The Thirty-Second AAAI Conference
on Artificial Intelligence (AAAI-18)

WiFi-Based Human Identification via Convex Tensor Shapelet Learning

Han Zou,* Yuxun Zhou,* Jianfei Yang,† Weixi Gu,‡ Lihua Xie,† Costas J. Spanos*
*Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, USA
† School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore
‡ Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, China
Email: {hanzou, yxzhou, spanos}@berkeley.edu, {yang0478, elhxie}@ntu.edu.sg, guweixigavin@gmail.com

Abstract

We propose AutoID, a human identification system that leverages the measurements from existing WiFi-enabled Internet of Things (IoT) devices and produces the identity estimation via a novel sparse representation learning technique. The key idea is to use the unique fine-grained gait patterns of each person revealed from the WiFi Channel State Information (CSI) measurements, technically referred to as shapelet signatures, as the "fingerprint" for human identification. For this purpose, a novel OpenWrt-based IoT platform is designed to collect CSI data from commercial IoT devices. More importantly, we propose a new optimization-based shapelet learning framework for tensors, namely Convex Clustered Concurrent Shapelet Learning (C^3SL), which formulates the learning problem as a convex optimization. The global solution of C^3SL can be obtained efficiently with a generalized gradient-based algorithm, and the three concurrent regularization terms reveal the inter-dependence and the clustering effect of the CSI tensor data. Extensive experiments are conducted in multiple real-world indoor environments, showing that AutoID achieves an average human identification accuracy of 91% from a group of 20 people. As a combination of novel sensing and learning platform, AutoID attains substantial progress towards a more accurate, cost-effective and sustainable human identification system for pervasive implementations.

Introduction

Human identification, which aims to automatically associate a person with his/her identity, is a critical underpinning not only for secure authentication but also for tailoring services to each individual. In particular, it has become an important, if not indispensable, tool for many emerging applications in human-in-the-loop cyber-physical systems. In addition to automatic access authorization that vastly improves user experience, an integrated human identification system can also help achieving preventive actions to protect children and elderly people, and realizing tailor-made services in smart home, such as playing customized music, sports or TV shows, based on the identity information made available by the system.

Conventional human identification systems (e.g., biometric-based and vision-based approaches) either require the deployment of dedicated infrastructure or the active cooperation of users to carry additional devices, which are expensive, inconvenient and privacy intrusive for pervasive implementation (Ngo et al. 2014; Yang et al. 2017). For instance, using biometric signatures (e.g. fingerprint, face recognition and iris) are able to provide high identification accuracy (Gafurov 2007). However, a specific and expensive hardware has to be deployed and users' physical interaction with the equipment is required. To alleviate the issue of high cost and user involvement, Ngo et al. (2014) proposes to use accelerometers embedded in wearable devices to collect gait information for identification purposes. But requiring users to carry these devices may introduce considerable inconvenience for the person. The most common device-free human identification system is based on video monitoring from cameras (Chen et al. 2017). Nonetheless, the performance of these vision based approaches rely on appropriate lighting conditions, and more critically, they also raise privacy concerns. Thus, a non-intrusive, privacy-preserving, cost-effective and accurate human identification scheme is desired urgently.

Recently, the ubiquity of WiFi infrastructure and WiFi enabled mobile devices (MDs) have enabled a myriad of applications in context-aware services and location-based services (Zou et al. 2016; 2017a). Furthermore, with the booming development of Internet of Things (IoT), billions of WiFi enabled IoT devices, such as thermostats, sound bar, and smart TV, are en route to being widely deployed in indoor environments. Because the body movements of a human introduce variations in WiFi Received Signal Strength (RSS) measurements, device-free occupancy sensing becomes feasible by analyzing the signals (Ghourchian, Allegue-Martinez, and Precup 2017). Being a coarse measurement, nevertheless, RSS usually fails to capture the multipath effects caused by complicated human motions. Alternatively, at the physical layer, Channel State Information (CSI) describes how WiFi signal propagates from a transmitter (TX) to a receiver (RX) through multiple paths at the granularity of Frequency Division Multiplexing (OFDM) subcarriers, which is more sensitive to the presence and movements of an object and is more robust to background noise. Recent literature has witnessed many successful employments of CSI measurements for various applications, such as crowd counting (Zou et al. 2017b) and human ac-

tivity recognition (Wang et al. 2015). Being an off-the-shelf and fine-grained sensing measurement without the introduction of any extra infrastructure or user involvement, CSI data is the ideal sensing recourse for device-free and low-cost human identification.

In this paper, we propose AutoID, a human identification system that leverages the CSI measurements from existing WiFi-enabled IoT devices and produces the identity estimation via a novel sparse representation learning technique. A key observation is that a person's unique gait and body movement can be characterized by a small continuous fraction, technically known as shapelets, of the CSI measurement. Hence human identification can be readily achieved by learning the shapelets from the multi-stream, time-dependent data and then using them as the signature, or "fingerprint" of a person.

To this end, two contributions are made. Towards a better sensing technology, a novel OpenWrt-based CSI enabled IoT platform is designed for commercial IoT devices to collect CSI data from regular data frames. Towards a more effective machine learning, a new optimization-based shapelet mining method, namely Convex Clustered Concurrent Shapelet Learning (C^3SL), is proposed. C^3SL formulates shapelet learning from tensors as a convex optimization problem and establishes an efficient generalized gradient-based algorithm. Moreover, the incorporation of three concurrent regularization terms enables the automatic learning of the inter-dependence and the clustering effect of time series CSI tensor data. To the best of our knowledge, AutoID is the first human identification platform realized with shapelets in CSI measurements. The C^3SL is the first convex shapelet learning framework that avoids the costly combinatorial search over subsequences, making the shapelet technique scalable to real-world big-data scenarios. It also extends shapelet learning to data sets in the form of 3-way tensors (multiple time series). Experiments are conducted in multiple real-world environments and the results justify the idea of using CSI shapelet as the fingerprint, as well as the effectiveness of the C^3SL framework.

Related Work

Human Identification based on Gait Information and RF Signals

Gait has been acknowledged as one of unique biometric signature for human identification (Gafurov 2007). Existing approaches utilize various sensing techniques (e.g., cameras (Yang et al. 2017) and wearable sensors (Ngo et al. 2014)) to infer gait information. The vision based scheme requires line-of-sight (LOS) for monitoring and their performances degrade severely under non-ideal situations (Yang et al. 2017). Moreover, image processing algorithms usually introduce high computational overhead and continuous monitoring with camera raises privacy concerns. The accelerometers equipped in MDs are also leveraged to capture gait signature (Ngo et al. 2014). The accuracy of these approaches decreases significantly when the user does not wear the sensing device correctly. All these schemes either require dedicated sensors installation or inconvenient wearable sensor attachment, which introduce additional cost and limit large-scale deployment. Thus, a device-free, cost-effective and robust scheme for pervasive human identification is of great interest.

Several RF based human identification systems have been proposed in recent decades by analyzing the Doppler shift of the reflected RF signal caused by human movement from dedicated RF platform, such as Frequency-Modulated Continuous-Wave (FMCW) radars. However, the high costs of dedicated infrastructure limit their practical implementation. With the pervasive availability of WiFi infrastructure in buildings and the proliferation of IoT, the ubiquitousness of WiFi communications among them provides a precious opportunity for device-free human identification by analyzing WiFi signals. Some researchers have explored CSI measurements in the physical layer from WiFi NIC cards for human identification. For instance, a system called WiWho is proposed by Zeng, Pathak, and Mohapatra (2016), that calculates 23 statistical features from time domain and frequency domain based on the CSI readings to extract human gait information for identification. It provides 92% to 80% recognition accuracy for 2 to 6 human subjects. However, it requires the user to walk on a pre-determined path that near to a TX-RX pair. Wang, Liu, and Shahzad (2016) generates spectrograms from CSI measurements to extract 170 features related to the walking pattern for human identification. It achieves 79.28% recognition accuracy but TX and RX have to be placed close to each other (1.6m), which is impractical in most indoor environments.

Shapelet-based Classification

Shapelet, first proposed by Ye and Keogh (2009), is defined as local, phase-independent sub-sequences of time series that features the maximal discriminative power for time series classification. Given a set of shapelets, one can compute the similarity between shapelets and time series, and then use the similarity as a discriminatory feature for a more accurate and interpretable time series classification. The rationale behind shapelets is that different classes of time series can usually be distinguished by their local variations instead of their global structure (Hou, Kwok, and Zurada 2016). Being a sparse representation of the original training data, it is also able to reveal the inherent structure of the data with intuitive interpretation and provide explanatory insights. Since the length of a shapelet is usually much shorter than the original time series, the time and space required for classification can be reduced significantly (Mueen, Keogh, and Young 2011). With the high prediction accuracy and the interpretability of features, shapelet-based classification has been implemented in various applications, such as gesture recognition (Hartmann and Link 2010) and clustering (Zakaria, Mueen, and Keogh 2012).

However, the learning of shapelets from a training dataset is a non-trivial task. To obtain the shapelet to represent a time series, the original work (Ye and Keogh 2009) performs searches over all the possible subsequences, and the candidate with the maximal discriminative power was selected into the shapelet set. The major issue with the combinatorial search is the computational cost, which prohibits the usage

of shapelet classification for large or even median-sized data set. Although numerous speed up schemes, such as early abandoning of distance computations (Lines et al. 2012), entropy pruning of information gain (Ye and Keogh 2009), and projections on the SAX representation (Rakthanmanon and Keogh 2013), have been designed to reduce the computational cost, those search-based methods are still not scalable in worst case with the length and the number of training series. In Hou, Kwok, and Zurada (2016), the shapelet discovery task is reformulated as a numerical optimization problem, which combines the generalized eigenvector method and fused lasso regularizer to learn the shapelet positions. However, the formulation is non-convex and the solution quality is not guaranteed. More importantly, all previous shapelet learning methods can only handle single time sequence data as a sample point, while the CSI data provides multiple time series measurement as a sample point for each person. All the above considerations motivate us to develop a new scalable and tensor compatible shapelet learning method.

System Design of AutoID

CSI enabled IoT Platform

With the pervasiveness of WiFi traffic among IoT devices, it is foreseeable that in the near future TX-RX networks will be ubiquitous allowing device-free non-intrusive human identification. In indoor environments, WiFi signals propagate through multiple paths from TX to RX due to reflection, scattering and diffraction introduced by walls, doors, furniture, as well as the presence and movements of occupants (Wang, Liu, and Shahzad 2016). Nowadays, most of the commercial off-the-shelf (COTS) WiFi devices are equipped with multiple antennas for multiple input, multiple output (MIMO) communication and at the physical layer adopt orthogonal frequency-division multiplexing (OFDM) that supports the IEEE 802.11n/ac standard. In order to increase the data rate, TX divides the data and sends it through a number of narrow subcarriers at different frequencies simultaneously. Different from the RSS which only captures the superimposition of multipath signals, CSI reveals fine-grained information about how the signal is propagated and interfered, including different time delays, amplitude attenuation, and phase shift of multiple paths on each subcarrier. Analyzing these signal propagation variations caused by human motions makes device-free human identification feasible. In a nutshell, the signal can be modeled as a Channel Impulse Response (CIR) $h(\tau)$ and the OFDM receiver is able to provide a sampled version of the signal spectrum of each subcarrier in the frequency domain, which contains both amplitude attenuation and phase shift as complex numbers. These measurements can be summarized as CSI:

$$H_i = \|H_i\| e^{j\angle H_i} \tag{1}$$

where $\|H_i\|$ and $\angle H_i$ denote the amplitude and the phase of the CSI at the i^{th} subcarrier respectively.

Conventional CSI-based sensing systems adopt the Intel 5300 NIC tool (Halperin et al. 2011) to extract the CSI data from PCs or laptops with modified WiFi NIC cards. Requiring laptops as receivers severely hinder the above systems from large-scale implementation. To overcome this bottleneck, we develop a CSI enabled IoT platform so that the CSI readings can be reported directly from the COTS IoT devices, such as commodity WiFi routers or thermometers. OpenWrt is chosen as the OS for our platform since it is a lightweight and widely used Linux OS for embedded devices. We upgrade the Atheros CSI Tool (Xie, Li, and Li 2015) and develop a new OpenWrt firmware for commercial IoT devices so that the CSI measurements from regular data frames transmitted in the existing traffic can be obtained directly from them. Moreover, the traditional CSI tool only provides CSI readings from 30 subcarriers. Our platform reports CSI data on all the 114 subcarriers for 40 MHz bandwidth on each TX-RX pair when both TX and RX IoT devices operate on 5 GHz, which provides much more information than conventional CSI tools. N_{TX} and N_{RX} represent the number of transmitting and receiving antennas, respectively. At each time instance, $N_{\mathrm{TX}} \times N_{\mathrm{RX}} \times 114$ CSI streams are available, characterizing the variations of WiFi communication links caused by human presence and movement in a much more refined manner.

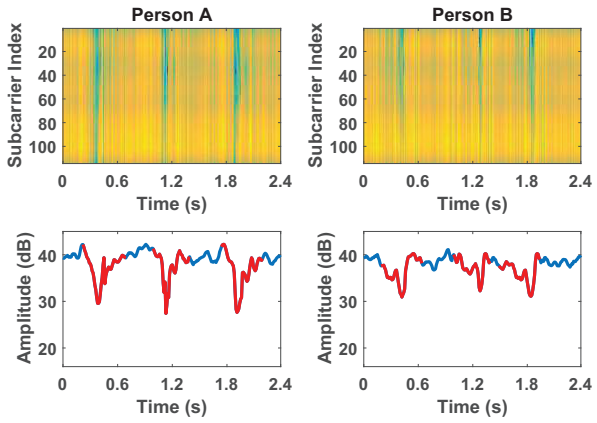


Figure 1: CSI amplitude readings of two human subjects.

CSI Shapelet based Human Identification: Intuition

We conducted a motivational experiment with multiple human subjects walking in a conference room. Two IoT devices (TPLINK N750 routers) were used to collect CSI measurements to explore the potentials of CSI data for human identification. Both devices were upgraded to our CSI enabled IoT platform and form a TX-RX pair. Figure 1 depicts the CSI amplitude readings (after Discrete Wavelet Transform (DWT) based de-noising) when Person A and Person B were walking three steps along the same path. As shown in Figure 1, the periodical patterns of peak and valley, which indicate the gait cycles, can be easily observed from the CSI readings for both human subjects. Moreover, the shape of each step (highlighted in red) for the same person displays evident similarity, while the shapes of the gait cycle and step lengths for disparate individuals exhibit noticeable distinction. The above observation implies that the unique

gait information of each individual can be extracted from the CSI time series data and characterized by sub-sequences at critical times, known as shapelets. CSI shapelet provides a sparse and unique representation of the high-resolution CSI data obtained from a person, like a fingerprint. Also according to the biometric research conducted by Ngo et al. (2014), the gait cycle contains unique information that can be used as a biometric signature to identify the person. All above suggests mining CSI shapelets and storing them in a database to build classifiers for human identification.

A Convex Clustered Concurrent Shapelet Learning (C^{3}SL) Framework

Despite the aforementioned benefits of using shapelet for human identification, the learning of effective shapelets from training data is always a challenging task. Most search-based shapelet learning algorithms are only feasible for a small number of short time series, even with various early stopping, pruning or lazy neighboring techniques incorporated. Hou, Kwok, and Zurada (2016) were the first to propose optimization-based shapelet learning, however, their method requires solving a non-convex problem. More importantly, all previous works only consider one time series as a sample point, in many practical applications and with the CSI dataset in particular, each sample consists of multiple sequences and the overall data essentially forms a 3-way tensor. Hence an additional challenge is to encompass the interaction among multiple sequences for shapelet learning.

In this part we propose a novel optimization-based shapelet learning framework, namely Clustered Concurrent Shapelet Learning (C^{3}SL) Algorithm. Compared to existing shapelet learning methods, C^{3}SL is advantageous in that (1) The formulation is convex and can be solved globally and efficiently with a generalized gradient-based algorithm, and (2) the incorporation of three concurrent regularization terms enables the automatically learning of the interdependence and the clustering effect of multiple time series.

A Convex Formulation

The CSI data, or more generally a 3-way tensor is denoted as follows:

$$X \in \mathbb{R}^{M \times N \times T} \tag{2}$$

where M is the number of subcarriers, N is the number of samples obtained and T is the length of the measurement in each experiment. Notation-wise, we further use $X_{(m)}$ to denote the $N \times T$ matrix containing all measurements from the subcarrier m. Its (i^{th}) row is denoted by $X_{(m),i}$ and its (i, j^{th}) element by $X_{(m),ij}$.

Our goal is to learn a "Shapelet Coefficient" matrix $W \in \mathbb{R}^{T \times M}$, which indicates the location and strength of all shapelets in X. For clarity, we also denote $W = [w_{(1)}, w_{(2)}, \ldots, w_{(m)}]$, i.e., each $w_{(m)}$ is a $T \times 1$ vector containing the shapelet coefficient for subcarrier m (or task m).

Now consider the following learning problem:

$$\begin{aligned}\min_{\substack{W \in \mathbb{R}^{T \times M} \\ \Sigma \in S^M}} & \sum_{m=1}^{M} \left\| 1 - y \circ X_{(m)} W_{(m)} \right\|_H \\ & + \lambda \text{trace} \left\{ W \Pi \Sigma^{-1} (\Pi W)^{\mathrm{T}} \right\} \\ & + \rho_1 \|W\|_1 + \rho_2 \|RW\|_1 + \rho_3 \|W\|_{2,1}\end{aligned}$$

where y is the label of the human subject, which is a binary class indicator variable $y \in \{-1, 1\}$, the $\circ$ operater indicates the element-wise product, $\|\cdot\|_H$ is the "Hinge Norm" for a vector of dimension N, i.e., $\|v\|_H = \sum_{i=1}^{N} \max\{0, v_i\}$. The centering projection matrix $\Pi \triangleq I - U$, where $U = \frac{\mathbb{I}\mathbb{I}^{\mathrm{T}}}{M}$ and $\mathbb{I}$ is a vector of all ones with dimension N. $\|W\|_1 \triangleq \sum_{m=1}^{M} \sum_{t=1}^{T} \|W_{tm}\|$ is the Lasso Regularization for a matrix, and $\|W\|_{2,1} \triangleq \sum_{t=1}^{T} \sqrt{\sum_{m=1}^{M} W_{tm}^2}$ is the Grouped Lasso Regularization. R is a $(T-1) \times T$ matrix in which $R_{i,i} = 1, R_{i,i+1} = -1$, and all the other elements are zeros. Hence we have,

$$\|RW\|_1 = \sum_{m=1}^{M} \sum_{t=1}^{T-1} \|W_{m,i+1} - W_{m,i}\|$$

The intuition for the above formulation is the following. We can regard the first term as the "classification loss" for the tensor data. Because a hinge loss is used for each sample, this loss term is similar to that of SVM classification and is convex. The second term induces "clustering" among learning tasks since it can be considered as an approximation of the "minimal within-cluster maximal between-cluster" criterion (Zhou, Chen, and Ye 2011). The matrix $\Sigma \in S^M$ is a $M \times M$ symmetric positive definite matrix. It has been shown in Jacob, Vert, and Bach; Zhou, Chen, and Ye (2009; 2011) that this penalty can automatically capture the grouping of tasks and ensure that only related tasks (subcarriers) are used to improve the shapelet learning. Moreover, it is known that the problem

$$\min_{\substack{W \in \mathbb{R}^{T \times M} \\ \Sigma \in S^M}} \text{trace} \left\{ W \Pi \Sigma^{-1} (\Pi W)^{\mathrm{T}} \right\}$$

is convex (Boyd and Vandenberghe 2004). The combined usage of the Lasso and Grouped Lasso Regularization allows the learning of "sparse grouped" features, and is widely used in Multi-task Learning literature for sparse feature selection (Argyriou, Evgeniou, and Pontil 2007; Zhou et al. 2011). Finally, the term $\|RW\|_1$, sometimes called fused Lasso Regularization, imposes the continuity of the learned patterns. This is important as shapelet by definition should be a subsequence instead of a set of disjoint features.

To sum up, the proposed learning objective minimizes the regularized classification loss, with regularization terms including (1) the grouping of tasks (2) the sparsity of learned shapelet, and (3) the continuity imposed by the definition of shapelet. In addition, the above learning problem is convex.

$$\min_{W\in\mathbb{R}^{T\times M}} \underbrace{\sum_{m=1}^{M}\left\|1-y\circ X_{(m)}W_{(m)}\right\|_H^2+\min_{\Sigma\in S^M}\lambda trace\left\{W\Pi\Sigma^{-1}(\Pi W)^{\mathrm{T}}\right\}}_{g(W)}+\underbrace{\rho_1\|W\|_1+\rho_2\|RW\|_1+\rho_3\|W\|_{2,1}}_{h(W)} \quad (3)$$

Generalized Gradient-based Optimization

To solve the above learning problem, we adopt the Accelerated Generalized Gradient Method (AGGM) to achieve fast learning for large scale problems. The AGGM is one of the best first-order methods that solves optimization of the form $\min_{x\in\mathbb{R}^N} g(x)+h(x)$ for $g(x)$ convex differentiable and $h(x)$ convex (Chen et al. 2009). With some initial x^0, x^{-1}, the AGGM repeats for $k=1,2,3,\ldots$ that

$$\begin{aligned} y &= x^{(k-1)}+\frac{k-2}{k-1}(x^{(k-1)}-x^{(k-2)}) \\ x^{(k)} &= prox_{\tau_k}(y-\tau_k\nabla g(y)) \end{aligned} \quad (4)$$

where $prox_t(x)=\arg\min_Z \frac{1}{2t}\|x-z\|^2+h(z)$ and ∇ is the gradient operator. This algorithm achieves a $O(\frac{1}{k^2})$ convergence rate.

In order to apply AGGM, we consider modifying the hinge loss into a squared hinge loss that ensures the differentiability, and rewrite the problem into (3). The gradient of the first term in $g(W)$ is straightforward and can be obtained in an explicit form. The gradient of the second term fits in the "Parametric Dual Maximization" framework (Zhou, Kang, and Spanos 2017), where the gradient can be obtained by regarding W as the parameter of the dual problem of the term. More specifically, we have with singular value decomposition,

$$\min_{\sum\in S^M}\text{trace}\left\{W\Pi\Sigma^{-1}(\Pi W)^{\mathrm{T}}\right\}=\min_{\substack{\lambda\in\mathbb{R}^M\\ \alpha\le\lambda_m\le\beta\\ I^{\mathrm{T}}\lambda=\gamma}}\sum_{m=1}^{M}\frac{\sigma_m^2}{\lambda_m} \quad (5)$$

and the dual problem reads

$$\begin{aligned} &\max_{v\ge 0}\Sigma_i 2\sigma_i v I(\alpha\sqrt{v}\le\sigma_i\le\beta\sqrt{v}) \\ &+\sum_i(\frac{\sigma_i^2}{\alpha}+v\alpha)I(\sigma_i\le\alpha\sqrt{v}) \\ &\sum_i(\frac{\sigma_i^2}{\beta}+v\beta)I(\beta\sqrt{v}\le\sigma_i)-v\gamma \end{aligned} \quad (6)$$

In addition, one can show that the gradient of this term can be obtained by using the solution of the dual problem through

$$\frac{\delta trace\left\{W\Pi\Sigma^{-1}(\Pi W)^{\mathrm{T}}\right\}}{\delta\sigma_i}=\frac{2\sigma_i}{\lambda_i}$$

.

Concerning the computation of the proximal operation associated with the non-smooth regularizations, we first observe that the problem

$$\begin{aligned} prox(V) &= \arg\min_W \frac{1}{2}\|W-V\|_F \\ &+\rho_1\|W\|_1+\rho_2\|RW\|_1+\rho_3\|W\|_{2,1} \end{aligned} \quad (7)$$

decouples in rows of W. Hence to get the ith row W_i, we only need to solve

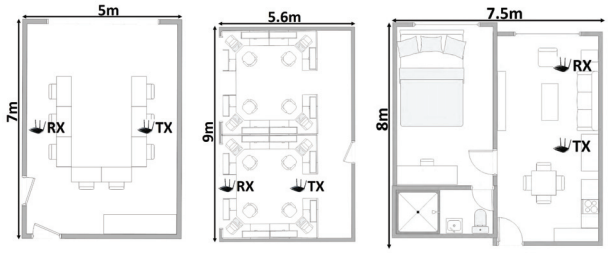
$$\begin{aligned} prox(V) &= \arg\min_W \frac{1}{2}\|W_i-V_i\|_F \\ &+\rho_1\|W_i\|_1+\rho_2\|RW_i\|_1+\rho_3\|W_i\|_{2,1} \end{aligned} \quad (8)$$

which can be solved efficiently using the RBCD algorithm proposed in Zhou (2017) or the decomposed projection method proposed in Zhou et al. (2012).

Experiments

We implemented AutoID using two TPLINK N750 routers: one serving as TX and the other as RX (each one with 3 external omnidirectional antennas). We upgraded their firmware to our CSI enabled IoT platform so that the CSI measurements from regular data frames are reported directly from the RX. TX operated in 802.11n AP mode and RX connected to the TX's network in client mode. TX was operated on 5 GHz frequency band because it has less interference and higher distance resolution due to shorter wavelength compared to 2.4 GHz band. Furthermore, we leveraged the 40 MHz channel bandwidth since a larger bandwidth introduces more CSI measurements (114 subcarriers), which increases the chance to capture the detailed small-scale fading effects caused by subtle human motions. There were no potential hazards to occupants because TX operated in default transmission power which totally met the FCC regulations. Existing WiFi networks such as campus network were operated as usual and other WiFi MDs coexisted during the entire experiments. The sampling rate was 700 packets/s and linear interpolation was adopted to ensure the stationary interval of consecutive CSI values when there was a packet loss.

To validate the human identification performance of AutoID, 20 human subjects (14 male and 6 female graduate students) with similar ages in the range of 23-28 years, participated in the experiments. We conducted the experiments in three typical indoor environments, including a conference room ($5m\times 7m$), an office zone ($5.6m\times 9m$) and a 1-bedroom apartment ($7.5m\times 8m$), to evaluate the generality of AutoID. Figure 2 demonstrates the floor plans of the



(a) Meeting room (b) Office (c) Apartment

Figure 2: Floor plans of the three indoor environments.

locations as well as the locations of TX and RX routers. As illustrated in Figure 2, three locations have distinct furniture layouts and sizes. In our experiments, TX and RX routers were 3 meters apart and placed on tripods at height of 1.5 meters. In each location, all the 20 subjects walked at arbitrary pace through the LOS of TX-RX pair for 10 times on one day (T1) to construct the model and another 10 times on another day (T2) to evaluate the performance of AutoID under both temporal and environmental dynamics. All the CSI data were collected anonymously to preserve occupants' privacy. Since the walking pattern of a human subject is related to his or her physiological characteristics (e.g. gender, height, and weight), these characteristics of each participant were also recorded in an anonymized database with their approvals. Due to the page limit, detailed participant information is presented in the supplementary.

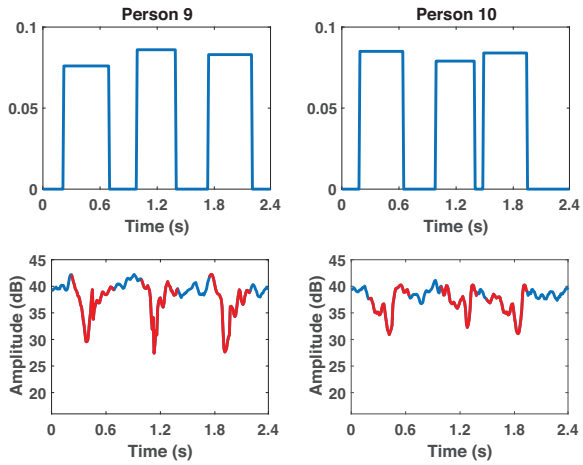


Figure 3: Shapelet indicator vectors and learned shapelets.

Evaluation

Shapelet Mining Results

We first evaluate whether the proposed C^{3}SL framework can extract unique shapelet from the high-resolution CSI measurements as a signature for each person. Figure 3 depicts both the learned shapelets and the shapelet indicator vectors of Person 9 and 10 as an example. As shown in Figure 3, the lengths of the most of shapelets are around 0.5s which is similar to the step cycle time for adults. justifying that C^{3}SL can precisely extract the gait cycles from the CSI data. Due to the page limit, Figure 4, only demonstrates the shapelets learned from 5 distinct human subjects on 5 different subcarriers in the three indoor environments (Figure 2). It is observed from Figure 4, that the learned shapelets of various persons are quite different from each other. Moreover, the shapelets learned by C^{3}SL are robust (almost invariant) across different time and environmental settings, hence can be used as a unique gait signature to identify each person.

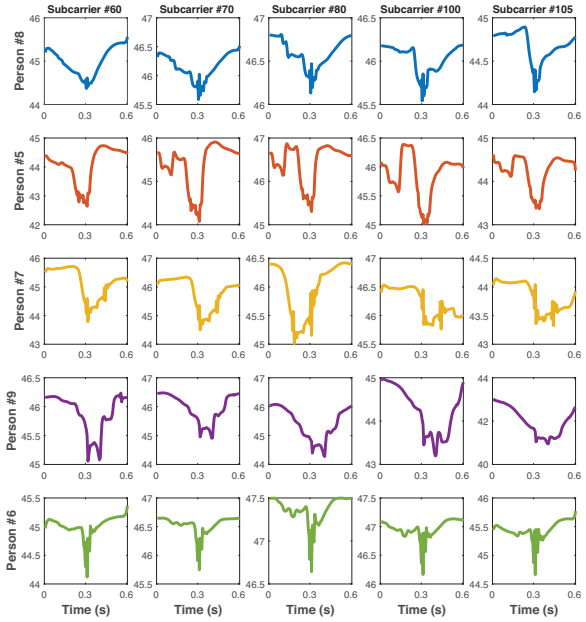


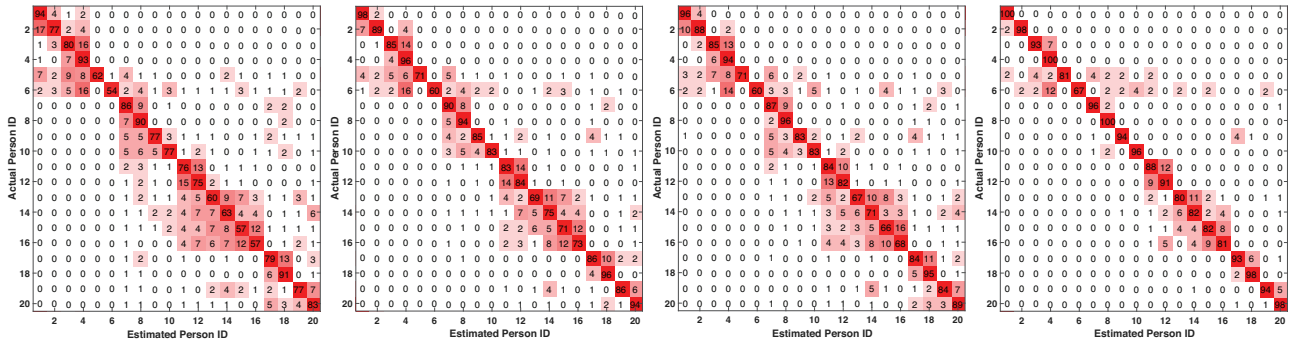
Figure 4: Shapelets of distinct people.

Identification Performance

To evaluate the human identification performance of AutoID, we first utilized the 200 walking traces of the 20 human subjects collected during T1 to train a classifier using C^{3}SL, and test its accuracy with the 200 walking traces obtained at T2. Two evaluation metrics are adopted, the overall accuracy, i.e., $P(\hat{y} == y)$, and the confusion matrix. Each row of the confusion matrix represents the estimated person ID while each column represents the actual (true) identities, i.e., entry (i, j) in the matrix represents the number of the times person j (the true identity) was classified as the person i.

Table 1: Performance evaluation of different methods.

Approach	Identification Accuracy (%)	Training Time (s)	Classification Time (s)
WiWho	73.18	76.72	1.33
WifiU	75.24	201.45	1.64
KDD14	83.42	966.86	0.95
AAAI16	81.59	127.31	0.91
AutoID	90.77	92.75	0.87



(a) WifiU (b) KDD14 (c) AAAI16 (d) AutoID

Figure 5: Confusion matrix of human identification with 20 people using different methods.

The performance of AutoID is compared to two state-of-the-art CSI-based human identification systems (WiWho (Zeng, Pathak, and Mohapatra 2016) and WifiU (Wang, Liu, and Shahzad 2016)), and two search-based shapelet learning methods (KDD14 (Grabocka et al. 2014), as well as the most recent work of AAAI16 (Hou, Kwok, and Zurada 2016)). Table 1 presents the average testing human identification accuracy of the five methods. AutoID achieves an average testing accuracy of 90.77% in a group of 20 people, which improves the overall accuracy by 17.6% and 15.5% over WiWho and WifiU, and 7.4% and 9.2% over KDD14 and AAAI16, respectively.

Figure 5 depicts the confusion matrices of different methods. It is seen that the performance of AutoID is the best. Since the gait information of a person is related to his or her weight, height and age, as shown in Figure 5(d), Person 12 is misclassified with 9% to be Person 11 because both of them are female students with similar height. Similarly, Person 15 and Person 16 are misclassified to each other with 8.5% because they have similar weight. In general, AutoID can provide satisfactory human identification accuracy in a device-free and privacy-preserving manner.

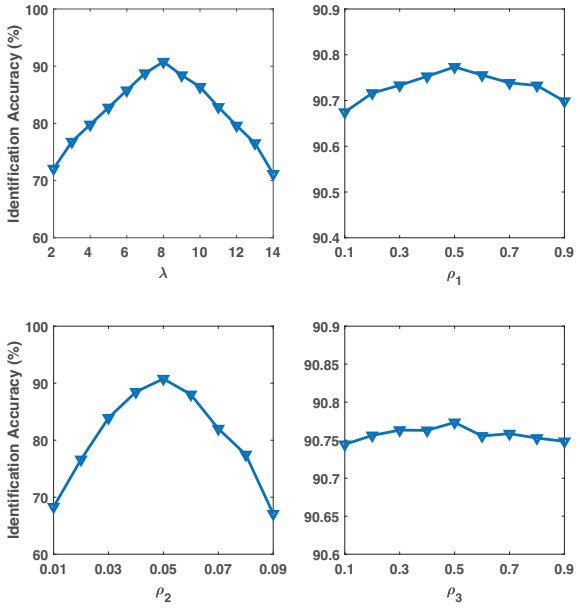


Figure 6: Sensitivity analysis of the key parameters.

Computation Time and Sensitivity Analysis

We also evaluated the efficiency of AutoID in terms of training and testing time. Table 1 illustrates these two performance metrics along with other methods. As far as shapelet learning is concerned, C^{3}SL outperforms KDD14 and AAAI16, mainly because it is an optimization based and is much more efficient than the search based KDD14. Besides, due to our convex learning formulation, it obtains the optimal shapelet and requires less time with the help of AGGD when compared to AAAI16. It also can be observed from Table 1 that shapelet-based learning methods consume less time for online classification than feature engineering based methods in general because they only need sparse subsequences of the entire time series (shapelet) for prediction.

The choice of the four hyperparameters, $\lambda, \rho_1, \rho_2, \rho_3$, in the objective equation 3 are crucial for the performance of C^{3}SL. In all the above experiments we have used an optimal value obtained from 10-folds cross validation (CV). To further illustrate their impact, Figure 6 depicts the identification accuracy with different values of the four hyperparameters. It appears that the testing performance is more sensitive to λ and ρ_2 while exhibiting less variation with different choices of ρ_2, ρ_3. This not only suggests that those four hyper-parameters should be tuned with different CV grid, but also implies that the first regularization term for "clustering" and the fused lasso regularization for continuity play a more important role in C^{3}SL.

Conclusion

In this paper, we proposed AutoID, a human identification system that leverages CSI data from existing WiFi-enabled IoT devices via a novel sparse representation learning technique. The designed CSI enabled IoT platform can collect CSI data from COTS IoT devices and the proposed C^{3}SL shapelet learning framework is able to extract the unique fine-grained gait patterns as shapelet signature of each per-

son for human identification. Experimental results have validated that AutoID can uniquely identify a person with average accuracy of 91% from a group of 20 people.

Acknowledgments

This research is funded by the Republic of Singapore National Research Foundation (NRF) through a grant to the Berkeley Education Alliance for Research in Singapore (BEARS) for the Singapore-Berkeley Building Efficiency and Sustainability in the Tropics (SinBerBEST) Program. BEARS has been established by the UC Berkeley as a center for intellectual excellence in research and education in Singapore.

References

Argyriou, A.; Evgeniou, T.; and Pontil, M. 2007. Multi-task feature learning. In *Advances in neural information processing systems*, 41–48.

Boyd, S., and Vandenberghe, L. 2004. *Convex optimization*. Cambridge university press.

Chen, X.; Pan, W.; Kwok, J. T.; and Carbonell, J. G. 2009. Accelerated gradient method for multi-task sparse learning problem. In *9th IEEE International Conference on Data Mining*, 746–751. IEEE.

Chen, W.; Chen, X.; Zhang, J.; and Huang, K. 2017. A multi-task deep network for person re-identification. In *AAAI*, 3988–3994.

Gafurov, D. 2007. A survey of biometric gait recognition: Approaches, security and challenges. In *Annual Norwegian computer science conference*, 19–21. Citeseer.

Ghourchian, N.; Allegue-Martinez, M.; and Precup, D. 2017. Real-time indoor localization in smart homes using semi-supervised learning. In *AAAI*, 4670–4677.

Grabocka, J.; Schilling, N.; Wistuba, M.; and Schmidt-Thieme, L. 2014. Learning time-series shapelets. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 392–401. ACM.

Halperin, D.; Hu, W.; Sheth, A.; and Wetherall, D. 2011. Tool release: gathering 802.11 n traces with channel state information. *ACM SIGCOMM Computer Communication Review*.

Hartmann, B., and Link, N. 2010. Gesture recognition with inertial sensors and optimized dtw prototypes. In *2010 IEEE International Conference on Systems Man and Cybernetics*, 2102–2109. IEEE.

Hou, L.; Kwok, J. T.; and Zurada, J. M. 2016. Efficient learning of timeseries shapelets. In *AAAI*.

Jacob, L.; Vert, J.-p.; and Bach, F. R. 2009. Clustered multi-task learning: A convex formulation. In *Advances in neural information processing systems*, 745–752.

Lines, J.; Davis, L. M.; Hills, J.; and Bagnall, A. 2012. A shapelet transform for time series classification. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, 289–297. ACM.

Mueen, A.; Keogh, E.; and Young, N. 2011. Logical-shapelets: an expressive primitive for time series classification. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1154–1162. ACM.

Ngo, T. T.; Makihara, Y.; Nagahara, H.; Mukaigawa, Y.; and Yagi, Y. 2014. The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication. *Pattern Recognition* 47(1):228–237.

Rakthanmanon, T., and Keogh, E. 2013. Fast shapelets: A scalable algorithm for discovering time series shapelets. In *Proceedings of the 2013 SIAM International Conference on Data Mining*, 668–676. SIAM.

Wang, W.; Liu, A. X.; Shahzad, M.; Ling, K.; and Lu, S. 2015. Understanding and modeling of wifi signal based human activity recognition. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, 65–76. ACM.

Wang, W.; Liu, A. X.; and Shahzad, M. 2016. Gait recognition using wifi signals. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 363–373. ACM.

Xie, Y.; Li, Z.; and Li, M. 2015. Precise power delay profiling with commodity wifi. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, 53–64. ACM.

Yang, Y.; Wen, L.; Lyu, S.; and Li, S. Z. 2017. Unsupervised learning of multi-level descriptors for person re-identification. In *AAAI*, 4306–4312.

Ye, L., and Keogh, E. 2009. Time series shapelets: a new primitive for data mining. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 947–956. ACM.

Zakaria, J.; Mueen, A.; and Keogh, E. 2012. Clustering time series using unsupervised-shapelets. In *IEEE 12th International Conference on Data Mining*, 785–794. IEEE.

Zeng, Y.; Pathak, P. H.; and Mohapatra, P. 2016. Wiwho: wifi-based person identification in smart spaces. In *Proceedings of the 15th International Conference on Information Processing in Sensor Networks*, 4. IEEE Press.

Zhou, J.; Yuan, L.; Liu, J.; and Ye, J. 2011. A multi-task learning formulation for predicting disease progression. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, 814–822. ACM.

Zhou, J.; Liu, J.; Narayan, V. A.; and Ye, J. 2012. Modeling disease progression via fused sparse group lasso. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1095–1103. ACM.

Zhou, J.; Chen, J.; and Ye, J. 2011. Clustered multi-task learning via alternating structure optimization. In *Advances in neural information processing systems*, 702–710.

Zhou, Y.; Kang, Z.; and Spanos, C. J. 2017. Parametric dual maximization for non-convex learning problems. In *AAAI*, 2956–2962.

Zhou, Y. 2017. Statistical learning for sparse sensing and agile operation.

Zou, H.; Huang, B.; Lu, X.; Jiang, H.; and Xie, L. 2016. A robust indoor positioning system based on the procrustes analysis and weighted extreme learning machine. *IEEE Transactions on Wireless Communications* 15(2):1252–1266.

Zou, H.; Jin, M.; Jiang, H.; Xie, L.; and Spanos, C. J. 2017a. Winips: Wifi-based non-intrusive indoor positioning system with online radio map construction and adaptation. *IEEE Transactions on Wireless Communications*.

Zou, H.; Zhou, Y.; Yang, J.; Gu, W.; Xie, L.; and Spanos, C. 2017b. Freecount: Device-free crowd counting with commodity wifi. In *2017 IEEE Global Communications Conference*. IEEE.