COMPARING SENSOR MODALITIES

FOR ACTIVITY RECOGNITION

By

YASAMIN SAHAF

A dissertation submitted in partial fulfillment of
the requirements for the degree of

MASTER OF SCIENCE IN COMPUTER SCIENCE

WASHINGTON STATE UNIVERSITY
School of Electrical Engineering & Computer Science

AUGUST 2011

To the Faculty of Washington State University:

The members of the Committee appointed to examine the thesis of YASAMIN SAHAF find it satisfactory and recommend that it be accepted.

Diane J. Cook, Ph.D., Chair

Lawrence Holder, Ph.D.

Behrooz Shirazi, Ph.D.

ACKNOWLEDGEMENTS

It is a pleasure to thank those who made this thesis possible. I am heartily thankful to my supervisor Dr. Diane Cook who introduced me to the research world and her expertise, advice and patient encouragement enabled me to develop an understanding of the subject.

Special thanks to the other members of my committee, Dr. Larry Holder, and Dr. Behrooz Shirazi for their useful comments on this thesis.

I am indebted to all of my colleagues and in particular I must acknowledge Parisa Rashidi and Narayanan Krishnan for their advice and suggestions throughout my study.

At the end, I would like to thank my family: my mother for all her love and support and my brother Amir for his support and encouragements through my entire life.

COMPARING SENSOR MODALITIES

FOR ACTIVITY RECOGNITION

Abstract

by Yasamin Sahaf, M.S.
Washington State University
August 2011

Chair: Diane J. Cook

Activity recognition has a promising role in various applications such as health care, psychology, and security. Choosing an appropriate sensing modality to gather data is one of the most important factors in effective activity recognition. The sensory modality has an impact both on the final results and also on the degree to which users will accept the technology. In this work we evaluate different sensing technologies (Environmental, Object and Wearable) for the purpose of activity recognition in smart environments based on the type of activities being recognized.

In this thesis, we introduce different sensing technologies and discuss both the positive points and the limitations of each. We also conduct experiments in a real home setting with participants performing common activities of daily living. Alternative data features and activity recognition algorithms are tested with the goal of determining an optimal sensor class for a type of activity. We discuss results based on different sensor combinations and provide suggestions about

which sensor technology is most suitable for recognizing a particular class of activity.

In addition, this study introduces the notion of a suffix tree to adapt pattern discovery techniques to the problem of activity recognition with wearable sensors. This model is evaluated using data gathered from wearable accelerometers.

Finally, we present a formal analysis of activity complexity. By defining measurements in terms of three dimensions, sensing, computational and performance, this analysis characterizes activities in terms of a complexity measure. Moreover, we introduce grammars as a formal representation of activities and propose such grammars as an approach for measuring the complexity of an activity.

TABLE OF CONTENTS

LIST OF TABLES

Dedication

I would like to dedicate this thesis to my mother for her unconditional love and support and my father, gone now but never forgotten who enlightened me the first glance of science in early years of my life.

CHAPTER ONE

1. INTRODUCTION

1.1 Motivation

The life style of humans is moving toward automation; computers and automated devices in our environment are used more and more in order to support our day-to-day tasks. Ubiquitous computing, as an extension of home automation, tries to integrate information processing into everyday objects and activities. It is defined as "machines that fit the human environment instead of forcing humans to enter theirs" (York & Pendharkar, 2004). Nowadays, automated devices can be found in common devices in every home, such as washing machines, dryers, ovens, refrigerators, TVs, and so forth. Yet calling a building equipped with automated devices a smart home is not appropriate. "Smart" should refer to anything that is equipped with Artificial Intelligence; something that can learn and reason about its surroundings.

Winston has defined Artificial Intelligence (AI) as "The study of the computations that make it possible to perceive, reason, and act" (Winston, 1992). An intelligent agent (Russell & Norvig, 2003) is one that perceives its environment through sensors and acts on the environment through actuators. Thus, smart environments need sensors and actuators and most importantly they need to map perceptions to actions through reasoning.

Smart environments can be found in different settings and applications. One setting is a smart home, which refers to any living space with characteristics of a smart environment. The study reported in this thesis has been conducted in the Center for Advanced Studies in Adaptive Systems smart home, which is part of a multidisciplinary project at Washington State University. The home has certain overall goals, such as minimizing the cost of maintaining the home and maximizing the comfort of its inhabitants. In order to meet these goals, the house must be able to reason about and adapt to information provided by system designers and residents.

Activity recognition aims to recognize the actions and goals of one or more agents from a series of observations on the agents' actions and the environmental conditions. Activity recognition in smart homes supports many different applications such as healthcare, human computer interaction, security and sociology; therefore, since the 1980s this research field has captured the attention of several computer science communities. In the past 6 -7 years, there has been significant advancement in the area of ubiquitous, pervasive and wearable computing resulting in the development of a variety of low bandwidth, data rich environmental and body sensor networks.

While activity recognition has different applications, this study is motivated by the need for assistive technologies for individuals with mild cognitive impairment or

Alzheimer's disease and in general for elder care. This study looks at two areas not addressed in current activity recognition research and tries to close some of the gaps in this area.

One of the issues to address is the lack in earlier research of using all sensor types in one study, which prevents researchers in this field from having a fair and complete comparison between different sensing technologies. Conducting experiments using a variety of sensing technologies helps researchers to compare results from different sensor combinations and ideally find the best one for activities in their interest. *We hypothesize that using more sensor types does not necessarily provide us with the best result; however, each sensor modality performs better in recognizing a particular subset of activities.*

The other important need is providing a means for researchers in the activity recognition area to choose proper activities for their study. The lack of having a meaningful measurement for classifying different activities has led researchers to do it based on their own needs and requirements. This measurement is useful for assessing older adults, individuals with cognitive disabilities and those with chronic diseases, in order to evaluate what type of health care services an individual may need. There are many Activity of Daily Living (ADL) lists published in the Psychology domain; however, each research group has concentrated on a subset of this list according to their own needs and requirements (Liao, Fox, & Kautz, 2005; Orr & Abowd, 2000).

In summary, with each recognition system considering its own set of activities and sensor types, it is difficult to compare the performance of these different systems and more importantly it makes the task of selecting an appropriate set of technologies and tools for recognizing an activity challenging.

1.2 Contribution of this Thesis

The main contribution of this thesis is to provide a comparative study on different sensor modalities for activity recognition. Experiments using different sensor types are conducted with multiple residents in the CASAS smart home. Different combinations of sensing technologies are tested and the results are evaluated based on each set. In this study we mainly focus on ADLs and iADLs and experiments are conducted with a number of these activities that are most common.

The sensing platform used in this thesis can be classified into three main categories: Environmental, Object and Wearable. Environmental sensors consist of infra-red motion detectors which are used for localization of an individual in the house, magnetic door sensors, ambient temperature sensors, stove burner sensors, water flow sensors, and whole-house power meters. Object sensors include item presence sensors and shake sensors and wearable sensors consist of accelerometers along with gyro sensors that participants wear during the experiments.

Having this almost complete sensing platform provides us with a means to compare activity recognition in different sensing technologies. Then we will be able to study the relationship between different activities and sensor types. We hypothesize that each activity class can be recognized most accurately with one particular class of sensor types.

In order to achieve this goal we add wearable and object sensors to the existing sensing platform in the CASAS smart home which was up to this point equipped with only environmental sensors. Previous studies with environmental sensors have been very successful in recognizing many activities, but still have trouble with distinguishing some classes of activities. In particular, activities which take place in the same area but perform different functions can be difficult to distinguish, such as washing hands and washing the dishes. In these cases the need for other sensing technologies such as wearable and object sensors becomes apparent.

In general, the more sensors you use, the higher accuracy you may possibly achieve, but there are some drawbacks. As discussed by Bao and Intille (Bao & Intille, 2004), using more sensors requires more computational power. More sensors are harder to deploy, they do not integrate well into the daily clothing and they can be expensive. In addition, if sensor values are fed directly as features into a machine learning model then more training data will be needed to learn the complex model. These problems are compounded in the context of technology-assisted living, since the target community consists of older individuals and persons with disabilities. As a result, our goal is to minimize the number of small and lightweight sensors used for activity recognition as well as to minimize the impact on the daily lives of individuals who use the technologies.

A second contribution of this study is an analysis of the complexity of activities and a formal representation of activities. This study attempts to characterize activities in terms of a complexity measure. We define activity complexity along three dimensions – sensing, computation and performance and illustrate different parameters that encompass these dimensions. We look at grammars for representing activities and use grammar complexity as a measurement for activity complexity. We then describe how these measurements can help evaluate the complexity of activities of daily living that are commonly considered by researchers.

CHAPTER TWO

2. ACTIVITY RECOGNITION

2.1 Overview

Activity recognition is a widely researched area with applications in health care, security and other domains which provides context-aware services. At one end of spectrum are certain activities such as ambulation that can be defined by movements or posture, while at the other end are activities that are defined by interacting with objects such as ironing and working on a computer. Activity recognition systems for detecting this complex subset of physical activities require a variety of sensing and processing methodologies. A single recognition system will not be sufficient for recognizing this variety of activities.

One major consumer of this technology would be elderly and disabled people. Activity recognition plays an important role in developing assistive technologies for aiding older adults in leading an independent life. The post-World War II baby boom has had an effect on many western countries including the United States. Baby boomers are now late middle age and are entering the senior years. In the economy, many are now retiring and leaving the labor force. According to the US Census Bureau projections, the relative size of the older population in 2015 will reach 14% as compared to its current level of 12.4%. In 2030, older adults are projected to comprise almost 20% of the total population (He, Sengupta, Velkof, & DeBarros, 2006).

The increasing health care and nursing costs place a tremendous stress on the society and the government. Assisting older adults to stay at their own homes is financially and emotionally beneficial. This is called "aging in place" and has been of significant interest recently. Smart homes help older adults and people with disabilities to stay at home and maintain their normal life without the need to stay in healthcare facilities. The aging population has also generated significant interest by the government as well as industry leaders to develop home automation systems for the elderly (Hawes, Phillips, Rose, Holan, & Sherman, 2003). The number of research projects in this field is still growing.

Many of these older adults will suffer from cognitive diseases such as dementia. According to gerontologists, identifying changes in everyday behavior such as sleeping, food preparation, housekeeping, entertainment, and exercise is often more valuable than biometric information for the early detection of emerging physical and mental health problems - particularly for the elderly.

Smart home technologies are able to help recognize early symptoms of dementia and other chronic conditions by detecting changes in patterns of behavior of residents. Medical professionals believe that one of the best ways to detect an emerging medical condition before it becomes critical is to look for changes in the activities of daily living (ADLs), instrumental ADLs (IADLs) (Lawton & Brody, 1969), and enhanced ADLs (EADLs) (Rogers, Meyer, Walker, & Fisk, 1998). Recognizing resident activities helps in understanding difficulties faced by the individuals in completing the activities or recognizing incomplete activities such as taking medication.

The other application of activity recognition in smart environments is designing new architectural tools which can assist normal people in day to day tasks. They can be used in task automation such as lighting and HVAC control. Saving energy is another useful feature of a smart home (Chen, Das, & Cook, 2010). Security systems that are able to create a model of people's activities and behavior over time could predict intent and motive as people interact with the environment.

2.2 Related Work

The significant potentials of automatic activity recognition have been realized for decades in the computer science communities. Since the 1980s researchers have been designing activity recognition techniques with increasing accuracy. However they have not been successful in evaluating these approaches in real applications until recent years. Even today, automated activity recognition systems have limited capabilities. They either focus on very basic activities or very unique activities in specific environments. It is very challenging to build an unintrusive sensing platform to collect the data from real users, annotate the data with ground truth labels for training and testing purposes, and design efficient algorithms to learn and accurately recognize activities. . Current approaches in the literature differ according to the type of sensor that is used for data collecting and the machine learning algorithm that is designed to model activities. In the following sections we present a literature survey of different technologies and algorithmic models and discuss challenges that are facing activity recognition methods in more detail.

2.2.1 Sensing Modalities

Here we discuss different sensor modalities which are available and commonly used by activity recognition researchers. These modalities mainly consist of environmental, object and wearable sensors.

A) Environmental Sensors

An environmental sensor in an activity recognition context is a passive sensor that are integrated into the environment itself. Since objects can be seen as part of the environment, some researchers consider object sensors in this category. However, we tend to have more fined grained classification of sensor types, thus we dedicate a separate category for object sensors.

The most common environmental sensors used in activity recognition are motion detectors (section 3.1.1). These sensors can be used to recognize room occupancy; they need to be modified if they are needed for capturing the location of the resident in one room.

The other common used environmental sensor is camera which has its own benefits and limitations. In the following subsections we present previous studies with the most common environmental sensing modalities used in activity recognition.

Robotic devices. Bennewitz et al. (Bennewitz, Burgard, & Thrun, 2002) proposes an algorithm that learns collections of typical trajectories that characterize a person's motion patterns. Data recorded by mobile robots equipped with laser range finders is clustered into different types of motion using the popular expectation maximization algorithm, while simultaneously learning multiple motion patterns. Experimental results obtained using data collected in a domestic residence and in an office building, illustrate that highly predictive models of human motion patterns can be learned. This robot can be seen in Figure 1.

Figure 1. Pioneer I robot used to record the data (left) and Person moving in the environment (right). Figure has been adapted from (Bennewitz, Burgard, & Thrun, 2002)

Combinations of audio and video. Oliver et al. (Oliver, Horvitz, & Garg, 2002) have described the use of the representation in a system that diagnoses states of a user's activity based on real-time streams of evidence from video, acoustic, and computer interactions in an office environment.

Vision based Systems. In the computer vision community there is considerable work on behavior recognition using probabilistic models, but it usually focuses on recognizing simple low-level behaviors in controlled environments (Jebara & Pentland, 1999). Recognizing complex, high level activities using machine vision has only been achieved by carefully engineering domain-specific solutions, such as for hand-washing (Mihailidis, Carmichael, & Boger, 2004; Hoey, Bertoldi, Poupart, & Mihailidis, 2007), or operating a home medical appliance (Shi, Huang, Minnen, Bobick, & Essa, 2004) or very basic interactions such as punch, hug, shake hands (Park, Park, & Aggarwal, 2004). Messing et al. (Messing, Pal, & Kautz, 2009) propose an approach for recognizing activities in a kitchen using velocity profiles from videos. Essa et al. (Essa, Yin, Criminisi, & Winn, 2010) have built vision-based sensors to track multiple individuals in their smart home called "The Aware Home". An extensive survey of vision based approaches for activity/gesture recognition can be found in (Gavrila, 1999).

Researchers at Georgia Institute of Technology have built different laboratories such as Classroom, Wearable Computing Project, Aware Home, and Augmented Offices. In their Aware Home, they have vision-based sensors to track multiple individuals and they try to use similar signal processing techniques to build a smart floor interface that can identify and track people walking across a large area. There are many compelling applications for these sensing technologies throughout a home, such as support for the elderly or finding lost objects, or in specialized spaces within the home, such as the front door or the kitchen (Patel, Kientz, Jones, Price, Mynatt, & Abowd, 2007).

The most important limitation of camera based sensors is intruding upon the privacy of the individual and hence is not an appropriate technology for monitoring patients due to privacy concerns by residents. Being costly and need of large space for recording video streams is the other disadvantage of this kind of sensor. Moreover, computer vision based techniques for extracting motion and object information from images and videos are expensive, cumbersome and inaccurate due to changes in poses and illustration. Finally, designers of vision-based activity recognition have to deal with difficult issues such as selecting locations for the camera, tracking residents, and compensating for occlusion and lighting conditions.

Eye tracking sensor. This is a very recent technology that has been used for activity recognition. It was first demonstrated by Bulling et al (Bulling, Ward, Gellersen, & Troster, 2009). It uses the principle of electro-iconography to track eye movements. The hypothesis is that eye movements of an individual varies depending on the activity that is being performed. In their work they have attempted at using this technology for recognizing activities such as working on a computer, reading, and writing.

Location Based Systems. One of the most important factors in activity recognition is recognizing the location of an individual. Many activities can be recognized only from knowledge about the location.

Satellite-based location systems. Currently GPS is the most widely used location based systems. The Global Positioning System (GPS) is a space-based global navigation satellite system (GNSS) that provides location and time information in all weather, anywhere on or near the Earth, where there is an unobstructed line of sight to four or more GPS satellites. It is maintained by the United States government and is freely accessible by anyone with a GPS receiver.

Other systems are Galileo, a global system being developed by the European Union, planned to be operational by 2014. However, the important limitation of satellite-based systems is a noticeable increase in error when receivers are indoor or close to tall buildings. That's why it is not practical for ADL recognition in smart homes.

WiFi-based location systems. The other widely used location based systems are Wifi-based ones which don't have satellite-based system's limitations. Another positive point about WiFi is that most notebook computers, PDAs and some mobile phones today are equipped with WiFi devices, so it is widely accessible. One example of an indoor location based system based on WiFi is RADAR (Bahl & Padmanabhan, 2000).

Triangulation can then be used with WiFi to determine the location of a person of object. However, signal strength is influenced by a number of factors other than distance, including obstacles, reflection, and refraction, so in practice it is virtually impossible to obtain an accurate propagation model.

Mobile-phone-based location systems. These systems have the benefit of working indoors and outdoors. They have two different architectures: station-based and client-based. Client-based systems are cheaper and give end users more controls, but the handsets have to maintain the location databases and support complex programming interface.

Pressure Sensors. These sensors measure changes in the pressure exerted on them. They are primarily force-sensing resistors that decrease resistance with increasing force. Multiple pressure sensors are typically used in combination for sensing pressure changes across a defined area. A well-known device that uses pressure sensing is the WiiFit balance board. For activity recognition, pressure sensors can be integrated into the floors and carpets in an environment to determine the location of an individual as discussed by Orr and Abowd (Orr & Abowd, 2000) and Richardson et al. (Richardson, Leydon, Fernstrˆm, & Paradiso, 2004). The resolution of the pressure-sensing units within a certain area determines the reliability of the mechanism for determining the location of an individual. Processing the data stream from these sensors to track the movement of an individual is considered to be a hard problem. With further advances in the hardware, this type of sensor will be an effective device for estimating the location of an individual. Moreover, pressure sensors can be integrated into objects of everyday use such as a bed, and a chair. Integrating them into a bed helps in assessing the sleep quality of individuals. Two examples of pressure sensors can be seen in Figure 2.

Figure 2. Mutlu et al. (Mutlu, Krause, Forlizzi, Guestrin, & Hodgins, 2007) have instrumented a chair with pressure sensors, in order to enable seating posture recognition in left. Honeywell's Trustability pressure sensor in right.

Passive Infrared Sensor. A passive infrared sensor (PIR) is an electric device that uses infrared rays to sense motion when heat changes in its field of view. These are commonly used in location-based systems. An example of infra-red motion detectors is shown in Figure 3. A network of PIRs embedded in an environment can be used to track the movement of people and identify their locations. Thus a PIR sensor provides an unobtrusive mechanism for sensing the location of an individual, thereby aiding in recognizing the activity at a high level.

Motion detectors are widely used in smart environments because they are cheap, easy to install, computationally inexpensive, require minimal maintenance and supervision, and do not have to be worn or carried (Singla, Cook, & Schmitter-Edgecombe, 2008; Wilson, 2005). While these sensors are easy to install and provide highly specific data to support activity recognition, they are difficult to use when there are multiple people in the environment. Crandall has studied this challenge in detail (Crandall, 2011). Furthermore, these sensors alone do not provide the fine level of information that is required to track the progress of activities.

Wilson (Wilson, 2005) has used binary sensors including wireless X10 Hawkeye motion detectors to perform location tracking. In his research, he has looked at two main problems: tracking individuals and activity recognition; the latter is in the interest of this study. The results strongly indicate that knowledge of location is a key to activity recognition. In almost every activity they saw a statistically significant increase in accuracy as they move from no location information up to perfect location information. Wilson realized that while using environmental and object sensors, more sensors will increase accuracy, regardless of the number of occupants. A low sensor density contributes to significant periods of time between readings (especially with only one occupant). During these "quiet" times no new information arrives to help the tracker recover from mistakes (such as the lag between entering a new room and triggering a sensor). Motion detectors are the most active sensors, and a lack of them hurts accuracy the most.

Singla et al. in (Singla, Cook, & Schmitter-Edgecombe, 2008) have used environmental motion detectors for activity recognition in CASAS smart home. They were able to classify 8 common ADLs. To overcome limitations of motion detectors they considered temporal features as well. Temporal features helped them to distinguish between activities that take place in the same location thus improved accuracy. Kim et al. (Kim, K. N. Ha, Lee, & Lee, 2009) uses the pyroelectric infrared sensor-based indoor location-aware system and presents an enhanced location-recognition algorithm using a Bayesian classifier for activity recognition.

The Adaptive House at University of Colorado is a real house, with environmental sensors such as temperature, ambient light levels, sound, motion, door and window openings, and actuators to control the furnace, space heaters, water heater, lighting units, and ceiling fans. Control systems in the residence are based on neural network reinforcement learning and prediction techniques (Mozer, Pashler, Wilder, Lindsey, Jones, & Jones, 2010).

Some of the features of the Adaptive House consist of predicting when the occupants will return home and determining when to start heating the house so that a comfortable temperature is reached by the time the occupants arrive; detecting statistical patterns of water usage, such that hot water is seldom if ever used in the middle of the day on weekdays, allowing the water heater to shut off at those times; inferring where the occupant is and in what activities the occupant is engaged - perhaps he is reading at the kitchen table - and controlling lighting patterns and intensities accordingly, even anticipating which rooms are about to be entered and turning on the lights before the room becomes occupied.

Working with these motion sensors can be challenging, particularly when they are used for exact location detection as opposed to presence detection sensors. Some modifications need to be made to the sensor in order to narrow the view angle. Different applications need different settings of how long they need to remain on after being triggered. As a result, accurate calculation is needed based on the sensor type and application.

Figure 3. Examples of common infrared motion detectors.

B) Object Sensors

In the activity recognition area, any sensor that provides us with information about the objects that an individual uses or manipulates can be considered as an object sensor. An effective way to understand what someone is doing is to collect and analyze information about the objects with which they are interacting (Hodges, Newman, & Pollack, 2009). There are a variety of sensor types that can be used as object sensors; including binary contact switches, shake sensors and RFID tags.

In an activity recognition study, Hodges and Pollack (Hodges, Newman, & Pollack, 2009) exploited the regularities between object usage and activity performance in a different way. Given data about object use collected during the performance of a known activity, they inferred the identity of the person performing the action, calling the regularities between a person's interaction patterns and their performance of an activity their "object-use fingerprint." They showed that even with very simple machine-learning techniques, they could identify subjects about three-quarters of the time, a rate of success that was well above chance.

Using object sensors is mainly done using three technologies: RFID tags, shake (movement detection) sensors and vision based techniques with the use of cameras. These sensors are discussed in more detail in the following subsections.

1) Radio Frequency Identification (RFID) tags

RFID tags can be used both for location detection (Wilson, 2005) and object interaction (Hodges, Newman, & Pollack, 2009). Studies using these tags collect information about object use by utilizing a glove or bracelet outfitted with a RFID antenna. Notably, RFID antennae are able to discriminate among specific instances of objects that are otherwise the same (e.g., two spoons), with a 0% false positive rate (Want, 2003).

Figure 4. This figure illustrates an example of RFID glove used in Patterson et al. study. (Patterson, Fox, Kautz, & Philipose, 2005)

Newman and Pollack (Hodges, Newman, & Pollack, 2009) have used RFID tags for objects in the kitchen. They did experiments with Traumatic brain injury (TBI) patients participants. A learning algorithm, C4.5, has been used with features such as: detect (if the sensor is triggered), count (amount of time interaction occurred with an entity), average duration and order. Using the full feature set obtains the best accuracy of 71%.

Activities under study can have shared objects which make it harder for the classifier to distinguish them. However, many studies have looked at activities which don't have shared objects (Philipose, Fishkin, Perkowitz, Fox, Kautz, & Hahnel, 2004). In these cases, inference engines can simply be used to recognize performed activities. Only a few researches have looked at activities with shared objects. One is (Patterson, Fox, Kautz, & Philipose, 2005) which examines the advantages and challenges of reasoning with globally unique object instances detected by an RFID glove. What distinguishes that work from most activity recognition studies is that selected activities have shared objects and the study includes interleaved activities, in contrast with the typical approach of examining only segmented, sequential activities.

In general, using RFID tags for activity recognition has the following drawbacks. First is lack of accuracy in detecting objects with little number of tags; sometimes the tag is on one side of the object and participant holds it from the other side. This prevents the glove or bracelet from firing the tag. The other problem is when a wrong tag is fired because the participant is close to the tag but is not really using the object. Second, participant has to wear RFID detector in the form of glove or bracelet which is usually uncomfortable and impractical. Moreover, it is not possible to put tags on many objects such as metallic. Finally, the theory of relying only on RFID tags leads to too many sensors on objects and as a result a lot of data needs to be annotated. Data annotation is one of the barriers in current activity recognition techniques due to requiring considerable amount of time and energy.

Patterson et al. in (Patterson, Fox, Kautz, & Philipose, 2005) have demonstrated an example in their study which shows some of these flaws: In this example, the participant turns out a light and goes to the kitchen, where he opens the cup cabinet with his right hand (wearing the RFID bracelet), but reaches in with the left hand. The tags under the shelves usually fire when the bracelet reaches in, but the participant used the "wrong" hand to grab his cup. Cups don't have tags because of the microwave. He puts the cup on the counter and opens the refrigerator with his right hand. No tags are on the front of the refrigerator because they did not work due to the metal surface. He reaches in with his right hand and a tag on one of the shelves fires. He grabs a bottle, which is untagged because it was recently purchased, and puts it on counter next to the cup. He leaves the refrigerator door open and walks out of kitchen into the hallway to speak to his spouse. He comes back and closes the refrigerator with his right hand and then walks to the living room to get a key chain that has a bottle opener on it. He reaches down to the table with his right hand, at which time a tag for another object on the table might have fired if he were just a few centimeters closer. He returns to the kitchen, opens the bottle and pours a glass. He takes the bottle to the untagged metal sink and rinses it several times holding the bottle in his right hand, without using the tagged soap. He takes a drink and then puts the glass down and carries the bottle down the hall to the recycling area. A tag could fire at the recycle bin, but the area is large and even with 2 tags nearby, his hand does not get sufficiently close. He walks back to the living room and starts cleaning up, leaving the full cup in.

Because of the issues mentioned above, even with having a one-to-one map between objects and activities, we don't usually see 100% accuracy using RFID technology alone.

2) Object Movement Detection Sensors

Although most studies interested in monitoring object usage have used RFID tags as their sensing technology, other technologies have been used. These studies do not detail the same problems encountered with RFID, but they still have their own drawbacks. Reed switches (Figure 5) or binary on/off sensors are primarily electric switches that operate by a magnetic field. These sensors typically consist of two surfaces that turn on the sensor on contact or vice versa. These components can be taped to objects in the environment for studies lasting up to several weeks. In the activity recognition context, these sensors are often used for determining the state of an object. For example, placing sensor on a door joint will facilitate detection of door being closed or opened. They can be placed near the stove knobs to detect if a stove is on or off, or even to determine if a light switch is turned on.

Tapia, et al. (Tapia, Intille, & Larson, 2004) looks at recognizing activities for medical applications such as toileting, bathing, and grooming with detection accuracies ranging from 25% to 89% depending on the evaluation criteria used. The study assumes all activities are sequential and only the primary activity is considered while a person is multitasking. The researchers define an algorithm which is able to merge and capture temporal relationships from data coming from multiple sensors efficiently. Two subjects perform activities off and on for a total study duration of two weeks. Object sensors are installed on many objects that are manipulated by the occupant, including light switches, doors, windows, cabinets, drawers, microwave ovens, refrigerators, stoves, sinks and kitchen containers

Kasteren et al. in (Kasteren, Noulas, Englebienne, & Krose, 2008) conducted a 28 day experiment using sensors on doors, cupboards, refrigerator and a toilet flush. They have compared two common probabilistic algorithms in activity recognition: HMM (Hidden Markov Model) and CRF (Conditional Random Fields) on different activities.

Akin to the PIR sensors, reed switches are inexpensive and easy to install. The sensors output binary streams that typically do not require any additional processing for extracting information related to activities. Reed switches are more versatile than PIR sensors, because that they can be used in many different ways to obtain activity information. For example, when attached to a door, these sensors can provide the location information of the individual, when it is attached to an object the sensor can act as an object recognizer.

However, there are certain limitations to the use of this technology for activity recognition. As discussed with PIR sensors, reed switches do not facilitate tracking of multiple resident activities in an environment. They also do not provide the fine level of activity information that is required to track the progress of an activity to its completion. In Tapia's work (Tapia E. M., 2008), one subject reported that she was able to hear when the sensors were activated (magnetic reed sensor closing). This is one of the reasons why Tapia recommended to use an accelerometer instead of the external magnetic reed sensor in future systems. By using an accelerometer as an object sensor, any movement of the object can be detected. These sensors are sometimes called shake sensors (Figure 6).

In his dissertation, Tapia (Tapia E. M., 2008) suggests that sensors might be built into architectural components and into furniture at the time of manufacture. However, we believe in order to have a practical smart home we need to have a sensing system that can be deployed in any house with their current furniture. Clearly, changing all furniture would not be a desirable approach for people who want to use this technology in their home or healthcare facilities.

Figure 5. Reed switches attached to different objects. The picture has been adapted from the work of Tapia et al. (Tapia E. M., 2008).

Figure 6. Examples of shake sensors that can be used for object movement detection.

3) Object Sensors Summary

Object sensors come in different types and shapes; each has its own positive and negative points. These sensors are more useful in detecting iADLs. In general, a number of limitations for using object sensors are as follows:

- Some activities (e.g., walking) do not involve interactions with objects.
- Many activities (e.g., dishwashing) involve objects which cannot use a sensor or tag.

- Placing sensors or tags on all objects is impractical. There are too many objects in a home, new objects are added every day and some objects are disposable (e.g., water bottles)
- Many activities have shared objects which make them hard to be distinguished.

C) Wearable Sensors

Studies using environmental and object sensors have been very successful in recognizing a number of activities, but still have trouble with distinguishing some classes of activities. In particular, activities that take place in the same area but have different body gestures, such as reading and eating at the dining table, can be difficult to distinguish. In these cases the need for wearable sensors becomes apparent. There has been a considerable amount of research in activity recognition using wearable sensors. Advances in miniaturization are leading to sensors being embedded within wrist bands, bracelets, adhesive patches, and belts and they can even send data to a mobile computing device wirelessly.

Accelerometers and gyros are the most commonly used wearable sensors in activity recognition because they provide motion information about different parts of the body such as hand, leg and hip. Other wearable sensors that are commonly used include microphones, GPS and light sensors. These provide additional information about the environment or location of a user that might help in the task of activity recognition.

Most previous works with wearable sensors have focused on basic gestures and low-level activities such as ambulation (Ermes, Prkka, Mantyjarvi, & Korhonen, 2008), running, sitting etc (Bao & Intille, 2004). The research literature demonstrates that forms of locomotion such as walking, running, and climbing stairs and postures such as sitting, standing, and lying down can be recognized at 83% to 95% accuracy rates using hip, thigh, and ankle acceleration (Mantyjarvi, Himberg, & Seppanen). Acceleration data of the wrist and arm are known to improve recognition rates of upper body activities (Chambers, Venkatesh, West, & Bui, 2002; Foerster, Smeja, & Fahrenberg, 1999).

Most previous works with multiple accelerometers have used accelerometers connected with wires, which may restrict subject movement. One of the few works to investigate performance of recognition algorithms with multiple, wire-free accelerometers on different subjects is done by Bao and Intille (Bao & Intille, 2004). Data is collected from five biaxial accelerometers placed on 20 subjects but has been collected under laboratory and semi-naturalistic conditions. Sensors are placed on each subject's right hip, dominant wrist, non-dominant upper arm, dominant ankle, and non-dominant thigh to recognize ambulation, posture, and other everyday activities. Mean, energy, frequency-domain entropy, and correlation features are extracted from the collected data. A decision tree classifier performed best among decision table, instance-based learning (IBL or nearest neighbor), C4.5 decision tree, and naïve Bayes classifiers. In their study, user-specific training resulted in an increase in recognition accuracy of 4.32% over recognition rates for leave-one-subject-out-training. This difference shows that as we expect given equal amounts of training data, training on user-specific training data can result in classifiers that recognize activities more accurately than classifiers trained on example data from many people.

Bao and Intille (Bao & Intille, 2004) claim that since activity recognition system need to be trained only once before deployment, the slow running time for decision tree training is not an obstacle. Nonetheless, there may be limitations to a pre-trained algorithm. Although activities such as "running" or "walking" may be accurately recognized, activities that are more dependent upon individual variation and the environment (e.g. "stretching") may require person-specific training.

Lester et al. (Lester, Choudhury, & Borriello, 2006) have developed a personal activity recognition system that is practical and reliable. They claim that their system has the following characteristics. Data only from a single body location is needed, and it is not required to be from the same point for every user. It should work out of the box across individuals, with personalization only enhancing its recognition abilities. It should be effective even with a cost-sensitive subset of the sensors and data features.

They have tried to detect basic activities such as walking, standing, walking up/down stairs, riding up/down elevator and brushing teeth. In their first experiment they used accelerometer alone. They tested their sensor on different parts of the body: wrist, waist and shoulder. They also conducted an experiment having sensors on all three locations. Interestingly, single body position performed better than all locations and the best result is from place a sensor on the wrist with an overall resulting accuracy of 45.81% and recall of 45.10%. In contrast, placing accelerometers on all three positions resulted in 41.15% accuracy and 38.96% recall.

In another experiment Lester et al. (Lester, Choudhury, & Borriello, 2006) added two other sensors (microphone and barometric pressure) to the accelerometer and resulted in overall 78.18% accuracy and 87.05% recall. On the other hand, having sensors on all locations resulted in 82.07% overall accuracy and 81.55% recall. In both experiments HMM has been used as the classifier. This increase in accuracy demonstrates limitations of using accelerometers alone for activity recognition. Although this only examined basic activities, results show how important environmental information can be in classifying activities. Figure 7 shows the sensor platform used for Lester et al. study (Lester, Choudhury, & Borriello, 2006).

Figure 7. Sensor platform for work of Lester et al.
(Lester, Choudhury, & Borriello, 2006).

One important factor while using wearable sensors is choosing the body position where the sensor is worn. The common body positions used in the literature are ankle, thigh, hip, waist, wrist and upper arm (Bao & Intille, 2004; Krishnan & Panchanathan, 2008; Kern, Schiele, & Schmidt, 2003; Lester, Choudhury, & Borriello, 2006). There have been a few studies on comparing different body positions for accelerometers, Bao and Intille in (Bao & Intille, 2004) evaluate the discriminatory power of each accelerometer location, recognition accuracy using the decision tree classifier (the best performing algorithm in their study). Their results show that the accelerometer placed on the subject's thigh is the most powerful for recognizing their set of 20 activities. Acceleration of the dominant wrist was more useful in discriminating these activities than acceleration of the non-dominant arm. Acceleration of the hip is the second best location for activity discrimination.

One bottleneck that is currently holding back assisted technology from being widely used is that many residents don't feel comfortable in current smart environments. For example, most studies on wearable sensors have used multiple sensors on different body positions (Lester, Choudhury, & Borriello, 2006), which would increase the dissatisfaction among users and therefore prevent the technology to become practical. On the other hand, using only environmental sensors is not sufficient for recognizing some of the important Activities of Daily Living (ADL).

The more sensors you use the higher accuracy you will get but there are some drawbacks. As discussed by Bao and Intille (Bao & Intille, 2004), using more sensors requires more computational power. More sensors are harder to deploy, they do not integrate well into the daily clothing and finally is expensive. These problems are compounded in the context of technology assisted living, since the target community consists of older individuals and persons with disabilities.

In general, sensors used in assisted livings and smart homes are required to be easy to use, have long battery life, beautiful design and many more to be accepted by residents. In particular, wearable sensors need more attention in this regard, because residents need to wear them most of the time. Having an uncomfortable and hard to use wearable sensor will discourage residents to wear them and subsequently results won't be as expected. That's why even with all current advances in this technology, a lot of research still needs to be carried out for easy adaptation of these sensors for activity recognition in real-life settings.

One main challenge in designing wearable sensors is the excessive power consumption of the sensor due to wireless signal transmission. The battery life of the sensor can thus be prohibitively short. Wearable sensors are attached to people who move constantly most of the times so sensors need to transmit data almost constantly.

The most common type of wearable sensor for activity recognition is an accelerometer. It provides a unique and clean mechanism for capturing movements, but it is more often used to recognize basic movements of body parts (Ermes, Prkka, Mantyjarvi, & Korhonen, 2008). There are many practical and usability issues that still need to be addressed by the research community to facilitate wide spread adoption of the modality.

Wearable sensors need perfect alignment and positioning of the sensor on body. In practical everyday use a fixed sensor position in relation to the body cannot be guaranteed. Even in the scenario where the sensor is attached directly to the body, the variations in the physical form of the person such as height, weight, and body mass index, in addition to age factors can result in different sensor outputs.

That's why finding a comfortable and practical garment for wearable sensors is still a challenging research.

Other limitations of using wearable sensors for activity recognition, in particular for health monitoring are as follows.

- The need for charging the batteries is an extra burden for residents, especially ones with cognitive disorders.
- Most current technologies don't provide completely waterproof sensors, as a result they need to be taken out during some activities such as bathing.
- People are often unwilling, forget, change clothes too often, or are not sufficiently clothed when at home to wear a badge, beacon, set of markers, or RF tag.

As a result, our goal in this study is to use small and lightweight wearable sensors and minimize the number of such sensors that are required. We have tried to investigate the usefulness of using wearable sensors in order to be able to realize the trade-off between drawbacks and effectiveness of using wearable sensors in activity recognition. Figure 8 illustrates some common wearable sensors often used for activity recognition.

Figure 8. This figure illustrates examples of
different wearable sensors.

C) Combination of All Sensor Types

Logan et al. (Logan, Healey, Philipose, Tapia, & Intille, 2007) look at activity recognition with a combination of sensor types. Their sensing technology consists of RFID tags, motion sensors and accelerometers. First of all, RFID has issues that are discussed in section 2.2.1.

Second, there are total of two residents in the home but only one is wearing the bracelet and activity recognition is done for just that person. In our experiment we take advantage of multiple subjects so there is no activity that is performed by one participant twice.

Their results show that infrared motion sensors have performed best among others, yet this is not a clear conclusion from the evidence. Since the study included a few infra-red sensors in the environment, for most activities there was a one-to-one mapping between sensors and activities they detected acceptably using infra-red sensors and their location was almost always where the targeted activities were performed in the house. As a result, infrared sensors performed very well in this study.

D) Summary

In this section we summarize related work in activity recognition with wearable, environmental and object sensors. Tables 1 and 2 summarize highlights of each of these sensor classes.

Table 1. Summary of related work on activity recognition using accelerometer.. "Body position" represents position on the body where the accelerometer was worn and "Activities" represents the activities that are targeted in that study.

Reference	**Body Positions**	**Activities**	**Accuracy**
(Veltink, Bussmann, de Vries, Martens, & Van Lummel, 1996)	Chest and thigh	Posture (standing, sitting and lying)	N/A
(Mathie, Coster, Lovell, & Celler, 2003)	Waist	Sit-to-stand, and stand-to-sit and walking for the	N/A

		purpose of fall detection	
(Bao & Intille, 2004)	Upper arm, wrist, thigh and ankle	Ambulation, posture and some ADLs	89%
(Chambers, Venkatesh, West, & Bui, 2002)	Wrist	Kung Fu movements	96.6%
(Lester, Choudhury, & Borriello, 2006)	Wrist, waist and shoulder	Ambulation, posture, riding elevator, walking up stairs	90%
(Mantyjarvi, Himberg, & Seppanen)	On the belt	Start/stop points, level walk, down/up stairs	83%-90%
(Al-ani, Ba, & Monacelli, 2007)	On the belt	Fall detection	N/A

Table 2. Summary of related work on activity recognition using environmental and object sensors. "Sensing technology" represents the types of sensors that were used, "Activities" represents the activity recognized in that study and "Accuracy" represents the overall accuracy achieved by that study.

Reference	**Sensing Technology**	**Activity**	**Accuracy**
(Wilson, 2005)	Environmental and Object(motion detectors, break-beam sensors, pressure mats, contact switches, water flow sensors, current sensors, and wireless object movement	Eating/drinking, housework, toileting, cooking, using a computer, watching Television and using the telephone	19.28% - 80.72%

	sensors)		
(Singla, Cook, & Schmitter-Edgecombe, 2008)	Environmental and Object (motion sensors, temperature sensors, humidity sensors, contacts switches in the doors, and item sensors on key items)	Phone call, cooking, wash hands and clean up	78.5%
(Tapia E. M., 2008)	Environmental and Object (Binary sensors on doors and objects)	Toileting, bathing and grooming	25%-89%
(Kasteren, Noulas, Englebienne, & Krose, 2008)	Object (Shake sensors)	Leaving, toileting, showering, sleeping, drinking and eating	79.4%
(Philipose, Fishkin, Perkowitz, Fox, Kautz, & Hahnel, 2004)	Object (RFID)	Toileting, oral hygiene, washing, telephone use, taking medication and etc	88%
(Patterson, Fox, Kautz, & Philipose, 2005)	Object (RFID)	Using bathroom, making meals/drinks, telephone use, set/clean table, eat and take out trash	81%
(Hodges, Newman, & Pollack, 2009)	Object (RFID)	Making coffee	71%

2.2.2 Algorithmic Models

Another important factor in activity recognition is the algorithmic model that is used for classification. In this section we present some of the most common machine learning algorithms used in this area.

A) Bayesian Network

A Bayesian network (Pearl, 1988) is a directed, acyclic graph whose nodes represent random variables and whose edges indicate direct influence between variables. Bayesian networks have provided acceptable results in areas requiring inference under uncertainty. There have been two general approaches in the literature for using Bayesian network. The bottom up approach has been used by Charniak and Goldman (Charniak & Goldman, 1993). Their approach manually translates activity knowledge into an associative network. Then it uses a number of rules to automatically convert an associative network to the corresponding Bayesian network. The problem with this model is that it relies on general Bayesian network inference engines to solve the problem and thereby it cannot utilize the special relations among the variables.

The other approach is a top-down approached, presented by Huber et al. (Huber, Edmund, & Wellman, 1994). In this model the Bayesian networks are constructed from the plan library before receiving any observations.

B) Logic Models

In general, a logic model sets out how an intervention (such as a project, a program, or a policy) is understood or intended to produce particular results (Rogers, Meyer, Walker, & Fisk, 1998).

In an activity recognition context, logic-based approaches keep track of all logically consistent explanations of the observed actions. Thus, all possible and consistent plans or goals must be considered. Kautz's event hierarchy (Kautz, 1987) is one of the earliest models for activity recognition. His model uses first-order logic to represent the abstraction and decomposition relations. However, the model does not take uncertainty into account and it has an exponential time complexity in worst case, measured in the size of input hierarchy.

As one step further, Lesh and Etzioni (Lesh & Etzioni, 1995) present methods in scaling up goal recognition to scale up his work computationally. Lesh and Etzioni's approach provides automatic plan-library construction from domain primitives, in contrast to Kautz's approach where the plan library is explicitly represented. Furthermore, they introduce compact representations and efficient algorithms for goal recognition on large plan libraries. In addition, they present methods for adapting a goal recognizer to handle individual idiosyncratic behavior given a sample of an individual's recent behavior.

Bouchard et al. (Bouchard, Giroux, & Bouzouane, 2006) have proposed a non-quantitative logical approach to ADL recognition in a smart home, dedicated to Alzheimer's patients. Their formal framework for the recognition process is based on lattice theory and action description logic. Their framework reduces the uncertainty about the prediction of the observed patient's behavior, allowing the assistant agent to anticipate the opportunities for assistance. This is achieved by dynamically generating the future potentially incoherent intentions of the patient, which result from the symptoms of their cognitive impairments (disorientation, memory lapse, etc.). This approach offers an effective solution to actual recognition of an ADL in a smart home, in order to provide assistance to persons suffering from Alzheimer's disease.

Rugnone et al. in (Rugnone, Poli, Vicario, Nugent, Tamburini, & Paggetti, 2007) focuses on the problem of automated recognition of sequences of events that may indicate critical conditions and unexpected behaviors requiring intervention and attention from caregivers. This work is based on a formal framework developed with temporal logic used for the specification of critical sequences of patterns and a behavior checking engine for automated recognition.

One important limitation of logic-based approaches is their inability or inherent infeasibility to represent uncertainty. They don't have any mechanism for deciding whether one particular plan is more likely than another, as long as both of them can be consistent enough to explain the actions observed. Moreover, they offer no ability of learning associated with logic based methods.

C) Decision Tree

A decision tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. Decision tree learning is one of the most widely used and practical methods for inductive inference. It is a method for approximating discrete-valued functions that is robust to noisy data and capable of learning disjunctive expressions. Decision Tree Learning algorithms include ID3, ASSISTANT, and C4.5. These decision tree learning methods search a completely expressive hypothesis space and thus avoid the difficulties of restricted hypothesis spaces. Their inductive bias is a preference for small trees over large trees. (Mitchell, 1997)

Some researchers including Maurer et al. (Maurer, Smailagic, Siewiorek, & Deisher, 2006) and Bao and Intille (Bao & Intille, 2004) have employed decision trees to learn logical descriptions of the activities. One advantage of using decision trees is that they generate expressive models. Moreover it can handle noisy data and has the added advantage of relative transparency of which sensors inputs contribute to classification. However, decision trees should be used when target function is discrete-valued. Moreover, it is prone to overfitting.

D) Artificial Neural Networks

An artificial neural network (ANN) is a mathematical model that is inspired by the structure and/or functional aspects of biological neural networks. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation.

ANNs provide a general, practical method for learning real-valued, discrete-valued, and vector-valued functions from examples. Algorithms such as back propagation use gradient descent to tune network parameters to best fit a training set of input-output pairs. ANN learning is robust to errors in the training data and has been successfully applied to problems such as interpreting visual scenes, speech recognition, and learning robot control strategies (Mitchell, 1997).

Kiani et al. (Kiani, Snijders, & Gelsema, 1998) have used probabilistic neural networks to recognize basic ADLs such as sitting, walking, lying and so forth. The average recognition rate of the trained neural networks is 95%, which is a good classification of all presented cases of the daily life activities. An automatic misclassification of 5% resulted from certain activities being too short or the occurrence of activities that were not included in the training set.

In another study (Rivera-Illingworth, Callaghan, & Hagras, 2005) Rivera et al. use a neural network agent based approach to recognize a few ADLs including eating, working on computer and sleeping.

One of the limitations of ANNs which has become the common criticism of this model is that it requires a large diversity of training for real-world operation. In addition, these models have high computation and time complexity due to the fact that the designer of NN systems will often need to simulate the transmission of signals through many of these connections and their associated neurons, which must often be matched with incredible amounts of CPU processing power and time.

E) Dynamic Bayesian models

A dynamic Bayesian network is a Bayesian network that represents sequences of variables. These sequences are often time-series or sequences of symbols (for example, protein sequences).

1) Hidden Markov Model

This model is explained in more in detail in section 5.2.3.

2) Conditional Random Fields (CRF)

A CRF is a finite state model with un-normalized transition probabilities. CRFs assign a well-defined probability distribution over possible labeling, trained by maximum likelihood or MAP estimation (Lafferty, McCallum, & Pereira, 2001). Recently some researchers such as (Vail, Veloso, & Lafferty, 2007; Kasteren, Noulas, Englebienne, & Krose, 2008) have used CRF in activity recognition.

2.3 Activity Types

2.3.1 Basic Activities

Activity recognition includes a wide range of activity types. One of the most common activity types in activity recognition research is basic activities. Basic activities consist of basic movements such as ambulation, sitting, laying down, shaking hands, riding and so forth. A considerable amount of research is dedicated to recognizing ambulatory movements (Bao & Intille, 2004; Lester, Choudhury, & Borriello, 2006). Wearable sensor is the most often used sensor for this purpose.

2.3.2 Activities of Daily Living (ADL)

Activities of Daily Living (ADLs) is a term used in health care to refer to daily self-care activities within an individual's place of residence, in outdoor environments, or both. In this study we only look at indoor activities. Health professionals routinely refer to the ability or inability to perform ADLs as a measurement of the functional status of a person, particularly in regards to people with disabilities and older adults (Meghan, 2002).

Different researchers have used different sets of activities for their studies. The most often used measure of functional ability is the Katz Activities of Daily Living Scale (Katz, Ford, Moskowitz, Jackson, & Jaffe, 1963). Katz gives a low level list of important daily activities: Feeding, Continence, Transferring, Toileting, Dressing and Bathing. In activity recognition we need a more specific list. Some groups have defined their own set of ADLs such as B-ADL (Bayer) (Hindmarch, Lehfeld, de Jonge, & Erzigkeit, 1998), LCADL (London Chest) (Garrod, Bestall, Paul, Wedzicha, & Jones, 2000). In this study, instead of defining our new set of activities we try to look at ADLs that are most common in activity recognition literature and categorize them based on the technology needed for recognizing each activity.

Basic ADLs consist of self-care tasks, including: (McDowell & Newell, 1996, 2nd edition)

- Personal hygiene and grooming
- Dressing and undressing
- Meal preparation and feeding oneself
- Functional transfers, e.g. Getting out of bed
- Voluntarily controlling urinary and fecal discharge
- Elimination
- Ambulation (Walking without use of canes or crutches or using a wheelchair, should be able to safely maneuver stairs.)

A more complete list of ADLs can be found in Appendix A. (Tapia E. M., 2008 ; Logan, Healey, Philipose, Tapia, & Intille, 2007)

2.4 Computational Challenges

In addition to sensing challenges which are discussed above, activity recognition involves some computational challenges as well. In the following subsections we discuss some of these challenges.

2.4.1 Sequential Order of Tasks in Complex Activities

Sequential order of tasks in complex activities means different possible order of tasks composing one particular activity makes the activity difficult to be recognized. When cooking a cup of noodle soap, for example, one might boil the water then pour noodles and at the end pour salt and pepper, while another individual might boil the water with salt and pepper, and then pour noodles, and another one might even boil the water in the microwave instead of using a conventional burner. Thus, modeling of the sequence of tasks in these uncertain scenarios is an interesting problem.

Furthermore, one might perform more than one activity concurrently. This makes the activity recognition problem even more challenging, since both activities might even share common tasks.

2.4.2 False Starts

A person may start an activity, and then suddenly begin a new task because something more important has caught his attention or because he simply forgot about the original task (Tapia E. M., 2008). In this case, correct start of the activity is not clear and start of the first activity does not have an end which is misleading for classification algorithm.

2.4.3 Lack of Sufficient Real-world Training Data

This is another general problem in activity recognition which might not look very important at first, but since most current activity recognition systems use supervised learning techniques, they require sufficient training data. The difficulty of providing training data consists of but not limited to finding participants, process of collecting, cleaning and annotating the data. Therefore some studies have used mock scenarios for collecting data (Krishnan N. K., 2010), some have used sample generating techniques and some try to come up with unsupervised techniques that don't need training data. In contrast, this study has used real-world annotated data in all of the experiments.

2.4.4 Occurrence of Irrelevant Actions

As described earlier, complex activities consist of sequences of actions. In real scenarios we have irrelevant actions along with a large variety of other tasks, which makes recognizing tasks at this level very hard. As an example, while executing the tasks in the activity making a cup of tea, the individual might look at different cupboards to find sugar or turn down the TV. This does not constitute interleaved activities, because turning down the TV is just an irrelevant action in making tea and it should not be considered as a complete activity. This problem is a particularly complex aspect of activity recognition and is still an open research challenge.

CHAPTER THREE

3. CASAS TESTBED

This chapter provides a detailed representation of the CASAS smart home technology and the conditions under which the study data was gathered. The technology behind each research study has an impact on the results, therefore it needs to be taken into consideration along with the results. As the main purpose of this thesis is to compare different sensing technologies in activity recognition, we devote a chapter to describe the sensing platform that we use for our study.

A collection of hardware and software tools is used in the CASAS project for the implementation of a smart home. The entire home is devoted to the study except one control room which is dedicated to experimenter use. That room is used in some studies where experimenters need to monitor participants throughout the study; they can look at the live video of participants and give them instructions through the microphone.

We have categorized CASAS sensing technology into three main classes. These sensors are listed in detail in Section 3.1. In addition, important issues related to their supporting infrastructure are provided.

The software built for the CASAS Technology Platform (CTP) is an agent-based system. CASAS Lightweight Middleware (CLM) has been designed to provide the communication infrastructure for sensors and different software tools. CLM is described in detail in section 3.2.

Section 3.3 talks about data representation and the database-backed architecture which is used for long term storage of the data gathered by the sensor platform.

3.1 CASAS Sensing Technology

The goal of CASAS is to be simple, reliable, energy efficient and user friendly. Many of the component devices are commercial products integrated into a variety of data buses to be read by the server. Data events from different sensors pass through middleware where they are processed and stored. We have tried to install sensors in a way that participants become comfortable with the new technology in their home after a few weeks.

Most sensors are attached to a Dallas 1-wire bus™. This bus helps in fast transfer of small data packets along a number of devices on a common serial bus. A custom board was made to attach the sensors to the 1-wire bus. This board is used in the PIR motion sensors, temperature sensors and door sensors of all of the CASAS testbeds. This board allows easier connection of different devices without serious modification. All sensors that are connected to the server and use the 1-wire bus share a single software agent which reports their activities. Some of the other sensors use RS232 (serial), USB (Universal Serial Bus) and power line signal injection to communicate with the server. Wireless sensors use ZigBee protocol to communicate with the server. However, each has a separate agent that reports events and provides an interface to communicate with them.

3.1.1 Environmental Sensors

Passive Infrared Motion Detector

The PIR Motion Detector used by the CTP is a Visionic™ model K-940. This device is designed for general purpose home security installations and usually is installed on the wall for a human-height field of view. In CTP, these sensors are used for two purposes: as area sensors or downward facing sensors.

The PIR Area sensors are usually positioned in a way that only views one room. They are often used for occupation recognition in a room. Thus, the sensor is triggered when someone enters the room but the exact position of them can not be realized. A stock area sensor is shown in Figure 9.

Figure 9. PIR sensor used in CASAS.

In order to obtain more detailed information about the position of residents, these sensors need to be attached to ceiling facing downward. As a result, it can only view the floor below it and more locality is achieved. By using this feature it can recognize if more than one person entered the room or where the person is located at each moment. Sensors can be modified for more or less focused view of the space.

Recently one wired PIR sensors have been replaced by Control4 / CardAccess motion detectors.

Control4™ Power Control System

Control4™ is used as wireless power control system in CTP. Control4™ uses ZigBee-based solutions for communication with wireless devices.

Ambient Temperature Sensor

For measuring the temperature in a room 1-wire temperature sensor including Dallas DS18S20 chip Dallas DS18S20 chip (Figure 10) has been used. This sensor can sense the surrounding temperature to within ½° C. We have installed them on the ceilings of rooms in the smart home. These sensor can recognize stove burner usage if it is installed above the stove. Since experiments were conducted in different times of year and day, we did not use data from Temperature sensors in current study.

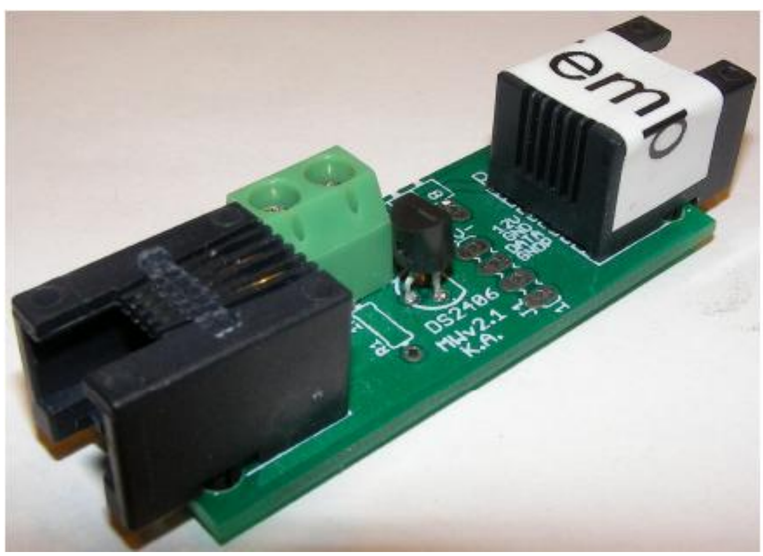

Figure 10. A CTP Ambient Temperature Sensor. Figure has been adapted from Crandall's work (Crandall, 2011)

Magnetic Door Sensors

Magnetic door sensors are one of the common and inexpensive home security items. It works with a simple magnet-driven reed switch, whenever the magnet moves away from the reed switch, it closes and the state on the 1-wire chip is changed. After sending this change to the server, server sends an "OPEN" event out over the middleware for processing and storage. When the magnet goes back into place, a "CLOSED" event is created. This sensor has been place on doors such as bedrooms, kitchen cabinets and refrigerators. Figure 11 illustrates examples of magnetic door sensor usage in Kyoto smart home. More detailed information about this sensor is provided in section 2.2.1.

Figure 11. Illustrates magnetic door sensors in CASAS smart home on a room door in left and cupboard door in right.

Item Presence Sensors

Item presence sensor (Figure 12) detects the presence of some items throughout the home. When the item is placed or removed from the plate, the switch is depressed and an event created. These item sensors are used for items that we cannot put object sensors on, such as glass and some medication bottles.

Figure 12. A CTP Item Sensor in top (Figure has been adapted from Crandall's work (Crandall, 2011)). Use of CTP in CASAS kitchen cupboard in bottom.

Stove Burner Power Meter

Stove burner power meter is designed to measure the use of the stove burner. The resulting voltage measurements are interpreted to give both duration and power setting when the residents are cooking. We did not use data from this sensor in current study.

Water Flow Sensor

In order to measure the use of the sink in the Kyoto (where the current study has been conducted) testbed's kitchen, a pair of water flow sensors were installed. These commercial products from Lake Monitors™, pictured in Figure 13, were placed on both hot and cold inflow pipes to the sink. Measuring water flow in smart home helps in detecting activities that need water usage. We did not use Water Flow Sensor in current study.

Figure 13. An example of the CTP water flow sensor.

OneMeter Power Metering

The OneMeter™ device monitors an inductive coil to determine the current wattage and cumulative kWh passed through a wire. In CTP it is installed in a way that lets the computer to poll for the current power status.

3.2 CASAS Middleware

CTP is an agent-based pattern and agents are communicating with each other through message passing in a distributed network. The CTP middleware is documented in depth in Kusznir's Master Thesis (Kusznir, 2009) under the title CLM as a Smart Home Middleware" where CLM stands for CASAS Lightweight Middleware".

The CTP middleware uses the XMPP protocol (Saint-Andre, 2004) as the messaging and presence layer. This means CTP can use any full featured XMPP server to manage the interconnection of agents and passing of messages. The agents implemented in the CLM system use XML formatted messages to communicate with one another.

3.2.1 Distributed Clocks and Event Timestamps

Synchronizing clocks in distributed system has always been an issue in message passing systems. CASAS has handled this problem with using the clock of the main server as the authoritative source of time. As events arrive at the

Manager agent, it stamps the current time on the event before passing it along for recording and processing.

3.3 CASAS Database and Data Representation

Data is first gathered from the sensor platform which is discussed in Section 3.1 then it is passed through the Middleware explained in Section 3.2. Output of CLM is in a standard format and is stored in a database. All events passed by the CTP from the sensor platform are stored in a database. These data are stored for future data mining and history building tools; it is also used for monitoring correct operation of testbeds. While the current implementation is an SQL database, any kind of structured repository can be used.

3.4 CASAS Testbeds

There have been total of 6 CASAS testbeds that have been implemented and deployed. The CPT infrastructure has been installed and tested in each of these testbeds for at least several months. A variety of studies such as detecting ADLs, detecting the number of residents, detecting the steps of an activity and providing resident activity prompting to aid older adults with dementia issues have been conducted using these tested. Subsets of the data gathered from these spaces are available from the CASAS shared data set web site (CASAS shared smart home datasets repository, 2011). We have only used the Kyoto testbed as source of data in this thesis.

3.4.1 Kyoto testbed

Kyoto (also known as the "smart apartment") is the primary research testbed for CASAS which is located on WSU campus. It is a three bedroom apartment as part of the WSU housing system. Data used in this thesis was gathered from this testbed. Layout of the apartment can be seen in Figure 14. Each resident has their own room with a bed, desk and closet. There is a shared bathroom, living room and kitchen. This apartment looks like a typical homes thus, it makes the results from the research performed here more applicable than partial smart home implementations or work done with specialized facilities.

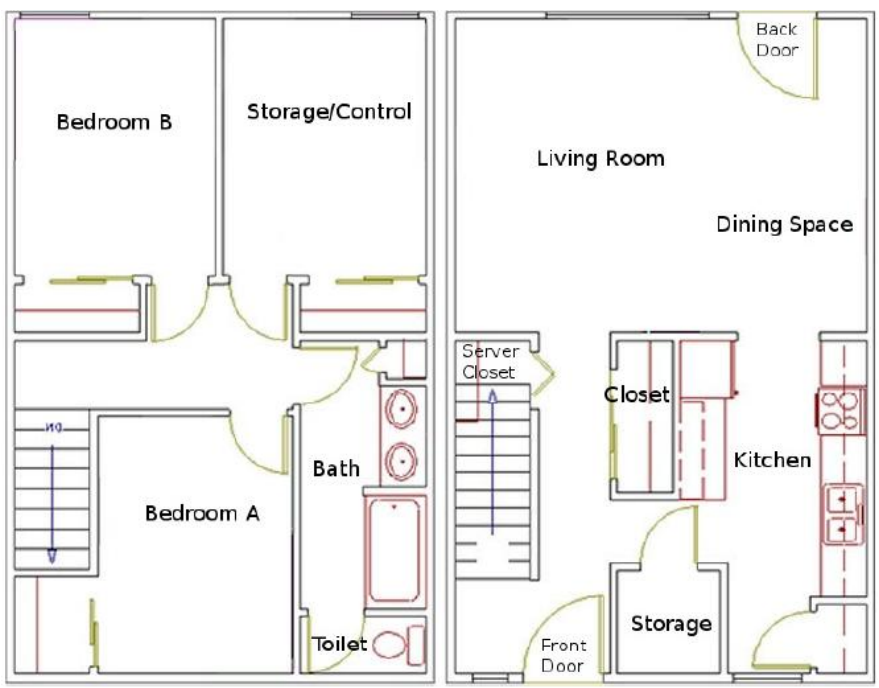

Figure 14. Labeled room map of Kyoto testbed

Kyoto has been data source for various types of studies such as Chen's study on associating of activities with energy consumption (Chen, Das, & Cook, 2010), recognition of ADLs with environmental sensors (Singla, Cook, & Schmitter-Edgecombe, 2008).

The Kyoto motion detectors are installed throughout the apartment except in the control room upstairs which is designated for experimenters. There is only one single area motion sensor installed in the room to monitor the occupancy of the room. Almost every door in the apartment has a magnetic closure sensor. Some of the cabinets in the kitchen, the microwave and the refrigerator are monitored in this manner. The kitchen sink has water flow sensors to monitor water usage.

Additionally there is a stove burner sensor, a water flow sensor, a number of item presence sensors, and a OneMeter power meter gathering data in the apartment. The Kyoso sensor layout is presented in Figure 15. In this figure, 'M' represents motion sensor, 'D' represents magnetic door sensor, ''I' represents item sensor. This layout doesn't show shake sensors which are used in the kitchen. Figure 36 illustrates kitchen layout with object shake sensor locations.

Kyoto residents can be classified in to two categories: full time and transient. Full time residents are students who live there but they are not part of CASAS study.

Transient residents are people who visit the apartment during the day for different purposes. These visitors may be performing maintenance, conducting different experiments, or participating in experiments. Based on the needs of each study sensor data generated by these individuals can be filtered out.

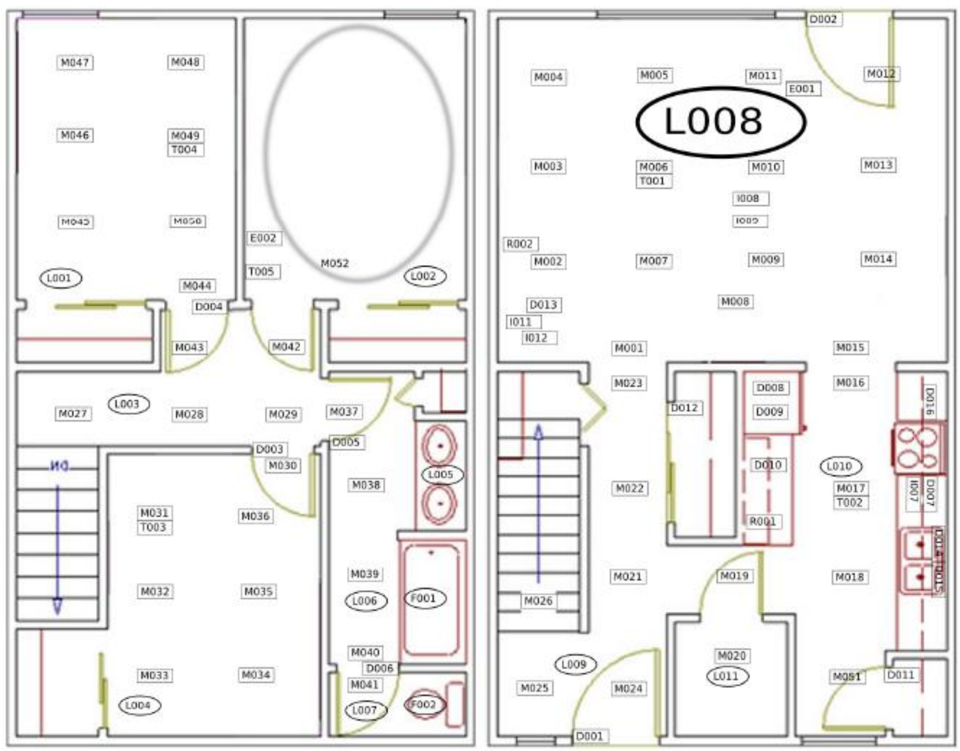

Figure 15. Sensor layout of Kyoto testbed

3.4 CASAS PyViz visualization tool

PyViz visualization tool (Thomas & Crandall, 2011) visualizes data steam. It is used in data annotation purposes. Annotators are observing events as they are replayed using this tool. They can see events at the map of the home, by using their common sense they annotate data samples. They labeled the start and stop event of each activity. Some features of PyViz can be seen at Figure 16 and Figure 17.

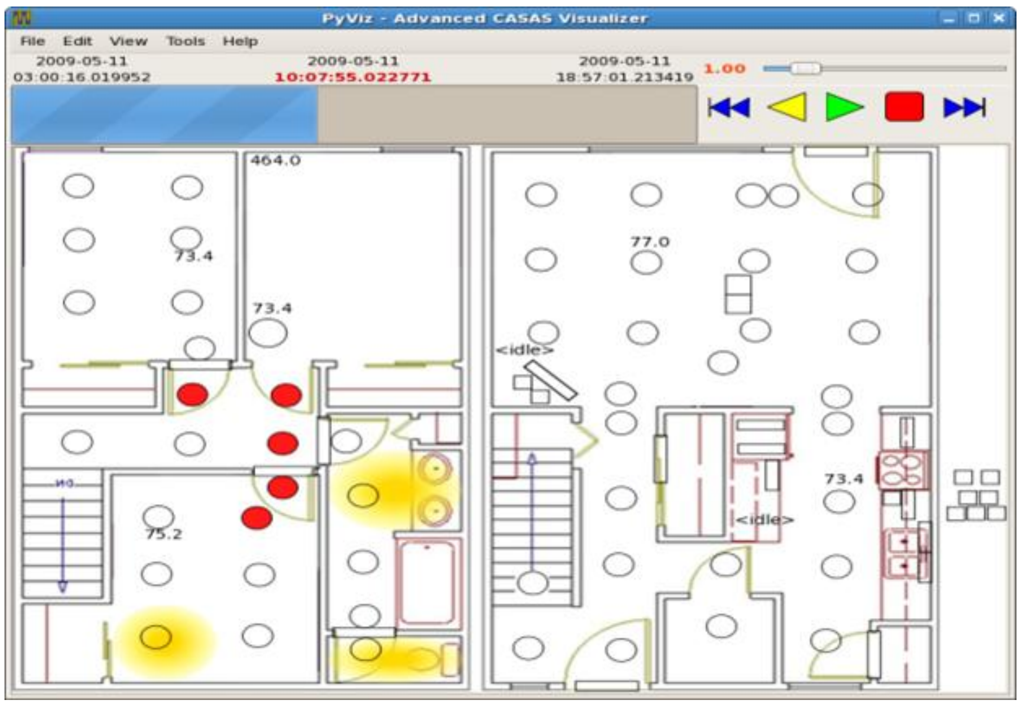

Figure 16. Kyoto layout with PyViz visualizer.

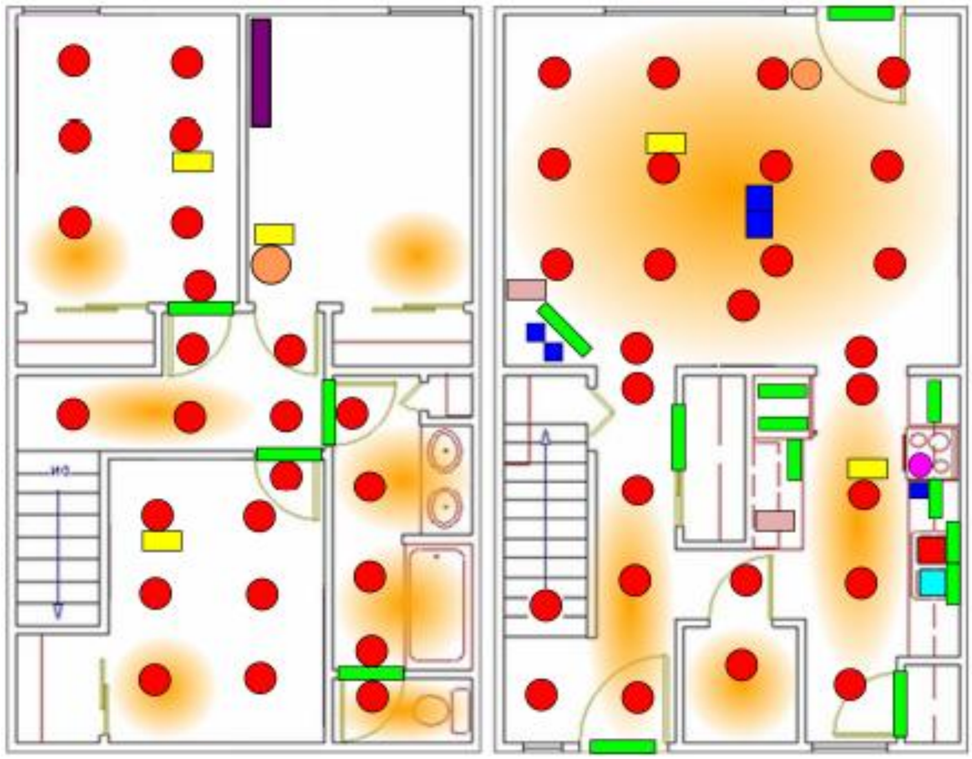

Figure 17. Graphically rendered SVG configuration file for Kyoto smart apartment. Figure has been adapted from (Thomas & Crandall, 2011)

CHAPTER FOUR

4. STUDY METHODOLOGY

In this chapter we present our approach to the problem of activity recognition. First, the sensor selection process and its challenges are discussed; lastly, the sensors used in this study are introduced.

4.1 Study Highlights

In addition to integrating different sensor types into one system and comparing their positive and negative points in one study, another factor which differentiates this work form previous studies is using a variety of everyday activities for a diverse sample population under conditions akin to those found in real-world settings.

A large number of previous works use data collected in lab or from one or few participants and often the subjects are the researchers themselves (Ravi, Dandekar, Mysore, & Littman, 2005; Srinivasan, Chen, & Cook, 2010). It is important to train and test activity recognition systems on data collected under real circumstances, because data collected in a laboratory usually has restricted or simplified activity patterns. However, little work has been done to validate the idea under real-world circumstances.

Moreover, research using naturalistic data collected from multiple subjects has focused on recognition of a limited subset of a few everyday activities consisting largely of ambulatory motions and basic postures such as sitting and standing (e.g. (Krishnan & Panchanathan, 2008; Lester, Choudhury, & Borriello, 2006)).

Foerster, et al. (Foerster, Smeja, & Fahrenberg, 1999) demonstrates a 95.8% recognition rate for data collected in the laboratory but recognition rates dropped to 66.7% for data collected outside the laboratory in naturalistic settings.

Although our study is not using data gathered in participants' real homes, but since experiments are conducted in a real home it can be considered as a realistic but controlled experimental setup. In addition, our participants have not been told what steps should an activity consist of and orders of steps were not specified.

There has been a considerable amount of research on algorithms for activity recognition in smart homes. We believe the performance of these algorithms depends heavily on the data and sensing technologies that are employed. In this study we do not intend to present a new algorithm, Instead, we focus on a comparison between different sensing technologies and their relation to recognizable activities.

One important point that this study tries to make is that no one can provide a unique list of sensor types that performs best for general activity recognition. A variety of factors including activity types, number and placement of sensors, sensor feasibility and purpose of the study affect the results and need to be taken into consideration when choosing a proper sensor type.

4.2 Sensor Selection

Adding new sensors to any system introduces challenges for the system infrastructure. In our case, we need to add wearable and object sensors to current CASAS sensing technology.

4.2.1 Challenges

(Wilson, 2005) has defined the following characteristics for sensor selection:

- Should fit into familiar forms, be inexpensive, preferably available off-the shelf and easy to install.
- Sensor data should be private and should not reveal sensitive information, especially identity.
- Arguably equally important sensors should not be perceived as invasive.

- Wireless sensors can be mounted to any surface, while wired sensors may require running cable to a central location.
- Processing sensor data should require minimal computational resources (e.g., a desktop computer). Sensors should be low-maintenance, easy to replace and maintain. Sensors will be neglected and should be robust to damage.

Finally, sensors should be low-power, requiring no external power or able to run as long as possible on batteries. As a last resort the device may need to be plugged in or powered by low voltage wiring.

One of the sensors we need to add to CASAS platform is wearable sensor that can assist us in recognizing ADLs. For this purpose, first we try to find a proper wearable sensor with the following features:

- **Accelerometer with 6 degrees of freedom.** There are many other wearable sensors which are being used in health care for monitoring respiration (Paradiso, Loriga, & Taccini, 2005), heart rate or blood oxygen levels. However, accelerometers are the best wearable sensors that can help us in recognizing day-to-day activities. An accelerometer is a sensor that measures the linear acceleration that is induced by gravity or by the movement of the sensor. It is sensitive to shock, orientation, and vibrations. There are different kinds of accelerometers based on type of construction and sensitivity range. The most informative ones have a gyro along with an accelerometer which provide us with 6 degrees of freedom.

- **Wireless.** Being wireless was an important factor for our system, since wired wearable sensors cannot be used in real home experiments. Performing ADLs requires the individual to move between different parts of the house, thus wearing wired sensors is not practical in this case.

- **Lightweight**. The comfort of participants is one of the most important factors to be considered when choosing appropriate technology. A big and heavy wearable sensor is likely to make participants unsatisfied, especially older adults.

By using existing, well tested off the shelf commercial products, the robustness, energy efficiency and profile of the system are often improved. Commercial products are also often packaged well, so their visibility profile after being installed is lower and the residents are less likely to notice the system after they become accustomed to it. In our case, finding a commercially available wearable accelerometer could solve many of our practical problems such as finding a way to wear the sensor on different parts of body, easily transmitting data to a PC and saving it in a familiar format.

4.2.2 Selected Sensors

We had to find a practical wearable sensor to add to CASAS sensing platform. As a result, we did one trial experiment with our first selected sensor (Sparkfun), after realizing its limitations we decided to replace it with another sensor (Shimmer). Ultimately, Shimmer sensor has been integrated into CTP. In this section we provide more detailed discussion on our sensor selection process for this study.

A. Sparkfun Accelerometer

The first sensor that we tried is an atomic IMU 6 Degree of freedom - XBee Ready chip marketed by Sparkfun Electronics (Figure 18)

1) Accelerometer Overview

This sensor is designed to give good performance at a low price. The unit can run as a hard-wired UART Interface (0-3V, 115200bps), or optionally with an XBee RF module, and is powered from a single LiPo (Lithium Polymer) battery. The processor is an Atmel ATMega328 running at 10MGz with 6 dedicated 10-bit ADC channels reading the sensors.

The 6-DOF Atomic uses these sensors:

- 1 x Free scale MMA7260Q TM triple-axis accelerometer, settable to 1.5g, 2g, 4g or 6g sensitivity

- 3 x ST Microelectronics LISY300AL TM single-axis, 300°/s gyros

All sensor readings are available through any terminal program in either ASCII or binary format, or with the 6DOF Atomic IMU Mixer demo application.

More features are as follows:

- Input voltage is 3.4 V to 10 V DC
- Current Consumption: 24mA (75mA with X-bee)
- Dimensions: 1.85 x 1.45 x 0.975 inches (47 x 37 25 mm)

Figure 18. Accelerometer with 6 degrees of freedom.

2) Battery

Since the accelerometer is wireless we need a battery along with it (Figure 19). We use Polymer Lithium Ion Battery - 2000mAh which is a very slim, extremely light weight battery based on the new Polymer Lithium Ion chemistry. This is the highest energy density currently in production. Each cell outputs a nominal 3.7V at 1000mAh. And comes terminated with a standard 2-pin JST connector - 2mm spacing between pins.

Features:

- 2C continuous discharge
- Excellent long-term self-discharge rates (<8% per month)
- Robust power source under extreme conditions (-25 to 60C)
- Dimensions: 2.09 x 1.3 x 0.225" (53 x 33 x 5.7 mm)
- Weight: 22g (0.77oz)

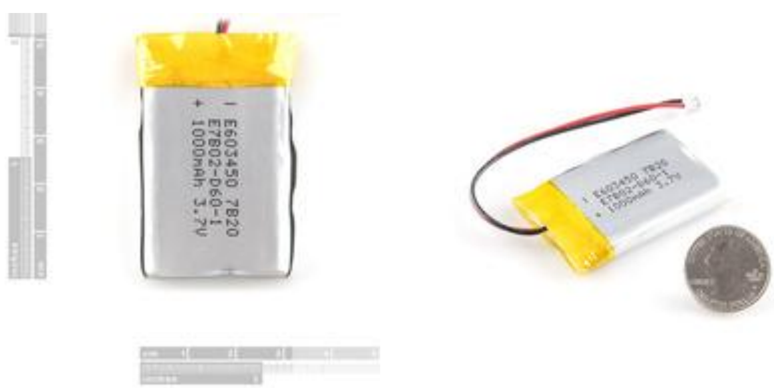

Figure 19. Polymer Lithium Ion Battery and its dimensions

3) Antenna

Another type of device that is used in this experiment is an XBee Pro 60mW Chip Antenna for wireless transmission of data from accelerometer to another XBee Pro 60mW Chip Antenna which is connected to a laptop USB. According to Sparkfun's website, this module allows a very reliable and simple communication between devices with serial port like microcontrollers, systems and computers as shown in Figure 20. Point to point and multi-point networks are supported.

Figure 20. Digi XBee Pro 60mW chip antenna.

Features:

- 3.3V @ 215mA
- 250kbps Max data rate
- 60mW output (+18dBm)
- 1 mile (1500m) range
- Built-in antenna
- Fully FCC certified
- 6 10-bit ADC input pins
- 8 digital IO pins
- 128-bit encryption

- Local or over-air configuration
- AT or API command set

Since it is just a chip and not placed in a predesigned box we couldn't fit it in a normal armband. In order to find a proper way to wear the sensor we faced a few challenges. First, one problem is the size of the battery which is relatively big. Second, the ON/OFF button is very tiny in the middle of the chip, thus it is not possible to place a box around it.

As a result, we wrapped the sensor and the battery with a bandage around the participant's arm. We performed one experiment with this sensor, which is explained in more detail in section 5.1. In that experiment we have one participant performing four activities of daily living.

4. Sparkfun Sensor Problems

During this experiment we faced some problems with the Sparkfun accelerometer.

- First of all, using it as a wearable sensor was not an easy task. Wrapping it with a bandage was giving us acceptable results but was not a practical solution for long term use.
- Second, since it is a chip which does not have any protection around it, data sometimes becomes noisy, especially when sensor is touched.

- Third, the collected data does not include a time stamp, therefore it was not a good option for our final intension which was integrating it into the CASAS sensor technology.

B. Shimmer Accelerometer

Our second approach is using Shimmer sensor. Shimmer is a low power wireless sensor platform that provides much of the functionality needed for wireless sensing. Designed for wearable sensing applications, the platform has an on-board microcontroller, wireless communication via Bluetooth or 802.15.4 low power radio, and the option of local storage to a micro SD card. The unit also has an integrated accelerometer for motion sensing, activity monitoring and inertia measurement applications.

The Shimmer unit acts as a baseboard for the full range of Shimmer wireless sensor modules. Shimmer can be connected to any one of a range of sensors such as Gyro, Magnetometer, ECG, EMG, GSR, GPS/Temperature, or Strain Gauge modules, whilst maintaining its small, lightweight, and wearable form factor. In this experiment we have used a gyro connected to shimmer units in order to monitor the Euler angles introduced by Leonhard Euler to describe the orientation of a rigid body. To describe such an orientation in 3-dimensional Euclidean space three parameters are required. Euler angles also represent three composed rotations that move a reference frame to a given referred frame. These three parameters are yaw, pitch, and roll.

Hardware Features:

- Processing: MSP430 microcontroller (8mHZ, 16 Bit)
- Communication: Bluetooth – RN-42 , 802.15.4 radio - TI CC2420
- Storage: Integrated 2GB microSD card slot
- Battery: 450mAh rechargeable Li-ion

Integrated Sensors:

- 3 Axis accelerometer – Freescale MMA7361
- Tilt/vibration switch

The shimmer sensor is designed for wearable and remote sensing applications. The shimmer unit is designed to be highly flexible and adaptable, easily integrating into existing systems and technologies. Shimmers are frequently used in activity monitoring, sport science, and intelligent building applications to name but a few applications. Due to its flexibility, the Shimmer platform is generally application agnostic.

Benefits:

Highly Configurable. A shimmer sensor can be programmed to meet any specific application, with configurable sensitivity, sampling rate, transmission rate and frequency, communication protocols, and packet formats. We programmed it in order to integrate it into the CASAS technology; so each sensor sends a packet to our existing server. The middleware will assign a timestamp to the packet and store it in the database. As a result, all data from different sensor types in CASAS are stored in one database with a synchronized timestamp.

Compatibility. Since Shimmer sensors use ZigBee protocol, it easily integrates and interacts with existing systems and technologies in CASAS smart home.

No Proprietary Software. As opposed to some commercial products that only provide processed data to users, with Shimmer we have full access to all raw sensed data to interpret data specific to any application, product or service requirements

Supporting Application. Supporting developer applications including labVIEW, ShimmerConnect, and a full range of TinyOS firmware.

Advantages over Sparkfun's sensors:

Size. Battery is embedded into the box. (Small size). (Figure 21)

Weight. Shimmer is Light Weight (baseboard and battery 15g; with enclosure 22g)

Style. Shimmer is stylish, functional enclosure with wearable straps available.

Figure 21. Shimmer accelerometer and its dimensions

C. Shimmer Gyro

The Shimmer Gyro (Figure 22) provides researchers with 3-axis angular rate sensing (gyroscopes) and additional features for enhanced operation and accuracy. Utilizing two integrated dual-axis angular rate gyroscopes, the Shimmer Gyro Board can perform complex motion sensing applications. The Gyro Board uses next-generation MEMs technology that offers higher performance, lower power and a rigid board implementation to ensure a perpendicular Z-axis. The Gyro Board is connected to the Shimmer main board via an internal connector pin, and is contained within the Shimmer enclosure. With fixed reference output, the Gyro Board runs off a secondary LDO for improved power-supply noise rejection.

The Shimmer Gyro Board has an enhanced user interface with a programmable button for application control (such as sampling start/stop, RF transmission start/stop, data marker), a programmable indicator, and a pinhole reset.

Figure 22. Shimmer gyro with shimmer accelerometer.

CHAPTER FIVE

5. EXPERIMENTAL RESULTS

In this chapter we present the setup and results of experiments performed in this study. We collected data from two experiments. The first experiment utilizes the Sparkfun accelerometer for four activities, where the goal is to choose the proper accelerometer to add to CTP. After realizing Sparkfun's limitations, our main data collection is performed with Shimmer sensor. It utilizes shimmer sensors together with all of the current sensors in CASAS Kyoto smart home and wearable sensors for six common ADLs. The goal is to compare these different technologies for activity recognition. All experiments and their results are detailed in the following sections.

5.1. Activity Recognition with Sparkfun Accelerometer

In this experiment one female participant age 24 performed four common ADLs: Eating, Reading, Hand Washing, and Sweeping. For *Eating*, we asked the participant to eat a bowl of oatmeal. *Washing hands* involved rubbing soap between two hands and rinse with water. For *Reading*, the participant was asked to read a magazine, which consists of repeatedly turning page after a few seconds of staying on one page. In real life the magazine could be replaced by a cookbook, telephone book, etc. Activity *Sweeping* was performed as sweeping the entire room with a broom. Each activity was repeated 10 times by the same participant. As a result, the total number of data samples is 40. Accelerometer frequency was 10 Hz with 6 output values: X, Y, Z, Pitch, Roll, Yaw. First three are acceleration in three dimensions, the rest are the Euler angles.

5.1.1. Problems with Pre-processing the Data

One of the problems we faced in working with this chip was having noisy and inconsistent data in some cases. In particular, because the chip does not have any protection layer, it is very sensitive to touch and its data becomes noisy very easily. As a result, the data has to be checked line by line and noisy data lines are ignored. It should be noted that because of the high sample frequency, ignoring these lines does not affect the data.

5.1.2. Mean Feature Approach

A) Trial One

In the first trial we generated the mean value for each column (so the new X is the mean of all X values of one activity run). In order to pre-process the data we need to transform total 40 sample files in to one input data file. Each line of this new file corresponds to one original output file which is one run of a particular activity.

Each data point is described by 7 attributes: mean of x, mean of y, mean of z, mean of roll, mean of pitch, mean of yaw and activity (the target class value). Different machine learning algorithms have been tested on the data and the best three were chosen.

B) Trial Two

By introducing a sliding window, we divide output files into equal-length segments. (For example, by dividing each output file into 5 equal segments, all activities are divided into 5 segments). In this way we are adding more features to our data and instead of having one mean value for each axis (x, y, z, roll, pitch, yaw), we will have *n* mean values. (*n* is the number of windows). This approach helped us to use the information in sequence of data and as a result boost the accuracy. Different segment numbers can be tried in the future. Results of these two trials are shown in Figure 23. Leave-one-out cross-validation has been used in these two trails.

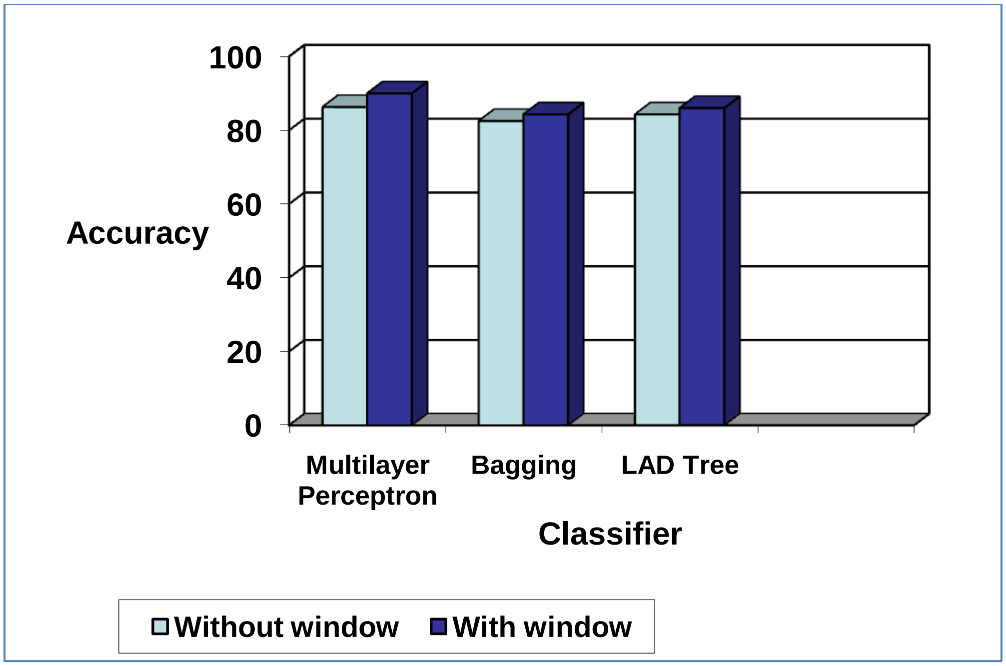

Figure 23. Comparison of results with and without windows.

5.1.3. Pattern Discovery Approach

In this section, we present a model which applies pattern discovery techniques to wearable sensor data in order to recognize daily activities. This model transforms raw accelerometer data into string sequences, and then uses suffix trees to generate structural patterns of different activities. This approach has been evaluated using data described in the previous section. Finally, we discuss the performance of the technique along with its advantages over other methods.

A) Background

In Bioinformatics, researchers have come up with many algorithms for pattern discovery in nucleotide sequences. In Computer Science we deal with time series data from different sensors. Converting it to symbolic data provides us with the opportunity to use algorithms from the text processing and bioinformatics communities.

In this approach, we demonstrate the use of a pattern mining structure called Suffix Tree. We use Suffix Tree as an activity representation to efficiently recognize the structure of activities (Gusfield, 1997) by analyzing their constituent event-subsequences over multiple temporal scales, which has already been widely applied into text indexing (Grossi & Vitter, 2000) and genome analysis (Abouelhoda, Kurtz, & Ohlebusch, 2002). Chen et al. have used suffix tree for monitoring, analyzing and predicting energy usage in CASAS smart home (Chen & Cook, 2011). Unlike other mining methods, which are exponential in their complexity, a suffix tree can be constructed in linear time O(n) for a data sequence of length *n*, and also spend O(m) time to search for a subsequence of length *m* in a sequence of length *n*, regardless of the length of *n*.

Suffix trees have been used in an unsupervised manner in the context of smart environments and assistive living. Minnen et al (Minnen, Starner, Essa, & Isbell, 2007) use Suffix trees for discovering frequently occurring motifs in the sensor streams and then apply a Hidden Markov Model (HMM) for finding all the occurrences of the motifs in the original time series. They validate their model using data from wrist-mounted sensor and were able to discover characteristic actions in dumbbell exercises with an overall accuracy of 86.7%. Suffix Trees have also been used in unsupervised activity analysis with environmental sensor data from a kitchen environment (Hamid, Maddi, Bobick, & Essa, 2007; Hamid, Maddi, Bobick, & Essa, 2006). To the best of our knowledge, no study on supervised learning in activity recognition has used Suffix tree. And no study with wearable sensors has used motif discovery methods for recognizing complex activities like ADLs.

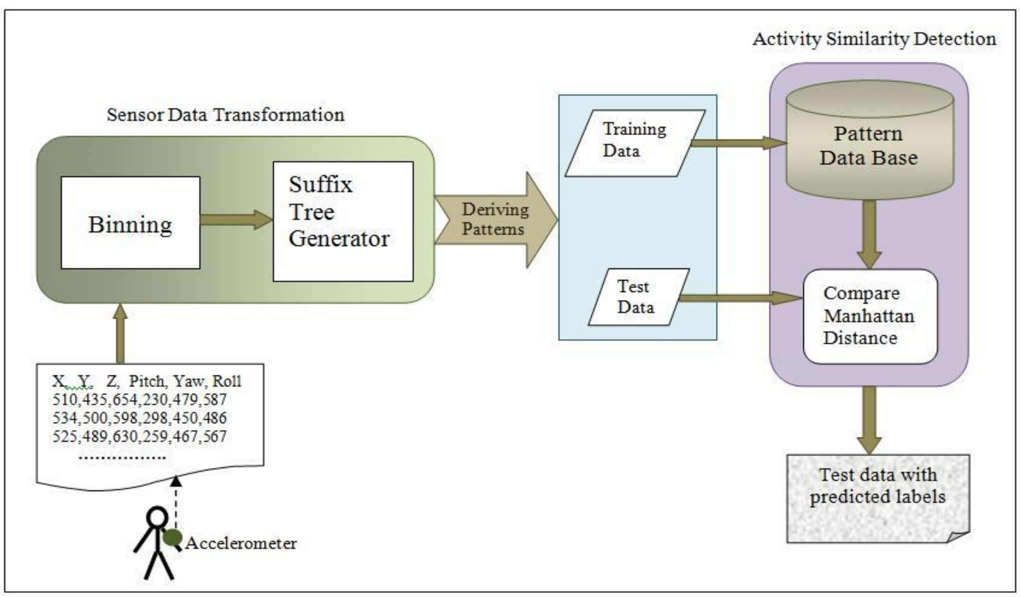

Figure 24. Main components of proposed model for recognizing activities from raw sensor data.

B) Model Description

Our model consists of 3 components that are shown in Figure 24, namely: sensor data collection, sensor data transformation, pattern database creation and activity similarity detection.

Sensor Data Transformation

In this module, data goes through binning and suffix tree generation phases. Here, the purpose of binning is to discretize continuous raw sensor data. Data values that fall in a given small interval, a bin, are replaced by a value representative of that interval, resulting in a sequence of symbols. This decentralization is done on each axis separately. The next step is to generate a Suffix tree for each sequence. After going through the Sensor Data Transformation component, we have 6 suffix trees for each sample representing the 6 dimensions of the data. Next, patterns are derived from each suffix tree and training samples are stored in Pattern Database component. This life cycle of data is shown in Figure 25.

Activity Similarity Detection

In this component test data is compared with samples from the Pattern Database to compute the pattern similarities. The activity that is more similar to test sample is chosen as the test sample predicted label. A more detailed description of the proposed model and its components are presented in the following subsections.

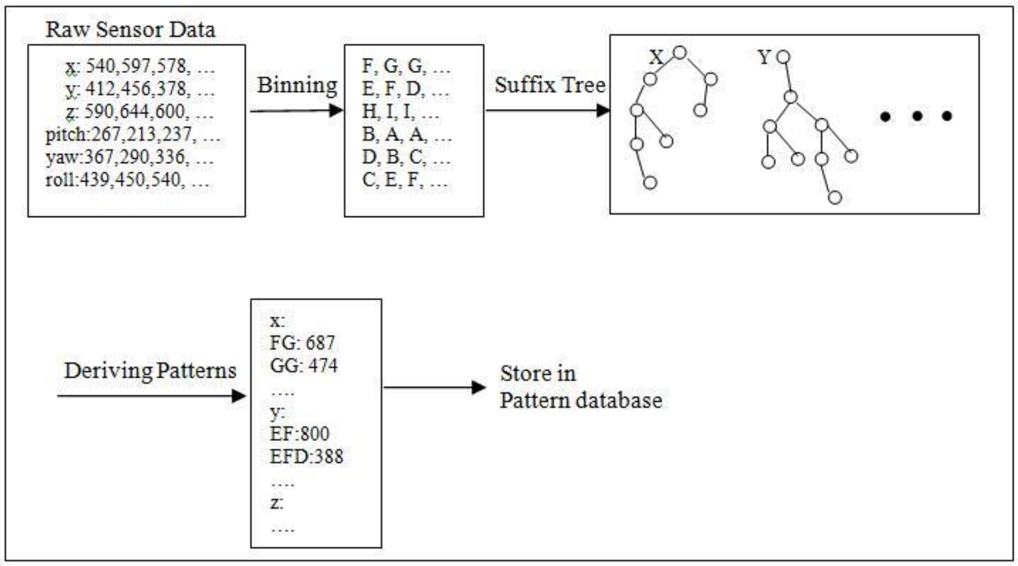

Figure 25. Preprocessing of data

Sensor Data Transformation

BINNING METHOD

The input data collected by the wearable accelerometer is a time series stream of sensor events, which are composed of six different features (x, y, z, pitch, yaw, Roll) as were described in previous section.

We define the wearable sensor sequence of n sensor events $E = < e_1, e_2, e_3, \ldots, e_n >$, where $e_i = \{x_i, y_i, z_i, pitch_i, yaw_i, roll_i\}$. To address the goal of mining sequence data, we discretize the wearable numerical values using equal width binning (Liu, Hussain, Tan, & Dash, 2002) which has been used to preprocess continuous-valued attributes by creating a specified number of bins, or numeric ranges.

For equal width binning, we define V_{max} as the maximal value of $\mathbf{E}$ and V_{min} as the minimal value of E. The number of bins is defined as k, which is assigned by the user and the interval of the bin I is defined as follows:

$$I = \frac{V_{max} - V_{min}}{k} \qquad (1)$$

The continuous range of a feature is evenly divided into equal-width intervals I and each interval represents a bin.

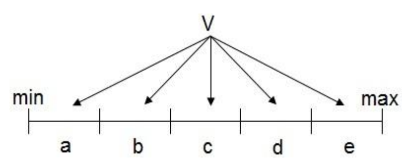

Figure 26. Transform sensor numerical data into character symbol (k=5, V is an accelerometer data)

We define alphabet $\Sigma = \{a, b, c, \ldots\}$ that represents the set of sensor pattern symbols. $\|\Sigma\|$ is equal to the number of the bins $\mathbf{k}$. Figure 26 depicts how an accelerometer data is transformed into a character symbol using equal width binning ($k = 5$). As shown in the graph, we can transform wearable sensor events $E = < e_1, e_2, e_3, \dots, e_n >$ to six different string sequences $S = < s_1, s_2, s_3, \dots, s_n >$ over the alphabet Σ.

Suffix Tree Generator

Let S' denote a string over the alphabet Σ. Let $\$ \notin \Sigma$ be a unique termination character, and $S = S'\$$ be the string resulting from appending $\$$ to S'. Let $|S| = n$ and $suff_i = S[i]S[i+1] \dots S[|S|]$ be the suffix of S starting at i^{th} position. The suffix tree of S is a compacted trie of all suffixes of string S over the alphabet Σ.

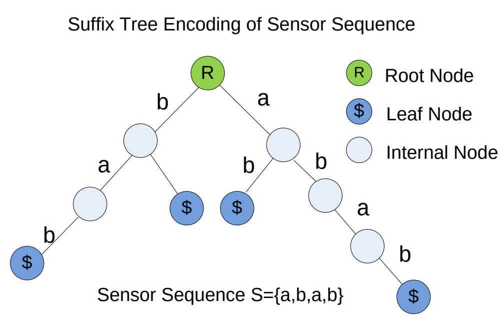

Figure 27. A Suffix Tree defined on a wearable sensor sequence S can represent every subsequence in S with at most 2m nodes where m is the length of S.

Traditional suffix tree construction algorithm starts from the root and follows the unique path matching characters in $suff_i$ one by one until no more matches are possible. If the traversal does not end at an internal node, it creates an internal node. The running time of this algorithm is proportional to $|suff_i| = n - i + 1$.

The total running time of the algorithm is $\sum_{i=1}^{n}(n-i+1)=O(n^2)$. In order to achieve $O(n)$ running time, we use McCreight's algorithm (McCreight, 1976) to construct suffix tree, which uses suffix links to speed up the insertion of a new suffix.

The figurative illustration of the transformation of a sensor sequence into its equivalent Suffix Tree is shown in Figure 27. No two edges of Suffix Trees begin with the same symbol, thus every unique subsequence in S, starting from the root node can be generated by traversing through the suffix tree.

Pattern Database

After generating suffix trees for each sample, patterns are recognized from them. Patterns of training set are stored in a database called Pattern Database. Each training sample in Pattern Database consists of a list of patterns and their frequencies for each axis (Figure 28). New training samples can be added to the database at any time.

X	Y	Z	Pitch	Yaw	Roll
AB: 1799	BC: 1576	AC: 2654	HD: 1053	GH: 987	BC: 2043
ABF: 1754	BCD: 1043	DE: 1043	FA: 674	AD: 432	FE: 1860
BD: 1232	FE: 498	FE: 876	DCG: 237	BCF: 254	ADC: 438
DA: 987	FED: 5	ADC: 198	EH: 58	EAC: 35	FDB: 176
DAA: 800	.	FGDE: 23	FBC: 32	.	AB: 3
.	.	AHF: 10	.	.	.
.	.	.	.	.	.
.		.	.		.
		.			

Figure 28. A training sample ready to be stored in Pattern Database

Activity Similarity Detection

In the Activity Similarity Detection component, a test sample is compared against all activities and the most similar activity to the test sample gets chosen as its predicted label. This comparison consists of 4 steps that are explained below.

The first step is to make clusters from samples with the same activity labels. Thus, each cluster represents an activity. For two samples q and r, we define their distance $D(q,r)$ as in (2). It measures the distance between two samples according to similarity in their patterns.

For two data samples q and r, let P_q and P_r denote sets of their patterns respectively, and x_i be the i^{th} pattern. Let $f(x_i \mid P_q)$ and $f(x_i \mid P_r)$ denote the frequency of pattern x_i in q and r respectively. Then we define $D(q, r)$ as follows:

$$D(q,r) = \sum_{x_i \in P_q} f(x_i | P_q) - f(x_i | P_r) + \sum_{x_j \in (P_r - P_q)} f(x_j | P_r)$$

(2)

For each activity (cluster) *A*, the average distance from test sample *t* is calculated using (3). Here *K* is the count of instances in that cluster (i.e. frequency of instances with activity label *A*); and s_A is a training sample with label *A*. Then the distance from activity (DFA) *A* for test data *t* will be as follows:

$$\mathrm{DFA}(t, A) = \sum_{s_A \in A} \frac{D(t, s_A)}{K} \tag{3}$$

Finally the activity that has the least DFA from test data will be chosen as predicted label for that sample. Thus, predicted label (PLabel) for test data *t* will be defined as follow:

$$\mathrm{PLabel}(t) = \mathrm{Min}_{A \in \text{activities}}\, \mathrm{DFA}(t, A) \tag{4}$$

3) Results

Leave-one-out testing is commonly used while experimenting with very small datasets (Eftestl, 2010). Figure 29 shows the accuracy of five different algorithms we used for this comparative study. We have used these algorithms since they are

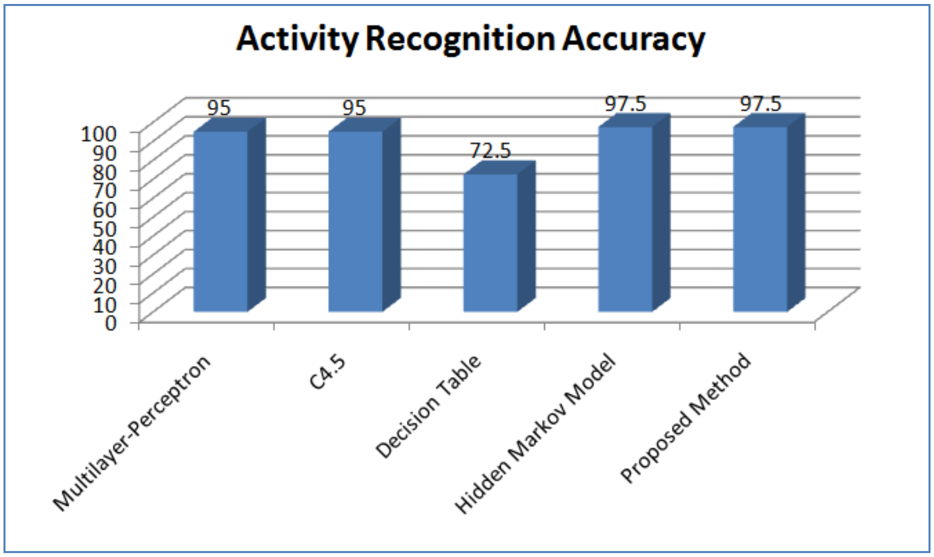

Figure 29. Activity recognition accuracy over different models.

commonly used in activity recognition studies. The models we compare are: Multilayer Perceptron (MLP) (Frank, 1961), C4.5 (Quinlan, 1993), Decision Table (Fisher, 1966) and Hidden Markov Model (HMM) (Rabiner, 1989).

For the first three algorithms, we represent the data using six features ($\overline{x}$, $\overline{y}$, $\overline{z}$, $\overline{pitch}$, $\overline{yaw}$, $\overline{roll}$). Each of these values represents the mean of the corresponding value over a fixed data collection time period. These features are commonly employed for activity recognition with wearable sensors

Some methods use the raw data directly, thus they consider all information hidden in data. We tried two approaches for this group: The first one is our proposed method, which uses pattern extraction and the second one is a hidden

Markov model, which is a probabilistic approach for modeling activities. As illustrated in Figure 29, both these methods from second group have achieved higher accuracy compared to the first group of algorithms. The proposed method and the HMM both achieved 97.5% accuracy.

Although our approach did not offer a significant accuracy advantage over HMMs, we believe our model has the advantage of not requiring any complex learning process. Time complexity of our model is $O(nlogn)$, where *n* is the length of each sequence. However, for the HMM, time complexity is $O(T \times |Y|^2)$, where *T* is the number of events and *Y* is the number of activities.

Figure 30 graphs the accuracy of these models for each activity. It can be seen that *Washing hands* was the easiest activity to be recognized by all five models and four of them achieved 100% accuracy. We think the reason is the unique rotational gesture involved in *Washing hands* that is not in the other three activities. Time might be another factor that helps recognizing this activity, since this is the shortest activity among the four different activities.

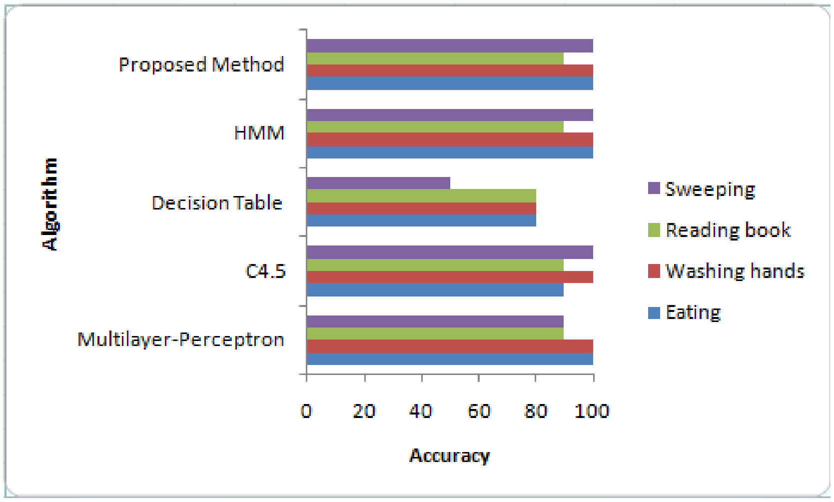

Figure 30. Activity recognition accuracy over different models specified for different activities.

Reading has been wrongly classified as *Eating*; that might be because they both have similar patterns (bringing spoon to mouth and turning pages both happen repeatedly with a considerable period of time in between).

D) Conclusion

Activity recognition is an essential component in enhancing health care in smart environments. This section has proposed a method for recognizing activities of daily living (ADLs) through supervised learning. Using a single, wearable accelerometer, the raw sensor data are first discretized into bins; each bin is given a symbol. Suffix trees are constructed based on training data and stored in a pattern database. A similarity measure is proposed for recognizing new patterns and used to compare the new pattern to those stored in the pattern database. By using a Suffix tree, this model adapts pattern discovery techniques to address the problem of activity recognition with wearable sensors. We have compared this model with 4 common machine-learning algorithms.

Our experimental results show that quantizing data and using pattern discovery methods can help in activity recognition since it is extracting beneficial information from raw data. Furthermore, we have shown that achieving high accuracy in activity recognition can be done with only one wearable sensor.

It can be seen that all four activities have been classified with relatively high accuracy. Usually accelerometers don't provide a very high accuracy in recognizing high level activities. In this study there are a number of elements that helped in achieving current results:

- Experiments have been conducted with only one subject, so we don't have variation in performing activities among samples.
- Moreover, it almost took the same amount of time for participant to do one activity in different rounds, and most of the activities have different duration by nature. Thus, we believe duration of activities was an important factor that helped the classifiers.
- Subject performing the activities was a researcher with the background knowledge of this study and its purpose. This had effect on how the subject performed activities.

5.2 Experiments with Shimmer Accelerometer

In this experiment, data is collected in the Kyoto smart home for 24 participants. Participants are aged from 18 to 35. We use shimmer accelerometer for two purposes: first, as wearable accelerometer and second, as shake sensor. More detailed information about shimmer sensors is provided in section 4.2.2 of this study.

As wearable sensor shimmer accelerometers are attached to wrist and hip of participants. Gyro sensor is added to the accelerometer in order to capture 3-axis angular rate sensing. Figure 31 and Figure 32 illustrate where the sensor was placed on participants.

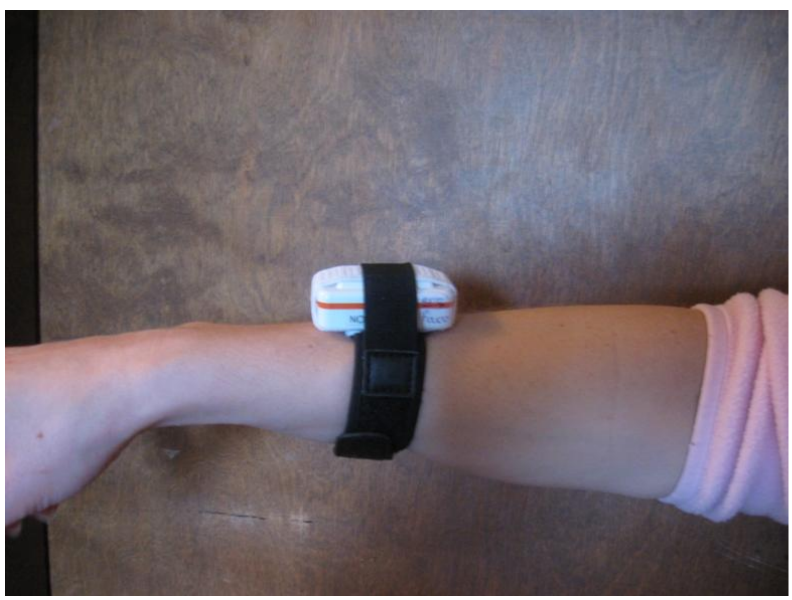

Figure 31. Participant wearing shimmer accelerometer.

Figure 32. Participant wearing shimmer accelerometer

As shake sensors, accelerometer with some modifications is used, so that it only sends ON/OFF messages. This modification and adding a threshold help us in using them as binary shake sensors. Shake sensors are attached to objects using Velcro straps so that they can easily be taken on/off.

5.2.1 Data Collection

A) Questionnaire

Participants spend approximately 5 minutes at the beginning of their sessions completing a paper-and-pencil based questionnaire about their age, education history and cognitive health. We require participants to be in a healthy cognitive state.

B) Activities

A total of 6 experiments have been used in this study. We will explain them in more detail in the following discussion.

Sweeping. In this task the participant is asked to sweep the kitchen floor and dust the dining room and living room. They are told that all the supplies they need for this task, including the broom, duster and dustpan and brush are located in the kitchen closet.

Medication. Participants are asked to fill a 7-day pill holder. They will find the 7-day pill holder, the pill bottles and the directions for filling the pill holder in the kitchen cupboard. The directions for filling the pill holder are taped to the inside of the cupboard door.

Cooking. For this task participants are asked to prepare a cup of noodle soup and get a glass of water for a friend. They are told that a glass, a measuring cup, the cup of noodle soup, and utensils are located in the cupboard. They have to fill the measuring cup with water and microwave for 3 minutes. Then follow the remaining directions on the cup of noodle soup to prepare the soup. In addition, they are asked to fill the glass with water using the pitcher of water located on the top shelf of the refrigerator. When they finish pouring the hot water into the cup of noodles and obtaining a glass of water, they have to bring all items to the dining room table for their friend.

Watering Plants. For this next task, they are asked to lightly water the apartment plants. There are 3 plants; two plants are located on the kitchen windowsill and the other plant is located on the living room table. The watering can is in the kitchen closet. Participants have to add water to the watering can using the kitchen faucet and lightly water the plants.

Hand Washing. For the hand washing task participants are asked to wash their hands at the kitchen sink using the hand soap. After they are done, they need to dry their hands with the cloth towel.

Washing Kitchen Countertops. For the last task they are asked to clean the kitchen countertops. They need to use the sponge and the dish washing soap to clean the countertops.

Participants perform activities in the first floor of the Kyoto smart home. While they are doing the experiments no one else is in the first floor so that no unwanted sensors get triggered. One or two experimenters stay in control room located in the second floor supervising participants by watching them through cameras and talking to them through microphone. They read the instructions for each activity and ask participants to do it with freedom in choosing orders of the steps of each activity.

Object and wearable sensors need to be charged everyday which is done by experimenters at the end of each day.

5.2.2 Data Annotation

In supervised methods data samples need to be labeled, thus annotators have to go through the samples and label them. This process of labeling data is called data annotation. One of the main concerns for supervised learning is the costly annotation process after collecting the data. Examples of data samples can be found in Table 3 and Table 4.

Table 3. Subset of data used in this study

Sensor Type	**Timestamp**	**Sensor ID**	**Value**
Motion	2011-03-16 12:41:21.913422	M014	ON
Shake	2011-03-16 13:14:13.14079	SS010	MOVED
Item	2011-03-16 13:12:07.425369	I006	ABSENT

Table 4. Format of shimmer accelerometer output used in this study.

Timestamp	**Sensor Name**	**Value (X, Y, Z, Pitch, Yaw, Roll)**
2011-03-16 13:12:07.236622	SG023	(1190,2495,2440,1849,1834,1938) (1230,2442,2458,1860,1825,1887) (1237,2423,2474,1833,1825,1887)

		(1193,2431,2485,1848,1793,1879) (1236,2409,2457,1828,1787,1853)

A) Common Approaches

Research community has used different approaches for annotating their data. A few of them is discussed in the following.

Experience sampling method

ESM refers to a method of data collection in which participants respond to repeated assessments at moments over the course of time while functioning within their natural settings. Bao has used worksheets for participants to fill out after performing each activity (Liao, Location-based activity recognition, 2006). The tools of ESM have evolved to allow greater ease of data collection for the researchers as well as the participant. At its nascence, participants carried pagers or alarm watches, along with a stack of paper on which they recorded their responses when signaled. Today, hand-held computers (a.k.a. personal digital assistants (PDAs) or palmtop computers) can be pre-programmed to signal participants at random moments.

The palmtop computers allow data to be directly transferred to statistical software packages or other programs for immediate analysis, and with no data entry, mistakes are minimized or eliminated altogether. Furthermore, participants cannot as easily fake their responses as with the paper–pencil measures.

Scollon et al in (Scollon, Kim-Prieto, & Diener, 2003.) provide a detailed discussion on ESM and its strengths and weaknesses. Some of the limitations mentioned in their work are as follows.

Participant interruption is one of the limitations of ESM; signaling participants a few times a day can be annoying. Moreover, each participant has to carry the device all the time which is another burden. Fake response is another weakness. With wrist watches or beepers and paper-pencil reports, participants can easily fake their responses by completing all their forms in one sitting.

Using Questionnaire and Interviewing after the experiments

Another common technique for data annotation is subject self-report recall surveys which can be done by using questionnaire or interviewing subjects after the experiments. However, these techniques are prone to recall errors and lack the temporal precision required for training activity recognition algorithms.

Using Video and Audio

Some studies record full experiments and have annotators to annotate the data by watching the videos (Chambers, Venkatesh, West, & Bui, 2002; Logan, Healey, Philipose, Tapia, & Intille, 2007). This approach is more reliable but has issues such as intruding the privacy of participants, need enormous storage for recording audio and video, being costly and so forth.

B) Data Annotation for Current Study

Data annotation for this study is performed by two experimenters who themselves were involved in conducting the experiments. Due to the limitations mentioned above, we didn't use ESM method nor did we use any video or audio recording techniques. The human annotators were taught to observe the events as they were replayed using PyViz visualization tool which is described in section 3.4.

5.2.3 Methodology

A) Features

Features used in this experiment can be divided in to three groups based on the type of sensor that generates that data. The features we used are discussed in the following discussion in more detail. Different feature settings are used and the best result is achieved by settings in experiment #3.

Experiment #1

Environmental: Motion sensors are divided in to four groups based on the area of the sensor. These features are: LivingRoom, DiningRoom, Kitchen, and KitchenWindow. We count the number of times that a sensor in one group is triggered.

Object: Object sensors are divided into groups based on the activity in which they are triggered. The features are: Sweeping (dustpan, broom, duster), Cooking (bowl of noodles, water pitcher, measuring cup, glass, fork), WateringCan, HandSoapContainer, DishWashingSoap, and Medication (pill dispenser, medicine bottles).

Wearable: Output data consists of 6 columns. For each column the following features are used: mean and standard deviation. These are considered for each arm and hip sensor separately.

Experiment #2

Environmental: Motion sensors are used directly as features. Each feature represents how many times a motion detector is fired during the activity.

Object: Object sensors are used directly as features.

Wearable: For each column the following features are used: Mean, Standard deviation, XY correlation, YZ correlation, XZ correlation, Pitch / Yaw correlation, Pitch / Roll correlation, Yaw / Roll correlation. These are considered for each arm and hip sensor separately.

Experiment #3

Environmental: Instead of recording the number of times a motion sensor was triggered during an activity and using this value as one of the data features, we compute the ratio of the number of times each individual motion sensor is triggered to the number of times all motion sensors are triggered in that activity.

Object: The two following features have been added:

- Total duration of each object use in each activity.
- Average duration of each object use in each activity.

Wearable: The same as experiment #2.

B) Models

In this section we try most common classification algorithms for discrete data. The best algorithm will be chosen as a classification model throughout the remainder of the thesis.

1) SVM

A support vector machine (SVM) is a set of related supervised learning methods that analyze data and recognize patterns, used for classification and regression analysis. Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.

Kernel-based methods and SVMs in particular, are among the best performing classifiers on many learning problems. By using Kernel-based methods we will have linear-speed learning in non-linear spaces. SVMs belong to a family of generalized linear classifiers. These classifiers simultaneously minimize the empirical classification error and maximize the geometric margin. In addition, support vector machines ignore all but the most differentiating training data (those on or inside the margin).

For linear SVMs, at training time a quadratic problem should be solved, and at test time prediction is linear in the number of features and constant in the size of the training data.

The limitations of kernel-based methods include the fact that choosing an appropriate kernel is challenging and must be done for each application. In addition, the high dimensionality of the original learning problem can pose a computational bottleneck for the learning algorithm.

For this study the SMO function found in Weka software is employed. This algorithm implements John C. Platt's sequential minimal optimization algorithm (Platt, 1998) for training a support vector classifier using polynomial or RBF kernels. This implementation globally replaces all missing values and transforms nominal attributes into binary ones. It also normalizes all attributes by default. In our study we deal with a multiclass problem. Weka handles multiclass learning problems using a combination of pairwise classifiers.

We considered this algorithm as one of our choices because Support Vector Machines are considered to be the best performing classifier among many learning problems.

2) HMM

The hidden Markov model can be considered as simplest simple dynamic Bayesian networks. A hidden Markov model (HMM) is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (hidden) states. In a hidden Markov model, the state is not directly visible, but output, dependent on the state, is visible. Each state has a probability distribution over the possible output tokens. Therefore the sequence of tokens generated by an HMM gives some information about the sequence of states. Hidden Markov models are especially known for their application in temporal pattern recognition.

Figure 33 represents the general architecture of a HMM where each circle represents a random variable. The random variable $x(t)$ is the hidden state and the random variable $y(t)$ is the observable state at time t. The arrows are used to denote conditional dependencies. The conditional probability distribution of any hidden state $x(t)$ at time depends only on the value of its preceding hidden state i.e. $x(t-1)$ i.e. the values at any time before time $t-1$ have no influence on the value of state at time t which essentially is the Markov property. Also, the value of the observable state $y(t)$ depends only on the value of the hidden state $x(t)$ given at time t (Singla, 2009).

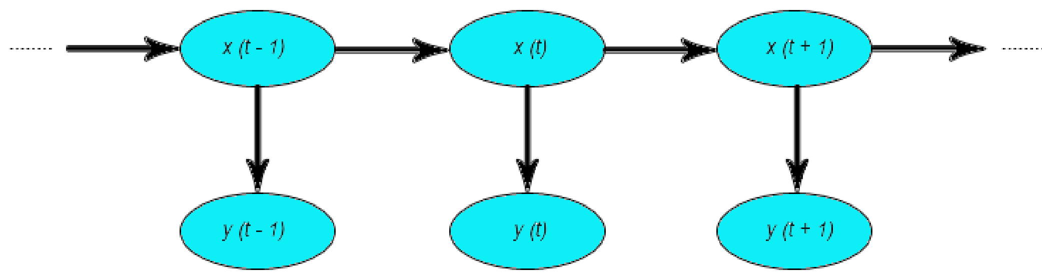

Figure 33. General architecture of Hidden Markov Model. Figure has been adapted from Singla's Master thesis (Singla, 2009)

Hidden Markov Models have been used in activity recognition studies (Singla, Cook, & Schmitter-Edgecombe, 2008; Kasteren, Noulas, Englebienne, & Krose, 2008). Moreover, many extensions are proposed, including layered hidden Markov models (Oliver, Horvitz, & Garg, 2002), quantitative temporal Bayesian networks (Colbry, Peintner, & Pollack, 2002), propagation networks (Shi, Huang, Minnen, Bobick, & Essa, 2004) and aggregate dynamic Bayesian models (Patterson, Fox, Kautz, & Philipose, 2005).

Tokuda et al. (Tokuda, Heiga, & Black, 2002) mentions three following limitations for Hidden Markov Model. 1) An HMM state relies on piece-wise constant statistics within an HMM state, 2) HMMs do not easily represent or reason about the duration of events.

3) Ensemble

Ensemble methods use multiple models to obtain better predictive performance than could be obtained from any of the constituent models (Opitz & Maclin, 1999). An ensemble is itself a supervised learning algorithm, because it can be trained and then used to make predictions. The trained ensemble, therefore, represents a single hypothesis. This hypothesis, however, is not necessarily contained within the hypothesis space of the models from which it is built. Thus, ensembles can be shown to have more flexibility in the functions they can represent. This flexibility can, in theory, enable them to over-fit the training data more than a single model would, but in practice, some ensemble techniques (especially bagging) tend to reduce problems related to over-fitting of the training data. As these algorithms conduct a weighted voting of many other algorithms, they usually perform better than many others.

For this study, we use implementations of some of the ensemble algorithms included in the Weka toolkit. These algorithms include AdaBoostM1, Bagging and Dagging. Bagging with LAD Tree as classifier achieved 97.33% accuracy, AdaBoostM1 with LADTree achieved 97.33% and Dagging with SMO achieved 95.33%.

4) Bayesian Network

A Bayesian network is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph (DAG). They are directed acyclic graphs whose nodes represent random variables in the Bayesian sense. Edges represent conditional dependencies; nodes which are not connected represent variables which are conditionally independent of each other. Each node is associated with a probability function that takes as input a particular set of values for the node's parent variables and gives the probability of the variable represented by the node.

Bayesian Network can be used to compute the conditional probability of one node, given values assigned to the other nodes; hence, it can be used as a classifier that gives the posterior probability distribution of the classification node given the values of other attributes. (Cheng & Greiner, 1999)

In theory, even approximate inference of probabilities in Bayesian networks can be NP-hard. For the special case of a polytree, has an efficient runtime of $O(Nq^{e})$, where e is the maximum number of parents on a vertex (MacKay, McEliece, & Cheng, 1998).

Fortunately, in practice approximate methods have been shown to be useful in many cases.

A Naive Bayes classifier is a simple probabilistic classifier based on applying Bayes' theorem with strong (naive) independence assumptions. A Naive Bayes classifier assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature. This classifier chooses as the output label class that yields the highest probability given the values of the data features. Probabilities are estimated based on frequency of data feature values and class labels in the available training data.

Training time complexity of Naïve Bayes is linear and essentially optimal. Test time complexity is also very efficient, linearly proportional to the time needed to just read in all the data.

5) Selected Model

All of the above models were evaluated on data from all sensor types and results shown in figure 34 illustrate that Bayesian Network perfoms best in classification of our activities. Although ensemble methods are usually expected to perform better than other models, but Bayesian Network still performs better in our experiments. In our model each feature is a child of the parent which is the activity label; as a result, Bayesian network is considering features as conditionally independent and it is working as a Naïve Bayes classifier.

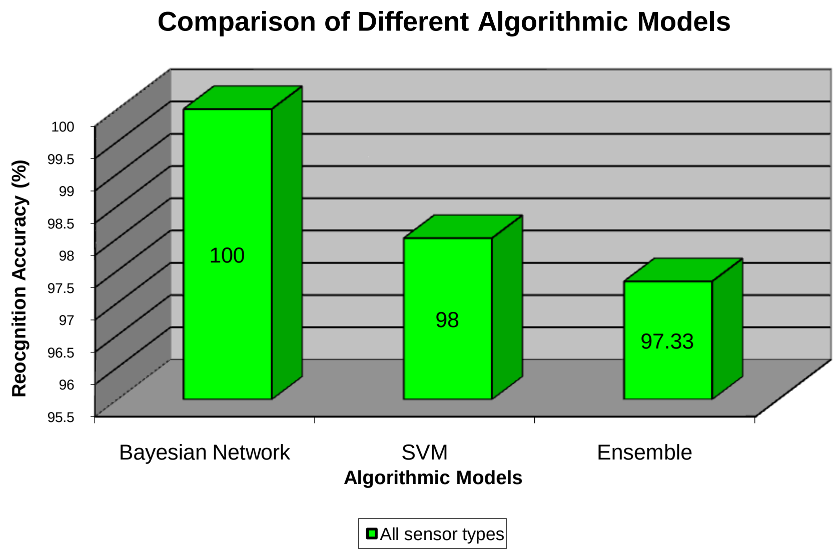

Figure 34. Illustrates comparison between different algorithmic models

HMM is usually used when dealing with accelerometer data because of the pattern-based nature of accelerometer data. That was the reason why we decided to run it on our wearable data. In order to use HMM for our data, we need to define a window size so that events in one window are considered as one new event. For each window we compute mean, standard deviation and correlation. As a result each event of new data has the following format. n shows the window number which is new data's event number.

Mean of window *n* (x, y, z, pitch, yaw, roll), std of window *n* (x, y, z, pitch, yaw, roll), correlations of window *n* (x/y, x/z, y/z, pitch/yaw, pitch/roll, yaw/roll)

Above features are calculated for both wearable sensors, as a result we have total of 36 features. We tried different window sizes. Results improved by decreasing the window size, which implies that HMM is not finding any particular pattern in data and it is not working as a pattern recognition model. This result is not very unexpected for this study, we believe because we have complex activities, data doesn't consist of repetitive and recognizable patterns.

Bayesian networks describe conditional independence among subsets of variables. And they allow combining prior knowledge about (in)dependencies among variables with observed training data. This model achieved the highest accuracy with 100%. As a result, all experiments in the rest of this thesis will use Bayesian Network as the classification algorithm.

5.2.4 Results

In this section, first we evaluate each sensor type individually and then combine classes of sensor types to recognize 6 chosen ADLs. Our purpose is to find a correlation between sensor modality and the targeted activity to recognize. Figure 33 illustrates the overall accuracy for each sensor modality using Bayesian Network classifier. Figures 34 through 41 visualize confusion matrix of each experiment. It shows how many samples of each activity were classified correctly and how many were classified as other activities.

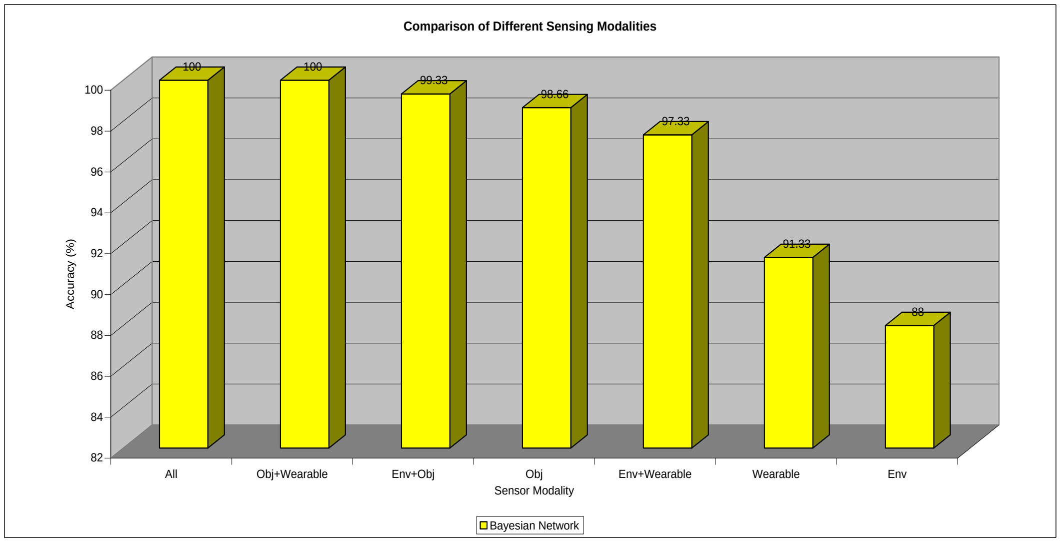

Figure 35. Comparison of different sensing modalities based on their overall accuracy in recognizing our 6 set of ADLs.

Figure 35 illustrates the results for activity recognition using only environmental sensors. *Sweeping & Dusting* and *Medication* have been classified with the same accuracy and 22 out of 25 samples were classified correctly for these activities. *Cooking* is classified correctly at all times. 23 samples out of 25 are correctly classified as *Watering Plants*. *Hand Washing* and *Countertop Washing* are classified with the same accuracy and they were the hardest activities to recognize with Environmental sensors.

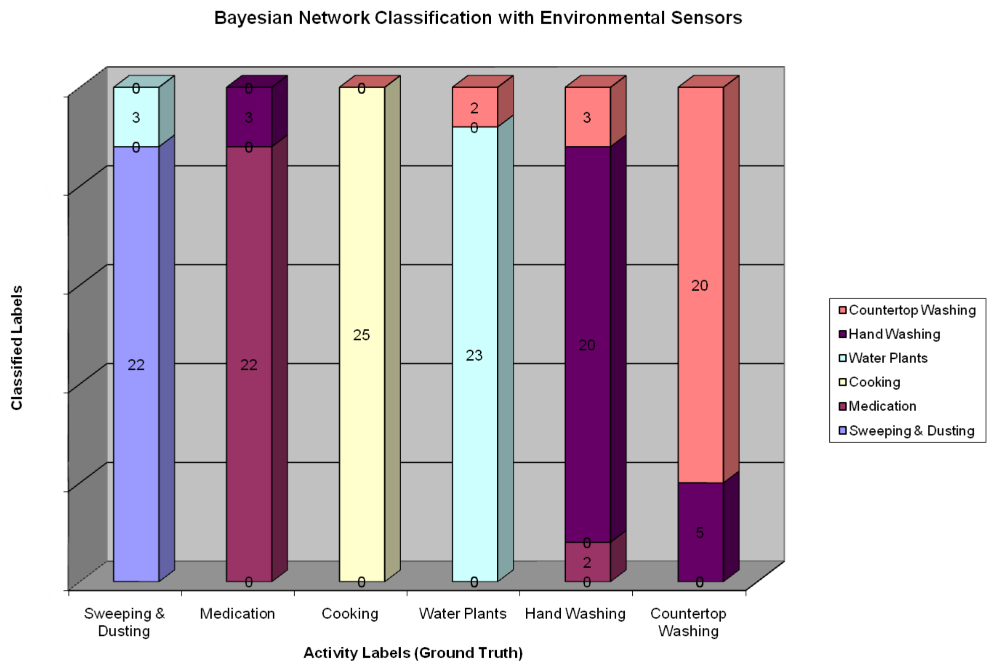

Figure 36. Bayesian Network classification with environmental sensors.

We believe that activities mentioned above that are classified with highest accuracy either had a unique pattern of changing locations or there was a contact switch sensor that helped in classification. For instance, in *Cooking,* participant spends some time near stove and then brings the food to dining room. This has a unique location movement that helps classifier algorithm to classify it with 100% accuracy. *Hand Washing* has been classified as *Medication* and *Countertop Washing.* We believe the reason is that these activities take place in almost the same location, near kitchen countertop. *Countertop Washing* was classified as *Hand Washing* 5 out of 25 times and we believe the reason is they both occur in exactly the same location and there is no environmental sensor other than motion sensors to differentiate between these activities. *Sweeping & Dusting* is classified as *Watering Plants* because they are the only activities that participant needs to visit dining room and all kitchen corners.

B) Object

This type of sensor modality achieved 98.67% accuracy in general. Although there was a one to one map between objects and activities because of the noise in real data we do not achieve 100% accuracy. Shake sensors have to be sensitive to shakes and vibrations and controlling all movements in a real home is not possible. For example, in our study, when the participant is shutting the cupboard door heavily, all of the shake sensors located in the cupboard are triggered. Or sometimes with heavier participants walking in the kitchen, sensors attached to items located on the floor are triggered. Figure 36 illustrates the location of each object sensor in our study.

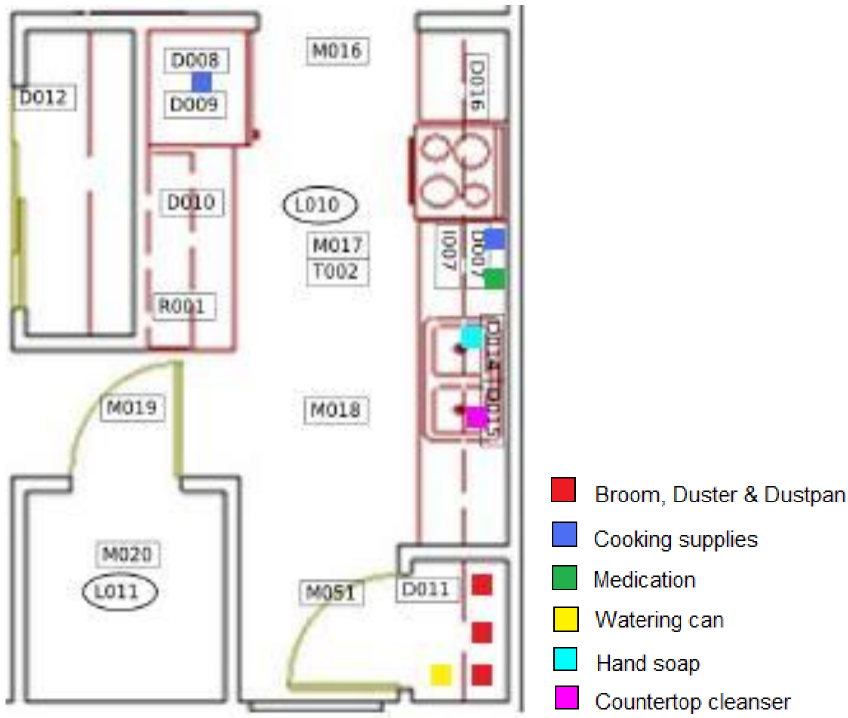

Figure 37. Kyoto kitchen layout with object
sensors' locations.

Figure 37 illustrates that *Medication* and *Washing Hands* have been classified with only one mistake. All other activities are classified correctly. *Medication* is classified as *Cooking* because both supplies are stored in the same closet and when participants close the door heavily, all of the objects in the cupboard get triggered. *Hand Washing* is classified as *Countertop Washing* in one sample, we believe because both hand soap and countertop cleanser are located on sink, using one might trigger the other as well.

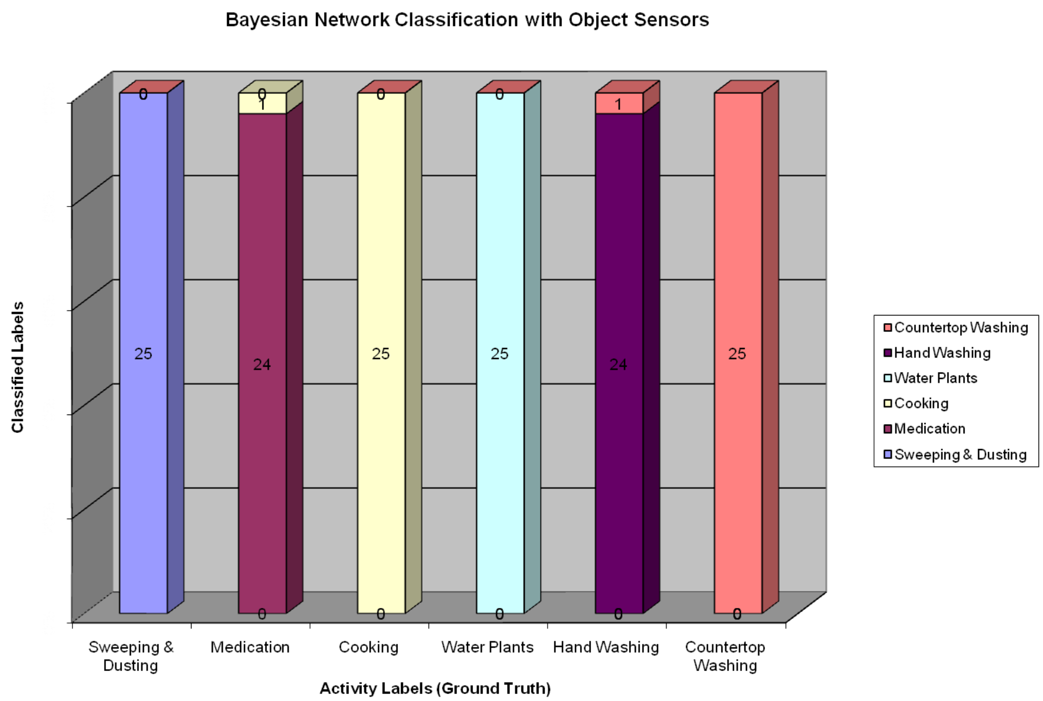

Figure 38. Bayesian Network classification with object sensors

C) Wearable

We expected wearable sensors to perform very poorly in this study. The reason for this is while performing complex activities, people don't move their hands at a particular range or speed. Patterns of movements are not easily distinguishable in complex activities. Moreover, as discussed in section 2.3.4, people often to include other, irrelevant subtasks while performing complex activities, which makes the classification problem even more challenging. In our study the wearable sensor alone performed better than we anticipated (Figure 38), and it outperformed Environmental modality. One factor that might have helped in this performance is the fact that different activities required different time durations. *Cooking* was the easiest activity to be recognized by our wearable sensing technology.

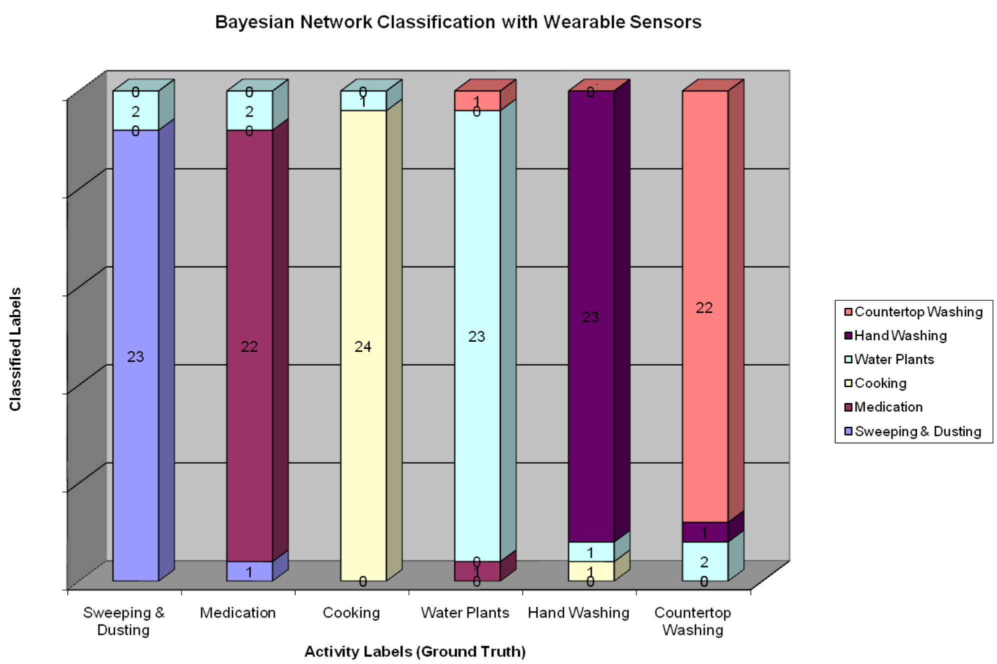

Figure 39. Bayesian Network classification with wearable sensors.

D) Environmental + Object

This modality performed very well overall with an average accuracy 99.33%, the second highest accuracy. Combining environmental and object sensors provides a rich amount of information needed for distinguishing most activities. Poorest performance is for *Hand Washing*. This activity was classified as

Countertop Washing, which is not unexpected. They take place in the same location and involved objects are affected by ambient vibration as we discussed earlier.

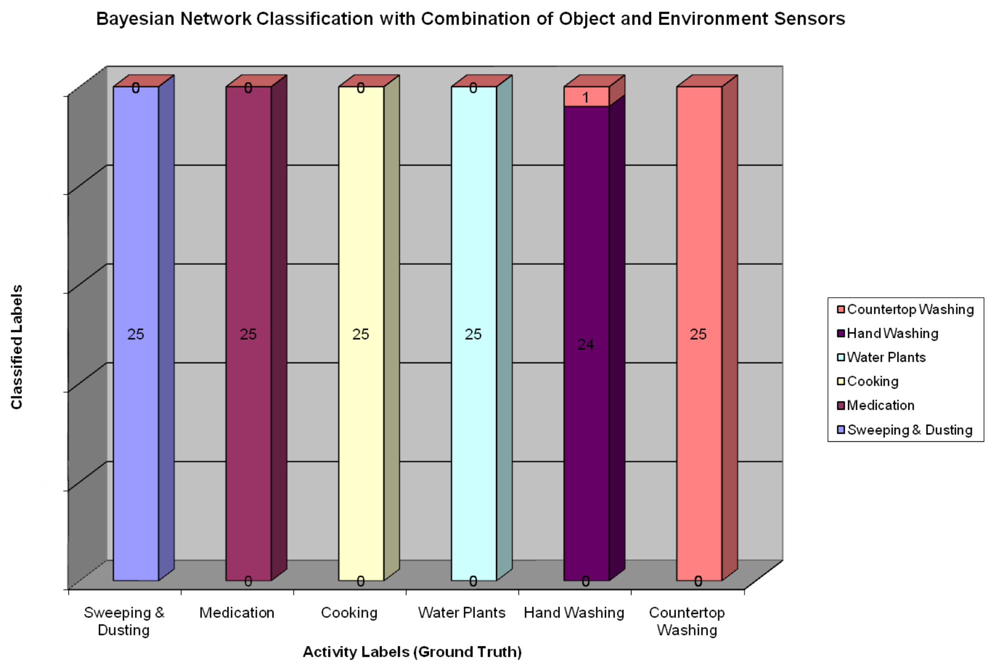

Figure 40. Bayesian Network classification with combination of environmental and objects sensors.

E) Environmental + Wearable

As Figure 33 illustrates, performance has been improved by combining these two sensing modalities, thus helpfulness of object sensors becomes apparent. Even though this model is still weaker than other combinations. Figure 40 illustrates *Watering Plants* is the most difficult activity to get recognized by this modality.

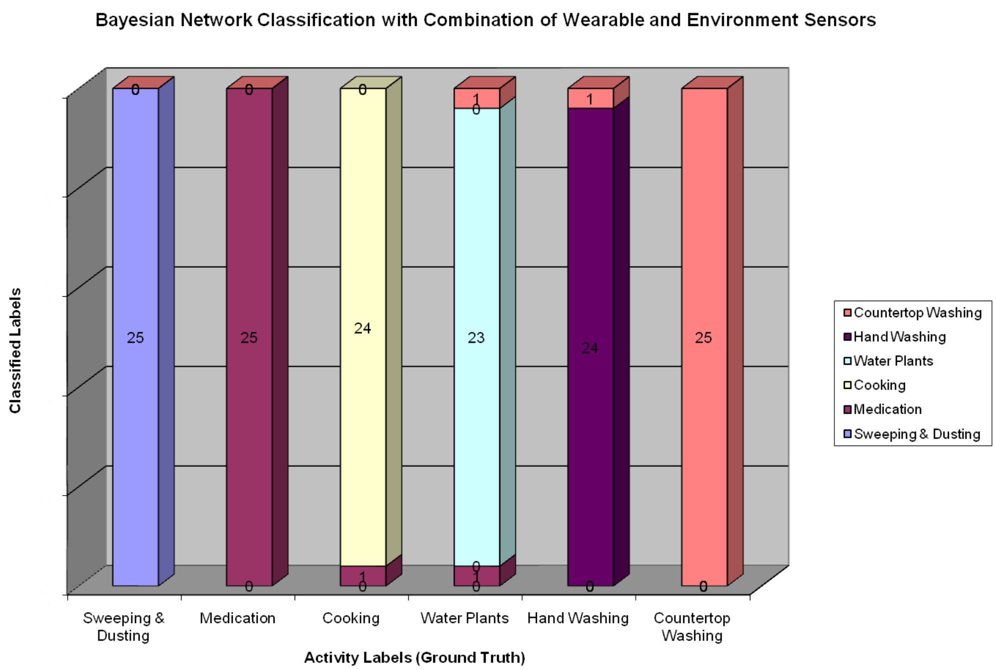

Figure 41. Bayesian Network classification with combination of wearable and environmental sensors.

F) Object + Wearable

This model achieved 100% accuracy.

G) All (Environmental + Object + Wearable)

This model and *Object + Wearable* model achieved the highest accuracy among all combinations that we considered, with 100% accuracy. This result is not far from our hypothesis, the more sensors you use, you have more information for training the algorithm and as a result you will get better results.

Moreover, if objects in interest are chosen correctly and used object sensors are reliable they can perform very well.

5.2.4 Discussion

As Figure 34 demonstrates, environmental sensors achieved the lowest accuracy (88%) among all other sensor modalities. As it can be seen in Figure 41 it performed very well for all activities except *Hand Washing* and *Washing Countertop*. The confusion between these two activities occurs because they have to be recognized with only motion detector and they happen at exactly the same location. As a result, we conclude that environmental sensors can perform very well for many activities as long as they don't take place in exactly the same location. If they do, there should be another sensor modality other than motion detector to provide additional information for classifying those activities. Alternatively, time of day and other contextual information may provide additional discriminating features for in-home naturalistic settings.

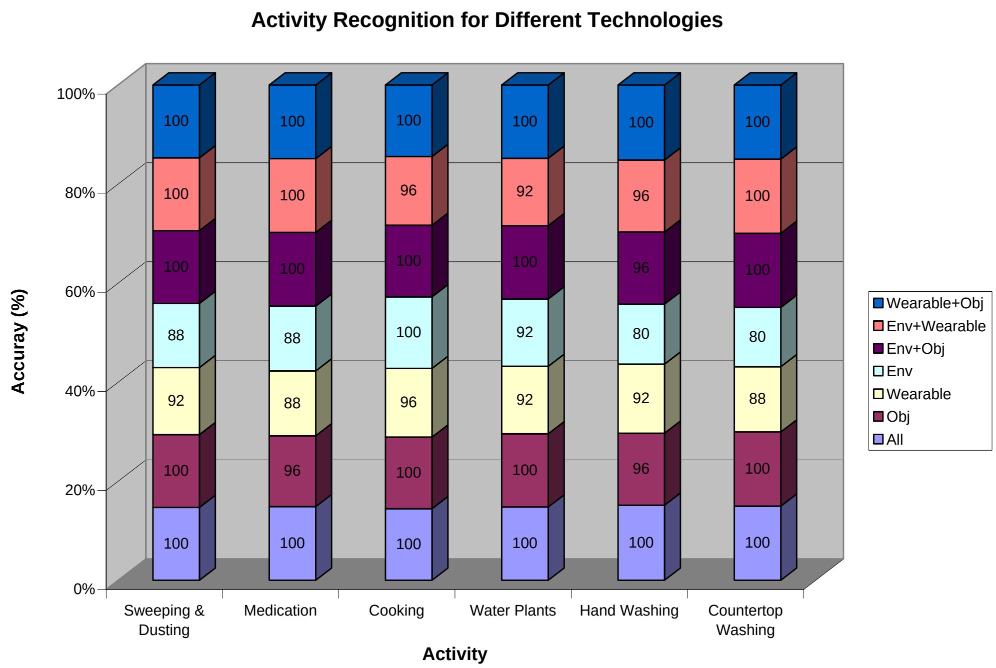

Figure 42. Activity recognition for different technologies per activity

The next lowest accurate modality after environmental was wearable sensors with 91.33% accuracy. Wearable sensors can be a good option for recognizing short movements of body parts such as raising hand, shaking hands, punching and so forth or for basic activities such as walking, running and sitting. But our results show that it can't perform very well with relatively long and complex activities such as our six selected ADLs. This sensor modality needs to be combined with other modalities. Figure 41 demonstrates that it has more difficulty recognizing *Cooking* and *Washing Countertops*.

The third lowest accuracy was achieved by a combination of environmental and wearable sensors with 97.33%. Figure 41 demonstrates that adding wearable sensors boosted environment sensors in recognizing most activities. This suggests that although wearable sensors do not perform very well in complex activities, combining them with other sensing technologies can be helpful in recognizing some activities.

Object sensors achieved the overall accuracy of 98.67%. As we discussed earlier having one to one mapping between objects and activities in addition to using a reliable sensing modality leads to achieving high accuracy.

Adding environmental sensors to object sensors helps in correctly classification of *Medication,* but it does not help in boosting the accuracy of *Washing Hands*. Because *Washing Hands* and *Countertop Washing* take place in the same location.

Sensor modality of object + wearable and combination of all sensor types both achieved 100% accuracy.

These results demonstrate that combining different sensing technologies outperforms using them separately. By adding low accurate modalities such as wearable or environmental sensors to others we can increase the performance of our sensing modalities. However, an important point is that all of the limitations of each technology should be taken into consideration while choosing the proper one for a study. There should be a tradeoff between accuracy and problems that using a particular sensing technology would make.

In addition, monitored activities play an important role in choosing the proper technology. For instance, if activities that we need to recognize do not have shared objects, object sensors would distinguish these activities quite well; or if the location movements are unique for each activity, environmental sensors would be a proper technology to use.

Sensor reliability is another important factor. Although shake sensors still have noise, but our results compared to studies with RFID tags mentions in section 2.2.1 suggests that they are more reliable than RFID technology.

In summary, our experimental results support our hypothesis and each sensing modality performs better with a subset of activities. Moreover, combining different sensing platforms can increase the overall accuracy, however, all limitations of each technology should be considered when selecting a sensing modality.

CHAPTER SIX

6. DEFINING THE COMPLEXITY OF AN ACTIVITY

6.1 Introduction

Activity recognition is a widely researched area with applications in health care, security and other domains. With each recognition system considering its own set of activities and sensors, it is difficult to compare the performance of these different systems. More importantly, it makes the task of selecting an appropriate set of technologies for recognizing an activity challenging. This chapter defines complexity measurements for activities and uses grammar complexity as the second approach. In addition, it uses grammars to formally represent different ADLs.

While sitting, standing, walking, etc., appear at one end of the spectrum of activities, the other end consists of complicated activities such as cooking and taking medication, which encompass ambulation, ADLs and instrumental ADLs (iADLs). From a computational standpoint, it is difficult to combine these different activities into a single category for the purpose of designing a recognition system. Having a standard way to classify these activities based on their complexities will help researchers in all fields who want to study activities. This is the primary motivation behind this paper, where we attempt to define a formal complexity measure for activities. The complexity of an activity can be defined in terms of different parameters such as the underlying sensing modality, the computational techniques used for recognition or inherent properties of the activity. We describe each of these parameters in greater detail. Defining such a complexity measure provides a means for selecting activities for conducting benchmarking experiments. Furthermore, it also helps in choosing the correct technology for recognizing a specific set of activities.

This study has used two approaches for activity complexity measurement. First, the complexity of an activity is defined in terms of three components: Sensing complexity, Computational complexity, and Performance complexity. Second, grammar complexity has been used as a measurement for complexity of an activity.

6.2 Defining Activity Complexity

In general, the complexity of an activity can be defined in terms of different factors. In this paper we attempt to define it in terms of three components: Sensing complexity, Computational complexity, and Performance complexity.

6.2.1 Sensing Complexity

Sensing complexity refers to complexity of sensors which are used in collecting data. Research advances in computing have resulted in the development of a wide variety of sensors that can be used for sensing activity.

On one hand there are sensors that have to be worn by individuals (Krishnan & Panchanathan, 2008) and on the other hand there are environmental and object sensors that have to be embedded in the environment for gathering activity related information (Singla, Cook, & Schmitter-Edgecombe, 2008). Each of these sensors provides a rich set of information on a certain set of activities. For example, as we discussed earlier in the thesis it is easier to recognize ambulation using wearable sensors over environmental sensors, while iADLs such as cooking and bathing are easier to recognize using environmental or object sensors. We define the sensing complexity of activities in terms of the following parameters: the number of distinct sensors fired, the number of sensor types fired, the number of objects involved to which a sensor can be attached, sensor size, sensor price, ease of use (Subject, Deployment), type of output data, battery life and type of sensor (wired or wireless). In the following paragraphs, we will discuss each of these parameters in more detail.

The number of sensors used is an important factor that defines this complexity, which in turn can be divided into two groups: the number of distinct sensors fired and the number of sensor types fired. For example, one particular sensor might be fired many times, but we count it as only one distinct sensor. Based on the technology used in each study, different sensor types can be seen, such as environmental sensors (motion, temperature, light, etc.), object sensors (RFID tags, accelerometers, shake sensors, etc.) and wearable sensors (accelerometers, RFID, health monitoring sensors, etc.). For example, if we are using environmental motion sensors, wearable accelerometers and shake sensors on objects, all three sensor types are fired during a cooking activity. For washing hands, however, only two of them are fired: environmental and wearable (assuming no sensor has been placed on the soap). The number of objects involved in an activity that can be sensed through some modality is another factor defining the sensing complexity. For some activities such as sweeping, placing sensors on the objects involved (broom) is possible, thus it can be considered simpler than reading books (placing a sensor on every book is impractical).

The price and form factor of a sensor is another component of the sensing complexity. An expensive sensor system would be harder to implement, so it can be considered more complex. The same is true with sensor size, especially for wearable and object sensors. Smaller sensors are easy to adopt, while bigger sensors are relatively difficult to deploy. The ease of use of a sensor can be seen from two perspectives: Subject and Deployment. Ease of use with respect to subject refers to ease of use and the level of acceptance with which participants will adopt and use sensors. For example, some wearable sensors may be easier and more comfortable for participants to wear. The deployment aspect of ease of use can be defined in terms of the ease with which experimenters install a particular sensor. A sensor might give us helpful data but working with it might be so hard for experimenters that they prefer alternative but less useful tools. This reasoning would be true about type of output of the sensor as well. Some sensor outputs need further complex computations and pre-processing which results in higher sensing complexity.

The battery life of a sensor is an important factor especially in the context of wireless and wearable systems. Choosing wired or wireless sensors depends on the requirements of the system and it has a corresponding effect on the sensing complexity.

While the values for some of these parameters (e.g., number of sensors, battery life) can be derived empirically, other factors (e.g., form factor and ease of use) require some kind of subjective evaluation. We would expect the measure derived from these parameters to be low for ambulatory activities for wearable sensors such as accelerometers, but will be high for environmental sensors such as motion sensors. In Table 6 we present some popular activities analyzed using these different factors.

6.2.2 Computational Complexity

Advances in machine learning and pattern recognition have resulted in a number of supervised and unsupervised techniques for recognizing activities. Discriminative classifiers such as SVMs (Krishnan & Panchanathan, 2008), Logistic regression (Krishnan & Panchanathan, 2008), CRFs (Nazerfard, Das, Holder, & Cook, 2010) and generative classifiers such as GMMs (Pansiot, Stoyanov, McIlwraith, Lo, & Yang, 2007), HMMs (Singla, Cook, & Schmitter-Edgecombe, 2008) are very popular for activity recognition. In addition to this, computational complexity also includes the algorithms that transform the raw data stream into a form that is used by some of the recognition algorithms. Examples of these algorithms are FFTs (Huynh & Schiele, 2005), wavelets, and other techniques that extract the statistical and spectral properties of the raw data. The main component of the computational complexity is the complexity of the underlying recognition/transformation algorithm. Other factors that affect the computational complexity include memory requirements of the algorithm and real-time performance. The relevance of the computational complexity of an activity depends on the computational resources available. For example, if the goal of the system is to perform recognition on a low power device such as mobile phone, the computational complexity plays an important role in selecting the appropriate set of algorithms.

6.2.3. Performance Complexity

We define the performance complexity to be an abstraction of some of the inherent properties of an activity that is independent of the underlying sensing and computational mechanisms. This complexity term can be defined using different parameters such as: average duration and deviation, duration of non repetitive patterns, number of activity steps, number of distinct location movements, and number of people and objects involved.

The average duration of an activity, even though an important component, does not clearly differentiate the complexity of activities. In other words there is no general rule that can say an activity with higher duration is more complex or vice versa. As an example, cooking is a relatively long and complex activity. At the same time sleeping is also long but not very complex from the perspective of recognition. Thus, this component should be taken into consideration along with other factors.

Perhaps one could look at how much time during the activity the person was active. For example, a person is not active for a large portion of time while sleeping and watching TV. Associated with the average duration of an activity is also the deviation in the duration in the performance of the activity.

The third component is the duration of non-repetitive patterns. Patterns that are inherent in activities give us useful information. Repetitive patterns are easier to recognize. For example, walking or running involve periodic movements of the human body that can be easily recognized, in contrast to movements such as pouring water, or scooping sugar while making a cup of tea. Some activities have a predefined time of occurrence during the daily routine of an individual. Such a unique characteristic of an activity can be effectively utilized by machine learning algorithms for recognition. An example of such an activity is taking medication. Many individuals take medicine at the same time each day or in the same context, such as while eating a meal.

Typically every activity is defined in terms of a number of steps. Some activities have a larger number of steps which make them more complex. An activity step can be defined as event that cannot be divided in to sub-events in the current technology. Defining the activity steps in this format facilitates different representations of the steps depending on the underlying technology. The next issue to be considered is the number of distinct location movements; an activity which is performed in different locations can be considered more complex in comparison with an activity that takes place in one location.

Other factors that define the performance complexity of an activity are the number of people and the number of objects involved in that activity. The activities get more complex with an increasing number of people and objects defining the activity.

6.3. Evaluating the Complexity

In Table 5 we represented 6 common activities and measured some of their complexity measurements that are discussed in this section. There are different ways to generate one total value from these measurements. One straight forward approach would be assigning numbers 1, 2, 3 to values low, medium and high respectively, and then summing up all the values for each activity. We can ignore the value of 'Number of people involved' in this case, since it is the same for all these activities. Following the above rules we will get 12 for 'cooking', 10 for 'sweeping', 9 for 'watering plants' and 'washing counter tops', 8 for 'hand washing' and 6 for 'medication'. Therefore, 'cooking' can be categorized as the most complex activity to recognize with this study's sensing technology and 'taking medication' as the easiest one.

For generating these examples we assumed sensing technology of WSU Center for Advanced Studies in Adaptive Systems (CASAS), which consists of three sensor types (environmental, wearable and object).

Table 5. Complexity measurement over activity based on WSU CASAS sensing technology

Activity	Number of objects involved that cannot put sensors on	Number of distinct sensors fired	Average Time	Duration deviation	Number of people involved	Has a predefined time?	Number of distinct location movements
Sweeping	Low	High	High	Medium	1	No	Low
Medication	Medium	Low	Low	Low	1	Yes	Low
Watering plants	Low	Medium	Medium	Low	1	No	Medium
Hand washing	Medium	Medium	Low	Low	1	No	Low
Washing kitchen countertops	Medium	Low	Medium	Medium	1	No	Low
Cooking	High	Medium	High	Medium	1	Yes	Low

6.4. Using Grammar Complexity

While the complexity values can be derived from pre-defined measures as described previously, another possible approach is making use of grammars for representing activities (Sahaf, Krishnan, & Cook, 2011). Grammar complexity can then be used for measuring the complexity of the corresponding activity. Using a grammar has different benefits. The grammar can formally define complex activities based on simple actions or movements. Rules are understandable by human. The grammar can also be extended and modified at any time and it can be used by systems with different technologies. In addition, grammars provides us with a formal representation of activities which helps researchers in different fields to have a benchmark while trying to choose and compare activities to monitor in their studies.

Researchers have used grammars for representing different activities. Ward et al. used wearable accelerometers and looked at wood workshop activities such as "grinding" and "drilling" (Ward, Lukowicz, Tr ¨oster, & Starner, 2005). Most of these studies use cameras for gathering data. For example, Ryoo and Aggarwal have defined grammars for activities such as "Shake hands", "Hug", "Punch", etc (Ryoo & Aggarwal, 2006). Chen et al. have used grammar in gesture recognition (Chen, Georganas, & Petriu, 2007). There are a few studies on using grammar for representing ADLs. In particular, Teixeira et al. has represented ADLs with hierarchical finite state machines (Teixeira, Jung, Dublon, & Savvides, 2009).

In other areas such as Human Computer Interaction (HCI), user tasks have been represented by means of task notations. A task defines how the user can reach a goal in a specific application domain. Paterno has defined CTT model which provides a rich set of operators to describe the temporal relationships among tasks and enables designers to describe concurrent tasks (Paternò, Mancini, & Meniconi, 1997). In addition, for each task, further information can be given; task is described by attributes including Name, Type (abstract, user, application, interaction), Subtask of, Objects, Iterative (a Boolean indicating whether the task is iterative), First action, and Last action.

Beyond these descriptive aspects, these notations can also be used to assess the complexity of the tasks. Palanque and Bastide have modeled tasks using the Interactive Cooperative Objects (ICO) formalism, which is based on Petri nets and on the object-oriented approach (Palanque & Bastide, 1970). In their quantitative analysis of task complexity they have considered the number of nodes (corresponding to the number of states in the task model) the number of actions (corresponding to the number of arcs with different labels) and the length of the path to come back to the initial state which are associated with weights.

To the best of our knowledge, no study has looked at grammar complexity to derive activity complexity. Different grammars such as CFG (Teixeira, Jung, Dublon, & Savvides, 2009), SCFG (Moore & Essa, 2001), DOP (Data Oriented

Processing), LFG (Lexical-functional Grammar) can be used for this purpose. In this study we will focus on Context-free Grammars, in which the left-hand side of each production rule consists of only one single non-terminal symbol, and the right-hand side is a string consisting of terminals and/or non-terminals. Human actions and interactions are usually composed of multiple sub-actions which themselves are atomic or composite actions and CFG is able to construct a concrete representation for any composite action (Ryoo & Aggarwal, 2006). On the other hand, context-free grammars are simple enough to allow the construction of efficient parsing algorithms (Chen, Georganas, & Petriu, 2007).

In this study we present a very simple CFG as a baseline for future work which can represent sequential behaviors. In order to define a CFG, we need to define terminals and non-terminals symbols. We can associate the atomic actions with the terminals and complex actions with non-terminal symbols. However, as discussed before, the definition of the atomic action can vary according to the underlying sensing technology. For example, if one is looking at walking patterns, atomic action can be each movement of legs and hands, if one is using accelerometers as the sensing modality. In contrast, in a study that only uses environmental sensors, moving from one part of the room to the other which results in triggering a new sensor is considered atomic. In this paper, we try to define a general definition in a way that any research study will be able to adopt it. Continuing with our previous discussion, we define an atomic action as an event that cannot be divided into smaller sub-events that is recognizable by the underlying sensing modality. If an action contains two or more atomic actions, it is classified as a composite action (Ryoo & Aggarwal, 2006). By using CFGs, we are able to define a composite action (Non-terminal) based on atomic actions (Terminals).

In order to formally represent an atomic action we follow the linguistic theory of "verb argument structure". Park's operation triplet is <agent-motion-target> (Park, Park, & Aggarwal, 2004), where agent refers to the body part (i.e. arm, head) directed toward an optional target. The motion set contains action atoms such as "stay", "move right", etc.

However, this triplet is too specific to their sensing technology which is using camera and image processing. As a more generic formal representation we define an atomic action as <agent – motion – location - target> where an agent is the person performing the action, motion represents the event of that atomic action which can be in any form based on the technology, location indicates the location of the event and target is the object or person in interaction. If the action does not contain any interactions, the target value will remain null. As an example, we chose two common activities and formalized them with this CFG scheme. The following examples focus on the 'Sweeping' and 'Dusting' activities. There is only one person involved in these activities which is represented by 'i'. Complex actions such as ‘Dusting’ is represented as ‘OR’ of two atomic actions ‘DustLivingRoom’ and ‘DustDiningRoom’. In order to generate these examples we assumed CASAS sensing technology which we have described before.

Sweeping:

RetrieveBroom(i) =

atomicAction(<i, RaiseHand, Near kitchen cupboard, Broom>)

SweepKitchenFloor(i) =

atomicAction(<i, Repetitive pattern & Raise, Kitchen, Broom>)

Sweep(i) -->

RetrieveBroom(i) and SweepKitchenFloor(i)

Dusting:

DustLivingRoom(i) =

atomicAction(<i, Repetitive pattern & Raise, Living room, Duster>)

DustDiningRoom(i) =

atomicAction(<i, Repetitive pattern & Raise, Dining room, Duster>)

Dusting(i) -->

DustLivingRoom(i) or DustDiningRoom(i)

RetrieveDuster(i) =

atomicAction(<i, RaiseHand, Near kitchen cupboard, Duster>)

DustRooms(i) -->

RetrieveDuster(i) and Dusting(i)

As discussed before, the complexity of each activity can be calculated from the corresponding grammar complexity. There are various ways for obtaining complexity of a grammar, such as considering nodes and edges of finite state machines for CFG grammars. Number of alphabets, rules, terminal and non-terminal symbols can be other factors in this regard.

6.5. Summary and Conclusion

In this chapter, we have defined the complexity of an activity using two approaches. First, we have proposed measurements along three dimensions sensing, computation and performance. We have illustrated some of the parameters that define each of these dimensions, and then categorized some of the popularly used ADLs using these measures. In addition, we propose to use grammars as a formal representation of activities and make use of grammar complexity for categorizing activities.

CHAPTER SEVEN

7. CONCLUSIONS AND FUTURE WORK

Most studies on activity recognition focus on enhancing recognition algorithms and evaluating recognition results under varying conditions. We believe that representing the strengths and limitations of different sensor types is an important point that has not been addressed adequately in the literature. In this study, we presented different sensing technologies that can be used for activity recognition, particularly for indoor activities. We discussed the positive points and the limitations of each technology. Moreover, we evaluated different sensing technologies on data gathered from the CASAS smart home. Our experiments show there is a close relationship between each sensing technology and a subset of activities that can be recognized best with that technology. These results can be beneficial for researchers who want to select the proper sensing modality for their study. In the future, more sensor types can be evaluated using more activities.

In the second part of this work, we presented two approaches for estimating the complexity of an activity. In the first approach we defined features along three dimensions and categorized some of the popularly used ADLs using these measures. In the second approach we proposed using grammars for representing activities and made use of grammar complexity for classifying ADLs.

In the future we intend to use wearable sensor data for recognizing actions that are part of activities as opposed to recognizing the whole complex activity. Then, we anticipate using this result in the grammar representation model that we presented in second part of this study. Defining grammars for activity recognition has the benefit of making the process understandable by humans, adaptable to different technologies, and easily editable. In the future, more complex grammar types, such as grammars that can represent parallel activities can be used.

APPENDIX A

A. List of Activities of Daily Living (ADLs)

Cleaning the Living Area
Vacuuming the floor
Dusting

Cleaning the Bedroom
Making the bed
Vacuuming the floor
Dusting
Changing the sheets

Cleaning the Bathroom
Cleaning the bathroom floors and walls
Cleaning the sink, tub and toilet

Cleaning the Kitchen
Washing dishes
Drying dishes
Cleaning the outside of the stove
Cleaning the oven
Gathering and taking out the trash
Cleaning inside/outside of refrigerator
Sweeping floor
Mopping floor
Washing countertop

Home maintenance
Cleaning the windows
Cleaning the ceiling fans
Replacing light bulbs
Replacing the batteries in smoke detectors

Laundry
Sorting laundry
Loading/unloading washer/drier
Folding laundry
Hand washing
Ironing
Meal Preparation and Eat

Eating
Drinking
Preparing food
Preparing drink

Grooming Personal Hygiene and Dressing
Bathing
Hand Washing
Combing hair
Brushing teeth
Flossing teeth
Using mouthwash
Shaving
Trimming nails
Outfit Selection
Putting on clothes

Pet Care
Feeding/watering pet
Walking pet
Cleaning birdcage or fish tank
Cleaning the kitty litter box or dog pen
Bathing pet
Grooming pet

General
Transferring in and out of bed
Taking medication
Watching DVD/TV
Statement Filling
Reading book/magazine
Writing checks/birthday cards
Filling picnic Basket
Use of the telephone
Working on computer

BIBLIOGRAPHY

Rogers, P. J. (2005). Logic models. *Sandra Mathison (ed) Encyclopedia of Evaluation* (p. 232). Beverly Hills, CA: Sage Publications.

Abowd, R. J. (2000). The smart floor: A mechanism for natural user identification and tracking. *Conference on Human Factors in Computing Systems, ed , 275276. ACM.* New York, NY.

Al-ani, T., Ba, Q. T., & Monacelli, E. (2007). On-line automatic detection of human activity in home using wavelet and hidden markov models scilab toolkits. *CCA 2007. IEEE International Conference on Control Applications*, (pp. 485–490).

Aoki, P., Syzmanski, P., Thornton, J., Wilson, D. H., & Woodruff, A. (2003). The mad hatter's cocktail party: A social mobile audio space that supports multiple simultaneous conversations. *In Human Factors in Computing Systems (CHI).*

Bahl, P., & Padmanabhan, V. (2000). RADAR: An in-building RF-based user location and. *In Proceedings of IEEE INFOCOM.*

Bajers, F., & Moeslund, T. B. (1999). Computer Vision-Based Human Motion Capture - A Survey. *Computer Vision-Based Human Motion Capture - A Survey.*

Bao, L., & Intille, S. S. (2004). Activity Recognition from User-Annotated Acceleration Data. *Pervasive Computing* , pp. 1-17.

Bennewitz, M., Burgard, W., & Thrun, S. (2002). Learning motion patterns of persons for mobile service robots. *In Proceedings of ICRA.*

Bouchard, B., Giroux, S., & Bouzouane, A. (2006). *A Logical Approach to ADL Recognition for Alzheimer's patients.*

Brody, M. L. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* .

Bucceri., R. N. (2004). *How to Automate Both New & Existing Homes.* Silent Servant,Inc.

Bulling, A., Ward, J. A., Gellersen, H., & Troster, G. (2009). Eye movement analysis for activity recognition. *InProceedings of the 11th international Conference on Ubiquitous Computing* (pp. 41-50). Orlando, Florida, USA: Ubicomp '09. ACM.

CASAS shared smart home datasets repository. (2011). Retrieved from WSU CASAS: http://casas.wsu.edu/datasets.html

Chambers, G. S., Venkatesh, S., West, G. A., & Bui, H. H. (2002). Hierarchical recognition of intentional human gestures for sports video annotation. *In Proceedings of the 16th International Conference on Pattern Recognition* (pp. 2:1082–1085). IEEE Press.

Charniak, E., & Goldman, R. P. (1993). A Bayesian model of plan recognition. *In Proc. of the National Conference on Artificial Intelligence (AAAI).*

Chen, C., & Cook, D. (2011). Energy Outlier Detection in Smart Environments. *Proceedings of the International Workshop on Smarter Living: Conquest of Complexity (AAAI workshop).*

Chen, C., Das, B., & Cook, D. (2010). Energy Prediction Based on Resident's Activity. *Proceedings of the International Workshop on Knowledge Discovery from Sensor Data (KDD workshop).*

Chen, Q., Georganas, N., & Petriu, E. (2007). Real-time vision-based hand gesture recognition using haar-like features. *In Proc. of the IEEE Instrumentation and Measurement Technology Conf*, (pp. 1-6).

Colbry, D., Peintner, B., & Pollack, M. E. (2002). *Execution monitoring with quantitative temporal Bayesian networks.* In 6th International Conference on AI Planning and Scheduling.

Consolvo, S., Roessler, P., Shelton, B. E., LaMarc, A., Schilit, B., & Bly, S. (2004). Technology for care networks of elders. *IEEE Pervasive Computing Magazine: Successful Aging. 3(2):22–29.*

Crandall, A. S. (2011). Behaviometrics for multiple residents in a smart environment. *Ph. D. Thesis, Washington State University* .

Ermes, M., Prkka, J., Mantyjarvi, J., & Korhonen, I. (2008). Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. *IEEE Transactions on Information Technology in Biomedicine.*

Essa, I. A., Yin, P., Criminisi, A., & Winn, J. (2010). Bilayer Segmentation of Webcam Videos Using Tree-Based Classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence* .

Foerster, F., Smeja, M., & Fahrenberg, J. (1999). Detection of posture and motion by accelerometry: a validation in ambulatory monitoring. *Computers in Human Behavior* , 15:571–583.

Garrod, R., Bestall, J. C., Paul, E. A., Wedzicha, J. A., & Jones, P. W. (2000). Development and validation of a standardized measure of activity of daily living in patients with severe COPD: the London Chest Activity of Daily Living scale (LCADL). *Respir Med* .

Gavrila, D. M. (1999). The visual analysis of human movement: a survey. *Computer Vision and Image Understanding* , Vol. 73, pp. 82-98.

Goldman, R. P., & Charniak, E. (1993). A Bayesian model of plan recognition. *In Proc. of the National Conference on Artificial Intelligence (AAAI).*

Hahnel, D., Philipose, M., Fishkin, K. P., Perkowitz, M., Patterson, D. J., Fox, D., et al. (2004). Inferring Activities from Interactions with Objects. *Pervasive Computing* , 50-57.

Hawes, C., Phillips, C. D., Rose, M., Holan, S., & Sherman, M. (2003). *A National Survey of Assisted Living Facilities.* The Gerontological Society of America.

He, W., Sengupta, M., Velkof, V., & DeBarros, K. (2006). *Current Population Reports.* U.S. Department of Health and Human services.

Helal, S., & Mann, W. (2002). Smart Phones for the Elders: Boosting the Intelligence of Smart Homes. *Proc. AAAI Workshop Automation as Caregiver: The Role of Intelligent Technology in Elder Care* (pp. 74–79). AAAI Press.

Hindmarch, I., Lehfeld, H., de Jonge, P., & Erzigkeit, H. (1998). The Bayer Activities of Daily Living Scale (B-ADL). *Dementia and Geriatric Cognitive Disorders* , 9(suppl2):20–26.

Hodges, M. R., Newman, M. W., & Pollack, M. E. (n.d.). Object-Use Activity Monitoring: Feasibility for People with Cognitive Impairments.

Hoey, J., Bertoldi, A. V., Poupart, P., & Mihailidis, A. (2007). Assisting persons with dementis during handwashing using a partially observable Markov decision process. *Proc. Int. Conf. on Vision Systems.*

Huber, M. J., Edmund, H. D., & Wellman, M. P. (1994). *The automated mapping of plans for plan recognition.* In Proc. of the National Conference on Artificial Intelligence (AAAI).

Huynh, T., & Schiele, B. (2005). Analyzing features for activity recognition. *In: Proc. Soc-EUSAI 2005. ACM Int. Conf. Proceeding Series,* (pp. 159–163). ACM Press.

Intille, S. S., Bao, L., Tapia, E. M., & Rondoni, J. (2004). Acquiring in situ training data for context-aware ubiquitous computing applications. *Proceedings of CHI 2004 Connect: Conference on Human Factors in Computing Systems.* ACM Press.

Iso, T., & Yamazaki, K. (2006). Gait Analyzer Based on a Cell Phone With a Single Three-axis Accelerometer. *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services,* (pp. 141-144).

Jebara, T., & Pentland, A. (1999). Action Reaction Learning: Automatic Visual Analysis and Synthesis of Interactive Behaviour. *In ICVS '99: Proc. of the First Intl. Conf. on Computer Computer Vision Systems* (pp. 273–292). London, UK: Springer-Verlag.

K. Tokuda, H. Z. (2002). An HMM-based speech synthesis system applied to English. . *Proc. of IEEE 2002 Workshop on Speech Synthesis.* Santa Monica, USA.

Kasteren, T. V., Englebienne, G., & Kr¨ose, B. (2008). Recognizing activities in multiple contexts using transfer learning. *AAAI Fall 2008 Symposium: AI in Eldercare,* (pp. Washington DC, USA). Washington DC, USA.

Kasteren, T. V., Noulas, A. K., Englebienne, G., & Krose, B. J. (2008). Accurate activity recognition in a home setting. *Proceedings of the 10th international conference on Ubiquitous computing*, (pp. 1-9). New York, NY, USA.

Katz, S., Ford, A., Moskowitz, R., Jackson, B., & Jaffe, M. (1963). Studies of illness in the aged. *JAMA 185* , p. 914-919.

Kautz, H. (1987). *A Formal Theory of Plan Recognition.* PhD thesis, University of Rochester.

Kern, N., Schiele, B., & Schmidt, A. (2003). Multi-sensor activity context detection for wearable computing. *Ambient Intelligence* .

Kiani, K., Snijders, C. J., & Gelsema, E. S. (1998). Recognition of daily life motor activity classes using an artificial neural network. *Arch Phys Med Rehabil* .

Kim, H. H., K. N. Ha, K. N., Lee, S., & Lee, K. C. (2009). Resident location-recognition algorithm using a Bayesian classifier in the PIR sensor-based indoor location-aware system. *Trans. Sys. Man Cyber Part C 39, 2* , 240-245.

Krishnan, N. C. (2010). A Computational Framework for Wearable Accelerometer-Based. *PhD Thesis, Arizona State University* .

Krishnan, N. C., & Panchanathan, S. (2008). Analysis of low resolution accelerometer data for continuous human activity recognition. *IEEE International Conference on Acoustics, Speech and Signal Processing.*

Kusznir, J. (2009). CLM as a smart home middleware. *Master's thesis, Washington State* .

Kwapisz, J. R., Weiss, G. M., & Moore, S. A. (2010). Activity recognition using cell phone accelerometers. *Proceedings of the Fourth International Workshop on Knowledge Discovery from Sensor Data*, (pp. 10-18).

Lafferty, J., McCallum, A., & Pereira, F. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. *Intl. Conf. on Machine Learning.*

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* , 9:179–186.

Lehfeld, H., Jongh, P. d., Erzigkei, H., & Hindmarch, I. (1998). The Bayer Activities of Daily Living Scale (B-ADL). *Dementia and Geriatric Cognitive Disorders* , 9(suppl2):20–26.

Lesh, N., & Etzioni, O. (1995). A sound and fast goal recognizer. *Proceedings of the International Joint Conference on Artificial Intelligence.*

Lester, J., Choudhury, T., & Borriello, G. (2006). A practical approach to recognizing physical activities. *Pervasive Computing* , 116.

Liao, L., Fox, D., & Kautz, H. A. (2005). Location-Based Activity Recognition using Relational Markov Networks. *Nineteenth International Joint Conference on Artificial Intelligence (IJCAI 2005).* Edinburgh, Scotland.

Lorincz, K., & Welsh, M. (2007). MoteTrack: a robust, decentralized approach to RF-based location tracking. *Personal and Ubiquitous Computing* .

MacKay, D. J., McEliece, R. J., & Cheng, J. F. (1998). Turbo decoding as an instance of pearl's belief propagation algorithm. *IEEE Journal on Selected Areas in Communications* , 140-152.

Manabe, K., Matsui, T., Yamaya, M., Sato-Nakagawa, T., Okamura, N., & Ari, H. (2000). Sleeping patterns and mortality among elderly patients in geriatric hospitals. *Gerontology* .

Mann, W., & Helal, S. (2002). Smart Phones for the Elders: Boosting the Intelligence of Smart Homes. *Proc. AAAI Workshop Automation as Caregiver: The Role of Intelligent Technology in Elder Care* (pp. 74–79). AAAI Press.

Mantyjarvi, J., Himberg, J., & Seppanen, T. Recognizing human motion with multiple acceleration sensors. *In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics* (pp. 747–52). IEEE Press.

Marek, K. D., & JRantz, M. (2000). Aging in place: A new model for long-term care. *Nursing Administration Quarterly* .

Mase, S. L. (2002). Activity and location recognition using wearable sensors. *Pervasive Computing, IEEE* .

Mathie, M., Coster, A., Lovell, N., & Celler, B. (2003). Detection of daily physical activities using a triaxial accelerometer. *Medical and Biological Engineering and Computing* , vol. 41, no. 3, pp. 296–301.

Maurer, U., Smailagic, A., Siewiorek, D., & Deisher, M. (2006). Activity recognition and monitoring using multiple sensors on different body positions. *Proc Int Workshop Wearable Implantable Body Sensor Netw*, (pp. 99–102).

McDowell, I., & Newell, C. (1996, 2nd edition). *Measuring Health: A Guide to Rating Scales and Questionnaires.* New York: Oxford University Press .

Meghan, M. G. (2002). Activities of Daily Living Evaluation. *Encyclopedia of Nursing & Allied Health. Ed. Gale Group, Inc.*

Messing, R., Pal, C., & Kautz, H. (2009). Activity recognition using the velocity histories of tracked keypoints. *ICCV "09: Proceedings of the Twelfth IEEE International Conference on Computer Vision.*

Meyer, G. (2005). *Smart Home Hacks: Tips & Tools for Automating Your House.* O'Reilly Media, Inc.

Mihailidis, A., Carmichael, B., & Boger, J. (2004). The Use of Computer Vision in an Intelligent Environment to Support Agingin- Place, Safety, and Independence in the Home. *IEEE Transactions on Information Technology in BioMedicine.*

Mitchell, T. M. (1997). *Machine Learning.* New York: McGraw-Hill.

Moeslund, T. B., & Bajers, F. (1999). Computer Vision-Based Human Motion Capture - A Survey. *Computer Vision-Based Human Motion Capture - A Survey* .

Moore, D., & Essa, I. (2001). Recognizing multitasked activities using stochastic context-free grammar. *In CVPR Workshop on Models vs Exemplars in Computer Vision.*

Mozer, M. C., Pashler, H., Wilder, M., Lindsey, R., Jones, M. C., & Jones, M. N. (2010). Decontaminating human judgments to remove sequential dependencies. *In J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel,*

& A. Culota (Eds.), Advances in Neural Information Processing Systems , (pp. 23 (pp. 1705-1713).

Mutlu, B., Krause, A., Forlizzi, J., Guestrin, C., & Hodgins, J. (2007). Robust, low-cost, non-intrusive sensing and recognition of seated postures. *In Proc. UIST*, (pp. 149–158).

Mynatt, E. D., & Rowan, J. (2005). Digital family portrait field trial: Support for aging in place. *In Proceedings of the SIGCHI Conference on Human Factors in*, (pp. 521 – 530).

Nazerfard, E., Das, B., Holder, L. B., & Cook, D. J. (2010). Conditional Random Fields for Activity Recognition in Smart Environments. *Proceedings of IHI.*

Newell, C., & McDowell, I. (1996). *Measuring Health: A Guide to Rating Scales and Questionnaires.* 2nd edition. New York: Oxford University Press.

Oliver, N., Horvitz, E., & Garg, A. (2002). Layered representations for human activity recognition. *Fourth IEEE Int. Conf. on Multimodal Interfaces*, (pp. 3-8).

Opitz, D., & Maclin, R. (1999). Popular ensemble methods: An empirical study. *Journal of Artificial Intelligence Research* , 11: 169–198.

Orr, R. J., & Abowd, G. D. (2000). The smart floor: A mechanism for natural user identification and tracking. *Conference on Human Factors in Computing Systems.* New York, NY: ed , 275276. ACM.

Pacifica, T. (2005). *Easy X10 Projects For Creating a Smart Home.* Indy-tech Publishing.

Padmanabhan, V., & Bahl, P. (2000). RADAR: An in-building RF-based user location and tracking system. *In Proceedings of IEEE INFOCOM.*

Palanque, P., & Bastide, R. (1970). Synergistic modelling of tasks, system and users using formal specification techniques. *Interacting With Computers, Academic Press* , 129-153 .

Pansiot, J., Stoyanov, D., McIlwraith, D., Lo, B., & Yang, G. Z. (2007). Ambient and Wearable Sensor Fusion for Activity Recognition in Healthcare Monitoring Systems. *In Proc. of BSN 07.*

Paradiso, R., Loriga, G., & Taccini, N. (2005). A wearable health care system based on knitted integrated sensors. *IEEE Transactions on Information Technology in Biomedicine* , 9:337--344.

Park, J., Park, S., & Aggarwal, J. K. (2004). Model-based human motion tracking and behavior recognition using hierarchical finite state automata. *Lecture Notes in Computer Science, Proceedings of ICCSA.*

Park, S., & Aggarwal, J. K. (2004). Semantic-level Understanding of Human Actions and Interactions using Event Hierarchy. *Computer Vision and Pattern Recognition* .

Patel, S. N., Kientz, J. A., Jones, B., Price, E., Mynatt, E. D., & Abowd, G. E. (2007). An Overview of the Aware Home Research Initiative at the Georgia Institute of Technology. *In the Proceedings of the International Future Design Conference on Global Innovations in Macro- and Micro-Environments for the Future*, (pp. 169-181.).

释Paternò, F., Mancini, C., & Meniconi, S. (1997). ConcurTaskTrees: A Diagrammatic Notation for Specifying Task Models. *In: Proc. of Interact'97. Chapman & Hall*, (pp. 362-369.).

Patterson, D. J., Fox, D., Kautz, H. A., & Philipose, M. (2005). Fine-Grained Activity Recognition by Aggregating Abstract Object Usage. *IEEE International Symposium on Wearable Computers* , 44--51.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann.

Philipose, M., Fishkin, K. P., Perkowitz, M., Fox, D., Kautz, H., & Hahnel, D. (2004). Inferring Activities from Interactions with Objects. *Pervasive Computing* .

Pineau, J., Montemerlo, M., Roy, N., Thrun, S., & Pollack, M. (2003). Towards robotic assistants in nursing homes: Challenges and results. *Robotics and Autonomous.*

Platt, J. (1998). *Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines.* Technical Report MSR-TR-98-14, Microsoft Research.

Rantz, K. D. (2000). Aging in place: A new model for long-term care. *Nursing Administration Quarterly* .

Ravi, N., Dandekar, N., Mysore, P., & Littman, M. L. (2005). Activity recognition from accelerometer data. *AAAI* , 1541–1546.

Richardson, B., Leydon, K., Fernstrˆm, M., & Paradiso, J. (2004). Z-Tiles: building blocks for modular, pressure-sensing. *Extended Abstracts of the 2004 conference on Human factors and computing systems*, (pp. 1529-1532). Vienna, Austria.

Rivera-Illingworth, F., Callaghan, V., & Hagras, H. (2005). A neural network agent based approach to activity detection in AmI environments. *IEE International Workshop, Intelligent Environments.* Colchester, UK.

Rogers, W. A., Meyer, B., Walker, N., & Fisk, A. D. (1998). Functional limitations to daily living tasks in the aged: a focus groups analysis. *Human Factors* , 40:111–125.

Rowan, J., & Mynatt, E. D. (2005). Digital family portrait field trial: Support for aging in place Computing Systems. *In Proceedings of the SIGCHI Conference on Human Factors in*, (pp. 521 – 530).

Rugnone, A., Poli, F., Vicario, E., Nugent, C. D., Tamburini, E., & Paggetti, C. (2007). *A Visual Editor to Support the Use of Temporal Logic for ADL Monitoring.*

Russell, S., & Norvig, P. (2003). *Artificial Intelligence: A Modern Approach.* Pearson Education Inc.

Ryoo, M. S., & Aggarwal, J. K. (2006). Recognition of Composite Human Activities through Context-Free Grammar based Representation. *IEEE Conference on Computer Vision and Pattern Recognition.*

Sahaf, Y., Krishnan, N. C., & Cook, D. J. (2011). Defining the Complexity of an Activity. *Proceedings of the AAAI Workshop on Activity Context Representation: Techniques and Languages.* San Francisco.

Saint-Andre, P. (2004). *Extensible messaging and presence protocol (XMPP): Core.* RFC 3920 (Proposed Standard).

Scollon, C. N., Kim-Prieto, C., & Diener, E. (2003.). Experience sampling: promises and pitfalls, strengths and weaknesses. *J. Happiness Stud 4:* , 5–34.

Shi, Y., Huang, Y., Minnen, D., Bobick, A., & Essa, I. (2004). Propagation networks for recognition of partially ordered sequential action. *In Proc. of* *the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR).*

Singla, G. (2009). Recognizing ADL initiation and completion in a smart home. *Masters Thesis, Washington State University* .

Singla, G., Cook, D., & Schmitter-Edgecombe, M. (2008). Incorporating temporal reasoning into activity recognition for smart home residents. *Proceedings of the AAAI Workshop on Spatial and Temporal Reasoning*, (pp. pages 53-61).

Srinivasan, R., Chen, C., & Cook, D. (2010). Activity recognition using actigraph sensor . *Proceedings of the International workshop on Knowledge Discovery from Sensor Data.*

Tapia, E. M. (2008). Using Machine Learning for Real-time Activity Recognition and Estimation of Energy Expenditure. *Ph.D. Thesis, Massachusetts Institute of Technology* .

Tapia, E. M., Intille, S. S., & Larson, K. (2004). Activity Recognition in the Home Setting Using Simple and Ubiquitous Sensors. *Pervasive computing* , 158–175.

Teixeira, T., Jung, D., Dublon, G., & Savvides, A. (2009). Recognizing activities from context and arm pose using finite state machines. *Third ACM/IEEE International Conference on Distributed Smart Cameras. ICDSC.*

Thomas, B., & Crandall, A. (2011). A demonstration of PyViz, a flexible smart home visualization tool. *Proceedings of the IEEE Conference on Pervasive Computing and Communication.*

Tokuda, K., Heiga, z., & Black, A. W. (2002). An HMM-based speech synthesis system applied to English. *Proceedings of 2002 IEEE Workshop on Speech Synthesis*, (pp. 227 - 230).

Vail, D. L., Veloso, M. M., & Lafferty, J. D. (2007). Conditional random fields for activity recognition. *In AAMAS* , 1331–1338.

Want, R. (2003). A Key to Automating Everything. *Scientific American* .

Ward, J., Lukowicz, P., Tr ¨oster, G., & Starner, T. (2005). Activity recognition of assembly tasks using body-worn microphones and accelerometers. *In PAMI* , vol. 28(10).

Wilson, D. H. (2005). Assistive Intelligent Environments for Automatic Health Monitoring. *Ph.D. Thesis, Robotics Institute, Carnegie Mellon University* .

Winston, P. H. (1992). *Artificial Intelligence 3rd ed.* Addison Wesley Publishing company.

Yin, P., Criminisi, A., Winn, J., & Essa, I. A. (2010). Bilayer Segmentation of Webcam Videos Using Tree-Based Classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence* .

York, J., & Pendharkar, P. C. (2004). Human–computer interaction issues for mobile computing in a variable work context. *Int. J. Human-Computer Studies 60* , 771–797.

Zazula, D., & Sprager, S. (2009). Gait identification using cumulants of accelerometer data. *The 2nd WSEAS International Conference on Sensors, and Signals and Visualization, Imaging and Simulation and Materials Science,.*