國立臺灣大學電機資訊學院資訊工程研究所

碩士論文

Graduate Institute of Computer Science and Information Engineering

College of Electrical Engineering and Computer Science

National Taiwan University

Master Thesis

基於智慧空間之銀髮族日常生活活動觀測照護系統

Activity of Daily Living-aware of Healthcare for Elderly in Pervasive Environment

陳雅虹

Ya-Hung Chen

指導教授：傅立成 博士

Advisor: Li-Chen Fu, Ph.D.

中華民國 104 年 7 月

July, 2015

中文摘要

由於醫療進步與致死率降低，人口老化現象逐步加劇，且隨著許多國家伴隨著少子化現象，導致年輕族群將面臨撫養多位長者巨大的壓力。且老年族群若伴隨著不良的生活型態將有較高的致死率，故在非臨床環境下能隨時隨地監測年長的的生理狀態將帶給家屬與醫護人員許多的幫助。故本研究將致力於開發一套具有自動監測環境中長者日常生活活動的智慧環境照護系統，提出一整套創新的非參數多階層之非監督式學習活動辨識模型，解決以往建立活動辨識方法時標記活動之困難。並於即時偵測日常生活活動之模組提供適應性功能，該功能能隨著年長者生活型態改變，發先新的活動且即時提供適當的服務內容，並重建分類模型。且為了提升活動辨識的準確度，本研究整合穿戴式感測器與環境感測器兩類型異質感測器整合分析方法，改良以往單種感測器之活動辨識方法，使得本系統能夠同時分析多樣性的日常生活活動，並與物聯網(Internet of Things)理念整合，簡易的使用介面亦使得智慧居家照護系統有機會布建於真實家庭中，本系統可輔佐瞭解居住者的生活型態，並可在未來發展用以監測異常狀態或異常生活型態，逐步提升居住者的健康狀態。且我們於實驗中證實本系統之活動辨識準確率高達 97.67%，並將大量的訓練集資料簡化成約 15 種分類，故使用者僅需標註這 15 種分類分別代表的日常生活活動，大幅降低標記資料的困難。

關鍵字:在地老化、活動辨識、適應性學習模型、智慧照護系統、物聯網

ABSTRACT

The high development of medicine causes the world's population aging quickly. To resolve the problem with limited medical resources, constant monitoring of elders' activity of daily living is important. We propose an activity recognition system for smart home, so elders can live alone and their children can monitor their parents' living activity to achieve the concept of "Aging in Place". The living activity monitoring model is powerful to recognize meaningful activities by using both ambient and wearable sensors. It's feasible to deploy in the real living environment because it is a non-parametric learning model. Elders need less effort to label activity in training part, and the model may have chance to find some special activities that the elders did not consider in the past. And the proposed activity recognition system can discover new activity that not appears in the training stage. We use the mechanism of case-based reasoning to achieve immediately providing service and the function of adaptive learning. The case-based reasoning will find the most similar known activity and provide the same service for user. If the elderly user occurs a serious abnormal situation, the system could notify caregiver immediately by the mechanism of case-based reasoning. And the system will ask user to confirm this new activity and re-train the online model. So, if the activity occurs again, the system could provide service without confirming. We invited several users to test the system, and the average precision of online activity recognition is up to 97.67%. The experiment result demonstrates the activity of daily living-aware elderly health system is feasible to be deployed in a real life home envi-ronment with high pre-cision performance of the activity recognition result. ~~So, if the activity occurs again, the system could provide service without confirming. We invited serveral users to test the system, and the average precision of online activity recognition is up to 97.67%. The~~

~~experiment result demonstrates the activity of daily living aware elderly health system is feasible to be deployed in a reallife home environment with high precision performance of the activity recognition result.~~

Keyword: Agin in Place, Activity Recognition, Adaptive Learning Model, Internet of Things

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Due to the high development of medicine and the success of reducing mortality, the world's population has aged quickly. As we all knew , the elderly are usually under high risk because of deteriorating health conditions, so the needs for monitoring their physiological state in non-clinical setting is critically important [1, 2]. With declining birthrate nowadays, taking care of elders will gradually bring burden and pressure to their children. Currently, a large portion of elders live independently. That is the reason why most elders would prefer to age in place or to remain in their accustomed home of choice as long as possible. About 72% of elders who are 85 years old or above live by themselves or with spouses in their own houses in United States (the 2012 American Community Survey[]). The institution "Centers for Disease Control and Prevention (CDC)" in United States defines aging in place as "the ability to live in one's own home and community safely, independently, and comfortably, regardless of age, income, or ability level." Although the concept of "Aging in Place" for elders to live in their own houses has been proposed for long time, the risks the elders are facing still exist, *e.g.*

fall, loss of autonomy, etc. Technology can be an important role for aging in place to assist elders in their own home. There are four categories of technology for aging in place, namely Communication and Engagement, Health and Wellness, Learning and Contribution, and Safety and Security. Due to advances in activity recognition technology development, one is able to build a daily activity monitoring system in home environment. It is generally known that activity of daily living (ADL) is an important factor to estimate the independent ability of elders. Barthel index is used to measure performance in activities of daily living and to assess whether the elders have independent ability [3]. Thus, monitoring the ADL of elders to measure their ability with the help from some advanced technology can improve the safe living conditions at home.

On the other hand, wireless sensor network (WSN) is a well-developed technology in recent years and a lot of interesting products here been produced, *e.g.* mobile pad, smart watch or raspberry Pi with ambient sensors. With extensive technological development of WSN, the Internet of Things (IoT)becomes overwhelming in applications comcerning humans' life. Specifically, the IoT related technologies include ubiquitous computing where computing is made to appear everywhere and anywhere [4], Ambient Intelligence (AmI) where devices can work in concert to support people in carrying out their daily life activities and tasks more easily, and wearable computing where the associated devices can provide specific, limited features like pedometer, and provide advanced smart function [5]. Conceivably, technology of IoT is an important factor to implement a smart environment to monitor residents' daily activities.

~~1.2 According to the high development of medicine and the success of reducing mortality, the world's popula-~~

~~tion has aged quickly. The elderly people suffer from high risk due to poor health conditions, so the needs for monitoring human's physiological state in non-clinical setting is critically importance [1, 2]. With the declining birthrate phenomenon, taking care of elder family will gradually bring burden and pressure for their children. Currently, a large portion of elders live independently. That is the reason of most adults and elders would prefer to age in place which means remain in their home of choice as long as possible. About 72% of elders who are 85 years old or above live by themselves or with spouse in their own houses in United States (the 2012 American Community Survey). The institution "Centers for Disease Control and Prevention (CDC)" defines aging in place as "the ability to live in one's own home and community safely, independently, and comfortably, regardless of age, income, or ability level." Although the concept of "Aging in Place" for elders to live in their own houses has been~~

~~proposed for long time, the risks the elders are facing still exist, *e.g.* fall, loss of autonomy, etc. Technology can be an important role for aging in place to assist elders in their own home. There are four categories of technology for aging in place Communication and Engagement, Health and Wellness, Learning and Contribution, and Safety and Security. The high development of activity recognition technology is able to build a daily activity monitoring system in home environment. Activity of daily living (ADL) is an important factor to estimate the independent ability of elders. Barthel index is used to measure performance in activities of daily living and to assess whether the elders have independent ability [3]. Monitoring the ADL of elders to measure their ability can improve the safe living conditions at home.~~

~~1.3 Wireless sensor network (WSN) is a well-developed technology in recent years and a lot of interesting products has been proposed, *e.g.* mobile pad, smart~~

~~**watch or raspberry Pi with ambient sensors. With the technological development of WSN, the Internet of Things (IoT) has more space to play well. The Internet of Things includes the ubiquitous computing where computing is made to appear everywhere and anywhere [4], Ambient Intelligence (AmI) where devices can work in concert to support people in carrying out their daily life activities and tasks more easily, and wearable computing where the associated devices can provide specific, limited features like pedometer, and provide advanced smart function [5]. The technology of IoT is an important factor to implement a smart environment to monitor residents' daily activity.**~~

~~1.4~~1.2 Challenges

In order to implement a friendly and autonomous system to monitor daily activities in smart environment, three primary challenges have to be addressed. The first challenge is to be able to recognizing more activities more precisely by integrating ambient and wearable sensors data. The second challenge is to pay lower efforts on labeling activities when users need to build the activity learning models. The final challenge is to propose an autonomous learning framework, *i.e.* this framework can automatically identify the activities that have not been seen, and then add them to the activity recognition model.~~In order to implement a friendly and automatic daily activity monitoring system in smart environment, three primary challenges have to be addressed. The first challenge is to achieve more activities to recognize with higher resolution of integrating ambient and wearable sensors data. The second challenge is to achieve lower cost on labeling activity when users build activity learning models. The final challenge is to propose an automatic learning framework, i.e. this framework can automatic identification of activity that has not been seen, and add it to the activity recognition model.~~

~~1.4.1~~1.2.1 Integrating Ambient Sensor Network and Body Sensor Network

In order to identify user's activities more precisely for our system, adopting both ambient sensors and body sensors at the same time is necessary. The ambient sensors generally records environment information to serve as objective viewpoints for monitoring user's daily activities; the body sensors, such as wearable sensors, pulse sensors, usually record human's vital signs , which serve as the subjective viewpoins for monitoring user's daily activities and his/her health status. The technology for analyzing ambient sensorinformation, named Ambient Intelligent (AmI); and the technology of analyzing body sensor information called Wearable Computing. In our work, we use both types of sensor to monitor user's daily activities. Although, ubiquitous computing has been proposed for a decade, relatively fewer researches try to combine both AmI and wearable computing.

It's hard to analyze AmI and wearable computing with the same methodology for two reasons. The first reason is that the patterns of ambient sensing data and wearable sensing data are significantly different. The ambient data are usually more static than the data extracted from human body with wearable devices. We can snapshot ambient sensors every-minute to consider the environment information, but we can hardly proceed the same on wearable sensors. The sensing data from wearable sensors usually have the characteristics of rapid change, so that the analysis on wearable computing needs a design of an efficient and statistic model. The second reason is that wearable computing usually uses only one or two sensors to retrieve information from human body, but the sensors in AmI are usually triggered by events, *e.g.* ”TV is on” is usually triggered by the event of watching TV.~~In order to identify more precise activity of user for our system, adopting both ambient sensors and body sensors in same time is necessary. The ambient sensors records environment information as an objective point of view to monitor users daily activity; the body sensors — wearable sensors, pulse sensors, etc. — records vital sign of body and they shows the subjective point of view to monitor users daily activity and health status. The technology of analysis ambient sensors called Ambient Intelligent (AmI); and the technology of analysis body sensors called Wearable Computing. In our work, we use both of two type sensors to monitor users' daily activity. Although, ubiquitous computing has been proposed for a decade, relatively fewer researches try to combine both AmI and wearable computing.~~

~~It's hard to analyze AmI and wearable computing in the same methodology with two reasons. The first reason is that the patterns of ambient sensing data and wearable sensing data are significantly different. The ambient data are usually more static than the data extracted from human body with wearable devices. We can snapshot ambient sensors every-minute to consider the environment information, but we can hardly do the same procedure on wearable sensors. The sensing data from wearable sensors usually have the characteristics of rapid change, so that the analysis on wearable computing~~

格式化: 縮排: 第一行: 0 公分

~~needs a design of an efficient and statistic model.~~

~~The second reason is that wearable computing usually uses only one or two sensors to retrieve information from human body, but the sensors in AmI are usually triggered by fixed events, *e.g.* TV on is usually triggered by the event of watching TV.~~

~~1.4.2~~1.2.2 High Cost on Labeling Activity

Supervised model is usually adopted in the methodologies for activity recognition in smart environment. Supervised learning is a procedure using labeled data to construct a mapping function, which can be applied to map new instances later on. Generally, the training data need to be labeled with their ground truth before one can build the supervised learning model. When the training data are abundant, the inferred results of the supervised learning model usually performs well. However, labeling large quantity of data is a heavy load and training data with wrong labeling makes a poor classifier. Besides, it is also hard to remember daily activities per day in real life which makes the task of labeling difficult. These unfavorable features renders the activity recognition in smart environment to be just in an experimental setting, hand to be realized pratically though it is a popular issue for Internet of Things. Some research try to resolve the high cost on labeling data. They adopt unsupervised learning method to build the activity recognition model, *e.g.* k-means algorithm, Gaussian mixture model, etc. Specifically, unsupervised learning is used to find hidden structure in unlabeled data. The drawback of unsupervised learning is that it needs knowledge of the specific number of activity to build unsupervised learning model. If the given parameter is incorrect, a poor unsupervised learning model. It is hard to choose a parameter that will yield a better unsupervised learning model in real life.~~Supervised model is usually used in the methodologies of activity recognition in smart environment. Supervised learning is a procedure using~~

~~labeled data to construct a mapping function, which can be used for mapping new instance. Its training data needs to be labeled their ground truth before building the supervised learning model. When the training data is abundant, the inferred results of supervised learning model usually performs well. However, labeling amount of data is a burden task and wrong labeling training data makes a poor classifier. It is also hard to remember daily activities per day in real life. These negative features make the activity recognition in smart environment becoming an experimental application though it is a popular issue for Internet of Things. Some research try to resolve the high cost on labeling data. They adopt unsupervised learning method to build the activity recognition model, *e.g.* k-means algorithm, Gaussian mixture model, etc. Unsupervised learning is used to find hidden structure in unlabeled data. The drawback of unsupervised learning is that it is necessary a specific number of activity to build unsupervised learning model. If given a wrong parameter to build unsupervised learning model, it would become a poor inference model. It is hard to choose a better performance parameter of unsupervised learning in real life.~~

~~1.4.3~~1.2.3 Adaptive Learning of Activity Recognition Model

Autonomous activity monitoring is the ideal goal for elderly home care. However, the function of adaptive learning is an important section for implementing a realistic autonomous activity monitoring system. With aging, elderly people may have new lifestyles as time goes by. For the established activity monitoring system, those new activities are unseen activities. If the system is without adaptive function, it may not be appropriate for elderly people to use. Adaptive learning is an educational method which uses computers as interactive teaching devices, and orchestrates the allocation of human and mediated resources according to the unique needs of each learner. In smart environment case, the adaptive learning algorithm of activity recognition model can identify unseen data and consider it as a new type of activities. The adaptive learning model includes two principal functions: discovering new activities and adding data of new activities into training data set to rebuild the learning model. These two functions are both important for elderly home care. If elderly people perform an abnormal behavior, the function of discovery can help their caregivers perceive the emergency alert from the monitoring system. The second function can record new activity as seen activity, so the monitoring system will not query the activity from the user again. If this activity takes place in the future, the monitoring system can recognize this activity instantly.

~~Automatic activity monitoring is the ideal goal for elderly home care. The function of adaptive learning is an important role for automatic activity monitoring system. With aging, elderly people may have new lifestyles. For built activity monitoring system, those new activities are unseen activities. If the system without adaptive function, it is not suitable for elderly people to use. Adaptive learning is an educational method which uses computers as interactive teaching devices, and to orchestrate the allocation of human and mediated resources according to the unique needs of each learner. In smart environment case, the adaptive learning algorithm of activity recognition model can identify unseen data and consider it as a new type of activities. The adaptive learning model includes two principle functions: discovering new activities and adding data of new activities into training data set to rebuild the learning model. These two functions are both important for elderly home care. If elderly people occur an anomaly behavior, the function of discovery can help his/her offspring get the emergency alert from the monitoring system. The second function can record new activity as seen activity, so the monitoring system will not ask user to identify this activity again. If this activity occurs in the fu-~~

~~ture, the monitoring system can recognize this activity automatically. However, the application of elderly home care system activity~~

~~1.5~~1.3 Related Work

Due to the telehealth grows rapidly, the more and more interesting healthcare applications are proposed. The original telehealth focuses on remotely monitoring patients' vital sign, such as blood pressure, heart rate variation, etc.   Physicians analyze their vital sign and give a health report. So patients do not need to go to the hosiptial by them selves [6]. With the wireless sensor network (WSN) developing, the telehealth are trying to build a complete smart care home to achieve the concept of aging in place [7]. The most popular application is automactiv monitoring user's activities [8]. The technologies of activity recognition are designed by machine learning methods, including supervised learning and unsupervised learning. The most of activity recognition models are built by supervised learning algorithms, so their training data are required to label activities. Labeling data is a difficult task for researcher, not to mention end users [9, 10]. The new trend is to design a more friendly human computer interaction interface, so every one is able to use the productions of smart care home. The unsupervised learning can preliminary category raw data into a number of clusters, and each cluster may represent one activity[11, 12]. This method reduces the quantity of labeling data. And some works try to design activity recognition model by semi-supervised learning algorithm. Some data can be automatically categorized by the learning algorithm, but some still need to be labeling ground truth[13]. Try to deploy the automatic activity monitoring telehealth system is an interesting and valuable research.

The activity recognition models are construsted by two types of sensors: ambient

Table 1-1 Comparison among different approaches in the literature and ours

	Sensor Network		Reduce Burden on Labeling	Discover Unknown Activity
	Vital Sign	Ambient		
Our Approach	✓	✓	✓	✓
Sun et al. [11]	✓	-	✓	-
Yuan et al. [17]	✓	-	-	✓
Cheng et al. [13]	✓	-	✓	-
Sanchez et al. [9]	-	✓	-	-
Cook et al. [12]	-	✓	-	✓
Zhang et al. [10]	-	✓	-	-

sensors and on-body sensors. The individual activity recognition models have good performances for specific activities [14, 15]. However, they are not considering in the activity recognition model in same time. The data characteristic are different between ambient sensors and on-body sensors. The data of ambient sensors are usually discrete, and one environment needs to deploy serveral ambient sensors [16]. In the construct, the data of on-body sensor are continuous and also called vital sign data [5]. They are hard to combine together and build activity recognition model directly.

The adaptive learning is an important role for activity recognition model in real life. People have different lifestyles in different stages, so the discovering new activity function and the learning new activity mechanism all are important. Some healthcare systems focus on designing well mechanism of adaptive learning[12, 17].

A complete solution of the home care system need to consider different aspect issues in the same time. Table 1-1 shows the differences between the previous works and our approach from several perspectives, including the usaged sensors, the mechanism of reducing burden on labeling data and the function of adaptive learning.

~~1.6~~1.4 Objective

In this thesis, we try to build a smart home environment based on the technique of IoT. The goal of this smart home is to provide a robust tele-healthcare system that is able to monitors resident's living activities real-time. For the purpose of addressing the aforementioned challenges, the objective of this thesis is to develop an Activity of Daily Living(ADL)-aware Healthcare system of the Elderlyat home, which apparently is a powerful smart environment application. Such system aims to monitor residents' daily activities anytime in their house, whereby the system aims to observe the lifestyle behaviors of residents and report the potential anomaly when it happens. The contributions of this thesis are listed as follow:

~~1.6.1 In this thesis, we try to build a smart home environment~~

~~1.6.2~~1.4.1 Better Recognized Performance by Sensors Fusion

The activity monitoring result is important for caring of the elderly because ADL is an important factor to estimate their health states. Different sensors can observe residents from different points of view in a home environment. In other words, it is necessary to fuse heterogeneous sensors to observe activities more precisely. For ambient sensors, they usually provide more objective observation meaning that they monitor residents' activities in indirect ways. The observations from ambient sensors are usually triggered by specific activities, so the ambient data analysis is event trigger oriented. For wearable sensors or other vital sign measurement sensors, they usually take more subjective obeservations, meaning that they monitor residents' activities in direct way.~~The activity monitoring result is important for elderly residents because ADL is an~~

~~important factor to estimate their health state. To analyze activity of daily living (ADL) of elderly peoples can help caregivers estimate their health conditions. Different sensors can observe different points of view on residents in home environment. In other words, it is necessary to fuse heterogeneous sensors to observe more precise activities. For ambient sensors, they usually observe objective point of view that means they monitor residents' activity in their home in indirect way. The observations from ambient sensors are usually triggered by specific activity, so the ambient data analysis belongs to event trigger orientation. For wearable sensors or other vital sign monitoring sensors, they usually observe subjective point of view that means they monitor residents' activity in direct way.~~ The observations from wearable sensors are unlike those from~~to~~ ambient sensors. To observer an activity usually needs a set of continuous sensing data from the wearable sensor ~~data. A~~. Nowadays, a few sensors can be ~~built~~ integrated into a~~on~~ wearable device~~s~~, and ~~each~~ this different kinds of sensor data have high degree of correlation. This characteristic makes analysis of wearable sensor data ~~analysis hardly using same methodology of~~very different from that ambient sensor data ~~analysis~~. ~~Both~~ Generally, ambient intelligence~~t~~ (AmI) and mobile computing are individually highly developed~~ing~~ ~~independently~~technologies, but ~~they~~ there are rare effects to fuse them together ~~with~~ given their totally different characteristics. ~~We~~ In this thesis, we have designed a hierarchical activity recognition (AR) model with two-layer structure, ~~and~~ which it has resolved the problem of combining AmI and mobile computing coherently. The first layer of AR model is to identify residents' behaviors from two activity recognition models: environment-based AR model and body-based AR model. The environment-based AR model determines the residents' living activities from ambient sensors data through an unsupervised learning algorithm with fuzzy learning. The body-based AR model collects wearable sensor data from residents' smart wearable devices through some wearable computing technique. Here, we~~I~~ propose a topic model to consider ~~meaningful~~ the underlying ~~actions~~ activities based on an unsupervised clustering method, ~~named~~namely, Dirichlet Process Mixture Model (DPMM) [18]. This~~e~~ topic model is used to retrieve meaningful information from a large ~~amount~~ quentity of temporal/sequential raw data. Such proposition~~The concept~~ is appropriate for recognizing the living activities~~y~~ ~~recognition~~ with the ambient sensing data ~~- and the meaningful actions~~ simply because the feature distances between two different activities should be ~~high~~large. Before building the second layer of the hierarchical AR model, the system will fuse results of the environment-based and body-based models as new format of training features. The second layer AR model is used to determine residents' living activity in an overall view. These two-layer AR model can determine precisely activities~~y~~ for residents, *e.g.* i~~I~~f one is watching TV and sweeping the floor at same time, then ~~analysis~~ analyzing the of ambient data can only determine the ~~active status of~~activity of watching TV ~~and analysis of~~whereas analyzing the wearable sensor adta~~s~~ can only determine ~~the resident is doing~~ the activity of sweeping. However, ~~Fusion~~ fusing ambient and wearable sensor data ~~analyses~~ can determine both ~~two~~ activities.

~~1.6.3~~1.4.2 Facilitated Activity-aware System for Elderly Healthcare

For the elder people, learning to use technology products is ~~greater obstacles~~more difficult than that by younger people. Simple and automated system for elderly people is a more plausible approach~~able~~. In order not to ~~reduce the~~impose too much burden on the elders~~ly~~ ~~people learning~~when they learn to use the activity-aware healthcare system, the system ~~has easily~~should have the function for easy labeling ~~function~~ and the function of discovering activities ~~discovery~~ new lifestyle function. These two novel functions ~~make elderly people independently use~~will enable the elders to employ the healthcare system by themselves independently.

To ~~achieve~~ realize the friendly labeling function, the system avoids complex and time consuming procedure on labeling all training data. In other words~~training mode~~, the AR model learns activities ~~by~~ with data-driven approach, *i.e.*, without labeling training data. Specially, ~~Data~~ data driven approach means that progress in an activity is compelled by data, rather than by intuition or personal experience. This is the main reason ~~that~~ why the hierarchical AR model adopts unsupervised learning methodology. ~~And~~ Also, it ~~also~~ has a special characteristic which is that such model does not require a prior specification of the ~~cluster~~ number of clusters to be learned from the raining data,~~.~~which is known as ~~The~~ non-parametric clustering algorithm ~~means users do not need to set the specific number as the number of activities to be recognized~~. This~~It~~ is an important factor for ~~elderly residents to set this system~~the underlying learning system to be realistic enough in a real living environment. ~~It is apparently very different from both~~Unlike the supervised learning method ~~and~~ on parametric clustering,~~;~~ such as k-means algorithm,~~.~~ ~~The~~ the non-parametric learning method will find the most appropriate number of clusters based on their training data,~~.~~ ~~It's~~ which is more suitable ~~to model the~~for problem modeling in the realistic world. ~~The~~ That is, which the system tries~~y~~ to monitor elders' living activities, ~~but~~ it's hard to ask them to label their activit~~ies for every~~y instances ~~in~~ during the training ~~course~~. Since the proposed AR model can find the significant living activity ~~of~~ associated with each cluster, a direct advantage is that the elders only need to specify ~~the~~ each resulting cluster~~s~~ in a very straightforward manner. For an elder who lives alone, deploying the hierarchical AR model is apparently easier than the parametric model.

格式化: 字型: 斜體

~~To achieve discovery~~In order to discover new lifestyle activities of home residents, the system needs to adopt~~s~~ adaptive learning technique to discover the activities~~y~~ which are~~is~~ out of the training data and then to rebuild the activity recognition model. In online mode, the activity-aware system has a mechanism called case-based reasoning (CBR)~~.~~ ~~Case-based reasoning~~, which is an artificial intelligence~~t~~ ~~technology~~technique, ~~it is the process of solving~~doomed to solve the new problems based on the solutions of similar past problems. Generally, CBR of the activity-aware system is used to discover~~y~~ unknown cases and immediately provide service based on the most similar case. After the system provides service, it will ask the resident ~~what~~ for the meaning of the activity ~~is~~. ~~When~~ Right after this activity ~~has been~~ is labeled, the AR model ~~is going to~~will be rebuilt. Later, t~~T~~he activity-aware system can identify this activity from the ~~re-new~~renewed AR model. ~~The~~ Such automatically discovering mechanism is suitable for the elderly ~~residents~~ because their aging~~ed~~ processes are usually accompanied~~y~~ with new habits and behaviors, yielding new activities.

~~1.7~~1.5 System Overview

In order to address the activity-aware healthcare system, We have~~I~~ proposed two modes of the system: training mode and online mode. The training mode is used to identify activities from data automatically ~~as~~ into a number of clusters. Each cluster represents a meaningful activity. Residents need to apply training mode before the system proceeds to monitor~~s~~ their daily activities~~y~~. Notice that the model for the t~~T~~raining mode ~~model~~ will provide an interface to label the activity associated with each cluster.~~'s mapping activity.~~ After residents label all clusters, the online model will be ~~generated by those~~established through these labeled data.

The online mode is used to real-time determine the activity of the elderly ~~residents~~ and ~~monitor~~ to detect the abnormal ~~anomaly~~ activity. If an unknown activity has been observed, the system will determine whether it is an anomaly activity or not. If the unknown is similar to those ~~anomaly~~ abnormal activities~~y~~, the system will send an alert message to the caregivers of the elderly ~~resident's caregiver~~. And the system will require

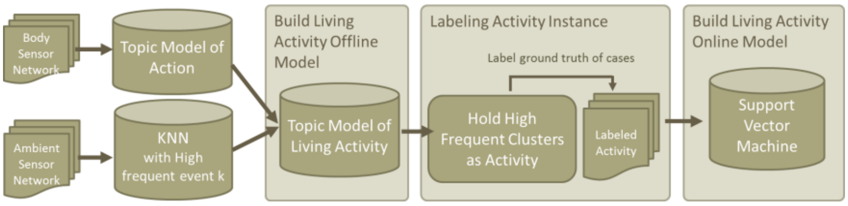


Fig. 1-1 System Overview of Training Mode

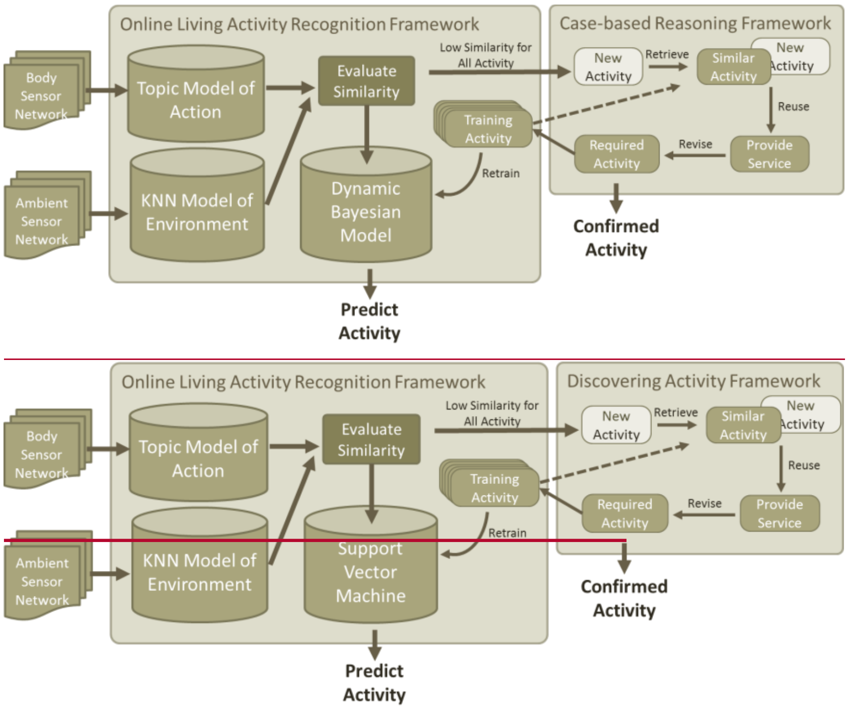


Fig. 1-2 System Overview of Online Mode

resident to label this activity and retrain the AR model. Fig. 1-1 and Fig. 1-2 show the system overview of training mode and online mode respectively.

Before building the hierarchical AR model, extracting feature is an important part. In machine learning, feature extraction starts from an initial set of measured data and ~~builds~~ derived values intended to be informative, facilitating the subsequent learning. The ambient sensor data ~~is~~ are extracted as Boolean ~~variable~~values, namely~~:~~ on and off, because they are triggered by human activities~~y~~, *e.g.*, current sensor on Television measures low power when Television is ~~close~~turned off,~~;~~ but measures~~ing~~ high power when Television is ~~open~~turned on. A simple classifier is proposed to determine the state of Television by its current sensor data. The wearable sensor data are~~is~~ extracted as some statistic variables or physical features, such as a second mean and variance data.

~~1.8~~1.6 Thesis Organization

This thesis consists of five chapters, and they are organized as follows. Chapter 2, introduces some preliminary knowledge of the thesis, includes pervasive environment, non-parametric statistic distribution and the methodologies of machine learning. Chapter 3 describes the details of our hierarchical activity recognition model. We fuse the heterogeneous sensors of ambient and vital sign sensors. The fusion mechanism of ambient and wearable also describes in detail. Chapter 4 proposed the healthcare system, and it is used to real-time ~~aware~~ recognizing residents' activities~~y~~ based on our proposed hierarchical activity recognition model. The adaptive learning mechanism of this healthcare system also describes in detail. Chapter 5 shows the details of the experimental environment and the evaluation metrics. The experimental results are discussed and analyzed in this chapter as well. Finally, in Chapter 6, conclusion and discuss the future work are provided.

Chapter 2

Preliminaries

2.1 Pervasive Environment

Over the past two decades, ubiquitous and pervasive computing have been evolving[19, 20]. Users are surrounded by many different devices capable of capturing and processing information. The importance of this area has been recognized by researchers and funding bodies alike. The most common system is the Smart Home. ~~It has been~~ There have learn focused efforts on developing techniques to implement a home that has ability to identify residents' demand and to automatically provide services. For example, it has ability to achieve intelligent light controls, window shutters, safety system or kitchen appliances, etc. In particular, there has been considerable interest in developing a smart home with healthcare ~~home~~ that can support ~~for~~ elderly or disabled residents. The goal of smart healthcare home is making it safe for them to live at their home. ~~This~~ Such approach is basically concerned with a fixed space that is require to provide intelligent features.

Beyond the fixed space of the Smart Home, the mobile user presents different and more challenging problems. The most different situation between fixed space and mobility is user’s location. The pervasive system always tries~~y~~ to provide access to devices and services in the user’s environment. However, it is hard to control the mobile user’s environment information. The research ~~into~~ on fixed spaces associated with buildings is generally quite independent of that being conducted on pervasive systems for mobile users. Those separated systems make the support for pervasiveness be~~is~~ limited. It is a challenge ~~that~~ for controlling~~s~~ all devices, including~~ed~~ ambient devices and mobile devices in the same system.

For our system, we have built a smart home environment at Room 313 of Barry Lam Hall (BL 313) at~~in~~ National Taiwan University. Although the established smart home ~~used~~ was to pursue the goal of~~to achieve~~ energy saving ~~in the~~ before, here we ~~try to~~ use the~~o~~se ~~constructed~~ already depolyed devices to build an activity-aware healthcare home. The energy saving system of this constructed smart home is called M-CHESS~~.~~ ~~M-CHESS~~, ~~is the~~ abbreviation of M2M-based Context-aware Home Energy Saving System[21]. The basic~~ed~~ function of M-CHESS is aware of the surrounding contexts, *e.g.*, the on-going users’ activities and the status of electronic appliances. However, the context structure of M-CHESS is not suitable for the robust activity-aware healthcare system, because it was developed for monitoring users’ activity of using electronic appliances. In this work, we want to monitor residents’ daily living activities, so M-CHESS is hard to identify some activities which without using electronic appliances, *e.g.*, doing exercise or cleaning home. We try to use mobile devices to monitor ~~these kind of~~the activities including these kinds. We ~~adopt~~ employ the smart watch, called “Zen–Watch”, to monitor users’ vital signs, *e.g.*, acceleration~~ometer~~ and notation through accelerometer and gyroscope ~~on the smart~~embedded in the watch. In the next chapter, we will describe the ~~fusion~~ method ~~between~~ of fusing data from ambient sensors and wearable sensors~~; that ambient sensors are the constructed~~ when the former

格式化: 字型: 斜體

格式化: 字型: 斜體

格式化: 字型: 斜體

devices ~~of~~ deployed for M-CHESS ~~and~~ whereas the ~~wearable~~ latter sensors ~~are~~ refers to thses built on Zen-Watch.

The layout of the home is as shown in Fig. 2-1. That environment specifically consists of 4 rooms, namely, living room, studying room, kitchen, and bedroom. ~~The~~ It is worthwhile to mantian that the current sensor is a simple “smart meter” built on the embedded device “Taroko” that is used monitor the on/ off status of each electronic appliance. Moreover, each room has been equipped with temperature, lumen and humidity sensors also built on Taroko. ~~The~~ A switch sensor in mounted on the entrance door ~~has also been mounted a switch sensor used~~ to monitor the activities of “go out” and “come home”. Those sensors in the previous research work ~~are developed~~were employed to research on how to save energy, unlike ~~this~~ the present work where we try to use them to monitor elders’ living activit~~y~~ies.

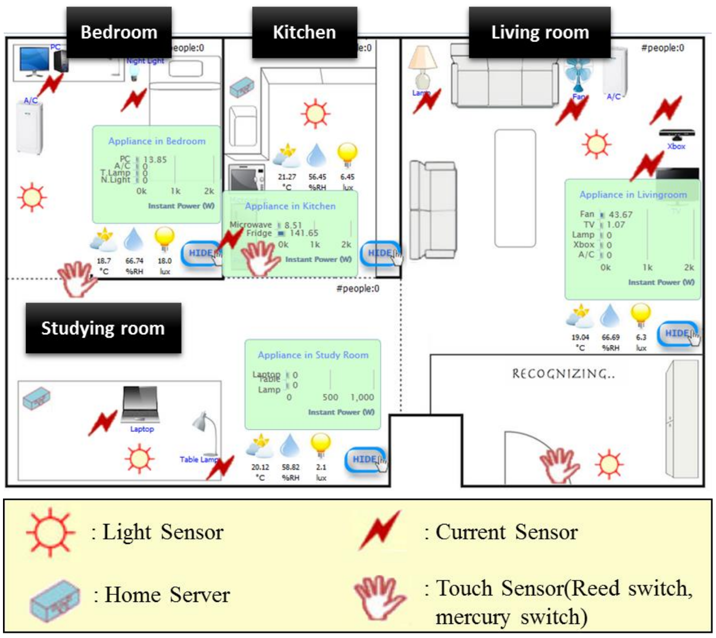
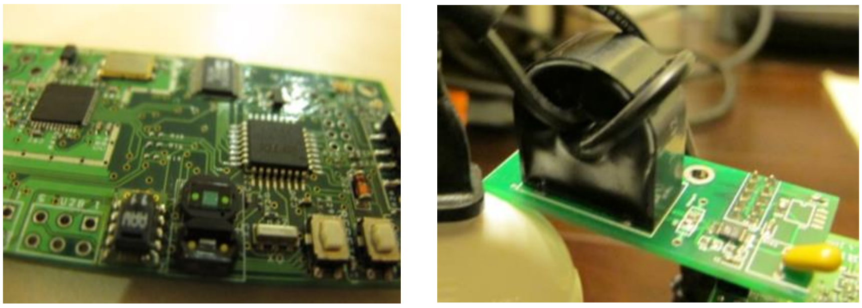


Fig. 2-1 The ~~2D ichnography~~ loyout of the smart home in BL 313

The ambient sensors ~~in our work have been~~discreased in this research are divided into two different categories, of which one is mainly to monitor the living environment, *e.g*. lumen sensors used to monitor light level for each room and switch sensors ~~equipped~~ on the entrance door used to ~~determine~~ tell whether the resident goes out, comes home, or else, and the other is mainly to monitor the active states of electric appliances so that the residents' activities will be revealed when the associated electric appliances are being used. An example of the latter is that the AR system recognizes that the resident is watching TV partly because the system~~it~~ finds the state of TV is "on". ~~First~~ The first type of sensor is ~~as~~ shown in Fig. 2-2(a),~~.~~ ~~Second~~ and thesecond type of shown in Fig. 2-2(b). We will use ~~their~~ the obtained sensor data to build a machine model ~~that uses to~~and then category residents' activitiesy~~y~~; the detailed processing described in C~~c~~hapter3.

In order to ~~realize~~ identify a specific physical activity of an elder, using the wearable sensors to collect ~~their~~ the elder's vital signs is necessary. ~~Note that~~Generally speaking, ambient sensors ~~themselves~~ are usually ~~embedded into~~installed in the environment, whereas wearable sensors are directly attached to the human body, which typically can run continuously and can be operated hands-free [22]. While those ~~embedded~~ environment devices and wearable devices communicate and cooperate with each other,



(a) (b)

Fig. 2-2 The sensor deployed in the home environment the original monitoring system is now enhanced to serve as a seamless monitoring system.

~~In this study~~As aforementioned, the wearable sensor we here adopted is a smart watch, named "Zen Watch",~~. It~~which is equipped with a tri-axial accelerometer, gyroscope, and low energy Blue-tooth with 30 Hz sampling rate. We have preprocessed those acceleration data ~~of~~against each axis to obtain mean and variance of this physical value every second,~~.~~ ~~Extracting~~ Besides, we extract the orientation, consisting of pitch, roll and yaw, from ~~its~~the embbedded accelerometer and gyroscope~~.~~, ~~Then~~ and then find its mean ~~orientation and variate of orientation~~and variance every second. ~~as machine learning features. Each pattern extracted from the accelerometer and gyroscope inside the wrist watch~~As time goes on, the preprocessed acceleration and orientation values from a pattern which normally indicates a waving motion of arm, and the numbers of consecutive waving motions can be associated with a specific activity. ~~We~~ Thus, we will use machine learning method to categorize the waving motions, and the results will serve as new features for activity recognition. The detailed processing procedure will be described in ~~chapter3~~Chapter3.

2.2 Non-parametric Statistical Distribution

Non-parametric statistics are statistics that are not based on parameterized families of probability distributions,~~.~~ ~~Says, the~~cohose typical parameters are ~~usually the~~ mean ~~or~~ and variance of the distribution. For non-parametric statistics, they do not assume about the probability distributions of the variables being assessed~~. They include both~~, and the examples are descriptive and inferential statistics. The difference between non-parametric and parametric model is that the former will ~~grows~~ increase the number of parameters ~~with their amount~~in proportion to the quantity of training data; ~~but~~ co-hereas the latter needs to give a fixed number of ~~their~~ parameters *a priori*. The major goal of our system is to reduce~~ing~~ the burden of labeling step~~.~~, ~~The~~ and the non-parametric statistic~~s~~ ~~is~~serves an important role ~~for~~ in our recognition model simply because ~~it makes~~ the reality recognition model does not ~~ask~~ require the user ~~assign the specific~~prespecify the number of ~~their~~interested daily activit~~y~~ies. ~~A~~in short, a recognition model ~~used~~ based on non-parametric statistic ~~concept~~notion can identify activities from ~~its~~ it's the existing quantuty~~amount~~ of training data [23].

The term, ~~of~~ “non-parametric statistics,” has two different meanings. ~~The first meaning of~~First, the “non-parametric” technique~~s~~ does ~~not rely on~~require the data belonging to any particular distribution, *e.g*. “distribution free” methods, ~~which~~ *i.e.* ~~do not rely on assumptions that~~one does not have to assume that the data are drawn from a given probability distribution. ~~The second meaning of~~Next, the “non-parametric” technique~~s~~ does not assume that the structure of model is fixed,~~.~~ *i.e.*~~Says~~, the model is growing in size to accommodate the complexity of training data.~~ And individual variables are assumed to belong to parametric distributions, and assumptions about the types of connections among variables are also made.~~ These type of technique~~s~~ include non-parametric regression and non-parametric hierarchical Bayesian models. The former refers to modeling where the structure of the relationship between variables is treated non-parametrically. The latter allows the number of latent variables to grow as necessary to fit the data. Individual variables still follow parametric distributions and the process controlling the rate of growth of latent variables follows a parametric distribution. The Dirichlet process is one of the distributions in this category.

In our system, we try to detect users’ activit~~y~~ies ~~in~~ real-time. For activity detection, such parameter specification may face several challenges. First, it is hard to find

格式化: 字型: 斜體

格式化: 字型: 斜體

格式化: 字型: 斜體

格式化: 字型: 斜體

the appropriated parameter values for personalized models that may be different for different users. Second, for a single user, user’s behavior patterns may change over time. Given a fixed parameter is not suitable to build the inference model. In other words, the most appropriate~~ing~~ parameter values must be adjusted accordingly. Hence, we try to achieve that the model has ability to automatically select parameter values based on individual users’ training data. Our inference activity model uses nonparametric statistic methodologies,~~.~~ ~~I~~including the non-parametric Bayesian methods, “Dirichlet process mixture model (DPMM),” and non-parametric lazy learning, “k-nearest neighbor (KNN)” algorithm. The former can avoid declaring the number of activities and routines in a person’s daily life beforehand in parametric settings~~.~~, ~~The~~ whereas the latter does not make any assumptions on the underlying data distribution.

2.2.1 Dirichlet ~~Distribution~~Process

Before we design the learning model of Dirichlet process mixture model, we ~~need~~ first review~~to realize~~ the ~~based~~basic distribution “Dirichlet”,~~.~~ ~~The Dirichlet distribution~~which in fact is a model ~~of~~ showing how proportions vary. In other words, the Dirichlet distribution is a family of continuous multivariate probability distributions parameterized by a vector α of positive reals~~.~~, ~~A~~and th~~e~~is Dirichlet distribution is denoted as $\mathrm{Dir}(\alpha)$. It can be seen as the multivariate generalization of the beta distribution [24]. ~~It’s a family of continuous multivariate probability distributions parameterized by a vector α of positive reals.~~ This means when the parameter of a data point is distributed as Dirichlet, the posterior distribution of the parameter will be a Dirichlet. ~~Its probability density function returns the probabilities of K rival events are π_i given that each~~

格式化: 字型: 斜體

~~event has been observed $\alpha_i - 1$ times. Dirichlet distribution is the conjugate prior distribution of multinomial distribution. There are existing the number of successes in a sequence of independent data that each data in one of k possible outcomes with probabilities $p_1, \ldots, p_k$; Multinomial models the distribution of the histogram vector which indicates how many time each outcome was observed over N trials of experiments.~~

$$P(x_1,\ldots,x_k|n, p_1,\ldots,p_k) = \frac{N!}{\prod_{i=1}^{k} x_i!} p_i^{x_i}$$

$$\sum_{i=1}^{k} x_i = N \text{ and } x_i \geq 0 \qquad (2\text{-}1)$$

~~Dirichlet distribution is parameterized by a vector of $A = \{\alpha_1,\ldots,\alpha_k\}$, where $\alpha_i \geq 0$ and $K \geq 2$. Define a random vector $\Pi = \{\pi_1,\ldots,\pi_K\}$, where $\sum_{i=1}^{K} \pi_i = 1$, $\pi_i > 0$ and $\pi_i \in [0,1]$. The Dirichlet distribution with vector A has a probability density function $P(A)$ shown as follows:~~

$$P(\alpha_1,\ldots,\alpha_K) = \frac{\Gamma(\sum_i^K \alpha_i)}{\prod_i^K (\Gamma(\alpha_i))} \sum_{i=1}^{K} \pi_i^{(\alpha_i-1)} \qquad (2\text{-}2)$$

~~So π is mapping to the probabilities p of multinomial. We say that Dirichlet is the conjugate prior of multinomial. And the support of K dimensional Dirichlet distribution is the K-1 dimensional probability simplex. Let $\Pi=(\pi_1,\ldots,\pi_K)$~Dir(α), meaning that the first K-1 components have the above density and the weight of π_K should be defined as follows:~~

$$P(\alpha_1,\ldots,\alpha_K) = \frac{\Gamma(\sum_i^K \alpha_i)}{\prod_i^K (\Gamma(\alpha_i))} \sum_{i=1}^{K} \pi_i^{(\alpha_i-1)} \qquad (2\text{-}3)$$

~~*e.g.*, a set of parameter $\Theta=(\theta_1,\ldots,\theta_K)$ describes a Dirichlet distribution with $P(\theta=\theta_i) = \pi_i$. When the set of concentration parameter is $(\alpha_1,\ldots,\alpha_K)=(1,\ldots,1)$, the distribution of $(\pi_1,\ldots,\pi_K)$ is uniform.~~

格式化: 本文, 間距 套用後: 0 點

格式化: 本文, 靠左, 不與下段同頁

格式化: 本文, 靠左

格式化: 本文, 靠左, 不與下段同頁

格式化: 本文, 靠左

格式化: 本文, 靠左

~~2.2.2 Dirichlet Process~~

The Dirichlet process assumes dimension of the Dirichlet distribution is infinite, so the non-parametric learning model can be built based on this characteristic. The Dirichlet process is an infinite-dimensional generalization of the Dirichlet distribution and it is denoted as DP(α,H), where α is a positive real number called the concentration parameter and H is a baseic distribution. The baseic distribution is the expected value of the process, *i.e.*, the Dirichlet process tries to draw distributions around the baseic distribution. And, the concentration parameter used to specify the strong level of the discretization, such as in the limit of $\alpha \to 0$ means the realizations are all concentrated on a single value. On the contrary, in the case where limit of $\alpha \to \infty$ means the realizations become continuous.

If a distribution $G(\theta)$ is $G(\theta) \sim DP(\alpha, H)$, it is a Dirichlet process.

$$G(\theta) = \sum_{i=1}^{\infty} {\pi_i}^{\delta(\theta=\theta_i)} \tag{2-1}$$

~~Where~~ where $\theta_K \sim H$ and a Dirac delta function $\delta(\theta = \theta_i)$. To construct infinity sequence of mixture weight π_i using the stick-breaking scheme, that represented as $G(\theta) = \sum_{i=1}^{\infty} \beta_i \delta_{\pi_i}(\pi)$, where $\{\pi_i\}_{i=1}^{\infty}$ are sampling from the base distribution H. The Dirac delta function δ_{π_i} centers on π_i. The β_i are defined by a recursive scheme from the beta distribution. The parameter α defines how concentrated the distribution is equation (2-2) and $i = 2,3, \dots, \infty$.

$$\begin{gathered} \beta_i \sim Beta(1, \alpha), \\ \pi_1 = \beta_i, \\ \pi_i = \beta_i \prod_{j=1}^{i-1}(1-\beta_j) \end{gathered} \tag{2-2}$$

In many applications, the infinite dimensional distributions appear only as an in-

格式化: 縮排: 第一行: 0 公分

termediary computational device. They are not required for the initial specification of prior beliefs or the statement of the final inference. The Dirichlet process is used to avoid infinite computational requirements.

2.3 Inference Model of Machine Learning Techniques

Machine learning is a subject of computer science and it evolves from the study of artificial intelligent and pattern recognition. Machine learning explores the construction and study of algorithms that learns from and make predictions on data. The machine learning algorithms build a model that make data-driven predictions or inferences from instance inputs. Machine learning tasks are classified into three main categories, depending on the learning "signal" or "feedback" available to the learning model. Each is supervised learning, unsupervised learning and reinforcement learning. And I will briefly introduce them in the below.

Supervised learning: Given the learning algorithm a "teacher", then the computer is presented with example inputs and their desired outputs. And the goal is to learn a general rule that maps inputs to outputs. Teacher means a set of specific labels of the training data. In other words, supervised learning analyzes the training data and produces an inferred function, which is called a classifier

Unsupervised learning: It does not necessary labels to implement the learning algorithm. Leaving it on its own to find structure in its input. Unsupervised learning is trying to find hidden structure in those unlabeled data.

Reinforcement learning: The environment is usually formulated as a Markov decision process (MDP). That the reinforcement learning does not require knowledge about the MDP and they target large MDPs where exact methods become infeasible. In other words, a program of reinforcement learning technique interacts with a dynamic environment in which it must perform a certain goal, without a teacher explicitly telling it whether it has come close to its goal or not.

In our activity-aware system, we use and modify some machine learning techniques in order to build the hierarchical activity recognition model, including two unsupervised learning algorithm, mixture model and k-nearest neighbor algorithm, and one supervised algorithm, support vector machine.

2.3.1 K-Nearest Neighbors Algorithm

K nearest neighbors (KNN) is a simple unsupervised learning algorithm that stores all available cases and classifies new cases based on a similarity measure, e.g., distance functions. KNN has been used in statistical estimation and pattern recognition already in the beginning of 1970's as a non-parametric technique [25]. A case is classified by a majority vote of its neighbors, with the case being assigned to the class most common amongst its K nearest neighbors measured by a distance function. If K = 1, then the case is simply assigned to the class of its nearest neighbor. Three distance functions are used to measure the similarity of each instance. They are Euclidean distance, Manhattan distance and Minkowski distance.

$$Euclidean\ distance\ D_E = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \tag{2-3}$$

$$Manhattan\ distance\ D_{Manhattan} = \sum_{i=1}^{k}|x_i - y_i| \tag{2-4}$$

$$Minkowski\ distance\ D_{Minkowski} = (\sum_{i=1}^{k}(|x_i - y_i|)^q)^{\frac{1}{q}} \tag{2-5}$$

It should also be noted that all three distance measures are only valid for continuous variables. In the instance of categorical variables the Hamming distance must be used [26]. It also brings up the issue of standardization of the numerical variables between 0 and 1 when there is a mixture of numerical and categorical variables in the dataset.

$$Hamming\ distance\ D_H = \sum_{i=1}^{k} |x_i - y_i|$$
$$\mathrm{x} = \mathrm{y} \rightarrow D_H = 0 \tag{2-6}$$
$$\mathrm{x} \neq \mathrm{y} \rightarrow D_H = 1$$

Choosing the optimal value for K is best done by first inspecting the data. In general, a large K value is more precise as it reduces the overall noise but there is no guarantee. Cross-validation is another way to retrospectively determine a good K value by using an independent dataset to validate the K value.

2.3.2 Mixture Model

A mixture model is a probabilistic model for representing the presence of subpopulations within an overall population, without requiring that an observed data set should identify the sub-population to which an individual observation belongs [27]. It corresponds to the mixture distribution that represents the probability distribution of observations in the overall population. A general mixture model, which is usually present the infinity-dimensional mixture model, is a hierarchical model consisting of the following components. *N* random variables corresponding to observations. It assumed to be distributed according to a mixture of K components. N corresponding random latent variables specifying the identity of the mixture component of each observation, each distributed according to a K-dimensional categorical distribution. A set of K mixture weights, each of which is a probability (a real number between 0 and 1 inclusive), all of which sum to 1. A set of K parameters, each specifying the parameter of the corresponding mixture component. In many cases, each "parameter" is actually a set of parameters. That a basic parametric mixture model can be described as follows:

K = number of mixture components

N = number of observations

θ_i = parameter of distribution of observation associated with component i, for $i = 1,\ldots,K$.

ϕ_i = mixture weight, *i.e.*, prior probability of a particular component i

Φ = K-dimensional vector composed of all the individual $\phi_{1\ldots K}$

z_i = component of observation i

x_i =observation i

$f(x|\theta)$ = probability distribution of an observation

$$z_i \sim \text{Categorical}(\phi)$$

$$x_i \sim f(\theta_{z_i})$$

In the case of Bayesian setting, the mixture weights and parameters are random variables, and prior distributions are placed over the variables. So the weights are usually viewed as a k-dimensional random vector that drawn from a Dirichlet distribution. In other words, all parameters are associated with random variables. The form of Bayesian setting shows as below. And the parameters of K, N, θ_i, ϕ_i, Φ, z_i, x_i and $f(x|\theta)$ are same to the general mixture model's parameters.

α = shared hyper-parameter for component parameters

β = shared hyper-parameter for mixture weights

$H(\theta|\alpha)$ = prior probability distribution of component parameters

$$\theta \sim H(\theta|\alpha)$$

$$\Phi \sim \text{Symmetric-Dirichlet}_{\mathrm{K}}(\beta)$$

$$z_i \sim Categorical(\phi)$$

$$x_i \sim f(\theta_{z_i})$$

Using $f(x|\theta)$ and $H(\theta|\alpha)$ to describe arbitrary distributions over observations and parameters. That H will be the conjugate prior of F. And in the field of mixture model, F is typically using Gaussian distribution that called Gaussian mixture model (GMM). But we will use Dirichlet process as the based distribution as F, and it is called Dirichlet process mixture model (DPMM). The detail implement process will describe in Chapter3.

2.3.3 Dynamic Bayesian Network

Observing a sequence of emissions, but do not know the sequence of states. The Dynamic Bayesian Network (DBN) is that went through to generate the emissions [28]. Analyses of DBN seek to recover the sequence of states from the observed data. It is similar to Hidden Markov Model. In a hidden Markov model, the state is not directly visible, but output, dependent on the state. Each state has a probability distribution over the possible output tokens [29]. In order words, a HMM represents the state of the world using a single discrete random variable $x_t \in \{1, \ldots, K\}$. A DBN represents the state of the world using a set of random variable $X_y = \{x_t^{(1)}, \ldots, x_t^{(D)}\}$. And the dimension D of random variable X_t is the number of activity in the recognition system. And DBN are especially known for their application in sequential pattern recognition. We choose DBN as the online mode because the temporal and sequential information can help rise the recognition performance.

In a graphical model, nodes represent random variables, and arcs represents conditional independencies. It considers directed graphical models is equaling to Bayesian networks or belief networks. And DBN is one of Bayesian network for dynamic processes. The graphic of DBN should be acyclic and an arc from X_i to X_j means "X_i causes X_j".

We have built a simple DBN likes HMM. It assumes that the observation at time t was generated by state S_t from the observer. Given S_{t-1}, the current state S_t is independent of all the states prior to previous state $t-1$. The variable Y_t of observation is independent of the states and observations at all other time indices. The joint distribution is formulated as follows:

$$P(Y, x_1, x_2, \ldots, x_T) = \prod_{t=1}^{T} \prod_{i=1} P\left(Y_{t,i} \middle| x_t\right) P(x_t | x_{t-1}) \tag{2-7}$$

where $y_{t,i}$ refers to each feature of one instance at time *t*. Since all variables are discrete; therefore, each conditional probability is assumed to have categorical distribution. The probability mass function (PMF) of multinomial categorical distribution for variable *x* is as follows:

$$\begin{gathered} \mathrm{f}(y = i | \mathrm{p}) = p_i \\ \sum_{i=1}^{K} p_i = 1 \end{gathered} \tag{2-8}$$

where p_i is the probability of *y* equaling *i*. The number of possible outcomes of *i* is de-

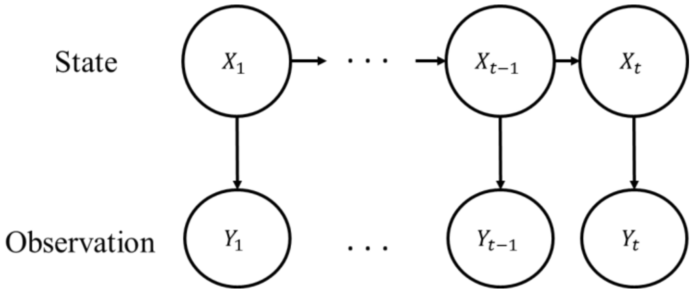


Fig. 2-3 The graph structure of Simple DBN

termined by the number of states in each random variable.

The arrow from node A to B in Bayesian networks represents the conditional probability $P(B|A)$. Through maximum likelihood estimation (MLE) using historical data, each conditional probability in the DBN model can be learned, even when hidden variables exist. It should be noted that all conditional probabilities are assumed to follow categorical distribution because all variables are discrete and no prior assumption is made about which distribution is followed.

The MLE algorithm is used to obtain the joint probabilities capable of maximizing the likelihood that training data will be obtained. For the purpose of estimating living activity, we apply the belief propagation algorithm to estimate the marginal probability of the interested variable X.

We use Dynamic Bayesian Network to classify users' daily activity in online mode. Before the system builds DBN, the system used unsupervised learning methodologies DPMM and KNN to build the hierarchical activity inference model of offline mode and asked user to label their activity of the processed training data that has been categories into a set of clusters by the hierarchical activity inference model. Because the training data are labeled in offline mode, the supervised learning algorithm is able to build its classification model. In order to considering features over previous steps, the system also automatically examine features that link state transitions in the model directly to observations. The detail implementation of our online AR model that built by DBN describes in the chapter3.

Chapter 3

Activity Recognition Model by Fusing Ambient and Vital Sign Fusion

3.1 The Architecture of Activity Recognition Model

I In order to real-time monitor the daily activity of the elderly, we need to design an activity recognition model. We have proposed an activity-aware healthcare system that can sperate in two modes: training mode and online mode. The training mode is used to identify activities from data automatically as a number of clusters. Each cluster is associated with a meaningful activity. The system will ask user to label each cluster as one activity. After all clusters are labeled, the online model will use those labeled data to build an inference model by dynamic Bayesian network.

~~n order to real-time monitor elderly residents' daily activity, we are necessary to design an activity recognition model. I have propose an activity-aware healthcare system that has two modes for utilizing: training mode and online mode. The training mode is used to identify activities from data automatically as a number of clusters. Each cluster presents a meaningful activity. The system will ask user to label each cluster as one~~

~~activity. After all clusters are labeled, the online model will use those labeled data to build an inference model by conditional random field.~~

~~T~~The goal of offline mode is trying to identify all observed environmental and wearable sensor data as living activities in the offline mode. To reduce the burden of labeling those observed sensor data, the proposed system categories sensor data as a set of clusters by~~he goal of offline mode is trying to identify all observed environmental and wearable sensor data as living activities in the offline mode. And reduce the burden of labeling those observed sensor data. The system categories sensor data as a set of clusters by~~ the non-parametric hierarchical activity clustering~~recognition model~~ (N~~HARM~~HAC). User only needs to label each clusters, so the efforts of labeling data is significantly reduced.~~User only needs to label each clusters, so the number of labeling data is decreasing.~~

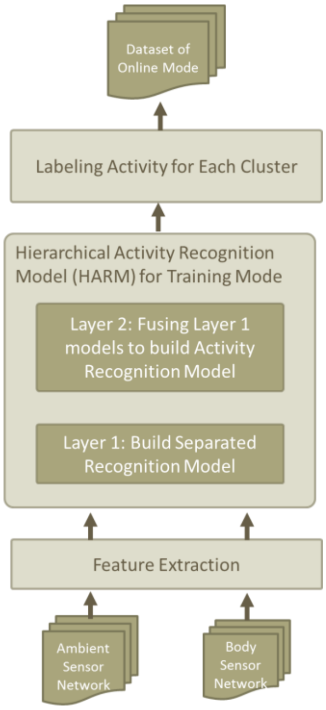


Fig. 3-1 Flowchart of Training Mode

To monitor elders' living activities, it's hard to ask them to label their activities for every instance in training section. Because the offline mode can find the significant living activity of each cluster, it help elders label the resulting clusters in a simpler way. Besides, the advantages of NHAC include that a simpler way is used to label data, and it provides a way to observe data and to categorize them automatically. A non-parametric learning method can find the most appropriate number of clusters based on their training data. For NHAC, it is a set of data-driven and non-parametric inference models, and it may discover some facts which are ignored by users, i.e., users usually remember significant activities, but ignores common or subconscious activities. Since the or-dinary activity recognition methods apply supervised learning algorithm, all the train-ing data need to labeled for then . If the user ignores those aforementioned common activities and label them into incorrect activities, those wrong labeled data will be-come noises in those supervised learning models, which in turn will mark the per-for-mances in prediction stage bad. Although one activity may not map to one cluster from our NHAC, one cluster can only be mapped to one activity. i.e., one activity can be mapped to one or more than one activity, and one cluster should be mapped only one activity. When users labels clusters, they can label different clusters as the same activ-ity. For example, if one has two habits of reading books: “turn on lamp” or “turn off lamp”, the NHAC formed two clusters for this reading behavior and both of them should be labeled as the same activity “Reading”. The following figure, Fig. 3 1, shows the flowchart of the training mode.~~To monitor elders’ living activities, but it’s hard to ask them labeling their activities for every instances in training section. Because the offline mode can find the significant living activity of each cluster, it help elders label the result clusters in the easy way. Besides, the advantages of NHARM is not only an easier way to label data, but also a way to observe data and category them automatically.~~

The online mode is used to real-time determine activities of the elder and to simultaneously monitor the anomaly activity. If an unknown activity has been observed, the system will determine whether it is an anomaly activity or not. The procedure of determining anomaly activity uses an artificial intelligent methodology called “case-based reason-ing (CBR)”, which is to solve new problems by adapting the previously additional successful solution to similar problems. So, before we design CBR in our activi-ty-aware healthcare system, we need to design a function to compute the similarity score of all known activities. This function can evaluate any new instance about how it is similar to all observed activities. And, we set a threshold T that is used to determine the instance which it is sufficiently similar to a known activity or is an unknown ac-

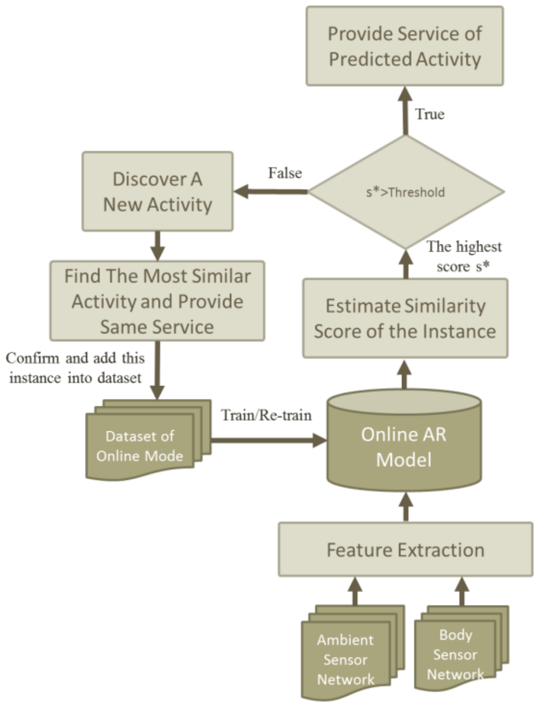


Fig. 3-2 Flowchart of Online Mode

tiv-ity.~~The online mode is used to real-time determine activity of elderly residents and monitor anomaly activity. If an unknown activity has been observed, the system will determine whether it is an anomaly activity or not. The procedure of determining anomaly activity is used an artificial intelligent methodology "case-based reasoning (CBR)". Case-based reasoning solves new problems by adapting previously successful solution to similar problems. So before we design CBR in our activity-aware healthcare system, we need to design a function to compute the similar score of all known activity. This function can evaluate any new instance how it is similar to all observed activities. And we set a threshold *T* that is used to determine the instance is belong to the most similar activity or an unknown activity.~~ If the highest similar score is still lower than threshold *T*, the system will consider this instance to be a new activity (or called a new problem). Then, the CBR mechanism will be triggered. If this unknown activity is similar to an anomaly activity, the system will send an alert message to the elder's caregiver. After the system provides service that sends an alert message or done nothing, it will ask resident to label this activity and retrain the AC model of online mode. The Fig. 3 2 shows the flowchart of the training mode.~~the system will consider this instance is a new activity (or called a new problem). Then, the CBR mechanism will be triggered. If this unknown activity is similar to an anomaly activity, the system will send an alert message to the elderly resident's caregiver. After the system provide service that sent an alert message or do nothing, it will require resident to label this activity and retrain the AR model of online mode. The Fig. 3-2 shows the flowchart of training mode.~~

In Chapter3, we will describe the algorithms and implementations of the AC model of both training and online mode in detil. Moreover, the architecture of the ac-tivity-aware healthcare system will be described in Chapter4, such as the labeling in-terface for training mode or the mechanism of CBR for online mode.

~~In the chapter3, I will detail describe the algorithms and implementations of AR model of both training and online mode. And the architecture of the activity aware healthcare system will describe in the chapter4, such as the labeling interface for training mode or the mechanism of CBR for online mode.~~

3.2 Activity Clustering~~Recognition Model~~ of Training Mode

~~3.3~~ The AC model in training mode has two main layers. Fusing information from different sensors to infer high level activities is also a hot topic. So the mechanism of fusing heterogeneous sensor data is proposed in the training mode. For the first layer, two separate AC models categorize training data from ambient and vital sign sensors; they identify activity individually. For the second layer, before we fuse the results from the first layer, it transforms the vital sign part to have a format where characteristic is event-triggered. The system considers the clusters of vital sign part from the first layer as new features. i.e., each cluster is considered as a sensor, so the number of clusters means the number of dimension of new features. Moreover each activity can be as-so-ciated with specific clusters, of which the nature is similar to that with ambient part. For example, the activity "watching TV" of ambient part can be determined by the "On" states of TV and lighting condition in the living room. Apparently, the activity is observed by the characteristic of event-trigger. For the part of vital sign model, when user watches TV and the AC model of vital sign part determines a cluster, the cluster is considered as a sensor which is used to monitor the user's activity of "watching TV". When the system determines a result of this cluster, the system infers that the activity, "watching TV", is taking place. Since the characteristics of both ambient part and vital sign

格式化: 本文

part are similar, they can be fused in an easier way.~~The AR model of training mode has two main layer. Fusing information from different sensors to infer high level activities is also a hot topic. So the mechanism of fusing heterogeneous sensor data is proposed in the training mode. For the first layer, two separated AR models categorize training data from ambient and vital sign sensors; they identify activity individually. For the second layer, before we fuse the results from the first layer, transforming vital sign part to have the characteristic of event-trigger format. The system considers the clusters of vital sign part from the first layer as new features. *i.e.*, each cluster is considered as a sensor, so the number of clusters means the number of dimension for new features. And each activity can map to specific clusters. This characteristic is similar to ambient part. For example, the activity "watching TV" of ambient part can be determined by the "On" states of TV and living light. Says, the activity is observed by the characteristic of event-trigger. For the part of vital sign model, when user watches TV and the AR model of vital sign part determines a cluster, the cluster is considered as a sensor which used to monitor user's activity of "watching TV". When the system determines a result of this cluster, the system can infer an activity "watching TV" occurred. The characteristic of both ambient part and vital sign part are similar, so they can fuse in an easier way.~~

~~3.3.1~~3.2.1 Activity Clustering ~~Recognition Model~~ from~~of~~ Ambient Sensor~~Part~~

Before building the AC model, extracting feature is an important task. In machine learning, feature extraction starts from an initial set of measured data and goal is to build the derived values intended to be informative, and to facilitate the subsequent learning. The ambient sensor data are extracted as Boolean variables: "on" and "off", because they are triggered by human activity, *e.g.* current sensor on Television measures low power when it is turned off; but measures high power when it is turned on. Generally speaking, the types of ambient sensor contain current sensor, lumen sensor, and switch sensor, and the state of each sensor is trained in the development stage. The feature of the current sensor presents the active status of the connected electronic appliance is “On” or “Off”. When a current sensor is attached to an electronic appliance, we will record its current data as ”Off”, namely the appliance isn’t turnd on , and compute mean m and standard deviation σ of the recorded data. A threshold T is used to determine whether the active status of the electronic appliance is “On” or “Off” by its current sensor data. The Threshold is $T = m + \sigma$. If a value of input sensor data is higher than T, the active status of its appliance is “On”; otherwise it is “Off”. The feature extracted from lumen sensor is likely to current sensor, whereas the switch sensor will response “On” or ”Off”, which explains why we do not need to preprocess the data of switch sensors.

The system collects ambient sensor data every 5 seconds as training data for a numbers of day, and uses k-nearest neighbor (KNN) algorithm to build the activity recognition model of ambient part. We need to decide cluster heads of the activity recognition model. The formats of features which are Boolean variables can be seen as binary patterns. Although the total number of cases(*i.e.* number of combinations) is two to the power equal to the number of sensors in theaory, the feature are usually sparse. An activity is usually related to two to four sensors, and the other sensors are void. In real life, actually the number of case types that appear is much less than two to the power equal to the number of sensors. We calculate the number of case types to find respective time-frequency (TF). If a case’ TF is higher than 1%, we regard this case as a cluster head. When the system has found all cluster heads, it uses k-nearest neighbor to build the activity recognition model, where k is generakky set to 3. In KNN, k is usually an odd number, and the meaning for “k=3” is that when k is 3, AC model will find up to three cased which are the most similar to the input case. Some difference in sensor data between the so-called similar cases and input case exist. For example, a new input case is observed when one is watching TV and turning on a fan. The KNN compares this case to all other cases, and find the three most similar cases are “watching TV”, “playing Kinect on TV”, and “watching TV and turning on the air conditioner”, respectively. The difference in obtained sensor data between “watching TV and turning on a fan” and “watching TV” is that the active status of the fan are the opposite; the difference in the obtained sensor data between “watching TV and turning on a fan” and “playing Kinect on TV” is that the active status of Kinect and the fan are the opposite; and The difference in the obtained sensor data between “watching TV and turning on a fan” and “watching TV and turning on the air conditioner” is that the active status of the air conditioner and the fan are the opposite. So, the nearest neighbor of “watching TV and turning on a fan” are two different patterns of watching TV and one pattern of playing Kinect. The instance of “watching TV and turning on a fan” should be just the activity of watching TV. The aforementioned process is a mechanism called majority voting, and we also would like to consider the difference between the case and its neighbor by designing a weight.

The features’ dimension is the number of sensors, m, and let the similarity measure be defined as Hamming distance. Let $T = \{x_1, \dots, x_N\}$ be the training set, and the training vectors $x_i \in R^m$ be vectors in the m-dimensional feature space. The following equation (3-1) is used to find the distance between new instance i and other training data j.~~Before building the AR model, extracting feature is an important part. In machine learning, feature extraction starts from an initial set of measured data and builds~~

~~derived values intended to be informative, facilitating the subsequent learning. The ambient sensor data is extract as Boolean variable: on and off, because they are triggered by human activity, *e.g.* current sensor on Television measures low power when Television is close; but measuring high power when Television is open. The types of ambient sensor contain current sensor, lumen sensor and switch sensor. The state of each sensor is training in the development stage. The feature of current sensor presents the active status of its attaching electronic appliance is "On" or "Off". When a current sensor attaches on one electronic appliance, we will record its current data in close status. And computing mean m and standard deviation σ of the recording data. A threshold T is used to determine the active status of the electronic appliance is "On" or "Off" by its current sensor data. The Threshold is $T = m + \sigma$. If a value of input sensor data is higher than T, the activity status of its appliance is "On"; otherwise it is "Off". The feature extracts from lumen sensor is likely to current sensor. And the switch sensor will response on and off, so we do not need to preprocess the data of switch sensors.~~

~~The system collects ambient sensor data every 5 seconds as training data for a numbers of day. And using k nearest neighbor (KNN) algorithm builds the activity recognition model of ambient part. We need to decide cluster heads of the activity recognition model. The formats of feature are Boolean variable that can be seen as binary patterns. Although the most type of case is two to the power of the number of sensors in theatrical, the feature are usually sparse. An activity is usually relative to two to four sensors, so other sensors are close. In real life, the number of type of appearing cases is much less than two to the power of the number of sensors. We calculate the number of each type of case to find their time frequency (TF). If a case' TF is higher than 1%, we regard this case as one of cluster head. When the system has found all cluster heads, it uses k-nearest neighbor to build the activity recognition model and sets k to 3. In KNN, k is~~

~~usually odd number, and the reason of "k=3" is that when k is 3, it means the AR model will find three most similar cases of the input case. A few difference of sensor data between those similar cases and input case. For example, a new input case is observed when one is watching TV and opening a fan. The KNN compares this case to all other cases, and find the most similar three cases. Each are "watching TV", "playing Kinect on TV" and "watching TV and opening air condition." The difference of observed sensor data between "watching TV and opening a fan" and "watching TV" is that the active status of the fan are opposite; the difference of observed sensor data between "watching TV and opening a fan" and "playing Kinect on TV" is that the active status of Kinect and the fan are opposite; and The difference of observed sensor data between "watching TV and opening a fan" and "watching TV and opening air condition" is that the active status of the air condition and the fan are opposite. So, the nearest neighbor of "watching TV and opening a fan" are two different patterns of watching TV and one patter of playing Kinect. The instance of "watching TV and opening a fan" should belong to the activity of watching TV. It is a mechanism called majority voting and we are also considering the difference between the case's neighbors by designed a weight.~~

~~The features' dimension is the number of sensors~~ $\cancel{m}$~~. And the similarity measurement uses Hamming distance. Let~~ $\cancel{T = \{x_1, \ldots, x_N\}}$ ~~be the training set, and the training vectors~~ $\cancel{x_i \in R^m}$ ~~are vectors in the m-dimensional feature space. The following equation (3-1) uses to find the distance between new instance~~ $\cancel{i}$ ~~and other training data~~ $\cancel{j}$~~.~~

$$dist(x_i, x_j) = \sum_{k=1}^{m} |x_{i,k} - x_{j,k}|$$
$$x_i = x_j \rightarrow dist(x_i, x_j) = 1 \qquad (3\text{-}1)$$
$$x_i \neq x_j \rightarrow dist(x_i, x_j) = 0$$

After finding all distances, the k nearest neighbors can be found. Let $f_i(x)$ be the cluster head for the *i-th* neighbor of x, and $\delta(\mathrm{c}, f_i(x))$ be the identity function. Let the function $\mathrm{g}(\mathrm{c})$ is used to present the number of neighbors with cluster c~~After finding all distances, the k nearest neighbors can be found. And let $f_i(x)$ be the cluster head for i-th neighbor of x. $\delta(\mathrm{c}, f_i(x))$ is the identity function. And a function $\mathrm{g}(\mathrm{c})$ is used to present the number of neighbors with cluster c.~~

$$\delta(\mathrm{c}, f_i(x)) = \begin{cases} \text{if } c = f_i(x), \text{then it is } 1 \\ \text{otherwise, then it is } 0 \end{cases} \tag{3-2}$$

$$\mathrm{g}(\mathrm{c}) = \sum_i \delta(\mathrm{c}, f_i(x)) \tag{3-3}$$

The function of weight voting is used to determine this which cluster instance belongs to. And the weight of each training data is the distance of this instance.~~The function of weight voting is used to determine this instance belongs to which cluster. And the weight of each training data is the distance of this instance.~~

格式化: 縮排: 第一行: 0.85 公分

$$w_i = \frac{1}{dist} \tag{3-4}$$

$$\mathrm{c}^* = \arg\max_c \sum_i w_i\, \delta(\mathrm{c}, f_i(x)) \tag{3-5}$$

which means that the instance will belong to the closet cluster c^*. Because the majority of instances whose TF are higher than 1% are considered to be cluster heads, only few instances whoseTF are lower than 1% are used to train the activity recognition model by KNN algorithm.

~~The instance will belong to the closet cluster c^*. Because the majority of instances which TF is higher than 1% are considering to cluster heads, only few instances which~~

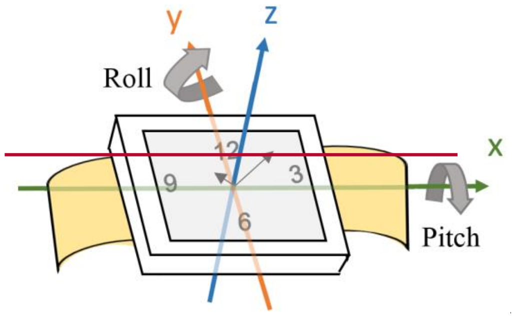
~~TF is lower than 1% are used to train the activity recognition model by KNN algorithm.~~

3.2.2 Activity Clustering Model from Vital Sign Sensor

~~3.3.2 Activity Recognition Model of Vital Sign Part~~

The appropriate features of vital sign data help the recognition model infer both static and dynamic behaviors. We choose two sensors to monitor user's vital sign, namely, accelerometer and gyroscope which are embedded in a wearable device called "ZenWatch". We try to monitor user's performed activites by his/her motion of wrist. An activity usually has its special wrist motion, e.g., when an elder reads a book, his wrist usually turns to a fixed direction and the associated motion is static. When this elder takes a walk, his/her wrist is regularly moved and the associated motion is dynamic. These two main kinds of activites are defined as "Posture type" and "Motion type" , respectively, where to the former usually refer to activites to some fixed directions but without moving hands,where the latter point to activites , not turning to fixed directions, whose hands are being moved most of the time.. Thus, we figure out that monitoring of turning direction and quantity of motion of user's wrist helps our recognition model infer user's activity correctly. In order to distinguished user's posture type and motion type activites, we monitor the orientation and acceleration of the wearable device and use them to build the AC model. The features of wearable sensor refer to mean and variance of tri-axis acceleration datameasured and computerd every second. Next, by using both tri-axis accelerometer and tri-gyroscope, we can get linear and angular accelerations whereby the sensor data of orientation (Pitch and Roll) can be derived. Here, Pitch and Roll are defined as follows: referring to Fig. 3-3, positive Pitch is defined when the wearable device starts by laying flat on a table and the positive Z-axis begins to tilt towards the positive Y-axis,whereas positive Roll is defined when the device starts by laying flat on a table and the positive Z-axis begins to tilt towards the positive X-axis.~~The appropriated features of vital sign data help the recognition model infer both static and dynamic behaviors. We choose two sensors to monitor user's vital sign each is accelerometer and gyroscope and they are built on a wearable device "Zen watch". We try to monitor user's behavior by his/her motion of wrist. An activity has its own action on wrist. e.g., when an elderly man reads a book, his wrist usually turns in a fixed direction and belongs to static action. When this man takes a walk, his wrist is regular moving and belongs to dynamic action. These two main kinds of behavior are defined as "Posture" and "Motion" actions. The behaviors of "Posture" are usually turning in a fixed direction but without moving hands. In contrast, the behaviors of "Motion" are always moving hands but not turning in a fixed direction. So we figure out that monitoring user's wrist turning direction and quantity of motion helps our recognition model infer right activity. In order to realize user's posture and motion actions, we monitor the wearable device's orientation and its acceleration variable and use them to build the AR model. The features of wearable sensor data are using tri-axis accelerometer to compute mean and variable value of each axis every second. And using both tri-axis accelerometer and tri-gyroscope to get their accelerations and angle accelerations. Each axis of orientation (Pitch and Roll) can be captured from acceleration and angle acceleration. That the definition of Pitch and Roll are as follow. Positive Pitch is defined when the wearable device starts by laying flat on a table and the positive Z-axis begins to tilt towards the positive Y-axis. Positive Roll is defined when the phone starts by laying flat on a table and the positive Z-axis begins to tilt towards the positive X-axis.~~

We need to compute mean and variance of acceleration at each axis every second. Let $A_l = \{a_{l1}, a_{l2}, \dots, a_{ln}\}$ denote the set obtained every second of acceleration data, $l = 1,2,3$, where $l = 1$ represents x-axis, $l = 2$ represents y-axis, and $l = 3$ represents z-axis. And for our wearable device, the sampling rate of accelerometer is about 30 Hz, so the number of acceleration data n is about 30.~~We need to compute mean and variance of acceleration at each axis every second. That~~ $\cancel{A_l = \{a_{l1}, a_{l2}, \dots, a_{ln}\}}$ ~~is the set of every second's acceleration data and~~ $\cancel{l = 1,2,3}$ ~~that~~ $\cancel{l = 1}$ ~~represents x-axis,~~ $\cancel{l = 2}$ ~~represents y-axis and~~ $\cancel{l = 3}$ ~~represents z-axis. And for our wearable device, the sampling rate of accelerometer is about 30 Hz, so the number of acceleration data~~ $\cancel{n}$ ~~is~~



~~Fig. 3-3 The Pith and Roll of a smart watch~~

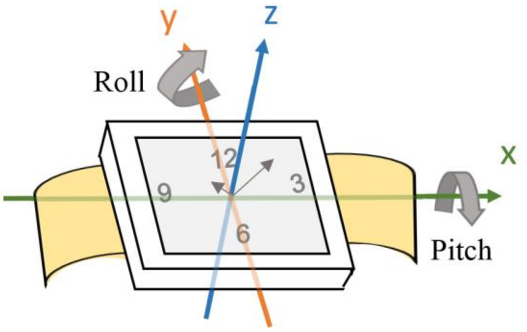


Fig. 3-3 The Pith and Roll of a smart watch

格式化表格

~~about 30.~~

$$Mean_l(\mathrm{A}_l) = \frac{\sum_{i=1}^{n} a_{li}}{\|A\|} \tag{3-6}$$

$$Var_l(\mathrm{A}_l) = \frac{\sum_{i=1}^{n}(a_l - Mean_l(\mathrm{A}_l))^2}{n} \tag{3-7}$$

We also choose signal vector magnitude (SVM) to measure this activity's strength[30].

$$SVM(\mathrm{A}_x, \mathrm{A}_y, \mathrm{A}_z) = \sqrt{\sum_{i=1}^{n} a_{xi} + \sum_{i=1}^{n} a_{yi} + \sum_{i=1}^{n} a_{zi}} \tag{3-8}$$

The orientation information is captured from wearable device's android API. This API can capture Pitch and Roll by the linear and angular accelerations derived from the sensor data as mentioned previously. The features of Pitch and Roll are their mean and variation. Let $O_l = \{o_{l1}, o_{l2}, \dots, o_{lm}\}$ denote the set of orientation data obtained every second , $l = 1,2$, where $l = 1$ represents Pitch, and $l = 2$ represents Roll. Both the sampling rates of accelerometer and gyroscope are about 30 Hz, but their sampling times are not synchronized. The number of orientation data m is not always equal to n, but it is still near to 30.~~The orientation information is captured from wearable device's android API. This API can capture Pitch and Roll by the sensor data acceleration and angle acceleration. The features of Pitch and Roll are their mean of direction and their mean of angle variation. That~~ $\cancel{O_l = \{o_{l1}, o_{l2}, \dots, o_{lm}\}}$ ~~is the set of every second's orientation data and~~ $\cancel{l = 1,2}$ ~~that~~ $\cancel{l = 1}$ ~~represents Pitch,~~ $\cancel{l = 2}$ ~~represents Roll. Both the sampling rates of accelerometer and gyroscope are about 30 Hz, but their sampling time not synchronize. The number of orientation data~~ $\cancel{m}$ ~~is not always equal to~~ $\cancel{n}$~~, but it is still near to 30.~~

$$Angle\mathrm{Mean}_l(O_l) = \frac{\sum_{i=2}^{m} o_{\mathrm{li}} - o_{\mathrm{l(i-1)}}}{m} \tag{3-9}$$

$$\mathrm{OrientationMean}_l(O_l) = \frac{\sum_{i=1}^{m} o_{\mathrm{li}}}{m} \tag{3-10}$$

We can extract 11 features from wearable sensor data every second. Each feature inside the wrist watch indicates a waving action of arm, and numbers of consecutive waving motions can be associated with a specific activity. So we propose a topic model to infer one minute behavior performing to a specific activity from wearable sensor data. Thus, hereby defined topic model can be used to extract activity pattern from the sensor data, and in turn to recognize those daily routines. We take each feature collected every second as "word" in the topic model, and let 11 words construct a vocabulary. Each topic must include several identical types of vocabularies, *i.e.*, one topic can be seen as one activity, such as walking, sitting, standing, eating meal, etc.

The topic model is constructed by two-layer Dirichlet process mixture model (DPMM), abbreviated as 2LDPMM. The 1st-layer 2LDPMM, can extract categories of features from raw data. For example, people have different kinds of hand's waving motion, such as drooping hands, horizontally waving hands, vertically waving hands or show of hands. Generally, it is hard to define the specific number of kinds of hand's waving motions, and the 1st-Layer 2LDPMM is suitable to learn hand's waving motions because DPMM is a non-parametric unsupervised clustering model. It can find different kinds of hand's waving motion from raw data without given a specific number of motion types. As we may know, temporal information is not applicable to extract in traditional unsupervised clustering methods, and the 2nd-layer 2LDPMM is going to remedy this drawback by grouping 60 continuous outputs of 1st-layer 2LDPMM into one new feature. We take statistics of the occurrence time of each waving motion from 1 minute result of first layer, and construct the statistic of result as the feature of the 2nd-layer

2LDPMM. The statistics of waving motion can be seen as a meaningful action of user's body behavior. For exanple, people usually do not change their waving motions much when they perform specific actions, *e.g.* while sitting, hands are usually placed on thigh fixedly; while sweeping, hands whip regularly; while having a meal, hands are put on table sometimes and take the bowl sometimes. Fig. 3-4 shows the histograms of waving motions of three meaningful actions. The aggregated 1st-layer 2LDPMM results as new feature for 2nd-layer 2LDPMM; the horizontal number means the waving motions found in 1st-layer 2LDPMM; the vertical number means the occurrence time of each waving motion. The daily activities of reading and watching TV have similar histograms because their body behavior are similar, and we refer to them as sitting.~~We can extract 11 features from wearable sensor data every second. Each feature inside the wrist watch indicates a waving action of arm, and numbers of consecutive waving motions can be associated with a specific activity. So we propose a topic model to infer one minute behavior belonging to which activity from wearable sensor data. The topic model can be used to automatically extract activity pattern from the sensor data and then to recognize those daily routines. We take each feature collected every second as "word" in the topic model, and 11 words construct a vocabulary. Each topic must include several identical types of vocabularies, i.e., one topic can be seen as one activity, such as walking, sitting, standing and eating meal, etc.~~

~~The topic model is constructed by two layers Dirichlet process mixture model (DPMM), called 2LDPMM. In the first layer of 2LDPMM, it can extract categories features from raw data. For example, people have different kinds of hand's waving motion, such as drooping hands, horizontally waving hands, vertically waving hands or show of hands. It's hard to define the specific number of kinds of hand's waving motions. 1st Layer DPMM is used to learn hand's waving motions because DPMM is a~~

~~non-parametric unsupervised clustering model. It can find different kinds of hand's waving motion from raw data without given a specific number of motion types. The temporal information is hardly extracting in traditional unsupervised clustering methods. The second layer of 2LDPMM is going to reduce this drawback. The training feature of 2nd layer DPMM are grouped into one new feature by 60 continuous output of 1st layer 2LDPMM. We statistic the occurrence time of each waving motion from 1 minute result of first layer, and construct the statistic result as the feature of 2nd layer 2LDPMM. The statistic of waving motion can be seen as a meaningful action of user's body behavior; says, people usually have changeless waving motions to do a specific action, *e.g.* while sitting, hands usually place on thigh fixedly; while sweeping, hands whip regularly; while having meal, hands put on table sometimes and take the bowl sometimes. Fig. 3-4 shows the histograms of waving motions of three meaningful actions. The aggregated 1st layer 2LDPMM results as new feature for 2nd layer 2LDPMM; the horizontal number means the waving motions found in 1st layer 2LDPMM; the vertical number means the occurrence time of each waving motion. The daily activities of reading and watching TV have similar histograms because their body behavior are similar and called sitting.~~

We use 2LDPMM to recognize body behavior with two reasons. First, Dirichlet process mixture model can find meaningful clusters without prior knowledge of the number of cluster; second, it's a powerful clustering method to retrieve latent information from raw data.

Before building the topic model 2LDPMM, the preliminary of Dirichlet process is described in Chapter2.~~We use 2LDPMM to recognize body behavior with two reasons; first, Dirichlet process mixture model can find meaningful cluster without given the number of cluster; second, it's a powerful clustering method to retrieve latent information from raw data. Before building the topic model 2LDPMM, the prior knowledge of Dirichlet process is describing in chapter2.~~ So I will briefly describe Dirichlet distribution and Dirichlet process here. A Dirichlet distribution Dir(α) is a distribution over multinomial variables, and it can be seen as the multivariate generalization of the beta distribution. It's a family of continuous multivariate probability distributions parameterized by a vector α of positive reals. Its probability density function returns the probabilities of K rival events are π_i given hat each event has been observed $\alpha_i - 1$

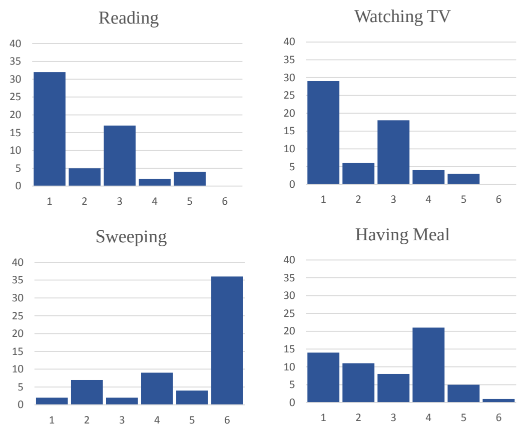


Fig. 3-4 Four types of activities' histograms

times. Says, Dirichlet distribution is the conjugate prior distribution of multinomial distribution.

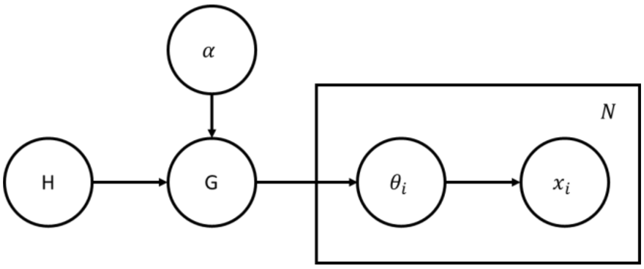
The Dirichlet process is an infinite-dimensional generalization of the Dirichlet distribution and it is denoted as DP(α,H), where α is a positive real number called the concentration parameter and H is a base distribution. The Dirichlet process tries to draw distributions around the base distribution. And, the concentration parameter used to specify the strong level of the discretization, such as in the limit of $\alpha \to 0$ means the realizations are all concentrated on a single value. On the contrary, in the case where limit of $\alpha \to \infty$ means the realizations become continuous.

The Dirichlet process mixture model generalizes a mixture model with infinite mixture components. A mixture model is a hierarchical model, and it's a probabilistic model for representing the presence of subpopulations within an overall population [31]. Mixture models are used to make statistical inference about the properties of the subpopulations given only observations on the pooled population, without subpopulation identity information. And, it consists of the following components. K is number of mixture components, and N is number of observations. The parameter θ_i is the distribution of observation that associated with component i, where $i = 1, \ldots, K$. The mixture weight ϕ_i is the prior probability of a particular component i. Φ is K-dimensional vector composed of all the individual $\phi_{1\ldots K}$. And z_i is the component of observation i; x_i is the observation i. Let $f(x|\theta)$ be the probability distribution of an observation. So z_i belongs to Categorical(ϕ) and $\mathrm{x_i}$ belongs to $f(\theta_{z_i})$.

A data point x_i is drawn from the distribution $P(x)$.

$$P(x) = \sum_{k=1}^{K} \phi_k f(x|\theta_k) \tag{3-11}$$

When the mixture weight $\Phi = \phi_1, \ldots, \phi_k$ is multinomial distribution, we can use



the Dirichlet distribution as its prior. In DPMM, the number of mixture components is infinite, so the original mixture model needs to be modified as follows:

$$\theta_{z_i} \sim H \; for \; z_i = \{1, \dots, K\},$$

格式化: 字型: 斜體

$$\phi_1, \dots, \phi_K \sim Dir(\alpha/K, \dots, \alpha/K),$$

格式化: 字型: 斜體

$$z_i \sim Multinomial(\phi_1, \dots, \phi_K),$$

格式化: 字型: 斜體

$$x_i \sim f(x|\theta_{z_i})$$

格式化: 字型: 斜體

In the limiting case, $k \rightarrow \infty$, the mixture model becomes

$$P(x) = \sum_{k=1}^{\infty \cancel{K}} \phi_k \, f(x|\theta_k) \tag{3-12}$$

The Dirichlet distribution becomes the Dirichlet process as shown in the following steps:

格式化: 縮排: 第一行: 0 公分

$$G \sim Dir(\alpha, H),$$

$$\theta_i \sim G,$$

格式化: 字型: 斜體

$$x_i \sim f(x|\theta_i)$$

格式化: 字型: 斜體

The prior distribution function G is drawn from a Dirichlet process and, α is the concentration parameter, and H is the base prior. Given G, x_i belongs to θ_i, and we sample θ_i to the components. Then, given θ_i, we generate each data point x_i from acceleration features. For implementation of DPMM, we use an open API called Dataumbox to train our arms' waving motions. The DPMM uses Gibbs sampling[32] algorithm and we build the Dirichlet process by Chinese restaurant process [33].

Fig. 3-5 The plate notation of DPMM

Using DPMM twice is seen as one topic model, called 2LDPMM. The 1^{st}~~first~~ layer ~~of~~ 2LDPMM is used to extract pattern from raw data; before processing the 2^{nd}~~second~~ layer ~~of~~ 2LDPMM, we aggregate those patterns from the first layer into a time-windows. The aggregated data represent the pattern distribution over every time-window and are consided~~r~~ as new features for the second layer DPMM. The result of each cluster means a meaningful topic,~~.~~ ~~In~~ and our work, it means the meaningful actions, *e.g.* sitting and sweeping.

~~3.3.3~~3.2.3 Activity ~~Recognition~~ Clustering from Fuing Ambient and Vital Sign Sensors~~Model of Ambient and Vital Sign Fusion~~

The result of vital sign part is obtained from the topic model 2LDPMM,~~.~~ ~~The 2LDPMM~~which categories the training data into n clusters and each cluster can represent a mapped~~ing~~ activity. We consider each cluster as a sensor used to monitor a specific activity, ~~so~~ and thus the result of each cluster becomes a feature for the second layer non-parametric hierarchal activity ~~recognition~~ clustering ~~model~~ (the 2^{nd}-layer ~~NHARM~~HACNHAC). The 2^{nd}-layer N~~HARM~~HAC transfers the result of 2LDPMM to~~as~~ $F = \{f_1, f_2, \ldots, f_n\}$, where n is the number of determined clusters from 2LDPMM. ~~And~~ Moreover, F is the set of features of the 2^{nd}-layer NHAC~~RM,~~, and the format of f_i is Boolean variable. For example, since 2LDPMM determines the input data as the 3^{rd} cluster, and there is totally 10 clusters in 2LDPMM,~~.~~ t~~T~~he number of features of 2^{nd}-layer N~~HARM~~HAC is 10 ($n = 10$) and the features of this instance are $f_1 = 0, f_2 = 0, f_3 = 1, f_4 = 0, f_5 = 0, f_6 = 0, f_7 = 0, f_8 = 0, f_9 = 0$ and $f_{10} = 0$.

Extracting features from ambient part of the 1st-layer ~~NHARM~~HAC~~-~~NHAC to the 2nd-layer ~~NHARM~~HAC~~—~~NHAC uses ~~alike~~ similar mechanism. The 2nd-layer ~~NHARM~~HAC~~–~~NHAC transfers the result of ambient part KNN model as $F =$

where m is the number of determined clusters from the ambient part KNN model. And F is the set of features of the 2nd-layer ~~NHARM~~HACNHAC, and the format of f_i is variable. So the total number of features is $n + m,$ and their formats are Boolean variable. The characteristic of feature in the 2nd-layer N~~HARM~~HAC is similar to the of 1st-layer N~~HARM~~HAC. We use same concept to build the 2nd-layer N~~HARM~~HAC by neighbor algorithm but using different weight voting mechanism.

The 2nd-layer N~~HARM~~HAC ~~inference~~ infers the resident's activity every 5 seconds, ~~so~~ and thus the training data from the first layer should send their result every 5 seconds. The ambient part snapshot, the situation of the environmental sensors~~' situation~~ every second~~.~~, and t~~T~~he vital sign part determines user's waving action every second also. For every 5 seconds, the system accumulates 60 successive result of waving actions from the 1st ~~first~~ layer ~~of~~ 2LDPMM, and the predicts the cluster which this successive waving actions belong~~sing~~ to ~~which cluster~~ in the ~~second~~ 2nd layer ~~of~~ 2LDPMM. Ather b~~B~~oth ambient and vital sign parts send their predicted results, we can then generate the features for the 2nd layer ~~HARM~~NHAC, and ~~the~~ such process has been described in the previous paragraph.

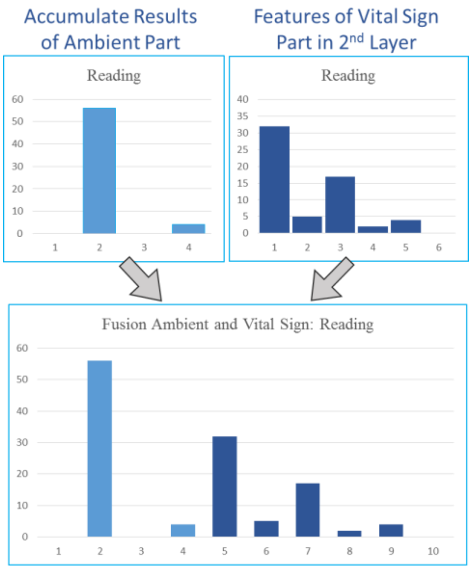


Fig. 3-6 The feature of 2^{nd} layer ~~HARM~~NHAC

For KNN, we need to decide cluster heads from the training data. The implemented flow is similar to that of the ambient part AR model. We calculate the times of occurrence of each type of training data, known ~~occurrence time~~ as time-frequency (TF). If ~~a case'~~ TF of a case is higher than 1%, we regard this case as one of cluster head. When the system has found all cluster heads, it uses k-nearest neighbor to build the activity recognition model~~ and also sets k to 3. F is used to find the instance of cluster head.~~ If an instance belongs to a cluster head, we label this instance as this cluster~~. A~~, and the format of KNN features is the joint result of ambient part KNN result and vital sign part 2LDPMM feature. ~~That~~ Let the number of feature ~~is~~ be n in ambient part KNN, and~~;~~ the number of feature be~~is~~ m in vital sign part 2^{nd} layer 2LDPMM, and then t~~. T~~he features' dimension is $n + m$. In order to normalize the distance between ambient and vital sign part, we also accumulate the results of ambient par~~t and multiply 5~~. Because the feature ~~which~~ extracted~~s~~ from vital sign part are the accumulation of the successive

60 instances ~~that the~~over one second time interval~~ interval time is one second. T~~, and the feature ~~which~~ extracted~~s~~ from the ambient part are the snapshot of environmental sensors ~~and~~ over the 5 second time interval.~~ time is 5 seconds. So we accumulate the result from ambient part and multiply to 5.~~ ~~And~~ Note that the similarity measurement also uses Manhattan distance. The format of fusion feature is shown~~s~~ in~~as~~ Fig. 3-6. the accumulated results of ambient part and the features of vital sign part. In this sample, the ambient part has found 4 clusters and the vital sign part has found 6 clusters in its 1st 2LDPMM.

Let $T = \{x_1, \dots, x_N\}$ be the training set. N is the number of training data, and the training vectors $x_i \in R^{n+m}$ are vectors in the (n+m)-dimensional feature space. The following equation (3-13) is used~~s~~ to find the distance between new instance i and training ~~data~~ datum j.

$$dist(x_i, x_j) = \sum_{k=1}^{n+m} |x_{i,k} - x_{j,k}| \tag{3-13}$$

After finding all distances, the k nearest neighbors can be found. And let $f_i(x)$ be the cluster head for i-th neighbor of x. $\delta(\mathrm{c}, f_i(x))$ is the identity function. And a function $\mathrm{g(c)}$ is used to present the number of neighbors with cluster c.

$$\delta(\mathrm{c}, f_i(x)) = \begin{cases} \text{if } c = f_i(x), \text{then it is 1} \\ \text{otherwise, then it is 0} \end{cases} \tag{3-14}$$

$$\mathrm{g(c)} = \sum_j \delta(\mathrm{c}, f_i(x)) \tag{3-15}$$

The function of weight voting is used to ~~determine~~ find the cluster to whcich this instance belongs~~ to which cluster~~. And the weight of each training data is the distance of this instance.

$$w_i = \frac{1}{dist} \tag{3-16}$$

$$c^* = \arg\max_{c} \sum_{i} w_i \, \delta\big(c, f_i(x)\big) \tag{3-17}$$

格式化: 縮排: 第一行: 0 公分

The instance will belong to the closest cluster c^*. Because the majority of instances ~~which~~ whose TFs ~~is~~ are higher than 1% are considered~~ing~~ ~~to~~ as cluster heads, only few instances ~~which~~ whose TFs ~~is~~ are lower than 1% are used to train the activity recognition model by KNN algorithm. This KNN model is used to ~~category~~ categorize all training data, and the categorized results are used to reduce the burden of data labeling ~~data~~. User only needs to label each category that each category is one cluster ~~of~~ in the 2nd Layer N~~HARM~~HAC. The labeling procedure will be described in C~~the c~~hapter 4.

~~3.4~~3.3 Activity Recognition Model of Online Mode

In the offline mode, the system helps user categorize~~y~~ their living ~~behaviors~~ activities to a number of classes. Each class represents an activity of daily living. And the user should label all classes as the training data for online mode AR model. Because the training data are labeled, we choose a supervised learning method "Dynamic Bayesian Network (DBN)" to build the online mode AR model. DBN is an extended version of a Bayesian network (BN), which is a probabilistic graphical model representing a set of variables and their conditional independencies via directed acyclic graph. The random variables in DBNs are identified in the previous section. The ordinary BN only considers the situation in which all random variables occur at same time, but the DBN is able to model the relationships associated with random variables in each time slice. BNs are suitable for data analysis in a clinical context because they allow for interpretation; i.e., one can construct causal relationships for each random variable by asking a clinician

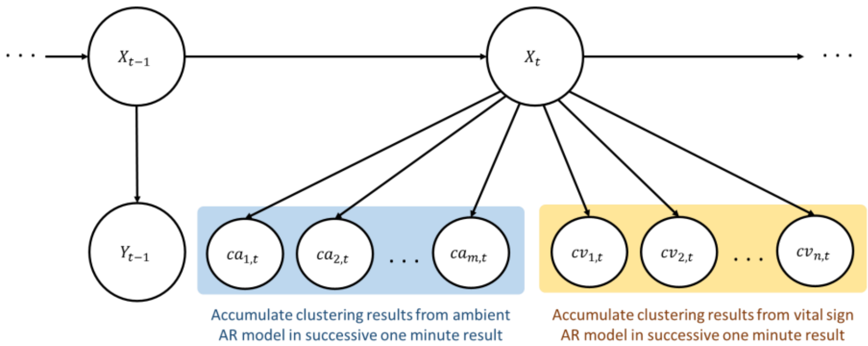


Fig. 3-7 The graph structure of online mode

whether it makes medical sense. DBN is also an interpretable model; however, it considers the temporal relationships of random variables. This is the reason that we selected DBN for estimating the health status of patients. In other words, variations in behavior are reflected in the patterns of fluctuating environment information and vital signs information; therefore, the DBN examines both ambient and vital sign sensor data from training data to estimate human activity. Fig. 3-7 illustrates the proposed DBN.

We consider the training data is sequential data $S = \{y_1, \dots, y_t, \dots, y_T\}$, $y_t \in R^{(n+m)}$ and T is the number of training data. Every y is generated by a state x_i where $i = 1, \dots, T$. The joint probability is formulated as follows:

$$P(Y, x_1, x_2, \dots, x_T) = \prod_{t=1}^{T} \prod_{i=1} P\left(Y_{t,i} \middle| x_t\right) P(x_t|x_{t-1}) \tag{3-18}$$

where $y_{t,i}$ refers to each feature of one instance at time *t*. Since all variables are discrete; therefore, each conditional probability is assumed to have categorical distribution. The probability mass function (PMF) of multinomial categorical distribution for variable *x* is as follows:

$$\mathrm{f}(y = i|\mathrm{p}) = p_i \tag{3-19}$$

$$\sum_{i=1}^{K} p_i = 1$$

where p_i is the probability of y equaling i. The number of possible outcomes of i is determined by the number of states in each random variable.

The arrow from node A to B in Bayesian networks represents the conditional probability $P(B|A)$. Through maximum likelihood estimation (MLE) using historical data, each conditional probability in the DBN model can be learned, even when hidden variables exist. It should be noted that all conditional probabilities are assumed to follow categorical distribution because all variables are discrete and no prior assumption is made about which distribution is followed.

The MLE algorithm is used to obtain the joint probabilities capable of maximizing the likelihood that training data will be obtained. For the purpose of estimating living activity, we apply the belief propagation algorithm to estimate the marginal probability of the interested variable X. In other words, the marginal probability of the current activity $P(X_t)$, is estimated by the observations of X_t and the node of X_{t-1} that propagates from the parent and the neighboring node of X. Once $P(X_t)$ is estimated for each possible value on X_t, the most likely activity estimate of the user c^* is:

$$c^* = \arg\max_{c} P(x_{c,t}) \qquad (3\text{-}20)$$

We can use DBN to real-time recognize user's activity by the current input and the previous recognized activity. The sequential pattern is considered in the online activity recognition model. If a sensor responses error data into the system, the online AR model may have ability to identify the true activity from its parent node.

Chapter 4

Activity of Daily Living-aware Elderly Healthcare System

4.1 System Overview

The healthcare system is used to real-time ~~aware~~ beware the residents' activity based on the proposed hierarchical activity recognition model. We describe the adaptive learning mechanism of this healthcare system in detail. The framework of online mode and offline mode are also described in this chapter. The offline mode is an important role to build the first generation AR model. In the offline mode, the system collects data from environment and wearable devices, and after ~~number of~~certain days the system will train the hierarchical AR model and provide an interface for residents to label their activities~~y~~. In the online mode, the system provides a real-time monitoring function and it has ability to discover unknown activity and retrain the activity recognition model. We use~~s~~ case-based reasoning (CBR) algorithm to implement the function of discovery. CBR is an artificial intelligent method and the process of solving new problems based on the solutions of similar past problems.

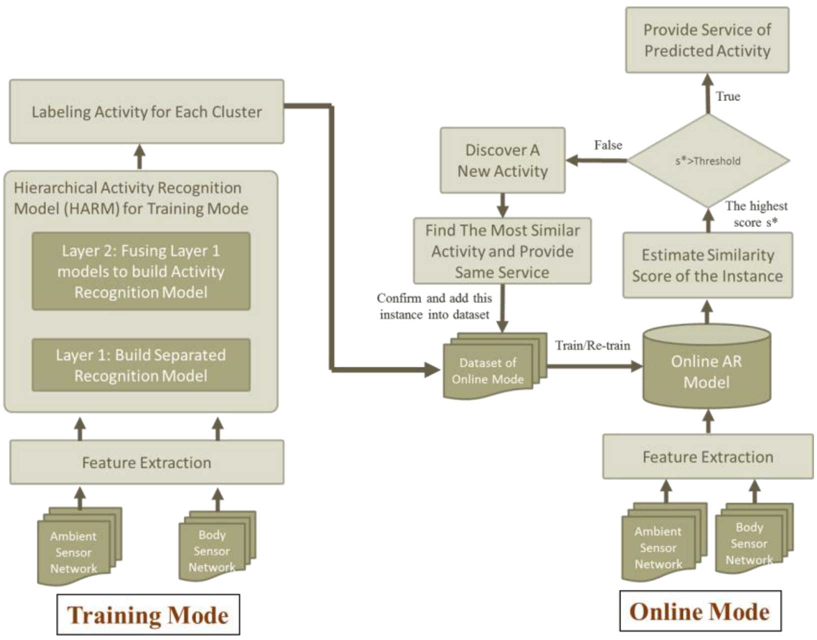


Fig. 4-1 The flowchart of activity-aware elderly healthcare system

Fig. 4-1 illustrates the flowchart of the healthcare system. The functions of activity recognition are described in ~~chapter3~~Chapter3. Based on the Hierarchical Activity Recognition Model in training mode, users can easily label training data with less burden. Those labeled data are ~~constructed~~ put to a dataset for building activity recognition model of online mode. And, the online mode AR model is constructed by dynamic Bayesian network, which is a sequential pattern recognition methodology. According to the previous predicted result and the input observations, the online mode AR model predicts the current activity. Before the system provides service, the system computes similarity score of the activity. If the similarity score is lower than threshold, the system ~~seems~~ takes this activity ~~is~~ as an unknown activity and triggers case-based reasoning mechanism. It finds the cases which are the most similar~~ity~~ ~~cases by~~to that unknown activity through the similarity score and provides the same service immediately. After

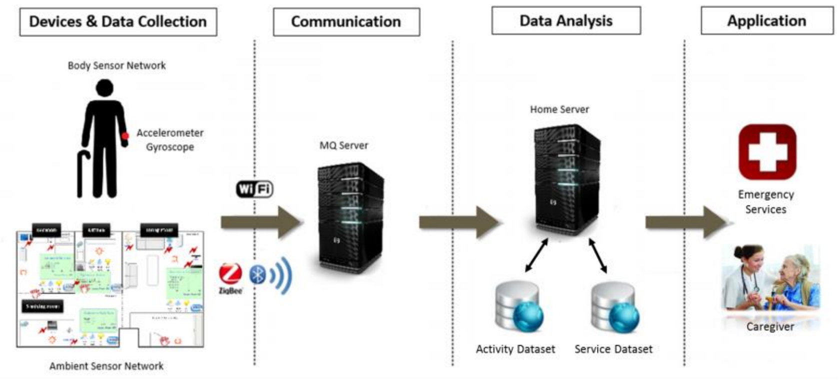


Fig. 4-2 The activity-aware elderly healthcare system overview

the system provides services, it will ~~require~~ ask user to confirm this new activity and give an associated~~a mapping~~ service. ~~While~~ After the system receives the confirmed message and updates the training dataset, the system will re-train the online mode AR model. For the procedure, the system in online mode is able to learning new lifestyle activities without ~~expert~~ professional maintaince.

The overall design of the activity-aware elderly healthcare system in pervasive environment is shown in Fig. 4-2. The elder~~ly people~~'s vital signs are collected by smart watch with a wireless body sensor network, and it can collect ~~different~~ variance kinds of body sensors in the future, *e.g.*, heart rate sensor, blood pressure measurement sensor, etc. And environmental sensors are deployed to monitor the pervasive environment with ambient sensor network. Ambient sensor network includes light sensor (lumen sensor), current sensor and switch sensor through wireless connection. In the communication layer, we use MQ server to transmit collected data to home server. The collected data are too many, so the communication layer reduces the burden on network traffic. In the analysis layer, collected data are used to real-time recognize activity of user. The activity dataset stores the training data and ~~new~~ the newly ~~discovering~~ discovered activity.

The service dataset stores the service of ~~mapping~~ mapped activity. When the activity recognition model detects an abnormal activity ~~that~~ and its corresponding service is to send~~ing~~ an alert message, the application layer will really send ~~an alert~~that message to user's caregiver. Based on the ability of activity recognition, many helpful application can be generated, *e.g.*, anomaly detection by analyzing user's activity of daily living. We record all observed activities into an ADL report, so user's caregiver can evaluate the elderly's health status or individual ability by the ADL report.

4.2 Labeling Interface

The ordinary activity recognitions are usually generated by supervised learning methods, but the procedure of labeling data is a high burden task for users. Even techniques of Internet of Things are flourishing, the difficulty of labeling makes the technology of activity recognition ~~is~~ still in the experimental stage. However, monitoring activity of daily living for elderly healthcare ~~are needed~~is increasingly important in the country with aging population~~ country~~. The automatic activity recognition system supports "aging place". An easier activity recognition is required for those elderly people and their caregivers. So the proposed activity-aware system provides a mechanism for easier data labeling~~ data~~. In the training mode, the activity recognition ~~categories~~ categorizes ~~amount~~ a large guantity of observed data ~~as~~ into a lower number~~s~~ of clusters,~~.~~ ~~Each~~ where each cluster represents one daily activity from those observed data. Users only need to label those clusters, so the burden of labeling data is reduced~~ing~~.

Fig. 4-4 A sample of the labeling interface that clusters are labeled

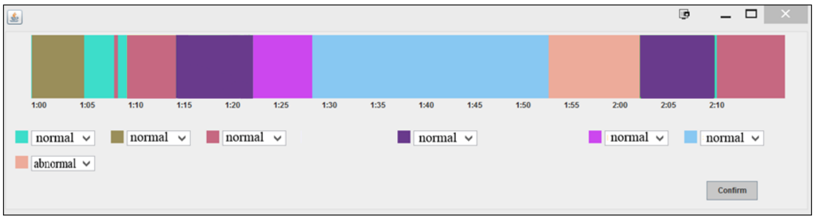


Fig. 4-5 A sample of the labeling interface that requires annotating services

Fig. 4-3 shows a sample of labeling interface. In the sample, the system collects sensor data from PM 1:00 to PM 2:20, and analyzes the data to identify different activities. There are nine clusters ~~are~~ extracted by the hierarchical activity recognition model.

Fig. 4-4 shows that the user has labeled each cluster through~~of~~ the labeling interface. We can observe that the cluster of Purple and Green are the same activity "Read". Because the user did different activity freely, the user did one activity with a ~~little~~ slightly different action~~s in~~ at different time. Although the system considers the activity "Read" as two clusters, user can label them into the same activity. While user put the

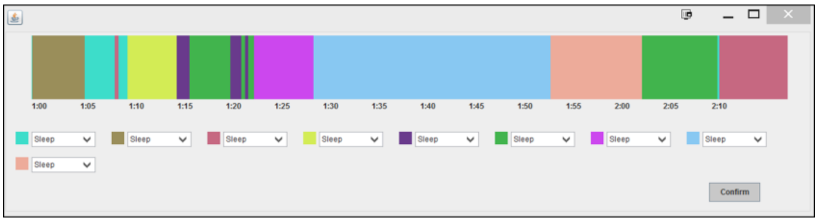


Fig. 4-3 A sample of the labeling interface that data are not labeled

button of confirm action in the interface, the system will generate the dataset for online mode. The dataset contains the processed training data of training mode and their labeling activity information. And, the system will send another ~~require~~ request to annotate the service for each activity. The color of repeated clusters of one activity ~~are~~ is replaced~~s~~ by that of the ~~first occurring cluster~~cluster that takes place first. Fig. 4-5 shows tating services. The proposed activity-aware healthcare system only provides a simple service ~~that activity are~~to categorize~~d~~ the activity to either of the two types: normal and While the system detects an activity that belongs to abnormal state, it will send an alert message to user’s caregiver.

4.3 Adaptive Learning

The function of adaptive learning makes the system individually working without ~~expert~~ professional maintaince. In online mode, the input sensor data are identified by DBN model. We can find the probability of each activity. The probability function $P(X_t)$ is estimated for each possible value on X_t~~. That X_t~~, which is the set of all possible activity, and it denotes as $X_t = \{x_{t,1}, x_{t,1}, \dots, x_{t,N}\}$, where ~~exists~~ N stands for the number of all possible activities. The system chooses top three~~3~~ possible activities to compute their similarity score by similarity function. If their similarity scores are lower than a threshold T, the system considers the current activity a~~i~~s an unknown activity. The mechanism of case-based reasoning is triggered~~ing~~ to provide a similar service and confirms the activity. Then, the system re-trains the DBN model as adaptive learning model.

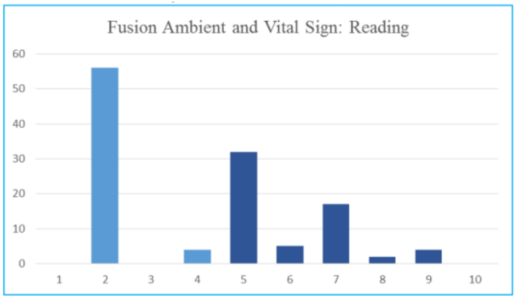


Fig. 4-6 The data format of similarity function

4.3.1 Similarity Function

We apply k-NN to calculate the distance between the current input data and all top three~~3~~ similar activities. The input data are simply transferred by two AR models of 1st layer ~~HARM~~NHAC. Fig. 4-6 shows a sample, ~~that~~ cohere the horizontal value represented actions from both ambient part and vital sign part models,~~.~~and t~~T~~he vertical represents the times of occurrence ~~time~~ in one minute and the total occurrence time of input data is 120. Let X_t ~~as~~ be the input data for similarity function,~~.~~ ~~That~~ $X_t = \{ x_t^{(1)}, x_t^{(2)}, \ldots, x_t^{(N)} \}$, ~~and~~ where N is the number of possible actions from both ambient and vital sign part.

The similarity function is generated by the difference of distances. The following equation (4-1) is used~~s~~ to find the distance between the current instance X and the training data of all top 3 similar activity Y. Then, we find the mean value of~~do~~ their distances ~~of each~~to that activity.

$$dist(X,Y) = \sum_{i=1}^{N} \left| x^{(i)} - y^{(i)} \right| \tag{4-1}$$

$$dist_l(X, Y_1, Y_2, \ldots, Y_K) = \frac{\sum_{j=1}^{K} \sum_{i=1}^{N} \left| x^{(i)} - y_j^{(i)} \right|}{K} \tag{4-2}$$

The equation (4-2) shows the mean distance of activity l. And K is the number of activity l in training data. The similarity score of activity l is the reciprocal value of the mean distance.

$$score_l(X, Y_1, Y_2, \ldots, Y_K) = \frac{1}{dist_l(X, Y_1, Y_2, \ldots, Y_K) + 1} \tag{4-3}$$

w~~W~~hen the similarity scores of top 3~~ top~~ similar activities are all lower than threshold $T = 0.75$, the current activity are considering as an unknown activity. Otherwise, the system considers the activity it the most similarity activity.

格式化: 縮排: 第一行: 0 公分

4.3.2 Case-based Reasoning Approach

Case-based reasoning (CBR) approach is a problem solving paradigm that in many respects is fundamentally different from other major AI approaches. Instead of relying solely on general knowledge of a problem domain, CBR is able to utilize the specific knowledge of previously experienced problem situations, also called cases. The underlying idea is the assumption that similar problems have similar solutions. Although there exists some traditional knowledge based system for healthcare, CBR still has several advantages over traditional knowledge based systems. CBR reduces the knowledge acquisition effort and less maintenance effort. It also improves over time and adapts to changes in environment.

In the proposed system, the problem means the service of mapping activity. While an abnormal activity detects, the system should send an alert message to user's caregiver. The mechanism of CBR is proposed in our system; even an unknown activity is detected, the system can immediately provide service for the user, *i.e.*, an unknown is detected and it is most similar to an abnormal activity. The system will send an alert message to

Table 4-1 The processes of case-based reasoning approach

Process	Description
RETRIEVE	Find the most similar case or cases
REUSE	The information and knowledge in that case to solve the problem
REVISE	Confirm the proposed solution
RETAIN	Update the dataset of cases and retrain the AR model to learn the current case

user's caregiver.

CBR application can be described by a cycle composed of four processes. The following Table 4-1 錯誤! 找不到參照來源。shows the processes of CBR.

The process "RETRIEVE" in the proposed system is used the highest similarity score to find the most similar activity. It considers the process is a 1-NN approach. In nearest neighbor retrieval, let CB denotes a set of input descriptions P for which a solution S exists, *i.e.*, (P, S) is in the case base. The similarity measurement function is a mapping similarity: $P \times \mathrm{CB} \rightarrow [0,1] \in \mathrm{R}$. The similarity score is computed by the equation (4-3). If the highest similarity score is lower than 0.75, it represents the activity is still too different to the most similar activity. So the activity is considered as a new activity.

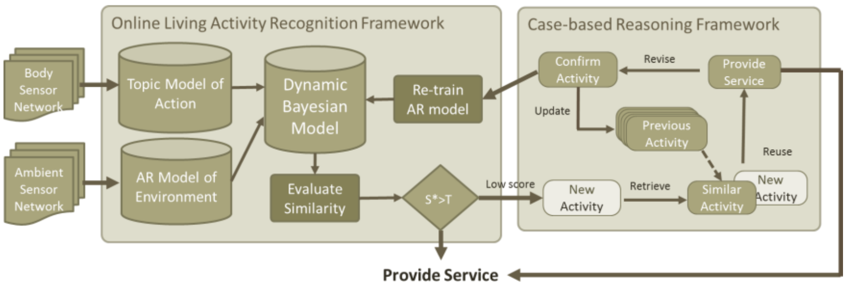


Fig. 4-7 The flowchart of online activity recognition and case-based reasoning

Chapter 5

System Evaluation

In this chapter, we will evaluate our recognition rate in a real environment at our home like lab and test the activity-aware healthcare system in a home simulator respectively. In order to get more realistic results, the simulated home environment will incorporate human behavior scenarios, which model a user's habit of appliance usage based on real-life scenarios, and scenarios are used to better reflect the behavior patterns of daily living. Therefore, we firstly describe the simulated home environment and the scenarios of activity of daily living. Next, we will present the evaluation of the proposed approach in terms of two factors: 1) the performance of activity recognition, 2) the ability of discovering new activity.

5.1 Experimental Environment

In order to simulate a real home environment, we built a simulated home where the layout is similar to a general home with multiple residents, as shown in Fig. 5-1. In the figure, the simulated home consists of 5 rooms: hallway, living room, kitchen, study room and bedroom. There are several electronic appliances attached current sensor and ambient sensors (lumen, temperature sensor, etc.) deployed in the home environment, and the power consumption of each appliance shows in Table 5-1. The power consumption of electronic appliances are have three modes: Off, Standby and On. Almost all appliances are usually connecting with electronic plugs, even the appliance are idle. When an appliance belongs to this status, it is on the mode of Standby. We have monitored the power consumption of all appliance in our simulated home environment, and Fig. 5-1 shows their power consumptions. For the proposed system, it inferences one appliance's usage mode by the mapping table.

The power consumption of electronic appliances are have three modes: Off, Standby and On. Almost all appliances are usually connecting with electronic plugs, even the appliance are idle. When an appliance belongs to this status, it is on the mode of Standby. We have monitored the power consumption of all appliance in our simulated home environment, and Table 5-1 shows their power consumptions. For the proposed system, it inferences one appliance's usage mode by the mapping table.

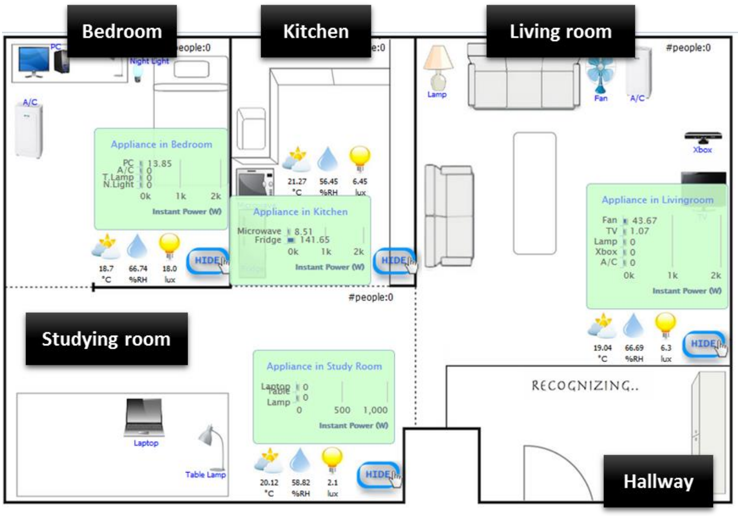


Fig. 5-1 The layout of the simulated home environment

We design 10 daily living activities which are commonly occurred in a real life, as illustrated in Table 5-2. The simulated home lacks bathroom, so it ignores activities in bathroom, *e.g.*, taking a bath, brushing tooth or using toilet, etc. Those activities contains both Posture and Motion actions. In the experiment, we only test single user for the home environment.

The user should wear a smart watch with operation system Android Wear and a mobile phone with Android 4.2. We have built a simple application for Android Wear to collect the sensor data of tri-acceleration and Pitch and Roll, and the application will send the sensor data to the mobile phone. Then, it sends the sensor data to the home server. Although the proposed system trains activity recognition model by unsupervised learning approaches, the application of mobile phone can label activity. We use those labels to evaluate the performance of activity recognition. And the application set the

Table 5-1 The power consumption of electronic appliances

App. Name	**Power Consumption**		**App. Name**	**Power Consumption**	
	Off/Standby	On		Off/Standby	On
Light	0.01	0.25	TV	0.09	1.3
Kinect	0.03	1.4	Fridge	0.1	2.8
Lamp	0.03	0.35	A/C	0.03	10
Fan	0.01	0.35	Microwave	0.03	9.3
PC	0.03	0.4	Water Heater	0.1	10

Table 5-2 Activity list in the simulated home

Location	**Activity**	**Location**	**Activity**
Living Room	Watch TV	Study Room	Read book
	Do exercise		Play pad
	Read newspaper		Sweep
	Meal	Kitchen	Wash dishes
Bedroom	Sleeping	Hallway	Go out

default label as “other”, so user does not need to label activity for realist life.

We have invited 3 subjects to test our system, and Table 5-3 shows their information. And the subject 2 did not do the activity “Go Out” in the experiment. The interval time of two continiuious instances is 5 seconds. So each subject has collected about 2 hours ADL data in our simulated home. In the experiment, they could do one kind of ADLs more than one times, i.e., the subject 1 did “play pad” two times in two hours experiment. According to the result of the subject 1, we will discuss the performance of training mode. And we will show the performances of online mode with all subjects.

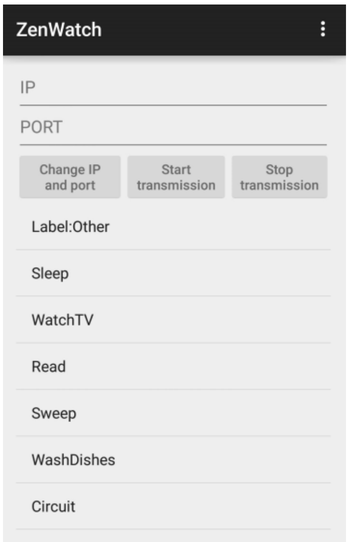


Fig. 5-2 The interface of mobile application with Android OS

Table 5-3 The information about experiment subjects

	Subject 1	Subject 2	Subject 3
Gender	Male	Male	Male
Age	23	61	21
ADL number	10 ADLs	9 ADLs	10 ADLs
Instances number	1351	1547	1193
Experiment time	113 minutes	130 minutes	101 minutes

5.2 Evaluation of Training Mode

In this thesis, the experiment is divided into two parts. The first part is the activity recognition result via group activity and the second is the ability of discovering new activity. In the part of assessment on activity recognition, we shows the results of two individual AR model: ambient part and vital sign part. After evaluating them individually, we fuse them into the training format of second layer non-parametric hierarchical activity recognition model. And we evaluate the performance of this fusion ambient and vital sign AR model.

5.2.1 Performance on Ambient Model

Some activities of ambient sensors are using same appliances or same state of environment. For example, the environment state between doing exercise and having meal are same. When people do exercise or have meal are usually staying in living room and only open the light. For the ambient AR model, it is hard to identify their different from its ambient sensors. So the evaluation of the part of ambient AR model only shows the observation from Boolean types of sensor data.

Table 5-4 shows light on/off state in each room and electronic appliances usage state, and each activity uses appliances are fewer than two appliances. We have collected about 2 hours ambient sensor data with 1349 instances, and the interval time between two successive instances is 5 seconds. And each instance is labeled by the mobile phone application when the user started doing the activity. Table 5-5 shows the predicted results of collected data with 10 activities and a label “Other” that the user consider the activity cannot belong to a specific activity.

Table 5-4 Light and appliance usage states of different activities

Activity	Light On/Off				Appliance On/Off					
	Living Room	Studying Room	Bedroom	Kitchen	Switch Door	TV	Lamp	Fan	Micro-wave	Pad
Watch TV	1	0	0	0	0	1	0	0	0	0
Read Newspaper	1	0	0	0	0	0	1	0	0	0
Exercise	1	0	0	0	0	0	0	0	0	0
Meal	1	0	0	0	0	0	0	0	0	0
Play Pad	0	1	0	0	0	0	0	0	0	1
Read Book	0	1	0	0	0	0	0	0	0	0
Sweep	0	1	0	0	0	0	0	0	0	0
Sleep	0	0	1	0	0	0	0	0	0	0
Wash Dish-es	0	0	0	1	0	0	0	0	0	0
Go Out	0	0	0	0	1	0	0	0	0	0

Table 5-5 Predicted Result of Ambient part AC model

	C1	C2	C3	C4	C5	C6	C7	C8
Watch TV	69	0	0	0	0	0	0	0
Read Newspaper	0	62	0	0	0	0	0	0
Exercise	0	0	114	0	0	0	5	0
Meal	0	0	220	0	0	0	5	0
Play Pad	0	0	0	133	4	0	0	0
Read Book	0	0	0	0	126	0	0	0
Sweep	2	0	0	0	63	1	0	0
Sleep	0	0	0	0	0	293	0	6
Wash Dishes	6	0	0	0	0	0	166	0
Go Out	0	0	0	0	0	0	0	59
Other	8	0	0	4	0	0	0	3

The ambient part AR model finds 8 clusters, and each cluster represents one to two activities. If two activities use same appliance and stays in same room, they are belonging to same cluster. For example, “Exercise” and “Meal” are both staying in living room and they are not using any electronic appliance, so they are belonging to the cluster “C3”. And this is the reason that only using ambient sensor is hard to identify various activities of daily living.

5.2.2 Performance on Vital Sign Model

To monitor the variation of vital sign data can extract some interesting results. Although we only use one wearable device to monitor user's vital sign all day long in the experiment, it can find different activity by extracted features from vital sign. Because the vital sign data are continuous, they are hard to use discriminated method to classify. We propose a 2 layer Dirichlet Process Mixture Model (2LDPMM) as the activity recognition model and it is a likely topic model. Topic model can find the topic of a document, and our 2 layer DPMM can find the meaningful activity of a sequence of vital sign data. Before building 2LDPMM, the feature extraction is processing from vital sign data (tri-acceleration, Pitch and Roll). We compute every second vital sign data to extract7 features from tri-acceleration and 4 features from Pitch and Roll. The features extracted from acceleration are mean of each tri-acceleration, variance of each tri-acceleration and signal vector magnitude. And the features extracted from Pitch and Roll are mean of Pitch, mean of Roll, mean of angle changes of Pitch and mean of angle changes of Roll.

The feature extraction is used to find the physical meaning and also used to reduce the quantity of data. And the other reason is that power consumption of wireless transmit is high, so we try to reduce the transmit time for monitoring vital sign data. The smart watch only needs to send collected vital sign data every second, so the power consumption of the smart watch can be reduced. When the server receives the vital sign data, it will extract features for training model or predicting activity. Fig. 5-3 illustrates the raw data values of Y-axis acceleration and its feature of mean of Y-axis acceleration about 10 second data of 4 activities: play pad, read, exercise and sweep. The 4 activities includes 2 ”Posture” (“Play pad” and “Read”) and 2 ”Motion” (“Exercise” and

"Sweep"). We find that activities of Motion are more significant, but activities of Posture are not. The value between play pad and read are similar, and it may make the unsupervised learning algorithm have a bad performance. So we consider other physical meaning that is Pith and Roll to distinguish the activities of Posture.

Fig. 5-5 and Fig. 5-4 show the feature values of signal vector magnitude (SVM) and mean of Pitch in one minute. The features belong to acceleration can easily distinguish activities of Motion, *e.g.*, the feature SVM is used to represent the activity energy magnitude, and the value of SVM between "Exercise" and "Sweep" are significant differences. But the value of SVM between "Read" and "Play pad" are similar.

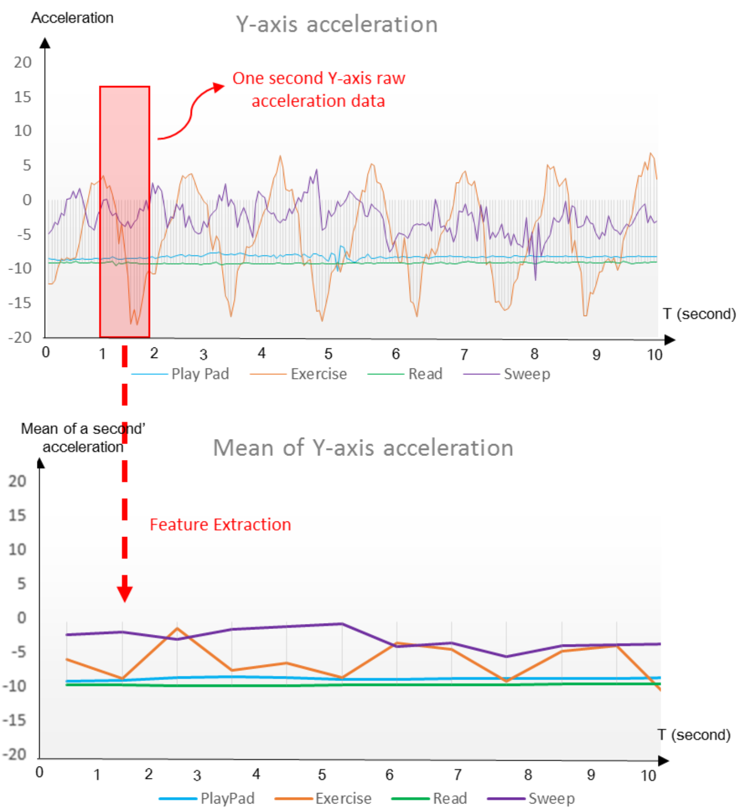


Fig. 5-3 The data value and features of Y-axis acceleration for four activities

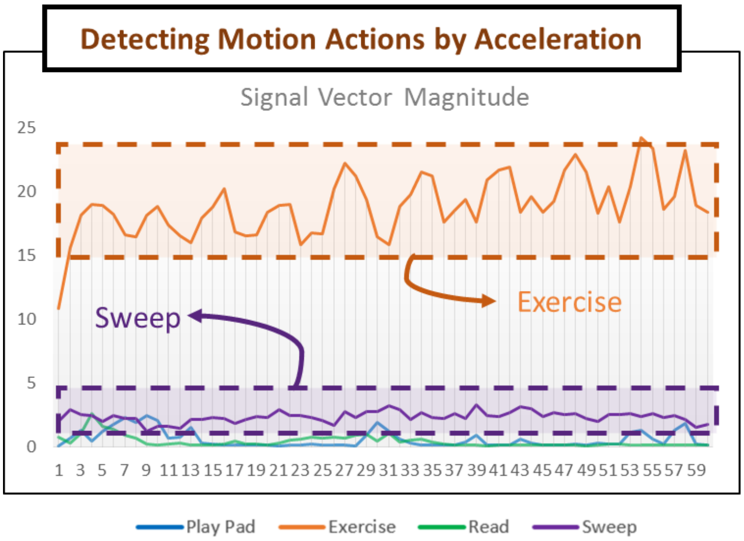


Fig. 5-5 The features of SVM for four activities

The features belong to orientation (Pitch and Roll) are used to distinguish activities of Posture because only acceleration is hard to determine them. The activities of Posture

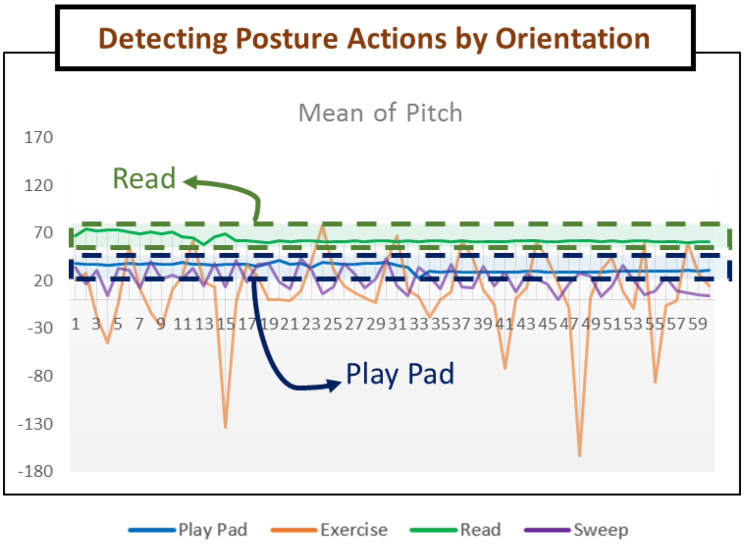


Fig. 5-4 The features of Mean of Pitch for four activities

are usually static and the weak hand of Posture activity is usually turning the face to a specific orientation. And different Posture activities have their own habits of the turning orientations, *e.g.*, the degree of mean of Pitch for "Play pad" is between $20°$ to $40°$; the degree of mean of Pitch for "Read" is between $60°$ to $70°$. This characteristic makes the data-driven method can distinguish them by their distribution of orientation. However, some activities are still hard to distinguish by the information of acceleration and orientation. Because we only ask user wear the smart watch on the wear hand, we can only monitor the variations of acceleration and orientations from user's weak hand. If two activities have same behavior of user's weak hand, we cannot distinguish them. In the experiment, the activities of "Read newspaper" and "Read Book" are same that both they are reading something in different rooms. And the behaviors between "Watch TV" and "Read" are similar, so the cluster C9 identifies three activities: "Watch TV" "Read newspaper" and "Read Book". And "Watch TV" sometimes is similar to "Play pad", so the cluster C8 identifies "Play pad" and "Watch TV".

Even though the vital sign part AR model can almost recognize both Posture and Motion activities, some activities are still hard to be correctly recognized. If we consider the appliance usage states or light on/off in different rooms, those ambiguous activities can be distinguished. We propose a method to fuse the results from ambient part AR model and vital sign part AR model, so the fused data can build a new AR model to correctly recognize those ambiguous activities.

2LDPMM is a likely topic model, so the function of 2LDPMM is similar to topic model. The results of 1st~~first~~ layer 2LDPMM represent the type of one second behavior of weak hand; and for the view point of topic model, each type can be considered as one vocabulary and its features are words. And we generate histogram of a sequence of weak hand's behaviors as the feature for 2nd~~second~~ layer 2LDPMM; for the view point

Table 5-6 Predicted Result of 2LDPMM

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16
Watch TV	0	0	0	7	0	0	0	21	41	0	0	0	0	0	0	0
Read Newspaper	0	2	0	2	0	0	0	0	54	4	0	0	0	0	0	0
Exercise	115	3	0	0	0	0	0	0	0	0	0	1	0	0	0	0
Meal	0	0	9	76	12	26	101	1	0	0	0	0	0	0	0	0
Play Pad	0	0	0	40	0	0	0	87	4	3	0	0	0	3	0	0
Read Book	0	2	0	9	0	0	0	0	108	3	0	0	0	0	0	0
Sweep	0	0	0	0	0	0	0	0	0	10	56	0	0	0	0	0
Sleep	0	0	0	0	0	0	0	0	0	0	0	13	286	0	0	0
Wash Dishes	0	0	4	0	0	0	0	0	0	0	0	0	0	4	164	0
Go Out	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	56
Other	0	1	2	2	0	0	0	1	0	5	0	3	0	1	0	0

Table 5-7 Predicted Result of hand's behavior in the ~~first~~ 1st layer 2LDPMM

格式化: 上標

	hb1	hb2	hb3	hb4	hb5	hb6	hb7	hb8	hb9	hb10	hb11	…	hb54	hb55	hb56
Watch TV	129	229	0	0	0	0	0	0	0	0	0	…	0	0	0
Read Newspaper	2	268	0	0	0	0	0	0	0	0	0	…	0	0	0
Exercise	0	0	0	151	19	87	67	13	235	25	13	…	0	0	0
Meal	184	78	2	0	0	0	0	0	0	0	0		2	124	170
Play Pad	541	103	0	0	0	0	0	0	0	0	0	…	0	8	5
Read Book	24	575	1	0	0	0	0	0	0	0	0	…	0	5	0
Sweep	0	1	1	0	0	0	0	0	0	0	0	…	0	1	0
Sleep	0	0	31	0	0	0	0	0	0	0	0	…	0	0	0
Wash Dishes	0	0	0	0	0	0	0	0	0	0	0	…	54	14	0
Go Out	0	0	1	0	0	0	0	0	0	0	0	…	0	1	0
Other	0	0	0	0	0	0	0	0	0	0	0		0	0	0

of topic model, this histogram can be considered as a document and the 2nd ~~second~~ layer 2LDPMM tries to find the topic of this document that one topic is mapping to a kind of activity. In the experiment, the ~~first~~ 1st layer 2LDPMM finds 56 clusters (the cluster denotes as "hb") and they represent different kinds of hand's behaviors.

格式化: 上標

格式化: 上標

Table 5-7 shows the predicted results from the ~~first~~ 1st layer 2LDPMM. There are 8 clusters (hb4 to hb11) belong to "Exercise" and only 2 clusters (hb1 and hb2) belong to

格式化: 上標

"Play pad". The hands' behavior of "Exercise" is various, so the types of "Exercise" is higher than other Posture activities. The hands' behavior of "Play pad" is relatively invariant than "Exercise", so the types of hand's behavior of "Play pad" are only be determined to 2 types. And the dominated activity of cluster hb2 is "Read" that is not "Play pad". We consider that hb2 represents the weak hand stays in a table and it may hold something (books, newspaper or pad). It obviously shows one cluster can represent one kind of hand's behaviors and different activities can have same hand's behavior in their completed behaviors.

Fig. 5-6 shows the result of the ~~second~~ 2nd layer 2LDPMM. The horizontal axis represents the cluster ID and the vertical axis represents the time and time interval of predicting input data is 5 seconds. The figure shows about 210 seconds successive input vital sign data. Most of clusters only contains one activity, e.g., cluster C4 contains "Go Out" and both clusters C2 and C3 contain "Sleep". "Sweep" can be represented to C5 and C6, and comparing to the real activity of doing sweep that C5 maps to sweep floor

格式化: 上標

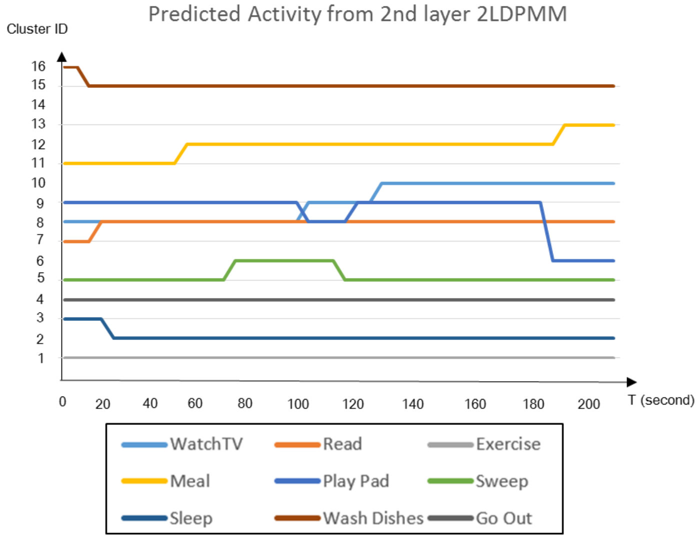


Fig. 5-6 The predicted results from the 2nd layer 2LDPMM

and C6 maps to take out the dusk from dustpan. These two different behaviors are categorized to different clusters. And the similar activities "Watch TV" and "Read" are overlapping to same cluster C8 for a long time. The second layer nonparametric hierarchical AR model can resolve the ambiguous clustering problem.

5.2.3 Performance on Fusion Ambient and Vital Sign Model

When we obtain the results from first layer nonparametric hierarchical AR model (N~~HARM~~HAC), we have get the predicted results of ambient and vital sign. Each instance is considered as the concatenation of 60 successive vital sign data, and the predicted result of vital sign data are the discrete data. Each predicted result represents an activity that the user does. Because the limitation of current sensor cannot sense environment in a very short interval time, the system snapshots the environment by ambient sensors every 5 seconds. The fused AR model also predicts user's daily living activities every 5 seconds.

The result of ambient AR model can represent the environment observation and the objective point of view to monitor user's activities. The result of vital sign model can represent the human behavior and the subjective of view to monitor user's activities. The 2nd layer N~~HARM~~HAC concatenates the results from ambient and vital sign AR models as the completed point of view to monitor user's activities. We compute different concatenated results occurs times from all results to find the time-frequency (TF) value of each type of concatenated result. When a TF value of a type of concatenated result is higher than 1%, this type of concatenated result becomes a cluster head. The other types of concatenated results, which TF value are lower than 1%, are used KNN approach to predict its cluster and add it into the cluster. For the KNN approach, it finds the distance between current instance and other instances by their features. If using the results from ambient and vital sign AR model for KNN approach, it cannot determine the difference between different activities. Since both the predicted results are different to ambient and vital sign AR models, the distance is fixed 2. And one predicted result is same but the other is different, then the distance is 1. We can only find the 2 different value of distances, the KNN is ineffective. So, we generate new features for KNN model, and the new features are extracting from ambient and vital sign part AR model. However, the time domain of features between ambient and vital sign are different. The time domain of ambient feature is the snapshot of ambient sensor's active status; and the time domain of vital sign feature is one minute successive hand's behaviors. It is hard to dismantle the features of vital sign AR model, so we transfer the data format of ambient predicted results. We consider the time domain of ambient AR model is same to hand's behaviors. So we build the histogram of one minute successive ambient predicted results as the new features. The occurred times of each cluster is multiple 5 to fill up the non-sensing data. For example, Table 5-8 shows the feature format from a sequence of activity "Watch TV".

For the vital sign AR model, the predicted results of first and second instances belong to cluster C4 and other instances in the example belong to cluster C9 (please see

Table 5-8 Feature of the 2nd layer N~~HARM~~HAC for a sequence of activity "Watch TV"

Vital Sign AR: features of ~~first~~ 1st layer 2LDPMM									Ambient AR: the successive predicted result for one minutes					
hb1	hb2	…	hb23	…	hb38	…	hb55	hb56	C1	C2	…	C5	…	C8
17	0	..	40	..	1	..	0	1	0	0	..	60	..	0
13	0	…	45	…	1	…	0	0	0	0	…	60	…	0
8	0	…	50	…	1	…	0	0	0	0	…	60	…	0
3	0	…	55	…	1	…	0	0	0	0	…	60	…	0
0	0	…	60	…	0	…	0	0	0	0	…	60	…	0
0	0	…	60	…	0	…	0	0	0	0	…	60	…	0
0	0	…	60	…	0	…	0	0	0	0	…	60	…	0

格式化: 上標

the Table 5-6). The concatenated results of first and second instances are “4,5”; other are “9,5”. The TF value of “9,5” is higher than 1%, so these kind of instances belong to one cluster. And the TF value of “4,5” is lower than 1%, so the first and second instances need to use KNN approach to find the most similar activity. Table 5-9 shows the clustering results from ambient and vital sign data. It observes 14 clusters, and each cluster represents a categorized training data. Each categorized training data is mapping to one daily living activity.

The system generates the labeling interface that each cluster has one color. User needs to label each cluster belong to which daily living activity. Fig. 5-7~~Fig. 5-7~~ illustrates the labeling interface and the mapping activities are labeling in the figure.

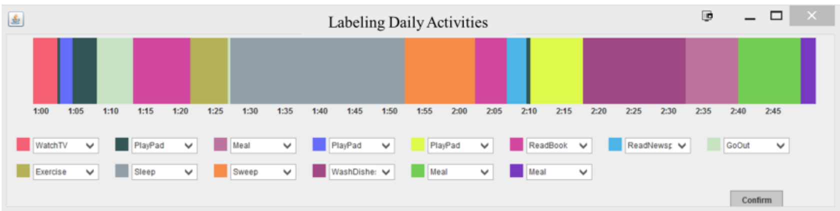


Fig. 5-7 The labeling interface of categorized training data

Table 5-9 Predicted result of daily living activity in the 2nd layer N~~HARM~~HAC

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
Watch TV	21	42	0	0	0	0	0	0	6	0	0	0	0	0
Read Newspaper	0	0	62	0	0	0	0	0	2	0	0	0	0	0
Exercise	0	0	0	119	0	0	0	0	0	0	0	0	0	0
Meal	0	0	0	0	26	107	90	0	0	0	0	0	2	0
Play Pad	0	0	0	0	0	0	0	89	45	3	0	0	0	0
Read Book	0	0	0	0	0	0	0	0	2	124	0	0	0	0
Sweep	0	0	0	0	0	0	0	0	0	0	64	0	0	2
Sleep	0	0	0	0	0	0	0	0	0	0	0	297	0	2
Wash Dishes	0	0	0	0	0	0	0	0	0	0	0	0	172	0
Go Out	0	0	0	0	0	0	0	0	0	0	0	0	0	59
Other	1	0	0	2	0	0	0	1	3	0	1	1	2	4

Table 5-10 the accuracy of the results from the proposed N~~HARM~~HAC for the subject 1

Activity	Watch TV	Read Newspaper	Exercise	Meal	Play Pad	Read Book	Sweep	Sleep	Wash Dishes	Go Out
Accuracy	0.9844	1	0.9834	1	0.9054	0.9764	0.9846	0.9966	0.9773	0.9403
Average Accuracy	0.974846									

Table 5-11 The number of clusters are found int the N~~HARM~~HAC for all subjects

	Subject 1	Subject 2	Subject 3

We compute the accuracy of each activity that considers the dominated clusters as the activity, so the different activities that belong to the dominated cluster are considered to false positives.The average accuracy is up to 97.48%. It examines that we can use the categorized clusters for user to label data and the mapping service to complete the healthcare system. The results of N~~HARM~~HAC for the subject 2 and the subject 3 shows in Table 5-11. Each number of clusters of the 2nd layer N~~HARM~~HAC is 14, 13 greatly reduces the burden of labeling instances.

Ambient observations	8	7	8
Hand's behaviors	56	43	21
Vital sign observations	16	15	17
ADLs (fusing ambient and vital sign)	14	13	17
Accuracy	97.485%	98.707%	97.049%

5.3 Evaluation of Online Mode

5.3.1 Performance of Discovering New Activity

The proposed system in online mode is not only recognizing activity, but also able to discover new activity from the input instance. This paragraph describes the performance of discovering new activity by the similarity function. We demonstrates the similarity function by testing some daily activities that are not in training data. The new activity is the other form of "Play pad" that user play pad on bed in the bedroom. We find that the input new activity is most similar to "Play Pad". And this activity will be a temporal new cluster. So, if other activities are similar to this activity, they will belong to this new cluster. *e.g.*, the successive input instance are usually the same activity, so the mechanism make sure the system will not generate more than two new clusters. And because the service of "Play Pad" is do nothing, the system will not send alert message to user's caregiver. The system will ask user to confirm this new activity belonging to "Play Pad" or not, and adding this data into the dataset. The online AR model will re-train with the incremental dataset.

Table 5-12 the distances between new activity and exist activities in the dataset

	Distance				Score	
	Hand behavior	Light	Appliance	Total dist	Score	Normalize Score
Watch TV	48.21	59.325	119.44	226.98	0.004386	0.308367
Read Newspaper	16.33	58.66	114.46	189.45	0.005251	0.369126
Exercise	56.63	59.61	58.64	174.88	0.005686	0.399704
Meal	44.48	58.97	59.56	163.01	0.006097	0.428632
Play Pad	10.28	58.12	0.9	69.3	0.014225	1

Read Book	14.73	58.66	58.21	131.6	0.007541	0.530166
Sweep	59.77	59.21	59.33	178.31	0.005577	0.392058
Sleep	59.61	56.25	59.97	175.83	0.005655	0.397557
Wash Dishes	52.24	57.84	59.29	169.37	0.00587	0.412631
Go Out	54.87	59.12	58.95	172.94	0.005749	0.404162

Fig. 5-8 The graph of adaptive activity recognition for each activity

格式化: 間距 套用前: 1 行, 套用後: 0.5 行, 行距: 單行間距

Fig. 5-8 illustrates the similarity score for adaptive activity recognition. Because the similarity function can find the most similar activity, we use the most similar activity to implement the case-based reasoning for the proposed healthcare system.

5.3.2 Performance of Online Activity Recognition Model

In the training mode, the system categorized training data into a number of clusters, and users were labeling each cluster to an activity. The labeling data stored in a dataset in order to train the online AR model. The online AR model is a dynamic Bayesian network (DBN) classifier. To assess the results of the DBN, we adopt 10-folds cross-validation. The dataset has been separated to 10 parts. The system concatenates 9 parts as training data, and the rest part becomes the testing data. And the system will build 10 models for different constitutes datasets. The following formulate is the definition of recall, precision and F-measure:

$$\text{Recall} = \frac{TruePositive}{(TruePositive + FalseNegative)} \tag{5-1}$$

$$\text{Precision} = \frac{TruePositive}{(TruePositive + FalsePositive)} \quad (5\text{-}2)$$

$$\text{F1 measure} = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (5\text{-}3)$$

For the online AR model, the precision of each subject is 97.8%, 98.7% and 96.5% and the recall is 97.8%, 99.4% and 97.6% and the F1-measure is also 97.8%, 99.0% and 97.0%. The F1-measure is used to balance the contribution of precision and recall. Table 5-13 shows the DBN result of subject 1 and only the activity "Play Pad" has a worse performance on recall (84.6%), other activities have both precision and recall higher than 90%. The number of instances of "Play Pad" is 78, and only 66 instances are true positive. There are 8 instances of "Play Pad" are predicted as "Read Newspaper", so the recall is descending. But the overall performance is up to 97.8%. The subject 2 had not done the activity of "Go out". The overall performance of subject is better thanothers. And the precision of each activity of the last subject is good, but the recall of "Read Book" is only 83.1%. Some instances of "Read Book" are predicted as "Play Pad" or "Meal", because the subject 3 had a meal before reading book and played pad after reading book. DBN predicts activity by the current input data and the previous state, so the previous state may cause some mistakes. But the overall precision and recall for the subject 3 is still good. The average precision of all subjects is up to 97.67%. It demonstrates the activity-aware healthcare system is able to predict and recognize activity in real-time. And the evaluation of discovering activity shows the system has adaptive activity recognition mechanism by the appropriated similarity function and case-based reasoning approach.

Table 5-13 the activity recognition rate of online AR model of subject 1

Subject 1							
Activity	Precision	Recall	F1-measure	Activity	Precision	Recall	F1-measure
Watch TV	100%	97.6%	98.8%	Go out	91.5%	97%	94.2%
Play Pad	95.7%	84.6%	89.8%	Exercise	97%	100%	98.5%

Meal	100%	99.6%	99.8%	Sweep	97.6%	99.2%	98.4%
Read Book	98.4%	97.4%	97.9%	Sleep	100%	99%	99.5%
Read Newspaper	91.4%	94.4%	92.9%	Wash Dishes	97.2%	100%	98.6%

Table 5-14 the activity recognition rate of online AR model of subject 2

Subject 2							
Activity	Precision	Recall	F1-measure	Activity	Precision	Recall	F1-measure
Watch TV	99.4%	98.9%	99.3%	Go out	-	-	-
Play Pad	99.7%	99.2%	99.5%	Exercise	95.2%	100%	97.6%
Meal	97.7%	100%	98.8%	Sweep	99.4%	98.9%	99.2%
Read Book	100%	100%	100%	Sleep	99.0%	99.8%	99.4%
Read Newspaper	91.7%	100%	95.7%	Wash Dishes	97.6%	100%	98.8%

Table 5-15 the activity recognition rate of online AR model of subject 3

Subject 3							
Activity	Precision	Recall	F1-measure	Activity	Precision	Recall	F1-measure
Watch TV	99.3%	100%	99.6%	Go out	87.1%	100%	93.1%
Play Pad	88.9%	99.0%	93.7%	Exercise	95.9%	100%	97.9%
Meal	94.6%	99.0%	96.7%	Sweep	98.1%	100%	99.0%
Read Book	100%	83.1%	90.8%	Sleep	100%	100%	100%
Read Newspaper	97.3%	100%	98.6%	Wash Dishes	96.9%	100%	98.4%

Chapter 6

Conclusion

6.1 Summary

We hereby proposed a healthcare system to monitor the activities of daily living for elders in their home environments. The material sensors for the model are roughly divided into two groups: one is used to monitor the environment called ambient sensors, whereas the other is used to monitor the human body called body sensors. The monitoring model has fused these two different types of sensors by using a hierarchical activity recognition model. The first layer of the activity recognition model is built by two separated AR model, each is KNN for ambient sensors and a topic model: two Layer Dirichlet Process Mixture Model (2LDPMM) for body sensors. For 2LDPMM, the first layer results represent some meaningful actions, such as sitting, standing or swinging hands; the second layer of the activity recognition model shows the results of living activity, *e.g.*, watching TV, sweeping. And we fuse the results of KNN and 2LDPMM to build an activity recognition mode, such as meal, reading, sleeping or sweeping. In this stage, some activities that is hard to recognize by only ambient sensor or only body sensors are able to be identify. Our activity monitoring model is more appropriate for real living environment because our activity recognition model does not need to specify priory the number of the cluster in order to train the model. The experiment has demonstrated the system's ability to build in real living environment. It can recognize more activities using both ambient and body sensors than using only one type of sensor and the accuracy is up to 97.48%. The activity-aware system is not only recognizing those learned activities, but also has ability to discovery unseen activity automatically and retrains the AR model. Even the elderly people has new life style that is caused by aging, the activity-aware system can monitor and learn those new behaviors automatically. That the proposed activity of daily living-aware elderly healthcare system is feasible to monitor elders' living activity in a smart home environment for the realization of the concept: "Aging in Place".

6.2 Future Work

For the proposed healthcare system can be done to improve in the future. Some of which are listed below:

- **A more friendly interface for elderly user**

The labeling interface is roughly in the work. It can be developed on mobile pad as App application. And we can invite some elderly people use our system and give some feedbacks. We can base on those feedbacks to improve the labeling interface. And for the case-based reasoning function, we also built an interface for given alert message and ask user to identify the unknown activity. This interface also needs to be improve, and we can give some questionnaire for caregivers to realize how to design an appropriated interface of alert message.

- **Developing more applications based on the activity-aware system**

In our system, the service is simple that we only give alert message when the system monitors the odd activity that labeled by user. There are more useful applications can imply in the smart healthcare environment based on the real-time monitoring user's activity. For example, an automatically detecting anomaly activity system or automatically reminding calendar system for dementia people.

REFERENCE

[1] H. Banaee, M. U. Ahmed, and A. Loutfi, "Data mining for wearable sensors in health monitoring systems: a review of recent trends and challenges," *Sensors*, vol. 13, pp. 17472-17500, 2013.

[2] A. Pantelopoulos and N. G. Bourbakis, "A survey on wearable sensor-based systems for health monitoring and prognosis," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 40, pp. 1-12, 2010.

[3] C. Collin, D. Wade, S. Davies, and V. Horne, "The Barthel ADL Index: a reliability study," *Disability & Rehabilitation*, vol. 10, pp. 61-63, 1988.

[4] G. D. Abowd and E. D. Mynatt, "Charting past, present, and future research in ubiquitous computing," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 7, pp. 29-58, 2000.

[5] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *Communications Surveys & Tutorials, IEEE*, vol. 15, pp. 1192-1209, 2013.

[6] A. Kailas and M. A. Ingram, "Wireless communications technology in telehealth systems," in *Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology, 2009. Wireless VITAE 2009. 1st International Conference on*, 2009, pp. 926-930.

[7] H. Alemdar and C. Ersoy, "Wireless sensor networks for healthcare: A survey," *Computer Networks,* vol. 54, pp. 2688-2710, 2010.

[8] E. M. Tapia, S. S. Intille, and K. Larson, *Activity recognition in the home using simple and ubiquitous sensors*: Springer, 2004.

[9] D. Sanchez, M. Tentori, and J. Favela, "Hidden markov models for activity recognition in ambient intelligence environments," in *Current Trends in Computer Science, 2007. ENC 2007. Eighth Mexican International Conference on*, 2007, pp. 33-40.

[10] Q. Zhang, M. Karunanithi, R. Rana, and J. Liu, "Determination of Activities of Daily Living of independent living older people using environmentally placed sensors," in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, 2013, pp. 7044-7047.

[11] F.-T. Sun, Y.-T. Yeh, H.-T. Cheng, C.-C. Kuo, and M. Griss, "Nonparametric discovery of human routines from sensor data," in *Pervasive Computing and Communications (PerCom), 2014 IEEE International Conference on*, 2014, pp. 11-19.

[12] D. J. Cook, N. C. Krishnan, and P. Rashidi, "Activity discovery and activity recognition: A new partnership," *Cybernetics, IEEE Transactions on,* vol. 43, pp. 820-828, 2013.

[13] H.-T. Cheng, M. Griss, P. Davis, J. Li, and D. You, "Towards zero-shot learning for human activity recognition using semantic attribute sequence model," in *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, 2013, pp. 355-358.

[14] C.-H. Lu and L.-C. Fu, "Robust location-aware activity recognition using wireless sensor network in an attentive home," *Automation Science and Engineering, IEEE*

Transactions on, vol. 6, pp. 598-609, 2009.

[15] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine," in *Ambient assisted living and home care*, ed: Springer, 2012, pp. 216-223.

[16] P. Cottone, G. L. Re, G. Maida, and M. Morana, "Motion sensors for activity recognition in an ambient-intelligence scenario," in *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2013 IEEE International Conference on*, 2013, pp. 646-651.

[17] B. Yuan and J. Herbert, "Context-aware hybrid reasoning framework for pervasive healthcare," *Personal and ubiquitous computing,* vol. 18, pp. 865-881, 2014.

[18] J. D. McAuliffe, D. M. Blei, and M. I. Jordan, "Nonparametric empirical Bayes for the Dirichlet process mixture model," *Statistics and Computing,* vol. 16, pp. 5-14, 2006.

[19] U. Hansmann, *Pervasive computing: The mobile world*: Springer Science & Business Media, 2003.

[20] M. Weiser, R. Gold, and J. S. Brown, "The origins of ubiquitous computing research at PARC in the late 1980s," *IBM systems journal,* vol. 38, pp. 693-696, 1999.

[21] C.-H. Lu, C.-L. Wu, T.-H. Yang, H.-W. Yeh, M.-Y. Weng, L.-C. Fu*, et al.*, "Energy-Responsive Aggregate Context for Energy Saving in a Multi-Resident Environment," *Automation Science and Engineering, IEEE Transactions on,* vol. 11, pp. 715-729, 2014.

[22] A. P. Pentland, *Wearable intelligence*: Scientific American, Incorporated, 1998.

[23] A. K. Jain, R. P. Duin, and J. Mao, "Statistical pattern recognition: A review,"

Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 22, pp. 4-37, 2000.

[24] T. Minka, "Estimating a Dirichlet distribution," ed: Technical report, MIT, 2000.

[25] J. M. Keller, M. R. Gray, and J. A. Givens, "A fuzzy k-nearest neighbor algorithm," *Systems, Man and Cybernetics, IEEE Transactions on*, pp. 580-585, 1985.

[26] H. Jegou, M. Douze, and C. Schmid, "Product quantization for nearest neighbor search," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, pp. 117-128, 2011.

[27] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, 2004, pp. 28-31.

[28] T. A. Stephenson, "An introduction to Bayesian network theory and usage," *IDIAP-RR 00-03*, 2000.

[29] Z. Zeng and Q. Ji, "Knowledge based activity recognition with dynamic bayesian network," in *Computer Vision–ECCV 2010*, ed: Springer, 2010, pp. 532-546.

[30] G.-C. Chen, C.-N. Huang, C.-Y. Chiang, C.-J. Hsieh, and C.-T. Chan, "A reliable fall detection system based on wearable sensor and signal magnitude area for elderly residents," in *Aging Friendly Technology for Health and Independence*, ed: Springer, 2010, pp. 267-270.

[31] C. E. Rasmussen, "The infinite Gaussian mixture model," in *NIPS*, 1999, pp. 554-560.

[32] S. G. Walker, "Sampling the Dirichlet mixture model with slices," *Communications in Statistics—Simulation and Computation®* , vol. 36, pp. 45-54, 2007.

[33] D. Griffiths and M. Tenenbaum, "Hierarchical topic models and the nested Chinese restaurant process," *Advances in neural information processing systems*, vol. 16, p. 17, 2004.