche de ré-identification de l'utilisateur. En d'autres termes, nous essayons d'atténuer leurs caractéristiques permettant la ré-identification de l'utilisateur

sans les supprimer complètement, car elles ont également un impact sur la reconnaissance des activités. Étant donné les données issues des capteurs notés S et de taille n, appliquer l'approche de normalisation sur S produira les «données normalisées» notées S^*. Dans ce travail, nous avons distingué cinq normalisations, chacune faisant référence aux descripteurs du domaine fréquentiel répertoriés dans le Tableau 2. En ce qui concerne les descripteurs temporels, nous les supprimons simplement.

5.2.1. Normalisation par la moyenne

(Y_gyro_mean)

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

avec μ and μ^* étant respectivement la moyenne des données avant et après la normalisation.

5.2.2. Normalisation par écart interquantile

(Y_gyro_irq)

L'écart interquartile (IQR) est une mesure de la dispersion statistique, égale à la différence entre le 75ème et le 25ème percentile.

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

avec IQR and IQR^* étant respectivement l'écart interquantile des données avant et après la normalisation.

5.2.3. Normalisation par l'écart-type

(Y_grav_std)

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

avec σ and σ^* étant respectivement l'écart-type des données avant et après la normalisation.

5.2.4. Normalisation par l'énergie

(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 par le maximum et minimum

(X_grav_max)

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

avec Max et Min étant respectivement le maximum et le minimum des données d'origine, et $newMax$ and $newMin$ le maximum et le minimum des données normalisées.

Les valeurs de référence après la normalisation par la moyenne, l'IQR, l'écart-type et newMin et newMax ont été choisies en prenant la moyenne des valeurs avant la normalisation.

6. Evaluation de notre approche

Nous avons effectué une évaluation approfondie de notre approche. Dans cette section, nous commençons par décrire les bases de comparaison (section 6.1) avant d'évaluer la performance de notre approche en termes de compromis entre utilité et confidentialité (section 6.2), de temps de calcul (section 6.3) et de sensibilité au bruit (section 6.4).

6.1. Comparaison à l'état de l'art

Pour mettre en évidence les avantages de notre approche, nous comparons ses performances à celles de deux alternatives de l'état de l'art. La première alternative suit un schéma de perturbation similaire à l'approche différentiellement privée décrite dans (Acs, Castelluccia, 2014) qui ajoute du bruit dans le domaine fréquentiel pour des séries temporelles dans le contexte de la confidentialité des localisations. Dans notre cas, cette alternative (appelée perturbation) ajoute un bruit gaussien au signal dans le domaine fréquentiel avant l'extraction des descripteurs. La deuxième alternative consiste simplement à supprimer les descripteurs menant à la ré-identification de l'utilisateur (section 4.4). La motivation derrière cette deuxième alternative (appelée suppression) est que sans ces descripteurs, la ré-identification est plus difficile.

6.2. Amélioration de la confidentialité

La Figure 9 indique, pour notre approche ainsi que celles de l'état de l'art, le compromis entre l'utilité capturée par la précision de reconnaître l'activité et la confidentialité capturée par la précision de réidentifier les utilisateurs. Pour l'approche basée sur la suppression de descripteurs, chaque point de la courbe correspond à la suppresion d'un descripteur (parmi les 8 sélectionnés pour la tâche de ré-identification). Pour l'approche basée sur la perturbation, chaque point se réfère à l'addition d'une quantité fixe de bruit croissante (le bruit est centré sur zéro et son écart type est augmenté de 2 pour chaque point). Enfin, concernant notre approche, chaque point correspond à la normalisation d'un nombre croissant de descripteurs (par ordre d'importance croissante). Les résultats montrent que l'approche de suppression (pente: 0,12) semble la plus avantageuse en termes de compromis entre utilité et confidentialité. Cependant, elle est très rapidement limitée par le nombre de descripteurs sélectionnés et donc par les mesures de confidentialité et d'utilité: par exemple, les meilleures performances

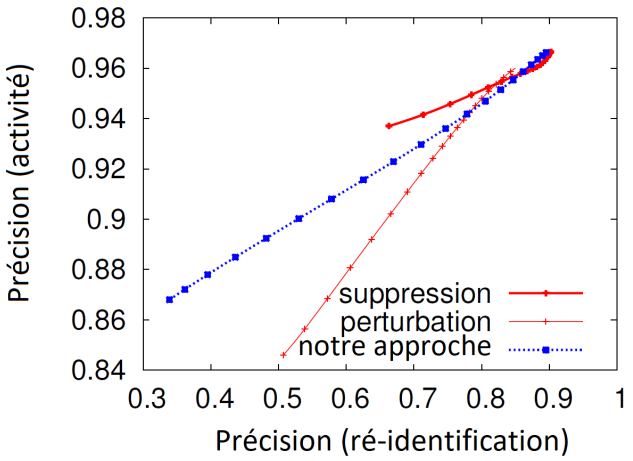


Figure 9. Notre approche fournit un meilleur compromis précision et confidentialité que nos approches de base.

obtenues sont respectivement 0,66 et 0,93. L'approche par perturbation (pente: 0,34) est très efficace en termes de perte d'identité, mais elle entraîne également une perte d'utilité très importante, avec les meilleures performances en termes de confidentialité et d'utilité, de respectivement 0,51 et 0,84. Notre approche se situe entre les deux (pente: 0,21) et fournit le meilleur compromis entre utilité et confidentialité (respectivement 0,87 et 0,33). Notre approche basée sur la normalisation permet de mieux contrôler le poids de chaque descripteur de la confidentialité, contrairement à l'approche de suppression pour laquelle leur impact est limité. Enfin, nous avons également considéré un adversaire qui entraîne un classifieur uniquement avec des descripteurs conduisant à la ré-identification (tableau 2), dans ce cas, la précision en termes de ré-identification est moins efficace qu'avec notre approche (0.17).

6.3. Amélioration des temps de calcul

Nous allons maintenant comparer les temps de calcul de notre approche avec la solution basée totalement sur un serveur centralisé. Dans cette solution, toutes les données collectées par les capteurs IoT sont envoyés au serveur applicatif qui va réaliser toutes les opérations incluant le pré-traitement des signaux, la segmentation, l'extraction des descripteurs et la classification comme décrite en Figure 1. Or, dans notre cas, c'est le smartphone de l'utilisateur qui s'occupe de réaliser le pré-traitement des signaux, la segmentation et l'extraction des descripteurs, laissant au serveur applicatif seulement la tâche de classification, ce qui réduit donc fortement les coûts de calcul au niveau du serveur.
La Figure 10 décrit le temps passé sur le serveur applicatif à effectuer tous les traitements et seulement réaliser la tâche de classification dans notre cas. Pour notre jeu de données (soit 30 utilisateurs et 15 minutes de données par utilisateur), le serveur

applicatif passe presque 52 secondes à effectuer tous les traitements contre seulement 10 secondes pour notre approche. Cela représente une réduction du temps de calcul de 81%. Avec un grand nombre d'utilisateurs, cette réduction permet donc de conserver des ressources pour les opérations sur le serveur. De plus, considérant le serveur applicatif basé sur un cloud tel que Amazon EC2 services (*Amazon Elastic Compute Cloud*, s. d.), cette réduction de temps de calcul souligne aussi l'avantage économique de notre solution.
Enfin, nous avons évalué le coût de calcul de notre approche au niveau du smartphone de l'utilisateur. Ces coûts prennent en compte les traitements sur les signaux uniquement – soit le pré-traitement, la segmentation et l'extraction des descripteurs. Sur un ordinateur de base, ces opérations appliquées sur les données d'un utilisateur prennent en moyenne 2.5 secondes. Ce coût (délivré toute les 15 minutes) reste bas. De plus, ces calculs peuvent être planifiés pendant la nuit quand l'utilisateur est inactif.

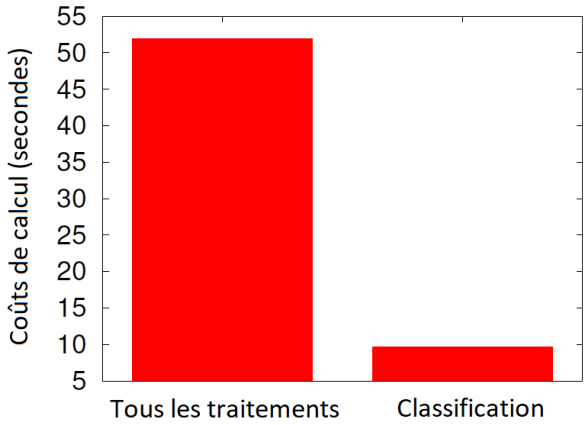


Figure 10. Coût de calcul pour le serveur applicatif: en traitant les signaux sur les smartphones des utilisateurs (soit pré-traitement, segmentation et extraction des descripteurs), notre approche réduit fortement ces coûts.

6.4. Sensibilité au bruit

Nous évaluons maintenant les performance de notre pipeline de traitements quand les signaux issus des acquisitions sont plus bruités. Dans la première base de données, l'acquisition avait été faite sur des volontaires portant le smartphone au niveau de la taille. Dans cette section, nous utilisons une nouvelle base de données (soit le "MotionSense Dataset" (Malekzadeh *et al.*, 2018)) qui fait l'acquisition des données avec un smartphone placé dans la poche de l'utilisateur. Ainsi, le smartphone étant moins contraint (mouvement possible dans la poche), les signaux acquis sont plus bruités que dans le cadre de la première base de données. En dehors de ce point, cette nouvelle base de données reste très similaire à la première. Elle contient des données issues d'un accéléromètre et d'un gyroscope en tri-axial avec une fréquence d'accquisition

de 50 Hz. Un groupe de 24 participants a réalisé 6 activités différentes: descendre des escaliers, monter des escaliers, marcher, courir, s'asseoir et se tenir debout. Toutes ces activités ont été répétées 15 fois par les utilisateurs, dont 9 fois sur une période longue (soit 2-3 minutes), et 6 fois sur une période courte (soit autour de 30 secondes à 1 minute).
Pour quantifier le bruit présent dans les signaux collectés, on a mesuré le rapport signal sur bruit (SNR) dans sa version la plus simple qui est le ratio de la moyenne du signal sur son écart-type (Welvaert, Rosseel, 2013). Ce ratio compare le niveau de signal désiré au niveau de bruit dans le fond. Le Tableau 4 montre le SNR pour chaque activité et dans les deux bases de données. Les résultats montrent une importante différence de SNR entre les deux bases qui traduit une forte présence de bruit pour chaque activité dans la 2ème base. Il est à noter que cette différence est signicativement plus importante pour les activités statiques (par exemple un SNR de 345.1 contre 1.1 pour l'activité debout). En effet, l'amplitude du signal de base dans les activités statiques est bien moins important, laissant plus d'impact au bruit.

Activité	**Base de données 1 (smartphone à la taille)**	**Base de données 2 (smartphone dans la poche)**
Descendre escaliers	3.2	2.4
Monter escaliers	4.6	1.0
S'asseoir	220.2	6.8
Être debout	345.1	1.1
Marcher	4.4	0.5

TABLE *4. Rapport signal sur bruit (SNR) pour toutes les activités dans chaque base de données: la 2ème base contient beaucoup plus de bruit, notamment pour les activités statiques.*

Nous évaluons maintenant l'impact de la présence de bruit dans les données sur les classification et donc les mesures de précision dans notre approche. Le Tableau 5 montre la mesure de précision pour la reconnaissance d'activité. Les résultats montrent que notre pipeline de traitements est encore capable de reconnaître avec de bonnes précisions les activités dynamiques (soit entre 79% et 89% de précision pour les activités de marcher et de course). La Figure 11 décrit la distribution cumulée de la précision dans le cas de la tâche de ré-identification sur la deuxième base de données. Cette distribution montre que les utilisateurs peuvent toujours être ré-identifiés même si le signal est perturbé par davantage de bruit, mais avec une précision plus faible que pour la base précédente (90% contre 48% de précision en moyenne pour respectivement la première et la nouvelle base de données).

L'impact de ce bruit peut être atténué en complétant notre pré-traitement avec un nouveau filtre. Par exemple, nous avons essayé d'ajouter un filtre de Savitzky-Golay (Savitzky, Golay, 1964) qui lisse nos signaux et augmente donc le SNR de chaque activité. Plus précisement, le filtre applique un processus de convolution en ajustant chaque sous-ensemble de données avec une courbe polynomiale de degré faible par la méthode des moindres carrés. De plus, ce changement dans l'étape de pré-traitement

Activité	Précision (activité)
Descendre escaliers	0.84
Monter escaliers	0.89
S'asseoir	0.16
Être debout	0.30
Marcher	0.80
Courir	0.79

TABLE 5. *Même si les données collectées contiennent plus de bruit, notre pipeline de traitements est encore capable de bien reconnaître les activités dynamiques, alors que l'impact du bruit réduit fortement la précision de la reconnaissance des activités statiques.*

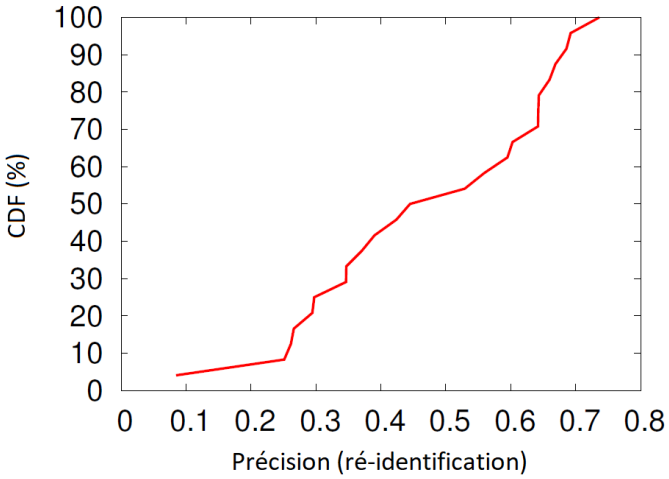


Figure 11. Distribution cumulative de la précision pour la tâche de ré-identification: les utilisateurs peuvent encore être ré-identifiés même si les signaux sont bruités, mais avec de moins bonnes précisions que sur la première base contenant moins de bruit.

a un faible impact sur les descripteurs identifiés comme les plus importants pour les tâches de classification. En appliquant la même méthodologie décrite en Section 2 pour identifier les nouveaux descripteurs les plus importants pour la reconnaissance d'activité et les ré-identification d'utilisateur (les descripteurs dans le domaine temporel restent prédominants pour discriminer les activités et les descripteurs dans le domaine fréquentiel restent prédominants pour discriminer les utilisateurs), et en utilisant notre protocole nous obtenons une précision moyenne pour la classification des activités de 97% et une précision moyenne de 37% pour la ré-identification. Ces résultats sont similaires à ceux obtenus sur la première base de données. En somme,

pour avoir les meilleures performances issues de notre pipeline de traitements, l'étape de pré-traitement doit être adaptée aux signaux d'entrée pour faire l'extraction des descripteurs. Cependant, même avec des descripteurs non optimaux, notre pipeline de traitements permet de garantir une bonne reconnaissance d'activité tout en réduisant le risque de ré-identification des utilisateurs.

7. Travaux similaires

Avec les progrès technologiques de ces dernières années, le domaine médical évolue rapidement et pose d'importants problèmes de protection de la vie privée. Par exemple, les nouvelles technologies de séquençage à haut débit ont considérablement réduit le prix et démocratisé l'analyse de l'ADN. En raison de la nature extrêmement sensible de ces données, un nouveau domaine de recherche a émergé pour traiter la quantification du risque associé à cette information et la protéger (Ayday, Humbert, 2017 ; Tramèr *et al*., 2015). L'adoption généralisée des IoT dans le domaine médical a également introduit de nouvelles questions et préoccupations en matière de sécurité et de confidentialité. Ces problèmes émergent à plusieurs étapes dans le cycle de vie des données (Rushanan *et al*., 2014). Dans la transmission de données, par exemple, (Wood *et al*., 2017) a proposé une méthode permettant de capturer le trafic réseau provenant de dispositifs IoT médicaux et de détecter automatiquement les informations en texte clair susceptibles de révéler des données médicales sensibles. Alternativement, (Aranki, Bajcsy, 2015) a présenté PDI, un cadre visant à empêcher un adversaire de déduire certaines informations sensibles sur des sujets utilisant des données cryptées qu'ils ont divulguées au cours d'une communication avec un destinataire prévu. Une autre approche, telle que l'architecture NeuroSENS (Gard *et al*., 2018), vise à améliorer la sécurité et la confidentialité du suivi de la démarche neurologique à plusieurs niveaux (stockage de données, applications mobiles et Web et transmission des données). Bien que la reconnaissance des gestes attire actuellement beaucoup d'attention (Watanabe *et al*., 2016), à notre connaissance, notre travail est le premier qui traite de la protection des données dédiées à la reconnaissance des activités au moyen de dispositifs portatifs dans le domaine médical. L'identification de descripteurs pertinents pour la reconnaissance d'activité et la ré-identification de l'utilisateur est également nouvelle. Plusieurs approches bien connues ont montré que cacher des informations d'identité explicites par pseudonymat ne suffisait pas pour garantir l'anonymat des utilisateurs (Lamberg, 2001). En effet, de nombreux critères conduisent à une identification unique des utilisateurs. Des recherches antérieures ont montré que des individus peuvent être identifiés à partir de leur mobilité (Boutet *et al*., 2016 ; Manousakas *et al*., 2018), de leurs gestes tactiles sur des appareils à écran tactile (Masood *et al*., 2018) ou de leurs navigateurs Web (Eckersley, 2010), pour n'en citer que quelques-uns. À la suite de ces études, nous démontrons également dans cet article qu'un utilisateur peut être facilement identifié à partir de son activité recueillie par des capteurs. Par rapport à d'autres approches qui masquent indépendamment chaque enregistrement (par exemple, par confidentialité différentielle (Assam *et al*., 2013)), seules les descripteurs conduisant à la ré-identification des utilisateurs sont masqués. De plus, bien que

ce masquage fondé sur une normalisation ne fournisse pas la même garantie de confidentialité que d'autres approches basées sur la généralisation (k-anonymat), l'utilité (c'est-à-dire la reconnaissance de l'activité) reste élevée tout en assurant une bonne confidentialité (c'est-à-dire un faible taux de ré-identification). Dans le cas de vérités terrain insuffisantes, (Gu *et al*., 2011 ; Yao *et al*., 2016) essayent de tirer parti des informations partagées entre différentes classes (c'est-à-dire des activités présentant des caractéristiques similaires en termes de signal) afin d'améliorer la classification. Enfin, le fractionnement des informations sensibles (l'identité des utilisateurs et leurs données) sur différents noeuds a déjà montré ses avantages en termes de confidentialité (Guha *et al*., s. d. ; Petit *et al*., 2015). En outre, en traitant les signaux à la marge du réseau sur le smartphone des utilisateurs, notre infrastructure réduit de manière inhérente les coûts opérationnels de l'application (Boutet *et al*., 2014) et renforce le contrôle des utilisateurs sur leurs données.

8. Conclusion

Nous présentons une approche préservant la confidentialité dans le contexte de la reconnaissance d'activités pour la surveillance de l'état de santé avec des appareils portables de type IoT. Notre architecture traite le signal et extrait les descripteurs pertinents localement sur le smartphone de l'utilisateur. En outre, conformément à l'observation selon laquelle le domaine de fréquence prédomine dans la tâche d'identification de l'utilisateur, une normalisation est effectuée sur les descripteurs basés sur la fréquence afin de masquer la ré-identification des utilisateurs. Enfin, seul un ensemble de descripteurs non liés à l'identité de son propriétaire est chargé dans le serveur d'application qui est alors capable de reconnaître l'activité des utilisateurs avec une grande précision tout en réduisant le risque de ré-identification de l'utilisateur. Une validation approfondie de notre approche a été réalisée sur des ensembles de données de référence, ce qui a donné de bons résultats en termes de compromis entre protection de la vie privée et utilité: une reconnaissance d'activités élevée avec peu de ré-identification de l'utilisateur.

Remerciements

Ce travail a été réalisé grâce au soutien financier du CNRS via le projet PEPS intitulé NEUROSENS dans la thématique « Objets communicants : algorithmes, architectures et applications » (OCA3).

Bibliographie

Acs G., Castelluccia C. (2014). A case study: Privacy preserving release of spatio-temporal density in paris. In *Kdd*, p. 1679–1688.

Amazon-EC2. (s. d.). *Amazon elastic compute cloud.* http://aws.amazon.com/ec2.

Anguita D., Ghio A., Oneto L., Parra X., Reyes-Ortiz J. L. (2013). A public domain dataset for human activity recognition using smartphones. In *Esann.*

Aranki D., Bajcsy R. (2015). Private disclosure of information in health tele-monitoring. *CoRR*, vol. abs/1504.07313.

Assam R., Hassani M., Seidl T. (2013). Differential private trajectory obfuscation. In *Mobiquitous*, p. 139–151.

Ayday E., Humbert M. (2017). Inference attacks against kin genomic privacy. *S&P*, vol. 15, n° 5, p. 29-37.

BenAbdelkader C., Cutler R., Davis L. (2002). Stride and cadence as a biometric in automatic person identification and verification. In *Fg*, p. 372–377.

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

Boutet A., Ben Mokhtar S., Primault V. (2016, octobre). *Uniqueness Assessment of Human Mobility on Multi-Sensor Datasets*. Research Report. LIRIS UMR CNRS 5205. Consulté sur https://hal.archives-ouvertes.fr/hal-01381986

Boutet A., Frey D., Guerraoui R., Kermarrec A.-M., Patra R. (2014). Hyrec: Leveraging browsers for scalable recommenders. In *Middleware*, p. 85–96.

Breiman L. (2001). Random forests. *Machine learning*, vol. 45, n° 1, p. 5–32.

Eckersley P. (2010). How unique is your web browser? In *Pets'10*, p. 1–18.

Frindel C., Rousseau D. (2017). How accurate are smartphone accelerometers to identify intermittent claudication? In *Healthyiot*, p. 19–25.

Gard P., Lalanne L., Ambourg A., Lesueur F., Frindel C. (2018). *Neurosens rehabilitation project.* https://neurosens.creatis.insa-lyon.fr/.

Goldreich O. (2003). Cryptography and cryptographic protocols. *Distrib. Comput.*, vol. 16, n° 2-3, p. 177–199.

Gramaglia M., Fiore M. (2015). Hiding mobile traffic fingerprints with GLOVE. In *Conext*, p. 26:1–26:13.

Gregorutti B., Michel B., Saint-Pierre P. (2017). Correlation and variable importance in random forests. *Statistics and Computing*, vol. 27, n° 3, p. 659–678.

Gu T., Wang L., Chen H., Tao X., Lu J. (2011, Nov). Recognizing multiuser activities using wireless body sensor networks. *IEEE Transactions on Mobile Computing*, vol. 10, n° 11, p. 1618-1631.

Guha S., Jain M., Padmanabhan V. N. (s. d.). Koi: A location-privacy platform for smartphone apps. In *Nsdi*, p. 183–196.

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

Han J., Pei J., Kamber M. (2011). *Data mining: concepts and techniques*. Elsevier.

James G., Witten D., Hastie T., Tibshirani R. (2013). *An introduction to statistical learning* (vol. 112). Springer.

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

Lamberg L. (2001). Confidentiality and privacy of electronic medical records. *JAMA*, vol. 285, n° 24, p. 3075-3076.

Malekzadeh M., Clegg R. G., Cavallaro A., Haddadi H. (2018, Feb). Protecting Sensory Data against Sensitive Inferences. *arXiv e-prints*, p. arXiv:1802.07802.

Manousakas D., Mascolo C., Beresford A. R., Chan D., Sharma N. (2018). Quantifying privacy loss of human mobility graph topology. *PETS*, vol. 2018, n° 3, p. 5–21.

Masood R., Zhao B. Z. H., Asghar H. J., Kâafar M. A. (2018). Touch and you're trapp(ck)ed: Quantifying the uniqueness of touch gestures for tracking. *PoPETs*, vol. 2018, n° 2, p. 122–142.

Mehrang S., Pietilä J., Korhonen I. (2018). An activity recognition framework deploying the random forest classifier and a single optical heart rate monitoring and triaxial accelerometer wrist-band. *Sensors*, vol. 18, n° 2, p. 613.

Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O. *et al*. (2011). Scikit-learn: Machine learning in python. *Journal of machine learning research*, vol. 12, n° Oct, p. 2825–2830.

Petit A., Cerqueus T., Ben Mokhtar S., Brunie L., Kosch H. (2015). PEAS: Private, Efficient and Accurate Web Search. In *TrustCom.*

Pires I. M., Garcia N. M., Pombo N., Flórez-Revuelta F. (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*, vol. 16, n° 2, p. 184.

Preece S. J., Goulermas J. Y., Kenney L. P., Howard D., Meijer K., Crompton R. (2009). Activity identification using body-mounted sensors—a review of classification techniques. *Physiological measurement*, vol. 30, n° 4, p. R1.

Reyes-Ortiz J. L. (2015). *Smartphone-based human activity recognition*. Springer.

Rushanan M., Rubin A. D., Kune D. F., Swanson C. M. (2014). Sok: Security and privacy in implantable medical devices and body area networks. In *S&P*, p. 524-539.

Savitzky A., Golay M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, vol. 36, n° 8, p. 1627–1639.

Scalvini S., Baratti D., Assoni G., Zanardini M., Comini L., Bernocchi P. (2014). *Information and communication technology in chronic diseases: a patient's opportunity.* Springer.

Sprager S., Juric M. B. (2015). Inertial sensor-based gait recognition: a review. *Sensors*, vol. 15, n° 9, p. 22089–22127.

Tang Y., Ono C. (2016). Detecting activities of daily living from low frequency power consumption data. In *Mobiquitous*, p. 38–46.

Tramèr F., Huang Z., Hubaux J.-P., Ayday E. (2015). Differential privacy with bounded priors: Reconciling utility and privacy in genome-wide association studies. In *Ccs*, p. 1286–1297.

Watanabe H., Terada T., Tsukamoto M. (2016). Gesture recognition method based on ultrasound propagation in body. In *Mobiquitous*, p. 288–289.

Welvaert M., Rosseel Y. (2013, 11). On the definition of signal-to-noise ratio and contrast-to-noise ratio for fmri data. *PLOS ONE*, vol. 8, n° 11, p. 1-10. Consulté sur https://doi.org/10.1371/journal.pone.0077089

Wood D., Apthorpe N., Feamster N. (2017). Cleartext data transmissions in consumer IoT medical devices. In *IoT S&P*, p. 7–12.

Yao L., Nie F., Sheng Q. Z., Gu T., Li X., Wang S. (2016). Learning from less for better: Semi-supervised activity recognition via shared structure discovery. In *Proceedings of the 2016 acm international joint conference on pervasive and ubiquitous computing*, p. 13–24. New York, NY, USA, ACM.