Towards Robust & Realtime Human Activity Recognition Using Wearable Sensors

by

Delaram Yazdansepas

(Under the Direction of Lakshmish Ramaswamy)

Abstract

With the proliferation of smartphones and fitness bands that have various sensors such as accelerometers, wearable sensor-based Human Activity Recognition (HAR) systems have gained wide popularity and researchers have proposed numerous techniques for recognition of these activities. Human activity recognition has many applications particularly in health care, cognitive assistance, city planning, indoor localization and tracking, and human-computer interaction. Although there has been some progress, a practical robust HAR system remains elusive because the collected data are affected by several factors such as noise, data alignment, and other constraints. In addition, the variability in the sensing equipment and their displacement is a practical challenge for implementing HAR in real-world applications.

This dissertation explores the twin problems of making wearable sensor-based HAR systems robust and real time. Towards enhancing the robustness of ML-based HAR systems, we adopt feature selection methods on time and frequency domain features and apply classifiers for evaluating the recognition performance. We show the effect of different feature sets on each of the classifiers and further demonstrate in our results the impact of decreasing the size of the training set on the accuracy of the classifiers. Towards building an Online HAR system, this thesis explores the concept of Shapelets to avoid complex feature extraction. We propose a procedure to find the most representative shapelet for each activity class based on time series distance metrics and dynamic time warping. Furthermore, we generate a personalized shapelet library database driven from users' activity time series.

We evaluate the proposed algorithm and techniques using a dataset comprised of accelerometer readings of 77 individuals performing various activities such as walking/jogging on treadmill, walking on different surfaces, climbing stairs, and non-ambulatory activities. Our experiments demonstrate that by using selected features from the time and frequency domain, we can achieve higher accuracy rates if we limit the training and testing sets to specific age groups. Furthermore, while we mainly use a single hip-worn accelerometer sensor as our sensing device, we show our method could support any wearable accelerometer sensor.

INDEX WORDS: Human Activity, Times Series, Shapelets, DWT, Wavelets, Data mining, Feature selection, Classification, Decision tree, Health Informatics

Towards Robust & Realtime Human Activity Recognition Using Wearable Sensors

by

Delaram Yazdansepas

B.S., Qazvin Azad University, 2011

A Dissertation Submitted to the Graduate Faculty
of The University of Georgia in Partial Fulfillment
of the
Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2017

Towards Robust & Realtime Human Activity Recognition Using Wearable Sensors

by

Delaram Yazdansepas

Major Professor: Lakshmish Ramaswamy

Committee: Hamid R. Arabnia
Khaled Rasheed

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2017

Dedicated to my parents,

Mastaneh & Amir,

who instilled in me the virtues of perseverance and

commitment, who encouraged me to strive for excellence

Acknowledgments

Persian poet, Attar of Nishapur, records the tale of a powerful king who asks his wise men to create a ring that will make him happy when he is sad. After deliberation the wise men hand him a simple ring with the words "This too shall pass" (Farsi: این نیز بگذرد) engraved inside it, which has the desired effect to make him happy when he was sad. This ring, however, became a curse for times he was happy. This simple phrase reminds us to look to the future during hard or challenging times and to also appreciate and be joyous during good times, because neither may last. During my years spent at the University of Georgia as a graduate student, I have been strongly supported by many people, some of them reminding me during particularly challenging times that indeed "this too shall pass." It is these individuals who I will lament parting with and who I wish to recognize them as this "good time" passes and this positive academic experience comes to a close. There is no way to truly express their contributions, but I would like to thank some of the wonderful people who contributed so much to my life and success in these past years.

This dissertation would not have come to fruition without the support of my advisor, Dr. Lakshmish Ramaswamy. I am grateful for his guidance, knowledge, and encouragement over the last few years. He helped me gain independence of thought and decision making by having confidence and trust in my ideas and abilities. I would also like to thank my kind and encouraging committee members: Dr. Khaled Rasheed provided insightful advice and was always willing to offer help when I needed it. Dr. Hamid Arabnia, for always providing thoughtful feedback and pushing me towards scientific research excellence. I am especially thankful to him for being my mentor each and every day in my Ph.D. journey and for helping me to accomplish so much more than I imagined, both personally and professionally. I hope to be able to pay their generosity forward by teaching and inspiring students for years to come.

I owe a special thanks to my family. My husband, Mohammadhossein, who always believed in me and my ability to do anything I set out to accomplish and for his continual optimism. Thank you to my parents, for always encouraging me to explore my own path and never settle for anything less than my aspirations. They raised me to believe there are no limits; I thank them for everything they have given me in my life - all of the support and love. To my sister, Delnavaz, for loving her big sister and for helping me to feel confident and brave even from afar.

The last three years of my Ph.D. program were funded by the National Park Service's Southeast Coast Network. I am grateful to my colleagues there, especially to Ms. Paula Capece, my supervisor and true friend. She challenged me enough to spark deep learning and provided valuable insights from a new and important perspective. I feel fortunate to have had the opportunity to work with her and learn valuable skills from her. I wish to thank her for everything she has done for me, both professionally and personally.

Next, I would like to thank members of my research group in DIPS Lab: Nitin Saroha for his passion, commitment, and hard work. To Grant Cooksey, with who I have thoroughly enjoyed a short research collaboration. I am also thankful to Dr. Maier, Dr. Gay, Dr. Buman and Anzah Niazi for their feedback and valuable contributions. I am grateful to all the faculty and staff members of the Computer Science department at UGA. I will forever be thankful for all the opportunities that I was given to learn and to grow.

Last, but definitely not least, I want to thank all of my friends who have helped me along the way and made graduate school a much more enjoyable place. Thank you to Neda Abolhassani, Faranak J.Fard, Maryam Pirmoradi (thanks for all the rides!), Mona Ahmadiani, Pourya Valizadeh, Sahar Voghoei, Navid Hashemi, Amna Basharat, Nasrin Rouhani and Saeedeh Zia. I would also like to specially thank Dr. and Mrs. Book, for being my second family and strong support during the past few years while I was far from my family. All of you have been fantastic friends and provided immeasurable academic and moral support over the years, and I am glad to have gotten to build a true friendship with you all.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it.
>
> — Mark Weiser [1999]

1.1 Sensor based Human Activity Recognition

Human Activity Recognition (HAR) has been a topic of interest within the ubiquitous computing research community for several years. With the development of low-priced wearable sensors, HAR is getting attention both in the research realm and in industries due to its applications in health care, cognitive assistance, indoor localization and tracking, and human-computer interaction. Many of these domains demand a robust HAR system that can distinguish ambulatory and non-ambulatory human activities with high accuracy.

HAR systems can be categorized into two broad classes: namely, those that are based on wearable sensors and those that use external devices such as video recorders [Turaga et al., 2008a], cameras and pressure sensors [Orr and Abowd, 2000]. The later has a number of disadvantages including expensive hardware, computationally intensive image and video analytics, limited monitoring and limited portability. On the other hand, wearable sensor-based HAR systems are largely devoid of the above problems. Furthermore, recent years have witnessed a tremendous proliferation of smartphones and fitness bands that have various energy-efficient sensors such as accelerometers. Driven by these trends, wearable sensor-based HAR systems have gained wide popularity and acceptability.

Although accelerometers have become fairly accurate over the years, robust and accurate identification of activities still poses a number of significant challenges. First, because many different brands of accelerometers are available on the market, HAR algorithms have to be resilient against device heterogeneities. Second, human activities are highly personalized in the sense that the salient characteristics of the activities vary from person to person. Many factors such as age, height, body mass index (BMI), injury histories and deformities may impact the manner in which a person walks, runs, or performs other activities. HAR algorithms have to be robust towards these factors. Third, accelerometer data is often noisy (for example, sensors might be shaken, moved, dropped during activities), and hence HAR algorithms have to be robust against noise. There are also challenges associated with online HAR systems where real-time sensor data processing and classification are required. In classifications systems the dimensionality of the feature space has a direct relation with the computational intensiveness of the systems. However, in online HAR systems the objective is to minimize computations and it is therefore important to use a minimum number of features. This may result in less accurate classification of every day living activities. There had been limited amount of research on online HAR classification algorithms that recognize activities in real-time, and most are based on a limited set of activities [Shoaib et al., 2015].

In this dissertation, we explore an offline as well as online approach to detect a set of human daily activities. Although several researchers have explored using machine learning algorithms for HAR, very few, to our best knowledge, have comprehensively studied the various facets of HAR [Bao and Intille, 2004; Lockhart and Weiss, 2014]. In many of these studies, the datasets are small and homogenous (for example, consisting of subjects in the same age group such as college students). Towards addressing these limitations, we have comprehensively studied the effectiveness of machine learning-based techniques for offline HAR systems.

Furthermore, most current research focuses on extracting complex features to achieve high classification accuracy. We apply shape-based time series classification such that complex feature extraction is avoided. We use annotated temporal data collected from a single tri-axial accelerometer sensor worn on the hip.

In this chapter we provide an overview of the dissertation by laying out the research context and describe the motivation behind this research. The research novelty and contributions are highlighted.

1.2 Research Context and Motivation

One of the major challenges in human activity recognition is the usability of such systems in real life applications. Unobtrusive wearable devices with limited number of sensors are more user-friendly, however, with fewer sensors attached to the body, the accuracy of such systems may decrease. In this thesis, we propose approaches for an effective human activity recognition system based on a single accelerometer sensor. Our proposed system could be embedded on any wearable device, including smart mobile phones and smartwatches. Many HAR systems are only capable of detecting a predefined limited set of simple activities, such as walking, running and laying down. Such systems are limited to applications in controlled environments. We try to address these limitations by having the system to train itself on a wide set of activities. To improve the accuracy of the system, we personalize the training phase. Each user will have a training phase and the system will adapt its learning based on the movements of a particular user. This would require that each user have a personalized copy of the application on their personal wearable device.

HAR systems are effective if they can recognize and detect human activities in a reasonable time. Many HAR systems are based on employing machine learning techniques on motion data on powerful servers. This requires the system to record data and send it to a server to be analyzed and detect the users' activities, thus cannot be achieved in real-time. Our method proposes that while the training phase can be done offline on a server, but the application recognizes users activities in real time. Such a feature adds many benefits to HAR systems and applications, one of which being in healthcare. Online HAR systems enable us to continuously monitor patients with physical or mental difficulties for their safety and recovery as discussed by Lara and Labrador [2013].

In addition, there have been very few comprehensive studies on various facets of HAR [Bao and Intille, 2004; Lockhart and Weiss, 2014] with large heterogeneous datasets. We study the effectiveness of machine learning-based techniques for offline HAR systems using datasets that consists data from users from a wide age range and body features. Our research is motivated by the demand of fulfilling the weaknesses of current HAR systems on topics mentioned above in order to develop more accurate and effective HAR system.

1.3 Research Objectives and Contributions

The main objective of this thesis is to perform human daily activity detection using a single wearable accelerometer sensor in an unobtrusive manner. From this general objective the following particular objectives are derived with respect to offline and online Human Activity Recognition systems:

- Comprehensive study on effective features to classify activities. We analyze the effectiveness of a combination of time domain and frequency domain features. We also propose an expert-selected feature set which improves the system's accuracy.

- Studying the effect of reducing the dataset size of the accelerometer readings. We show that decreasing the sampling rate down to 20% does not result in significant degradation of classifier accuracies.

- In achieving a HAR system with high accuracy we limit the training and testing sets to specific age groups. We show that this leads to a significant increase in the accuracy of activity recognition.

- We generate a personalized shapelet library database driven from users activity time series for a shapelet-based online implementation of human activity classification. This database is small in terms of size and can be stored on mobile phones and wearable devices.

- We propose a procedure to find the best shapelet which represents an activity class based on time series distance metrics and DTW. For demonstration we use real human activity data and show our system is independent of the sensor device.

1.4 Dissertation Organization

The rest of the dissertation is divided into the following chapters. Chapter 2 presents an analysis of relevant background material. Chapter 3 focuses on offline human activity recognition systems. Offline activity recognition systems have high accuracy in detecting human activities. However, their application may be limited compared to online HAR systems. In chapter 4 we show an online HAR system that can determine users activities in real time. Chapter 5 is dedicated to related work in the field of activity recognition and its application in the medical domain. The dissertation concludes in chapter 6 with reflection on the implications of our work encompassing this dissertation. The chapter ends with future directions for this line of research.

Chapter 2

Background

2.1 Introduction

In the past decade, Human Activity Recognition (HAR) has attracted interest in academia and industry due to its potentials in human-centric applications and usefulness for context-aware computing. Mainly, the purpose of human activity recognition systems are to recognize user movements and behaviors from low-level data gathered from sensors (usually wearable sensors or mobile phone sensors). HAR systems can be exploited to great benefits especially in health and medical domains, as an example in smart home environments for aged care monitoring [Benmansour et al., 2015], based on the information provided by cameras and other pervasive sensors, the system would monitor the occupant and determine when they need assistance, raising an alarm if required. HAR also allows for a continuous evaluation of users physical and cognitive capabilities by monitoring the performed activities of daily living. Previously, most of the work in HAR has been done using computer vision [Efros et al., 2003]. However, privacy and ethics are very important therefore the camera based solution would not be suitable. Recently, mobile phones and wearable sensors have been used for this purpose, these devices range from accelerometers to magnetic field sensors and are unobtrusive and suitable for users to wear throughout the day.

Various approaches are used to acquire useful information and knowledge from such sensors and one of the key components of any HAR system are Machine Learning techniques. To have a system that could automatically infer what activity is being performed using the accelerometers and other wearable sensors, it must have a detailed model of the activity [Guan et al., 2007]. Currently a variety of machine learning methods have been proposed for human activity recognition applications, namely as neural networks, Bayesian networks, hidden Markov models, K-nearest neighbors, decision tree and more. Most of the machine learning approaches used in HAR systems are supervised and need labeled activity samples for training purposes. However, in real activity recognition systems, labeled samples usually require human subjects to annotate the activity data and this is an expensive and time consuming process, therefor, recently there has been some attention drawn towards semi-supervised applications which is beyond the scope of this dissertation. In this chapter we will focus on machine learning techniques used in offline and online HAR systems. We try to show that there is no preferred machine learning technique used for all activity recognition system types and the selection of the algorithm and approach depends on the HAR application, sensors used and system design. We review the most commonly machine learning methods used in activity recognition and show a few of many of the HAR applications that have been developed in the recent years. We mainly focus on applications that use wearable sensors or embedded sensors on mobile phones.

2.2 Wearable Sensors

Wearable sensors have the key advantage to record human motions and movements regardless of the users' location. Such sensors are small sized and can be attached to different body parts such as the hip, wrists, chest, ankle and head. In the recent years sensors have been built into every day garment [1], making HAR systems less obtrusive and enabling the user to record skin temperature, heart rate, heat flux, conductivity, GPS location and body motion data without carrying a sensor object with them.

Some challenges with regards to HAR systems based on wearable sensors are preserving battery life, minimizing obtrusiveness and privacy. We will discuss each issue in detail in Chapter 5 along with approaches to overcome the challenges. In the following section we will describe the key features for accelerometer wearable sensors that are used in most HAR systems.

2.2.1 Accelerometer

Accelerometer sensor is a wearable device that measures the physical acceleration of the user. Accelerometer sensors have been used in many applications in science, medicine, engineering and industry. These sensors are particularly effective in recognizing ambulation activities such as walking, running, climbing stairs, lying and etc. They are very

[1]Hexoskin Smart Shirts, https://www.hexoskin.com/

inexpensive and require low power, therefore they have been embedded in smart devices such as mobile phones and smart watches in the recent couple of years.

Traditionally three accelerometer sensors were used to measure the acceleration magnitude and direction as a vector by placing sensors orthogonally in the three spatial dimensions namely as x, y and z. Nowadays wearable accelerometer sensors are a single chip called triaxial accelerometers which can measure acceleration along the three principle axes. In modern smart devices triaxial accelerometer sensors are used.

The main issue in HAR systems based on a single accelerometer sensor is that they get confused for non-ambulatory activities, such as brushing teeth, talking on the phone, eating, working at the desk and etc. Since such non-ambulatory activities have similar acceleration patterns, accuracy rates may decrease noticeably when a single accelerometer sensor is used. It would be helpful to include data from other sensors to detect such non-ambulatory activities with higher accuracy.

2.3 Data Preprocessing

Feature construction and extraction is the key action in preprocessing the data and it is important not to lose any information in this process. To improve the accuracy of the Machine Learning algorithms, statistical calculations are performed on raw accelerometer data before using the data for training the classifier. In many HAR systems preprocessing transformations include normalization, scaling and statistical analysis. We will briefly review some of these methods in the following sections. It is worth mentioning that choosing appropriate features of accelerometer data is not always clear cut and requires experience and trial-and-error. A good choice of features can significantly increase classifier performance and on the other hand naive and un-wise selection of features could lead to results with high error rates. Because activities are preformed over a long period of time (in seconds or minutes) compared to their sampling rate (which is 100 Hz in our experiments), analyzing a single sample point would be meaningless, therefore an important initial step in feature extraction is *Windowing*.

2.3.1 Feature Extraction

Some authors extract features based on mean, standard deviation, median, dynamic time warping, mean between axis, energy, characteristic frequencies, Pearson correlation, magnitude and angular degree. In the following sections we discuss the different feature domains. Features can be extracted from the raw signal data or from the signal processing features on the raw accelerometer signals.

Time Domain Features

In analyzing time domain features we study the activity accelerometer signals with respect to time. Time domain features are extracted from each axis acceleration signal. Basic statistical and inter-relationship metrics among the data points in the time domain are some of the most commonly used features.

In addition to the raw accelerometer data, Vector Magnitude $\|v\|$ is calculated where $d_x(i)$, $d_y(i)$, and $d_z(i)$ are the i^{th} acceleration sample of the x, y, and z axis in each window respectively.

Table 2.1: Summary of Feature Extraction methods for accelerometer signals as presented by Lara and Labrador [2013].

Group	Methods
Time domain	Mean, standard deviation, variance, interquartile range, mean absolute deviation (MAD), correlation between axes, entropy, and kurtosis.
Frequency domain	Fourier Transform (FT) and Discrete Cosine Transform (DCT).
Others	Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Autoregresive Model (AR), and HAAR filters.

$$\|v(i)\| = \sqrt{d_x(i)^2 + d_y(i)^2 + d_z(i)^2}$$

We will discuss the various features extracted from the time domain in Chapter 3.

Frequency Domain Features

Frequency domain refers to the analysis of accelerometer signals with respect to frequency, rather than time [Broughton and Bryan, 2011]. The Frequency domain of the accelerometer signals reveals useful information for discriminating various activities. This information that cannot be easily extracted from the time domain. Due to the high sensitivity of accelerometer devices, noise (which is high frequency signals) cause the 3D accelerometer signals to jump from high frequencies to low. Analyzing the signals from a frequency domain perspective, would allow filtering such noise by omitting the high frequencies, resulting in cleaner signals and therefore higher accuracy for activity recognition.

2.3.2 Window Length

As mentioned in the previous sections, Segmentation or Windowing is the process of grouping streaming time series raw sensor data into manageable chunks that contain clueful information describing activities [Bashir et al., 2016]. The selection of the window length depends on the sampling rate of the device, the activities being performed and the computational complexity of the method. Although having short windows may enhance the feature extraction performance, but would cause higher computation overhead due to the recognition algorithm being triggered more frequently. Furthermore, short windows may not provide sufficient information to fully describe the characteristics of the pattern of the performed activities. On the other hand, if the window length is too long, there may be more than a single activity occurring within a single window, causing difficulties in the learning phase.

There are three common approaches for windowing [Bashir et al., 2016]; first approach is to collect chunks of data defined within every fixed time period. The second approach is to use overlapping windows where the previous window overlaps with the current window by a fixed percentage. Overlapping time windows are intended to handle transitions more accurately. Lara and Labrador [2013] propose that using small non-overlapping windows, misclassifications due to transitions are negligible. The third approach does not employ time but collects a predefined amount of data points from the streaming sensor data. It is worth mentioning that different types of sensor data may require a different approach for finding the best window lengths. As an example heart rate signal requires at least 30s time windows Tapia et al. [2007]. Controversially, some other activities such as non-ambulatory activities may need shorter windows as 1.5s [Lara and

Labrador, 2013].

2.3.3 Feature Selection

The purpose of feature selection is to select which features of a large feature set to include in the final feature subset. If we have n number of features in the initial set, we may have up to 2^n possible subsets, and trying each and every one of the subsets to find the best one is not a rational and feasible option. Many approaches exist to search the feature subset space and find optimal features. In this section we describes some of the most common feature selection methods as well as the methods we have used in our approach.

Relief-based

[Kira and Rendell, 1992]: This feature selection method is essentially based on how well the features values can differentiate similar data points in different classes. Generally it works by randomly selecting an instance, and then finding the nearest instance of the same class and of the opposite class. The attributes are then weighted based on how well their values can distinguish the sampled instance from the nearest hit and nearest miss, it is likely to receive higher weight if it can differentiate between instances from different classes and it would be updated for each data point as needed. In this method, features that carry more information about the target class will typically have closer value to their nearest neighbor of the same class, and be further away from instances of the other class. The major drawback of Relief-based feature selection is that it does not take feature dependencies into account and therefore cannot discard redundant features if they exist.

Correlation-based

Unlike Relief-based, Correlation Feature Selection (CFS) ranks feature subsets rather than individual features [Hall, 2000]. CFS evaluates the relevance of features based on the correlation heuristic and scores each of the feature subsets. It takes the inter-correlation among features into consideration as well as their ability to differentiate classes. Subsets of features that are highly correlated within the class while having low inter-correlation are preferred in this method. CFS selects attributes that are highly correlated within the class and uncorrelated with each other. If two features are perfectly correlated, only one should make it into the final selected subset.

PCA-based

Principle Component Analysis (PCA), is a well known technique used for de-correlation and dimensionality reduction of data. PCA is a basic form of feature learning in its nature since it discovers meaningful representations of raw data without the need of relying on the domain knowledge. PCA reduces the original data to lower dimensional feature vectors by constructing a linear combination of the variables. Jolliffe [2002] note that since PCA does not consider class labels of the data, an accurate class separation in the direction of the high variance principal components is not guaranteed.

ICA-based

Independent Component Analysis (ICA), is another commonly used method for creating spatially-localized features. Unlike PCA which generates linearly combined features, in ICA basis vectors that are statistically independent are generated [Comon, 1994]. The algorithm is based on minimizing mutual information between the variables for judging independence between them.

Deep Learning

As an alternative to the mentioned methods for feature selection are deep learning methods that are based on a feed-forward artificial neural network consisting of an input layer, an output layer and a number of hidden layers. The innermost layer of the network has lower dimensionality, therefore there is a bottleneck for transmitting a signal through the innermost layer. This can be solved by a meaningful encoding of the input. This non-linear low-dimensional encoding is hence an automatically learned feature representation as described by Plotz et al. [2011].

2.4 Learning & Activity Recognition

In recent years, the development of sensing devices (e.g., accelerometers, cameras, GPS, etc.) has enabled the ability to collect attributes related to individuals movements. These applications require additional challenges of knowledge discovery since raw accelerometer data are not useful. Various Machine Learning tools are used in Activity Recognition Systems to analyze, and predict data [Lara and Labrador, 2013]. In HAR machine learning classifiers each instance is a feature vector extracted from signals within a time window as described in the previous section. Activity Recognition can be viewed as a classification problem such that each class corresponds to an activity. Generally the instances in the training set are labeled by the application, user or an external source, however in some cases, labeling data is not feasible since it requires an expert to manually label the examples and assign a label based upon their decision. In this section, we provide a brief overview on the the different classification algorithms that are commonly used in for classifying different activities in HAR systems.

2.4.1 Decision Tree

Decision Tree classifiers are based on predictive models that determine the class of a new sample from attributes of the data. These attributes values are denoted by the branches of the tree. The classes are represented by the terminal leaves. Many HAR systems use decision trees because their models are easy to read. In this model a new decision tree is first constructed based on the attributes that discriminates samples in the training data. Decision trees can be evaluated in $O(\log n)$ for n attributes. The advantages of decision trees are that they are simple and fast however, need to hold considerable data in each of their non-terminal leaves, which would increase the memory space required and increase computation time.

2.4.2 Random Forests

Random Forest is essentially an ensemble model for decision trees and it corrects the over fitting behavior of decision trees. They are also considered as form of a nearest neighbor predictor, that construct a number of decision trees at training time and outputs the mode of the classes as the final output class [Chetty et al., 2015]. Random Forests reduce the existing bias and variance in decision trees by computing an average, and balancing the two extremes. Moreover, Random Forests have a few number of parameters to tune. Due to all the mentioned advantages, Random Forests can be used without much adjusting compared to other classifiers and yield a reasonable model that is fast and efficient to use.

2.4.3 K-Nearest Neighbor

Similar to Random Forest, k-NN is also an instance-based learner, and is the basis of many lazy learning algorithms. K-NN tends to be very fast in training because they simply store the entire training set and postpone inductive generalization to classification time [Wettschereck et al., 1997]. An instance would be classified using a majority vote method of its neighbors, then the instance is assigned to the class most common among its k nearest neighbors.

2.4.4 Baysian Network

This classifier is based on Bayes theorem which uses probabilities in order to perform Bayesian inferences. The simplest Bayesian method is Naive Bayes, which is based on supervised learning and is straightforward to train the model. Naive Bayes performs well in terms of accuracy [Chetty et al., 2015]. Although Naive Bayes are known to be good models for comparison, however it may not be the best choice for HAR systems since it assumes that for any class value all features are independent, but because accelerometer signal values are heavily correlated, classifiers that work in this manner would not result in promising results [Lara and Labrador, 2013].

2.4.5 Support Vector Machine

SVM has become one of the popular classification methods in Machine Learning field especially for activity recognition. SVM is a two-class classifier however, can be extended to multi-class problems by combining multiple binary SVM classifiers. A multi-class classification problem can be solved by dividing of the problem into several two-class problems. Most HAR systems based on SVM use One-versus-One Strategy (OVO), where a set of binary classifiers vote on the class, and the class with the most number of votes will be the output [Rifkin and Klautau, 2004].

2.4.6 Artificial Neural Networks

Artificial neural networks resemble the brain neural network. Multi-Layer Perceptrons, which are a class of feed-forward Artificial Neural Networks [Roy et al., 2005], are commonly used in HAR classifiers and have proven to produce relatively high accuracy rates due to their learning capabilities. A neural classifier consists of an input layer for all of the signal features as discussed in the previous section, connect them to hidden layers with the activation function. By comparing the position of the maximum value in output vector and label vector we can determine the class of the activity.

2.4.7 K-means Clustering

Clustering is an unsupervised learning approach where the training set does not need to be labeled as in the supervised learning methods. If the instances are related to each other they are placed in a group and those who are not related would be placed in a different group. K-Means is the simplest algorithm which works without having previous knowledge based on a distance metric (Euclidean distance or Manhattan distance) to analyze if the instances can be grouped together. Due to its simplicity in functionality, and its capability to work with unlabeled data, it is a good method for examining classifier performance.

2.4.8 Lazy IBk Classifier

IBk is a Lazy Learner based on the principle of learning during classification time. They store the training instances during training time. IBk classifier is very similar to k-nearest neighbor classifier. As most of the learning happens during classification phase, they tend to be slow, therefor a variety of different search algorithms are used to find the nearest neighbors .

Hidden Markov Models

HMM are based on discrete state variables which are linked using a state transition matrix. When classifying time-based sequences of human activities, observing signals generated from complex or unfamiliar activities can be used to build a model of the activity indirectly. San-Segundo et al. [2016] state in their work that HMMs are an effective technique for activity classification, because they offer dynamic time wrapping, have clear Bayesian semantics and are well-understood training algorithms.

2.4.9 Ensemble Models

Ensemble learning approach is based on the assumption of improving performance by combining the output of several classifiers. Bagging, boosting, and stacking are some examples of this method. By using the strengths of several individual classifiers it applies a combination rule for the final decision of the classifiers. As an example, minimum probability, maximum probability, majority voting, product of probabilities, and average of probabilities are different examples for the mentioned combination rules. Classifier ensembles are computationally expensive compared to single classifiers, as they require several models to be trained and evaluated [Catal et al., 2015].

2.5 Existing HAR Approaches & their Limitations

In this section we will present some of the state-of-art HAR applications that have been recently developed. We only consider studies that use accelerometer data gathered by the mobile phone or another wearable device. It is important to note that there are a variety of HAR applications that use multiple sensors however, we only mention those parts of these studies that fit the scope of this dissertation. We discuss these studies in the context of the following aspects: Online HAR classifiers and Offline HAR Classifiers. Online HAR systems can classify activities on the mobile phones in real time and Offline systems are trained beforehand, usually on a server. We see that most studies have used the offline method. One reason that offline systems are more popular could be because the training process is computationally expensive. Moreover, it is easy to implement only the classification part on the mobile phone. There is only a limited number of studies where classifiers can be trained on mobile phones in real time.

2.5.1 Online

The goal of this section is to see the potential of mobile phones and wearable devices in running activity recognition systems locally. In healthcare, continuously monitoring patients is an important factor for their protection, safety, and recovery [Shoaib et al., 2015]. In the rest of this section we will describe some of the online state-of-the-art activity recognition approaches.

Energy Efficient SVM-based HAR

Anguita et al. [2013] propose an energy efficient approach for human activity recognition using mobile phones as wearable sensing devices. In their method they use a fixed-point arithmetic for human activity recognition of instead of the conventionally used floating-point arithmetic algorithms. They claim there method is capable of preserving the phone's battery while maintaining high accuracy levels. Their dataset consists of a total of 30 people with age range of 19 - 48 years who performed a set of activities including: standing, sitting, laying, walking, walking upstairs and walking downstairs. Features extracted from raw data were Signal Magnitude Area (SMA), mean, standard deviation (STD), entropy, signal-pair correlation (Corr) and Fast Fourier Transform (FFT) was used to find the frequency components for each of the windows.

MARS

Mobile Activity Recognition System (MARS) [Gomes et al., 2012] is another HAR application that its model is built and continuously updated on the mobile device using data stream mining. The advantages of this model is that it is personalized and thus increases privacy as the data is not sent to any external site. Furthermore, training/updating the model takes less than 30 seconds per activity and quickly adapts to user profile changes while being scalable and efficient in terms of the devices resource consumption. Naive Bayes classifier is used because it provides a simple and incremental learning approach.

actiServe

Berchtold et al. [2010] proposed an Activity Recognition service for mobile phones called ActiServ. They use a fuzzy system to classify human movement based on accelerometer signals gathered by the phone. The systems accuracy is in the range of 71% and 97%. However, in order to obtain the top accuracy level, the system requires a very long runtime duration. When the algorithms are executed to meet a real-time response time, the accuracy drops to 71%. On the other hand accuracy increases up to 90% after subject-dependent analysis and personalization.

2.5.2 Offline

Besides HAR Online learners, in Offline learners the user does not need to receive immediate feedback. As an example we can refer to applications that analyze the users daily habits and exercise routines, or in other health applications which the physician can review the patient's movement and exercise habits. Another example of an offline HAR system is an application to discover commercial patterns for advertisement [Lara and Labrador, 2013]. For instance, if an individual exercises a lot, certain advertisement could be presented to them such as sport gear. In such cases data gathered from human movements can be analyzed on longer periods of time, like daily or weekly, to draw conclusions on the person's behavior. In the remainder of this section we will summarize some state-of-art work in offline human activity recognition based on body wearable sensors.

Orientation Independent

Mobile phones by their nature are not fixed wearable sensors, carrying location of phones is often affected by the carrier's gender and garment style. Therefore, the location of the device needs to be independent of the algorithm. Guiry et al. [2012] designed and implemented a mobility monitor algorithm across a range of Android-based smartphones based in a test set with 6 subjects. Activities in their data set consists of sitting, standing, cycling walking, jogging and running. Their results appear promising, with average accuracies of 88.8% produced by the real-time mobility monitor, and a custom personalized classifier. They deploy a method to existing fixed position based algorithms to make HAR systems work in an orientation independent manner.

Bao et al.

One of the highly cited works in this domain is Bao and Intille [2004] work. Their system can recognize 20 activities, including human daily activities such as scrubbing, vacuuming, watching TV, and working on the PC. They used the aid of the system's users to label their activities. The sensors they use for recording data was bi-axial accelerometers which were placed on the user's knee, ankle, arm, and hip. However, in their results they concluded that only two accelerometers that are on the wrist and hip were actually useful and by omitting the other sensors the recognition accuracy is not significantly diminished. They use decision tree classifiers on time and frequency domain features and achieve 84% overall accuracy.

Obtrusive System

In the work of Parkka et al. [2006], seven activities are considered namely as: lying, rowing, riding a bike, standing still, running, walking, and Nordic walking. A combination of accelerometers, vital sensors and environmental sensors were used in their HAR system, which made their system a relatively invasive approach. The sensors had to be attached to the users chest, wrist, finger, forehead, shoulder, upper back, and armpit, all the sensors were integrated into a package that the users would carry in a backpack. Time and frequency domain features were extracted from most of the sensor signals along with a speech recognizer that was applied to the audio signal captured from the microphone in the package. The classification methods used were decision tree and artificial neural network. A limitation of their system is the need for high processing power, and it also raises privacy concerns since the microphone constantly records audio.

2.5.3 Limitations of Existing Applications

One of the limitations in most systems discussed in the previous sections is the variability in sensor device characteristics. Sensor devices may have different sensitivities, particularly being more sensitive to motion, where another sensor may be less sensitive. This difference should be considered in designing HAR systems to implement a system that would be independent to variabilities in sensors. Also, another limitation in the systems mentioned in this chapter is sensor displacement. HAR systems that are based on mobile phone's accelerometer sensors are prone to changes in sensor orientation and location becuase the user may carry the device in different locations. HAR systems should be designed such that changes in sensor orientation or location would not effect the classification of the users activities. Another important challenge in the existing HAR systems is the tradeoff between accuracy and response time. Depending on the real-world application of the HAR system either response time, or accuracy would become crucial. By reviewing many of the HAR applications we observe that HAR systems that offer a real time classification often tend to be less accurate compared to offline HAR systems, which analyze and classify activities offline. Another limitation in current HAR systems is scalability. HAR systems that are heavily based on feature extraction from time and frequency domain will lack scalability. In offline HAR systems scalability can be addressed by having the system rely on less features or adding hardware. However scalability issues are pertinent with respect to online HAR systems as well. In such systems addressing scalability is more complicated compared to offline systems becuase adding hardware components is not an option. We will address scalability in Online HAR systems in chapter 4.

2.6 Chapter Summary

In this chapter we survey the state-of-the-art applications in human activity recognition based on wearable sensors and mobile phone. We consider studies that use supervised machine learning for activity classification either locally on mobile phones in real time or on a remote server on a periodic basis. Prepossessing techniques have an important role in the accuracy of the classifier, therefor we provide an overview of popular feature extraction and feature selection methods currently used. As we have stated in this chapter, there is no clear definition on what techniques would better suit HAR systems. The characteristics of the systems and devices used can be different in activity recognition applications and thus different approaches should be considered for different systems. We reviewed a number of Human Activity Recognition systems and compare them with regards to their response time, learning approaches, obtrusiveness, data segmentation, feature extraction and selection, recognition accuracy, and other important design issues.

Chapter 3

Offline Multi-Featured Approach for Human Activity Recognition

Chapter Overview [1]

In this chapter, we present an offline multi-featured approach for recognition of various everyday activities using a single tri-axial accelerometer under real-world conditions. Although several researchers have explored using machine learning algorithms for HAR, very few, to our best knowledge, have comprehensively studied the various facets of HAR [Bao and Intille, 2004; Lockhart and Weiss, 2014]. Furthermore, in many of these studies, the datasets are small and homogenous (for example, consisting of subjects in the same age group such as college students). Towards addressing these limitations, we aim to comprehensively study the effectiveness of machine learning-based techniques for HAR.

[1]This chapter partially appears as:
D. Yazdansepas et al., "A Multi-featured Approach for Wearable Sensor-Based Human Activity Recognition," 2016 IEEE International Conference on Healthcare Informatics (ICHI).

We extract features a from the combination of time and frequency domain. We adopt two feature selection methods, as well as an expert-defined feature selection method, to the datasets to extract the most effective features for discriminating different activities.

Six classifiers are used for evaluating recognition performance. As the primary focus, we show the effect of different feature sets on each of the classifiers. We further demonstrated in our results the impact of decreasing the size of the training set on the accuracy of the classifier. In addition we also analyze time and frequency domain features and their combination on the accuracy of the classifiers. As a secondary contribution we tested our system on age based training data to observe if the system can gain higher accuracy if the training mode is personalized for the subjects' age group. In conclusion, we show several findings from our empirical experiments and discuss outstanding challenges and propose open research directions.

3.1 Overview of Offline HAR system

In this study we perform a comprehensive study with six machine learning algorithms along with three different feature selection methods. Our experimental study considers HAR systems that rely on a single tri-axial accelerometer sensor, which can be embedded in the smart phone, fitness band or small wearable devices. Compared to similar work, we have a relatively large dataset with 77 participants of distributed age range (from 18 to 65 years old) performing 25 activities. We perform wavelet analysis on accelerometer signals, then extract a set of 176 features from both time and frequency domain. A combination of both domain features are used to achieve higher accuracy for classifying the activities. One question we answer in this chapter is whether we can increase the speed of the system, limit the communication overhead and preserve battery life by reducing the dataset size without a degradation in accuracy. Figure 3.1 illustrates our offline HAR system overview.

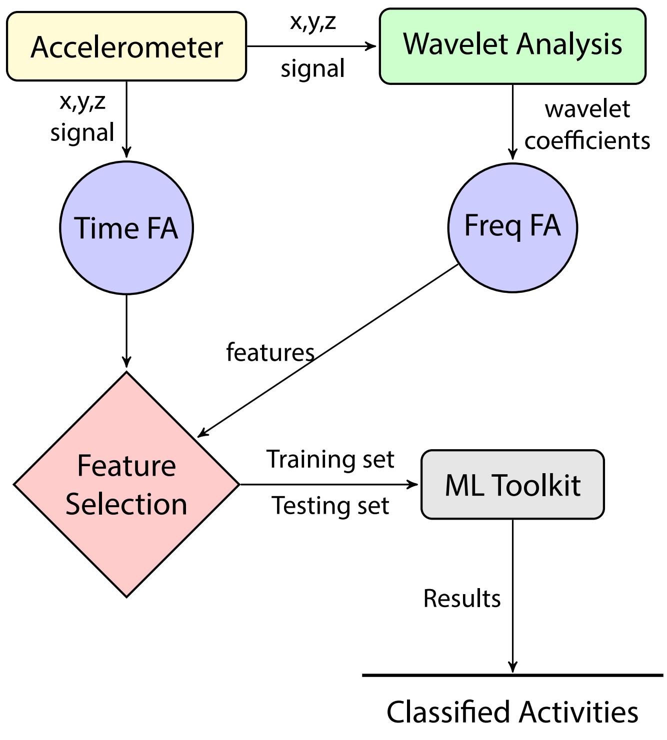


Figure 3.1: HAR system overview

Table 3.1: Summary of Activities

Activity Description	No of Subjects	Duration/Distance
Treadmill at 1 mph @ 0% grade	29	3 min
Treadmill at 2mph @ 0% grade	21	3 min
Treadmill at 3mph @ 0% grade	28	3 min
Treadmill at 3mph @ 5% grade	29	3 min
Treadmill at 4mph @ 0% grade	33	3 min
Treadmill at 5mph @ 0% grade	21	3 min
Treadmill at 6mph @ 0% grade	34	3 min
Treadmill at 6mph @ 5% grade	26	3 min
Seated, folding/stacking laundry	74	3 min
Standing/Fidgeting with hands while talking	77	3 min
1 minute brushing teeth + 1 minute brushing hair	77	2 min
Driving a car	21	-
Hard surface walking w/sneakers	76	400m
Hard surface walking w/sneakers hand in front pocket	33	100m
Hard surface walking w/sneakers while carry 8 lb. object	30	100m
Hard surface walking w/sneakers holding cell phone	24	100m
Hard surface walking w/sneakers holding filled coffee cup	26	100m
Carpet w High heels or dress shoes	70	100m
Grass barefoot	20	134m
Uneven dirt w/sneakers	23	107m
Up hill 5% grade w high heels or dress shoes	27	58.5m x 2 times
Down hill 5% grade w high heels or dress shoes	26	58.5m x 2 times
Walking up stairs (5 floors)	77	5 floors x 2 times
Walking down stairs (5 floors)	77	5 floors x 2 times

3.1.1 Data Set

Participants and Procedures

Participants were recruited from the Phoenix, AZ and surrounding areas through community sources, email distribution lists, and social media outlets. The participants were selected from a broad age range of 18-64 years old and were all free of any contraindications for exercise. Participants were fitted with a single hip-worn accelerometer and completed a series of activities for three minutes in duration (see Table 3.1). All participants completed the following activities: seated, folding/stacking laundry, standing/fidgeting with hands while talking, 1 min of brushing teeth and 1 min brushing hair, hard surface walking, carpet walking, and walking up and down stairs. An additional three treadmill activities and three other activities were randomly assigned. Time stamps for the beginning and end of activities were captured using a custom-built Android application which was synced to the same server as the activity monitor.

Activity monitor

Participants were fitted with the ActiGraph GT3X+ (ActiGraph, LLC, Pensacola, FL) activity monitor positioned along the anterior axillary line of the non-dominant hip. The monitor was fixed using an elastic belt. The ActiGraph GT3X+ is a lightweight monitor (4.6cm x 3.3cm x 1.5 cm, 19g) that measures tri-axial acceleration ranging from -6g to +6g. The device was initialized to sample at a rate of 100 Hz, then accelerometer data were sent to the server and extracted using Actilife 5.0 software (ActiGraph, LLC, Pensacola, FL).

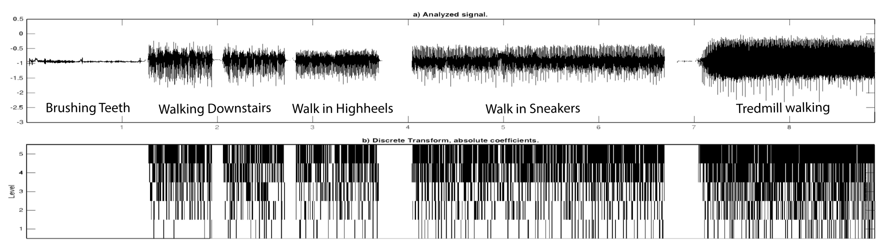


Figure 3.2: a) Human activity acceleration signal. The subject performed various activities with a pause between each of them. b)DWT absolute details coefficients on the five levels for the same activities.

3.1.2 Data Preprocessing

Feature construction is the key action in preprocessing the data and it is important not to lose any information in this process. To improve the accuracy of the Machine Learning algorithms, we perform statistical calculations on raw accelerometer data before using the data for training the classifier. In our study, pre-processing transformations include normalization, scaling, statistical analysis and discrete wavelet transforms. It is worth mentioning that choosing appropriate features of accelerometer data is not always clear cut and requires experience and trial-and-error. However, a good choice of features can significantly increase classifier performance. To achieve maximum performance we have used a combination of Time Domain features and Frequency Domain Features. Because activities are preformed over a long period of time (in seconds or minutes) compared to their sampling rate (which is 100 Hz in our experiments), analyzing a single sample point would be meaningless, therefore an important initial step in feature extraction is *Windowing*. We have set non-overlapping windows to be 2 seconds, yielding 200 data points in every window for each axis. In order to be able to effectively compare windows we need to perform *Feature Extraction*. Combining features extracted from the time and frequency domain we have a total of 176 features. In the following sections we discuss our different feature domains.

Time Domain Features

Time domain features are extracted from each axis acceleration signal, to process the data we look into some basic statistical and inter-relationship among the data points.

$$\|v(i)\| = \sqrt{d_x(i)^2 + d_y(i)^2 + d_z(i)^2}$$

In addition to the raw accelerometer data we calculate Vector Magnitude $\|v\|$ where $d_x(i)$, $d_y(i)$, and $d_z(i)$ are the i^{th} acceleration sample of the x, y, and z axis in each window respectively. The definition of the time domain features for a given window $W = \{d_1, d_2, ..d_n\}$ are listed below:

- $Max\{d_1, d_2, ..d_n\}$ and $Min\{d_1, d_2, ..d_n\}$
- *Arithmetic mean* ($\bar{w}$) for each accelerometer axis and $\|v\|$ (Equation 3.1)
- *Standard deviation* (std_W) for each axis and $\|v\|$ (Equation 3.2)
- *Median crossing rate* which is the number of times the signal changes from below the median to above the median or vice versa.
- The *10th, 25th, 50th* (Equation 3.3) , *75th, 90th percentile*
- Correlations between Accelerometer axes and Vector Magnitudes

$$m_W = \frac{1}{n}\sum_{i=1}^{n} W[i] \tag{3.1}$$

$$std_W = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(W[i] - m_W)^2} \tag{3.2}$$

$$50^{th} percentile = (\frac{n+1}{2})^{th} term \tag{3.3}$$

Frequency Domain Features

Frequency domain features are derived by transforming the raw accelerometer time series data into frequency domain by using Fast Fourier transform (FFT). The FFT coefficients obtained from the transform serve as the amplitudes of the frequency components of the signal and its energy distribution. A number of these coefficients can serve as features and other features such as energy can be derived from these coefficients [Bashir et al., 2016]. On the other hand wavelet analysis can show spectral information of non-stationary signals. Spectral analysis methods such as Fourier Transform (FT) can only provide the frequency content of a stationary signal [Bidargaddi et al., 2007]. In ambulatory activity signals, we observe that frequency changes rapidly over time. This highlights the importance of Wavelet Transforms, which can describe the intensity level of the signal's different base bands at a certain point of time. There has been a few number of Activity Recognition systems based on features extracted from wavelets.

He [2010] select Daubechies 3 wavelet by comparing the decomposition level required while keeping the energy as much as possible. Haar wavelet [Graps, 1995] filtering techniques are proposed as a low calculation cost feature extraction method suitable for 3D acceleration signals based human activity recognition as noted by Hanai et al. [2009]. Most Activity Recognition systems use Daubechies 3 wavelet because it resembles the pattern of human motion and therefore would provide more accurate wavelet features.

The usefulness of wavelet transform is in its trade off between time and frequency resolution. Components in the activity signal that have low frequency are not as easy to resolve compared to signals of higher frequencies, on the other hand high frequencies are difficult to resolve in the time domain. So by applying Wavelet transforms we can achieve better frequency resolution at low frequencies and better time resolution at high frequencies [Chau, 2001].

Figure 3.2 shows the original signal of a few daily human activities in the time domain, and the absolute coefficients of the discrete wavelet transform of the corresponding signal. After applying discrete wavelet transform on the raw accelerometer signal, the outputs are *Wavelet Coefficients* which describe the power of the signals at the given 5 levels. Below is a list of features derived from the wavelet coefficients and dominant frequency magnitude of the Fourier transform for a given window $W = \{d_1, ..., d_n\}$

- *Maximum* and *Minimum* of $\{d_1, d_2, ..d_n\}$
- *Arithmetic mean* $\bar{w}$ (Refer to equation 3.1)
- *Standard deviation* std_W (Refer to equation 3.2)
- *Median crossing rate*
- The *10th, 25th, 50th, 75th, 90th percentile* (Refer to equation 3.3)

Note that Wavelet transforms requires multiple components in order to be able to discriminate different human activities. Hence it will increase computation and may not be suitable for real time applications. Time domain features can be easily extracted in real time, they are more popular in many practical acceleration activity recognition systems.

3.1.3 Feature Selection

The purpose of feature selection is to decide which features of a large feature set to include in the final feature subset. If we have *n* number of features in the initial set, we can have

Table 3.2: List of features for each Feature Selection method.

Feature Selection	Time	Fourier	Wavelet
Expert-selection	X_{mean}, X_{stdev},$X_{50thPer}$,Y_{mean}, Y_{stdev},$Y_{50thPer}$, Z_{mean}, Z_{stdev}, $Z_{50thPer}$, $VecMag_{mean}$, $VecMag_{stdev}$, $VecMag_{50thPer}$	-	$XLev1_{mean}$ - $XLev5_{mean}$, $YLev1_{mean}$ - $YLev5_{mean}$
Correlation-based	X_{max}, X_{mean}, X_{stdev},$X_{50thPer}$,Y_{stdev}, $Y_{10thPer}$,$Y_{25thPer}$, $Y_{75thPer}$,Z_{mean}, $Z_{90thPer}$, $VecMag_{min}$, $VecMag_{max}$, $VecMag_{mean}$, $VecMag_{stdev}$, $VecMag_{10thPer}$,$VecMag_{25thPer}$, $VecMag_{50thPer}$, $VecMag_{75thPer}$, $VecMag_{90thPer}$	Xfr_{min}, Xfr_{stdev}, $Xfr_{25thPer}$, $Xfr_{50thPer}$,$Xfr_{90thPer}$, Yfr_{min} $Yfr_{50thPer}$,$Yfr_{75thPer}$,Zfr_{stdev} $Zfr_{25thPer}$,$Zfr_{50thPer}$,$Zfr_{90thPer}$	$XLev4_{max}$,$XLev5_{max}$, $XLev5_{50thPer}$, $XLev5_{75thPer}$, $YLev3_{stdev}$
Relief-based	$VecMag_{10thPer}$,$VecMag_{25thPer}$, $VecMag_{50thPer}$,XZ_{corr},YZ_{corr}, XY_{corr},Y_{stdev},X_{mean}, $VecMag_{stdev}$,Z_{mean}, $Z_{50thPer}$, X_{stdev}, $X_{90thPer}$,$VecMag_{75thPer}$, $X_{50thPer}$,$Z_{25thPer}$,$Z_{75thPer}$, $Y_{75thPer}$, $X_{75thPer}$,$VecMag_{90thPer}$	$XfrYfr_{corr}$, $XfrZfr_{corr}$, $YfrZfr_{corr}$	$XLev5_{max}$, $XLev5_{90thPer}$, $XLev5_{75thPer}$, $XLev4_{max}$, $YLev10_{max}$, $YLev1_{MedCross}$, $XLev1_{MedCross}$

up to 2^n possible subsets, and trying every one of them to find the best subset is not a rational option. Many approaches exist to search the feature subset space and find the optimal features. This section describes the feature selection methods we have used in our approach. A summary of the features selected in each Feature Selection method can be found in Table 3.2.

Relief-based [Kira and Rendell, 1992]: This feature selection method is based on how well the features values can differentiate similar data points in different classes. It works by randomly selecting an instance, and finding the nearest instance of the same class and of the opposite class. The attributes are assigned higher weights if they can differentiate between instances from different classes. Features that carry more information about the target class will have closer value to their nearest neighbor of the same class, and be further away from instances of the other class. The major drawback of Relief-based feature selection is that it does not take feature dependencies into account and therefore cannot discard redundant features.

Correlation-based: Unlike Relief-based, Correlation Feature Selection (CFS) ranks feature subsets rather than individual features [Hall, 2000]. CFS evaluates the inter-correlation among features along with their ability to differentiate classes. Highly correlated subsets with the class are preferred, while having low inter-correlation. If two features are perfectly correlated, then only one will make it into the final subset of selected features.

Expert-selection: Having run multiple feature selection methods and looking at the selected features in each case, we manually suggested a selection set of reasonable features on the dataset. By looking at the trends in the type of features selected by the majority of feature selection methods, we found the features that occur most frequently in the different feature subsets and construct a new subset from the top N that are most frequently chosen. This hand-tuning resulted in a reduced and standardized feature subset which also had improvements in generalization performance, pre-processing time and streamline the feature extraction processes. It was observed that many of the feature selection methods agreed on the following properties most prominently: Mean, Standard Deviation and the 50th Percentile. In addition to this, we observed that features of the Fourier transform features were lowly ranked by most methods and therefore were omitted in the Expert-selection feature set. Furthermore some of the wavelet coefficients were also discarded observing these trends. Our results show that applying the Expert-selection feature selection method reduces the pre-processing time significantly and has shown generally better results than the features sets extracted by the recommendation of the traditional feature selection methods.

3.2 Empirical Evaluation

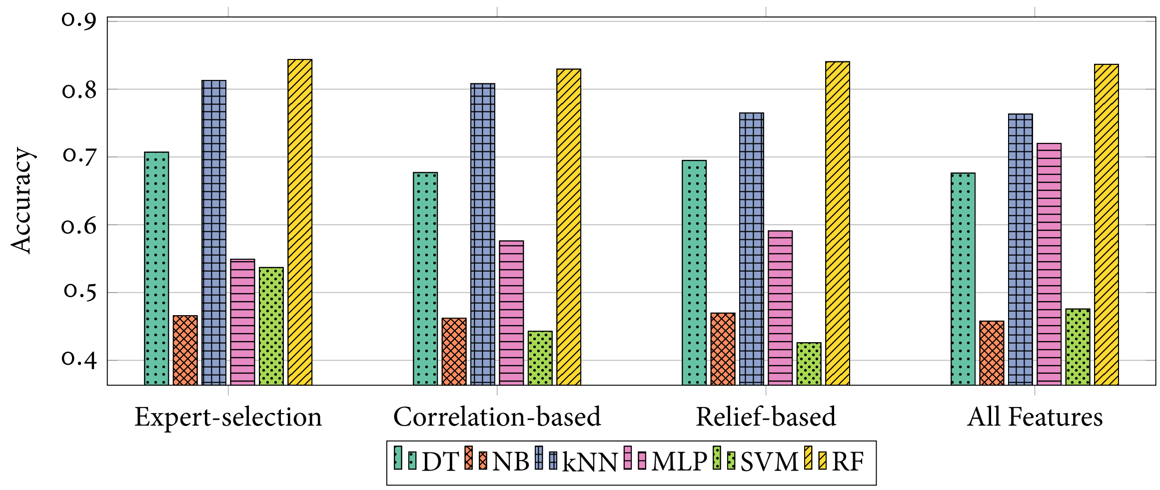


Figure 3.3: Comparison of the performance of different feature selection methods.

3.2.1 Classification Algorithms Comparison

Activity Recognition can be viewed as a classification problem in that each class corresponds to an activity. A selection of features were chosen from the combination of time domain and frequency domain as described in the previous sections. We tried six different machine learning algorithms: Random Forest (RF), K-Nearest Neighbor (k-NN), Decision Tree (DT), Multi-Layer Perceptron (MLP), Support Vector Machine (SVM), and Naive Bayes (NB) to classify the activities performed by subjects. All classifiers can be found in the Weka Machine Learning Algorithms toolkit [Hall et al.], an open source machine learning suite. Each classifier that was chosen has its own strengths and weaknesses in terms of activity recognition on accelerometer data. In order to determine the accuracy of each classifier the data set was divided into a training and testing set with ratio of 9 to 1 with no overlapping instances. In the first step of machine learning experiments, we compared the classification accuracy of the six machine learning algorithms using the features selected from the tri-axial accelerometer in time domain and frequency domain.

The results on the different feature selection methods are shown in Figure 3.3. As it can be seen, Random Forest ranked as the best classifier through all the experiments on different feature selection methods while having an average accuracy of 83.7%. Random Forest is essentially an ensemble model for decision trees and it corrects the over fitting behavior of decision trees, therefore we observe an improved accuracy rate compared to Decision Trees. The second best classification algorithm is k-Nearest Neighbor (k-NN). Similar to Random Forest, k-NN is also an instance-based learner, and is the basis of many lazy learning algorithms. K-NN tend to be fast in training because they simply store the entire training set and postpone all effort towards inductive generalization until classification time [Wettschereck et al., 1997]. Decision Tree ranked third among all feature selection methods. Multi-Layer Perceptrons, which are a class of feedforward Artificial Neural Networks, performs slightly better than random (50% accuracy rate). Naive Bayes and SVM have not provided any promising results in this set of experiments. We can claim that the poor performance of Naive Bayes is because it assumes that given any class value all features are independent and since accelerometer signal values are heavily correlated, classifiers that work in this manner would not result in promising

Table 3.3: Condensed Summary Matrix of Intra and Inter Activity Categories

Activity	$Treadmill_{4mph}$	$Treadmill_{othrs}$	Stand	$NonAmb_{othrs}$	$Walk_{hiheel}$	$Walk_{othrs}$	Downstairs	$Graded_{othrs}$
$Treadmill_{4mph}$	**0.84**	0.04	0.0	0.01	0.01	0.05	0.0	0.02
Stand	0.0	0.0	**0.91**	0.02	0.0	0.0	0.0	0.0
$Walk_{hiheel}$	0.1	0.06	0.02	0.01	**0.68**	0.14	0.02	0.03
Downstairs	0.0	0.04	0.01	0.02	0.0	0.05	**0.81**	0.03

results [Lara and Labrador, 2013].

By looking at the results in Figure 3.3, it may seem that the accuracy rates of the classifiers are lower than other similar HAR systems. However, it is important to reiterate that we perform classification on a large range of activities, many of which are in the same category (for example we have a variety of walking activities, in each one the subject is asked to hold a different object while walking, see Table 3.1). Naturally this yields a lower accuracy rate compared to HAR systems that classify activities in a lower granularity. The confusion matrix obtained with the Random Forest classifier and Expert-selected feature set is reported in Table 3.3. We limited the activities to be displayed in the confusion matrix because of space constraint. Note how some activities in the same category like walking in heels and other walking activities have confusions between them. From Table 3.3 it can be concluded that the RF classifier has accuracy above 82% while recognizing activities in different categories (for example treadmill walking and standing).

3.2.2 The Effects of Feature Selection

In section 3.1.3 we have introduced three different feature selection methods in order to extract the most dominating features. For the purpose of observing the effect of each feature selection method on the accuracy of the classifier, a dataset containing all 176 features is used in addition to the 3 feature selected datasets in our experiments. Figure 3.3 shows the accuracy of the classifiers over the different feature sets. As it can be easily observed in the graph, in Expert-selected feature set all the classifiers have better accuracies compared to Correlation-based, Relief-Based and All-Feature set. There is only one exception to this case, Multi-Layer Perceptrons perform better on the data set that includes all 176 features. In other words, feature selection did not have the same effect on recognition accuracy for Multi-Layer Perceptron as it did on the other classifiers. We believe the reason is that for very large datasets the neural network training algorithm is able to learn the relevance of individual features, and therefore no feature selection is necessary [Spence and Sajda, 1998]. This explains the phenomenon that in our study the accuracy of the MLP increases when we increase the number of features in the dataset we give into the network to train on. Correlation-based feature selection method has slightly lower accuracy rates for all the classifiers compared to Expert-selection feature set with an exception of SVM classifier. There is a significant difference of accuracy in SVM for Expert-selection and Correlation-based subsets, showing that the features in the Expert-selection subset have better effect on the accuracy of this classifier.

3.2.3 Impact of Training set Size on Accuracy

We have generated multiple datasets from the original dataset to observe the effect of dataset size on the accuracy of each classifier. As we have stated in the previous section, our Expert-selection feature dataset had the best accuracy for most of the classifiers among the others, therefore we have used this dataset and selected 50%, 25% and 10% of it's instances randomly. Although this is not the same as decreasing the sampling rate, however reducing the dataset size can result in a shortened response time to classify a new activity. In general, when the size of dataset decreases, the accuracy of the classification algorithms also decreases. However the degrading slope is different for each classifier. For Decision Tree, k-NN and Random Forest algorithm the accuracy inclines moderately when we shrink the dataset size to 20% of the original dataset, meanwhile some other algorithms such as Naive Bayes and MLP roughly maintain the same accuracy rate while the dataset is minimized to 20%. However we observe a sharp drop in the accuracy rates for the the majority of the algorithms once their dataset size is reduced to 10%.

Figure 3.4 illustrates the impact of the dataset size on the classification accuracy. All of the classifiers maintain the same ranking for the different dataset sizes. In our experimental studies, using four different dataset sizes, we observed that despite the fact that decreasing the dataset size would negatively impact the accuracy, but may be worth disregarding since the smaller datasets are more efficient in terms of space and computational time, thus the algorithms that are trained on smaller datasets can be implemented on devices with lower computational power such as mobile phone devices. This would also help preserve the battery power if the classification algorithms would be to run on

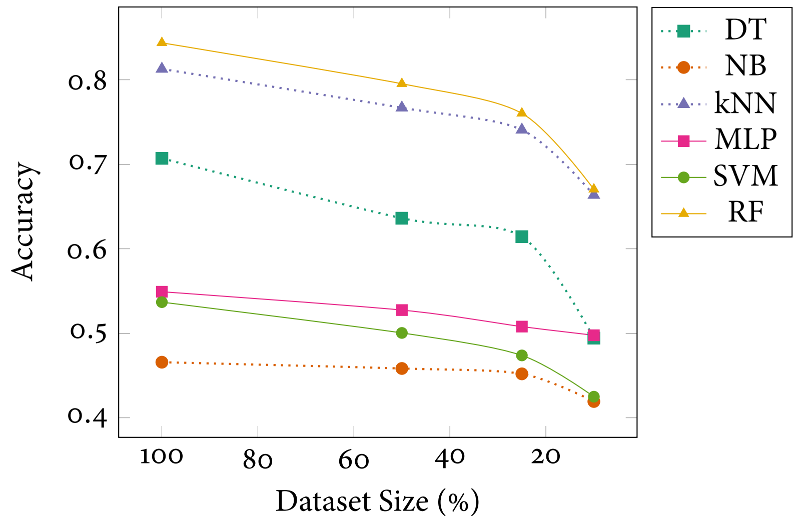


Figure 3.4: Accuracy of the six classifiers with varying dataset sizes

mobile application independent of an external server.

3.2.4 Effect of Combining Frequency and Time domain features

In section 3.1.2 we have claimed that we believe the Frequency Domain features add valuable information for recognizing human activity using accelerometer data. In order to show the impact of such features on the accuracy of the classifiers we have generated two separate datasets, one only consisting of time domain features, and the other one containing a combination of time and frequency domain features. Adding the frequency domain feature had different impacts on each classification algorithm. Figure 3.5 illustrates the accuracy of the classifiers in each dataset.

It is evident that for SVM classifier adding the frequency features had positive impact, increasing the accuracy 12% compared to the dataset where only the time domain features existed. MLP also had a 6% increase of accuracy rate when frequency features were added to time domain features. For the rest of the classifiers there was not a noticeable difference in the accuracy rates for the two datasets.

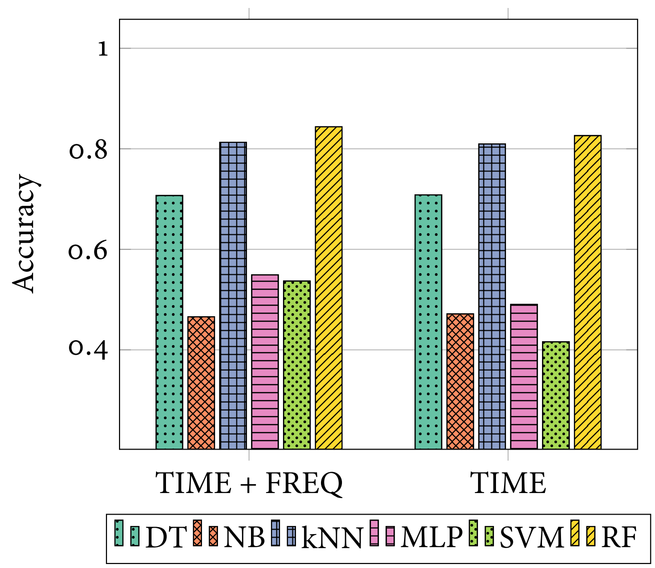


Figure 3.5: Classifier accuracy rates for time domain vs frequency domain.

3.2.5 Impact of Age-based Training

Human Activity Recognition systems that are trained on the data of other subjects can suffer from a loss of accuracy. To minimize the error rate we tried to group the subjects based on their ages, then train/test the classifiers on individual age group datasets. We have divided the dataset into four different subsets according to the subjects ages. Table 3.4 summarizes the age distribution across four age groups. We selected this age break down particularly for the following reasons: first, subjects in each age group have similar health and body characteristics representing their age group and second, based on the distribution of the subjects age, this break down was the most balanced in terms of number of subjects in each group.

We studied two approaches to compare the accuracy of the classifiers. First we used the Random Forest model we previously described to train the classifier on the individual age groups dataset and tested on the same age group using 5-fold cross validation. We compare these results to the accuracy of the general Random Forest Classifier, which was trained and tested on the entire dataset instead of specific age groups. However for the purpose of having comparison grounds, we down sampled the number of instances in the general dataset to have the same number of instances in the age grouped training sets.

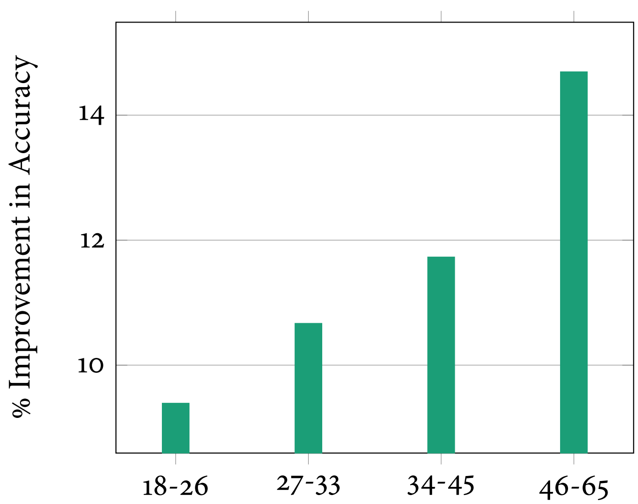


Figure 3.6: Accuracy improvement in age-specific training set over generic training set

As seen in Figure 3.6 for all of the age groups, training on a specific age group dataset results in an improved accuracy. We observe a significant increase in the accuracy of the system when we train the classifier on individual age groups. We see that the age group consisting of participants ages 45-65 has the highest accuracy improvement. It can be inferred that this age group have the most similarity in body movements within their group and are distinguishable from subjects in other age groups and therefore obtain higher accuracy rate. This suggests that we can design more accurate systems if we personalize the training phase based on the users age.

Table 3.4: Dataset Age Ranges Summary

Age Range	18-26	27-33	34-45	46-65
Number of subjects	17	27	24	9

3.3 Chapter Summary

In this study, we focus on the recognition of various everyday activities using a single triaxial accelerometer attached to the belt of the participants. The data was acquired from over 75 subjects under real-world conditions performing 25 activities. A total of 176 features were extracted from the combination of time and frequency domain. Two feature selection methods, as well as a self-defined feature selection method, were applied to the datasets to extract the most effective features for discriminating different activities. Six classifiers were used for evaluating recognition performance. We showed the effect of different feature sets on each of the classifiers. In order to improve the energy efficiency of the system, we further demonstrated in our results the impact of decreasing the size of the training set on the accuracy of the classifier. In addition, we tested our system on age based training data to observe if the system can gain higher accuracy if the training mode is personalized for the subjects' age group. Our preliminary results showed that training on specific age groups would be effective in increasing the accuracy of the activity recognition system, therefore training the classifiers based on the participants age would result in more accurate HAR models. We plan to extend this further and include demographic features in our datasets in our future work.

Chapter 4

Online Human Activity Recognition Using Time Series Shapelets

Chapter Overview

In this chapter, we explore an online HAR system to detect a set of daily activities using shape-based time series classification such that complex feature extraction is avoided. Mobile phones, activity trackers, and many other current mobile devices incorporate various sensors such as GPS, accelerometer and gyroscope. These sensors can be used to study and analyze human physical activities. In our study we use annotated temporal data collected from a single tri-axial accelerometer sensor worn on the hip. Such sensor data are noisy and therefore in preprocessing the data, we eliminate the noise and distortion. When comparing activity time series, we should consider the shape similarities within the time series regardless of the magnitude. Therefore, we consider applying normalization and standardization to make the time series independent of magnitude. Further in this chapter we propose a method for classifying activity time series based on shapelets. Shapelets enable us to compare time series in terms of their shape similarities, without the need of extracting features. Shapelets are small segments of time series that can capture important characteristics such as pattern of the time series. We believe shapelets are extremely useful for HAR systems because of the repetition nature of human motions and activities. Shapelets can easily capture and present activity patterns, and therefore represent a class of human activity. Because shapelets are small in terms of data points compared to time series, they are computationally effective for classifying time series which makes our proposed system fast and computationally efficient.

4.1 Challenges of Online HAR Systems

Most existing HAR systems are based on machine learning techniques that require many features to be extracted from the data. Feature extraction is computationally expensive and therefore systems that use this technique detect human activities need todo human activity classification offline [Bao and Intille, 2004]. By offline we mean that the recognition system would not detect the activity in real time, rather the process and computations will be performed on a server after the data has been collected. In offline HAR systems sensor data will be processed and analyzed as batches of data. In these systems the computation overhead of the classifiers is not as crucial as it is in online HAR systems. Online HAR systems give immediate feedback about the individual's activity, and they are usually embedded in a wearable device. Because these wearable devices are limited in data storage, computation capabilities and battery life, an important concern is to create online activity classification systems that would be fast and computationally efficient without requiring large amounts of data. We therefore study classifying the time series activities using shapelets in real time. We believe that time series shapelets would enable fast classification of data without the need to extract features at classification time. We will address the limitations of most online human activity classification systems by using time series shapelets and similarity metrics. In the remainder of this chapter we will discuss the different components of the proposed system.

4.2 Approach

Our method is based on extracting shapelets from activity time series where ideally each shapelet is a representation of each activity class. By using several distance metrics we compare the time series to the shapelets, each time series that exhibits similar patterns to the shapelet will be labeled in the same class as the shapelet. The ultimate purpose of our proposed method is to classify activities using raw data, therefore, there is no need to define features that separate different activities. Using raw data makes the method more general in terms of being applicable to other sets of activities.

The main component in our proposed method is extracting and selecting shapelets from activity time series. Ideally each shapelet is a representation of each activity class capturing the dominant pattern of the activity. We evaluate all sub-patterns in the time series and select those of which are representative of the activity. Finding the best of such patterns requires examining every single sub-pattern of a particular size. A few years ago, many authors of related work [Geurts, 2001] believed this task would be tedious to test and slow in execution, but we believe with the recent improvements in CPU time and parallelization techniques the proposed algorithm for discovering time series sub-patterns is achievable in a reasonable time. We then select the activity class representative shapelets by selecting the most representative shapelet for each activity which captures the dominant pattern in that activity time series. We then classify time series based on distance or similarity to the shapelet of each class. Cross correlation, Euclidean distance, Pearson correlation and Dynamic Time Warping (DTW) are a few examples of distance metrics popularly used for time series.

We generate a personalized shapelet library database driven from users activity time series for a shapelet-based online implementation of human activity classification. This database is small in terms of size and can be stored on mobile phones and wearable devices. To build this library in the training phase each individuals activity time series goes through the different components in our method defined below:

- **Average Peak Distance**. Given a time series, P is the average number of data points between two consecutive local peaks. We use a peak detection algorithm explained in section 4.3.2 to find the local upper and lower peaks. We denote the average distance of the upper peaks with P_{upper} and the lower peak distance with P_{lower}.

- **Shapelet Extraction**. We extract all shapelets of a certain length l_{sh} using a sliding window. The size of the sliding window is set based on the lower and upper peaks of the activity times series. We call the set of all shapelets of a time series the candidate shapelets. The total number of possible candidate shapelets for all the time series in the dataset are:

$$\sum_{TS_i \in D} (n_i - l_i + 1)$$

where n_i is the length of TS_i, D is the set of all time series and l is the length of the shapelets for the *i*th time series.

- **Distance Metric**. The distance between each time series and a shapelet is represented with *dist*. The distance function takes *TS* and *SH* as inputs and returns a total distance value. In this study we measure distance using Euclidean distance and DTW, which are discussed and evaluated in section 4.4.

- **Class Shapelet Representative**. Given a set of candidate shapelets we find the single best shapelet which is representative of the activity class from the set of candidate shapelets. There are several methods used to select the single best shapelet that would represent an activity class, we discuss this further in section4.3.4.

We propose a procedure to find the best shapelet which represents an activity class based on time series distance metrics and DTW. For demonstration we use real human activity data and show our system is independent of the sensor device.

4.2.1 Notation

In this section we introduce key terms used in this study. Table 4.1 summarizes the notations used in the this chapter. We expand on each definition below.

Definition 1: *Time Series*. A time series $TS = t_1, t_2, t_3, \ldots, t_n$ is an ordered set of n real-valued variables that are recordings of an accelerometer sensor measured in meters per

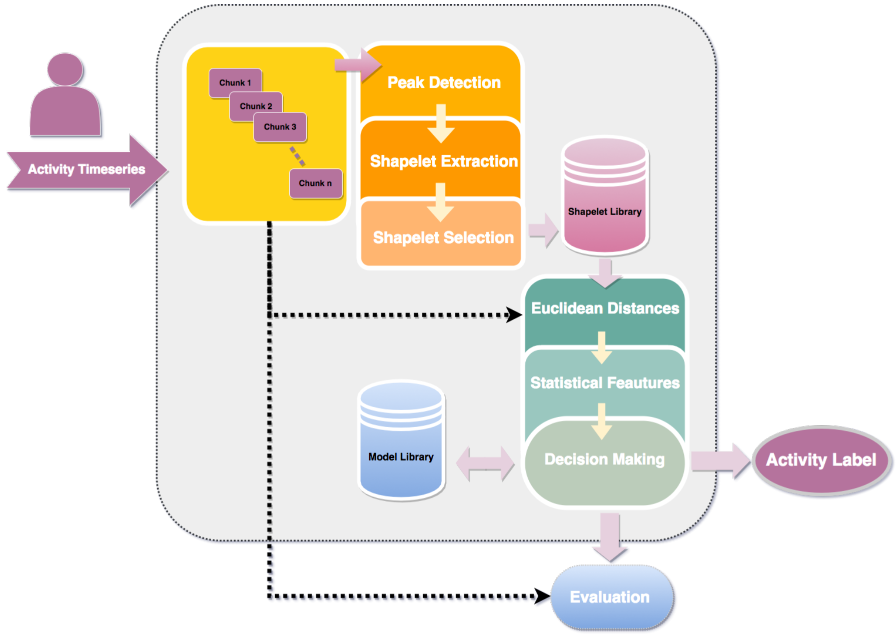


Figure 4.1: Online HAR System Overview

second squared (m/s^2). Each value t_k in the time series represents the rate of change of the velocity of the user wearing the sensor.

Definition 2: *Shapelet*. Given an activity time series TS of length n, a sub-pattern or a shapelet is a partition of length m where $m \leq n$. A shapelet is a continuous set of sensor readings that are spaced at the same rate of the time series. We present a shapelet as $SH = t_l, t_{l+1}, ..., t_{l+m-1}$ for $1 \leq l \leq n - m + 1$.

Definition 3: *Recording Rate*. Data points in activity time series are arranged in temporal order spaced at equal time intervals $1/r$.

Table 4.1: Summary of notations used in Chapter 4

Symbol	Explanation
TS	time series
TS_k	*k*th data point in the time series
SH	shapelet
SH_j	*j*th data point in the shapelet
$\|TS\|$	length of time series
l_{sh}	length of shapelet
r	sensor recording rate

4.2.2 Data Representation

Participants and Procedures

We use a subset of ambulatory activities performed by users in this study. Refer to section 3.1.1 for a description of participants and procedures of the collected dataset. We select activities that are similar in nature but are preformed in different environments to show that shapelets can capture the differences in activities that may appear similar. For instance, walking on a treadmill may seem to have a similar pattern of walking bare foot on grass. However, we show the shapelets that represent each class have different patterns.

Activity monitor device

Participants were fitted with a hip worn tri-axial accelerometer sensor. Refer to section 3.1.1 for description of the wearable device used in collecting the dataset.

4.3 Shapelet Extraction

4.3.1 Time Series Normalization

Before finding the candidate shapelets in the activity time series, we need to perform an important preprocessing step to reduce noise and highlight the pattern in the time series. Normalizing activity time series is necessary in order to compare time series with shapelets using any distance measure [Keogh and Kasetty, 2003]. Normalizing time series enables us to eliminate differences of overall magnitude of two time series. Therefore, with normalization we can correctly measure the true similarity of a shapelet and a segment of an activity time series that that may be similar in shape but have different offsets along the accelerometer axis.

There are various methods of normalization techniques, common ones are exponential smoothing, Holt's linear smoothing [Holt, 2004], single moving average, differencing, splines [Silverman, 1984], and LOESS technique [Cleveland, 1979]. In related literature standardization and normalization are sometimes used inter-changeably. However, there is a slight difference between the two mentioned methods. Standardizing a time series requires the data to have a mean of 0 and a standard deviation of 1. In normalization of time series we scale the data in a specific range which is usually between 0 and 1, but other ranges could also be used.

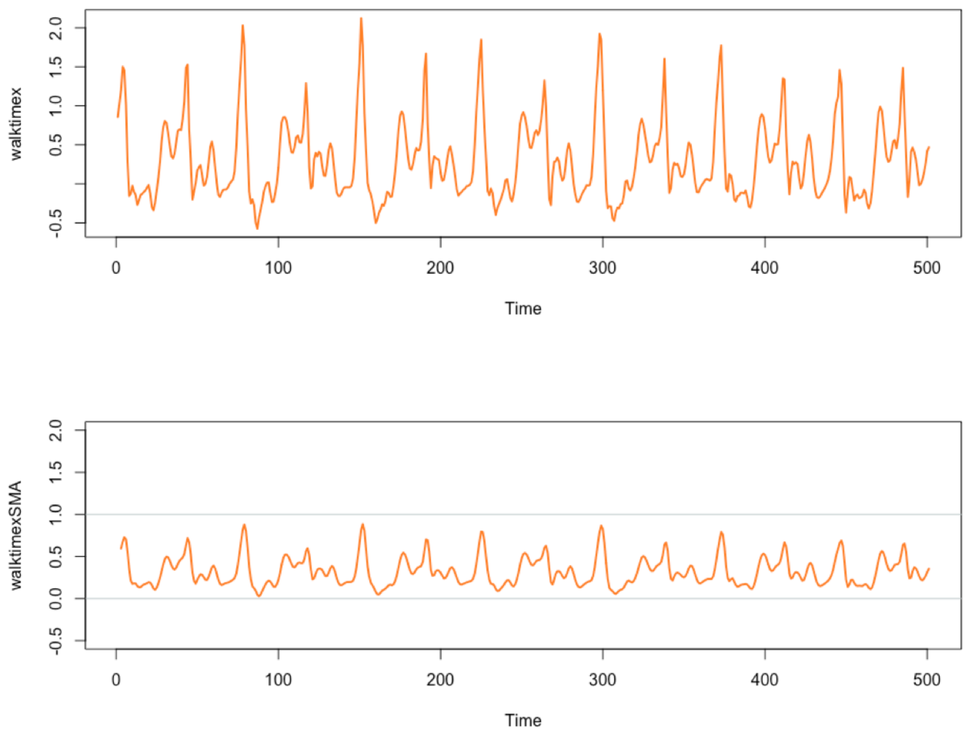


Figure 4.2: Activity time series before and after applying SMA technique

Our method is based on single moving average smoothing (SMA) technique. SMA avoids noise and smoothens the pattern in the times series. The formula is shown below.

$$SMA_t = \sum_{i=t-N+1}^{t} \frac{x_i}{N}$$

SMA_t is single moving average for the t^{th} data point, x_i is the i^{th} time series data point, and N is length of the moving average window.

4.3.2 Peak Detection

The nature of human motion is based on repetition of movement phases [Ignatov, 2016]. This is an important feature of human motion data which helps us analyze and classify activities. The data point values that are recorded by the accelerometer signal in every point of time my change in each cycle, but the general shape of the pattern in the time series stays the same. For example, when a person is walking, each leg goes through a stance phase, a swing phase and then returns to the stance phase again [Shultz et al., 2009]. When a person is performing a non-ambulatory activity, such as brushing teeth, the repetition phase in the time series data is less visible. Usually in non-ambulatory activities repetition is connected with respiration phases [Ignatov, 2016].

We can find the repetition pattern in activity time series signal and based on the average size of these patterns we measure *repetition periods*. We denote the average number of data points in the repetition period with P. Repetition period is a segment of time series, measured during one cycle of motion [Ignatov, 2016], such as a step. Figure 4.3 shows that finding peaks can help us segment the time series into individual steps or pattern repetitions. Every peak in the accelerometer time series denotes that there has been a sudden increase in acceleration, followed by a sudden drop. In every step when the lower limb goes through a swing phase we can see a sudden rise in the accelerometer value. We can therefore use such peaks to determine the step size (number of data points in each step).

In our method we use a common peak detection algorithm to detect the peaks. We use Python[1] PeakUtils [Negri, 2014] package to identify peaks and find their indexes. This

[1] https://www.python.org/

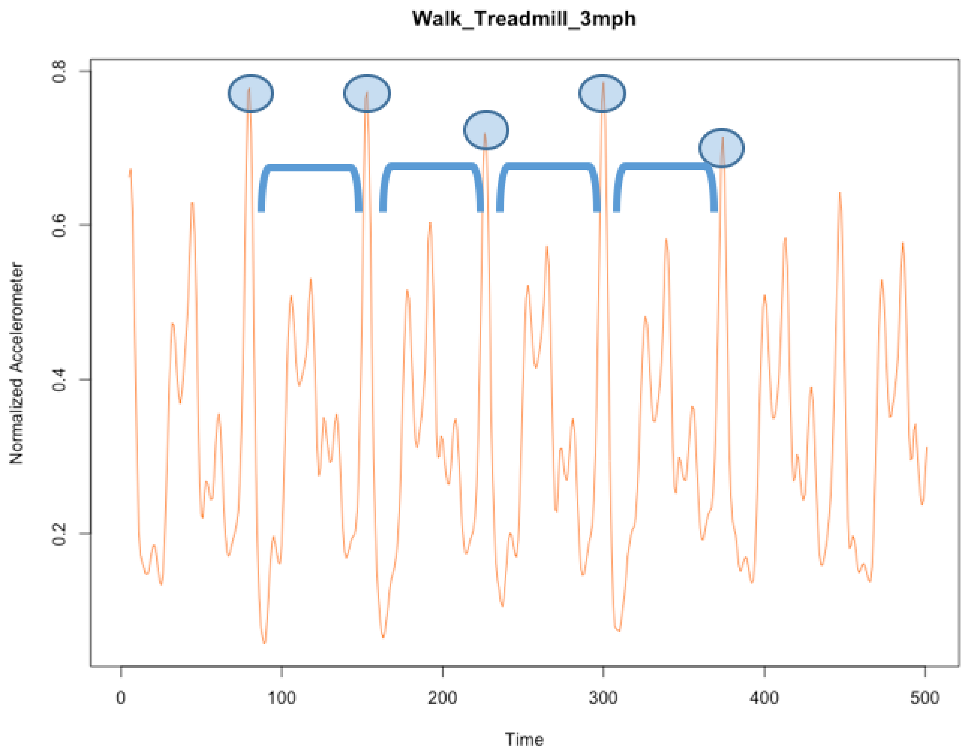


Figure 4.3: Finding peaks on walking activity time series

package provides utilities related to the detection of peaks on 1-Dimensional data by taking the first order difference and applying a normalized threshold. We find the indexes of the peaks from the activity time series as illustrated in Figure 4.3. Once we have found all peaks we find the number of data points between each two consecutive peaks in the time series. As the peaks identified by this package are rough approximations, there may be some outliers in the distance between consecutive peaks. Therefore we take the median of all peak distances instead of average. Using median instead of average enables us not consider peak distances that are too short or too wide and therefore eliminate any outlier peak distances. We apply peak detection algorithm to the data once for finding positive peaks, which we call *upper* peaks, and once for negative peaks. We call the negative peaks *lower* peaks. For finding the lower peaks, the same peak detection algorithm should be applied to the negative of the time series. The median of the number of data points in between upper and lower peaks are denoted by P_{upper} and P_{lower} respectively.

4.3.3 Shapelet Extraction

Given a time series TS with a median peak distance of P, we can extract candidate shapelets from the time series. Candidate shapelets are the set of all overlapping time series subsequences. The Candidate shapelet sizes will all have the same size, P. We start from the first data point of the time series and select the first P data points as our first candidate shapelet, we then move to the next data point and select the next P points. We continue these steps until we reach $(||TS|| - P)$th data point of the time series, every time jumping only one data point. All theses time series subsequences will be stored in an unordered list called the Candidate Shapelets Dataset. We will then select the best shapelet among the candidate shapelets.

4.3.4 Shapelet Selection

In this section we present our proposed algorithm to select a single shapelet from the list of candidate shapelets, such that it has the closest pattern to the time series. In other words we will select the shapelet which has a pattern that is most similar to the dominant pattern in the times series. The proposed algorithm checks each shapelet against the entire time series and computes a similarity value. We then use this similarity value to select the shapelet that is most similar to the fundamental or dominant pattern in the time series. This algorithm is based on a brute-force method inspired by Template Matching method in image processing.

In comparing the shapelet to the time series we use a sliding window to measure the similarity of each Candidate Shapelet to a time series. In every step, the sliding window selects a set of P consecutive points in the times series, and finds the distance to the shapelet, which is also of size P. We use two different methods for measuring the distance of the time series segments and shapelets, which we will go through in detail in the section 4.4. The brute force algorithm and the sliding window guarantee that we will go over every single possible segment of the time series and measure its similarity with the shapelet. Once the sliding window passes through all the data points in the time series, a total value of similarity will be returned by the method. This value denotes the similarity of a particular shapelet and the activity time series. However, the brute force method takes a long time and is computationally expensive operation. Therefore, we parallelized the algorithm using PySpark[2] which is the Python API for Apache Spark [Zaharia et al., 2016]. Parallelizing the brute force method enables us to run the main function and execute various parallel operations on a cluster. We successfully reduced the run time of the brute force shapelet selection method by 75%.

[2]http://spark.apache.org/docs/latest/api/python/pyspark.html

4.4 Matching Time Series with Shapelets

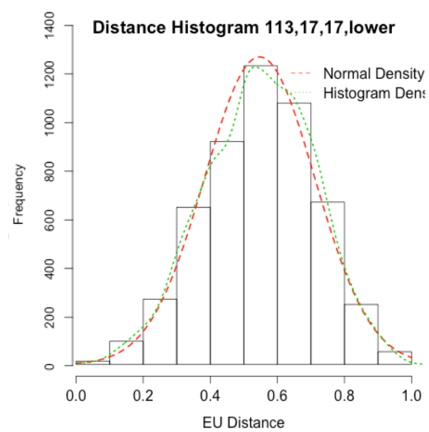
An important task in time series analysis is the estimation of similarity among different time series [Cassisi et al., 2012]. In activity time series a similarity measure is a relation between a shapelet and a time series. The algorithm requires comparing the time series to each candidate shapelet by evaluating the distance function and keeping track of the shapelet with the lowest distance to the time series. Shapelet matching requires that the shapelet SH be placed at every possible offset within the time series. In the next sections we describe two methods we have experimented in this study to estimate the similarity between a shapelet and a time series.

4.4.1 Euclidean Distance

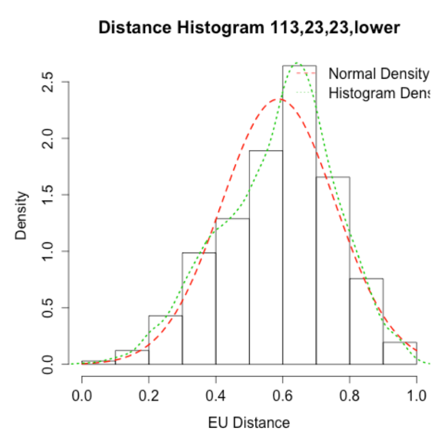
A common way to compare time series data involves the concept of distance measures. Let two time series x and y be the length of n, and x_i and y_i the i^{th} values of x and y, respectively. Euclidean Distance of the two time series is the sum of the point-to-point distances along all the time series data points.

$$||\bar{x} - \bar{y}||^2 = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}$$

Shapelets have smaller lengths compared to time series, therefore when we compare a shapelet to a time series, we use a sliding window in the time series that is the same length of the shapelet. We compute the distance of the shapelet to the part of the time series in the sliding window, we store the distance in a variable, and shift the sliding window to the next data point, we then calculate the euclidean distance and add the value to the previous value. As the sliding window goes through all the data points in the time series, the total distance gets accumulated each time. Once the sliding window reaches the end of the time series the total euclidean distance is a single value representing the distance between the shapelet and the entire time series. Figures 4.4 and 4.4 show the histogram of the euclidean distances for comparing a shapelet to a time series. Each euclidean distance between the shapelet and the sliding window of the time series is plotted in the histogram. It can be seen that when the shapelet is compared to a time series of the same activity (of same user) the histogram resembles a bell shape distribution. This is usually not the case when we compare a shapelet to a time series of another activity. Figure 4.5 illustrates this.



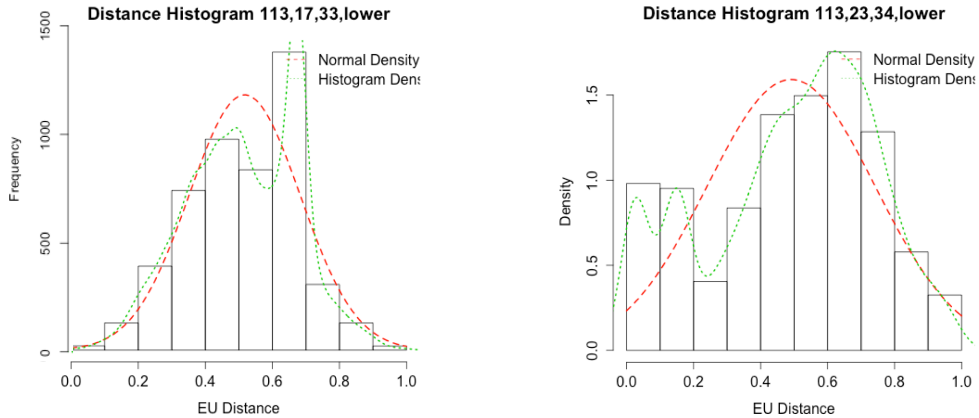
(a) Jogging at 6 mph of user 113



(b) Normal walking of user 113

Figure 4.4: Histogram of euclidean distances of a shapelet compared to the same activity time series windows

Once every candidate shapelet is compared to the activity time series and a total euclidean distance is computed for each shapelet, we will select a single or a set of shapelets as our final representatives for each activity class. We experimented two approaches to



(a) Jogging compared to walking upstairs (b) Walking compared to walking downstairs

Figure 4.5: Histogram of euclidean distances of a shapelet compared to another activity time series per window

select a final shapelet from the candidates and we compare these selected shapelets to the shapelets selected in the method described in the next section. In the first approach we select a single shapelet based on the least euclidean distance. We sort the shapelets based on their euclidean distances and select the shapelet which its euclidean distance is the median of the first decile of best distances (least distance). We use the median of the first decile of euclidean distances instead of the shapelet with the least euclidean distance because after performing many experiments we observed that the shapelets with the least euclidean distances were not expressing any visible pattern. Usually such shapelets resembled vertical lines with no trend or pattern, which are definitely not a practical representation of the time series. In light of these findings in the first approach, we explored another approach for selecting a shapelet. This approach selects a set of 10 shapelets instead of a single shapelet, aligns and averages the shapelets to generate a new shapelet. This approach also returns a single shapelet, but the shapelet is built upon 10 shapelets that had minimum euclidean distance with the time series. Figure 4.6 presents our second approach to select a shapelet. We preform a simple averaging on the shapelets to create the new shapelet.

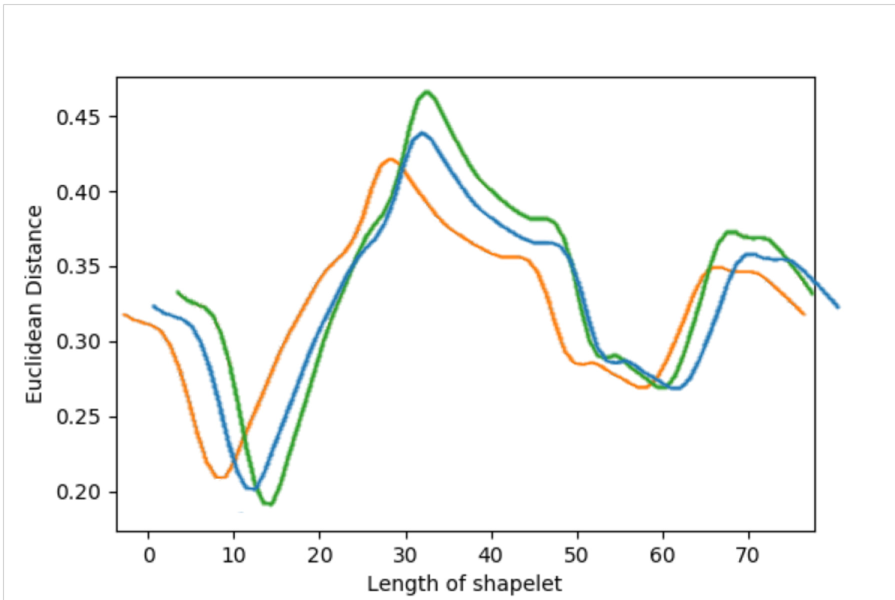


Figure 4.6: Aligning and averaging multiple shapelets to generate a new shapelet representing an activity class. Blue shapelet is the result of aligning and averaging the green and orange shapelets.

Euclidean distance is a simple and commonly used method for finding the similarity of time series, however it has several drawbacks. Activity time series may be similar in pattern but they may not be aligned in the time phase. This will cause the euclidean distance to measure the similarity incorrectly. Euclidean distance and its variants are not robust in phase differences in time of time series. In order to reduce the error associated with the Euclidean distance metric we study another similarity measure explained in the next section.

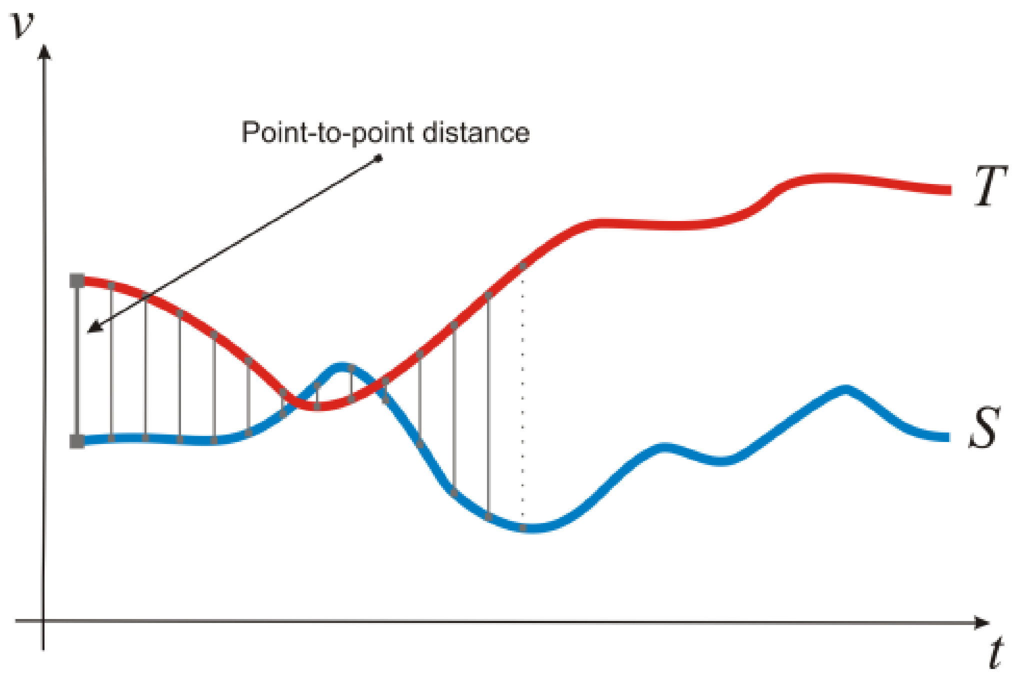


Figure 4.7: T and S are two time series along the time axis t. The Euclidean distance results in the sum of the point-to-point distances along all the time series [Cassisi et al., 2012]

4.4.2 Dynamic Time Warping

As we mentioned in the previous section Euclidean distance is an efficient distance measurement that is the sum of squared distances from each point in the time series to each point of the shapelet. One of the disadvantages of Euclidean distance is that if two time series are identical, but one is slightly shifted along the time axis, then their Euclidean distance will show them being different or distant from each other. Most algorithms used to compare time series use the Euclidean distance or some variation of this technique. However, since Euclidean distance is sensitive to distortion [Ratanamahatana and Keogh, 2004] we need another technique for calculating the distance between time series that would ignore the shifts in the time dimensions of the time series.

Dynamic Time Warping (DTW), is a widely used algorithm for computing the distance and alignment of time series [Seto et al., 2015]. DTW is less sensitive to time series shiftings, thus allows us to measure the similarity of time series even if they are out of phase in time. Although DTW has a time complexity of $O(n^2)$, but it is still the best solution known for time series problems in a variety of domains [Ratanamahatana and Keogh, 2005]. This method is much more robust compared to other similarity measures such as euclidean distance. It finds the best alignment between time series by finding the path through the grid that minimizes the total distance between them. Given two time series $x = x_1, x_2, ..., x_n$ and $y = y_1, y_2, ..., y_m$ of length n and m, respectively, an alignment by DTW uses information in a $n \times m$ distance matrix [Cassisi et al., 2012]:

$$distMatrix = \begin{bmatrix} d(x_1, y_1) & (x_1, y_2) & \dots & (x_1, y_m) \\ d(x_2, y_1) & (x_2, y_2) & \dots & (x_2, y_m) \\ \vdots & & \ddots & \\ d(x_n, y_1) & & & (x_n, y_m) \end{bmatrix}$$

where $distMatrix(i, j)$ is the the distance of ith point of x and jth point of y, with $1 \leq i \leq n$ and $1 \leq j \leq m$. The objective is to find the path $W = \{w_1, w_2, ..., w_k, ..., w_K\}$ of continuous datapoints on the *Distance Matrix* such that it minimizes the following function [Cassisi et al., 2012]:

$$DTW(x,y) = min\left(\sqrt{\sum_{k=1}^{K} w_k}\right)$$

It is worth noting that although with DTW distance there may be several warping paths of minimal total cost, DTW is well defined and much more effective in finding similarity of time series. In this study, our experiments show that DTW yields the most precise similarity between activity time series and shapelets.

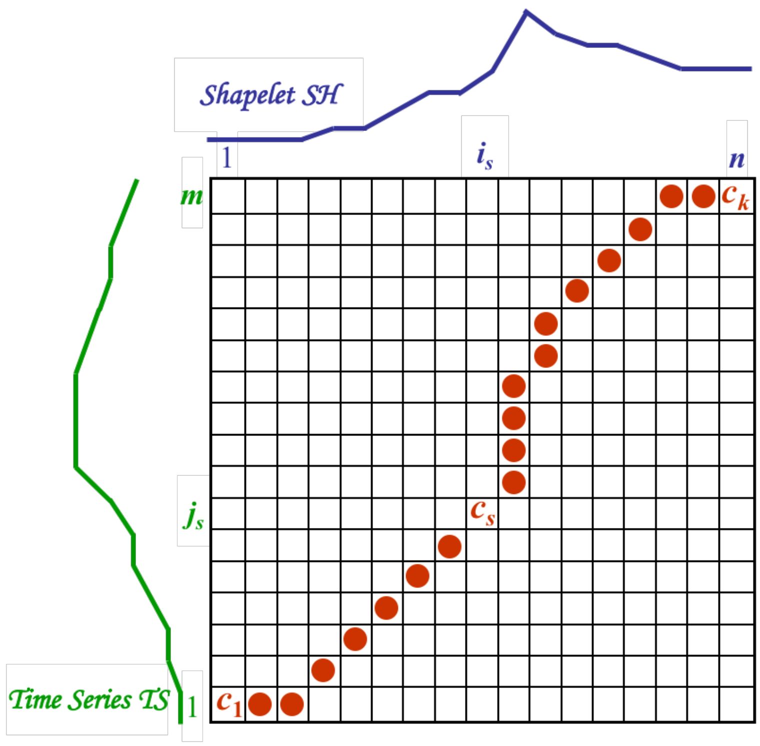


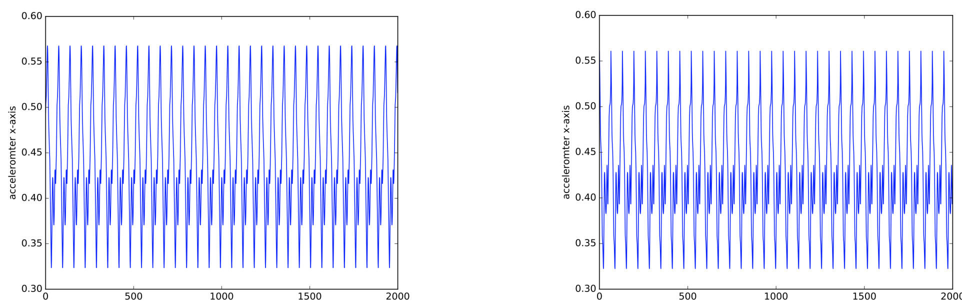
Figure 4.8: Dynamic Time Warping Alignment Path

DTW can handle time series of different lengths, and there is no evidence in literature to support the claim that comparing variable length time series has less accuracy

[Ratanamahatana and Keogh, 2005]. However for visualization reasons we linearly reinterpolate the shapelets to have the same length of the time series. We call this *expanding shapelets*, and it requires a shapelet to be repeated t times where t is:

$$\left\lceil \frac{\|TS\|}{l_{sh}} \right\rceil$$

Figure 4.9 illustrates how a shapelet for a particular activity is expanded.



(a) Expanded shapelet from lower peaks (b) Expanded shapelet from upper peaks

Figure 4.9: Jogging on treadmill expanded shapelets for user 113

4.5 Experimental Evaluation

The previous section we discussed how shapelets for each activity are extracted from the activity time series. It is worth noting that the proposed system is a personalized model system, every user would have a personal dictionary of activity shapelets based on their activity data. In this section we will discuss the details of training and testing the system.

4.5.1 Training Phase

The training phase of our proposed system will take place offline. In this phase we will extract shapelets from activity time series of each user, using methods described in section 4.3. We will select the best shapelet that we believe is most representative of the users activity among all the shapelet candidates based on similarity measure metrics. Each activity shapelet will be stored in a local database for a specific user. Figure 4.10 shows a graph of euclidean distances when comparing treadmill jogging time series of user 118 to different shapelets of the same user. It can be clearly seen that when we compare the activity time series to the shapelet of the same activity (treadmill jogging), the euclidean distances are lower compared to when the time series is compared to shapelets of other activities. We find a threshold for euclidean distances to classify a time series based on the euclidean distance with shapelets. When a time series is compared to a shapelet of the same activity, the euclidean distances are comparably smaller. This is the key component for classifying the time series class. In the next section we describe how we train a model to find this threshold and classify the time series based on the euclidean distances it has with shapelets.

Classifying Time Series

We will then compare each activity time series of a particular user to the set of its activity shapelets. We break each activity time series into chunks. A chunk CH_i is a partition of a time series with a limited length. In our experiments we have set the chunk size to be 3 seconds, and since the sampling rate is 100 Hz, every chunk will have 300 data points. The concept of a chunk is helpful for simulating the nature of stream sensor data. A chunk

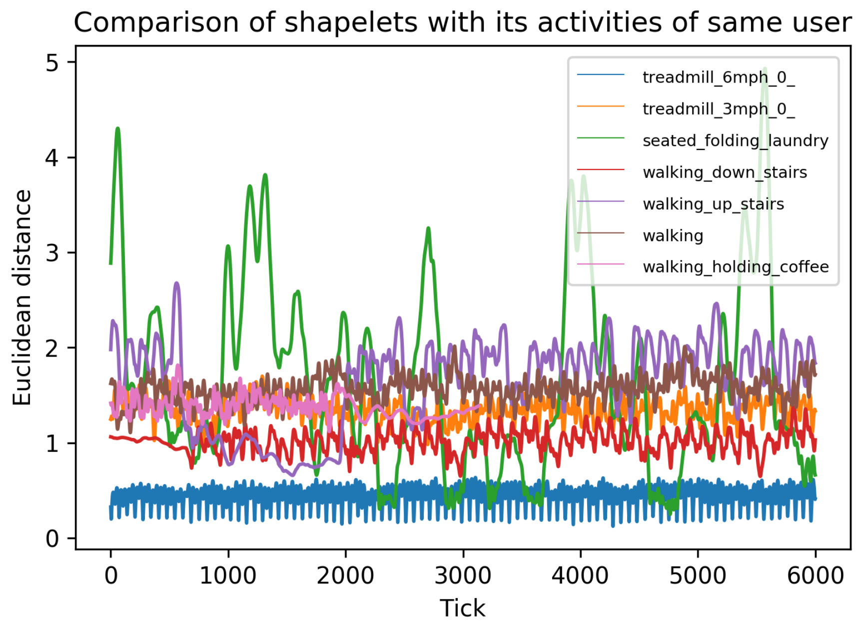


Figure 4.10: Euclidean distances graph of comparing treadmill jogging time series to different shapelets of the same user. It can be seen that comparing the time series to treadmill jogging shapelet (blue line) results in lower euclidean distances.

of 3 seconds infers that the program would delay 3 seconds in recognizing activities since it would need to wait for 300 data points to be buffered, then analyze them to predict the class. The activity time series will get compared to each shapelet in chunks, and the euclidean distance will get recorded. Once the buffered chunk of time series is compared against each shapelet, we will analyze the euclidean distances. For each shapelet we have a series of euclidean distance values for which we find the *five number summary* statistics in addition to variance and standard deviation. These statistic summaries provide a base line to compare how the time series is being matched to each shapelet.

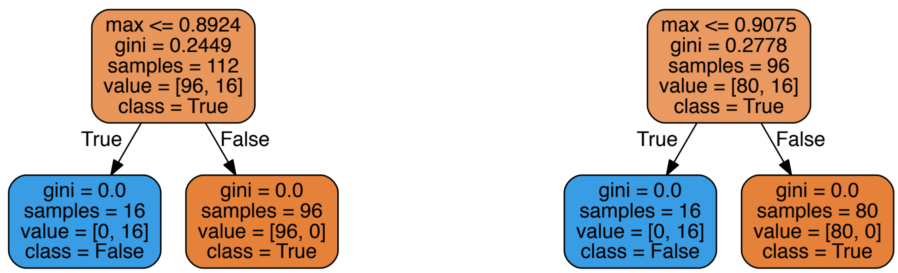
Table 4.2: Summary Statistics of Euclidean Distance of Time series and Shapelets

	Median	Min	Max	Var	STD	p25	p75
	User 118 - Treadmill Jogging Activity						
Activity Shapelet							
Treadmill walking	1.564	1.369	1.768	0.009	0.097	1.494	1.636
Treadmill jogging	**0.592**	**0.310**	**0.694**	**0.008**	**0.094**	**0.501**	**0.633**
Walking hard surface	1.139	0.940	1.321	0.008	0.0932	1.070	1.208
Upstairs	1.708	1.612	1.847	0.002	0.0542	1.681	1.758
Downstairs	0.456	0.327	0.556	0.004	0.063	0.391	0.501
	User 126 - Treadmill Jogging Activity						
Activity Shapelet							
Treadmill slow walking	1.826	1.695	2.0477	0.009	0.099	1.786	1.943
Treadmill jogging	**0.587**	**0.148**	**0.729**	**0.016**	**0.129**	**0.535**	**0.644**
Walking hard surface	0.750	0.623	1.074	0.014	0.121	0.684	0.904
Upstairs	1.004	0.737	1.241	0.0164	0.128	0.887	1.119

Table 4.2 shows the summary statistics for two treadmill jogging activities performed by two different users. For each user we have compared the time series to their set of

shapelets. It can be seen that when the activity is compared to the shapelet of the same activity the summary statistics are significantly distinguishable from other shapelets. Based on these summary statistics we train a decision tree to classify the time series. Decision Trees are a non-parametric supervised learning method used for classification and regression. We create a model for each shapelet to predict the class of a time series by learning simple decision rules inferred from the statistic summaries seen in Table 4.2. The Decision Tree Classifier is implemented using the well-known Python machine-learning library, Scikit Learn [Pedregosa et al., 2011]. The classifier takes as input an array holding the training samples, and another array holding the binary class labels of TRUE or FALSE, True being the state if the shapelet and the time series are of the same activity and false when they differ. After fitting the tree, the model could be used to predict the class of activities of time series which are not labeled.

We use decision tree model because they have features which make them favorable over other classifiers. First becuase the trees do not require much of data preparation. Second, because the cost of using the tree is logarithmic in the number of data points used to train the tree. Figure 4.11 illustrates a visualization of the models for treadmill jogging shapelet for two users. The decision tree models in figure 4.11 are very simple and easy to understand. The reason is because, based on the values in Table 4.2, the summary statistics for the shapelet that matches the activity are distinguishable from the other shapelets statistics, therefore a simple decision rule can easily classify it. However other shapelets may have a more complicated tree as seen in Figure 4.12. We can infer that the summary statistics for this shapelet (treadmill slow walking), were not as distinguishable from the other shapelet, and therefore a more sophisticated tree is needed to classify the state. The model for the shapelet of treadmill slow walking is more sophisticated becuase the activity is very similar to a non-ambulatory activity and therefore the shapelet cannot capture a distinguishable pattern. Shapelets which capture the pattern of the activity will have a simpler model since they are easy to distinguish.



(a) DT model for shapelet 17 user 118 (b) DT model for shapelet 17 user 126

Figure 4.11: Decision Tree Model for Treadmill Jogging Shapelet of users 118 and 126

4.5.2 Testing Phase

For each user we have extracted a shapelet from each of the individual's activity time series. We believe the selected shapelets are the best representative for each activity. Based on the shapelets and the activity time series we have also created decision tree models for each shapelet. Each user will have a personalized library of shapelets and decision tree models. Figure 4.1 shows the architecture of stages and components of our proposed online HAR system, using a library of shapelets and models. The shapelets and models are extracted and created in the training phase which can be offline. In this section we will describe the testing phase which is online with a 3-5 second delay in detecting the

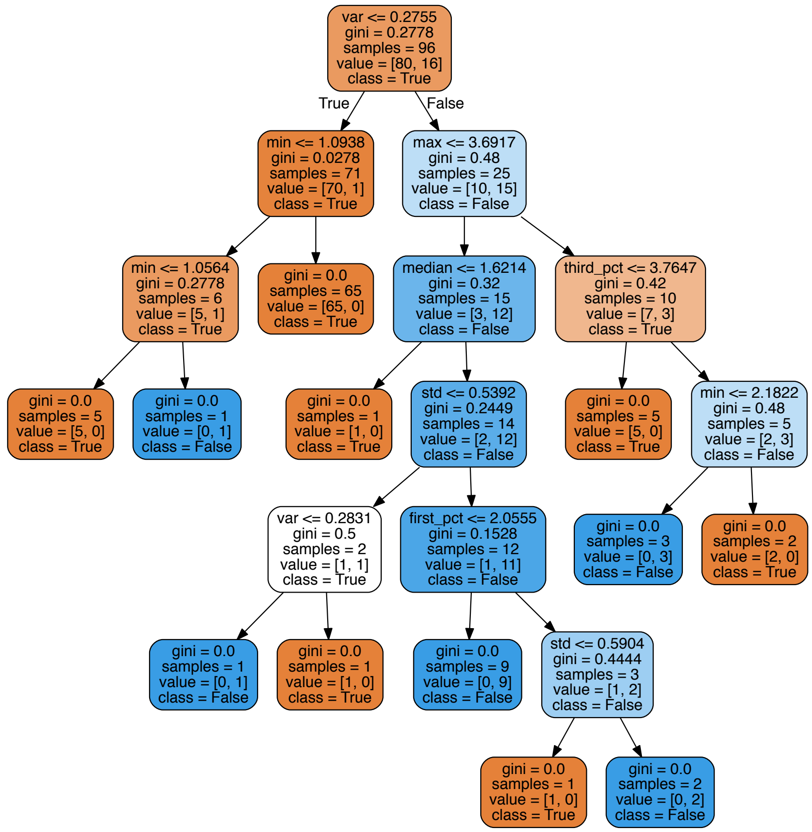


Figure 4.12: Decision Tree Model for User 126 performing treadmill slow walking activity

first activity.

Because of the stream nature of the accelerometer sensor data, we need to simulate the testing phase to generate time series similar to streamed data. We therefore take advantage of the concept of chunks again. We will read the activity time series data in chunks of 300 data points (3 seconds), buffer it, then analyze and classify it with the correct activity label, then the next chunk will be read and analyzed similarly. We set the chink size to be 300 and read 2 chunks at a time to assure that chunk sizes are larger than the shapelet size. We also need to make sure that each chunk contains only one activity type. Studying human daily activities time series we see that a 3 second window frame is an ideal chunk size, since usually activities are performed at least 3 seconds, or have a recurring pattern length less than 3 seconds. In other words, if a user is walking, then decides to jog, in most cases we see that jogging will last at least 3 seconds before the user decides to perform another activity. It is also worth noting that the activity time series in the testing phase may consist of several activities of the same user. Since this is a personalized system, each time series that is being analyzed will belong to a single user. Once the first two chunks of the activity time series are buffered in the system, they will be compared to all the shapelets of that particular user. Each time the chunks are being compared to a shapelet, euclidean distance statistics are created. These statistics include the minimum, maximum, median, standard deviation, variance, first and third percentile of the euclidean distances as the sliding window goes over the chunks and compares the chunk window to shapelets. Using the statistics provided for each shapelet and the shapelet decision tree models, we can predict the class of the chunk. The system will behave similar to a binary classifier on each model, outputting a label of yes or no for each shapelet.

Ensemble of Binary Classifier

Since there are multiple classes of activities, we need to have a multi-classifier to detect and label activities. Our system breaks the multi-classification problem into several binary classification problems with a divide and conquer strategy. We do this because binary problems are simpler to solve than the original multi-category problems [Galar et al., 2011], however, the outputs of each binary classifier have to be combined to decide the predicted class of activity. It is important to correctly combine the result of each classifier to produce a correct prediction. There are several strategies to manage the combination of binary classifiers, we use the One-vs-all method to combine the binary classifiers predictions. One-vs-all (OVA) uses a binary classifier to distinguish between a single class and the remaining ones, then with a voting strategy each classifier votes for the predicted class and the one with the largest amount of votes is predicted [Rifkin and Klautau, 2004].

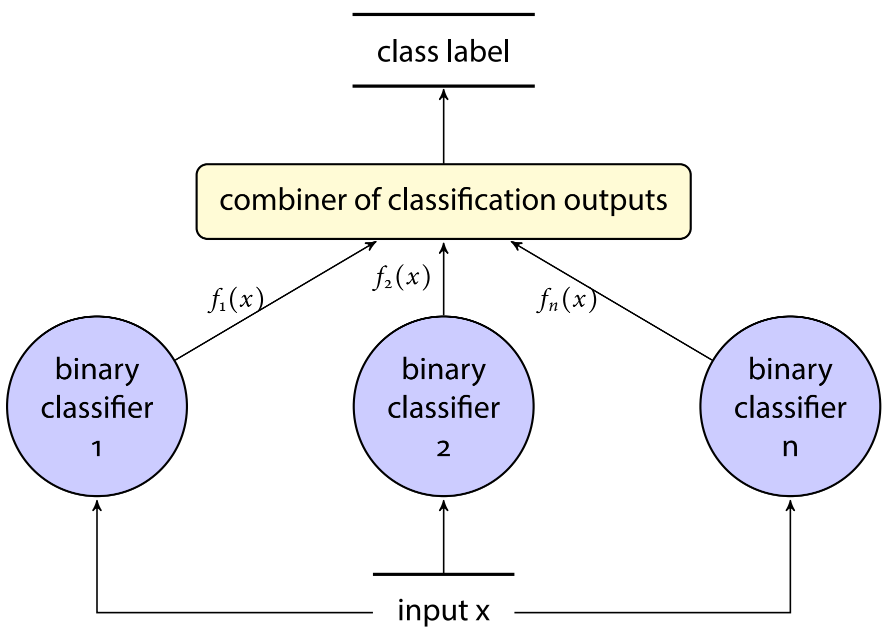


Figure 4.13: Diagram of multi-class classifier via combining binary classifiers.

We decompose the classification problem to n binary classification problems where n is the number of activity classes. When an activity time series chunk is presented to each one of the binary classifiers, the classifier which gives a positive output indicates the output class. Figure 4.13 illustrates this method. However, there exists an issue in such a system. In some cases the problem is that more than one of the classifiers outputs a positive label for a chunk of activity and a tie-breaking technique is required to decide on the activity label. In such cases we use the common approach of maximum confidence. The output class is determined by the classifier with the largest confidence score [Galar et al., 2011].

$$Class = arg\ \ max\ \ r_i\ ,\ \ i = 1, \ldots, n$$

4.5.3 Results

In this section we discuss the results of our system. We randomly select 5 users to detect their activities. Each user has an activity time series which consists of multiple activities performed in an arbitrary order. In the training phase we have already created a shapelet for each of the activities preformed by the user and we store these shapelets in the shapelet library for each user. We have decision tree classifiers for each shapelet, and we store all the models in the model library in our personalized system. Each activity time series will be read chunk by chunk and analyzed. The system will classify each chunk and display the label of the activity. We use a single time series for each user. Once activities are classified by the system we evaluate the results using an automated program. The program checks each chunk and its label against the ground truth and provides performance metrics. Table 4.3 summarizes the performance metrics of our system for each user. Each row in this table summarizes a performance metric. These metrics are the result of the average performance metrics over all the activities performed by the user. In order to make the results comparable between users we have selected users that have performed the same activities for the same amount of time.

Table 4.3: Summary of classification performance metrics for random users

Description	User 102	User 113	User 118	User 126	User 143	Total
Sensitivity	0.77	0.20	0.59	0.40	0.25	0.44
Specificity	0.94	0.88	0.95	0.97	0.91	0.93
Precision	0.68	0.20	0.65	0.73	0.31	0.51
False Positive Rate	0.06	0.12	0.05	0.03	0.09	0.07
False Discovery Rate	0.32	0.80	0.35	0.27	0.69	0.49
Accuracy	0.92	0.79	0.90	0.88	0.83	0.86
F1 Score	0.72	0.20	0.62	0.52	0.28	0.47

From the results of the study presented in Table 4.3, we demonstrate that the accuracy of the system can be variable between different users. Users who have higher accuracy in activity detection have shapelets that are more representative of an activity class. The system can improve by retraining shapelets for a user who has low accuracy. Different real life situations may affect the users time series and shapelet. As an example a user may be wearing uncomfortable shoes when training data was being collected. This will effect their normal movement patterns and shapelets would not be representative of their normal movement signature. Therefore an important step in the system would be to repeat the training phase for users who show low accuracy rates.

Table 4.4 provides the summary of classification performance metrics per activity. Activities such as activity 17 -treadmill jogging- has better classification performance

Table 4.4: Summary of classification performance metrics per activity.

Description	11	13	17	23	33	34
Sensitivity	0.30	0.71	0.80	0.30	0.39	0.33
Specificity	0.91	0.94	0.93	0.88	0.95	0.89
Precision	0.50	0.77	0.76	0.41	0.69	0.36
False Positive Rate	0.09	0.06	0.07	0.12	0.05	0.11
False Discovery Rate	0.50	0.23	0.24	0.59	0.31	0.64
Accuracy	0.77	0.89	0.90	0.75	0.83	0.80
F1 score	0.38	0.74	0.78	0.35	0.50	0.35

Activity 11: Slow Treadmill Walking; 13: Normal Treadmill Walking; 17: Treadmill Jogging; 23: Normal Walking; 33:Stairs Up; 34:Stairs Down.

metrics compared to other activities. Studying the results of this table we come to the fact that the more body movement there is in a specific activity, the classifier will classify that activity with higher accuracy. This means that an activity such as jogging, which involves several singular moving components, has a defined pattern which is much more visible and dominant than other patterns in the time series. Compare jogging to an activity that does not have many moving components, such as brushing teeth; it would be a far more difficult task to extract patterns from such an activity, and therefore extracted shapelets would not contain valuable information about the pattern of the activity. Looking at shapelets of non-ambulatory activities, we see that they all look very similar to a horizontal line. These shapelets would not be effective in finding a common pattern in the time series, therefore our proposed system has low accuracy rates when recognizing non-ambulatory activities. Activity 11, which is slow treadmill walking at a pace of 1 mile per hour resembles a non-ambulatory activity and therefore has low classification performance rates. We will further address this issue in our future work. However, it is important to reiterate that we perform classification on a large range of activities, many of which are in the same category (for example we have a variety of walking activities, in each one the subject is asked to hold a different object while walking, see Table 3.1). Naturally this yields a lower accuracy rate compared to HAR systems that classify activities in a lower granularity.

In order to show that our method is sensor-type independent, we also discuss the accuracy of human activity recognition using different sensors. We have used an iPhone 6s and Samsung Galaxy S8+ to gather motion data from 5 new individuals in Athens GA local area with a frequency of 100Hz. The activities these individuals performed was a subset of the the same activities included in the original dataset, including treadmill 1 mph, treadmill 3mph, treadmill 6mph, normal hard surface walking, walking upstairs, walking downstairs. Table 4.5 describes the results of activity recognition accuracy for the auxiliary dataset described above.

The results in Table 4.5 show that using other devices for recording data has comparable performance results to the original dataset. In the original dataset an accelerometer sensor was attached to the non-dominant hip. In the auxiliary dataset we attached a mobile phone to the individuals non-dominant hip and they performed the activities in a similar but not identical setting. In this dataset the activities were performed in a less controlled environment hence more similar to real-life activities. The results presented in table 4.5 proves that our proposed method can work on any dataset regardless of the make and model of the sensor used to collect motion data.

Table 4.5: Summary of classification performance metrics for users in auxiliary dataset

Description	Performance Rates
Sensitivity	0.62
Specificity	0.86
Precision	0.59
False Positive Rate	0.14
False Discovery Rate	0.41
Accuracy	0.80
F1 score	0.61

4.6 Discussion and Analysis

We proposed an online HAR system to detect human activities in real time. First, we showed how to create shapelets that would represent the pattern of each activity time series based on the the x-axis accelerometer data. It is worth mentioning that by leveraging other axes of the accelerometer sensor such as y-axis and z-axis it is possible to compute a more precise shapelet length and therefore extract better candidate shapelets. We also show how we would select the best shapelet from the candidate shapelets using various distance metrics.

Another possible way to select shapelets is through transforming the time series to the frequency domain and finding similarities in the frequency pattern. However, the drawback of this approach is that if the shapelets are smaller than the entire activity time series (which usually are), transformation of the time series into the frequency domain will not assist in detecting the patterns because they may appear at any point of time in the time series. Therefore, we believe that extracting shapelets from the time domain is more efficient and applicable. However, spectral analysis of the activity time series will show the dominant frequency and by leveraging this we find a proper shapelet length for extracting candidate shapelets.

The euclidean distance of each shapelet to a time series was used to train a decision tree model. The model is designed to classify activity time series in real-time. Our model is unique in its ability to classify an activity using a small portion (chunk) of the time series. It can perform near real time, since both the chunk and the shapelet are very small in size and the classifier is not not relying on any preprocessed features.

In addition to Decision Trees another possible learning algorithm that can be applied are Multi-Layer Perceptrons (MLP). MLP is a supervised learning algorithm that trains on a dataset and learns a function. The advantage of using MLP in online HAR systems is that they have the capability to classify the activities in real time. We can classify the activities by giving the set of features which are the five-number summary of the euclidean distances of shapelets and activities time series. By providing the set of features and the activity labels to MLP, it can learn a function for classifying new activities by having one or more non-linear layers. The disadvantage of using MLP is that the training phase requires many iterations with many training epochs and this will cause the training phase to be slow. However, since in our online HAR system the training phase is offline we can easily disregard this disadvantage.

4.7 Chapter Summary

We introduced an online human activity recognition system to detect activities in real time based on a single accelerometer sensor time series. The system's training phase is offline, and after training is completed, it detects human activities near real-time using time series shapelets. The proposed system is a personalized activity recognition system, therefore the system has to be trained on each individual's motion data. We train and test our proposed system on motion data from five users, while performing a variety of activities. Activities performed by the users are similar in nature, as they are all walking activities. However they are performed under different conditions with different speed, and different footwear. We have shown with extensive experiments that we find the most representative shapelet for each class, which can then provide accurate and fast classification decisions in activity time series. The results validate that it is possible to detect human activities based on the time series shapelets. The results obtained were promising for the following reasons. Firstly, they represent a baseline for practical realtime activity recognition using a device with with a single accelerometer sensor. Secondly, this study explored using a single sensor for activity recognition, but many other sensors could also be utilized to improve activity recognition such as blood pressure sensor, gyroscope and GPS. And thirdly, this work suggests that it may be possible to build systems that would be capable for recognizing more complex human movements while minimizing the computation overhead and time complexity to achieve promising results in real-world conditions. Despite the promise of this method, it is worth considering that the data was gathered under controlled settings. The sensor was secured to the users hip and an student research assistant monitored the users to make sure the activities are being recorded correctly. We would expect a slightly lower performance in real-world conditions. One of the challenges of such a HAR system is that the device may not be properly attached to the body. If the device is placed in a bag or pocket, it will have additional movements that would generate noise in the data being recorded. The existence of such noise will block the movement patterns in the data that correspond to the activities. We will discuss this issue in future work section.

Chapter 5

Related Work

Chapter Overview

In this chapter, we present recent work that pertains to activity recognition while leveraging sensors and ubiquitous computing techniques. Motivated by the increasing healthcare costs, research on Human Activity Recognition systems in health and medical domains has attracted attention in both scientific community and industry. The recent technology advances has also played an important role in the growth of such systems. In this chapter we review the current research and development of healthcare mobile applications, context-aware assistive devices, and activity monitoring systems. We believe advances in HAR research will transform the future of healthcare by enabling ubiquitous monitoring of human physical and mental health state. The main hardware component of HAR systems are sensing devices that are small, low in price and very accurate, therefor facilitate low-cost continuous activity monitoring. We review major approaches and methods based on vision sensors and wearable sensors.

5.1 HAR in Health and Medical Applications

The design and development of gadgets for health monitoring has attracted a lot of attention in the scientific community in the recent decades. Such advances in technology combined with medical breakthroughs has improved the quality of life and also increased life expectancy. Technological advances in small sized bio-sensors, smart textiles, microelectronics and wireless communications, will potentially transform the future of healthcare by enabling proactive personal health management and monitoring of a patientfhs health condition [Pantelopoulos and Bourbakis, 2010]. In the last couple of years, Human Activity Recognition (HAR) has became an emerging field of research within multiple disciplines in computer science. With the help of HAR systems health monitoring is now more accessible, and cheaper than it has ever been. Various communication technologies such as mobile devices, devices with embedded sensors and vision sensors can help improve health care systems and reduce medical staff activities. Such systems would allow patients to get discharged from the hospital sooner, as the physician would be able to monitor them remotely and would be quickly notified if any abnormalities occur for the patient. Implementing such systems in a large scale would eventually reduce healthcare costs as well. Human Activity Recognition applications have spread beyond the scope of health, and many real world problems rely on activity recognition. As an example in security and surveillance HAR systems would be able to detect any abnormal activities in public areas such as airports and train stations. City Planning [Feng and Timmermans, 2013], Sport coaching and Fitness Assessment [Ermes et al., 2008], Internet of Things and Smart Homes [Das et al., 2012] are a just few of many domains that HAR can play an effective role in. In this section of the dissertation we focus on the current state in research and development of various Human Activity Recognition Systems in the field of Health and Medical sciences.

5.1.1 Activity Recognition Types

Research on Human Activity Recognition involves the use of different sensing and vision technologies. Some work has been based on the integration of many different heterogeneous sensors into the system to allow the model itself to choose the most effective ones for any given situation [Wang et al., 2010]. On the other hand many work has been based on reducing the complexity of the HAR systems, in order to work more efficiently on lower power consumption. In the following sections we will review the two main categories of HAR applications. Vision-based and Sensor-based HAR systems.

Vision-based HAR

In the past decade we have witnessed a rapid growth in the quality of cameras and video cameras resulting interest in analysis of human activities in videos and images. Human Activity Recognition using vision based systems has applications in security and surveillance, entertainment and health monitoring. In this chapter we focus on applications in the health domain. Although there is a large variety of approaches in Vision-Based activity recognition however, the problem they are solving can be summarized as : *given a sequence of images of a person performing an activity, can a system be designed to auto-*

matically recognize what activity is being performed? [Turaga et al., 2008b]. The process overview is also similar in most vision-based systems as the human has to be detected in the image or video and then the activity can be recognized. Bodor et al. [2003] present the overview of the process as shown in Figure 5.1 to capture the key stages in human activity recognition. Most previous work on vision-based activity classification has focused on using 2D video [Ning et al., 2009; Gupta et al., 2009] and single images. Different vision-based methods for activity recognition are discussed in more detail below.

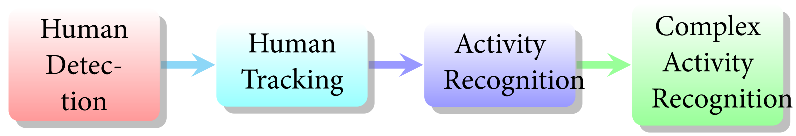


Figure 5.1: Activity Recognition Process using Vision-based Systems proposed by Bodor et al. [2003].

Video Efros et al. [2003] propose a system that can recognize human activities at a distance using low resolution video. Having stabilized human figure motion sequence, they compute the spatio-temporal motion descriptor centered at each frame using optical flow measurements. The descriptors are then matched to a database of pre-classified and annotated video sequences actions using the k-nearest-neighbor algorithm. The retrieved matches obtain the correct classification label, as well as other associated information with the human activity being performed by the subject.

In a more traditional video-based HAR , Stauffer and Grimson [2000] have developed a vision system that monitors activity in a site over extended periods of time, and detects patterns of motion and interaction demonstrated by human and objects in the site. They explore the monitoring of an outdoor site by using a set of video cameras. The main focus of their work is on the algorithmic processing of the data, rather than on the video camera locations. The forest of cameras can learn patterns of activities in a site, then monitor and classify activities based on these learned patterns. This method can handle lighting changes by slowly adapting the values of the Gaussians.

Siddiqi et al. [2014] proposed a method that has higher accuracy rates compared to other state-of-art vision-based HAR methods. They propose an accurate and robust HAR system, called WS-HAR that has high recognition rate. After the preprocessing stage they have images which have normalized intensity, size and shape. Symlet wavelet has been to extract the features from the activity frames. Features are selected by technique called stepwise linear discriminant analysis (SWLDA) that focuses on selecting the localized features from the activity frames and discriminating their class based on regression values. Finally, Hidden Markov Model (HMM) is used to classify the activities. To validate the performance of their method, two publicly available standard datasets was used to show the effectiveness of each approach.The average recognition rate for the WS-HAR was 97%.

Image In Image-base HAR systems, the image is actually a stack of video frames which demonstrates the flow of movement. Dobhal et al. [2015] have extended this method for 3-D depth maps. They remove the background of each frame using Gaussian Mixture Model, to obtain the foreground that contains the image of the individual. After combining these frames the Binary Motion Image (BMI) is calculated. Finally they use Convoluted Neural Network for training and testing their system. The importance of binary images is that it allows their method to be independent of the dress style worn by the individuals as well as the speed they perform their actions.

Infrared One of the low cost non-wearable sensors that is broadly used for activity recognition are Infrared-based motion sensors. The advantages of using such sensors are general versatility (available in darkness), low cost, small size, privacy protection (low resolution) and availability Mashiyama et al. [2015]. These sensors are very powerful in location-based activity recognition for indoor spaces such as GPS are for outdoor locations Krishnan and Cook [2014]. In recent work Mashiyama et al. Mashiyama et al. [2015] have introduced a fall detection system that uses an array of infrared sensors. Each sensor has two or more infrared detectors inside, and the temperature is achieved on a two dimensional area. They used the data from the temperature distribution, and classified activities as fall or non-fall. In their paper they mention that, knowing daily fundamental activities of elderly people is also important to prevent future fall incidents. In another published work the same authors Mashiyama et al. [2015] propose a novel activity recognition using a low resolution infrared array sensor. They analyze the temperature distribution obtained from the sensor and classify activities into five classes: no event, stopping, walking, sitting and falling. Figure 5.2 shows the algorithm of their method is divided into three steps: body detection, feature extraction, and classification. First, in human body detection, activities are classified into two classes based on the temperature difference between a person and the background. Second, in feature extraction phase, four features of human motions are extracted from temperature distribution data. Finally, in classification, support vector machine, nearest neighbor search or neural networks are used to classify the activity. Although they report their system to have very high accuracy (100% accuracy for most activities), but they have many false positives for no activity and fall classes.

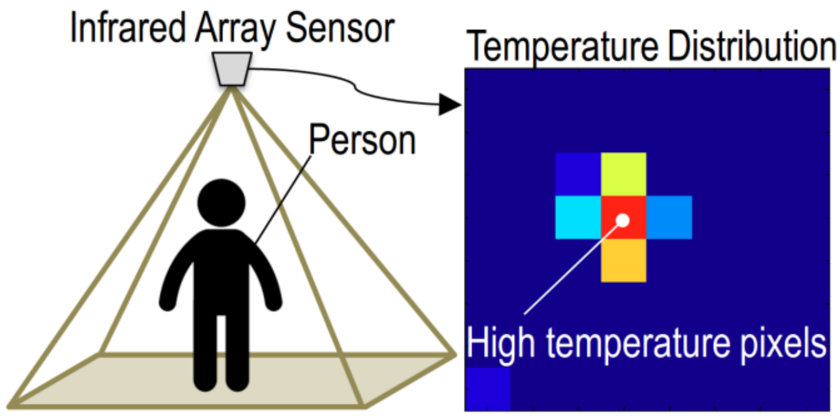


Figure 5.2: System model of Infrared HAR. Courtesy of Mashiyama et al. [2015].

Kinect (RGBD) RGB-D sensors combine RGB color information with per-pixel depth information. Sensors that provide such data have existed for years, however Microsoft[1] released the Kinect RGB-D sensor as a new Natural User Interface (NUI) for its XBOX 360 gaming platform. The Kinect, like other RGB-D sensors, provides color information as well as the estimated depth for each pixel. The Kinect is much cheaper than similar sensors therefor it has attracted the interest of researchers in HAR domain. Sung et al. [2011] use a supervised learning approach in which collected ground-truth labeled data is used for training their model. The input is RGB image as well as depths for each point collected by Kinect device. To compute the human pose features, each subject is described by a rigid skeleton that can move at fifteen joints. The skeleton is described by two features, three-dimensional Euclidean coordinates and the orientation matrix for each joint with respect to the sensor. Information extracted from these features are fed

[1]www.xbox.com/en-US/kinect

as input to the classification algorithms. A two-layered Markov model is trained which captures different properties of human activities, their hierarchical nature, transitions between activities, and the correspondence between activities and the skeletal features.

In a more recent work, Jalal and Kamal [2014] use RGBD cameras and extract depth silhouettes to create a real-time life log for elder individuals. Their system includes training routines and real-time recognition for life logging. In the training phase, depth silhouettes of human activities are captured using a depth camera. Then, body points are extracted from these depth silhouettes and computed using a set of magnitude and direction angle features which are used for training the HAR classifier. Finally, after training, the system is used to recognize learned activities via the trained Hidden Markov Models (HMMs) for activities and stores life log information in a database.

Sensor-based HAR

One of the visions of Weiser, father of ubiquitous computing, has been the prospect of disappearing technologies that "weave themselves into the fabric of everyday life until they are indistinguishable from it" [Weiser, 1999]. Sensors have made a breakthrough in the field of Ubiquitous Computing by making a shift from low-level data collection and transmission towards high-level information integration and context processing [Chen et al., 2012]. Since late 90's, various sensors have been used to monitor human activities. Since then, extensive research has been carried out to investigate the use of sensors in various application scenarios of ubiquitous and mobile computing, and considerable work has been done in the field of context awareness, smart appliances, and activity recognition [Chen et al., 2012]. Sensors used in such systems can either be attached to human body or onto objects that human interact with such as mobile devices or smart objects in the house. Before we give more details on Sensor-based HAR system, we would like to highlight that such systems can better address sensitive issues in activity recognition compared to Vision-based HAR. The main concerns in vision-based systems are privacy and obtrusiveness, which are not a problem in sensor-based systems. Traditionally Sensor-based HAR have been divided into two main subgroups in terms of the way they are deployed.These are Wearable Sensors and Dense Sensors and are described in detail in the following sections.

Wearable Sensing Systems In the past couple of years there has been an exceptional development of microelectronics and computer devices in a small size with high computational power enabling it to be the perfect sensor for attachment to human body. Wearable sensors have helped ubiquitous computing research area by allowing researchers to extract knowledge acquired by body worn wearable devices. We can group the sensors used in wearable sensing systems into three categories as Lara and Labrador [2013] have explained in their survey. The three groups of attributes measured by these wearable sensors in a HAR context are : Acceleration, Location, Physiological Signals. We provide a description of each sensor used in Wearable Sensing Systems below.

Accelerometer Without doubt, accelerometers are the most commonly used sensors in Activity Recognition. Triaxial accelerometers are particularly effective in recognizing ambulation activities such as walking, running, climbing stairs, lying and etc. These sensors are very inexpensive and require low power, therefore they have been embedded in smart phones in the recent couple of years. Several papers have been published using accelerometers to detect human daily activities, in the majority a very high accuracy is reported , as well as real-time or near real-time activity recognition. Accuracy rates have been as high 98% in Khan et al. [2010]'s work where the system is capable of recognizing a broad set of daily physical activities. Their proposed work only includes a single tri-axial accelerometer, therefore it is feasible to be used by free-living subjects throughout the day. Khan et al. [2010] report that their system is able to distinguish between sitting and standing postures and their transitions, walking-upstairs and downstairs. A main issue in accelerometer sensing systems is that they get confused for non-ambulatory activities, such as brushing teeth, talking on the phone, eating, working at the desk and etc. Becuase of similar motions in many of the non-ambulatory type of activities, accuracy rates may decrease noticeably when they are included in the activity set. Lara and Labrador [2013] mention that the placement of sensors is also an important factor in HAR systems, however there is not an optimal position where to place the sensor. Depending on the application and the type of activities being performed the optimal accelerometer position can vary. In another similar work, Najafi et al. [2003] use Gyroscope in addition to Accelerometers to detect body postures (sitting, standing, and lying) and periods of walking in elderly people. The principle operation of the gyroscope is the measurement of the rotational angular velocity. The purpose of their study is to test the performance of their system, based on one kinematic sensor and to classify body postures (sitting, standing, and lying) and locomotion (walking).

Global Positioning Systems (GPS) GPS sensors are another widely used wearable device for monitoring location-based activities. A variety of applications have used GPS sensors to recognize human behavior at a certain location, as an example Patterson et al.

[2003] use GPS sensor stream for detecting human activities like boarding a bus or traveling. Current smart phones are equipped with GPS sensors, making them an ideal device to track human activities and transportation mode. Using a GPS sensor in HAR systems along with the accelerometer sensor can help infer the users activity using ontological reasoning [Riboni and Bettini, 2011]. In order to make this concept more clear, if GPS shows the location of a user in a mall, we are sure that that person is not taking a shower, instead they are probably walking or sitting. Therefor we can limit the number of recognizable activities depending on the location of the user, increasing the accuracy of the system. Although GPS sensors have a huge impact on activity recognition but they also have a few drawbacks as well. One disadvantage of this sensor is the energy consumption which is relatively high compared to other sensors like accelerometer. GPS sensors do not work at indoor locations or in places that there are high surrounding buildings or mountains that can block the signal, so it would not be very useful to use this sensor when the user is indoors. At last but not least there has been a lot of concern about the privacy issues for GPS, because it can constantly track the location of the users. There has been some work done to encrypt the location of the user as discussed by Christin et al. [2011].

Bio-sensors Vital signs of human body could be very helpful in detecting the type of activity the individual is performing. A wide range of bio-sensors exist that could measure vital signs such as blood pressure, heart rate, heart sound, respiratory information, oxygen saturation, body temperature, blood glucose, skin conductivity, ECG and EEG. Just as GPS sensors, bio-sensors would need to be combined with other sensors in order to be able to make more accurate recognitions of human activities. Tapia et al. [2007]

mention in their paper that using heart rate individually is not helpful in detecting activities because the heart rate may remain high for some time after the user has performed an exhaustive activity. Smart Garment is a new interesting field of research being carried out by Material Scientists and Computer Scientists. It is the most pervasive HAR among all the existing since the user will not need to wear any extra device other than the shoes or garment they already have on. Harms et al. [2008] use a smart shirt, the shirt is equipped with acceleration sensors in order to determine the posture of human body. Their classification performance is analyzed on data from 8 users, with 12 postures, relevant for shoulder and elbow joint rehabilitation.They report accuracy rates of 89% for a user-independent evaluation.

Limitations

Wearable Sensor-based HAR systems suffer from limitations that is currently the focus of many researchers. One of the main issues with wearable sensors is that they need to be run continuously and in hands-free mode [Chen et al., 2012]. This would cause some difficulties for some individuals who have to wear them and not all people are willing to accept to wear a sensing device. This may be because the device is invasive and interrupts with their daily activities, or that they just feel they are giving up their privacy while having the sensing device on them. I believe educating the users of how such systems function would assure them of the privacy and more people would be willing to have them on. A part from the acceptability issues there also exists technical limitations for such systems. Below we describe some of the possible limitations:

- **Size:** Although there has been technological advances in reducing the size of sensors, many HAR systems use multiple sensors and other peripherals and attach them to the human body, making the device relatively large in size. Shrinking the size of the wearable device and limiting it to a single small sensor would be more attractive for users to try out. However, the more sources of information the more accurate the system would operate.

- **Ease of Use:** HAR systems that do not require the user to wear many sensors or to interact too much with the device are more successful. . As a general principle, there is always a trade-off between reducing the complexity of the system and the accuracy of HAR system.

- **Battery life:** Energy consumption is an important feature in wearable HAR systems because the devices used are energy constrained. Processing, communication and visualization are the most power draining tasks in activity recognition. Limiting the data process, aggregating data to be sent and omitting visualized reports on the device are a few efforts that could be made to reduce the energy consumption.

- **Effectiveness in real life:** There is an open debate on whether or not the projects and research on wearable HAR systems are effective in real life scenarios. Becuase of all the limitations discussed above, it may not be applicable to wear these devices at all times during the day. Some devices are not water resistant, some need to be recharged every couple of hours and some may be intrusive for the human subject to be willing to wear at all times. All these limitations may interrupt the effectiveness of these approaches in real life.

There is an extensive amount of research being done to address these limitations.

Figure 5.3: Stress monitoring using intrusive sensors.Courtesy of J. Choi [2010].

Smart Garments are becoming more popular and have the ability to embed the sensors in them so they would become less invasive [Harms et al., 2008]. Other ways to overcome the mentioned limitations are to take advantage of the existing gadgets that are being carried regularly by the users, such as Mobile Phones. Shoaib et al. [2015] provide an outline of relevant research done on activity recognition using smart phone sensors. They conclude in their paper that there are more than 30 study/projects dedicated to online activity recognition on mobile phone, many of which have the capability to recognize the human activity in real time with a relatively low cpu usage. They do not report directly about accuracy rates for such systems.

Dense Sensing Systems

In the previous section details about wearable HAR systems revealed that such systems may not be suitable for complex physical activities consisting motions in multiple limbs of the body. Another key weakness in such systems is the ability to distinguish between similar activities, such as making tea and making coffee. It is obvious that relying only on wearable sensors would not be sufficient for detecting such simple physical daily activities. Therefore recently a new field of research has became popular being known as Dense Sensing. Dense Sensing-based HAR refers to the practice that sensors are attached to devices that human interact with in their daily actives like the refrigerator , the door and etc. Simply said human activities are detected by the user-object interactions. The name Dense is used because there are many low-cost sensors attached to objects in a location that is going to be monitored.

RFID An important application of Dense Sensing systems are in Smart Homes. In earlier work on smart home environments, it was shown that with sufficiently accurate data a reliable in-home activity can be acquired [Sung et al., 2011]. As Philipose et al. [2004] have described in their paper, they tag objects of interest using radio frequency-identification (RFID) tags, which are small and can be easily attached to small objects as well. The user would wear an RFID reader fixed to a glove and could detect when users interact with the objects that had tags attached to them. As mentioned before, having this information would be helpful in the the process of distinguishing between similar activities. Please note that Wearable sensors discussed in the previous section is not mutually exclusive with Dense sensing systems, rather they are complementary of each other to have an accurate system. In a very interesting work Patterson et al. [2005] have introduced a system that can not only report the users activity (for example cooking), but also report the details of the activity being performed like the food that is being cooked by the user. This is done based on the interaction of the user with the RFID-tagged objects.

Motion Detectors As mentioned in the limitation section of wearable devices, an important factor of HAR systems is that it must be economical to manufacture, install and maintain. The methodology must also be efficient and scalable and most importantly privacy-sensitive. To address these requirements Wren and Tapia [2006] propose a system based on passive infrared (PIR) motion detectors with an appropriate analysis methodology, to detect human activities. As discussed in section 4.3, IR sensors are inexpensive, reliable, and require very little bandwidth. Wren and Tapia [2006] report accuracy rates of over 90% recognition. For more information on this method please refer to section 5.1.1.

Audio Detectors In HAR research area, a large amount of attention has been focused on using simple sensor data (wearable sensors, touch sensors, RFID tags) or camera information to detect the activities. There have been a few attempts to include audio information into the recognition process. In a very interesting study Hollosi et al. [2010] describe a system for detection of acoustic events in various acoustic situations that uses a voice activity detection mechanism to obtain low level information. Next the event detection stage extracts a midlevel representation from the input audio data and finally this representation is interpreted to high-level semantics to notify the need for help when necessary (eg. serious cough). This system can be expanded to recognize abnormal voice occurring in an environment such as a human fall, or shout for help. This would have huge benefits in the health care-industry for patients.

5.1.2 Summary

In this section we discussed the current gadgets available for Human Activity Recognition with the focus on applications in the context of health and medical sciences. We identified a few of many real-life applications of HAR in two main categories namely as Medical applications and Assisted Living applications. The recognition of human activities can be approached in two different ways: Vision-based and Sensor-based recognition systems. In the former approach, visual sensing devices are used to monitor user behavior and environmental changes and since they are usually fixed in predetermined points of interest, recognition of actives is limited to specific locations or rooms. In Sensor-based systems, sensors are either deployed on the users's body or on the devices they interact with. Finally, by reviewing a number of systems using both approaches we can conclude that sensor-based HAR's can better address important issues such as privacy, obtrusiveness, flexibility and cost.

5.2 Mobile-based Health Applications

The creation of healthcare mobile applications, context-aware assistive devices, and activity monitoring systems provide great opportunities to improve quality of life. Mobile health systems encompass new types of sensing and communication of the users health information and help integrate health monitoring to everyday lives, regardless of location and time. In the remainder of this chapter we discuss various technological and applicability aspects of mobile health systems that make them a promising platform for health applications. Furthermore we present critical challenges faced by the building and development of mobile health systems.

The ubiquitous nature of mobile phone devices, along with the embedded technology and sensors, processing power and low cost of the device have lead the healthcare industry to make a shift towards replacing traditional medical and health solutions with mobile health systems. Mobile Health also known as mHealth is defined as "a medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, PDAs, and other wireless devices" [Kay and Misha, 2011]. If done right, mHealth projects can enable better access to health services, raise the quality of care, and reduce health costs. With the growing number of affordable mobile network plans, free WiFi available in the community, and affordable smart mobile devices, mHealth has the potential to become a medium for intervention that is extremely affordable and easily adaptable [Shellmer et al., 2016]. In this section we will review Mobile-based health systems that are designed to enhance remote patient care in hospitals, care facilities or at home. Each applications uses one or many of the embedded sensors in the mobile phone then aggregate the collected data and send them to a remote server for further processing. Some applications have the capability to process the users data on the mobile device and would only send periodic reports or notify the physician if an abnormal state is detected. The mHealth applications differ in the technology they rely on, and also the algorithms and methods they use to aggregate and process the data. We provide a brief overview of the different technology that are commonly used in such applications. Certainly, there are some challenges/issues regarding the use mobile phones in healthcare such as; feasibility, reliability, stability, security and privacy, accuracy, user-friendliness and power consumption. Some of these issues and challenges will be addressed by referencing related studies and papers.

5.2.1 mHealth Technology

There are a variety of hardware and methods used in different mobile health interventions and some applications even rely on multiple types of technology. Now a days the smart mobile phones have capabilities varying from voice and text messaging to the support for third party applications, sensing, Internet access, and wireless connectivity [Klasnja and Pratt, 2012]. Many of the mobile health interventions take advantage of the hardware technologies and technology capabilities available on the mobile phones that are popular among people. In the following section, we briefly review some of the technical features of mobile phones and ways in which those features have been used in health and medical applications.

Text Messages

Text Messaging is broadly used for communication between two mobile phones. Since it is supported by all phones text messaging has been widely used in many health applications as well. The main reasons why text messaging is so popular in health-related apps is firstly because they have a push technology, allowing intervention messages to be delivered without any effort on the part of the recipient [Klasnja and Pratt, 2012]. Such applications could be used for sending reminders about taking or applying medicine. Armstrong et al. [2009] designed a text messaging based system to remind individuals to apply sunscreen when they are exposed to sunlight. Second, because text messages are not just limited to phones as they can also be sent and received by computers. Therefor, users can log their health-related activities and physiological parameters, maintain awareness of their health goals and receive customized feedback based on these data. Franklin et al. [2006] propose Sweet Talk, a text-messaging support system designed to enhance self-efficinacy, facilitate uptake of intensive insulin therapy and improve glycemic control in pediatric patients with Type One diabetes by sending scheduled and tailored text messages. Text messages can be processed automatically and this makes information exchange possible using text messages, so reminders would be sent to users to log relevant data. Users can reply to the reminder messages with the requested information, and the system can process the responses and re-send customized feedback, this process is known as SMS information loop [Klasnja and Pratt, 2012]. There is a large variety of different applications using text messaging and this provides evidence for how flexible text messaging can be as a part of a health application.

Mobile Camera

Cameras have become a standard in all smart phone devices. In the recent years the quality of the phones cameras has become comparable (in some cases better) with dedicated digital cameras. Constant availability of cameras that are embedded in mobile phones has made it a useful tool for creating health journals and collecting health related data. Although images cannot be processed as easily as text messages can be, in some cases the goal is supporting reflection or learning through active engagement with the user's data, so phone cameras can be a valuable tool for low-effort collection of health-related information. Mamykina et al. [2008] have proposed a distributed mobile application that individuals with diabetes use to record their blood sugar levels and diabetes-related questions using phone's camera and microphone and share their records with a diabetes educator. Phone Cameras have been used in several different ways for health related applications: Firstly as a health journal as presented by Lungu et al. [2015], secondly as a way to provide the physician with additional information about a medical condition, such as in [Schreier et al., 2008], where the patient can provide images of psoriasis lesions to the health care provider for more accurate treatments. Third way is for documenting for self-management process, such as in [Smith et al., 2006] where images are used to self-manage routines and augment glucometer data and facilitate the sharing of experiences that affect long-term health.

Mobile Sensors

Smart Mobile Phones have many sensors embedded in them which can be a great benefit to health monitoring applications. A few number of these sensors are: GPS, accelerometer, gyroscope, magnetometer, barometer, proximity and light sensor. Besides the built-in sensors, mobile phones can connect to other external sensing devices such as digital scales, blood pressure monitors, glucose meters, portable electrocardiograms(ECG), pedometers, and gym equipment. Being able to store, analyze, receive and send data collected from the mentioned sensors makes mobile phones a portable, easy to use , and low cost device to monitor and improve user's health status. The application RunKeeper

[2] uses the phone's embedded GPS to track how long users run or bicycle, and create maps of their exercise routes, and calculates how many calories were burned during these periods [Klasnja and Pratt, 2012]. Fall Detection applications are mostly built based on accelerometer, gyroscope and GPS. Such applications have been received a lot of attention from elderly population to detect the occurrence of fall accidents and help the injured person receive first aid. Shi et al. [2012] present a novel fall detection technique which describes the state change of the user's motion during the fall. Built-in sensors in mobile phones has the advantage of freeing users from the need to wear an additional device [Consolvo et al., 2008].

Native Applications

Most smartphone platforms like iOS, Android, Symbian, Blackberry, webOS, and Windows Phone provide developers with application programming interfaces (APIs) that is used to develop various applications. These APIs can have access to various features of the phone, giving the developer the power to access the phones hardware (e.g., accelerometers, cameras) and to other data and applications on the phone. Having this ability would lead to the creation of powerful health applications that researchers and industry companies have leveraged. In section 3, we will provide a detailed review of Mobile Health Application that are currently being used or are under study.

[2] http://runkeeper.com

Internet

Without doubt, Internet has played a key role in the advancement of health applications. Nowadays mobile phones have the ability to connect to the internet via cellular network from nearly everywhere [Klasnja and Pratt, 2012], and this ubiquitous connectivity means that users data can be sent or received on servers as soon as they are captured via the device. This phenomenon would lead to early detection of critical events such as an occurrence of an elderly fall or an abnormal health status of a patient. The data could also be further analyzed and processed for yearly, weekly or daily reports when it is uploaded to servers. Lastly, always-on connectivity makes it possible to include web pages and online audio and video as part of phone interventions. The use of online resources makes it easier to keep the content of an intervention up to date without requiring users to repeatedly install updated versions of the application.

5.2.2 mHealth Applications

In the year 2012 they estimated that the number of health-related apps are about 40,000 [Boulos et al., 2014], and this number will increase significantly in the years to follow. Mobile health applications can be viewed in a number of different ways, many categorize them as applications for health care professionals and applications for patients. In this review we would focus more on applications developed for patients, since they have a broader utilization among the users. Although some patient health applications cover a broad spectrum of general medical utilizations, others may be tailored to specific specialties. In the following sections we will briefly review mobile health applications in different categories. We must note that issues such as cost, connectivity, coverage, low literacy, and high diversity of users exists in many of the current applications and we will address these issues in section 5.2.3.

Daily Activity Monitoring

Mobile applications for activity monitoring are a very important and necessary field in the improvement of health and medical care. Physical activity has a correlation with overweight, obesity, and metabolic-related syndromes [Sherwood et al., 2013] therefor daily activity monitoring provides physicians with the ability to monitor and diagnose patients using continuously generated data and help detect if a patient has deviated from a typical routine. Activity Recognition based on acceleration data enables the usage of smart mobile phones for measurement and detection of physical activities performed by the user who is carrying the phone in a pocket. The mobile device can provide information about the type, intensity and duration of the performed activity. Therefore, human activity monitoring and recognition in mobile phones can have many applications such as evidence for medical diagnosis, treatment of diseases and recognition of unhealthy habits.

Khan et al. [2010] present a state-of-art recognition method which uses a hierarchical scheme. At the lower levels, the state of activity, i.e., static, transition, or dynamic, is recognized by statistical signal feature analysis and artificial-neural networks. At the upper level, autoregressive (AR) modeling of the acceleration signals is applied. Using the AR-coefficients with the signal-magnitude area and tilt angle results in a feature vector which is fed into a Neural Network to recognize a particular human activity. Using a single triaxial accelerometer attached to the subject's chest, their system recognizes three

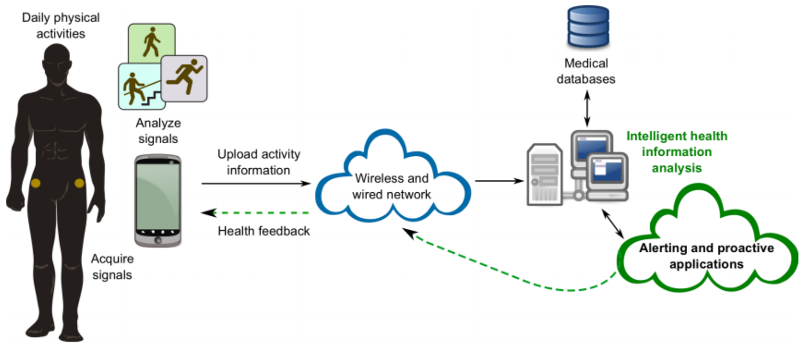


Figure 5.4: A general overview of a smartphone-based activity recognition for health monitoring. Courtesy of Torres-Huitzil and Alvarez-Landero [2015].

states and fifteen activities with an average accuracy of 97.9%. Gao et al. [2009] have presented a novel health-aware smart phone system named HealthAware, which utilizes the embedded accelerometer to monitor daily physical activities and the built-in camera to analyze food items. The system is composed of an on-device database which holds the user specific data and food item information.The physical activity analysis system works at background to obtain the users physical activity by analyzing the accelerometer data. The food item classification system is responsible to take a food picture, extract meta data from the picture and index into the database for further use.

An existing challenge in Activity Recognition systems that are designed for healthcare domains is that they need realtime assistive feedback, and not many of the current system have this feature. Limitations of mobile phone battery life, processing power and connectivity are a few other reasons that realtime feedback is challenging and there is room for further research in this area.

Fall Detection

Fall Detection is another major challenge in the healthcare domain, especially for the elderly population. Statistics in [Griffiths et al., 2005] present that falls are the primary reason of injury-related death for seniors aged 79 or more and also the second leading cause of injury related (unintentional) death among all ages. Therefor the demand for fall detection systems has increased rapidly. It has become very important to develop mobile-based intelligent fall detection systems which can automatically monitor and detect falls. The maturity of cameras and sensors built-in mobile phones make it feasible to deploy fall detection as they are available in both indoor and outdoor environment, user-friendly, requires no extra hardware and service cost, it is lightweight and power efficient. An effective fall detection system is required to provide urgent support and to significantly reduce the medical care costs associated with falls [Mubashir et al., 2013]. It is obvious that the medical consequences of a fall are dependent on the rescue response time, therefor a highly-accurate and fast fall detection system is likely to have a significant role in raising the confidence levels of supportive living for elder population. Fall detection approaches are divided into three main categories: wearable device based, ambience device based and vision based. In this section we only focus on wearable fall detection and more specifically on mobile-based fall detection systems.

Dai et al. [2010] proposed a fall detection application named PerFallID, which utilizes mobile phones as a platform for pervasive fall detection using the phones built-in accelerometer. Advantages of PerFallD is that it has has few false positives and false negatives. They collect fall data in different directions (forward, lateral and backward), different speeds (fast and slow) and in different environments (living room, bedroom, kitchen and outdoor garden). Data of activities of daily living (ADL) including walking, jogging, standing and sitting are also collected. In their system the user has a dedicated profile which is loaded when the application starts. A user dependent profile contains basic fall detection configuration such as the default sampling frequency, default detection algorithm, emergency contact list and etc. If information collected in real time satisfies a certain preset condition, the pattern matching process begins to determine if a fall occurs. If a fall is detected an alarm is triggered and starts a timer. If the user does not manually turn off the alarm within a certain time period, the system automatically calls emergency contacts.

Wearable devices have their advantages as well as disadvantages. The biggest advantage remains the cost efficiency of wearable devices. Mobile-based fall detection systems are relatively easy to operate. The disadvantages are that mobile phones are not fixed to the human body and therefore may be disconnected to the body and cause false positives detections. Also elder people may not remember to have their mobile phones with them at all times this disadvantages make mobile devices an unfavorable choice for the elderly. At last the battery life of the phone is a limitation and for long durations that the user does not have access to a power source this may cause to phone to not be operable and thus the fall detection would not work.

Location Tracking

Advances in location based services on mobile phones have lead to opportunities in real time patient tracking for healthcare applications. Chew et al. [2006] present a mobile-based location technique using the Global Positioning System (GPS) and cellular mo-

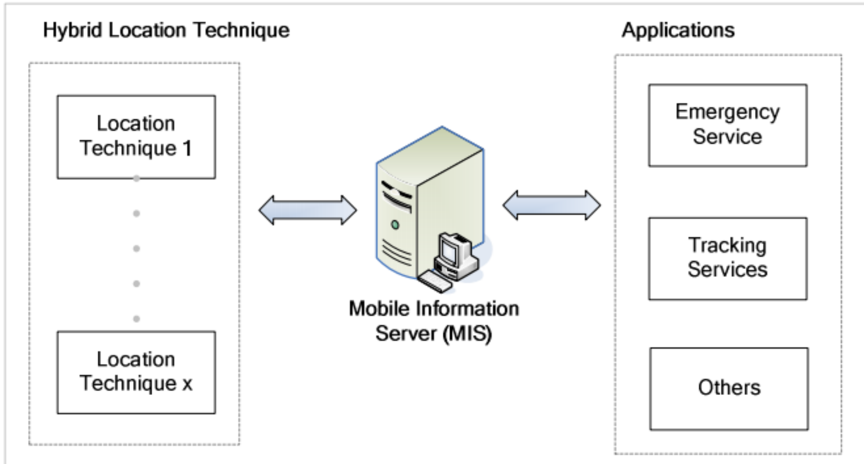


Figure 5.5: Overview of patient location tracking system architecture. Courtesy of Chew et al. [2006].

bile network infrastructure to provide the location tracking capability to assist caregivers or family members in locating patients such as elderly or dependents. Their model has shown relatively good accuracy for position tracking and can potentially use wireless to enhance the existing personal healthcare communication system through location based services. Figure 2 shows an overview of their system.

Another very recent and similar application is MedNav[3], a location-based mobile health application that bridges the gap between patients and clinics with real-time availability of information. The application allows patients to view live wait times, review verified ratings and schedule appointments, while eliminating excess capacity and increasing patient satisfaction for health care providers.

[3]http://www.mhealthnews.com/press-release/location-based-mobile-health-app-launches-washington

The main drawback of location-based Mobile Health Applications is privacy. Location tracking and monitoring aspects of such applications may have a negative effect on the privacy of the user, and therefor would prevent many of the users to use these applications. To solve this issue, the location of the user can be encrypted at all times, except when the application detects an unusual situation to report, or when the user agrees to share their location. Another disadvantage of applications based on location sensor data is that they consume a lot of the phones battery power therefor smart power management systems are required to allow the application to be operating on the phone throughout the day.

Medication Intake

Medication adherence, or taking medications correctly, is defined by FDA[4] as: *"The extent to which patients take medication as prescribed by their doctors. This involves factors such as getting prescriptions filled, remembering to take medication on time, and understanding the directions*". Therefor there is a huge need for mobile applications that would serve as a reminder, dosing monitor and management systems for prescribed medications. This application could be further expanded and be deployed on an item (pill container) that could visually remind people to take their medication. Pill bottle caps that change color to indicate that a medication should be taken are now equipped with the ambient technology. Smart medicine cabinets and pill bottles that work together to track movement of bottles in and out of the cabinets to ensure medication compliance, detect potential interaction hazards between medications sharing the shelf, when a prescription

[4] http://www.fda.gov/ForConsumers/ConsumerUpdates/ucm164616.htm

needs to be refilled, or when it is nearing its expiration date, automatically notifying the local pharmacy to enable delivery are under development [Coughlin and Pope]. More creative applications use mobile phones as portals for drug safety advisory to avoid pharmaceutical accidents as proposed by Ramaswamy et al. [2014]. In their work they present a personalized, context-aware, and user-friendly drug safety advisory framework built on mobile phones to provide timely advice to users whether the medicine they intend to take is safe under current circumstances or not.

There are currently issues in Medicine Intake Management Applications that need to be addressed. First, these applications are likely only to benefit those willing to use such reminders and who are already smart phone users, unless applications that are run on a dedicated devices are developed. Second, we have to consider that users may provide incorrect or incomplete data about their state or medications and thus the system has to be able to detect such common mistakes.

Medical Status Monitoring

The benefits of mobile health applications are clearly seen in most aspects of health care. However we believe that medical state monitoring using mobile phones has had a wider impact in this domain. With the high costs of health care services many people living in developing and third world countries either do not have access to medical care or do not have enough funds to use such services. With the emerge of mobile phones and its high availability in most countries, Mobile-based Medical Status Monitoring systems can play a significant role in reducing medical costs, and providing service to patients in rural areas that would not have access to clinics otherwise. In this section we will review a few of many state-of-art medical monitoring applications that have replaced expensive devices in the hospitals and clinics.

Vital Signs Monitoring applications, refer to applications that monitor heart rate, blood pressure, oxygen saturation, body temperature and respiratory rate. Airstrip Technologies[5] have developed a patient monitoring solution which is compatible with mostly all handheld smart phones and tablets. This system is able to continuously monitor the patients vital signs. In another work, Oresko et al. [2010] developed a smart phone based cardiovascular disease detection system called "HeartToGo" by integrating Holter monitor with mobile phone. Specifically, they developed two smart phone-based wearable Cardiovascular Disease (CVD)-detection platforms capable of performing real-time ECG display, feature extraction, and beat classification. In another recent work de Greef et al. [2014] present "BiliCam", a mobile phone application to detect newborn jaundice. BiliCam uses the phone's built-in camera to take an image of the newborn's skin. After confirming that the images are usable, the system uploads the relevant portions of the images to a server, which analyzes the newborn's skin color to estimate the bilirubin level, it then recommends a course of action. This application is very low in cost compared to other current medical devices and besides the smart phone and the calibration card, this non-invasive solution requires no additional hardware.

5.2.3 Challenges and Open Problems

As we have tried to show in the previous sections of this chapter, mobile-based health and medical applications would benefit patients and physician by providing access to health

[5] http://www.airstrip.com/

Figure 5.6: BiliCam smartphone application detects jaundice in newborns. Courtesy of de Greef et al. [2014].

information especially in emergency situations. However, there are challenges that must be addressed in such systems. Advantage of mobile-based monitoring has only been proved for data capture and transferring, but data analysis on the mobile phone is still a major concern because data processing on mobile devices has serious disadvantage in terms of accuracy, delay and power/battery life [Donker et al., 2013]. Continuous data transmission by mobile devices can significantly reduce battery life. The situation would get more complex in case of low signal strength in rural areas or in case of data transmission charges. Another important challenge in this domain is the security and privacy, especially in remote monitoring systems where the patient or users data would be send to a remote sever over any type of connection. Extra caution must be made to protect patient identification and confidentiality of medical information. These issues have not been fully addressed yet and there is need for improvement in the design and structure of these systems to comply with medical and ethical standards [Baig et al., 2014]. Despite the advantages of mobile phones in patient monitoring, education, and management there are still some critical issues and challenges related to acceptability, reliability and cost that need to be addressed. In the remainder of this section we will address some of the critical challenges and open problems in mobile healthcare technology.

Interoperability

Because of the multiple clinical needs of patients at the same time, interoperability has became a critical issue for mobile health applications. Medicine [2013] define interoperability as: "Interoperability refers to those properties of systems (whether software, communications, or other systems), that enable the exchange of data among systems in common formats, the use of common protocols, and ultimately the ability to work together." Although many mobile health applications may appear simple (such as medicine or appointment reminder systems), however, they can have potential for broader applications if they have the ability to easily and accurately exchange information with other systems. This would only become possible when we create standards that govern health data concepts, patient identity, data processing protocols, and mechanisms for secure sharing of patient data that preserve confidentiality [Medicine, 2013]. This is challenging becuase many of these common standards do not yet exist, therefore with the development of open standards the lack of interoperability between systems can be solved. Closed standards create a knowledge barrier for system developers in low- and middleincome countries [Estrin and Sim, 2010]. Without access to the open standards, it would be very unlikely to understand how the mobile health systems work and to develop the technical capacity. Figure 5 shows the narrow waist of the open hourglass that will include health-specific syntactic and semantic data standards; patient identity standards; core data processing functions such as feature extraction and analytics; and data stores that allow for selective, patient-controlled sharing.

Energy Consumption

It is very important to have a low energy consumption device for battery operated systems, such as mobile phones. Therefor it is critical to reduce energy consumption for achieving a longer battery life. In order for the mobile phone to send large amounts of raw data captured from the phone sensors to remote servers for processing and analyzing, a considerable amount of energy is required that is supplied by the phones battery. Some applications need to send continuous data to servers, activity recognition applications may be in this category. Another example is blood pressure measurement, applications which would need to transmit blood pressure measurements of the user to a server every 10 minutes, this means that every 10 minutes the application requires 35 mA/h for data transmission. Because of the need of sending high quality of data in real-time to multiple devices long term use of such systems can pose a serious threat to mobile device's battery life and seriously compromise the transmission of essential data [Ren et al., 2010]. There is a need for mobile applications and sensors that would be low-power and low-energy consumption which can be used in long time monitoring and gives more battery life. There has been a large amount of research carried out to develop such sensors and algorithms that would reduce the energy consumption [Bayilmis and Younis, 2010], these algorithms could be modified to be used in the mobile applications as well. Other attempts have been made to overcome the mobile power barrier such as where Pathak et al. [2012] present ePROF, a fine tuned energy profiler. Their system captures asynchronous power behavior of modern smartphone components in mapping energy activities to the responsible program entities. This research highlights the fact that most of the energy in smartphone apps is spent in I/O, and I/O events are clustered, often due to a few routines, therefor 20% to 65% energy consumption can be achieved by controlling I/O events within application.

Security & Privacy

Perhaps one of the most important aspects of Mobile Health Systems is security and privacy and unfortunately the importance of this aspect is underestimated and usually left behind. Mobile health application deals with personal information therefore there is a huge need for data protection in order to have a safe and secure system. It can be said that security in such systems is as important as safety, and because transmission of data in mHealth based application is wireless, it may result in various security threats [Baig et al., 2014]. In the recent years many researchers have focused on the security of wireless sensor networks and more specifically have addressed security issues with respect to healthcare applications. Carrion et al. [2011] have studied the need for psychological acceptability in privacy and security protection mechanisms for mobile health application users. They state that any privacy and security mechanism must be acceptable from a usability perspective, thus some improvements could be made to current privacy policies to enhance the management of users health and personal data.

Mainly security issues can be classified into two categories: system security and information security. System security includes administrative, physical and technical level security, and information security includes data encryption, data integration, authentication and freshness protection [Ameen et al., 2010]. Kargl et al. [2008] state in their paper that attacks in health monitoring can be: modification of medical data, forging of alarms on medical data, denial of service, location and activity tracking of users, physical tampering with devices and jamming attacks. In order to preserve privacy in mobile health systems it is essential to consider policymakers, certification bodies, manufacturers, public-key infrastructure, distribution and management [Baig et al., 2014].

Some of the successful steps toward more secure mobile health systems include the following work; Ren et al. [2010] use encryption and decryption to create secure mobile health applications. In another work a two tier architecture is used to secure mobile wireless-networked sensors [Mughal et al., 2010]. Finally Deng et al. [2006] build a robust and secure system using three main key elements: data protection on the device, secure authentication and data encryption. Such a secure system can be adopted for a mobile health application so it could be considered to have basic security standards and therefore get high acceptance in the general public.

Data Communication

Smart mobile devices mostly support 3G, 4G and LTE mobile networks for data transport and communication and therefore mobile devices are known to be better than other devices (i.e. laptop, desktop) in mobile health applications because they can enable doctors to get up-to-date information from the patients situation, and vice-versa, the patient can get up-to-date feedback from the physician. However, the adoption of wireless networking as the normal method in healthcare is slow and limited because of a number of reasons. In emergency situations when patient needs to communicate with health provider network communication may not be available. Although most mobile phones are able to support network connection but the infrastructure is still not available in some rural areas, hence the application would not be of any use to users who live in those areas. Communication could be both costly and energy consuming on a mobile phone, especially when the data needs to be sent to a server frequently throughout the day. One way to reduce the data being transmitted is to offload some of the data processing to the mobile phone. Today's smartphones have considerably high computing power and memory which could be compared to desktop PCs from only 10 years ago or less. Therefor many of the simple processing tasks of the users data could be done on the mobile platform. There has been a considerable amount of research done on machine learning services that are executed on mobile platform [Aradhye et al., 2013]. Another way to reduce data transmission and communication costs is to partially rely on the mobile device for data processing and transmit the portion of data that could not be processed on the mobile device, to a server. By balancing the amount of data being transferred and the processing done on the mobile platform, the battery power of the device could be used in an efficient way , while reducing the costs of data communication and relying less on remote sources. This would also lead us to applications that could be used in rural areas where there is no reliable network connections available.

Chapter 6

Conclusions

In this thesis, we present approaches for robust offline and online human activity recognition using wearable sensors. We show that human activities can be automatically detected in real-world settings by leveraging sensors embedded in practical, low-priced wearable devices such as mobile phones. The work in this thesis spans around a variety of research contributions in activity recognition using accelerometer sensors. We propose two systems, one robust offline system using a number of features in time and frequency domain, and another online real-time method which is based on time series shapelets. In total, we conducted 2 major studies and several experiments on a dataset with 77 participants, which has resulted in 3 conference publications so far [Yazdansepas et al., 2016; Niazi et al., 2017, 2016].

First, we presented an offline multi-featured approach for recognition of various everyday activities using a single tri-axial accelerometer under real-world conditions. In this system features were extracted from the combination of time and frequency domain.

Next, we adopt several feature selection methods, as well as an expert-defined feature selection method to the datasets to extract a number of the most effective features for detecting different activities. To examine the potential of the proposed system we used a variety of machine learning classifiers for evaluating recognition performance. As the primary focus of this study, we show the effect of different feature sets on each of the machine learning classifiers. We further demonstrate the impact of decreasing the size of the training set on the accuracy of the classifier by downsampling the sensor readings. We show that despite the fact that decreasing the dataset size would negatively impact the accuracy, but may be worth disregarding since the smaller datasets are more efficient in terms of space and computational time. Thus the algorithms that are trained on smaller datasets can be implemented on devices with lower computational power such as mobile phone devices. In addition we also analyze time and frequency domain features and the effect of their combination on the accuracy of the classifiers. As a secondary contribution in this study we tested our system on age based training data. We observed that training on specific age groups would be effective in increasing the accuracy of the activity recognition system, therefore training the classifiers based on the participants age would result in more accurate HAR models. Despite promising results, a limitation of this system is that it requires complex features to be extracted. This would prevent the system to be used in realtime and therefore would not have many real-world applications. Also the classifiers would need to be ran on a strong server, and therefore a mobile phone's operating system would not be capable of running such systems.

Leveraging shapelet based time series classification motivated the study of an alternative system that would be capable of detecting and classifying human activities in real-time. In the second study of this thesis, we introduce an online human activity recognition system to detect activities in real time based on a single accelerometer sensor time series. The system's training phase is offline, after training is completed, it detects human activities near real-time using time series shapelets. The proposed system is a personalized activity recognition system, therefore the system would be trained on each individual user's motion data. We train and test our proposed system on motion data from five users, while performing a variety of activities. Activities performed by the users are similar in nature, as they are all walking activities. However they are performed under different conditions with different speed, and different footwear. We have presented with extensive experiments that we find the most representative shapelet for each class, which can then provide accurate and fast classification decisions in activity time series. The results validate that it is possible to detect human activities based on time series shapelets. The results obtained were promising for the following reasons. First, they represent a baseline for practical realtime activity recognition using a device with with a single accelerometer sensor. Second, this study explored using a single sensor for activity recognition, but many other sensors could also be utilized to improve activity recognition such as blood pressure sensor, gyroscope and GPS. And third, this work suggests that it may be possible to build systems that would be capable for recognizing more complex human movements while minimizing the computation overhead and time complexity to achieve promising results in real-world conditions. Despite the promise of this method, it is worth considering that the data was gathered under controlled settings. The sensor was secured to the users hip and an student research assistant monitored the users to make sure the activities are being recorded correctly. We would expect a slightly lower performance in real-world conditions. One of the challenges of such a HAR system is that the device may not be properly attached to the body. If the device is placed in a bag or pocket, it will have additional movements that would generate noise in the recorded data. The existence of such noise will block the movement patterns in the data that correspond to the activities and decrease the system's accuracy.

6.1 Future Work

In this section, we outline opportunities and ideas that we have identified for potential future work. By expanding and improving upon the current proposed Human Activity Recognition work we can achieve a HAR system that addresses many of the limitations described above.

We presented results showing how a single hip-worn accelerometer sensor can help detect human daily activities. We showed how the x-axis data can be used successfully to detect activities. However, one direction we did not explore in our work, which represents a large opportunity to improve HAR systems, is to use multiple sensors. In other words, combining data recorded by multiple accelerometer sensors attached to different parts of the body, such as wrist, ankle and chest is likely to result in prediction accuracies that exceed those obtained with a single sensor. Although one may suggest HAR systems be designed in a simple fashion with fewer sensors, but there is also a great amount of interest in understanding how to detect activities in light of multiple streams of data. It would be promising to incorporate more sensors in HAR systems to recognize more complex human activities. There is an enormous variability in the pattern of how individuals move, despite it seeming similar. This infers that different movement patterns that differ from user to user but are unique for an individual, require building a HAR model that is personalized. We have proposed a personalized HAR system that is implemented at the level of one individual. However, we can further generalize the system by training it at the level of a group of users. This would require us to cluster users based on their body features, or by their activity patterns. Expanding this study, we observe that individuals that have similar body features typically adopt movement habits that are similar. We can therefore take advantage of this and reduce the necessary training time for each individual, by assigning them to a group of users and adopting the model of that particular group. Clearly, one of the challenges of such a semi-personalized system is acquiring a variety of data for different groups of individuals with different movement patterns and body characteristics.

Building a system for automatic human activity recognition represents significant challenges. We believe such systems would provide the foundation for a new tier of applications in health, security and many more domains. By continuing this research and overcoming the limitations of current HAR systems, we hope this work would benefit individuals and health researchers. Despite the limitations and opportunities for improvement, we believe the work outlined in this thesis provides evidence that a fast and practical HAR system based on low-costing sensors can play an important role towards achieving a variety of novel applications which were not possible in past.

Bibliography

Moshaddique Ameen, Jingwei Liu, and Kyungsup Kwak. Security and privacy issues in wireless sensor networks for healthcare applications. *Journal of Medical Systems*, 36(1):93–101, 2010. ISSN 1573-689X. doi: 10.1007/s10916-010-9449-4. URL http://dx.doi.org/10.1007/s10916-010-9449-4.

Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra, and Jorge Luis Reyes-Ortiz. Energy efficient smartphone-based activity recognition using fixed-point arithmetic. *J. UCS*, 19(9):1295–1314, 2013.

H. Aradhye, W. Hua, and R.S. Lin. Native machine learning service for user adaptation on a mobile platform, April 23 2013. URL http://www.google.com/patents/US8429103. US Patent 8,429,103.

Armstrong, Watson, Makredes, Frangos, Kimball, and Kvedar. Text-message reminders to improve sunscreen use: A randomized, controlled trial using electronic monitoring. *Archives of Dermatology*, 145(11):1230–1236, 2009. doi: 10.1001/archdermatol.2009.269. URL +http://dx.doi.org/10.1001/archdermatol.2009.269.

Mirza Mansoor Baig, Hamid GholamHosseini, and Martin J. Connolly. Mobile healthcare applications: system design review, critical issues and challenges. *Australasian Physical & Engineering Sciences in Medicine*, 38(1):23–38, 2014. ISSN 1879-5447. doi: 10.1007/s13246-014-0315-4. URL http://dx.doi.org/10.1007/s13246-014-0315-4.

Ling Bao and StephenS. Intille. Activity recognition from user-annotated acceleration data. In Alois Ferscha and Friedemann Mattern, editors, *Pervasive Computing*, volume 3001 of *Lecture Notes in Computer Science*, pages 1–17. Springer Berlin Heidelberg, 2004. ISBN 978-3-540-21835-7. doi: 10.1007/978-3-540-24646-6_1. URL http://dx.doi.org/10.1007/978-3-540-24646-6_1.

Sulaimon A Bashir, Daniel C Doolan, and Andrei Petrovski. The effect of window length on accuracy of smartphone-based activity recognition. *IAENG International Journal of Computer Science*, 23(1), 2016.

Cuneyt Bayilmis and Mohamed Younis. Energy-aware gateway selection for increasing the lifetime of wireless body area sensor networks. *Journal of Medical Systems*, 36 (3):1593–1601, 2010. ISSN 1573-689X. doi: 10.1007/s10916-010-9620-y. URL http://dx.doi.org/10.1007/s10916-010-9620-y.

A. Benmansour, A. Bouchachia, and M. Feham. Human activity recognition in pervasive single resident smart homes: State of art. In *Programming and Systems (ISPS), 2015 12th International Symposium on*, pages 1–9, April 2015. doi: 10.1109/ISPS.2015.7244997.

Martin Berchtold, Matthias Budde, Dawud Gordon, Hedda R Schmidtke, and Michael Beigl. Actiserv: Activity recognition service for mobile phones. In *Wearable Computers (ISWC), 2010 International Symposium on*, pages 1–8. IEEE, 2010.

N. Bidargaddi, A. Sarela, L. Klingbeil, and M. Karunanithi. Detecting walking activity in cardiac rehabilitation by using accelerometer. In *Intelligent Sensors, Sensor Networks and Information, 2007. ISSNIP 2007. 3rd International Conference on*, pages 555–560, Dec 2007. doi: 10.1109/ISSNIP.2007.4496903.

Robert Bodor, Bennett Jackson, and Nikolaos Papanikolopoulos. Vision-based human tracking and activity recognition. In *Proc. of the 11th Mediterranean Conf. on Control and Automation*, volume 1. Citeseer, 2003.

Maged Kamel Boulos, Ann Brewer, Chante Karimkhani, David Buller, and Robert Dellavalle. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. *Online Journal of Public Health Informatics*, 5(3), 2014. ISSN 1947-2579. URL http://128.248.156.56/ojs/index.php/ojphi/article/view/4814.

S Allen Broughton and Kurt M Bryan. *Discrete Fourier analysis and wavelets: applications to signal and image processing*. John Wiley & Sons, 2011.

Inma Carrion, Jose L. Fernandez-Aleman, and Ambrosio Toval. *Human-Computer Interaction – INTERACT 2011: 13th IFIP TC 13 International Conference, Lisbon, Portugal, September 5-9, 2011, Proceedings, Part IV*, chapter Usable Privacy and Security in Personal Health Records, pages 36–43. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011. ISBN 978-3-642-23768-3. doi: 10.1007/978-3-642-23768-3_3. URL http://dx.doi.org/10.1007/978-3-642-23768-3_3.

Carmelo Cassisi, Placido Montalto, Marco Aliotta, Andrea Cannata, and Alfredo Pulvirenti. Similarity measures and dimensionality reduction techniques for time series data mining. In *Advances in data mining knowledge discovery and applications*. InTech, 2012.

Cagatay Catal, Selin Tufekci, Elif Pirmit, and Guner Kocabag. On the use of ensemble of classifiers for accelerometer-based activity recognition. *Applied Soft Computing*, 37:1018 – 1022, 2015. ISSN 1568-4946. doi: http://dx.doi.org/10.1016/j.asoc.2015.01.025. URL `http://www.sciencedirect.com/science/article/pii/S1568494615000447`.

Tom Chau. A review of analytical techniques for gait data. part 2: neural network and wavelet methods. *Gait & Posture*, 13(2):102–120, 2001.

L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu. Sensor-based activity recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(6):790–808, Nov 2012. ISSN 1094-6977. doi: 10.1109/TSMCC.2012.2198883.

Girija Chetty, Matthew White, and Farnaz Akther. Smart phone based data mining for human activity recognition. *Procedia Computer Science*, 46:1181–1187, 2015.

S. H. Chew, P. A. Chong, E. Gunawan, K. W. Goh, Y. Kim, and C. B. Soh. A hybrid mobile-based patient location tracking system for personal healthcare applications. In *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, pages 5188–5191, Aug 2006. doi: 10.1109/IEMBS.2006.259579.

Delphine Christin, Andreas Reinhardt, Salil S Kanhere, and Matthias Hollick. A survey on privacy in mobile participatory sensing applications. *Journal of Systems and Software*, 84(11):1928–1946, 2011.

William S Cleveland. Robust locally weighted regression and smoothing scatterplots. *Journal of the American statistical association*, 74(368):829–836, 1979.

Pierre Comon. Independent component analysis, a new concept? *Signal processing*, 36 (3):287–314, 1994.

Sunny Consolvo, David W McDonald, Tammy Toscos, Mike Y Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, Louis LeGrand, Ryan Libby, et al. Activity sensing in the wild: a field trial of ubifit garden. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1797–1806. ACM, 2008.

Joseph F Coughlin and James Pope. Innovations in health, wellness, and aging-in-place.

Jiangpeng Dai, Xiaole Bai, Zhimin Yang, Zhaohui Shen, and Dong Xuan. Perfalld: A pervasive fall detection system using mobile phones. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2010 8th IEEE International Conference on*, pages 292–297, March 2010. doi: 10.1109/PERCOMW.2010.5470652.

Barnan Das, Adriana M Seelye, Brian L Thomas, Diane J Cook, Larry B Holder, and Maureen Schmitter-Edgecombe. Using smart phones for context-aware prompting in smart environments. In *Consumer Communications and Networking Conference (CCNC), 2012 IEEE*, pages 399–403. IEEE, 2012.

Lilian de Greef, Mayank Goel, Min Joon Seo, Eric C. Larson, James W. Stout, James A. Taylor, and Shwetak N. Patel. Bilicam: Using mobile phones to monitor newborn jaundice. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '14, pages 331–342, New York, NY, USA, 2014.

ACM. ISBN 978-1-4503-2968-2. doi: 10.1145/2632048.2632076. URL `http://doi.acm.org/10.1145/2632048.2632076`.

Jing Deng, Richard Han, and Shivakant Mishra. Decorrelating wireless sensor network traffic to inhibit traffic analysis attacks. *Pervasive and Mobile Computing*, 2(2):159–186, 2006.

Tushar Dobhal, Vivswan Shitole, Gabriel Thomas, and Girisha Navada. Human activity recognition using binary motion image and deep learning. *Procedia Computer Science*, 58:178 – 185, 2015. ISSN 1877-0509. doi: http://dx.doi.org/10.1016/j.procs.2015.08.050. URL `http://www.sciencedirect.com/science/article/pii/S1877050915021614`. Second International Symposium on Computer Vision and the Internet (VisionNetfh15).

Tara Donker, Katherine Petrie, Judy Proudfoot, Janine Clarke, Mary-Rose Birch, and Helen Christensen. Smartphones for smarter delivery of mental health programs: a systematic review. *Journal of medical Internet research*, 15(11):e247, 2013.

A. A. Efros, A. C. Berg, G. Mori, and J. Malik. Recognizing action at a distance. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 726–733 vol.2, Oct 2003. doi: 10.1109/ICCV.2003.1238420.

M. Ermes, J. Parkka, J. Mantyjarvi, and I. Korhonen. Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. *Trans. Info. Tech. Biomed.*, 12(1):20–26, January 2008. ISSN 1089-7771. doi: 10.1109/TITB.2007.899496. URL `http://dx.doi.org/10.1109/TITB.2007.899496`.

Deborah Estrin and Ida Sim. Open mhealth architecture: an engine for health care innovation. *Science*, 330(6005):759–760, 2010.

Tao Feng and Harry J.P. Timmermans. Transportation mode recognition using {GPS} and accelerometer data. *Transportation Research Part C: Emerging Technologies*, 37:118 – 130, 2013. ISSN 0968-090X. doi: http://dx.doi.org/10.1016/j.trc.2013.09.014. URL `http://www.sciencedirect.com/science/article/pii/S0968090X13002039`.

V. L. Franklin, A. Waller, C. Pagliari, and S. A. Greene. A randomized controlled trial of sweet talk, a text-messaging system to support young people with diabetes. *Diabetic Medicine*, 23(12):1332–1338, 2006. ISSN 1464-5491. doi: 10.1111/j.1464-5491.2006.01989.x. URL `http://dx.doi.org/10.1111/j.1464-5491.2006.01989.x`.

Mikel Galar, Alberto Fernndez, Edurne Barrenechea, Humberto Bustince, and Francisco Herrera. An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes. *Pattern Recognition*, 44(8):1761 – 1776, 2011. ISSN 0031-3203. doi: https://doi.org/10.1016/j.patcog.2011.01.017. URL `http://www.sciencedirect.com/science/article/pii/S0031320311000458`.

C. Gao, F. Kong, and J. Tan. Healthaware: Tackling obesity with health aware smart phone systems. In *Robotics and Biomimetics (ROBIO), 2009 IEEE International Conference on*, pages 1549–1554, Dec 2009. doi: 10.1109/ROBIO.2009.5420399.

Pierre Geurts. Pattern extraction for time series classification. In *PKDD*, volume 1, pages 115–127. Springer, 2001.

J. B. Gomes, S. Krishnaswamy, M. M. Gaber, P. A. C. Sousa, and E. Menasalvas. Mars: A personalised mobile activity recognition system. In *Mobile Data Management (MDM), 2012 IEEE 13th International Conference on*, pages 316–319, July 2012. doi: 10.1109/MDM.2012.33.

A. Graps. An introduction to wavelets. *Computational Science Engineering, IEEE*, 2(2): 50–61, Summer 1995. ISSN 1070-9924. doi: 10.1109/99.388960.

Clare Griffiths, Cleo Rooney, and Anita Brock. Leading causes of death in england and wales–how should we group causes. *Health Statistics Quarterly*, 28(9), 2005.

D. Guan, W. Yuan, Y. K. Lee, A. Gavrilov, and S. Lee. Activity recognition based on semi-supervised learning. In *Embedded and Real-Time Computing Systems and Applications, 2007. RTCSA 2007. 13th IEEE International Conference on*, pages 469–475, Aug 2007. doi: 10.1109/RTCSA.2007.17.

John J Guiry, Pepijn van de Ven, and John Nelson. Orientation independent human mobility monitoring with an android smartphone. In *Proceeedings of the IASTED International Conference on Assistive Technologies, Innsbruck, Austria*, pages 15–17, 2012.

A. Gupta, A. Kembhavi, and L. S. Davis. Observing human-object interactions: Using spatial and functional compatibility for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(10):1775–1789, Oct 2009. ISSN 0162-8828. doi: 10.1109/TPAMI.2009.83.

Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. The weka data mining software: An update.

Mark A. Hall. Correlation-based feature selection for discrete and numeric class machine learning. In *Proceedings of the Seventeenth International Conference on Machine Learning*, ICML '00, pages 359–366, San Francisco, CA, USA, 2000. Morgan Kaufmann Publishers Inc. ISBN 1-55860-707-2. URL http://dl.acm.org/citation.cfm?id=645529.657793.

Y. Hanai, J. Nishimura, and T. Kuroda. Haar-like filtering for human activity recognition using 3d accelerometer. In *Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop, 2009. DSP/SPE 2009. IEEE 13th*, pages 675–678, Jan 2009. doi: 10.1109/DSP.2009.4786008.

Holger Harms, Oliver Amft, Gerhard Tröster, and Daniel Roggen. Smash: A distributed sensing and processing garment for the classification of upper body postures. In *Proceedings of the ICST 3rd International Conference on Body Area Networks*, BodyNets '08, pages 22:1–22:8, ICST, Brussels, Belgium, Belgium, 2008. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering). ISBN 978-963-9799-17-2. URL http://dl.acm.org/citation.cfm?id=1460257.1460287.

Zhenyu He. Activity recognition from accelerometer signals based on wavelet-ar model. In *Progress in Informatics and Computing (PIC), 2010 IEEE International Conference on*, volume 1, pages 499–502, Dec 2010. doi: 10.1109/PIC.2010.5687572.

Danilo Hollosi, Jens Schröder, Stefan Goetze, and Jens-E Appell. Voice activity detection driven acoustic event classification for monitoring in smart homes. In *Applied Sci-*

ences in Biomedical and Communication Technologies (ISABEL), 2010 3rd International Symposium on, pages 1–5. IEEE, 2010.

Charles C Holt. Forecasting seasonals and trends by exponentially weighted moving averages. *International journal of forecasting*, 20(1):5–10, 2004.

Vadim V. Ignatov, Andrey D.and Strijov. Human activity recognition using quasiperiodic time series collected from a single tri-axial accelerometer. *Multimedia Tools and Applications*, 75(12):7257–7270, Jun 2016. ISSN 1573-7721. doi: 10.1007/s11042-015-2643-0. URL `https://doi.org/10.1007/s11042-015-2643-0`.

R. Gutierrez-Osuna J. Choi, B. Ahmed. Technical report, 2010. URL `http://www.cs.tamu.edu/academics/tr/2010-11-1.pdf`.

A. Jalal and S. Kamal. Real-time life logging via a depth silhouette-based human activity recognition system for smart home services. In *Advanced Video and Signal Based Surveillance (AVSS), 2014 11th IEEE International Conference on*, pages 74–80, Aug 2014. doi: 10.1109/AVSS.2014.6918647.

Ian T Jolliffe. *Introduction*. Springer, 2002.

F. Kargl, E. Lawrence, M. Fischer, and Y. Y. Lim. Security, privacy and legal issues in pervasive ehealth monitoring systems. In *Mobile Business, 2008. ICMB '08. 7th International Conference on*, pages 296–304, July 2008. doi: 10.1109/ICMB.2008.31.

Kay and Misha. mhealth: New horizons for health through mobile technologies. 2011.

Eamonn Keogh and Shruti Kasetty. On the need for time series data mining benchmarks: A survey and empirical demonstration. *Data Mining and Knowledge Discovery*, 7(4):

349–371, Oct 2003. ISSN 1573-756X. doi: 10.1023/A:1024988512476. URL `https://doi.org/10.1023/A:1024988512476`.

A.M. Khan, Young-Koo Lee, S.Y. Lee, and Tae-Seong Kim. A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. *Information Technology in Biomedicine, IEEE Transactions on*, 14(5):1166–1172, Sept 2010. ISSN 1089-7771. doi: 10.1109/TITB.2010.2051955.

Kenji Kira and Larry A. Rendell. A practical approach to feature selection. In *Proceedings of the Ninth International Workshop on Machine Learning*, ML92, pages 249–256, San Francisco, CA, USA, 1992. Morgan Kaufmann Publishers Inc. ISBN 1-5586-247-X. URL `http://dl.acm.org/citation.cfm?id=141975.142034`.

Predrag Klasnja and Wanda Pratt. Healthcare in the pocket: Mapping the space of mobile-phone health interventions. *Journal of Biomedical Informatics*, 45(1): 184 – 198, 2012. ISSN 1532-0464. doi: http://dx.doi.org/10.1016/j.jbi.2011.08.017. URL `http://www.sciencedirect.com/science/article/pii/S1532046411001444`.

Narayanan C. Krishnan and Diane J. Cook. Activity recognition on streaming sensor data. *Pervasive and Mobile Computing*, 10, Part B:138 – 154, 2014. ISSN 1574-1192. doi: http://dx.doi.org/10.1016/j.pmcj.2012.07.003. URL `http://www.sciencedirect.com/science/article/pii/S1574119212000776`.

O.D. Lara and M.A. Labrador. A Survey on Human Activity Recognition using Wearable Sensors. *IEEE Communications Surveys Tutorials*, 15(3):1192–1209, 2013. ISSN 1553-877X. doi: 10.1109/SURV.2012.110112.00192.

Jeffrey W Lockhart and Gary M Weiss. Limitations with activity recognition methodology and data sets. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, pages 747–756. ACM, 2014.

Codrin Lungu, Bona Kim, Mark Hallett, Kara Lukasiewicz, John Harrington, Anne Altemus, and Jeff Baldwin. ipad app for daily parkinson diary (p1. 193). *Neurology*, 84(14 Supplement):P1–193, 2015.

Lena Mamykina, Elizabeth Mynatt, Patricia Davidson, and Daniel Greenblatt. Mahi: Investigation of social scaffolding for reflective thinking in diabetes management. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 477–486, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-011-1. doi: 10.1145/1357054.1357131. URL `http://doi.acm.org/10.1145/1357054.1357131`.

S. Mashiyama, Jihoon Hong, and T. Ohtsuki. Activity recognition using low resolution infrared array sensor. In *Communications (ICC), 2015 IEEE International Conference on*, pages 495–500, June 2015. doi: 10.1109/ICC.2015.7248370.

PLOS Medicine. A reality checkpoint for mobile health: three challenges to overcome. *PLoS Med*, 10(2):e1001395, 2013.

Muhammad Mubashir, Ling Shao, and Luke Seed. A survey on fall detection: Principles and approaches. *Neurocomputing*, 100:144 – 152, 2013. ISSN 0925-2312. doi: http://dx.doi.org/10.1016/j.neucom.2011.09.037. URL `http://www.sciencedirect.com/science/article/pii/S0925231212003153`. Special issue: Behaviours in video.

Afsheen Mughal, Mohammed Kanjee, and Hong Liu. *Mobile Wireless Middleware, Operating Systems, and Applications: Third International Conference, Mobilware 2010, Chicago, IL, USA, June 30 - July 2, 2010. Revised Selected Papers*, chapter Mobile Healthcare Infrastructure with Qos and Security, pages 462–473. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. ISBN 978-3-642-17758-3. doi: 10.1007/978-3-642-17758-3_36. URL http://dx.doi.org/10.1007/978-3-642-17758-3_36.

B. Najafi, K. Aminian, A. Paraschiv-Ionescu, F. Loew, C. J. Bula, and P. Robert. Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly. *IEEE Transactions on Biomedical Engineering*, 50(6): 711–723, June 2003. ISSN 0018-9294. doi: 10.1109/TBME.2003.812189.

Lucas Hermann Negri. Peakutils, 2014. URL http://pythonhosted.org/PeakUtils. [Online; accessed ¡today¿].

Anzah H. Niazi, Delaram Yazdansepas, Jennifer L. Gay, Frederick W. Maier, Lakshmish Ramaswamy, Khaled Rasheed, and Matthew P. Buman. A hierarchical meta-classifier for human activity recognition. In *15th IEEE International Conference on Machine Learning and Applications, ICMLA 2016, Anaheim, CA, USA, December 18-20, 2016*, pages 81–86, 2016. doi: 10.1109/ICMLA.2016.0022. URL https://doi.org/10.1109/ICMLA.2016.0022.

Anzah H. Niazi, Delaram Yazdansepas, Jennifer L. Gay, Frederick W. Maier, Lakshmish Ramaswamy, Khaled Rasheed, and Matthew P. Buman. Statistical analysis of window sizes and sampling rates in human activity recognition. In *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technolo-*

gies (BIOSTEC 2017) - Volume 5: HEALTHINF, Porto, Portugal, February 21-23, 2017., pages 319–325, 2017. doi: 10.5220/0006148503190325. URL https://doi.org/10.5220/0006148503190325.

Huazhong Ning, Tony X Han, Dirk B Walther, Ming Liu, and Thomas S Huang. Hierarchical space-time model enabling efficient search for human actions. *Circuits and Systems for Video Technology, IEEE Transactions on*, 19(6):808–820, 2009.

J. J. Oresko, Z. Jin, J. Cheng, S. Huang, Y. Sun, H. Duschl, and A. C. Cheng. A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing. *IEEE Transactions on Information Technology in Biomedicine*, 14(3):734–740, May 2010. ISSN 1089-7771. doi: 10.1109/TITB.2010.2047865.

Robert J. Orr and Gregory D. Abowd. The smart floor: A mechanism for natural user identification and tracking. In *CHI '00 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '00, pages 275–276, New York, NY, USA, 2000. ACM. ISBN 1-58113-248-4. doi: 10.1145/633292.633453. URL http://doi.acm.org/10.1145/633292.633453.

A. Pantelopoulos and N. G. Bourbakis. A survey on wearable sensor-based systems for health monitoring and prognosis. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(1):1–12, Jan 2010. ISSN 1094-6977. doi: 10.1109/TSMCC.2009.2032660.

J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen. Activity classification using realistic data from wearable sensors. *IEEE Transactions on In-*

formation Technology in Biomedicine, 10(1):119–128, Jan 2006. ISSN 1089-7771. doi: 10.1109/TITB.2005.856863.

Abhinav Pathak, Y. Charlie Hu, and Ming Zhang. Where is the energy spent inside my app?: Fine grained energy accounting on smartphones with eprof. In *Proceedings of the 7th ACM European Conference on Computer Systems*, EuroSys '12, pages 29–42, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1223-3. doi: 10.1145/2168836.2168841. URL `http://doi.acm.org/10.1145/2168836.2168841`.

D. J. Patterson, D. Fox, H. Kautz, and M. Philipose. Fine-grained activity recognition by aggregating abstract object usage. In *Wearable Computers, 2005. Proceedings. Ninth IEEE International Symposium on*, pages 44–51, Oct 2005. doi: 10.1109/ISWC.2005.22.

Donald J. Patterson, Lin Liao, Dieter Fox, and Henry Kautz. *UbiComp 2003: Ubiquitous Computing: 5th International Conference, Seattle, WA, USA, October 12-15, 2003. Proceedings*, chapter Inferring High-Level Behavior from Low-Level Sensors, pages 73–89. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003. ISBN 978-3-540-39653-6. doi: 10.1007/978-3-540-39653-6_6. URL `http://dx.doi.org/10.1007/978-3-540-39653-6_6`.

Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.*, 12:2825–2830, November 2011. ISSN 1532-4435. URL `http://dl.acm.org/citation.cfm?id=1953048.2078195`.

Matthai Philipose, Kenneth P Fishkin, Mike Perkowitz, Donald J Patterson, Dieter Fox, Henry Kautz, and Dirk Hähnel. Inferring activities from interactions with objects. *Pervasive Computing, IEEE*, 3(4):50–57, 2004.

Thomas Plotz, Nils Y Hammerla, and Patrick Olivier. Feature learning for activity recognition in ubiquitous computing. 2011.

L. Ramaswamy, M. D. Scott, and S. K. Dhillon. Towards a mobile cloud-enabled personalized and context-aware drug safety advisory framework for mitigating pharmaceutical accidents. In *Healthcare Informatics (ICHI), 2014 IEEE International Conference on*, pages 220–225, Sept 2014. doi: 10.1109/ICHI.2014.38.

Chotirat Ann Ratanamahatana and Eamonn Keogh. Making time-series classification more accurate using learned constraints. In *Proceedings of the 2004 SIAM International Conference on Data Mining*, pages 11–22. SIAM, 2004.

Chotirat Ann Ratanamahatana and Eamonn Keogh. Three myths about dynamic time warping data mining. In *Proceedings of the 2005 SIAM International Conference on Data Mining*, pages 506–510. SIAM, 2005.

Yonglin Ren, Richard Werner Nelem Pazzi, and Azzedine Boukerche. Monitoring patients via a secure and mobile healthcare system. *Wireless Commun.*, 17(1):59–65, February 2010. ISSN 1536-1284. doi: 10.1109/MWC.2010.5416351. URL `http://dx.doi.org/10.1109/MWC.2010.5416351`.

Daniele Riboni and Claudio Bettini. Cosar: Hybrid reasoning for context-aware activity recognition. *Personal Ubiquitous Comput.*, 15(3):271–289, March 2011. ISSN 1617-

4909. doi: 10.1007/s00779-010-0331-7. URL http://dx.doi.org/10.1007/s00779-010-0331-7.

Ryan Rifkin and Aldebaro Klautau. In defense of one-vs-all classification. *Journal of machine learning research*, 5(Jan):101–141, 2004.

K Roy, C Chaudhuri, M Kundu, M Nasipuri, and DK Basu. Short paper_. *Journal of information science and engineering*, 21:1247–1259, 2005.

Rubn San-Segundo, Jaime Lorenzo-Trueba, Beatriz Martnez-Gonzlez, and Jos M. Pardo. Segmenting human activities based on {HMMs} using smartphone inertial sensors. *Pervasive and Mobile Computing*, pages –, 2016. ISSN 1574-1192. doi: http://dx.doi.org/10.1016/j.pmcj.2016.01.004. URL http://www.sciencedirect.com/science/article/pii/S1574119216000067.

G. Schreier, D. Hayn, P. Kastner, S. Koller, W. Salmhofer, and R. Hofmann-Wellenhof. A mobile-phone based teledermatology system to support self-management of patients suffering from psoriasis. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 5338–5341, Aug 2008. doi: 10.1109/IEMBS.2008.4650420.

Skyler Seto, Wenyu Zhang, and Yichen Zhou. Multivariate time series classification using dynamic time warping template selection for human activity recognition. In *Computational Intelligence, 2015 IEEE Symposium Series on*, pages 1399–1406. IEEE, 2015.

Diana A. Shellmer, Mary Amanda Dew, George Mazariegos, and Annette DeVito Dabbs. Development and field testing of teen pocketpath, a mobile health application to improve medication adherence in adolescent solid organ recipients. *Pediatric Transplantation*, 20(1):130–140, 2016. ISSN 1399-3046. doi: 10.1111/petr.12639. URL http://dx.doi.org/10.1111/petr.12639.

Nancy E. Sherwood, Meghan M. Senso, Claire K. Fleming, and Alison M. Roeder. Chapter 26 - behavioral risk factors for overweight and obesity: Diet and physical activity. In Ann M. CoulstonCarol J. BousheyMario G. Ferruzzi, editor, *Nutrition in the Prevention and Treatment of Disease (Third Edition)*, pages 479 – 499. Academic Press, third edition edition, 2013. ISBN 978-0-12-391884-0. doi: http://dx.doi.org/10.1016/B978-0-12-391884-0.00026-3. URL http://www.sciencedirect.com/science/article/pii/B9780123918840000263.

Y. Shi, Y. Shi, and X. Wang. Fall detection on mobile phones using features from a five-phase model. In *Ubiquitous Intelligence Computing and 9th International Conference on Autonomic Trusted Computing (UIC/ATC), 2012 9th International Conference on*, pages 951–956, Sept 2012. doi: 10.1109/UIC-ATC.2012.100.

Muhammad Shoaib, Stephan Bosch, Ozlem Durmaz Incel, Hans Scholten, and Paul J. M. Havinga. A Survey of Online Activity Recognition Using Mobile Phones. *Sensors*, 15 (1):2059–2085, January 2015. doi: 10.3390/s150102059. URL http://www.mdpi.com/1424-8220/15/1/2059.

Sandra Shultz, Peggy Houglum, and David Perrin. *Examination of Musculoskeletal Injuries with Web Resource-3rd Edition*. Human Kinetics Publishers, 3rd edition, 2009. ISBN 0736076220, 9780736076227.

Muhammad Hameed Siddiqi, Rahman Ali, Md Sohel Rana, Een-Kee Hong, Eun Soo Kim, and Sungyoung Lee. Video-based human activity recognition using multilevel wavelet decomposition and stepwise linear discriminant analysis. *Sensors*, 14(4):6370–6392, 2014.

Bernard W Silverman. Spline smoothing: the equivalent variable kernel method. *The Annals of Statistics*, pages 898–916, 1984.

Brian K. Smith, Jeana Frost, Meltem Albayrak, and Rajneesh Sudhakar. Integrating glucometers and digital photography as experience capture tools to enhance patient understanding and communication of diabetes self-management practices. *Personal and Ubiquitous Computing*, 11(4):273–286, 2006. ISSN 1617-4917. doi: 10.1007/s00779-006-0087-2. URL `http://dx.doi.org/10.1007/s00779-006-0087-2`.

Clay D Spence and Paul Sajda. Role of feature selection in building pattern recognizers for computer-aided diagnosis. In *Medical Imaging'98*, pages 1434–1441. International Society for Optics and Photonics, 1998.

Chris Stauffer and W. Eric L. Grimson. Learning patterns of activity using real-time tracking. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(8):747–757, August 2000. ISSN 0162-8828. doi: 10.1109/34.868677. URL `http://dx.doi.org/10.1109/34.868677`.

Jaeyong Sung, Colin Ponce, Bart Selman, and Ashutosh Saxena. Human activity detection from rgbd images. *plan, activity, and intent recognition*, 64, 2011.

E. M. Tapia, S. S. Intille, W. Haskell, K. Larson, J. Wright, A. King, and R. Friedman. Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. In *Wearable Computers, 2007 11th IEEE International Symposium on*, pages 37–40, Oct 2007. doi: 10.1109/ISWC.2007.4373774.

Cesar Torres-Huitzil and Andres Alvarez-Landero. *Mobile Health: A Technology Road Map*, chapter Accelerometer-Based Human Activity Recognition in Smartphones for Healthcare Services, pages 147–169. Springer International Publishing, Cham, 2015. ISBN 978-3-319-12817-7. doi: 10.1007/978-3-319-12817-7_7. URL http://dx.doi.org/10.1007/978-3-319-12817-7_7.

P. Turaga, R. Chellappa, V.S. Subrahmanian, and O. Udrea. Machine recognition of human activities: A survey. *Circuits and Systems for Video Technology, IEEE Transactions on*, 18(11):1473–1488, Nov 2008a. ISSN 1051-8215. doi: 10.1109/TCSVT.2008.2005594.

Pavan Turaga, Rama Chellappa, Venkatramana S Subrahmanian, and Octavian Udrea. Machine recognition of human activities: A survey. *Circuits and Systems for Video Technology, IEEE Transactions on*, 18(11):1473–1488, 2008b.

Yi Wang, Bhaskar Krishnamachari, Qing Zhao, and Murali Annavaram. Markov-optimal sensing policy for user state estimation in mobile devices. In *Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks*, IPSN '10, pages 268–278, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-988-6. doi: 10.1145/1791212.1791244. URL http://doi.acm.org/10.1145/1791212.1791244.

Mark Weiser. The computer for the 21st century. *SIGMOBILE Mob. Comput. Commun. Rev.*, 3(3):3–11, July 1999. ISSN 1559-1662. doi: 10.1145/329124.329126. URL `http://doi.acm.org/10.1145/329124.329126`.

Dietrich Wettschereck, David W Aha, and Takao Mohri. A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms. *Artificial Intelligence Review*, 11(1-5):273–314, 1997.

Christopher R Wren and Emmanuel Munguia Tapia. Toward scalable activity recognition for sensor networks. In *Location-and context-awareness*, pages 168–185. Springer, 2006.

Delaram Yazdansepas, Anzah H. Niazi, Jennifer L. Gay, Frederick W. Maier, Lakshmish Ramaswamy, Khaled Rasheed, and Matthew P. Buman. A multi-featured approach for wearable sensor-based human activity recognition. In *2016 IEEE International Conference on Healthcare Informatics, ICHI 2016, Chicago, IL, USA, October 4-7, 2016*, pages 423–431, 2016. doi: 10.1109/ICHI.2016.81. URL `https://doi.org/10.1109/ICHI.2016.81`.

Matei Zaharia, Reynold S. Xin, Patrick Wendell, Tathagata Das, Michael Armbrust, Ankur Dave, Xiangrui Meng, Josh Rosen, Shivaram Venkataraman, Michael J. Franklin, Ali Ghodsi, Joseph Gonzalez, Scott Shenker, and Ion Stoica. Apache spark: A unified engine for big data processing. *Commun. ACM*, 59(11):56–65, October 2016. ISSN 0001-0782. doi: 10.1145/2934664. URL `http://doi.acm.org/10.1145/2934664`.