Activity Recognition in the Home Setting Using Simple and Ubiquitous Sensors

by
Emmanuel Munguia Tapia
Bachelor of Science
Instituto Politécnico Nacional
Escuela Superior de Ingeniería Mecánica y Eléctrica
Mexico DF, Mexico
2000

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning in partial fulfillment of the requirements for the degree of
Master of Science
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
September 2003

Author . . . .
Program in Media Arts and Sciences
August 15, 2003

Certified by . . . .
Kent Larson
Principal Research Scientist
MIT Department of Architecture

Certified by . . . .
William J. Mitchell
Dean of the MIT School of Architecture and Planning
Head of the MIT Program in Media Arts and Sciences

Accepted by . . . .
Andrew B. Lippman
Chair, Department Committee on Graduate Students
Program in Media Arts and Sciences

Activity Recognition in the Home Setting Using Simple and Ubiquitous Sensors

by

Emmanuel Munguia Tapia

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning on August 15, 2003, in partial fulfillment of the requirements for the degree of Master of Science at the MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Abstract

In this work, a system for recognizing activities in the home setting that uses a set of small and simple state-change sensors, machine learning algorithms, and electronic experience sampling is introduced. The sensors are designed to be "tape on and forget" devices that can be quickly and ubiquitously installed in home environments. The proposed sensing system presents an alternative to sensors that are sometimes perceived as invasive, such as cameras and microphones. Since temporal information is an important component of activities, a new algorithm for recognizing activities that extends the naive Bayes classifier to incorporate low-order temporal relationships was created. Unlike prior work, the system was deployed in multiple residential environments with non-researcher occupants. Preliminary results show that it is possible to recognize activities of interest to medical professionals such as toileting, bathing, and grooming with detection accuracies ranging from 25% to 89% depending on the evaluation criteria used. Although these preliminary results were based on small datasets collected over a two-week period of time, techniques have been developed that could be applied in future studies and at special facilities to study human behavior such as the MIT Placelab. The system can be easily retrofitted in existing home environments with no major modifications or damage and can be used to enable IT and health researchers to study behavior in the home. Activity recognition is increasingly applied not only in home-based proactive and preventive healthcare applications, but also in learning environments, security systems, and a variety of human-computer interfaces.

KEYWORDS: Activity recognition, home, experience sampling, preventive healthcare, pattern recognition, machine learning, naive Bayes, sensors, ubiquitous.

Thesis Supervisor: Kent Larson, Principal Research Scientist in Architecture

Activity Recognition in the Home Setting Using Simple and Ubiquitous Sensors

by

Emmanuel Munguia Tapia

Master of Science Thesis

September 2003

Advisor . . . .

Kent Larson
Principal Research Scientist
MIT Department of Architecture

Advisor . . . .

William J. Mitchell
Dean of the MIT School of Architecture and Planning
Head of the MIT Program in Media Arts and Sciences

Reader . . . .

Dr. Stephen S. Intille
Research Scientist
MIT Department of Architecture

Reader . . . .

Dr. Joseph A. Paradiso
Associate Professor of Media Arts and Sciences
MIT Media Laboratory

Acknowledgments

This thesis would not have been possible without the contributions, support, and help of the following people.

Stephen

I am grateful for the time, patience, support, encouragement, and inspiration that you have provided me over the last two years. This thesis owes much of its contents to your ideas, guidance and wise advice.

Kent

You have given me gifts of your patience, respect, and wisdom that I cannot repay. Thank you for offering me the opportunity of pursuing research in the area I am most interested in, and guiding me through the process.

Jennifer

I am indebted to you for the help, encouragement, suggestions and valuable ideas that you have contributed during the development of this work.

Margarita

Words cannot express my gratitude to my wonderful fiance. You have shown me how wonderful life is, and how insignificant it would be without love, without you.

Mom and Dad

Thank you for all the love, encouragement, and support throughout my life and my education. You have always supported my dreams, aspirations, and shown me that nothing is impossible if you really wish it. I am especially grateful for your financial support during all those difficult times.

Edelmar and Anaid

I dedicate this thesis to my siblings Edelmar and Anaid, who provided me with inspiration, strength, and hope. Through your eyes, I have seen so many things that otherwise I would never realize myself. Thank you for opening my mind to the importance of early detection of physical and mental medical conditions at home. I have hope that this will be a reality in the near future.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The U.S. health care system is under severe financial stress. As the first wave of baby boomers reaches retirement age by 2010, the situation is expected to deteriorate rapidly [6]. One partial solution to this problem involves the development of systems that shift the emphasis from disease treatment in hospitals to health promotion and quality of life conservation at home. Keeping people at home and out of the hospital reduces the financial burden on the system. 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 [58]. A fine grain recognition of activities of daily living will be essential to implement many of the proposed strategies for encouraging healthy behavior related to diet, exercise, and medication adherence. Unfortunately, health researchers currently do not have the means to collect the necessary sensor data to detect activities and patterns of behavior in actual homes. Researchers do not know if it is possible to recognize human activities using a set of simple, easy to install ubiquitous sensors nor do they understand what modifications are necessary in conventional pattern recognition algorithms to achieve this. Two interesting questions are: What living patterns can be detected using simple inexpensive sensors installed in homes? How can this information be utilized?

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) [54], and enhanced ADLs (EADLs) [68]. These activities include eating, getting in and out of bed, using the toilet, bathing or showering, dressing, using the telephone, shopping, preparing meals, housekeeping, doing laundry, and managing medications. If it is possible to develop computational systems that recognize such activities, researchers may be able to automatically detect changes in patterns of behavior of people at home that indicate declines in health. This early data could be available to the individual or designated caregiver, such as a family member or a physician.

Preventive and proactive health care systems are not the only applications of activity recognition systems. Systems that understand user activity patterns could be used in new architectural design tools to identify and rank user needs and preferences over time. Activity recognition in real-time could also allow the development of just-in-time learning environments that educate and inform people by presenting information at the right time as they move through the environment. 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. Finally, knowing people's activities could enable more responsive systems to automate tasks such as lighting and HVAC control.

1.2 Challenges of Activity Recognition

To create algorithms that detect activities, computational models that capture the structure of activities must be developed. The behavior of an individual can be characterized by the temporal distribution of his activities such as patterns in timing, duration, frequency, sequential order, and other factors such as location, cultural habits, and age [78]. Below are human behavior attributes that present challenges for recognition:

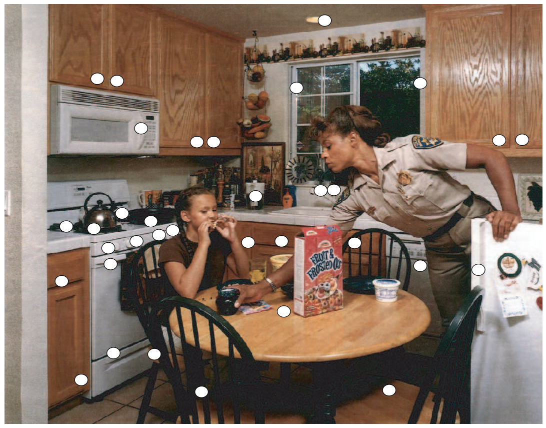


Figure 1-1: A typical median income American kitchen (from [77]). Every location where a state-change sensor might be placed in this particular kitchen is marked. Ideally, no damage would be done to the cabinets or furniture during the installation, and all rooms of the home could have such measurement devices.

Multitasking. Individuals often perform several activities at the same time when they do any kind of work that does not fully engage their attention.

Periodic variations. Everyday activities are subject to periodic daily, weekly, monthly, annual, and even seasonal variations. For example, a person might typically prepare breakfast in 15 minutes on weekdays and for one hour during weekends.

Time scale. Human activities also occur at different time scales. For example, cooking lunch can take 25 minutes, while grooming may only take a couple of minutes or even seconds.

Sequential order complexity. Sequential order, the position in time that an activity has in relation to those activities preceding and following it, is particularly important. The choice of what to do next as well as how that activity is performed is strongly influenced by what one has already done and what one will do afterwards [78]. For example, preparing breakfast is very likely followed by eating.

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.

Location. Human behavior is also affected by location. For example, cleaning the kitchen involves a different sequence of actions than cleaning the bathroom.

Cultural habits. Some cultural habits may be expressed by individuals in typical sequences of activities. For example, some cultures are used to taking a nap after lunch while others are used to having a cup of tea before having breakfast, lunch or dinner.

1.3 Activity Recognition System Goals

In this thesis, a system for recognizing activities in the home setting is presented. The system combines electronic experience sampling, a set of small, easy-to-install, and low-cost state-change sensors to monitor the interaction of people with the environment, and supervised pattern classification algorithms to recognize everyday activities.

Figure 1-1 illustrates how simple sensors might be installed ubiquitously in the home. Locations where a sensor could be placed to provide useful information in this particular kitchen is marked. Ideally, the sensors are unintrusive and removable without damage to the cabinets or furniture, and all rooms of the home may have such measurement devices. The sensors are "tape on and forget" devices that can be quickly and ubiquitously installed in home environments.

One of the goals of this work is to present an activity recognition system that would use such sensors to recognize activities of daily living. The proposed sensors require no major modifications to existing homes and can be easily retrofitted in real home environments. A second goal is to illustrate how a sensing system that does not use sensors that users typically perceive to be invasive such as cameras and microphones can be used to recognize activities.

The structure of the thesis is as follows. Chapter 2 presents extended examples of how the activity recognition system could be used in different applications domains.

In Chapter 3, background in previous research on methods of recognizing activities is introduced as well as some design considerations. Chapter 4 explains how the components of the activity recognition system were designed and implemented. In Chapter 5, the two studies that were carried out to collect data from real home environments and some preliminary results are presented. Finally, Chapter 6 summarizes the most important conclusions, and offers some suggestions for the modification of the current system in future work.

Chapter 2

Extended Examples

2.1 Home-based Preventive and Proactive Healthcare

In 2030, nearly one out of two households will include someone who needs help performing basic activities of daily living [3]. Sensor-based technologies in the home, for example, may help people proactively detect emerging medical problems before they become crises [43, 1, 40]. Detection of everyday activities would enable systems to monitor and recognize changes in patterns of behavior that might be indicators of developing physical or mental medical conditions. Similarly, it could help to determine the level of independence of elderly people, to understand side effects of medication, and to encourage medication adherence.

A sensor system that can detect changes in everyday activities in the home could enable a new generation of home-based and institutionally-based services for the aging [4, 62]. Activities of interest fall into three categories: activities of daily living, instrumental activities of daily living, and enhanced activities of daily living. *Activities of daily living*, or ADLs are the primary daily personal care activities that are necessary for people to be able to live independently. Even though they are body-centric activities, they often involve movement around a space and interactions with other objects. For the elderly, this includes walking supports, tubs, beds and bed pans, etc. If the activities can be regularly detected, then changes in these activities can also be monitored over time. Fundamentally, if there is a *routine* that involves objects that are moved or shaken, then detection may be possible.

Instrumental activities of daily living, or IADLs, are more complex actions such as using the telephone, preparing adequate meals, and doing housework [53]. Some IADLs can be broken into sub-activities such as vacuuming, washing dishes, storing dishes, etc. This work does not deal with detection of *enhanced activities of daily living*, or EADLs, because they are activities that are unlikely to create routine sensor firing patterns in the home. EADLs include actions such as accepting new challenges, continuing education, part-time work, and volunteering [68]. For a complete list of ADLs, IADLs and other activities important for medical applications, refer to table A.1.

Once the activities important for medical applications are detected by the system, the information could be available for: (1) *the family* to help them determine what kind of support their elderly or impaired relative needs, particularly as a disease progresses or as a recovering person becomes increasingly ready to cope with the world, (2) *the physician* to help him determine the appropriate doses of medications or the most effective therapy, and (3) to *the individual* so he/she can have a better understanding of his/her condition or impairment and how to cope with it.

Preventive and proactive healthcare is not the only application of systems that automatically detect activities. In the next section, other applications of home-based activity recognition are introduced.

2.2 Other Applications of Home-Based Activity Recognition

Changing Behavior at the Point of Decision Information about activity can also be used to motivate a positive behavior change by showing information at the point of decision [39]. Visualizing nutritional facts while shopping or preparing breakfast can strongly influence food choice [83]. Medication adherence could be motivated by identifying the best time to remind the person to take his medication. Determining when the individual is receptive to a reminder requires the understanding of the occupant's activities over time.

Learning Environments Activity recognition in real time could allow the development of just-in-time learning environments that educate and inform people by presenting information at the right time as they move through the environment. Knowing what a person is doing will help determine the best time to interrupt the occupant to present them with useful information or messages. Someone preparing dinner represents a good opportunity for a teaching system to show words in a foreign language related to cooking.

Architectural Design Tools Tools that understand activity patterns could be used in an architectural design process to identify and rank user needs and preferences over time. For example, information about mobility patterns and multitasking while cooking in the kitchen can be used to help users make design decisions while designing a new kitchen space. Knowing that a person moves from the stove to the refrigerator two times on average while preparing meals could help to arrange the furniture in a kitchen more efficiently [14].

Security and Surveillance Systems If a surveillance system can create a model of behavior over time, it could be able to predict intent and motive as people interact with the environment. Moreover, the system may be able to determine people's identity by observing activities and interactions with the environment over time.

Automation Systems In order to effectively automate tasks such as lighting and HVAC control, it is necessary to accurately predict activities, tasks and mobility patterns over time.

Chapter 3

Theory/Rationale

3.1 Approaches to Activity Recognition

There are at least four ways for a computer to automatically acquire data about people's activities using sensor systems: (1) ask the individual, as in experience sampling [24], (2) remotely observe the scene using audio, visual, electromagnetic field, or other sensors and interpret the signal readings, (3) attach sensors to the body and interpret the signal readings, and (4) attach sensors to objects and devices in the environment and interpret the sensor readings.

Directly asking questions is a powerful technique but one that must be used sparingly. Frequent interruption will annoy people.

Although progress is being made on algorithms that monitor a scene and interpret the sensors signals, the acquired sensor data is often highly unconstrained. For example, audio is currently being used to detect activity transitions in a monitoring system for domestic environments(ListenIn) [82]. The variety of sounds found in a home environment is, however, considerable, and it could be particularly difficult to differentiate sounds generated by multiple individuals.

More complex sensors such as cameras in computer vision have also been used for recognizing activities. Computer vision sensing for tracking [42, 73, 36] and action identification [26, 59, 76]) often works in the laboratory but fails in real home settings due to clutter, variable lighting, and highly varied activities that take place in natural environments. Little of this work has been extensively tested in the field due to the complexity of dealing with changes in the scene, such as lighting, multiple people, clutter, etc. Finally, because sensors such as microphones and cameras are so general and most commonly used as recording devices, they can also be perceived as invasive by some people.

Attaching sensors to the body is a promising and relatively inexpensive technique to acquire data about certain types of human movement [12]. In addition, they can be easily deployed in actual environments. Posture, for example, can be detected automatically from accelerometers [31, 57, 80], as can some types of physical activity such as walking, sitting, standing, lying down, and inclining [56, 10, 81, 45, 55, 12]. The following techniques have been used to detect various activities: A wearable sensor jacket using stretch sensors [29] to detect walking and running; an audio processing wearable computer to recognize social interaction in the workplace [18, 51]; a video processing wearable computer to identify some everyday patterns of activity [22, 21]; biometric wearable sensors to visualize exercise in everyday activities [37], and a context-aware PDA with GPS to recognize activities such as "grocery shopping", "going to work", "going to class" and "doing laundry" among others [19].

Many activities however, involve complex physical motion and more interaction with the environment. Signals from highly variable arm and leg movements during activities such as cooking and cleaning may not allow these activities to be differentiated from other activities. Movement is dependent upon objects in the environment. On the other hand, if sensors are embedded in the environment, activities such as cooking may be indicated by a distinct set of sensor firings (e.g the stove, cabinetry) with only minor day to day variations. Simple sensors can often provide powerful clues about activity. For instance, a switch sensor in the bed can strongly suggest sleeping, [2] and pressure mat sensors can be used for tracking the movement and position of people [64, 5, 13].

Previous work where sensors have been placed on objects in the environment include an adaptive control of home environments(ACHE) system developed to control the basic residential comfort systems (air, heating, lighting, ventilation and water heating) by observing the lifestyle and desires of the inhabitants [60]. Simple sensors in a kitchen (temperature on stove, mat sensors, and cabinet door sensors) have also been used to detect meal preparation activities in one specially-wired home kitchen [13]. Managing an intelligent versatile home (MavHome) is a system that seeks to maximize inhabitants comfort and minimize operation costs by predicting mobility patterns and device usage [25]. A major difference between this prior work and this thesis work is that these systems have not been deployed in multiple residential environments with actual occupants. They have typically been used in laboratories or homes of the researchers themselves and their affiliates. Further, all of these systems have required careful (and usually painstaking) installation and maintenance by research staff and students (e.g. [60, 2, 63, 17]).

Ultimately, systems that attach sensors to devices in the environment and to the human body in combination with sensors that observe the scene using audio, visual, or magnetic sensors, and ask the user for information will be most powerful for recognizing human activities, because each method has shortcomings and complementary strengths. In this work, the use of simple and small sensors distributed in devices around the environment for recognizing activities is explored.

3.2 Algorithms for Recognizing Activities

There has been some previous research in algorithms for recognizing activities and patterns of activities from sensor data on data collected from living environments. Mixture models and hierarchical clustering have shown some promise with a small set of sensors in a kitchen environment for clustering the low-level sensor readings into cooking events using temporal information [13]. Clustering low level sensor data by time and location is useful for recognizing activities happening at different times and in different locations in the home environment. However, choosing the number of clusters to use, and correlating the clusters of sensor readings to the activities is difficult. Furthermore, since different people perform activities in different ways, supervised learning with an explicit training phase offers a promising approach to the activity recognition problem.

Combination of neural nets(NNs) and lookup tables have been used to predict occupancy-mobility patterns and expected usage of comfort systems such as heating, lighting and ventilation in the Neural Network House [60]. However, since neural nets are like "black boxes" that do not provide information about the underlying model of the process, it is difficult to extend them to incorporate prior knowledge. Prior knowledge and underlying information about the model are particularly important in health related applications. Furthermore, it is not clear how NNs will scale to a large number of features and deal with small training sets. Large number of features might be required in order to encode time information, a primary feature of human activities.

Dynamic Bayesian networks (DBNs) [61] are a specific type of Bayesian network that graphically encode dependencies among sets of random variables which evolve in time. Hierarchical hidden semi-Markov models (HHSMMs), specific types of DBNs, have been used to track the daily activities of residents in an assisted living community [50]. The algorithm can distinguish different activities solely based on their duration and noisy information about the location of the residents. Even though DBNs have proven to be one of the most powerful representations for temporal events and efficiently fusion information from multiple sensors [35], the complexity of the networks and learning algorithms make it difficult to apply them in problems involving hundreds of low-level sensors. For example, the CPU time required to process one day of data from a couple of sensors in [50] was one hour. This suggests that algorithms that encode time in an effective but computationally efficient way need to be created.

Quantitative temporal Bayesian networks (QTBNs) have been used to monitor if a person is following an action plan appropriately from sensors observations [23]. QTBNs extend temporal reasoning approaches such as time nets [48], DBNs, and dynamic object oriented Bayesian networks [34] to model fluents and quantitative temporal relationships. Since this representation is used to monitor plans of actions, it requires manually encoding the relationship between different activities in the network structure. Furthermore, each activity is assumed to occur only once and if the plan contains two instances of the same action, each instance is considered different and modelled separately. This make it particularly difficult to capture recurrent activities. In addition, this approach is computationally expensive for hundreds of sensors since the time required for learning and inference increases with the number of nodes (sensors or features).

Sequence matching approaches have also been applied to predict inhabitant's actions in order to automate the routine and repetitive tasks of inhabitants. The Smart Home Inhabitant Prediction (SHIP) algorithm [25] matches the most recent sequence of events with collected histories of actions to predict inhabitant future actions. Since the inhabitant commands are encapsulated in actions and the predicted actions correspond to the matched sequence most frequent in the inhabitant history, it has difficulties modelling ambiguous and noisy information from multiple sensors. Sensors with noisy outputs would generate incorrect predictions due to incorrect sequence matches.

Nearest neighbor and decision trees have been used to detect everyday activities such as walking, watching TV, and vacuuming from accelerometers [12]. Nearest neighbors [7] would result in a poor real-time performance and a large storage requirements when detecting activities from hundreds of sensors over long periods of time since all instances are required in the prediction stage. Unlike most other techniques, decision trees [65] often generate understandable rules. One problem is that they overfit data and may not combine probabilistic evidence as well as other methods. They are also not suitable for representing activities happening in at the same time since the classes are assumed to be mutually exclusive. However, if high-level information about activities is required to weight attributes, decision tree results may be easier to adapt than other methods.

Naive Bayesian network structures have shown to be sufficient to recognize complicated activity involving the actions of 11 coordinated people given noisy input data [41]. Naive Bayesian classifiers make strong (and often clearly incorrect) independence assumptions that each class attribute is independent given the class. They also assume that all attributes that influence a classification decision are observable and represented. For these reasons, they are sometimes assumed to perform poorly in real domains. On the contrary, experimental testing has demonstrated that naive Bayes networks are surprisingly good classifiers on some problem domains, despite their strict independence assumptions between attributes and the class. In fact, simple naive networks have proven comparable to much more complex algorithms, such as the C4 decision tree algorithm [52, 20, 46, 27]. One theory is that the low variance of the classifier can offset the effect of the high bias that results from the strong independence assumptions [32]. To apply naive Bayes classifiers to the activity recognition problem, however, temporal dependencies must be considered. Therefore, one approach would be to encode large numbers of low-order temporal relationships in the networks [41].

One possible improvement over naive classifiers perhaps would be to encode key dependencies by augmenting simple networks with a few key links using a technique called tree-augmented Bayesian networks (TAN) [33].

TAN classifiers have recently been used to recognize activities such as “grocery shopping”, “going to work”, “going to class”, and “doing laundry” among others from GPS data [19]. One potential problem is that they sometimes perform poorly with small training sets and that they do not capture sequences of events, so an extension to weight features and encode temporal information must be considered.

In this work, the naive Bayes classifier is extended to incorporate temporal relationships among sensor firings and recognize activities in the home setting. The new algorithm created is introduced in section 4.

3.3 Activity Labelling Techniques

While studying human activities and behavior, it is important to use reliable methods for labelling people’s activities. Labels provide a way to validate the results of the study. In this work, activity labels were used to train and test the accuracy of the activity recognition algorithms. The following section presents some of the methods that could be used to label subject’s activities. Figure 3.1 shows a summary of all

Method	Sub-method
-Direct observation	
	-person following subject
	-continuous video
	-image-based experience sampling
-Indirect observation	
	-via sensor activations
	-via photo diaries
	-via audio notes
-Self report	
	-time diaries
	-recall surveys
-Experience sampling	
-End of study interview	

Table 3.1: Methods for labelling the subject's activities

the labelling methods introduced.

Direct observation Direct observation, considered the "gold standard" for assessment in medical and psychological research studying behavior in natural settings, does not suffer from selective recall if performed by trained observers. Even though direct field observation can provide helpful qualitative and quantitative measures, it is costly, time-consuming, and disruptive. This technique raises privacy concerns since researchers need to invade private settings such as the home in order to label the participants' activities. Therefore, it is not practical for labelling the sensor data required for the activity recognition system that are proposed in this work.

Indirect observation The researcher with or without the subject's help could be able to self-infer the activity labels by observing the sensor activation signals, diary photographs of the sensors activated, or audio notes recorded during the activities among others.

Self report: time diaries To minimize selective recall and selective reporting bias, time diaries can be used. Participants write down what they do during the day either as they do it or at regular, finely-spaced intervals [67]. Diaries provide better data than recall surveys but are burdensome for the user.

Self report: recall surveys Despite the widespread use of self-report surveys for assessment of behavior in naturalistic settings, these surveys are known to suffer from recall and selective reporting biases - users can often not remember what they did and/or do not report what they actually did. Furthermore, they often report what they did incorrectly [74].

ESM/EMA The experience sampling method (ESM), also known as ecological momentary assessment (EMA), involves using a timing device to trigger self-reported diary entries. In electronic ESM, questions can be answered on a survey on a portable computing device that "samples" (e.g. via a beep) for information about the participant's activities. Sampling can occur using fixed, statistical, or self-report schedules. With a sufficient number of subjects and samples, a statistical model of behavior can be generated. The ESM is less susceptible to subject recall errors than other self-report feedback elicitation methods [24, 74], but its high sampling rates interrupt activities of interest and irritate subjects. There can be subject-selection bias (e.g. busy people feel they are too busy for interruption) [24].

End of study Interviews At the conclusion of a study, the participant can be interviewed by the researcher. Interviews have shown to be particularly effective for critiquing ideas or gathering information about the participants' tasks and activities if performed properly. Often however, participants know more than they say in a single or even several interviews [70], and will tend to have difficulty understanding and recalling how context impacts their behavior (i.e. exhibiting selective recall and selective reporting [74] biases).

Chapter 4

Design and Implementation

4.1 Overview

The proposed system consists of three major components. (1) *The environmental state-change sensors* used to collect information about use of objects in the environment, (2) *the context-aware experience sampling tool (ESM)* used for labelling the activities, and (3) *the pattern recognition and classification algorithms* for recognizing activities.

Figure 4-1 shows the block diagram of the activity recognition system. During the data collection stage, the state-change sensors are installed in the home environment while the ESM is provided to the user to help train the system by making a detailed record of his/her activities. After data has been collected for up to 14 days, the sensor data and the activity labels are correlated in a data integration stage. Binary features that capture temporal relationships can then be calculated over the sensor activations. The activity classifier is now trained using the features and the activity labels. During the learning stage, the activity classifier creates a model of the user's activities based on the features calculated over the sensor firings. At this point, the activity classifier is able to predict activities based on new sensor firing observations. Figure 4-2 shows a simplified version of the study design. The following sections explain the different components of the activity recognition system in more detail.

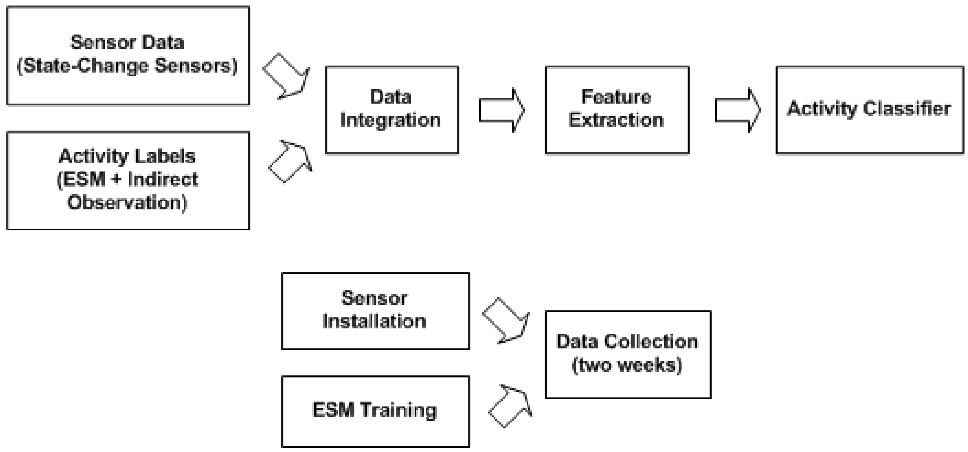


Figure 4-1: Block diagram of the activity recognition system.

4.2 Environmental State-Change Sensors

Although other low-cost wireless sensing systems have been developed, notably Berkeley Motes [38], Smart Dust [47], and Smart-ITS [49], their size, power, and cost points are not appropriate for distributing hundreds of units in a single apartment-size space to collect data reliably. They are not "tape on and forget" devices that can be deployed at a *low cost* in real homes.

Therefore, a low-cost computing technology to study human activity in home environments was developed. These sensors passively collect data via the measurement of objects in the environment. Each device consists of a small data collection board and a sensor(reed magnet or piezo film). These components can be taped to objects in the environment for studies lasting up to several weeks. For example, if a data collection circuit board and sensor are taped to a cabinet door, they will store a timestamp each time the cabinet is opened or closed. An example of the format of the data stored by the state-change sensors is shown in Table 5.1.

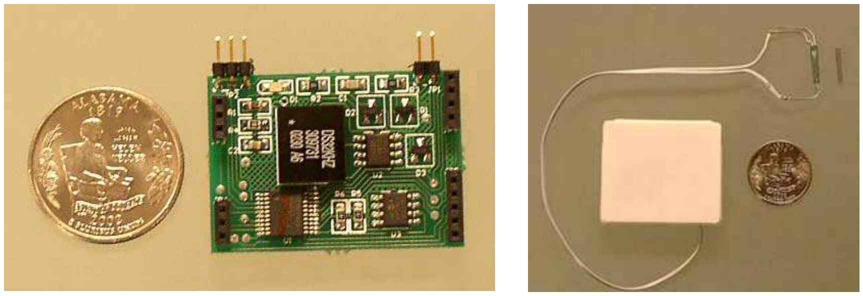
4.2.1 Design Criteria

The design goals were as follows: to permit several hundred sensors to be installed in an environment for at least two weeks, left unattended, and then to recover them with synchronized data.

Study Design

1. Design of data collection board and sensors
2. Development of methodology to collect and label the data
3. Collect data in real home environments for two-weeks
4. Label the data using ESM + indirect observation of sensor activations
5. Train the naive Bayes Classifiers
    - a) Determine the number of activity labels
    - b) Calculate the average duration of each activity Li (feature windows)
    - c) Calculate the features from the beginning to the end of each activity for each day
        - i) Sensors that fired for each activity
        - ii) Pairings: sensor A fires before sensor B
        - iii) Pairings: sensor in object type A fired before sensor in object type B
        - iv) Pairings: sensor in location A fires before sensor in location B
    - d) Train the multiclass and multiple naive Bayes classifiers with the training examples
6. Predict activities
    - a) Divide the day into time intervals Δt (3 min. for studies, 5 sec. for real time)
    - b) At each time t, calculate the features from t-Li to t for each feature window
    - c) Using the multiclass and multiple binary naive Bayes classifiers, calculate the probability for each feature window (activity) for current time t.
7. Evaluate the results using three different criteria
    - a) Calculate percentage of time activity was detected
        - i) Divide each activity label into time intervals Δt
        - ii) Predict activity label for each interval Δt
        - iii) Compare predicted label for interval Δt and activity label
        - iv) Fill confusion matrix according
    - b) Calculate if activity was detected at least once "around" the end of the activity or with a delay (best interval of detection)
        - i) Create an interval at the right most edge (E) of each activity $[E-\phi, E+\phi]$ and divide it into time intervals Δt (ϕ = time delay allowed)
        - ii) Predict activity label for each interval
        - iii) Compare predicted label with activity label and fill confusion matrix
        - iv) The activity is detected in best interval if at least for one interval Δt, the predicted label is equal to the activity label
    - c) Calculate if activity is detected at least once for the duration of the activity label
        - i) Divide each activity label in time intervals Δt
        - ii) Predict activity label for each interval Δt
        - iii) Compare predicted label for each interval Δt with activity label and fill confusion matrix
        - iv) The activity is detected at least once if for at least one interval Δt, the predicted label is equal to the activity label

Figure 4-2: Simplified diagram containing the steps followed in the study design.



(a) (b)

Figure 4-3: The state-change sensors that can be installed ubiquitously throughout an environment such as a home. Each device consists of a data collection board and a small sensor. (a) Data collection board (b) Reed magnet sensor and data collection board in protective case.

In order to achieve this goal, the data collection circuit boards need to have the following characteristics: (1) relatively inexpensive so that it is affordable to install hundreds of them in the home environment. (2) low power consumption, since they are powered by a small coin battery and are left unattended for weeks. (3) small size so that they are easy to install and hide,(4) adequate non-volatile memory size, so that they can collect data over weeks and, finally (5) high reliability.

4.2.2 Implementation

Figure 5-1 shows the sensor device which consists of the sensor itself connected by a thin wire to a 27mm x 38mm x 12mm data collection board. The devices are robust and easy to work with. Each fits snugly in a small plastic case of dimensions 37mm x 44mm x 14mm. A small hole is drilled in the case through which a thin wire is run out to the sensor. The boards can use either reed switches, which are activated when brought into contact with a small magnet, or piezoelectric switches, which detect movement of a small plastic strip. These sensors were chosen because they are relatively inexpensive (under $2), small, and do not consume much power.

To achieve well-synchronized measurements, the most precise temperature compensated crystal oscillator available on the market was used to drive the real time clock of the data collection board: the DS32kHz from Dallas Semiconductor. This achieves a time accuracy of +/- 1 minute per year if operated from 0 to +40C. To further improve synchronization(prior to installation), all the boards are synchronized in a single session with a single computer clock. When the data collection boards are recovered, the signals from each board are linearly interpolated to better match the reference clock. In boards installed in our laboratory, we have measured the synchronization after this correction to be +/- 2 seconds over a two-week period of time. The boards can record up to 3 activations per second and can record a total of 2666 activations in non-volatile memory. The total cost for parts and fabrication (in quantities of 150) for each data collection board as of February, 2003 was $24.66, with an additional $2 required for each sensor (e.g. magnetic reed). The complete hardware and software specifications for the devices are available upon request.

Figure 4-4 shows the schematic of the data collection board. The circuit consists of a PIC16F628 microcontroller, a DS1302 real-time clock with a temperature-compensated crystal oscillator DS32kHz to keep track of time, a 24LC128 EEPROM non-volatile memory chip for data storage, and circuitry for driving an external piezo film sensor and magnetic reed switch. The circuit is powered by a CR2430 coin battery. In order to achieve a low power consumption, the microcontroller is typically in the sleep mode and only wakes up when the external sensor is activated. The external magnetic reed or piezo film switches are internally connected to the external interruption pin (RBO) of the microcontroller so that when they are activated, they wake up the microcontroller from sleep mode. Once awakened, the microcontroller reads the time and date from the real-time clock, and stores the information in the 24LC128 EEPROM memory. The microcontroller is able to distinguish between opening and closing events since signals edges are different for closing (rising edge) and opening (falling edges) events [1]. The estimated battery life of the data collection board is one year if the external sensor is activated an average of 10 times per day for 30 seconds.

[1]This requires installation to be done consistently

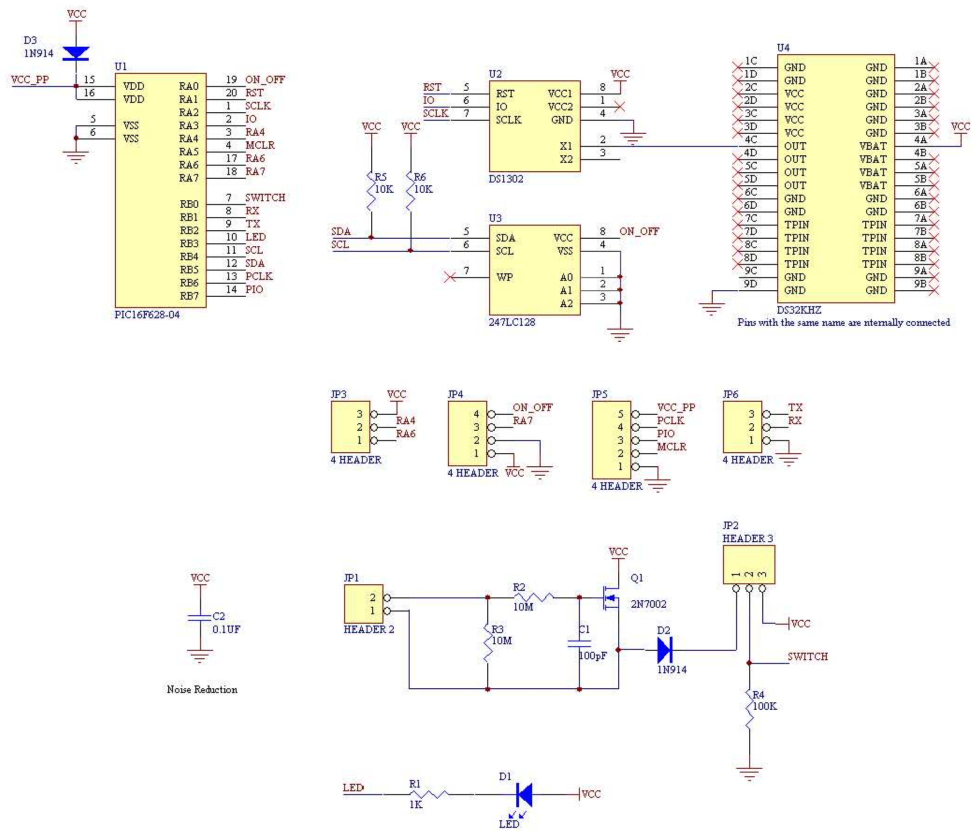
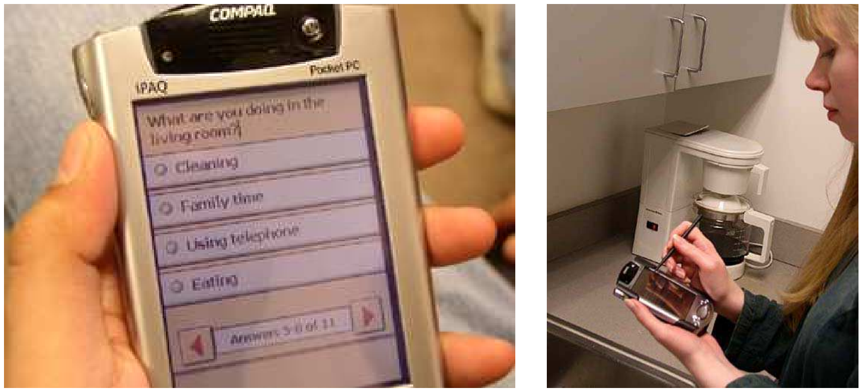


Figure 4-4: Schematic of the data collection board.

4.3 Context-Aware Experience Sampling

The context-aware experience sampling tool(ESM) is a robust data collection tool for acquiring self-reported data from subjects in experiments developed by the House_n research group at MIT [44]. It consists of a personal digital assistant (PDA) used as a timing device to trigger self-reported diary entries. The PDA samples (via a beep sound) for information by asking questions. Questions can be answered by the user on the PDA by selecting from multiple choice menus and triggered using plug-in sensors. Sampling can also occur using fixed, statistical, or self-report schedules. Researchers can load a new protocol by modifying a comma-delimited text file. In this work, only the (15 minute) basic fixed interval sampling functionality of the ESM tool was used to help in the labelling process of the subject's activities. Figure 4-5a shows a screen shot from the sampling protocol used in the studies. The advantages of using

ESM to label the data is that it eases the subject's burden, improves the accuracy of the time-stamps acquired, and reduces the data entry and coding burden of the researcher. The protocol used to collect subject self report labels of activity in this work is described in section 5.1.2.



(a) (b)

Figure 4-5: (a) One screen from the sampling protocol used in the studies for collecting training data on activities in the home setting and used to develop algorithms to detect everyday activities from sensor data automatically. (b) The ESM being used in a kitchen.

4.4 Activity Recognition Algorithms

The purpose of the state-change sensors and ESM was to provide the necessary data to create machine learning algorithms that can identify *routines* in activities from sensor activations alone. In order to accomplish this goal, new algorithms that correlate the sensor firings and activity labels, and predict activities from new sensor firings was required. In this section, the algorithm developed in this work for recognizing activities is presented.

4.4.1 Algorithm Requirements

When designing algorithms for recognizing activities in real home environments, it is important to consider factors such as the ease of setup and training the system, how privacy concerns are addressed, and real-time reliability performance. The following design goals motivated the activity recognition algorithms developed in this work.

Supervised learning. Homes and their furnishings have highly variable layouts, and individuals perform activities in many different ways. In one home a person may store the toothbrush in a medicine cabinet. In another, the toothbrush may be stored in a drawer. Sensing systems must use supervised learning because the same activity (e.g. brushing teeth) may result in a significantly different sensor activation profile based upon the habits, or *routines*, of the home occupant and the layout and organization of the particular home. An elderly woman's "preparing dinner" will cause different sensors to fire and with a different temporal characteristic than a young urban professional's "preparing dinner." One approach to handle such variability is to use supervised learning with an explicit training phase. Supervised learning is similar to learning with a teacher. The desired inputs (sensor firings) and outputs (activity labels) are presented to the classifier and its internal model is updated during the training step to produce the desired outputs. This allows activities to be represented by different sensor firings for multiple individuals.

Probabilistic classification. Probabilistic reasoning offers a way to deal with ambiguous and noisy information from multiple sensors. Further, it is an efficient way to represent different activities happening at the same time (multitasking). Rather than a binary "yes"/"no" determination for a lunch-preparation event, the system could recognize that a person is preparing lunch with a likelihood of 60% and washing the dishes with a likelihood of 40% at a given time. Such probabilistic classification output can then be interpreted as needed by other systems that use the activity information.

Model-based vs instance-based learning. There are different approaches to learning such as instance-based learning and model-based learning. Instance-based learning algorithms store all the attribute examples (training examples) for each class, thus postponing any generalization effort until a new instance needs to be classified. Model-based algorithms on the other hand, use the training examples to construct a global representation or model of the target classification function. When a new classification is required, only the model previously generated is used. In a system that recognizes activities, this means that all the raw sensor data could be eliminated as soon as the user model has been learned. This could reduce users privacy concerns.

Sensor location and type independent. Table 4.1 shows the four levels of information in the sensor data that could be used for developing activity recognition algorithms. Each sensor installed has a unique ID assigned at time of assembly. Ideally, it would not be necessary to record the position (e.g kitchen) and installation object type (e.g drawer) of each sensor in the home: a time-consuming process. Ideally, robust activity recognition would not require this information. Some sensors can record duration in the active state, but since information is dependent on the type of sensor it is not always available. For example, duration information is available for a reed magnet contact sensor attached to a cabinet (time cabinet was open) but not for a vibration or acceleration sensor attached to a TV remote control to monitor movement of the device.

Real-time performance. A system that recognizes activities in the home setting is most useful if it performs in real-time. This may require a trade-off between model, feature, and computational complexity to achieve.

Online learning. An additional feature of a system that recognizes activities is to be able to adjust its internal model in real-time as new examples of activities are available. This will allow the algorithm to adapt to changes in the user's routines over time.

In order to allow the activity recognition algorithms to produce probabilistic classification outputs, adapt to user's variabilities using supervised learning, build a model of the user's activities, learn new examples and perform in real-time, the Naive Bayes classifier (NB) was used in this work. The next section describes the naive Bayes classifier in more detail.

Information level	Description	Example
Level 1	Sensor ID	sensor *66*
Level 2	Duration	activated for *9 Sec*
Level 3	Type	installed in a *drawer*
Level 4	Location	located in the *kitchen*

Table 4.1: Different levels of information in the sensor data that can used to develop activity recognition algorithms.

4.4.2 Naive Bayes Classifier

For its computational efficiency and low classification error rate on many datasets, naive Bayes classifiers (NBs) are one of the most commonly used classifiers. The naive Bayes classifier [28] shown in figure 4-6a is a Bayesian network in which the class node is the parent to all attribute nodes, and its main assumption is that all the attribute variables are conditionally independent given the class variable. As a classifier, it calculates the most probable class given the data using Bayes rule. Among its most important advantages are the following:

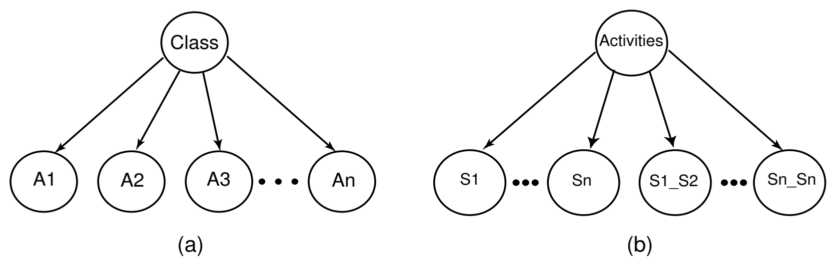


Figure 4-6: (a) The generic naive Bayes classifier. (b) A simplified version of the naive Bayes classifier used in the activity recognition algorithms. The activities to recognize are represented by the class node and all the features calculated by the child nodes. Each Sn node represents the *exist* feature for each sensor, and the Sn_Sn nodes represent the low-level *before*(sensor ID, type or location) features calculated over sensor pairs to incorporate temporal information.

Probabilistic hypothesis. The output is a probability distribution over the classes rather than a simple classification. This also makes it easier to incorporate utility functions to make optimal decisions. Probabilistic outputs are useful for representing the likelihood of two activities happening at the same time.

Combines prior knowledge and observed data. It combines prior probabilities $P(h)$ and the probability of the hypothesis given the training data $P(h|D)$

using Bayes rule. This could allow prior knowledge such as the prior likelihood of occurrence of each activity to be easily incorporated.

Simple learning algorithms. The independency assumption greatly simplifies the learning procedure, particularly when all variables are observed in the training data (frequency counting). Simple learning algorithms are important for achieving real-time performance.

Online learning with fast update. The prior probability $P(h)$ and likelihood $P(D|h)$ can be updated dynamically with each new observed example. This makes it possible to build systems that learn and update its models in real-time. Online learning offers a possibility for adapting to a person's activity variations over time.

Features/examples ratio. For its computational efficiency, naive Bayes is used in applications in which there are many features compared to the number of examples. For example, in natural language processing and information retrieval. The features/examples ratio is important in this work because of the number of sensors involved and the features of sensors pairwise combinations considered.

Meta-classification. The output of several classifiers can be combined to produce more accurate classifiers.

For those reasons, in this work the naive Bayes classifier is used for activity recognition. The network represents a soft, probabilistic summary of a class (the action to recognize). The class (activity) is detected via a set of attributes that are assumed to be independent given the class. The network encodes $P(attribute|class)$ and $P(class)$. The result of using a naive Bayes classifier is a computationally tractable system that recognizes certain types of activities.

4.4.3 Feature Extraction

It is possible to calculate features over (1) single-sensors, (2) multiple-sensors, and (3) multiple-sensor relationships. Multiple-sensor and multiple-sensor relationships

Feature description	Example
exist(*sensorA*, *start*, *end*)	Sensor A fires within time interval
before(*sensorA*, *sensorB*, *start*, *end*)	Sensor A fires before sensor B within time interval
before(*sensorTypeA*, *sensorTypeB*, *start*, *end*)	Sensor in a drawer fires before a sensor in the fridge within time interval
before(*sensorLocationA*, *sensorLocationB*, *start*, *end*)	Sensor in kitchen fires before sensor in bathroom within time interval

Table 4.2: Features calculated and evaluated

features were primarily used in this work to capture temporal relationships among sensor firings. The discriminant power of the features shown in table 4.2 was explored and evaluated over the dataset.

Incorporating Temporal Information

One of the challenges in recognizing activities as discussed in section 1.2 is to encode or capture temporal information such as sequential order, periodic variations and time scale. The main idea of this work is to encode large numbers of low-order binary temporal relationships in the NB network classifier. Thus, two temporal features have been used to recognize (1) whether a sensor activation *exists* during some time period and (2) whether sensor A fires *before* sensor B. These two feature nodes encode the belief that a particular temporal ordering in the sensor firings has been observed during a specific activity. Table 4.2 shows the binary features calculated over the sensor data. The last two features in the table incorporate high level contextual information about the *type* of object in which the sensor was installed (e.g cabinet) and *location* of the sensor (e.g bathroom). The number of *exist* features that will become nodes in the NB networks is equal to the number of sensors present in the system (77 and 84 for subject one and two respectively). The number features that become nodes for the *before sensorID*, *before type* and *before location* features is equal to the number of all pairs of sensors, object types, and locations existent in the home environment. For example, for the first subject, the number of *before sensorID* features is 77x77=5929, 27x27=729 for the *before type*, and 6x6=36 for the *before location*. In order to save memory and computation time, the pair of sensors that were never activated in the dataset were not used in the training and prediction steps.

Incorporating activity duration

Different activities have different lengths of time. Therefore, the duration of the activities has to be considered while developing algorithms for recognizing activities. In order to incorporate the activity duration, one feature window per activity to recognize was used, and the length of each window corresponded to the activity duration as carried out by the subject. Thus, if M is the number of activities to recognize, there were M different feature windows with lengths $L_i \cdots L_m$. The duration or length L_i for each feature window was the average duration for each activity calculated from all the activity labels generated by ESM + indirect observation. For example, the feature window for toileting was 7 minutes and 27 seconds and for preparing lunch was 37 minutes and 54 seconds for the first subject.

Generation of Training Examples

In the training stage, training examples are generated by calculating the features from the start to the end time of each activity label. Figure 4-7a shows an example of how training examples are generated. Examples for washing hands, toileting, and grooming are generated whenever a label for washing hands, toileting, and grooming is found in the dataset respectively.

Originally, there was no *unknown* activity, but examples of this class were created by generating an example of it whenever no activity labels for other activities were found in the dataset. Figure 4-7a also shows an example of how two examples for the unknown class were generated.

Predicting the activity labels

In the prediction step, each feature window (of length L_i) is positioned in the current time to analyze t. The features are then calculated from time $t - L_i$ to time t. Once the features are calculated, the probability for the current activity is calculated using the naive Bayes classifier.

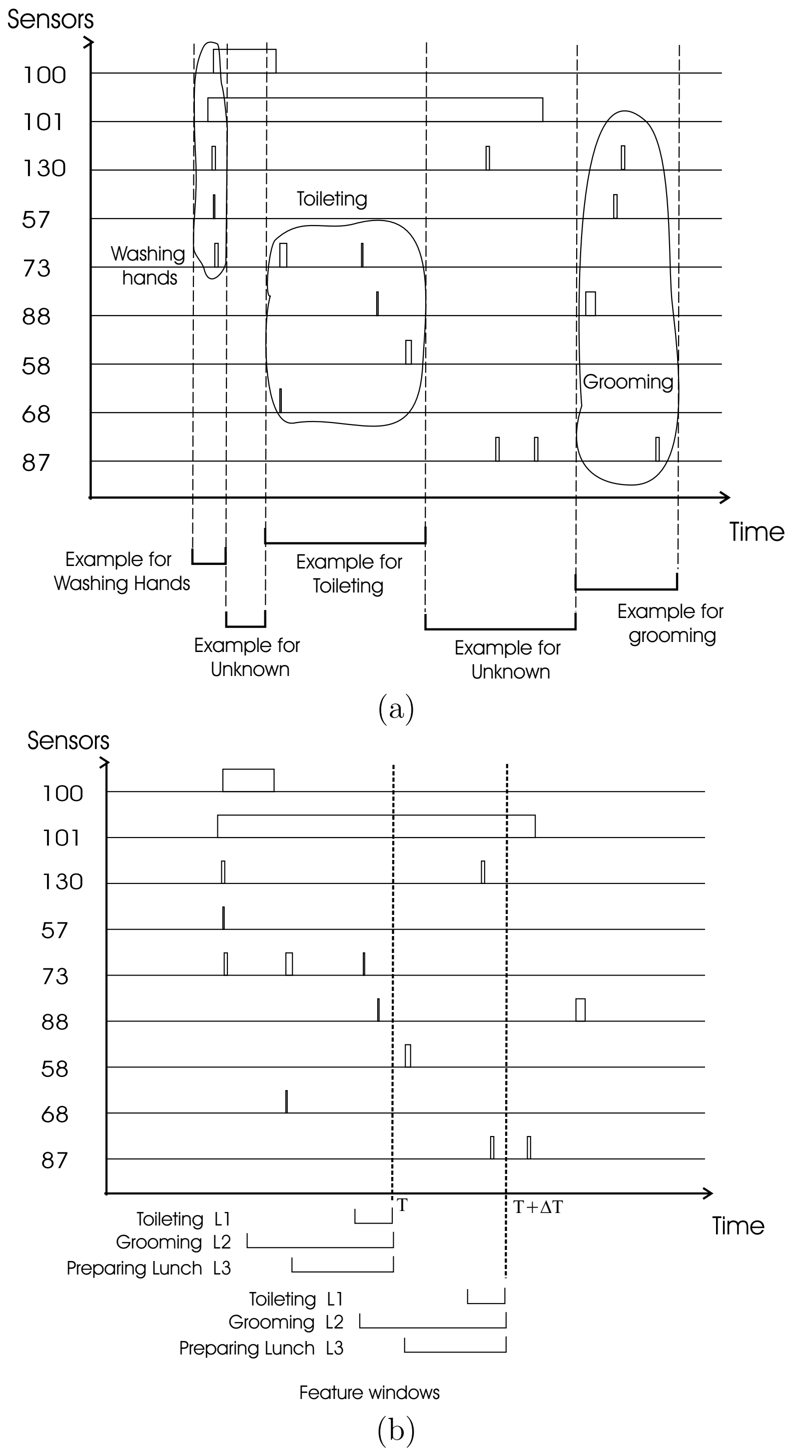


Figure 4-7: (a) Example of how training examples are created for "washing hands", "toileting", "grooming" and two "unknown" activities. (b) Example of how features are extracted from sensor firings using different feature window lengths for each activity for time t and the next time to analyze $t+\triangle t$ in the prediction step.

Figure 4-8 shows an example of how the probability for each activity is generated in the prediction step by shifting the feature window for each activity over the sensor activations. Note that the probability is maximum when the feature window aligns with the duration of the activity represented by sensor activations (activity label). This indicates that the classifier is more likely to detect the activities when they are ending.

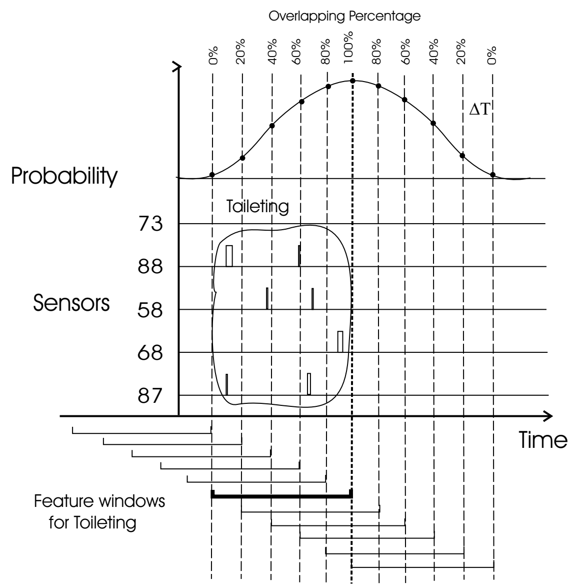


Figure 4-8: Example of how the probability for the "toileting" activity is generated in the prediction step by shifting the feature window for "toileting" over the sensor activations with increments of $\triangle t$ (3 minutes for this study). Note that the probability is maximum when the feature window aligns with the duration of the activity represented by the sensor activations.

Figure 4-7a shows an example of how the feature windows for each activity are positioned in the current time to analyze t and in the next time to analyze $t+\triangle t$ over simulated sensor data. The $\triangle$t increment in time used in the experiments was three minutes, which was half of the duration of the quickest activity. In a real-time application, however, the $\triangle$t can be chosen to be as small as required, for example 5 seconds. While predicting an activity label for new observed sensor firings, the activity with the maximum likelihood at any given time is considered to be the classification result.

4.4.4 Implementation

Two versions of the activity recognition classifier were implemented: (1) a multi-class naive classifier in which the class node represents all the activities to recognize and (2) multiple binary naive Bayes classifiers, each of them representing an activity to recognize. Both versions were implemented using the Java WEKA toolkit [84]. Matlab 6.0 was used to display the sensor activations and activity labels, to calculate the features, generate the training examples and to test the accuracy of the algorithms.

Multi-class Naive Bayes Classifier

A simplified version of the multi-class naive classifier used in the system is shown in Figure 4-6b. The class node represents all the activities to recognize and its child nodes are the *exist* and *before* attributes. In this configuration, all the activities are considered to be mutually-exclusive, which means that the probabilities for all activities sum up to one at any given time. Since the multi-class NB classifier consists only of one network, the training and prediction time are small compared to the time required in the multiple binary classifiers approach. Another advantage is that the evaluation is easier since there is only one answer at any given time (the class with maximum likelihood given the evidence).

Multiple Binary Naive Bayes Classifiers

A multi-class classification problem can be decomposed into a set of binary classification problems. Sometimes, it is easier to train algorithms to distinguish only between two classes and algorithms such as C.45, CART, and support vector machines require this decomposition in order to handle the multi-classfication problem [9]. In activity recognition however, the main advantage of binary decomposition is that the representation does not enforce mutual exclusivity. In this way, detection of *listening to music* does not preclude detection of *preparing breakfast*.

In this approach, one binary naive Bayes classifier is used to represent each activity to recognize. The structure of each classifier is the same as the one shown in Figure

4-6b with the difference that the class node only represents two classes: *activity happening* and *activity not happening*. The binary classifiers were trained using the one-against-all procedure. In the one-against-all training procedure, each classifier is trained using as positive examples the examples that belong to that class, and as negative all the other training examples [9]. The time necessary for training and prediction is longer compared to the multi-class NB approach, and the evaluation becomes more difficult since multiple classes can have high likelihoods simultaneously. This violates the assumption that there is only one answer at a time commonly used while calculating confusion matrices and evaluating machine learning algorithms.

Prior Probabilities and Parameter Estimation

In this work, the prior probabilities for all the activities to classify were assumed to be equal. This means that all the activities are considered to occur with the same likelihood. Moreover, the maximum likelihood or maximum a posteriori (MAP) approach was used to learn the parameters of the networks.

Chapter 5

Evaluation

5.1 Study and Data Collection

Two studies were run in two homes of people not affiliated with our research group to collect data in order to develop and test the activity recognition algorithms. Both subjects granted informed consent and were compensated with $15.00 dollars per day of participation in the study. The first subject was a professional 30-year-old woman who spent free time at home, and the second was an 80-year-old woman who spent most of her time at home. Both subjects lived alone in one-bedroom apartments. 77 state-change sensors were installed in the first subject's apartment and 84 in the second subject's apartment. The sensors were left unattended, collecting data for 14 days in each apartment. During the study, the subjects used the context-aware ESM to create a detailed record of their activities.

5.1.1 State-Change Sensors Installation

The state-changes sensors described in section 4.2 were installed on doors, windows, cabinets, drawers, microwave ovens, refrigerators, stoves, sinks, toilets, showers, light switches, lamps, some containers (e.g water, sugar, and cereal), and electric/electronic appliances (e.g DVDs, stereos, washing machines, dish washers, coffee machines) among other locations. For a complete list of the state-changes sensors installation places please refer to Appendix C.1. The plastic cases of the data collection boards were simply placed on surfaces or adhered to walls using non-damaging adhesive selected according to the material of the application surface. The sensor components (e.g. reed and magnet) and wire were then taped to the surface so that contact was measured. Figure 5-1 shows how some of the 77 sensors were installed in the home of the first subject. The devices were quickly installed by a small team of researchers: an average of about 3 hours is required for the sensors installation in a small one-bedroom apartment of typical complexity. A trained researcher can usually install and test a single sensor in less than 3 minutes.

When sensors were installed, each data collection board (which has a unique ID) was marked on a plan-view of the environment so that when the sensor data was collected, the location (e.g kitchen) and type (e.g cabinet) of each sensor was known. Figure 5-2 shows the top view of the two apartments and the sensor distribution. Once installed, the sensors were left unattended for up to 14 days in each apartment, continuously and passively collecting data about object use. Table 5.1 shows an example of the type of data acquired by the state-change sensors. Appendix F shows plots of the sensors activations for one day of data for subject one and subject two. From the plots, it is easy to see that the sensors are sometimes noisy. Most of the time, the source of the noise is the bad alignment between the magnet and reed switch. Table G.1 shows where the reed switch and magnet were installed in each of the devices or objects in the studies as well as information on how reliable was the installation procedure.

5.1.2 Labelling Subject's Activities

In the laboratory, obtaining annotated data is a straightforward process. Researchers can directly observe and label activity in real-time, or later through observation of video sequences. In the home environment, direct observation is prohibitively time-consuming and invasive. After carefully considering possible methods for labelling the subjects' activities (shown in Table 3.1), two were selected.

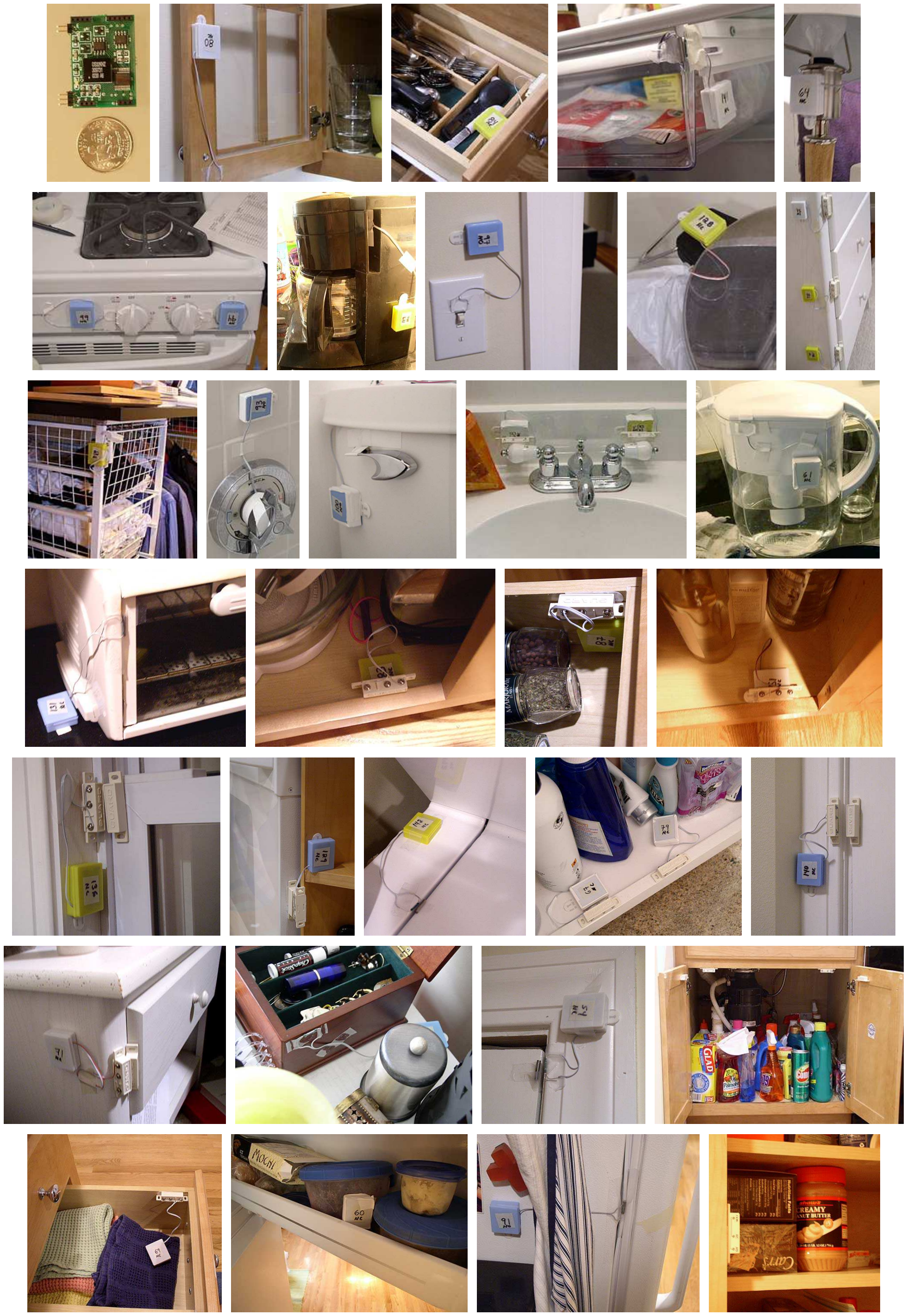
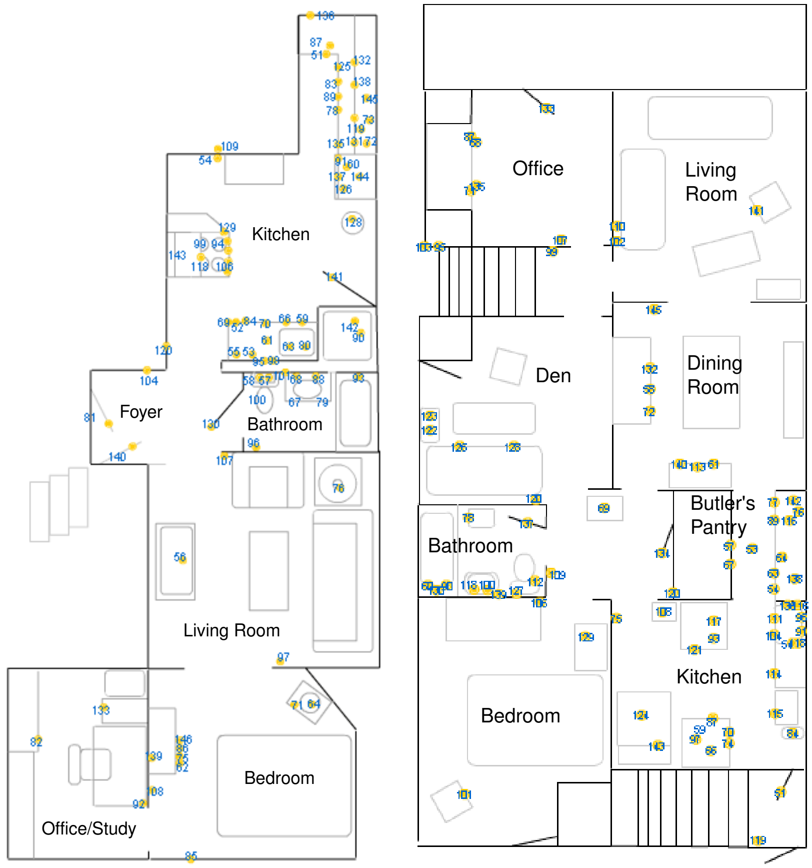


Figure 5-1: Examples of some of the 77 sensors that were installed in the home of the first subject. The sensors and data collection boards were literally taped to objects and surfaces for the duration of the data collection period. Sensors can be installed on many objects that are manipulated by the occupant, including light switches and kitchen containers.



(a) (b)

Figure 5-2: (a) The top view of the apartments and the sensors distribution for subject one. (b) Sensor distribution for subject two.

Experience sampling

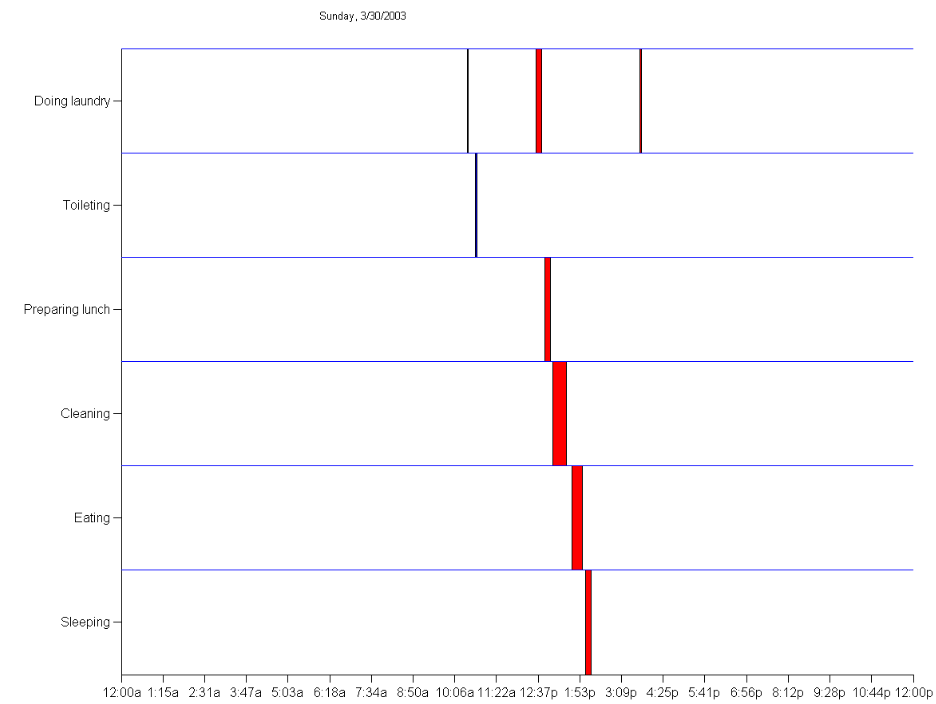
The first strategy was to use the context-aware experience sampling tool described in section 4.3. In this approach, the subjects were given a PDA running the experience sampling software at the start of the study. As the state-change sensors recorded data about the movement of objects, the subjects used experience sampling to record information about their activities. A high sampling rate was used, where the subject was beeped once every 15 minutes for 14 days (study duration) while at home. At the beep, the subject received the following series of questions. First the user was asked "what are you doing at the beep (now)?". The subject selects the activity that best matches the one that he/she was doing at the time of the beep from a menu showing

Activity	sensor ID	day	activation time	deactivation time	duration (sec)	room (opt)	object type (opt)
Preparing breakfast	PDA	12/1/02	08:23:01		10 min		
	23	12/1/02	08:23:03	08:23:07	4	kitchen	drawer
	18	12/1/02	08:23:09	08:23:17	8	kitchen	cabinet
	89	12/1/02	08:24:49	08:24:59	10	kitchen	fridge door
		⋮	(many readings)				

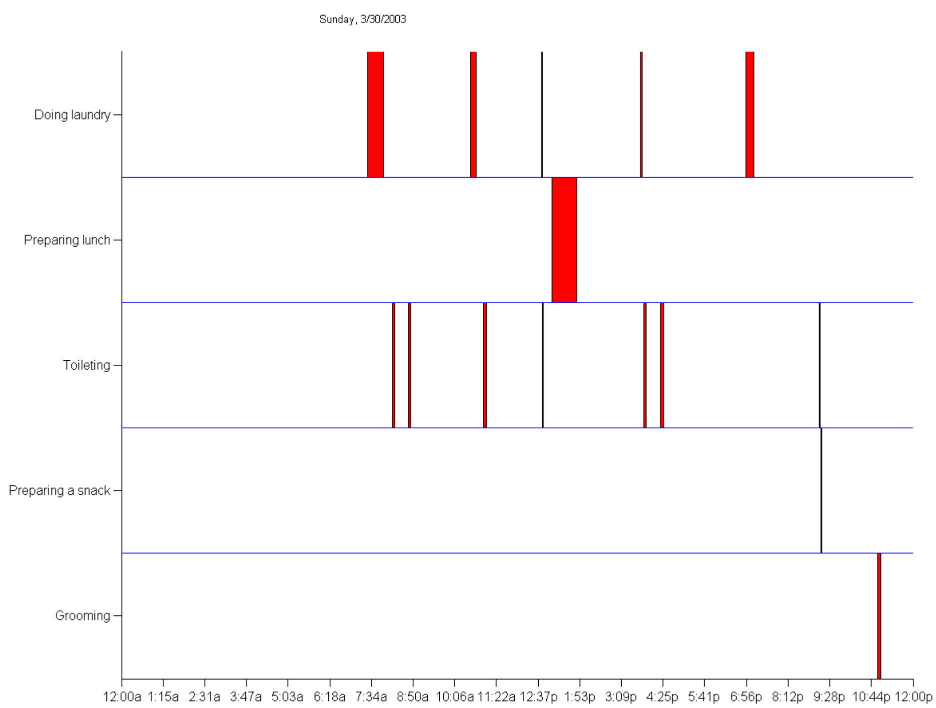
Table 5.1: An example of the type of data that was acquired by the state-change sensors and ESM. The activity attributes are acquired using experience sampling during a training period. The sensor activations are collected by the state-change sensors distributed all around the environment

up to 35 activities. Next, the following question was “For how long have you been doing this activity?” The subject could select from a list of four choices: less than 2 min., less than 5 min, less than 10 min., and more than 10 min. Then, the user was asked, "Were you doing another activity before the beep?". If the user responded positively, the user was presented with a menu of 35 activities once again. For the studies, an adaptation of the activity categories used by Szalai in the multi-national time-use study [78] were used. Activities from the compendium of physical activities were also added [8]. Appendix B shows the list of activities used in the study. The purpose of the *heading* and *category* columns is to make it easier to analyze activities that are important for medical applications. For example, one may be able to judge the level of independency of an elderly person by comparing the number of *personal needs* activities carried out versus *domestic work*. The activities in the *subcategory* column were the ones actually included in the ESM software for the studies. Figures 5-3a and 5-4a show examples of the activity labels recorded by subject one and two subjects using the ESM for one day.

The advantages of using ESM to label the data during the study is that it is less burden than diaries for the participant, improves the accuracy of time-stamps acquired, and reduces the data entry and coding burden for the researcher. Some of the disadvantages include annoyance with high sampling rates and missing data.

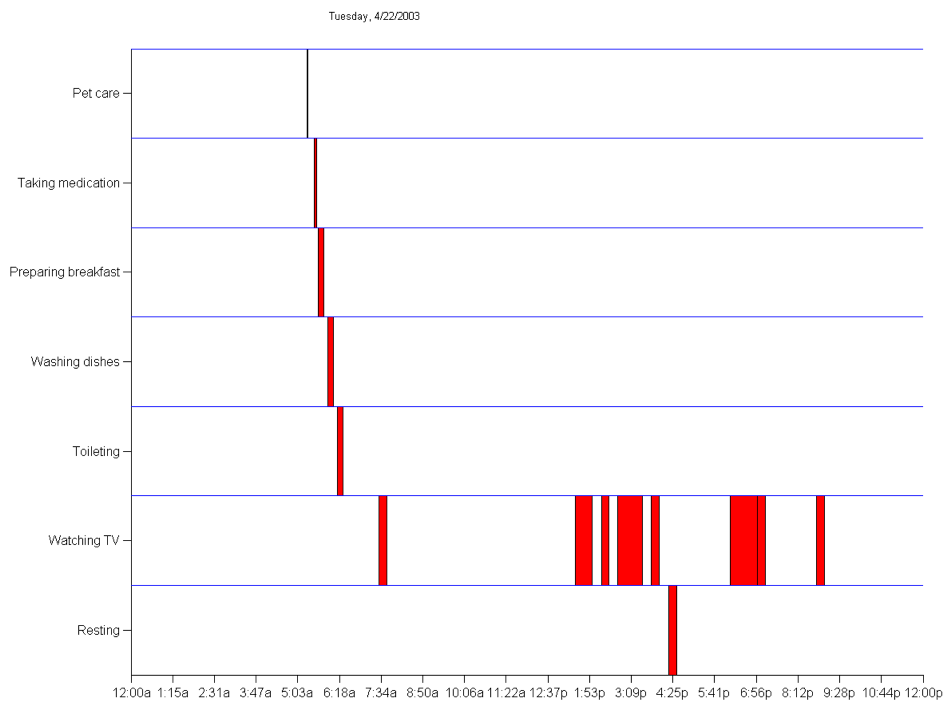


(a)

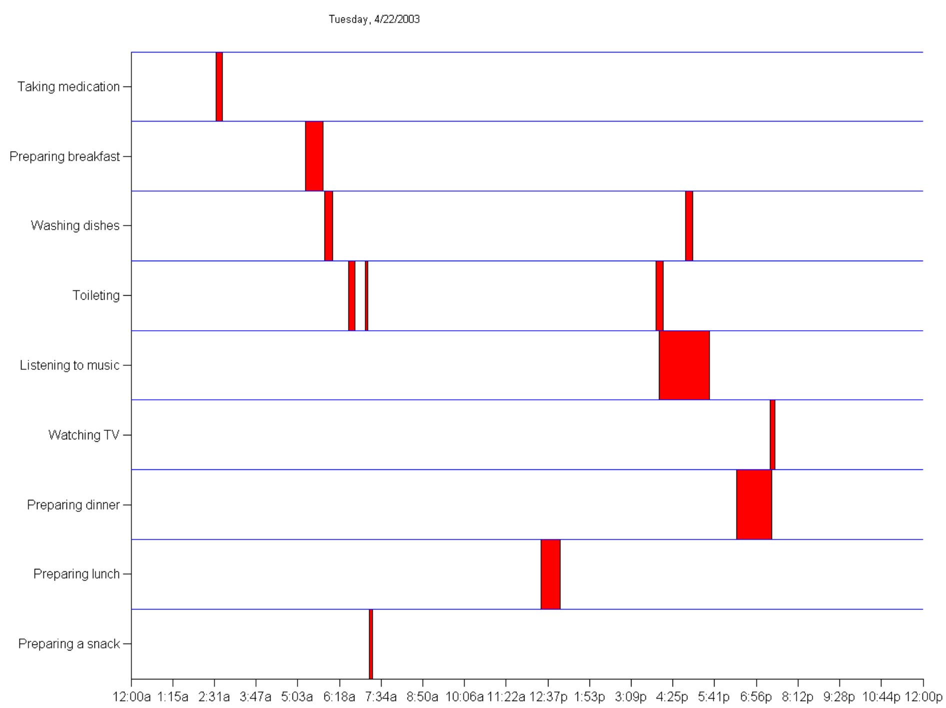


(b)

Figure 5-3: Examples of the activity labels for subject one: (a) acquired using the ESM and (b) by manually labelling using indirect observation via sensor activations.



(a)



(b)

Figure 5-4: Examples of the activity labels for subject two: (a) acquired using the ESM and (b) by manually labelling using indirect observation via sensor activations.

It seems clear that a system for recognizing activities in real home environments will have to be trained by the user, and ESM is one of the promising techniques to do it. Table 5.1 shows an example of the type of data acquired by the ESM.

After carefully studying the labels acquired using electronic experience sampling during the studies, some of the problems that were found are the following:

Human error. The subjects specified an activity that they were not actually carrying out by clicking on the wrong answer box.

False starts. The subject specified an activity that they began to carry out but they did not finish.

Activities with no sensor activations. The subjects specified activities for which there were no sensor activation associated with them. For example, the first subject specified *talking at the telephone* several times but there was no sensor on her cellular telephone. She also specified sleeping when there was no sensor in the bed.

Multitasking. While multitasking, the subject reported the primary activity, but the sensor firings for secondary activities were also recorded by the state-change sensors.

Short duration activities not captured. Activities with duration shorter that the ESM sampling rate (15 min) such as toileting and preparing a beverage proved to be difficult to capture.

Delays. There were delays between the sensor firings and the labels of the activities specified in the ESM. This was probably because the subjects had to stop their activities to answer the ESM questions.

Number of labels collected. Tired of being interrupted each 15 minutes, the subject sometimes did not answer the ESM questions at the beep. Table 5.2 shows a summary of the number of activity labels collected by the ESM.

Measure	Subject 1	Subject 2
Average activities captured per day using ESM	9.5	13
Average activities per day generated by I.O	17.8	15.5
Different activities captured using ESM	22	24
Different activities generated by I.O	21	27
Average ESM Prompts answered to per day	18.7	20.1

Table 5.2: Average number of labels collected by the ESM and indirect Observation (I.O) per day during the study.

Indirect observation of sensors activations

The second method used to label the subject's activities was indirect observation via sensor activations. In this method, the author, with the help of each subject, used self-inference to label the sensor data by visualizing the sensor activations clustered by location, time of activation, and type (description) of each sensor. Photographs of the sensors were also used to help the subject remember his activities during the sensor firings. Figures 5-3b and 5-4b show the manually-generated activity labels for one day of sensor data for subject one and two.

This method results in high quality labels since the subject and the researcher interact in the labelling process. The main disadvantage is that manual labelling all the sensor activations is very time consuming(about three hours per day of data).

Criteria used to label subject's activities Before introducing the criteria used to label the subject's activities, it is important to understand that activities can occur in complex relationships. According to the data collected in the studies, activities can occur sequentially, in parallel, alternating, and even overlapping. Figure 5-5 shows some examples of the different relationships in which activities occur.

Sequentially. Activities that occur consecutively, one after another. For example, preparing a snack followed by bathing. The activities do not happen at the same time and do not overlap with one another.

In Parallel. Activities that can be carried out simultaneously. This activities usually involve activities that do not fully engage people's attention such as preparing breakfast and listening to music.

Alternating. Activities that people switch between in a relatively short period of time. Someone can be washing dishes, washing hands, and cleaning appliances for short periods of time while preparing lunch. The main difference between sequential and alternating activities is that people return to finish one of the activities several times. This usually involves long duration activities and complex activities.

Overlapping. There are cases in which the boundaries between activities are fuzzy. It is difficult to know when one activity ends and another starts.

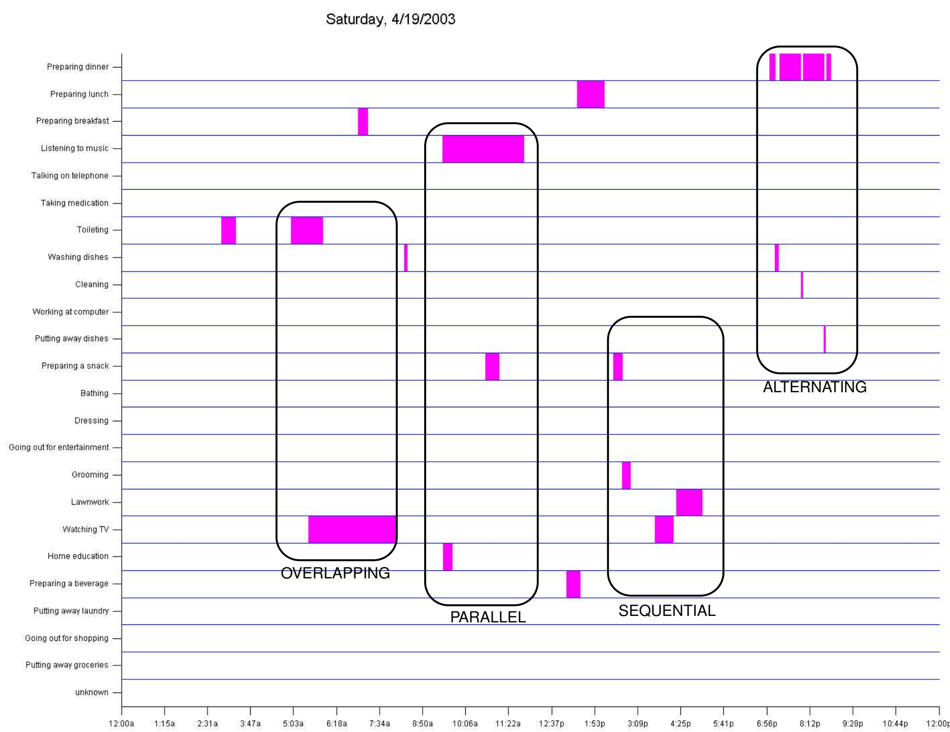


Figure 5-5: Activity relationship cases. (A) The end and start intervals of the activities overlap each other (B) multiple activities occur simultaneously given a period of time (C) activities occur sequentially, but do not overlap, and (D) activities occur sequentially, but a complex activity is carried out most of the time.

Some activities can be divided in starting and ending events. For example, while dressing, a person takes his clothes from the drawers and put them in the bathroom. The person then takes a shower, and after that, he gets dressed putting the dirty clothes in the hamper. From the sensor firings, we can say that the person started dressing while activating the drawer sensors and ended dressing while activating the hamper sensor.

The following criteria were used to label activities from the sensor data, to increase the likelihood that researchers labelling the data could come up with similar results.

1. Activities occur sequentially.
2. The only activities allowed to occur in parallel with other activities are *Listening to Music* and *Watching TV*.
3. Overlapping activities are not allowed. They are assumed to be sequential.
4. Only the primary activity is considered while a person is multitasking.
5. Activities with starting and ending events are divided in two activities and labelled using the same label for the starting and ending events.
6. Only activities for which there exist sensor activations were considered.

5.1.3 Experiments: ESM + Indirect Observation + NB classifiers

The first goal was to test the discriminant power of the attributes as well as the accuracy of the algorithms. In order to achieve this task, it was necessary to use the most accurate activity labels available. The number of labels acquired using the ESM method was not sufficient for training the machine learning algorithms, and they presented the problems mentioned in section 5.1.2. ESM labels, therefore, were only used to improve the manually-generated labels by indirect observation.

Once the ESM + Indirect observation labels were available, they were used to train and test the machine learning algorithm as discussed in section 4.4.3. All activities containing less than six examples were eliminated before training [1]. The average duration for each activity (duration of each feature window) was calculated from all the activity labels, and the $\triangle$t increment used in the prediction step was three minutes.

Two experiments were carried out for each subject's data: One for testing the performance of the multiclass naive Bayes classifier and another for testing performance of the multiple naive Bayes classifiers.

In the next section, the methods developed for evaluating the performance of the algorithms are presented as well as some preliminary results.

5.2 Algorithm Recognition Performance

5.2.1 Validation Criteria

The evaluation of the activity recognition algorithms was difficult. Unlike other machine learning and pattern recognition problems, there is no "right" answer when recognizing activities. The boundaries when activities begin and end are fuzzy. Furthermore, as mentioned in section 5.1.2 activities can happen sequentially, in parallel, alternating, and even overlapping. Finally, there is significant variation in the way observers would label the same activities.

Three methods were used to evaluate and measure the accuracy of the activity recognition algorithms. Which method is most informative depends upon the type of application that would be using the activity recognition data. The methods consider different features of the system that could be important for different applications, for example: (1) is the activity detected at all? and (2) for how long is the activity detected?

In the following section, each of the methods used for evaluating the algorithms is explained. All the methods consider that the classified or predicted activity is the

[1] The threshold of six was chosen arbitrary. Given the complexity of the activities and the large amount of variation possible due to day of the week, time, and other factors, to expect an algorithm to learn patterns with less than six examples did not seem reasonable.

activity with highest likelihood at any given time.

Percentage of time that activity is detected. This measures the percentage of time that the activity is correctly detected for the duration of the labelled activity. Figure 5-6 shows some cases to exemplify the evaluation using this method: (1) activity not detected at all (0% percentage of detection) since the predicted activity fall completely outside the period of the labelled activity. (2) Activity detected 50% of the time since the overlapping region between the predicted and activity label corresponds to half of the duration of the activity label, and (3) activity detected for 20% of the time.

To calculate the percentage of time that the activity is detected, each activity label is divided into time intervals of length $\triangle t$. For each interval, the predicted label (activity with highest probability) is calculated and compared with the original activity label. The rows and columns of the confusion matrix are filled with the comparison result respectively.

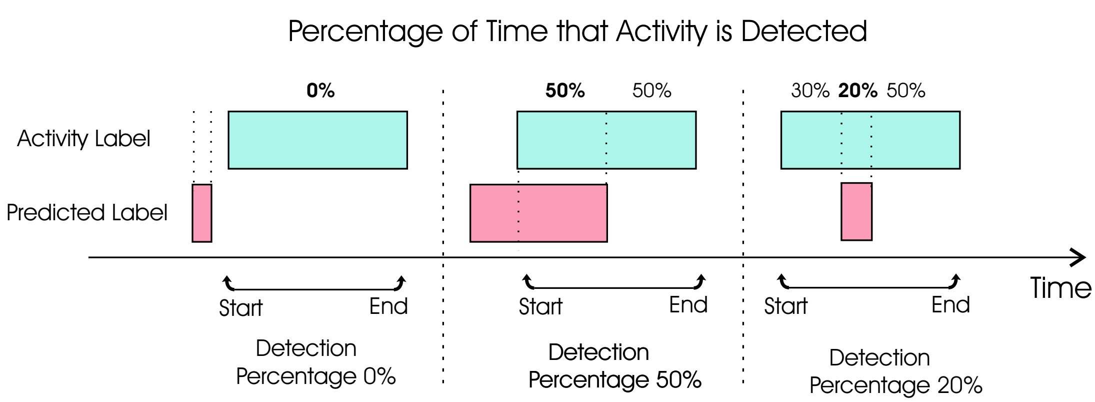


Figure 5-6: Example cases of the "percentage of time that activity is detected" method used to evaluate the activity recognition algorithms. This measures the amount of time that the activity is correctly detected during the duration of the labelled activity.

Activity detected in best interval. This measures whether the activity was detected "around" the end of the real activity or with some delay ϕ for at least one interval($\triangle$t), to allow the detection of activities with a slight delay. As discussed in section 4.4.3, the end of the activity is "the best detection interval".

Thus, the right most edge of each activity (E) is analyzed within an interval of $\pm\phi$. The $\pm\phi$ interval allows the detection of the activities when they are recognized before or after(with a delay) the right most edge. A Detection delay is introduced by the use of the feature windows that capture features back in time. In this work, the ϕ was chosen to be the 7.5 minutes. Thus, activities are allowed to be evaluated as "detected" even when they are are predicted 7.5 minutes after the end of the real activity label. Different applications would require different detection delays, thus different values of the $\triangle\phi$ interval to analyze could be used. Figure 5-7 shows four important cases that could occur while evaluating the algorithms using this method: (1) activity not detected because it occurred outside the duration of the activity label, (2) activity missed because it occurred at the beginning of the activity label, (3) activity detected "around" the end of the activity, (4) activity detected with a delay, and (5) activity not detected because of long delay.

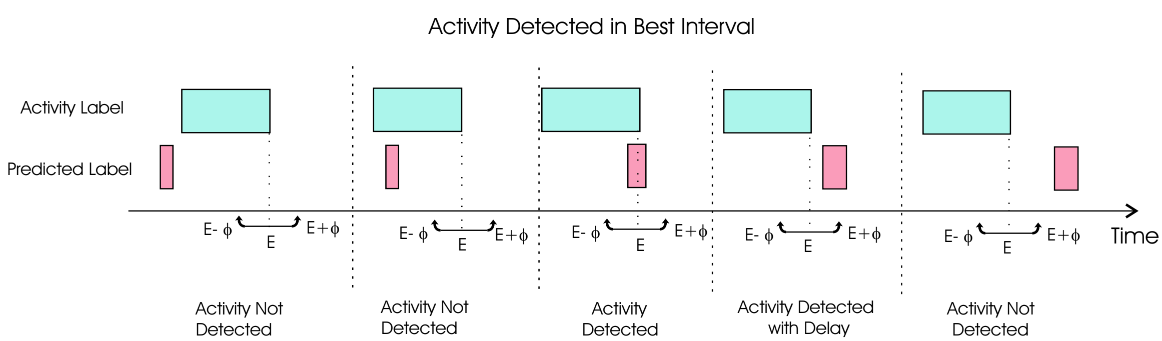


Figure 5-7: Example cases of the "activity detected in the best detection interval" method for algorithm evaluation. This measures whether the activity was detected "around" the end or with some delay after the end of the labelled activity.

In order to calculate this evaluation measure, the right most edge interval $[E-\phi, E+\phi]$ was divided into intervals of duration $\triangle$t. The activity was detected at least once if at least one of the predicted labels for a $\triangle$t interval was equal to the original activity label. If it was not detected at least once, the second most predicted label is used to fill the confusion matrix respectively.

Activity detected at least once. This measures if an activity was detected at least once for the duration of the activity label(no delay allowed). Figure 5-8 shows examples of when is the activity detected or not using this evaluation method: (1) activity not detected because it happens outside the duration of the activity label, (2) activity detected inside duration of activity label, and (3) activity also detected, note that is not important for how long was the activity detected, only if it occurred or not. To calculate it, all the activity labels are split in intervals of length $\triangle$t. The activity was detected at least once if at least one of the predicted labels for a $\triangle$t interval was equal to the original activity label. If the real activity was not detected at least once, the second most predicted label is used to fill the confusion matrix accordingly.

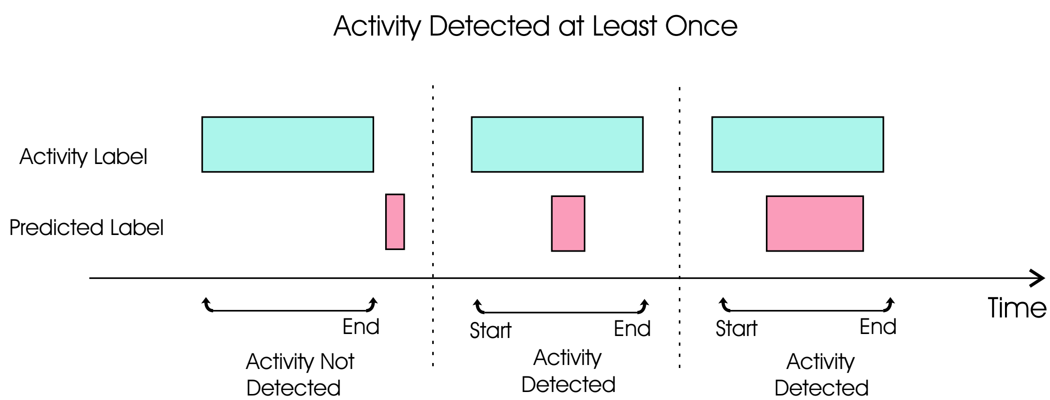


Figure 5-8: Example cases of the "activity detected at least once" method used to evaluate the activity recognition algorithms. This measures if the activity was detected at least once for the duration of the labelled activity.

The selected evaluation method depends upon the desired application. For example, an application intended to judge the level of independence of elderly people may simply determine whether specific activities such as cleaning and bathing are occurring. Conversely, a system designed to detect changes in behavior over time may require an understanding of how long activities are carried out in a daily basis.

5.2.2 Leave-one-day-out Crossvalidation Results

Leave-one-out cross-validation was used in each evaluation method in order to calculate the confusion matrix and measure the classification accuracy. This means that

Number of Examples per Class		
Activity	**Subject 1**	**Subject 2**
Preparing dinner	8	14
Preparing lunch	17	20
Listening to music	-	18
Taking medication	-	14
Toileting	85	40
Preparing breakfast	14	18
Washing dishes	7	21
Preparing a snack	14	16
Watching TV	-	15
Bathing	18	-
Going out to work	12	-
Dressing	24	-
Grooming	37	-
Preparing a beverage	15	-
Doing laundry	19	-
cleaning	8	-

Table 5.3: Reduced table of number of examples per activity generated by indirect observation once the activities with less than 6 examples were removed.

13 that days of data were used for training and one single day for testing the algorithm. The process was repeated 14 times (number of days in datasets) and the accuracies produced in each run were averaged to produce the final accuracies. Using this procedure, the use of training data was maximized. The activity with the highest likelihood at a given time was considered the classification result by all the evaluation methods.

Experiments to determine the discrimination power of the attributes were performed by running the multi-class and multiple binary NB classifiers with some of the possible combinations of attributes shown in Table 4.2. Table 5.4 shows the accuracies per class for the combination of attributes that performed the best for the multiclass naive Bayes (NB) classifier and Table 5.5 for the multiple binary NB classifiers. For a complete listing of the accuracies per class for all the attribute combinations computed, refer to Appendix H.

Discussion of Results

Accuracies vs number of examples. Table 5.3 shows the number of training examples per activity generated by indirect observation once the activities with less than six examples were eliminated. For a complete list of all the number of examples per activity please refer to Appendix J.

Multiclass Naive Bayes Classifier							
	Subject 1			Subject 2			
Activity	**E**	**E+BT**	**Random Guess**	**E**	**E + BT**	**Random Guess**	**Evaluation**
Preparing dinner	0.00	0.00	0.07	0.14	0.13	0.10	
Preparing lunch	**0.25**	**0.29**	0.07	**0.22**	**0.22**	0.10	
Listening to music	-	-	-	**0.20**	0.09	0.10	
Taking medication	-	-	-	0.00	0.00	0.10	
Toileting	**0.27**	**0.31**	0.07	0.20	**0.23**	0.10	Percentage of Time
Preparing breakfast	0.08	0.06	0.07	**0.30**	**0.24**	0.10	Activity is Detected
Washing dishes	0.00	0.00	0.07	0.05	0.11	0.10	
Preparing a snack	0.04	0.01	0.07	0.00	0.00	0.10	
Watching TV	-	-	-	0.04	0.16	0.10	
Bathing	**0.25**	**0.29**	0.07	-	-	-	
Going out to work	0.03	0.00	0.07	-	-	-	
Preparing dinner	0.00	0.00	0.07	-	-	-	
Dressing	0.07	0.03	0.07	-	-	-	
Grooming	**0.26**	**0.26**	0.07	-	-	-	
Preparing a beverage	0.07	0.13	0.07	-	-	-	
Doing laundry	0.09	0.07	0.07	-	-	-	
cleaning	0.01	0.01	0.07	-	-	-	
Preparing dinner	0.25	0.00	0.30	0.23	0.23	0.40	
Preparing lunch	0.59	**0.78**	0.30	0.51	0.48	0.40	
Listening to music	-	-	-	**0.61**	0.44	0.40	
Taking medication	-	-	-	0.10	0.00	0.40	
Toileting	0.71	0.71	0.30	**0.52**	0.48	0.40	Activity Detected
Preparing breakfast	0.45	0.45	0.30	**0.68**	**0.59**	0.40	in Best Interval
Washing dishes	0.00	0.00	0.30	0.51	**0.54**	0.40	
Preparing a snack	0.28	0.15	0.30	0.00	0.00	0.40	
Watching TV	-	-	-	0.25	**0.52**	0.40	
Bathing	**0.87**	**0.79**	0.30	-	-	-	
Going out to work	0.12	0.00	0.30	-	-	-	
Preparing dinner	0.00	0.25	0.30	-	-	-	
Dressing	0.64	0.41	0.30	-	-	-	
Grooming	**0.89**	**0.86**	0.30	-	-	-	
Preparing a beverage	0.36	0.36	0.30	-	-	-	
Doing laundry	**0.86**	**0.78**	0.30	-	-	-	
cleaning	0.00	0.00	0.30	-	-	-	
Preparing dinner	0.00	0.00	0.09	0.38	0.30	0.24	
Preparing lunch	**0.50**	**0.68**	0.17	**0.48**	**0.61**	0.26	
Listening to music	-	-	-	**0.66**	0.45	0.38	
Taking medication	-	-	-	0.00	0.00	0.08	
Toileting	0.42	0.43	0.03	0.46	0.43	0.10	Activity Detected
Preparing breakfast	0.20	0.12	0.07	**0.75**	**0.65**	0.16	at Least Once
Washing dishes	0.00	0.00	0.03	0.15	0.28	0.09	
Preparing a snack	0.08	0.05	0.03	0.00	0.00	0.12	
Watching TV	-	-	-	0.08	**0.45**	0.30	
Bathing	**0.70**	**0.75**	0.11	-	-		
Going out to work	0.12	0.00	0.02	-	-	-	
Preparing dinner	0.00	0.00	0.09	-	-	-	
Dressing	0.21	0.07	0.02	-	-	-	
Grooming	**0.68**	**0.71**	0.05	-	-	-	
Preparing a beverage	0.22	0.31	0.04	-	-	-	
Doing laundry	0.27	0.23	0.05	-	-	-	
cleaning	0.06	0.06	0.09	-	-	-	

Table 5.4: Leave-one-day-out crossvalidation accuracies per class for the multiclass NB classifier using the two combination of features that performed the best from the ones listed in table 4.2. E stands for the *exist* feature, and BT stands for the *before type* feature.

Multiple Binary Naive Bayes Classifiers							
	Subject 1			Subject 2			
Activity	**E**	**E+BT**	**Random Guess**	**E**	**E + BT**	**Random Guess**	**Evaluation**
Preparing dinner	0.00	0.00	0.07	0.11	0.13	0.10	
Preparing lunch	**0.38**	**0.44**	0.07	**0.25**	**0.26**	0.10	
Listening to music	-	-	-	**0.27**	**0.25**	0.10	
Taking medication	-	-	-	0.00	0.00	0.10	
Toileting	**0.26**	**0.28**	0.07	0.23	**0.26**	0.10	Percentage of Time
Preparing breakfast	0.08	0.03	0.07	**0.32**	0.21	0.10	Activity is Detected
Washing dishes	0.00	0.00	0.07	0.02	0.06	0.10	
Preparing a snack	0.04	0.0	0.07	0.00	0.00	0.10	
Watching TV	-	-	-	0.09	**0.22**	0.10	
Bathing	**0.28**	**0.43**	0.07	-	-	-	
Going out to work	0.03	0.00	0.07	-	-	-	
Preparing dinner	0.00	0.00	0.07	-	-	-	
Dressing	0.03	0.03	0.07	-	-	-	
Grooming	0.23	0.18	0.07	-	-	-	
Preparing a beverage	0.05	0.05	0.07	-	-	-	
Doing laundry	0.08	0.05	0.07	-	-	-	
cleaning	0.00	0.00	0.07	-	-	-	
Preparing dinner	0.25	0.00	0.30	0.15	0.26	0.40	
Preparing lunch	**0.84**	**0.90**	0.30	0.51	**0.55**	0.40	
Listening to music	-	-	-	**0.65**	**0.68**	0.40	
Taking medication	-	-	-	0.10	0.10	0.40	
Toileting	0.70	0.62	0.30	**0.58**	0.44	0.40	Activity Detected
Preparing breakfast	0.45	0.37	0.30	**0.68**	**0.59**	0.40	in Best Interval
Washing dishes	0.00	0.00	0.30	0.25	0.27	0.40	
Preparing a snack	0.28	0.10	0.30	0.00	0.00	0.40	
Watching TV	-	-	-	0.27	0.37	0.40	
Bathing	**0.83**	**0.83**	0.30	-	-	-	
Going out to work	0.18	0.00	0.30	-	-	-	
Preparing dinner	0.0	0.00	0.30	-	-	-	
Dressing	0.38	0.26	0.30	-	-	-	
Grooming	**0.86**	**0.66**	0.30	-	-	-	
Preparing a beverage	0.13	0.27	0.30	-	-	-	
Doing laundry	0.51	0.44	0.30	-	-	-	
cleaning	0.00	0.00	0.30	-	-	-	
Preparing dinner	0.00	0.00	0.09	0.30	0.30	0.24	
Preparing lunch	**0.78**	**0.87**	0.17	0.48	**0.61**	0.26	
Listening to music	-	-	-	**0.63**	**0.66**	0.38	
Taking medication	-	-	-	0.00	0.00	0.08	
Toileting	0.40	0.37	0.03	**0.52**	0.43	0.10	Activity Detected
Preparing breakfast	0.20	0.12	0.07	**0.75**	0.46	0.16	at Least Once
Washing dishes	0.00	0.00	0.03	0.10	0.15	0.09	
Preparing a snack	0.08	0.05	0.03	0.0	0.0	0.12	
Watching TV	-	-	-	0.31	**0.56**	0.30	
Bathing	**0.79**	**0.79**	0.11	-	-	-	
Going out to work	0.12	0.00	0.02	-	-	-	
Preparing dinner	0.00	0.00	0.09	-	-	-	
Dressing	0.11	0.07	0.02	-	-	-	
Grooming	**0.56**	**0.51**	0.05	-	-	-	
Preparing a beverage	0.18	0.18	0.04	-	-	-	
Doing laundry	0.25	0.23	0.05	-	-	-	
cleaning	0.06	0.00	0.09	-	-	-	

Table 5.5: Leave-one-day-out crossvalidation accuracies per class for the multiple binary classifiers using the two combination of features that performed the best from the ones listed in table 4.2. E stands for the *exist* feature, and BT stands for the *before type* feature.

As expected, the activities with higher accuracies were those with more examples. For subject one, they were “toileting”, “grooming”,“bathing”, and “doing laundry”, and for subject two “preparing lunch”, “listening to music”, “toileting” and “preparing breakfast”.

Best discriminant attributes. In order to compare the discriminant power of the attributes, the algorithms were run using the following combinations of attributes:

1. Exist
2. Before ID
3. Before Type
4. Before Location
5. Exist + Before ID
6. Exist + Before Type
7. Exist + Before Location

For a complete list of the results per attribute combination, refer to Appendix H. The *exist* attribute showed the best discriminant power and the combination of *exist + before type* attributes presented the second best discriminant power. The low-level *before* features did not perform with the discrimination power expected due to the following reasons: (1) The quality of the labels used in the training, (2) the relatively small data set collected of only two weeks, and (3) the variability in the sequence of pair of sensor activated during each activity in the small number of examples available. However, it is expected that when “ground truth” video is available for labelling the subject activities, this feature would show a higher discrimination power. Also, the *before* feature is expected to help in discriminating among activities with high complexity such as cooking events.

Optional attributes *Type and Location.* Preliminary results show that adding the attributes using the type of object in which the sensor was installed and location information such as the *before type* and *before location* features to the *exist* attribute did not represent a significant improvement in accuracy. This indicates that the development of activity recognition algorithms that do not use the "type" and "location" information is possible. This ultimately represents a speed-up in the installation of the activity recognition system in real home environments since it is not necessary to record or specify the location and description information for each sensor.

Accuracies vs evaluation method used. The activities show lower accuracies for the "percentage of time" evaluation method. Since activities are being detected from sensor firings, some activities such as watching TV, listening to music and dressing are represented by sensors firing only at the beginning or at the end of the activity. This means that there is no information other than the average duration of the activity represented by the feature windows to detect the activities during these "dead intervals" of sensor firings.

The evaluation method with the highest detection accuracies per activity was the "best interval detection". One explanation is that since a delay detection is introduced by the use of feature windows, some of the activities are detected with a delay. The delay interval (ϕ) allowed by this evaluation method was 7.5 minutes, the average delay detection observed in the dataset. Since the importance of delays vary according to the application in mind, different delay intervals can be specified while evaluating the recognition accuracy.

The classes with the highest "best interval detection" accuracies (over 70%) were "bathing", "toileting", "grooming", and "preparing lunch" for the first participant. For the second participant the higher "best interval detection" accuracies (over 50%) were "preparing breakfast", "preparing lunch", "listening to music" and "toileting".

Accuracies vs number of sensors. Since activities such as "going out to work" and "doing laundry" are represented by sensor firings from a single sensor (door and washing machine respectively), it was expected that they would show higher detection accuracies than other activities. However, this sensors were also activated in other activities which decreased their discriminant power.

Accuracies vs sensors installation locations Since only state-change sensors were used in the studies, it was not possible to install sensors in all of the most important locations that could help in discriminating among the activities. For example, sensors were not installed in pans, dishes, chairs, containers and other locations that could improve the recognition accuracy for preparing dinner. As the variety of sensors installed in the home environment increases, more activities will be recognizable. For example, installing a pressure switch under the bed would improve the recognition of someone sleeping.

Accuracies vs activity location. Most of the sensors installed in the in the subjects' apartments were located in the kitchen and in the bathroom. Therefore, more examples of activities usually carried out in these locations were available for training. With more examples available, the accuracy for these activities is higher.

Multiclass vs multiple binary classifiers. Tables 5.4 and 5.5 also show that the multiclass naive Bayes (NB) classifier and the multiple binary classifiers are performing approximately with the same accuracy (±5%). Plots comparing the activity labels and the output probabilities per activity for the multiclass and multiple binary NB classifiers can be found in Appendix F. The main advantage of using multiple binary NB classifiers is its ability to represent multiple activities happening at the same time. The main disadvantage is the number of false positives generated in the classification.

In general, a number of the false positives produced by both classifiers can be explained by the fact that since no video was available for labelling the subject's activities, a considerable number of activities carried out by the subjects were not labelled. Thus, the algorithms could be correctly predicting some of the unlabelled activities.

Subject one vs Subject two results. Overall, the accuracies for the experiments ran over the first subject data were higher than the ones for the second subject data. This was mainly for two reasons: (1) the number of sensors that were dislodged, failed and were noisy was higher in the second subject apartment, and (2) the quality of the labels for the first subject was higher because the sensor firings for each activity were easier to label. One possible explanation is that the first subject spent less time at home and the sensor firings were not as complex as the ones for the second subject.

Improvement over the random guess baseline. The results shown in Tables 5.4 and 5.5 represent a significant improvement for some activities over the random guess baseline.

Finally, while evaluating machine learning algorithms, it is important to consider the cost of false positive versus false negative predictions. A false positive means detecting an activity that is not happening while a false negative involves not detecting an activity that is happening. In the activity recognition problem, the importance of false positives and false negatives depends on the application. For example, in a system designed to interrupt the user to show information, false positives involve interrupting the user when not necessary. Similarly, in a system to judge the level of independence, false positives represent detecting activities that an elderly person is no longer carrying out. False negatives on the other hand, would involve not interrupting the user when required or judging an elderly persons as dependents when they are independent.

5.2.3 Representation Strengths and Weaknesses of the Algorithm

In this section, some of the representation strengths and weaknesses of the algorithm are discussed. The following list is a detailed description of them.

Multitasking. Two of the assumptions while hand-labelling the activity labels were (1) activities occur sequentially (except for “listening to music” and “watching TV”) and (2) while multitasking, the subject reported only the primary activity carried out. Therefore, the algorithms created were not trained on sufficient examples involving multitasking so that a better representation for multitasking could be learned. In addition, the methods used to evaluate the algorithms assume that the classification result is the activity with maximum likelihood (only one answer at a time). This suggest that other labelling conventions as well as new evaluation techniques that handle multitasking need to be created and explored.

Periodic variations. Since data was collected for two weeks, only day to day periodic variations could be considered. The high volatility of the results indicate that the daily variations in the activities was high. However, it is not possible to confirm this without high quality labels produced by video observation.

Time scale. The algorithms were able to detect activities such as “toileting” and “preparing lunch” happening at different time scales. For example, the “best interval” detection accuracy was between 48-90% for “preparing lunch” and between 48-90% for “toileting”.

Sequential order complexity. Sequential order was incorporated by the use of the low-order *before* features. The fact that the *before type* feature allowed maximum “best interval detection” accuracies of 90% and 86% for “preparing lunch” and “grooming” respectively indicate that these features are able to encode sequential or temporal characteristics in the classifiers. However, experiments over sensor data collected over longer periods of time (in the order of months) need to be performed to further explore how well the "before" features are encoding the time relationships among sensor firings.

False starts. False starts were not considered in the studies because it is difficult to study them without accurately knowing what the subject was doing. The only information available about the subject's activities is the sensor firings and the ESM labels.

Cultural habits. This work does not deal with cultural habits since both subjects were North Americans. Further, in order to study cultural differences, data from several subjects from different cultures would need to be collected and compared.

feature space size. Since each attribute is represented by a child node in the NB classifier, the low-level *before ID* feature represents the higher computational cost. The number of *before* features that become nodes in the naive Bayes network is N^2, where N is the number of sensors in the system (84 and 77 for subject one and two respectively). In this work, only binary *before* attributes were considered in order to reduce the computational cost involved in training and prediction.

Scalability in classes. In the multiclass NB classifier case, as new activities to recognize are incorporated, the size of the conditional probability table for the class node increases. This does not represent a major increase in the computational cost in the training and prediction step. On the contrary, the training and prediction in the multiple binary classifiers is considerable since a new binary classifier is needed for each new activity incorporated.

5.2.4 Areas of Improvement

Even though the accuracy for some activities is considerably better than chance, it is not as high as expected. The main problems faced were: (1) the quality and number of activity labels, and (2) the small training set of two weeks. It is expected that training sets collected over months, better labels generated by video observation or other methods, and improved versions of the data collection boards and sensors would improve the detection accuracies. Although these preliminary results were based on small datasets collected, techniques have been developed that could be applied in future studies and at special facilities designed to study human behavior such as the MIT Placelab.

5.3 Privacy Issues in Activity Recognition Systems

There are at least three important privacy issues to consider when designing inconspicuous systems that sense or monitor human activities or behavior in environments [15]. (1) How will the user know or be reminded that invisible devices are present? Since the devices are small or out or view, there is little or any indication for the user that data is being collected. (2) What are the devices capable of knowing/doing? and (3) How do the devices process and distribute the information collected?

On the user's point of view, privacy perception consists of the sensitivity and usage of the information involved, and who is the receiver of that information [15].

Information sensitivity. This is the user estimate of the sensitivity of the data being shared. In the proposed activity recognition system, this can be addressed by allowing the users to install the sensors only in the locations about which they are willing to share information.

Information usage. Information usage involves how the information collected will benefit the user. We think that once the user or family members realize how this technology could help to improve the user's quality of life, they will be less sensitive to privacy issues.

The information receiver. Who will have access to the information? The government spying? Members of the Defense Advance Research Projects(DARPA) Total Information Awareness (TIA) agency looking for suspicious behaviors?

The big brother cameras? Ultimately, any sensing system that detects activities could be used to violate privacy. However, there is always a trade-off between a person's privacy and the level of help or monitoring someone or a system can offer. In this work, it is assumed that the activity information will be primarily used by the user himself, caregivers, health professionals, and sometimes, family members and that this information will represent considerable advantages for the patient and his/her family.

Since the proposed system consist of sensors that need to be installed, the user and/or family members will be aware that the data collection devices are present. Another assumption is that the privacy decisions such as the share of data will be taken by the user of the system or close relatives that have a minimum understanding of how the system works and are able to take decisions about user's quality of life versus possible privacy risks and violations.

The following strategies could be consider to protect the user's privacy rights [15]:

Privacy decisions by giving informed consent. One way to let the user make privacy decisions by giving informed consent is to allow the user to stop the data recording at will. The simplest way to do this is to use a PDA to wirelessly receive the information coming from the sensors. The user can just turn off the PDA to stop the data recording.

Dispose of raw data as soon as possible. Raw data will be disposed as soon as the activity recognition algorithms have created a model of the user activities and behavior. Since the algorithms will use online learning Bayesian networks classifiers, the time that the raw data needs to be kept is minimum.

Use of security access files. Another strategy is to create a mechanism to restrict the access to the information by creating security access files. The user or relatives of the user could specify different levels of access for family members, caregivers, hospitals, insurance companies among others.

Finally, as mentioned before, any sensing system that detects activities could be used to violate privacy. Users, therefore, would ultimately need to be educated about the capabilities and privacy risks of the proposed system.

In this work, a system for recognizing activities that uses a sensing system less-threatening than optical(cameras) and auditory(microphones) sensors is presented.

5.4 Subject Experiences During the Studies

5.4.1 Subjects and State-Change Sensors

The participants felt that they became "oblivious" to the presence of the sensors after a few days of the installation. The first subject reported forgetting that the sensors were there after two days. The second subject was not even sure where some of the sensors were installed. A similar event was experienced by Summers [75] when cameras where used to study subjects in a workplace. The first subject reported that she was able to hear when the sensors where activated (magnetic reed sensor closing). This is one of the reasons why in future systems, an accelerometer will be used instead of the external magnetic reed sensor. Three sensors were dislodged in the first subject apartment and six in the second subject's apartment due to pets (cats), poor installation, and a careless cleaning crew. Table G.3 shows the sensors that failed and showed sporadic behavior during the studies and some possible causes of the problems.

5.4.2 Subjects and ESM

One of the main difficulties in making this type of system work in real home environments is finding effective ways to label people's activities in an automatic or easy procedure.

During the studies, the participants started to feel disrupted by the ESM between the second and third day of the study. This was probably because of the high sampling rate (15min). The second subject, who was an elderly woman, was more tolerant to continuous disruptions than the first, younger subject. The percent of total ESM prompts responded to was 17% and 24% for the first and second subject respectively. Table 5.2 shows the average number of labels produced by ESM and indirect observation per day.

Some "instantaneous" activities such as "toileting" and "going out" were difficult to capture using the ESM because their duration is smaller than the resolution of the sampling time(15min). The first subject stated that she was able to "eat, toilet, groom, and leave the apartment within a 15 min period". The ESM also gave the subjects an unwelcome sense of the passage of time. One of the examples was that the ESM went off three times while the first subject was installing mini-blinds, letting her realize how much time she was wasting in the difficult process. The same subject also described the ESM as "parental" and admitted that she felt a desire to give it incorrect information. Both subjects learned how to use the ESM system relatively fast (approximately in 30min).

5.4.3 Subject Changes in Behavior

The first participant reported that that being sensed did cause her to alter some behaviors. For example, she always made sure to wash her hands after using the bathroom. The second subject did not report any changes in behavior due to the sensors.

5.4.4 Subject Privacy Concerns

Subjects reported that they did not feel that their privacy was violated by installing the state-change sensors in their homes even when some of the sensors were installed in the toilet. When the participants were asked if they would have been comfortable having "time-lapse" images taken during the study they expressed some concerns. The first participant strongly asserted that she would not have been comfortable because she is often in various states of undress, living alone in her apartment. The second participant said that she would feel comfortable as long as they were not installed in the bathroom. Such images would have dramatically improved the density and quality of researcher annotated data.

Chapter 6

Conclusion

6.1 Summary

This work is a first step towards showing that everyday routines in the home are sufficiently structured so that activities can be detected in readings from ubiquitous sensors via supervised learning algorithms.

Some of the activities detected in this work include activities important for medical applications such as toileting, bathing, and grooming with detection accuracies ranging from 25% to 89% depending on the evaluation criteria used. Therefore, systems like the one developed in this work may enable IT and health researchers to study behavior in the home, to recognize a significant number of the activities of interest to medical professionals, and ultimately to detect specific changes in behavior automatically.

An algorithm for recognizing activities was created by extending the naive Bayes classifier to incorporate low-order temporal relationships. The algorithm makes a trade-off between feature and model complexity to achieve real-time performance. This algorithm is able to merge and capture temporal relationships from data coming from multiple sensors efficiently.

Unlike prior work in sensor systems for recognizing activities, the system developed in this thesis was deployed in multiple residential environments with actual occupants. Moreover, the proposed sensing system presents an alternative to sensors that are perceived by most people as invasive such as cameras and microphones. Finally the system can be easily retrofitted in existing home environments with no major modifications or damage.

Although these preliminary results were based on small datasets collected over a two-week period of time, techniques have been developed that could be applied in future studies and at special facilities to study human behavior such as the MIT Placelab.

6.2 Level of Information Inferred from Sensor Activations

What level of information or "context extraction level" can be inferred from sensor activations? Can it infer that a person is preparing Chinese food? Preparing dinner? Preparing food? Or just that cabinets are being opened and closed? Initially, the only thing that is known is that sensors are being activated.

Imagine a sensor installed in a chair located in an office. The sensor is a weight switch that is activated each time it detects a weight over a specific threshold on the chair. It might be possible to draw some conclusions from this sensor. Can it be concluded that a person is sitting in the chair each time the sensor is activated? If the chair is located in front of a computer, could it be concluded that a person is working at the computer each time the sensor is activated? Of course, sometimes this inference will be correct and others times incorrect. By placing more sensors and collecting more information, the inferences may be improved. If a sensor is added in the entry door of the office and another in the computer keyboard, the prediction of someone working at the computer would be more precise. Adding more sensors helps in gaining higher level understanding of the activities carried out while the sensors were activated. This work suggests that the quality of information inferred from the sensors will depend ultimately in the number of sensors placed in the environment, the level of provided contextual information (e.g sensor is in the kitchen, refrigerator sensor never activated while bathing), and the granularity of the information provided in the training phase of the activity recognition system.

6.3 Future Work

6.3.1 Improving State-Change Sensors

Redesign the sensors for wireless transmission. Wireless transmission [30], would enable real-time activity detection and system operation for long (e.g. months or years) periods of time. This can be done in two ways, transmitting data wirelessly to a small set of receivers in the home or to the home occupant's PDA device. This last option may be less expensive, but with the disadvantage of requiring the user to carry the PDA.

Embedding the sensors into common architectural components. Strategies for embedding the state-change sensors in objects such as cabinet hinges and light switches could further simplifying installation in new environments. Ultimately these sensors might be built into the architectural components, and furniture at time of manufacture [11].

Sensors capable of person identification If the sensors were able to distinguish the identity of the person activating them, it would be possible to create systems that recognize activities in multiple-person environments. This could be done by incorporating systems for person identification such as the crystal IR active tag system (Squirts) developed at MIT [71].

Replace the contact sensors with accelerometers. One of the most time consuming tasks while installing the state-change sensors is taping the external sensor to the surface of the object (e.g reed switch and magnet). Thus, one way to reduce the installation time is to incorporate the sensor inside the data collection board by, for example, using an acceleration sensor to sense when a cabinet is moved. In this way, no external components are required.

Incorporate new type of sensors in the environment. New type of sensors may help to recognize sets of more complex activities, or activities not detected before. For example, a sensor which is able to monitor current consumption of electric/electronic appliances is likely to provide powerful hints about device usability. Table D.1 shows a list of some of the possible sensors to incorporate.

Build Laboratory for studying activities. A full scale living laboratory called Placelab is currently under construction in Cambridge, Massachusetts and will be completed by the Fall of 2003. This will consist of a two-bedroom apartment that is fully instrumented with the sensor infrastructure required to implement real-time versions of the activity recognition system. In particular, the environmental sensing system will be physically wired so that PDA-based sampling or image-based sampling protocols can be developed where particular actions (e.g. opening a cabinet, turning a faucet) can trigger sampling. The facility will be a shared scientific resource available to researchers.

6.3.2 Improving the Data Labelling

Experiment with different labelling methods One of the subjects suggested that allowing the user to self-initiate the record of his/her activities in the PDA would improve the number of labels acquired. In this case, the PDA would only beep at the user if he/she does not introduce labels for a predefined amount of time (user self-labelling the data). Another possible way to improve the data labelling is to use audio annotation or speech recognition since it does not fully engage the user's attention and the user does not have to interrupt his activities in the labelling process. The technique has been used by the Human Design Group of the MIT Media Lab.

Incorporate sensors in the PDA to help in the labelling process. Another approach to improving the data labelling is by incorporating sensors in the PDA so that a set of activities could be recognized by the PDA. For example, including an accelerometer in the PDA [12] could help it to recognize when someone is walking, sitting, lying down, and this information could be used to better label the activities by triggering questions at more appropriate times [69].

Personalized lists of activities in PDA. The subjects suggested that having software that allows them to generate a personalized list of activities before starting the study would be useful. By doing this, the PDA activities list can be personalized by showing only those activities that the user performs, saving time in the navigation process while answering questions.

6.3.3 Improving the Activity Recognition Algorithms

Develop activity recognition algorithms to explore the following questions:

- Can the recognition algorithms themselves be used to reduce the burden of maintaining the models by querying about activities only during transitions between activities, which may be less disruptive than querying in the middle of an activity?
- Can algorithms be developed with these datasets that recognize not only action but the style of action [79] and intent? For instance, can we develop recognizers that can detect not only "preparing dinner" but "preparing dinner slowly"?
- Can multi-tasking activities be detected?
- How is high level knowledge about activities obtained without explicit expert knowledge engineering?
- If data from sensors such as cameras were available, how would that information be used by the recognition algorithms?
- How can algorithms that work for one home occupant be extended to multi-user households with intertwined activities?
- How can dynamic Bayesian networks and other model representations of sequential events that have proven powerful at recognizing patterns in auditory and visual data (e.g. [66, 72, 16]) be used to improve the recognition algorithms?

- How can lexical and knowledge databases such as WordNet and Lexica freenet (lexfn) incorporate high level knowledge that could be useful for activity recognition algorithms?

Appendix A

Activities Important for Medical Applications: ADLs, IADLs, and Others

ADLs			
Ambulation	Bathing	Continence	Dressing
Eating	Transferring	Toileting	
IADLs			
Doing housework	Managing money	Preparing adequate meals	Traveling
Vacuuming	Using the phone	Cooking breakfast	Shopping
Dusting	Taking medications	Cooking lunch	
Washing dishes	Accessing meds	Cooking dinner	
Storing dishes		Making snack	
Storing food		Drinking fluids	
Mopping floors			
Cleaning kitchen			
Cleaning windows			
Cleaning fridge			
Disposing trash			
Cleaning laundry			
Other			
Personal grooming	Caring for pet	Watching TV	Reading
Brushing teeth	Walking pet	Decorating	Using computer
Brushing hair	Playing with pet	Visiting with friends	Desk work
Applying makeup	Cleaning litter box	Washing hands	Writing a letter
Shaving	Changing bed	Gardening	Household repairs
Ironing	Sleeping	Napping	Getting the mail

Table A.1: Activities of daily living (ADLs), instrumental activities of daily living (IADLs), and other activities of interest for medical applications.

Appendix B

List of Activities Used in the Studies

Heading	Category	Subcategory
Employment related	Employment work at home	Work at home
Employment related	Travel employment	Going out to work
Personal needs	Eating	Eating
Personal needs	Personal hygiene	Toileting
Personal needs	Personal hygiene	Bathing
Personal needs	Personal hygiene	Grooming
Personal needs	Personal hygiene	Dressing
Personal needs	Personal hygiene	Washing hands
Personal needs	Personal medical	Taking medication
Personal needs	Sleeping	Sleeping
Personal needs	Talking on telephone	Talking on telephone
Personal needs	Resting	Resting
Domestic work	Preparing a meal	Preparing breakfast
Domestic work	Preparing a meal	Preparing lunch
Domestic work	Preparing a meal	Preparing dinner
Domestic work	Preparing a snack	Preparing a snack
Domestic work	Preparing a beverage	Preparing a beverage
Domestic work	Meal Cleanup	Washing dishes
Domestic work	Meal Cleanup	Putting away dishes
Domestic work	Putting away groceries	Putting away groceries
Domestic work	Clean house	Cleaning
Domestic work	Clean house	Doing laundry
Domestic work	Clean house	Putting away laundry
Domestic work	Outdoor chores	Taking out the trash
Domestic work	Outdoor chores	Lawnwork
Domestic work	Pet care	Pet care
Educational	Home education	Home education
Educational	Travel study	Going out to school
Leisure	Leisure	Watching TV
Leisure	Leisure	Listening to music
Entertainment and social life	Travel social	Going out for entertainment
Entertainment and social life	Working at computer	Working at computer
Sports	Travel sports	Going out to exercise
Purchasing goods and services	Travel services	Going out for shopping
Other	Other	Other

Table B.1: List of activities used in the study. Those in subcategory were included in the ESM software.

Appendix C

Installation Location of State-Change Sensors

Object	**APT 2**	**APT 1**	**Object**	**APT 2**	**APT 1**
Living room			**Bathroom**		
DVD	0	1	Toilet flush	1	1
TV	1	0	Medicine cabinet	2	2
Lamp	0	1	Sink faucets	2	2
Drawers	0	0	Shower faucets	3	1
Cabinets	0	0	Cabinets	0	4
Light switches	2	2	Drawers	0	0
			Hamper	1	0
Dining room			Exhaust fan	0	1
Drawers	5	0	Doors	1	1
Cabinet	2	0			
Light switch	1	0	**Bedroom**		
			Lamp	0	1
Kitchen			Cabinets	0	0
Microwave	1	1	Drawers	0	5
Toaster	1	1	TV	1	0
Fridge	1	1	Jewelry box	0	1
Freezer	1	1	Window	0	1
Fridge containers	4	3	Light switches	1	1
Drawers	3	7			
Cabinets	6	14	**Den**		
Sink faucets	3	0	TV	1	0
Coffee Machine	0	1	Stereo	3	0
Containers	1	1	Drawers	2	0
Burners	2	4	Doors	1	0
Oven	0	1	Telephone	1	0
Doors	2	2	Light switch	2	0
Garbage disposal	1	1			
Window	0	1	**Office/Study**		
Washing machine	0	1	Door	1	0
Laundry dryer	0	1	Drawers	4	2
Dish washer	0	1	Light switch	1	1
Light switch	1	2			
			Foyer		
Butler's pantry			Door	0	1
Cabinet	6	0	Closet	0	1
Drawer	5	0	Light switch	0	1
Light switch	1	0			
			Porch		
Hallway			Light switch	0	1
Light switch	3	0			
Door	2	0			
Total				84	77

Table C.1: List of locations in which state-change sensors were installed and number of sensors installed per location.

Appendix D

Possible Sensors to Incorporate in Future Systems

Sensor type	Functional description
Mechanical switch	A conventional mechanical switch such as the push-button. Switches made of plastic or metal that when pressed, generates logic "1s", and "0s" otherwise.
Piezo switch	Flexible components comprising PVDF polymer film with screen-printed electrodes laminated to a polyester substrate. As the piezo film bends, the sensor triggers an output of logic "1s" and "0s" otherwise.
Piezo vibration switch	The same piezo film sensor described for the piezo switch where a miniature mass is used to allow the switch to sense vibration. As the piezo film vibrates, the sensor triggers an output of logic "1s" and "0s" otherwise.
Temperature switch	The temperature switch consists of a temperature sensor that asserts logic "1s" whenever its temperature crosses the user-programmed threshold and "0s" otherwise.
Photo switch	A photo switch such as a photoelectric sensor is a device that detects a light source, and generates logic "1s" whenever the light intensity is greater than the selected threshold and "0" otherwise.
Current monitor switch	The current monitor switch consists of a current sense transformer/chip connected to special circuitry to assert logic "1s" when the current flowing through its terminals crosses the user-programmed threshold and "0s" otherwise. To use the current monitor switch, an electronic/electric appliance is plugged into the switch (much as one plugs in an outlet extender) and then the switch is plugged into the wall socket.
Load cell and pressure switches	Consist of load cells or pressure sensors connected to special circuitry so that they generate logic "1s" when the pressure/weight being measured crosses the user-programmed threshold, and "0s" otherwise.
Movement sensors	For example the pyroelectric sensor. It is made of a crystalline material that generates a surface electric charge when exposed to human body infrared heat radiation. When the amount of radiation striking the crystal crosses the pre-defined threshold because someone is moving near the sensor, it generates logic "1s" and "0s" otherwise.
Acceleration switch	Consists of an accelerometer (Piezo, MEMS, etc) that generates logic "1s" when the acceleration that it measures crosses the user-programmed threshold and "0s" otherwise.
Level switches	consists of level sensors that generate logic "1s" when the water level crosses the pre-defined level and "0s" otherwise.

Table D.1: Description of some of the sensors to use in future activity recognition systems

Appendix E

Sensor Types and Installation Locations

Property	Mechanism	Anticipated locations/roles
Proximity/contact	Reed magnet push-button	Doors closed/open, drawers closed/open, cabinets closed/open, windows closed/open, telephone on/off charger or cradle, trash can open/closed, refrigerators open/closed, microwave open/closed, toaster open/closed, medical cabinet open/closed, faucets open/closed, toilet seat open/close, toilet flush moved/not moved, clothes dryer door open/close, washing machine door open/closed, and roof lights on/off)
Acceleration	accelerometer	Doors closed/open, drawers closed/open, cabinets closed/open, windows closed/open, telephone on/off charger or cradle, trash can open/closed, refrigerators open/closed, microwave open/closed, toaster open/closed, medical cabinet open/closed, faucets open/closed, toilet seat open/close, toilet flush moved/not moved, clothes dryer door open/close, washing machine door open/closed, table used/not used, chairs used/not used, couch used/not used, bed used/not used, remote control used/not used.
Vibration	Piezo film sensor	Table used/not used, chairs used/not used, remote control used/not used, door mats person on/not person on.
Movement	pyroelectric sensor	Couch used/not used, bed used/not used, person movement yes/no
Temperature	Temperature sensor	Electric stove burners on/off, light bulbs on/off.
Water level	level sensor	toilet flushed/not flushed.
pressure/weight	load cells/pressure sensors	door mats person on/not person on, couch used/not used, bed used/not used, chairs used/not used.
Light	Photoresistance sensor	Lamps on/off, roof lights on/off, cabinets open/close, drawers open/closed, computer monitor on/off.
Current change	Current sensing transformer/chip	Television on/off, stereo on/off, DVD player on/off, VCR on/off, vacuum cleaner on/off,lamps on/off, coffee machine on/off, toaster on/off , dish washer on/off, electric can opener on/off, electric steam/slow cooker on/off, electric fry pan on/off, electric popcorn maker on/off, electric pasta machine on/off, electric pressure cooker on/off, trash compactor/disposal on/off, electric stove on/off, electric juicer on/off, electric knife on/off, electric mixer on/off,computer on/off, air purifier on/off, electric tooth brush on/off, electric shaving machine on/off, electric hair dryer on/off.

Figure E-1: List of sensor types and examples of home locations to install them

Appendix F

Sensor Activations and Activities Probability Plots

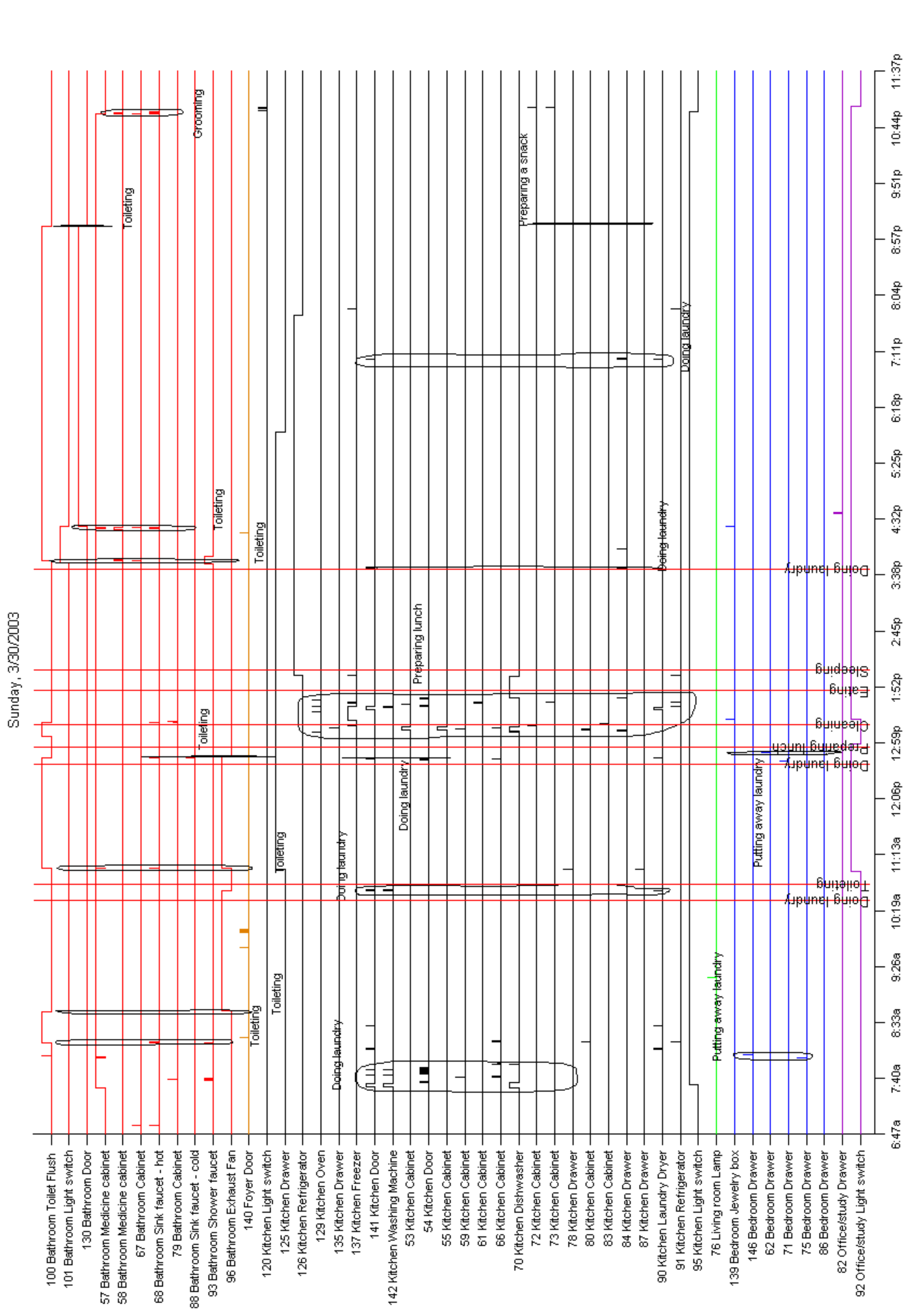


Figure F-1: Plot of one day of sensor data (03/30/2003) for subject one. The labels generated by indirect observation are shown by the round areas and the center time of the ESM labels by the vertical red lines.

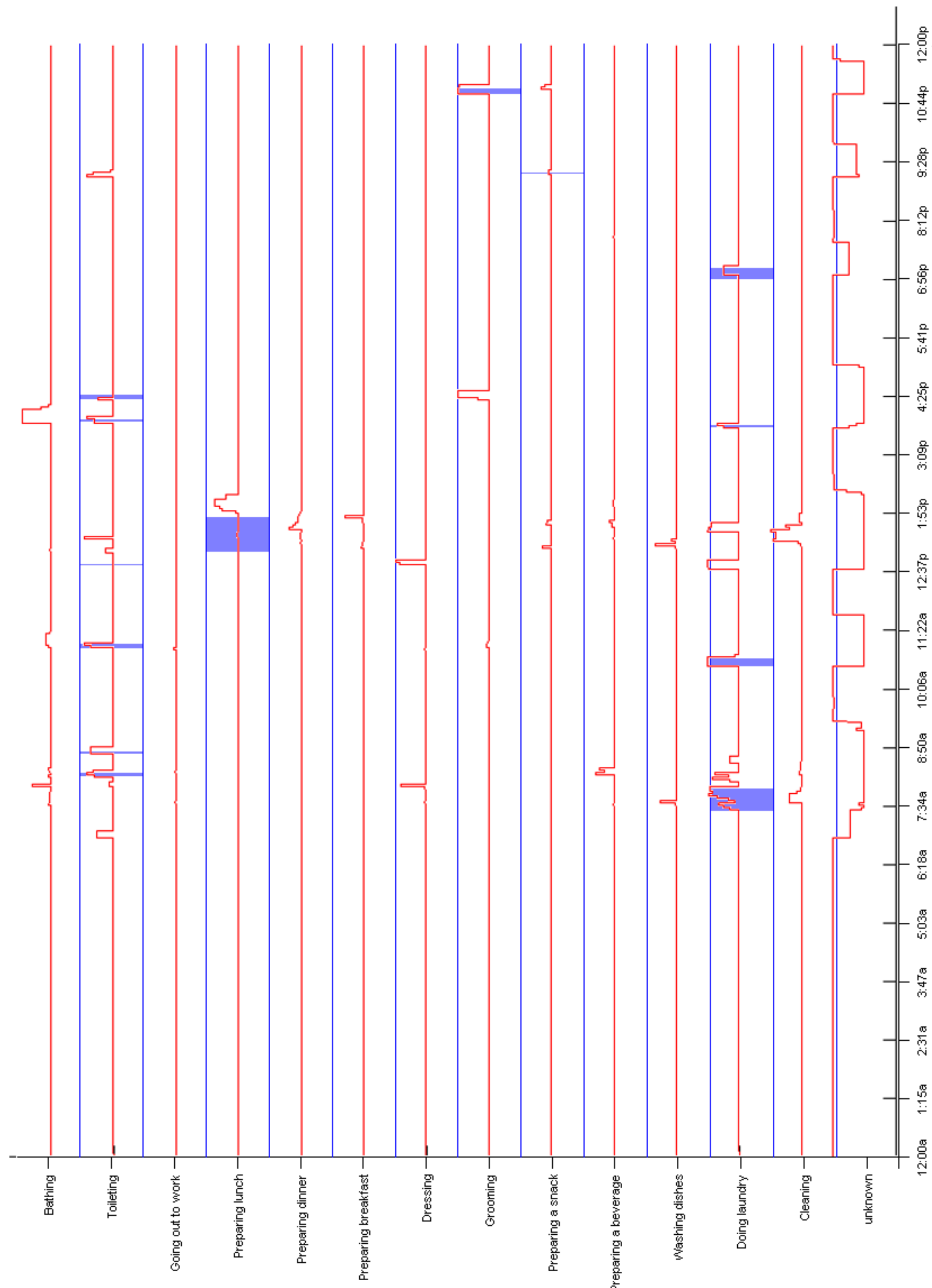


Figure F-2: Plot of the probability classification output of the multiclass naive Bayes classifier for one day of subject's one data (03/30/2003). The blue areas are the activity labels produced by ESM + indirect observation.

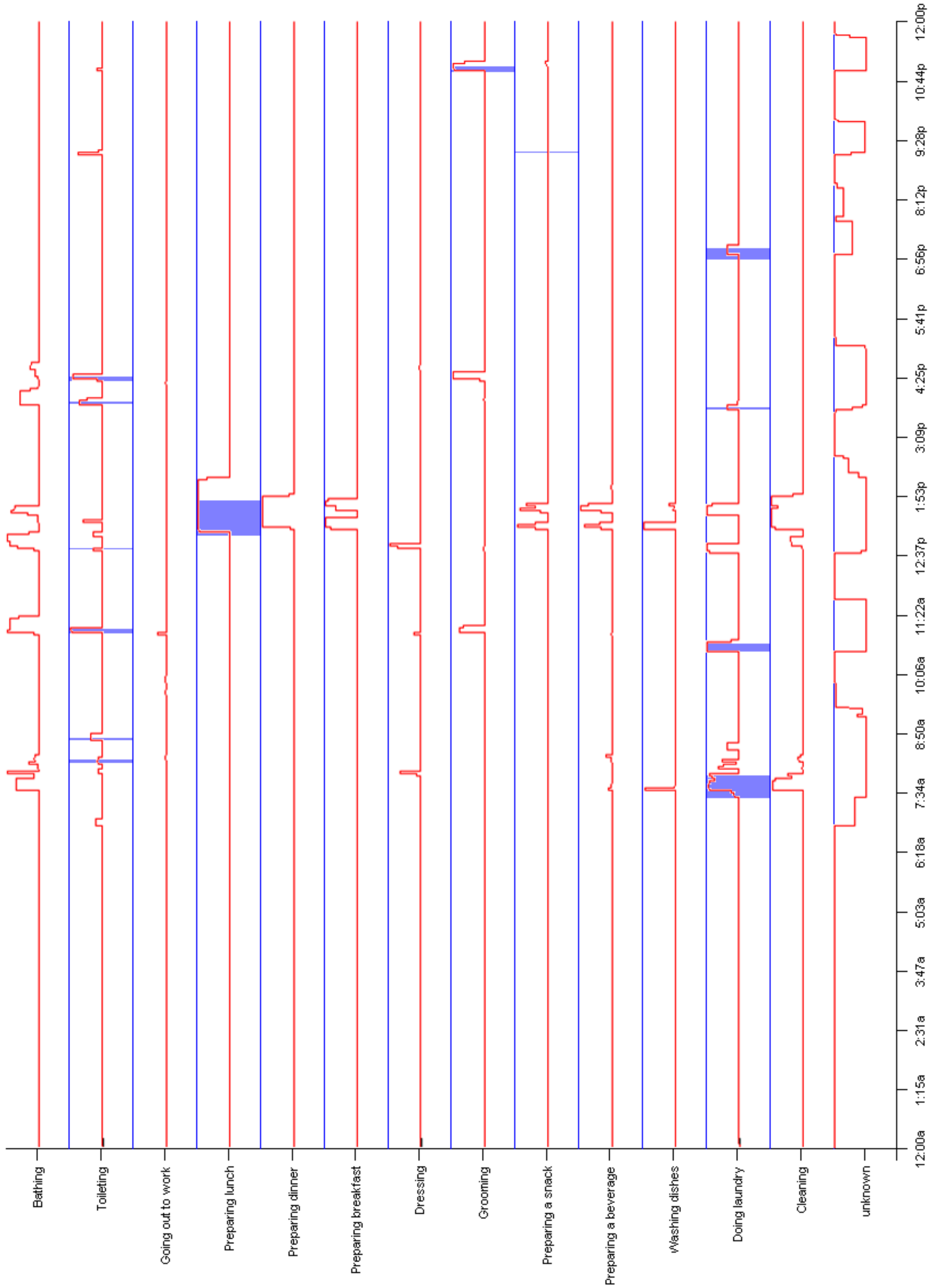


Figure F-3: Plot of the probability classification output of the multiple binary naive Bayes classifiers for one day of subject's one data (03/30/2003). The blue areas are the activity labels produced by ESM + indirect observation.

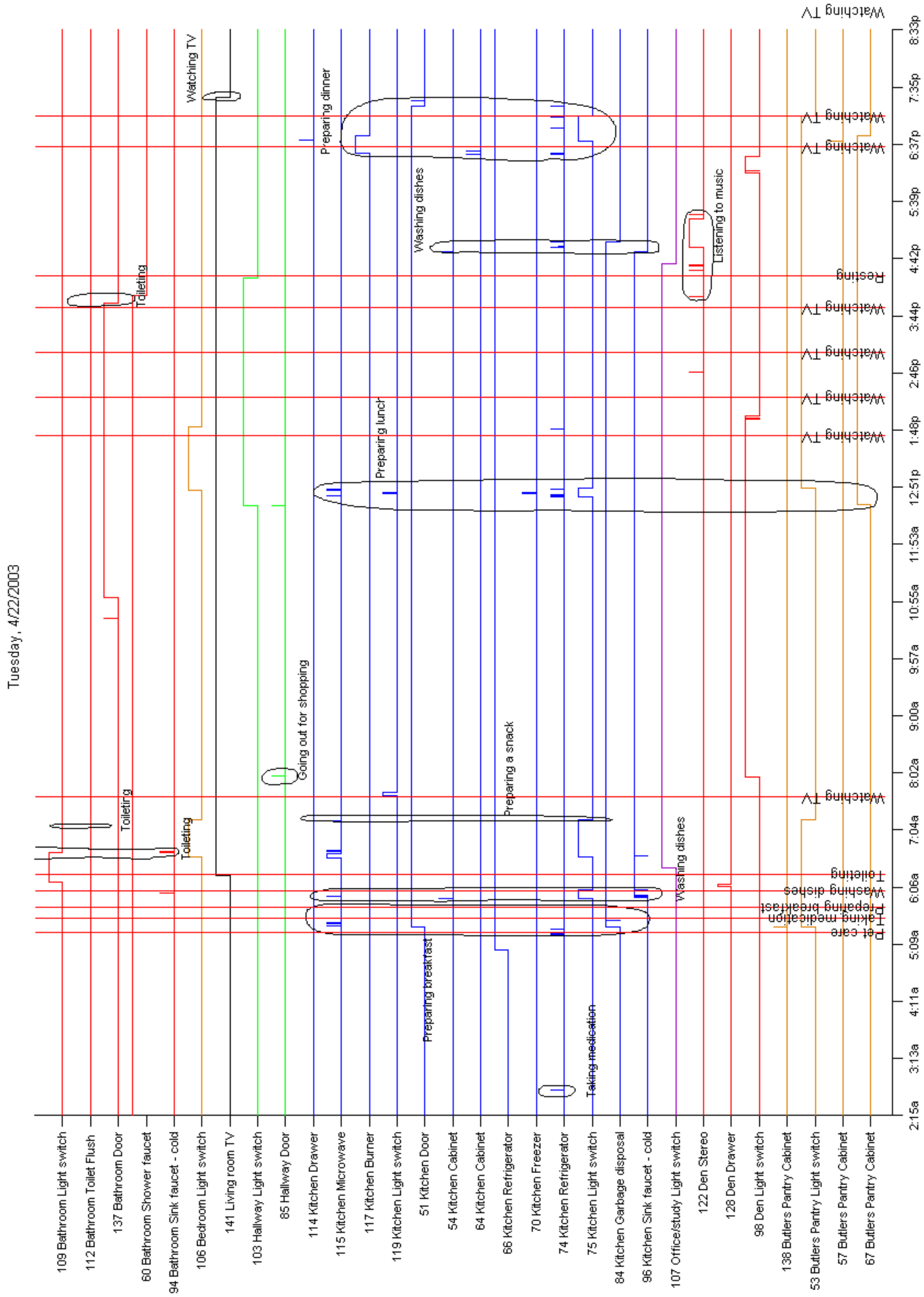


Figure F-4: Plot of one day of sensor data (04/22/2003) for subject two. The labels generated by indirect observation are shown by the round areas and the center time of the ESM labels by the vertical red lines.

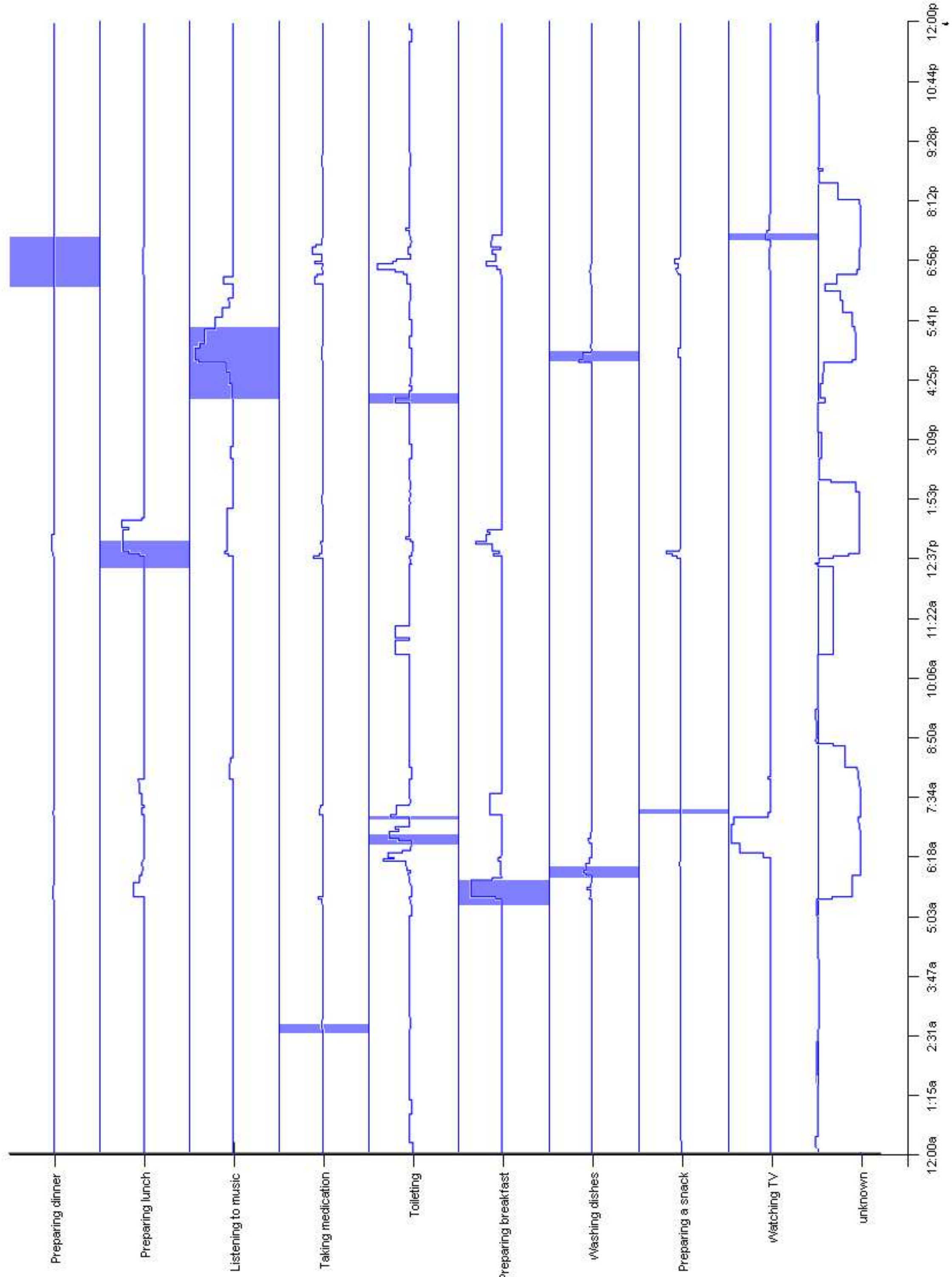


Figure F-5: Plot of the probability classification output of the multiclass naive Bayes classifier for one day of subject's two data (04/22/2003). The blue areas are the activity labels produced by ESM + indirect observation.

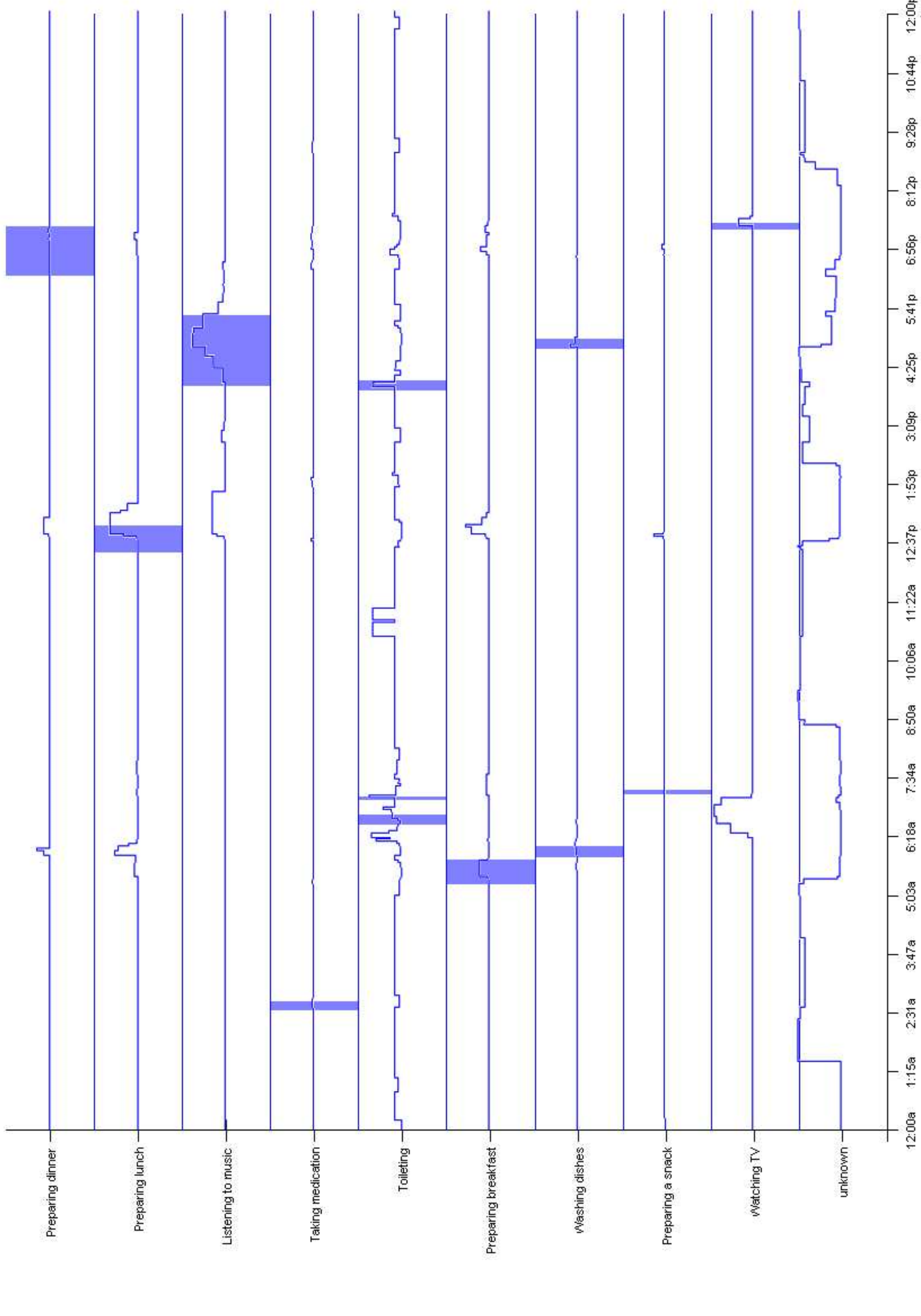


Figure F-6: Plot of the probability classification output of the multiple binary naive Bayes classifiers for one day of subject's two data (04/22/2003). The blue areas are the activity labels produced by ESM + indirect observation.

Appendix G

Sensors Installation Details

Object/Device	Reed Switch Location	Magnet Location
Door	Frame	Door
Window	Frame	Window
Drawers	Frame	Drawer
Cabinets	Frame	Door
Light switch	Switchplate	Switch
Microwave	Body	Door
Refrigerator	Body	Door
Freezer	Body	Door
Toaster	Body	Handle
Oven	Control panel	Dial
Burners	Control panel	Dial
Sink Faucet	Backsplash	Handle
Garbage Disposal	Switchplate	Switch
Washing machine	Body	Door
Dishwasher	Body	Door
Telephone	Receiver	Cradle
TV	Body	Switch
Stereo	Body	Deck
Lamp	Socket	Switch
DVD	Chassis	Tray
Jewelry box	Box	Lid
Containers	Container	Lid
Coffee Machine	Base	Carafe
Toilet	Tank	Flush lever
Medicine Cabinet	Frame	Door
Shower Faucet	Faucet neck	Faucet handle
Hamper	Container	Lid
Closet	Body	Door
Exhaust Fan	Switchplate	Switch

Table G.1: Sensors installation details.

Sensor Location	Number Failures	Possible Cause
Sink Faucet	2	Small installation space + difficulty of installation + high usage
Shower Faucet	2	Difficulty of installation + moisture
Cabinet	2	Difficulty of installation
Drawer	3	Difficulty of installation

Table G.2: Sensors failure during the studies

Sporadic Behavior	**Possible Cause**
Shower Faucet	Magnet/switch alignment + moisture
Faucet Sink	Magnet/switch alignment + moisture + movement
Light Switches	Magnet/switch alignment + movement
Burners	Magnet/switch alignment + movement
Toilet Flush	Magnet/switch alignment + movement
Cabinets	Magnet/switch alignment + movement
Drawers	Magnet/switch alignment + movement

Table G.3: Sensors sporadic behavior details

Appendix H

Result Details

SUBJECT ONE MULTIPLE BINARY CLASSIFIERS

Accuracies per class

Activities	E	BID	BT	BL	E + BID	E + BT	E + BL	Evaluation Method
'Bathing'	0.2822	0.2626	0.3874	0.2027	0.3261	0.4350	0.3162	Percentage of Time that Activity is Detected
'Toileting'	0.2651	0.2206	0.2705	0.1956	0.2640	0.2800	0.2342	
'Going out to work'	0.0313	0	0	0	0	0	0.0313	
'Preparing lunch'	0.3899	0.3077	0.4429	0.2286	0.3291	0.4456	0.5064	
'Preparing dinner'	0	0	0	0.0625	0	0	0	
'Preparing breakfast'	0.0837	0.0774	0.0399	0	0.0607	0.0399	0.0599	
'Dressing'	0.0346	0	0.0192	0	0.0192	0.0346	0.0192	
'Grooming'	0.2361	0.2275	0.1869	0.0582	0.2317	0.1813	0.1512	
'Preparing a snack'	0.0402	0	0.0059	0	0.0118	0.0059	0.0235	
'Preparing a beverage'	0.0555	0.0130	0.0523	0	0.0221	0.0562	0.0386	
'Washing dishes'	0	0	0	0	0	0	0	
'Doing laundry'	0.0865	0.0234	0.0276	0.0138	0.0331	0.0572	0.0816	
'Cleaning'	0.0087	0	0	0	0	0	0	
'Bathing'	0.8333	0.7500	0.8333	0.3750	0.7917	0.8333	0.7500	Activity Detected in Best Interval
'Toileting'	0.7052	0.6759	0.6702	0.6167	0.7015	0.6228	0.6770	
'Going out to work'	0.1875	0	0	0	0	0	0.1875	
'Preparing lunch'	0.8438	0.7188	0.9063	0.4063	0.7813	0.9063	0.8438	
'Preparing dinner'	0.2500	0	0	0.1250	0	0	0.1250	
'Preparing breakfast'	0.4583	0.3750	0.2917	0	0.3750	0.3750	0.3750	
'Dressing'	0.3846	0.1538	0.0769	0.1154	0.2308	0.2692	0.1923	
'Grooming'	0.8698	0.8021	0.7135	0.3385	0.8021	0.6615	0.6771	
'Preparing a snack'	0.2833	0	0	0	0.1500	0.1000	0.2500	
'Preparing a beverage'	0.1364	0.0909	0.2727	0	0.1818	0.2727	0.1364	
'Washing dishes'	0	0	0	0	0	0	0	
'Doing laundry'	0.5111	0.1556	0.3889	0.0667	0.4222	0.4444	0.6000	
'Cleaning'	0	0	0	0	0	0	0.2000	
'Bathing'	0.7917	0.6667	0.7500	0.5417	0.7083	0.7917	0.7083	Activity Detected at Least Once
'Toileting'	0.4050	0.3321	0.3521	0.3532	0.3979	0.3724	0.3757	
'Going out to work'	0.1250	0	0	0	0	0	0.1250	
'Preparing lunch'	0.7813	0.6563	0.8750	0.6875	0.7188	0.8750	0.8125	
'Preparing dinner'	0	0	0	0.2500	0	0	0	
'Preparing breakfast'	0.2083	0.2917	0.1250	0	0.2083	0.1250	0.2083	
'Dressing'	0.1154	0	0.0385	0	0.0385	0.0769	0.0385	
'Grooming'	0.5625	0.6146	0.5104	0.2396	0.6146	0.5104	0.4583	
'Preparing a snack'	0.0833	0	0.0500	0	0.0500	0.0500	0.0500	
'Preparing a beverage'	0.1818	0.0455	0.1818	0	0.0909	0.1818	0.1818	
'Washing dishes'	0	0	0	0	0	0	0	
'Doing laundry'	0.2556	0.1556	0.0667	0.1556	0.1556	0.2333	0.2556	
'Cleaning'	0.0667	0	0	0	0	0	0	

Total Final Accuracies

E	BID	BT	BL	E + BID	E + BT	E + BL	Evaluation Method
0.2384	0.1914	0.2316	0.1150	0.2159	0.2421	0.2198	Time
0.5724	0.4490	0.4788	0.3130	0.5102	0.4832	0.5100	Best Interval
0.3672	0.2977	0.3226	0.2350	0.3365	0.3430	0.3283	Detected

Figure H-1: Detailed results for subject one using multiple binary naive Bayes classifiers. E=*exist*, BID = *before ID*, BT = *before type*, and BL = *before location*.

SUBJECT ONE MULTICLASS CLASSIFIER

Accuracies per Class

Activities	E	BID	BT	BL	E + BID	E + BT	E + BL	Evaluation Method
'Bathing'	0.2534	0.2006	0.3021	0.1365	0.2363	0.2935	0.2307	Percentage of Time that Activity is Detected
'Toileting'	0.2763	0.2400	0.2828	0.1888	0.2676	0.3141	0.2647	
'Going out to work'	0.0313	0	0	0	0	0	0.0313	
'Preparing lunch'	0.2536	0.2229	0.2601	0.1484	0.2166	0.2992	0.3385	
'Preparing dinner'	0	0	0	0.0625	0	0	0.0278	
'Preparing breakfast'	0.0837	0.0726	0.0685	0	0.0518	0.0637	0.0837	
'Dressing'	0.0782	0.0385	0.0192	0.0667	0.0346	0.0346	0.0969	
'Grooming'	0.2678	0.2635	0.2219	0.0732	0.2736	0.2649	0.1793	
'Preparing a snack'	0.0402	0	0.0059	0	0.0176	0.0176	0.0235	
'Preparing a beverage'	0.0737	0.0260	0.0782	0.0114	0.0692	0.1328	0.0964	
'Washing dishes'	0	0	0	0	0	0	0	
'Doing laundry'	0.0930	0.0838	0.0427	0.0155	0.1024	0.0710	0.0930	
'Cleaning'	0.0174	0	0.0174	0.0133	0	0.0174	0	
'Bathing'	0.8750	0.6667	0.7917	0.2083	0.6667	0.7917	0.7083	Activity Detected in Best Interval
'Toileting'	0.7108	0.7448	0.7479	0.6627	0.7203	0.7114	0.7443	
'Going out to work'	0.1250	0	0	0	0	0	0.1250	
'Preparing lunch'	0.5938	0.6563	0.7188	0.2813	0.7188	0.7813	0.7188	
'Preparing dinner'	0.2500	0	0.1250	0.1250	0.1250	0.2500	0.2500	
'Preparing breakfast'	0.4583	0.4583	0.5417	0	0.5417	0.4583	0.3750	
'Dressing'	0.6410	0.1923	0.2179	0.3590	0.3846	0.4103	0.5128	
'Grooming'	0.8906	0.8385	0.8698	0.5104	0.8698	0.8698	0.8073	
'Preparing a snack'	0.2833	0	0	0	0.1833	0.1500	0.2833	
'Preparing a beverage'	0.3636	0.1364	0.3636	0.0455	0.2727	0.3636	0.4545	
'Washing dishes'	0	0	0	0	0	0	0	
'Doing laundry'	0.8667	0.2667	0.7667	0.3444	0.6667	0.7889	0.8444	
'Cleaning'	0	0	0	0.1000	0	0	0	
'Bathing'	0.7083	0.6667	0.7083	0.4583	0.6667	0.7500	0.7083	Activity Detected at Least Once
'Toileting'	0.4228	0.3437	0.3964	0.3812	0.3979	0.4307	0.4050	
'Going out to work'	0.1250	0	0	0	0	0	0.1250	
'Preparing lunch'	0.5000	0.5313	0.7500	0.4375	0.5938	0.6875	0.6875	
'Preparing dinner'	0	0	0	0.2500	0	0	0.1250	
'Preparing breakfast'	0.2083	0.2083	0.1667	0	0.1250	0.1250	0.2083	
'Dressing'	0.2179	0.0769	0.0385	0.1282	0.0769	0.0769	0.2051	
'Grooming'	0.6875	0.6875	0.6354	0.2760	0.7188	0.7188	0.5104	
'Preparing a snack'	0.0833	0	0.0500	0	0.0500	0.0500	0.0500	
'Preparing a beverage'	0.2273	0.0455	0.2273	0.0909	0.1818	0.3182	0.2273	
'Washing dishes'	0	0	0	0	0	0	0	
'Doing laundry'	0.2778	0.2667	0.2111	0.1556	0.2667	0.2333	0.2778	
'Cleaning'	0.0667	0	0.0667	0.1000	0	0.0667	0	

Final Accuracies

E	BID	BT	BL	E + BID	E + BT	E + BL	Evaluation Method
0.2185	0.1784	0.1988	0.1007	0.1957	0.2273	0.2047	Time
0.6062	0.4818	0.5578	0.3755	0.5532	0.5780	0.5910	Best Interval
0.3879	0.3130	0.3558	0.2495	0.3509	0.3861	0.3599	Detected

Figure H-2: Detailed results for subject one using a multiclass naive Bayes classifier. E=*exist*, BID = *before ID*, BT = *before type*, and BL = *before location*.

SUBJECT TWO MULTIPLE BINARY CLASSIFIERS
Accuracies per Class

Activities	E	BID	BT	BL	E + BID	E + BT	E + BL	Evaluation Method
'Preparing dinner'	0.1121	0.0236	0.0554	0.1045	0.1327	0.1327	0.1107	Percentage of Time that Activity is Detected
'Preparing lunch'	0.2517	0.2547	0.2439	0.1699	0.2646	0.2646	0.2306	
'Listening to music'	0.2779	0.2428	0.2666	0.3277	0.2599	0.2599	0.3139	
'Taking medication'	0	0	0	0	0	0	0.0333	
'Toileting'	0.2322	0.1822	0.2550	0.1796	0.2666	0.2666	0.2947	
'Preparing breakfast'	0.3205	0.2426	0.1973	0.1447	0.2152	0.2152	0.2376	
'Washing dishes'	0.0281	0.0133	0.0190	0.0280	0.0655	0.0655	0.0341	
'Preparing a snack'	0	0.0036	0	0.0250	0	0	0	
'Watching TV'	0.0980	0.1321	0.2431	0.3185	0.2291	0.2291	0.2974	
'Preparing dinner'	0.1538	0.1538	0.1923	0.1538	0.2692	0.2692	0.2308	Activity Detected in Best Interval
'Preparing lunch'	0.5167	0.5833	0.5500	0.3833	0.5500	0.5500	0.4167	
'Listening to music'	0.6528	0.6389	0.5972	0.5694	0.6806	0.6806	0.5972	
'Taking medication'	0.1000	0	0	0	0.1000	0.1000	0.1000	
'Toileting'	0.5885	0.3490	0.3333	0.2760	0.4427	0.4427	0.5833	
'Preparing breakfast'	0.6875	0.5625	0.5313	0.4063	0.5938	0.5938	0.5313	
'Washing dishes'	0.2556	0.2333	0.1889	0.1111	0.2778	0.2778	0.3222	
'Preparing a snack'	0	0.0333	0	0.1000	0	0	0	
'Watching TV'	0.2708	0.3958	0.5625	0.6250	0.3750	0.3750	0.5000	
'Preparing dinner'	0.3077	0.1538	0.2308	0.2308	0.3077	0.3077	0.3077	Activity Detected at Least Once
'Preparing lunch'	0.4833	0.5500	0.5833	0.4500	0.6167	0.6167	0.4167	
'Listening to music'	0.6389	0.6389	0.6667	0.6389	0.6667	0.6667	0.6389	
'Taking medication'	0	0	0	0	0	0	0.1000	
'Toileting'	0.5208	0.2813	0.3281	0.2917	0.4375	0.4375	0.5313	
'Preparing breakfast'	0.7500	0.6250	0.4688	0.4063	0.4688	0.4688	0.5938	
'Washing dishes'	0.1000	0.0667	0.0667	0.0889	0.1556	0.1556	0.1667	
'Preparing a snack'	0	0.0333	0	0.1000	0	0	0	
'Watching TV'	0.3125	0.3958	0.5208	0.5625	0.5625	0.5625	0.5625	

Final Accuracies

E	BID	BT	BL	E + BID	E + BT	E + BL	Evaluation Method
0.2083	0.1691	0.2060	0.1764	0.2314	0.2314	0.2293	Time
0.3987	0.3399	0.3333	0.2700	0.3805	0.3805	0.3975	Best Interval
0.3829	0.3098	0.3115	0.2945	0.3646	0.3646	0.3789	Detected

Figure H-3: Detailed results for subject two using multiple binary naive Bayes classifiers. E=*exist*, BID = *before ID*, BT = *before type*, and BL = *before location*.

SUBJECT TWO MULTICLASS CLASSIFIER
Accuracies per Class

Activities	E	BID	BT	BL	E + BID	E + BT	E + BL	Evaluation Method
'Preparing dinner'	0.1411	0.0560	0.0565	0.1375	0.0482	0.1373	0.1680	Percentage of Time that Activity is Detected
'Preparing lunch'	0.2280	0.1862	0.2133	0.1477	0.2339	0.2235	0.1636	
'Listening to music'	0.2096	0.1416	0.0624	0.2726	0.1652	0.0981	0.2344	
'Taking medication'	0	0	0	0	0	0	0.0333	
'Toileting'	0.2080	0.1835	0.2569	0.2847	0.1749	0.2386	0.2670	
'Preparing breakfast'	0.3040	0.2411	0.2077	0.1749	0.2703	0.2445	0.2441	
'Washing dishes'	0.0548	0.0896	0.0958	0.0300	0.0896	0.1176	0.0797	
'Preparing a snack'	0	0.0102	0	0.0317	0	0	0.0100	
'Watching TV'	0.0472	0.1045	0.1917	0.2631	0.1088	0.1613	0.2515	
'Preparing dinner'	0.2308	0.1923	0.1923	0.1538	0.1923	0.2308	0.1538	Activity Detected in Best Interval
'Preparing lunch'	0.5167	0.5167	0.5000	0.3500	0.5167	0.4833	0.3667	
'Listening to music'	0.6111	0.4861	0.4444	0.5417	0.5278	0.4444	0.5694	
'Taking medication'	0.1000	0	0	0	0	0	0.1000	
'Toileting'	0.5260	0.3698	0.3854	0.5052	0.3646	0.4844	0.5573	
'Preparing breakfast'	0.6875	0.6875	0.5313	0.5938	0.6250	0.5938	0.5938	
'Washing dishes'	0.5111	0.4222	0.4444	0.2556	0.4889	0.5444	0.5778	
'Preparing a snack'	0	0.0667	0.0333	0.1000	0	0	0	
'Watching TV'	0.2500	0.3542	0.5208	0.3958	0.4167	0.5208	0.3958	
'Preparing dinner'	0.3846	0.2308	0.2308	0.3077	0.2308	0.3077	0.3077	Activity Detected at Least Once
'Preparing lunch'	0.4833	0.5167	0.5500	0.3667	0.5833	0.6167	0.3667	
'Listening to music'	0.6667	0.5139	0.4306	0.6111	0.5833	0.4583	0.6389	
'Taking medication'	0	0	0	0	0	0	0.1000	
'Toileting'	0.4635	0.3385	0.3333	0.4688	0.3385	0.4323	0.5365	
'Preparing breakfast'	0.7500	0.6875	0.4688	0.4688	0.6875	0.6563	0.5938	
'Washing dishes'	0.1556	0.2667	0.2444	0.1000	0.2667	0.2889	0.2556	
'Preparing a snack'	0	0.0667	0	0.1333	0	0	0.0333	
'Watching TV'	0.0833	0.3958	0.5208	0.4375	0.3958	0.4583	0.4375	

Final Accuracies

E	BID	BT	BL	E + BID	E + BT	E + BL	Evaluation Method
0.2041	0.1491	0.1807	0.1878	0.1683	0.2087	0.2172	Time
0.4109	0.3739	0.3719	0.3295	0.3743	0.4125	0.4066	Best Interval
0.3785	0.3579	0.3351	0.3275	0.3636	0.3814	0.3975	Detected

Figure H-4: Detailed results for subject Two using a multiclass naive Bayes classifier. E=*exist*, BID = *before ID*, BT = *before type*, and BL = *before location*.

Percentage of Time Activity is Detected

a	b	c	d	e	f	g	h	i	j	k	l	m	n	<-- Predicted class
42	47	0	11	11	1	1	7	2	1	0	8	6	26	a
6	97	2	17	0	11	2	2	2	1	0	11	1	148	b
4	10	2	2	0	1	1	12	0	0	0	0	0	4	c
6	11	1	75	2	8	2	0	12	6	3	7	9	91	d
6	3	1	8	0	4	1	4	5	2	0	4	0	26	e
17	20	4	0	0	7	6	11	3	0	0	0	1	13	f
19	11	0	5	0	1	5	8	0	0	0	0	1	10	g
20	33	0	4	1	7	6	54	2	0	0	0	2	50	h
0	13	0	5	0	4	0	3	6	0	0	0	1	19	i
4	4	1	2	0	2	1	4	0	6	0	2	0	31	j
0	4	0	8	0	0	1	0	1	1	0	1	0	9	k
4	14	1	4	0	3	2	4	0	1	1	16	5	34	l
3	10	0	18	0	1	1	0	1	5	5	8	2	10	m
0	0	0	0	0	0	0	0	0	0	0	0	0	0	n

Activity Detected in Best Interval

a	b	c	d	e	f	g	h	i	j	k	l	m	n	<-- Predicted class
13	2	0	1	1	0	0	0	0	0	0	0	0	1	a
2	39	0	3	0	2	1	1	1	0	0	3	0	33	b
1	2	1	1	0	0	0	5	0	0	0	0	0	2	c
0	0	0	9	0	0	0	0	1	0	0	1	1	5	d
1	0	0	0	0	0	0	1	1	0	0	1	0	4	e
1	2	0	0	0	3	2	3	1	0	0	0	0	2	f
6	4	0	1	0	1	5	4	0	0	0	0	0	3	g
1	2	0	2	0	3	1	23	0	0	0	0	0	5	h
0	6	0	2	0	0	0	0	2	0	0	0	0	4	i
0	1	0	1	0	1	0	1	0	4	0	0	0	7	j
0	2	0	1	0	0	1	0	0	0	0	0	0	3	k
0	3	0	1	0	1	1	0	0	0	0	7	1	5	l
0	1	0	2	0	0	0	0	0	0	0	2	1	2	m
0	0	0	0	0	0	0	0	0	0	0	0	0	0	n

a -> Bathing
b -> Toileting
c -> Going out to work
d -> Preparing lunch
e -> Preparing dinner
f -> Preparing breakfast
g -> Dressing
h -> Grooming
i -> Preparing a snack
j -> Preparing a beverage
k -> Washing dishes
l -> Doing laundry
m -> Cleaning
n -> unknown

Activity Detected at Least Once

a	b	c	d	e	f	g	h	i	j	k	l	m	n	<-- Predicted class
15	1	0	0	0	0	0	1	0	0	0	0	0	1	a
3	62	0	1	0	0	0	2	1	0	0	3	1	12	b
1	1	1	1	0	0	0	6	0	0	0	0	0	2	c
0	0	0	10	0	0	0	0	0	0	0	2	1	4	d
1	0	0	2	2	0	0	0	1	1	0	1	0	0	e
2	1	0	0	1	6	1	1	0	0	0	0	1	1	f
4	1	0	1	0	1	14	2	0	0	0	0	0	1	g
1	2	0	0	0	0	0	32	0	0	0	0	0	2	h
0	3	0	1	0	1	0	0	4	0	0	0	0	5	i
0	1	0	0	1	0	0	1	0	6	0	0	0	6	j
1	1	0	2	0	0	0	0	0	0	0	0	0	3	k
1	0	0	0	0	0	0	0	0	0	0	17	0	1	l
1	1	0	3	0	0	0	0	0	0	0	2	0	1	m
0	0	0	0	0	0	0	0	0	0	0	0	0	0	n

Figure H-5: Confusion Matrices for Subject One Using a Multiclass Naive Bayes Classifiers and the *exist* feature.

Percentage of Time Activity was Detected

a	b	c	d	e	f	g	h	i	j	<-- Predicted class
38	47	5	4	11	6	7	0	0	79	a
18	87	31	3	26	25	18	1	0	103	b
1	45	105	2	28	9	10	3	0	236	c
0	0	0	0	15	0	5	0	0	51	d
6	17	11	1	57	20	1	1	9	123	e
0	3	0	6	5	60	10	2	6	81	f
3	27	14	0	25	6	9	0	1	30	g
3	2	10	0	12	13	5	0	2	68	h
9	3	18	0	51	12	3	0	27	153	i
0	0	0	0	0	0	0	0	0	0	j

Detected in Best Interval

a	b	c	d	e	f	g	h	i	j	<-- Predicted class
5	1	0	0	1	0	0	0	0	7	a
1	10	1	0	0	1	1	0	0	6	b
0	0	12	0	0	0	1	0	0	5	c
0	0	0	0	3	0	0	0	0	11	d
1	2	2	0	18	1	0	0	2	14	e
0	0	0	0	0	14	0	0	1	3	f
1	4	3	0	5	1	4	0	0	3	g
0	0	1	0	2	2	0	0	0	11	h
0	1	0	0	3	1	1	0	2	7	i
0	0	0	0	0	0	0	0	0	0	j

a -> Preparing dinner
b -> Preparing lunch
c -> Listening to music
d -> Taking medication
e -> Toileting
f -> Preparing breakfast
g -> Washing dishes
h -> Preparing a snack
i -> Watching TV
j -> unknown

Detected at Least Once

a	b	c	d	e	f	g	h	i	j	<-- Predicted class
3	4	1	0	0	2	0	0	0	4	a
1	11	1	0	1	1	0	0	0	5	b
0	2	9	0	0	1	0	0	0	6	c
0	0	0	1	2	0	2	0	0	9	d
1	0	2	0	20	1	1	0	1	14	e
0	2	0	0	0	13	1	0	1	1	f
1	2	1	0	3	0	11	0	0	3	g
1	1	4	0	2	1	2	0	0	5	h
1	0	0	0	3	0	1	0	3	7	i
0	0	0	0	0	0	0	0	0	0	j

Figure H-6: Confusion Matrices for Subject Two Using a Multiclass Naive Bayes Classifiers and the *exist* feature.

Percentage of Time Activity is Detected

a	b	c	d	e	f	g	h	i	j	k	l	m	n	<-- Predicted class
46	44	0	12	9	1	1	4	2	2	0	7	2	33	a
5	114	1	16	0	3	2	2	3	2	1	7	2	142	b
0	17	0	2	0	0	0	11	0	0	0	0	2	4	c
1	16	0	77	2	13	0	0	19	8	1	4	12	80	d
8	2	0	9	0	7	1	4	4	7	0	3	0	19	e
13	33	0	0	0	6	7	3	3	0	0	1	3	13	f
16	17	0	6	0	2	2	9	0	2	0	0	0	6	g
13	33	0	7	2	4	5	53	1	0	1	0	4	56	h
1	16	0	4	0	4	0	2	3	1	0	0	2	18	i
0	12	0	2	0	1	0	4	0	9	0	1	0	28	j
0	5	0	7	0	0	0	1	2	0	0	1	0	9	k
2	21	0	5	0	3	1	4	2	0	0	12	3	36	l
4	14	0	14	0	1	2	0	2	2	4	5	2	14	m
0	0	0	0	0	0	0	0	0	0	0	0	0	0	n

Activity Detected in Best Interval

a	b	c	d	e	f	g	h	i	j	k	l	m	n	<-- Predicted class
13	1	0	2	0	0	0	0	0	0	0	0	0	2	a
2	41	0	4	0	1	0	0	0	1	1	2	0	33	b
0	4	0	1	0	0	0	4	0	0	0	0	1	2	c
0	0	0	12	0	0	0	0	1	0	0	0	2	2	d
1	0	0	1	0	0	0	1	1	0	0	1	0	3	e
4	3	0	0	0	2	1	1	0	0	0	0	1	2	f
5	7	0	2	0	1	2	5	0	1	0	0	0	1	g
0	1	0	2	1	1	1	24	0	0	0	0	0	7	h
0	7	0	0	0	0	0	1	1	0	0	0	0	5	i
0	3	0	1	0	0	0	1	0	5	0	0	0	5	j
0	2	0	1	0	0	0	1	0	0	0	0	0	3	k
0	4	0	1	0	2	0	0	0	0	0	5	1	6	l
0	1	0	3	0	0	0	0	0	0	0	1	1	2	m
0	0	0	0	0	0	0	0	0	0	0	0	0	0	n

a -> Bathing
b -> Toileting
c -> Going out to work
d -> Preparing lunch
e -> Preparing dinner
f -> Preparing breakfast
g -> Dressing
h -> Grooming
i -> Preparing a snack
j -> Preparing a beverage
k -> Washing dishes
l -> Doing laundry
m -> Cleaning
n -> unknown

Activity Detected at Least Once

a	b	c	d	e	f	g	h	i	j	k	l	m	n	<-- Predicted class
14	2	0	0	0	0	0	1	0	0	0	0	0	1	a
2	63	0	3	0	0	0	0	0	0	0	3	2	12	b
0	4	0	1	0	0	0	4	0	0	0	0	1	2	c
0	0	0	13	0	0	0	0	1	0	0	1	0	2	d
1	0	0	3	2	1	0	0	0	1	0	0	0	0	e
2	1	0	1	1	6	1	0	0	0	0	0	0	2	f
6	3	0	1	0	0	9	4	0	0	0	0	0	1	g
0	2	0	0	0	0	1	31	0	0	0	0	0	3	h
0	5	0	1	0	0	0	0	2	1	0	0	0	5	i
0	2	0	1	0	0	0	1	0	6	0	0	0	5	j
0	2	0	1	0	0	0	1	0	0	0	0	0	3	k
1	1	0	0	0	0	0	0	0	0	0	15	0	2	l
1	1	0	3	0	0	0	0	1	1	0	1	0	0	m
0	0	0	0	0	0	0	0	0	0	0	0	0	0	n

Figure H-7: Confusion Matrices for Subject One Using a Multiclass Naive Bayes Classifiers and the *exist+Before Type* feature.

Percentage of Time Activity was Detected

a	b	c	d	e	f	g	h	i	j	<-- Predicted class
37	38	5	10	27	15	11	3	3	48	a
33	84	36	0	42	26	20	4	0	67	b
15	55	60	0	103	0	23	1	3	179	c
0	0	0	0	15	0	0	0	0	56	d
9	9	3	7	68	22	2	3	11	112	e
7	4	1	9	22	48	4	1	7	70	f
3	21	12	0	35	4	15	1	4	20	g
6	3	10	2	15	7	7	0	5	60	h
8	1	3	3	85	8	7	3	59	99	i
0	0	0	0	0	0	0	0	0	0	j

Detected in Best Interval

a	b	c	d	e	f	g	h	i	j	<-- Predicted class
4	1	1	0	2	1	0	0	1	4	a
2	13	1	0	0	0	1	0	0	3	b
1	3	8	0	4	0	0	0	0	2	c
0	0	0	0	3	0	0	0	0	11	d
2	1	0	1	19	1	0	0	2	14	e
1	0	0	0	0	12	0	0	1	4	f
1	3	1	0	7	0	6	0	1	2	g
1	1	1	0	2	1	1	0	0	9	h
0	0	0	0	3	0	1	0	6	5	i
0	0	0	0	0	0	0	0	0	0	j

a -> Preparing dinner
b -> Preparing lunch
c -> Listening to music
d -> Taking medication
e -> Toileting
f -> Preparing breakfast
g -> Washing dishes
h -> Preparing a snack
i -> Watching TV
j -> unknown

Detected at Least Once

a	b	c	d	e	f	g	h	i	j	<-- Predicted class
3	4	1	1	1	1	0	0	0	3	a
3	11	1	0	1	1	1	0	0	2	b
1	3	7	0	4	0	1	0	0	2	c
0	0	0	0	3	0	0	0	0	11	d
1	0	0	1	21	2	1	0	1	13	e
1	2	0	1	1	11	0	0	1	1	f
1	3	1	0	2	0	11	0	1	2	g
1	1	3	1	2	0	3	0	0	5	h
1	0	0	0	2	0	1	0	7	4	i
0	0	0	0	0	0	0	0	0	0	j

Figure H-8: Confusion Matrices for Subject two Using a Multiclass Naive Bayes Classifiers and the *exist+Before Type* feature.

122

Appendix I

Hardware and Data Availability

All hardware design specifications, raw and annotated data collected for this work is freely available to the public for research purposes. Please contact Emmanuel Munguia Tapia [emunguia@media.mit.edu] or Stephen Intille [intille@mit.edu] for information on acquiring data. The entire dataset is approximately 25MB.

Appendix J

Number of Examples per Activity

	Subject 1		Subject 2	
Activity	ESM	I.O	ESM	I.O
Work at home	-	-	-	-
Going out to work	5	12	-	-
Eating	20	-	31	-
Toileting	2	85	14	40
Bathing	12	18	1	3
Grooming	6	37	5	3
Dressing	8	24	8	5
Washing hands	-	1	-	-
Taking medication	-	-	10	14
Sleeping	5	-	3	-
Talking on telephone	7	-	1	4
Resting	-	-	24	-
Preparing breakfast	3	14	12	18
Preparing lunch	11	17	12	20
Preparing dinner	3	8	10	14
Preparing a snack	5	14	7	16
Preparing a beverage	5	15	1	1
Washing dishes	-	7	16	21
Putting away dishes	-	2	1	3
Putting away groceries	3	2	1	1
Cleaning	9	8	3	3
Doing laundry	4	19	-	-
Putting away laundry	-	2	1	1
Taking out the trash	-	-	-	-
Lawnwork	1	1	1	1
Pet care	-	-	8	-
Home education	-	-	7	2
Going out to school	-	-	-	-
Watching TV	5	3	80	15
Listening to music	-	-	1	18
Going out for entertainment	-	1	-	1
Working at computer	-	-	5	5
Going out to exercise	1	-	-	-
Going out for shopping	3	2	5	3
unknown	79	274	225	165

Table J.1: Number of examples per activity generated by ESM and indirect observation(I.O) for subject 1 and subject 2.

Bibliography

[1] Clinical and financial analyses of programs in congestive heart failure. Technical report, Health Hero Network. files Available from http://www.hhn.com.

[2] Elite care's oatfield estates. http://www.elite-care.com/oatfield-tech.html.

[3] Chronic care in america: A 21st century challenge. Technical report, Robert Wood Johnson Foundation, 1996.

[4] Center for aging services technology meeting. American Association of Homes Services for the Aging, 2002.

[5] G.D. Abowd. Director of the AwareHome initiative, Georgia Insitute of Technology, November 2002.

[6] U.S. Senate Special Committee on Aging Hearing. Long term care: Who will care for the baby boomers?, June 2001.

[7] D. Aha and D. Kibler. Instance-based learning algorithms. *Machine Learning*, 6:37–66, 1991.

[8] B. E. Ainsworth, W. L. Haskell, M. C. Whitt, M. L. Irwin, A. M. Swartz, S. J. Strath, W. L. O'Brien, Jr. Bassett, D. R., K. H. Schmitz, P. O. Emplaincourt, Jr. Jacobs, D. R., and A. S. Leon. Compendium of physical activities: An update of activity codes and met intensities. *Medicine & Science in Sports & Exercise*, 32(9 Suppl):S498–504, 2000. 7998 United States ISSN: 0195-9131; NLM Unique Journal Identifier: 8005433 English Grant Number: HL-362-72; Grant Info: HL; Agency: NHLBI Grant Number: RFA-86-37561; Agency: PHS Grant

Number: SIP 22W-U48/CCU 409554-03; Agency: PHS Grant Number: etc Index Medicus; Spaceline Journal Article Citation: Status: Completed Owner: NLM.

[9] E. L. Allwein, R. E. Schapire, and Y. Singer. Reducing multiclass to binary: A unifying approach for margin classifiers. *Machine Learning Research*, pages 113–141, December 2000.

[10] K. Aminian, P. Robert, E. Jequier, and Y. Schutz. Estimation of speed and incline of walking using neural network. *IEEE Transactions on Instrumentation and Measurement*, 44(3):743–746, 1995.

[11] S. Antifakos, F. Michahelles, and B. Schiele. Proactive instructions for furniture assembly. In G. Borriello and L.E. Holmquist, editors, *Proceedings of UbiComp 2002: Ubiquitous Computing*, LNCS 2498, pages 351–360. 2002.

[12] L. Bao. *Physical activity recognition from acceleration data under semi-naturalistic conditions.* M.Eng thesis, EECS, Massachusetts Institute of Technology, 2003.

[13] T. Barger, M. Alwan, S. Kell, B. Turner, S. Wood, and A. Naidu. Objective remote assessment of activities of daily living: Analysis of meal preparation patterns. Poster presentation, Medical Automation Research Center, University of Virginia Health System, 2002.

[14] J. S. Beaudin. *From personal experience to design: Externalizing the home-owner's needs and assesment process.* MS thesis, MIT Media Laboratory, Massachusetts Institute of Technology, 2003.

[15] R. Beckwith and S. Lederer. Designing for one's dotage: Ubicomp and residential care facilities. In *Home Oriented Informatics and Telematics (HOIT)*, Irvine, CA, 2003.

[16] M. Brand, N. Oliver, and A. Pentland. Coupled hidden markov models for complex action recgonition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 994–999. Los Alamitos, CA, 1997.

[17] CHF. Smart home diagram, 2002.

[18] T. Choudhury and A. Pentland. The sociometer: A wearable device for understanding human networks. MIT media lab technical report 554, Massachusetts Institute of Technology, 2002.

[19] N. Chungfat. Context aware activity recognition using tan classifiers. M.Eng thesis, EECS, Massachusetts Institute of Technology, 2002.

[20] P. Clark and T. Niblett. The cn2 induction algorithm. *Machine Learning*, 3(4):261–283, 1989.

[21] B. Clarkson and A. Pentland. Predicting daily behavior via wearable sensors. Technical report, MIT Media Lab, Massachusetts Institute of Technology, July 2001.

[22] B.P. Clarkson and A. Pentland. Unsupervised clustering of ambulatory audio and video. In *Proceedings of the International Conference of Acoustics, Speech and Signal Processing*. 1999.

[23] B. Peintner Colbry, D. and M. E. Pollack. Execution monitoring with quantitative temporal bayesian networks. *Sixth International Conference on AI Planning & Scheduling*, March 2002.

[24] M. Csikszentmihalyi and R. Larson. Validity and reliability of the experience-sampling method. *The Journal of Nervous and Mental Disease*, 175(9):526–36, 1987. 9026 United States ISSN: 0022-3018; NLM Unique Journal Identifier: 0375402 English Abridged Index Medicus; Index Medicus Journal Article Citation: Status: Completed Owner: NLM.

[25] S. K. Das, D. J. Cook, A. Bhattacharya, E. O. Heierman, and T.Y. Lin. The role of prediction algorithms in the MavHome smart home architecture. In IEEE Wireless Communications, editor, *IEEE*, 2002.

[26] J.W. Davis and A.F. Bobick. The representation and recognition of action using temporal templates. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 928–934. IEEE Press, 1997.

[27] P. Domingos and M. Pazzani. Beyond independence: Conditions for the optimality of a simple bayesian classifier. In L. Saitta, editor, *Proceedings of the Thirteenth International Conference on Machine Learning*, pages 194–202. Morgan Kauffman, 1996.

[28] R.O. Duda and P.E. Hart. *Pattern Classification And Scene Analysis*. New York, 1973.

[29] J. Farringdon, A. Moore, N. Tilbury, and J. Church P.D. Biemond. Wearable sensor badge & sensor jacket for context awareness. In *3rd International Symposium on Wearable Computers*. San Francisco, CA, 1999.

[30] M. Feldmeier, M. Malinowski, and J.A. Paradiso. Large group musical interaction using disposable wireless motion sensors. In *Proceedings of the ICMC 2002 Conference*, pages 83–87. San Francisco, CA, 2002.

[31] F. Foerster and J. Fahrenberg. Motion pattern and posture: correctly assessed by calibrated accelerometers. *Behavior Research Methods, Instruments, & Computers*, 32(3):450–7, 2000. Find United States ISSN: 0743-3808; NLM Unique Journal Identifier: 8413015 English Index Medicus Journal Article Citation: Status: Completed Owner: NLM.

[32] J. Friedman. On bias, variance, 0/1 - loss, and the curse-of-dimensionality. *Data mining and knowledge engineering*, 1:55–77, 1997.

[33] J. Friedman, D. Geiger, and M. Goldszmidt. Bayesian network classifiers. *Machine Learning*, 29:131–163, 1997. 9801.

[34] N. Friedman, D. Koller, and A. Pfeffer. Structured representation of complex stochastic systems. *Proceedings of the 15th National Conference on Artificial Intelligence (AAAI)*, pages 157–164, 1998.

[35] A. Garg, V. Pavlovic, and J.M. Rehg. Audio-visual speaker detection using dynamic bayesian networks. *IEEE International Conference in Automatic Face and Gesture Recognition*, 2000.

[36] I. Haritaoglu, D. Harwood, and L. Davis. W4: Who, When, Where, What: A real time system for detecting and tracking people. In *Third International Conference on Automatic Face and Gesture*. Nara, 1998.

[37] Timothy David Hirzel. Visualizing exercise hidden in everyday activity. B.Sc Thesis, 2000.

[38] S. Hollar. *COTS Dust*. Ph.D. thesis, University of California, Berkeley, 2001.

[39] J. Hyman. *Computer vision based people tracking for motivating behavior in public spaces*. M.Eng thesis, EECS, Massachusetts Institute of Technology, 2003.

[40] S.S. Intille. Preventative health care: An MIT proposal for supporting lifelong healthy behavior, 2002. Notes from a meeting presenting a "straw man" proactive healthcare proposal.

[41] S.S. Intille and A.F. Bobick. Recognizing planned, multi-person action. *Computer Vision and Image Understanding (1077-3142)*, 81(3):414–445, 2001.

[42] S.S. Intille, J. Davis, and A. Bobick. Real-time closed-world tracking. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 697–703. IEEE Press, 1997.

[43] S.S. Intille, K. Larson, and C. Kukla. Just-in-time context-sensitive questioning for preventative health care. In *Proceedings of the AAAI 2002 Workshop on Automation as Caregiver: The Role of Intelligent Technology in Elder Care*, AAAI Technical Report WS-02-02. AAAI Press, Menlo Park, CA, 2002.

[44] S.S. Intille, J. Rondoni, C. Kukla, I. Anacona, and L. Bao. A context-aware experience sampling tool. In *Proceedings of the Conference on Human Factors and Computing Systems: Extended Abstracts*. ACM Press, 2003.

[45] J. Himberg J. Mantyjarvi and T. Seppanen. Recognizing human motion with multiple acceleration sensors. *IEEE International Conference on Systems, Man, and Cybernetics.*, pages 747–52, 2001.

[46] G.H. John and P. Langley. Estimating continuous distributions in bayesian classifiers. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*. San Mateo, 1995. 1090, continuous belief networks.

[47] J.M. Kahn, R.H. Katz, and K.S.J. Pister. Mobile networking for smart dust. In *ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom 99)*. 1999.

[48] K. Kanazawa. A logic and time nets for probabilistic inference. *Proceedings of the 9th National Conference on Artificial Intelligence(AAAI)*, 5:142–150, 1991.

[49] O. Kasten and M. Langheinrich. First experiences with bluetooth in the smart-its distributed sensor network. In *Workshop on Ubiquitous Computing and Communications, PACT*. 2001.

[50] H. Kautz, O. Etziono, D. Fox, and D. Weld. Foundations of assisted cognition systems. technical report cse-02-ac-01, University of Washington, Department of Computer Science and Engineering, 2003.

[51] N. Kern, B. Schiele, H. Junker, and P. Lukowicz. Wearable sensing to annotate meeting recordings. In *International Symposium on Wearable Computers (ISWC)*. IEEE Press, 2002.

[52] P. Langley, W. Iba, and K. Thompson. An analysis of bayesian classifiers. In *Proceedings of the Tenth National Conference on Artificial Intelligence*. AAAI Press, San Jose, CA, 1992.

[53] M.P. Lawton. Aging and performance on home tasks. *Human Factors*, 32:527–536, 1990.

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

[55] S.W Lee and K. Mase. Activity and location recognition using wearable sensors. *IEEE Pervasive Computing*, 1(3):24–32, 2002.

[56] M. Makikawa and H. Iizumi. Development of an ambulatory physical activity monitoring device and its application for categorization of actions in daily life. In *MEDINFO*, pages 747–750. 1995.

[57] M. Makikawa, S. Kurata, Y. Higa, Y. Araki, and R. Tokue. Ambulatory monitoring of behavior in daily life by accelerometers set at both-near-sides of the joint. In *Proceedings of Medinfo*, volume 10(Pt) 1, pages 840–3. 2001. 6077 Netherlands ISSN: 1569-6332; NLM Unique Journal Identifier: 7600347 English Index Medicus Journal Article Citation: Status: Completed Owner: NLM.

[58] K. Manabe, T. Matsui, M. Yamaya, T. Sato-Nakagawa, N. Okamura, H. Ari, and H.Sasaki. Sleeping patterns and mortality among elderly patients in geriatric hospitals. *Gerontology*, 46:318–322, 2000.

[59] D.J. Moore, I.A. Essa, and M.H. Hayes III. Exploiting human actions and object context for recognition tasks. In *Proceedings of the 7th IEEE International Conference on Computer Vision*. IEEE Press, 1999.

[60] M. Mozer. The neural network house: an environment that adapts to its inhabitants. In *Proceedings of the AAAI Spring Symposium on Intelligent Environments*, Technical Report SS-98-02, pages 110–114. AAAI Press, Menlo Park, CA, 1998.

[61] K. P. Murphy. *Dynamic Bayesian Networks: Representation, Inference and Learning.* PhD thesis, University of California, Berkeley, 2002.

[62] B.D. Mynatt and W.A. Rogers. Developing technology to support the functional independence of older adults. *Aging International*, 27(1):24–41, 2002.

[63] B. Ornstein. Care technology: Smart home system for the elderly. In *Proceedings of NIST Pervasive Computing*. 2001.

[64] R.J. Orr and G.D. Abowd. The smart floor: A mechanism for natural user identification and tracking. In *Proceedings of the 2000 Conference on Human Factors in Computing Systems (CHI 2000)*. ACM Press, 2000.

[65] J. Quinlan. *C4.5: Programs for machine learning*. 1993.

[66] L.R. Rabiner and B.-H. Juang. Fundamentals of speech recognition. 1993.

[67] J.P. Robinson. The validity and reliability of diaries versus alternative time use measures. In F.T. Juster and F.P. Stafford, editors, *Time Goods and Well-Being*, pages 33–62. Ann Arbor, MI, 1999. Find Methodological Features of the Time Diary Method.

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

[69] J. C. Rondoni. *Context-Aware Experience Sampling for the Design and Study of Ubiquitous Technologies*. M.Eng thesis, EECS, Massachusetts Institute of Technology, 2003.

[70] D. A. Scheon. *The Reflective Practitioner: How Professionals Think in Action*. Basic Books, New York, 1983. Donald A. Schon. 22 cm.

[71] SQUIRTS. The crystal tag system. http://acg.media.mit.edu/people/rich/, 1999.

[72] T. E. Starner and A. Pentland. Visual recognition of american sign language using hidden markov models. In *Proceedings of the International Workshop on Automatic Face and Gesture Recognition*. 1995.

[73] S. Stillman, R. Tanawongsuwan, and I. Essa. A system for tracking and recognizing multiple people with multiple cameras. In *Proceedings of the Second International Conference on Audio-Vision-based Person Authentication*. 1999.

[74] A.A. Stone and S. Shiffman. Ecological momentary assessment (ema) in behavioral medicine. *Annals of Behavioral Medicine*, 16(3):199–202, 1994. 6070.

[75] M. Summers, G. Johnston, and F. Capria. "don't tell me. show me: How video transforms the way we analize and report data from field studies". In *11th Anual UPA Conference- Huminizing Design*, Orlando, Florida, 2002.

[76] X. Sun and C.W. Chen. Probabilistic motion parameter models for human activity recognition. *IEEE International Conference on Pattern Recognition (ICPR02)*, 1, 2002.

[77] C. Sweeney. The middle of the middle class. *New York Times Magazine*, page 74, 2002.

[78] S. Szalai. *The Use of Time. Daily Activities of Urban and Suburban Populations in Twelve Countries*. Mouton, The Hague, 1973. Edited by Alexander Szalai. In collaboration with Philip E. Converse [and others] 25 cm.

[79] J.B. Tenenbaum and W.T. Freeman. Separating style and content. *Advances in Neural Information Processing Systems*, 9, 1997.

[80] J. H. Tulen, D. L. Stronks, J. B. Bussmann, L. Pepplinkhuizen, and J. Passchier. Towards an objective quantitative assessment of daily functioning in migraine: A feasibility study. *Pain*, 86(1-2):139–49, 2000. 6082 Netherlands ISSN: 0304-3959; NLM Unique Journal Identifier: 7508686 English Index Medicus Clinical Trial; Journal Article Citation: Status: Completed Owner: NLM.

[81] M. Uiterwaal, E. B. Glerum, H. J. Busser, and R. C. van Lummel. Ambulatory monitoring of physical activity in working situations, a validation study. *Journal of medical engineering & technology.*, 22(4):168–72, 1998.

[82] G. Vallejo. Listenin: Ambient auditory awareness at remote places. MS thesis, MIT Media Laboratory, Massachusetts Insitute of Technology, September 2003.

[83] R. A. Winett, K. D. Kramer, W. B. Walker, S. W. Malone, Steven W., et al. Modifying food purchases in supermarkets with modeling, feedback, and goal-setting procedures. *Journal of Applied Behavior Analysis*, 21(1):73–80, 1988. Find.

[84] I.H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufmann, 1999.