Robust Location-Aware Activity Recognition Using Wireless Sensor Network in an Attentive Home

Ching-Hu Lu, *Student Member, IEEE*, and Li-Chen Fu, *Fellow, IEEE*

***Abstract*—This paper presents a robust location-aware activity recognition approach for establishing ambient intelligence applications in a smart home. With observations from a variety of multimodal and unobtrusive wireless sensors seamlessly integrated into ambient-intelligence compliant objects (AICOs), the approach infers a single resident's interleaved activities by utilizing a generalized and enhanced Bayesian Network fusion engine with inputs from a set of the most informative features. These features are collected by ranking their usefulness in estimating activities of interest. Additionally, each feature reckons its corresponding reliability to control its contribution in cases of possible device failure, therefore making the system more tolerant to inevitable device failure or interference commonly encountered in a wireless sensor network, and thus improving overall robustness. This work is part of an interdisciplinary Attentive Home pilot project with the goal of fulfilling real human needs by utilizing context-aware attentive services. We have also created a novel application called "Activity Map" to graphically display ambient-intelligence-related contextual information gathered from both humans and the environment in a more convenient and user-accessible way. All experiments were conducted in an instrumented living lab and their results demonstrate the effectiveness of the system.**

***Note to Practitioners*—This system aims to achieve non-obtrusive and location-aware activity recognition, and the authors have successfully prototyped several AICOs to naturally collect interactions from residents or status from the environment. In addition, these AICOs have the potential to be commercialized in the future due to practicability and near-term advances in embedded systems. Furthermore, other potential advantages of an AICO lie in its applicability to other domains beyond just the home environment. Our initial work has yielded high overall accuracy, therefore, suggesting that it is a feasible approach that may lead to practical ambient intelligent applications (such as the Activity Map in this work). The limitations are that some currently available sensors cannot measure specific desired observations, or, in some cases, require users to carry them to operate.**

***Index Terms*—Location-aware activity recognition, attentive home, ambient-intelligence compliant object (AICO), wireless sensor network.**

Manuscript received February 03, 2008. First published July 10, 2009; current version published September 30, 2009. This paper was recommended for publication by Associate Editor A. Heil and Editor M. Wang upon evaluation of the reviewers' comments. This work is supported in part by the National Science Council of Taiwan, R.O.C., under Grant NSC96-2752-E-002-007-PAE and in part by the Excellent Research Projects of National Taiwan University, 95R0062-AE00-05.

C.-H. Lu is with the National Taiwan University, Taipei 106, Taiwan (e-mail: jhluh@ieee.org).

L.-C. Fu is with the Department of Computer Science and Information Engineering and the Department of Electrical Engineering, National Taiwan University, Taipei 106, Taiwan (e-mail: lichen@ntu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASE.2009.2021981

I. Introduction

Wireless SENSOR NETWORK (WSN) techniques exhibit a promising future for smart homes aiming to fulfill the vision of ambient intelligence. To achieve the true promise and potential of ambient intelligence in the smart homes, resultant systems or designs should be responsive, sensitive, interconnected, contextualized, transparent, and intelligent [1]. The building blocks for such intelligent systems must be economical enough to manufacture, deploy and maintain, thus making them widely acceptable to the public. Recognizing high-level contexts (such as residents' current locations and ongoing activities) is one of the essential bases for realization of abundant context-aware applications which foster many innovative and attentive services, and recent advances in embedded systems make real-time high-level context recognition from compact and pervasive devices possible. Thus, we believe that compact and versatile devices have great potential to retrieve the contexts and realize Weiser's vision in ubiquitous computing [2]. Here, he proposed the use of profound technologies to investigate how information technology can be widely diffused into everyday objects and how these technologies can lead to creative ways of supporting people's everyday lives. However, to achieve such potential, several practical issues [3] must be addressed when utilizing and deploying these resource-constrained and interference-prone devices in order to obtain long-term and reliable contexts in a typical home environment.

Although vision- and audio-related technologies are commonly used in human tracking and activity recognition, they are generally considered undesirable in a home environment due to potential privacy-violation, which compromises the desired transparency in a ubiquitous or pervasive environment [4], [5]. To be more practical, sensors in a smart home must be incapable of either directly identifying residents or capturing privacy-sensitive information. These devices must be capable of recognizing contexts of interest more naturally without directly revealing or compromising privacy.

To address above-noted concerns and to fulfill the requirements of real human needs, a multiyear interdisciplinary project, named "Attentive Home," (kicked off in 2006) is currently underway at National Taiwan University (NTU) with the goal of designing a practical, human-centric, attentive environment. Members of the project team comprise specialists from a variety of fields, such as computer science, sociology, electrical and civil engineering, psychology, mechanical engineering, etc. Additionally, we have been leveraging a variety of technologies to construct an open-standards-based platform (OSGi [6]) in order to integrate mobile devices, wireless sensor networks, home robots and other state-of-the-art technologies. Our "CoreLab," a living lab at NTU, has been set up to develop and test numerous feasible methodologies that might facilitate deployment of practical ambient intelligence applications that could soon be realized in the real world.

During the first-year of the project, we discovered that unobtrusive and location-aware activity recognition is one of the key enablers for many practical scenarios in an attentive home. This finding has motivated us to aim for taking advantage of a minimum number of unobtrusive sensors to perform accurate location-aware activity recognition without requiring occupants to wear any sensor devices. This design provides a more accommodating, natural, and pleasant living experience for residents.

The task of accurate and reliable activity recognition is very interesting and challenging due to a variety of factors. First, time variance is highly unpredictable. For example, the time that a resident spends watching TV could range from less than thirty minutes to a couple of hours during weekdays, or, might even stretch to four hours on weekends. Second, activity composition varies in sequences and hierarchies. For instance, watching TV might start with turning on the TV first via a remote control or it might, instead, begin by pressing a button on the front panel. It is likely then followed by flipping through different channels (via the remote control or the buttons on the panel) or could simply be followed by the resident sitting through an entire program. At last, the activity could end with the resident switching off the TV consciously or might, instead, end with the resident falling asleep. Moreover, a complex activity often comprises many related sub-activities which residents might perform in arbitrary order. Third, interleaved (or interrupted) activities are common in our activities of daily living (ADL), and residents likely will need to pause their ongoing activities and may or may not resume them afterwards. Fourth, the system becomes much more complicated when the task of accurate activity recognition involves a multiresident environment. Finally, interference or uncertainties from error-prone sensors pose additional challenges to the task.

In this study, rather than considering all challenges for multiple residents, we focus primarily on detecting interleaved activities and whereabouts for a single person based on reliable features extracted from error-prone devices. The reliable features also allow us to effectively tackle the challenges of both time variance and activity composition under uncertainties.

II. Related Work

Human activity recognition and tracking are among the fundamental issues in an ambient intelligence environment, and many recent studies have focused on solving these crucial problems via a variety of sensors. Compared with much of the vision-based activity recognition research using Bayesian Network ([21]–[23]) or HMM (Hidden Markov Model [8], [24]–[26]), less attention has been paid to wireless sensor networks deployed in living labs [27]. Table I selectively lists some related research using various sensors along with the methods of their deployment to achieve the goal of accurate activity recognition.

TABLE I
SUMMARY TABLE OF ACTIVITY RECOGNITION USING VARIOUS SENSORS AND SENSOR DEPLOYMENT

Related research	Sensor used	Sensor deployment
Intelligent Workspace [7], Smart Homes [8]	camera	in a living lab or outdoors
Xpod [9], Probabilistic Adaptive Computing [10]	wearable sensor	on human body
The Aware Home [11]	Binary/analog sensor	in a living lab, or medical centers
Long-Term Care [12], Intelligent kitchen[13]	RFID	on human body or on objects
Smart House [14]	ultrasound	on objects in an instrumented house
Long-Term Care [15], Activity Recognition [10]	GPS	on human body
PlaceLab [5], House_n [16]	wireless sensor	in a living lab
e-Motion [17]	laser scanner	in a living lab
House_n [18]	accelerometer	on human body
Information Percolator [19] [20]	microphone and others	in a living lab (or a real house)

Tapia *et al.* [5], [16] have shown that raw data from simple binary sensors along with a simple inference mechanism have solid potential for solving activity recognition problems in regular home settings; however, the stability of resultant accuracies among various activities needs further improvement. We utilize multimodal sensors to generate many informative features with small intraclass variation to improve overall accuracy. To improve overall robustness, we also calculate real-time reliability factors to detect possible malfunction commonly encountered in resource-constrained wireless devices.

Wilson *et al.* [28] have demonstrated that multiple residents' current locations can assist with activity recognition and vice versa; however, their work provides coarse-grained results merely showing which rooms the residents are in (room-level) and whether they are currently moving or not (binary motion type). This has motivated us to establish location-aware activity recognition for a single user and, simultaneously, to increase the resulting granularity by providing more fine-grained outcomes. We also propose a flexible architecture such that our current work may easily be extended and applied to multiple resident environments.

Lester *et al.* [29] have utilized a hybrid inference model to demonstrate that multimodal wearable sensors can generate useful features to unobtrusively distinguish activities; however, residents must continuously wear sensors or detection devices, which can be quite unwieldy and inconvenient ([30]–[32]), particularly for the cognitively impaired elderly or in an environment targeting maximal comfort. Furthermore, annoying issues caused by frequent and necessary battery replacement will eventually deter residents from using such methods. Our approach embeds sensors throughout the lab and does not require residents to wear or carry sensors all the time.

As for interleaved/interrupted activity recognition, most previous works directly omitted this possibility in order to substantially decrease overall complexity. Philipose *et al.* [12] and Patterson *et al.* [30] have taken advantage of radio frequency identification (RFID) to detect such interleaved events by associating a set of detected objects (attached with RFID tags) with an activity. Nevertheless, wearing RFID gloves all the time to

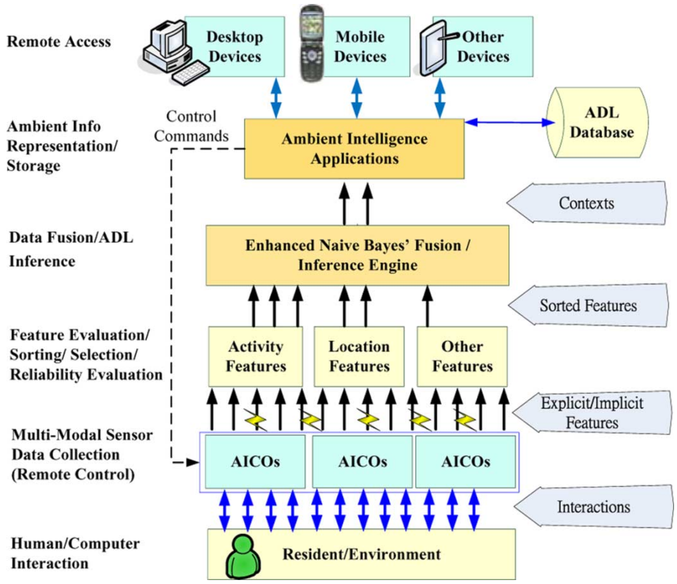


Fig. 1. Hierarchical system architecture for location-aware activity recognition in an attentive home.

detect objects and requiring that all objects in the environment be tagged is currently not feasible in our living space; moreover, it is also contrary to our goal of designing a comfortable and natural living environment. These considerations motivate us to detect interleaved/interrupted activities by generating or utilizing more informative features without requiring residents to wear sensors all the time.

To achieve the above-noted improvements upon previous works, the primary focus of this paper is on inferring a single resident's interleaved activities assisted by his or her location context. This work also fuses informative features extracted from multimodal wireless devices instrumented in a regular home rather than on the residents, themselves.

III. The Proposed Approach

A. System Overview

Fig. 1 illustrates the multiple-layered hierarchical architecture for location-aware activity recognition in our attentive home system. Starting from the lowest layer, instead of directly using simple sensors [5], [33], we propose using ambient-intelligence compliant objects (AICOs) to collect interactions from residents and status from the environment. The design of an AICO follows both general ambient intelligence visions and Weiser's vision; that is, an AICO is a building block for ambient intelligence applications and, meanwhile, facilitates weaving profound technologies [34] into the fabric of everyday life. As for implementation, an AICO is an ordinary household object (or a virtual object) overlaid by a virtual layer, where this additional virtual layer will capture interactions without interfering with natural manipulation of the object from residents. The usefulness of an AICO also lies in its augmented abilities to naturally capture the contexts or the interactions such that they become reliable feature sources to the upper layers and to actuate devices to stimulate more interactions via interoperability among AICOs.

AICOs can generate various features based on the captured data. We propose to divide the features into those which are explicit (in the time domain) and those which are implicit (in the frequency or other domains) in order to explore them separately. Since both activity and location features can work together in a reciprocal way [28], AICOs currently focus on generating explicit features for estimating a resident's activities, which will be further improved by his or her location information. Since highly discriminative features are more resistant to uncertainties and insensitive to interference inherent in battery-powered wireless devices, the system needs to concentrate on these useful features to reduce the overall computational burden. Additionally, the importance of more useful features is their potential for facilitating the feasibility of recognizing more complicated activities, especially in a multiple-resident environment.

After the extraction of various features, effective fusion of them can generate more reliable estimates. Fusion of multiple-source data has been proven effective in improving the accuracy of an estimate and we adopt feature level fusion such that the system is exempted from the responsibility of knowing the detailed implementations of the various multimodal sensors. This means the system can only concentrate on those features with better usefulness after sorting. Additionally, another advantage of feature-level fusion is that newly selected features (such as residents' IDs) can readily integrate into the system as long as they can help with the distinguishing tasks of location-aware activity recognition.

In order to deal with inferring activities under uncertainties, probabilistic reasoning offers an effective way to deal with unpredictable ambiguities from multiple sensors. Although [35] and [36] have demonstrated that naive Bayes classifiers work quite well in some domains with low variances of the classifiers in lieu of their strict independence assumptions among features and the classes, they did not consider overall efficiency and robustness of the classifiers, especially in a cluttered environment using error prone devices. We have implemented multiple naive Bayes classifiers, each of which represents an activity to be recognized; furthermore, we enhanced each classifier by incorporating both ranking features and reliability factors to detect interleaved activities and unexpected malfunction respectively. Since these classifiers benefit from not enforcing mutual exclusivity, they will not preclude the possibility of, for example, the detection of studying while listening to music. The inferred activities will be stored in a central ADL database as the source for later development of ambient intelligence applications.

Given the contexts stored in the ADL database, the system can represent real-time or historical activities in an innovative way. Furthermore, ambient intelligence and ubiquitous computing scenarios can be set up to demonstrate the usefulness of the technologies. How to effectively utilize the database and represent the stored data, however, becomes a critical issue, because good utilization of the information fosters the acceptance of technologies and provides helpful assistance to residents, caregivers or other users. For example, Bonanni *et al.* [2] provided awareness of water temperature by projecting a colored light, thus creating an intuitive ambient interface without distracting the resident's

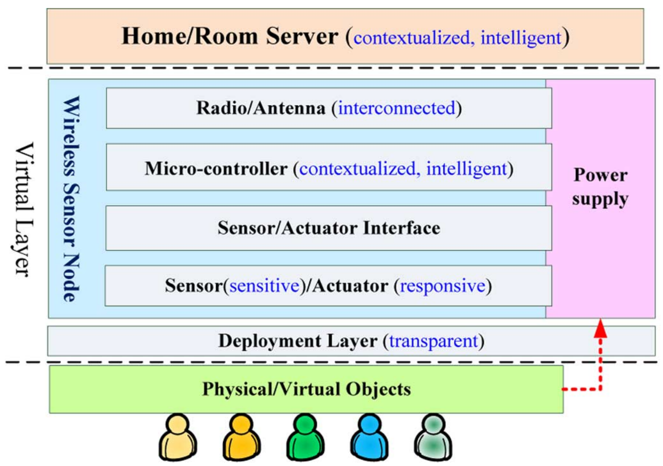


Fig. 2. Layered abstraction for an AICO. Functions enclosed in parentheses indicate component's characteristics for meeting the requirements of ambient intelligence.

current attention. Finally, remote access of the system provides information tailored to the capabilities of a client's device types.

One of the major advantages of this hierarchical architecture is its flexibility in replacing the mechanism on each layer whenever necessary. For instance, we can change the upper layer inference engine to a conditional random field [37], if necessary, to model undirected dependencies among features. In the future we could, for example, alter the fusion methods while continuing to use the sorted features generated from the AICOs on the lower layer.

B. Data Collection for Location-Aware Activity Recognition

1) AICO Design and Implementation: Fig. 2 shows the abstraction of our AICO design that can foster location-aware activity recognition. A wireless sensor node acts as part of the virtual layer described in the previous section; the node can sense changes from residents or the environment, and can also accept remote control signals to activate actuators, if necessary.

A sensor data collection can be triggered through manipulation of physical objects by residents (e.g., turning on a TV) or by some virtual mediums (e.g., measuring distance from a resident). Results from [33] show that residents may consider their houses as "dirty" when visible sensors are installed. This motivates us to establish an additional deployment layer and adopt "seamlessly embedded" strategy while designing the virtual layer in an AICO. Namely, we will try to seamlessly embed sensors/actuators with objects and hide them as much as possible during the design phase in order to avoid distracting residents' attention. This design can further ensure safety for children and the elderly, or can avoid unexpected damage from curious guests. The sensor/actuator interface can be customized by using a daughter board (with necessary interfaces like ADC, UART, SPI, etc.) to fit the connected sensors/actuators, thus further increasing overall flexibility.

The microcontroller can process raw data to extract higher level features or contexts to increase overall efficiency. It can also execute on-board self-diagnostic functions to reduce maintenance costs. It is preferred that the power supply for the virtual layer be derived from a stable source (e.g., using USB power directly or any energy-harvesting technology that can efficiently collect ambient sources of energy) to mitigate concerns regarding power consumption or battery replacement; however, the virtual layer can also be powered by batteries if necessary. Additionally, AICOs can further be programmed to cooperate with one another via various network topologies to accomplish diverse tasks. The features or contexts of interest will be transmitted wirelessly to a remote home/room server for higher level processing.

The design of an AICO includes three objectives. First, an AICO demands that the sensors/actuators should be seamlessly or inconspicuously integrated with a regular object such that residents can interact with the enhanced objects just as they did before. Second, smart home industries need more of a driving force to promote human-centric technologies so that better attentive applications can drive more demand. An AICO motivates embedding computing power into regular objects (or appliances) and can greatly increase their functionality and value in the market. Third, an AICO stresses the significance of good design in facilitating smart objects in a smart house application. From our ongoing "Attentive Home" project at NTU, we have learned that establishing a satisfactory attentive home goes beyond just using advanced technologies. Stronger interdisciplinary cooperation can perfect a smart home that humans will really desire. An AICO can be one of our first steps toward the goal of implementing human-centric applications. More and more regular and familiar objects will be retrofitted in our project in order to obtain contexts of interest in a more natural, inconspicuous, and elegant way.

2) Sensor Selection and Deployment for Activity Data Collection: Currently, we have successfully integrated various sensors with NTU "Taroko" wireless nodes [38] to generate informative features without revealing privacy information. The Taroko wireless node is a Tmote Sky [39] compatible wireless super node with routing ability, and it can transfer the preprocessed information wirelessly to a remote super node (also a Taroko) plugged into an OSGi-based home/room server. As of this point in time, we have successfully verified integration of various commonly used analog sensors [40] with the Taroko wireless nodes to detect current flow, voltage, pressure, vibration, motion, acceleration, distance, contact (via reed/mercury switches) along with IDs (from passive RFID tags attached on objects).

Utilizing domain knowledge, sensors/actuators in an AICO are selected and installed so that they can be triggered whenever activities of interest are performed. For electric-appliance-related activities, we have prototyped the virtual layer of a new mobile AICO for power monitoring/control (referred to from this point on as a "power-AICO"), as shown in Fig. 3. This device consists of a power socket and an outlet which can be used in a series connection with a regular appliance to measure power usage. In other words, the physical object that a power-AICO enhances refers to any regular electric appliance. We can instrument the power-AICO on the back/bottom of an appliance (such as a TV or a microwave) so that the connected appliance will also play a role as the deployment layer of the power-AICO. Furthermore, this power-AICO obtains its own power directly from a bundled regulator. With the integration of a passive RFID

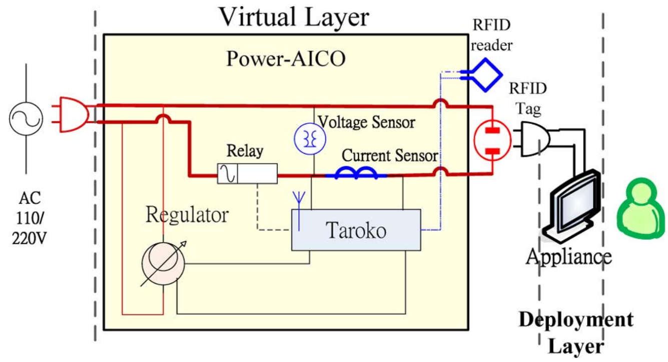


Fig. 3. The prototype of a power-AICO for monitoring power usage. It can be used in a series connection with any electric appliance (right) and a power source (left).

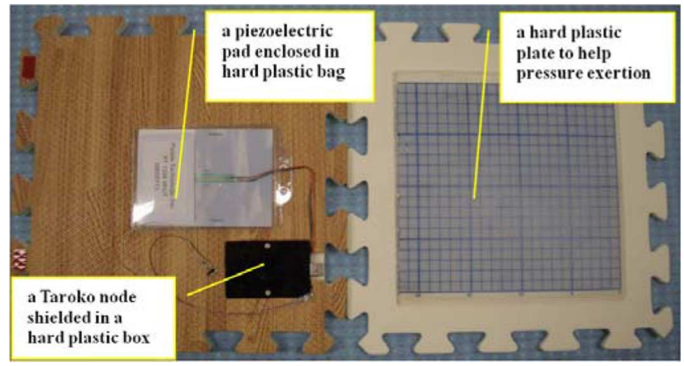


Fig. 4. A prototype of a floor-AICO (30×30 cm^2) and image of its inside.

reader, an appliance with an attached passive RFID tag can connect to the power-AICO so that any appliance-related activities that commonly occur in our ADL can successfully be recognized. We will soon add a relay to this device to enable remote control via a mobile device.

Therefore, the wireless node along with the sensors inside the power-AICO can enhance the functionality of a regular appliance, allowing it to become a building block for ambient intelligence applications, just like the "music bottles" demonstrated in [34]. Moreover, the residents will not notice the existence of a power-AICO; it will collect useful information silently (or transparently) without interfering with users' interactions or distracting their attention.

3) Indoor Human Location Data Acquisition: Since piezoelectric straps or pads can detect exerted pressure on the floor and can easily be located out of sight, they become good candidate sensors to embed in an AICO for human location detection. We have also prototyped a smart floor block (called a "floor-AICO") to detect residents' current locations for assisting activity recognition. Our system currently obtains high-level location information (e.g., in the living room or on a chair) by searching the IDs of a floor-AICO in a lookup table. In prototyping this AICO, we aim to design a cost-effective, accurate sensor device that is not only easy to deploy, but which can also be readily replaced after any malfunction; moreover, a floor-AICO can also lay the groundwork for later integration with other AICOs to provide location-aware services.

Fig. 4 depicts a floor-AICO containing one piezoelectric pad installed on the center of a mat with dimensions of approximately 30×30 cm ^{2}. The output voltage of the pad is roughly proportional to the input force exerted on the sensor and is converted to digital readings by an ADC bundled in the microcontroller on a Taroko.

The advantage of implementing floor-AICOs is threefold: 1) These blocks are easy and flexible to deploy. 2) Their accuracy and cost can be determined by the size of each mat as well as the total cost considerations of the floor. One could even deploy some dummy blocks, which have no sensors or wireless sensor nodes, under furniture or appliances; or, one could place floor-AICOs only in areas of interest like [7] to save further expense. 3) These floor-AICOs make various interactive or interoperable scenarios feasible. For example, the attentive home can automatically close the curtain if someone is studying beside a window and the sunlight is too strong (which could be detected by another AICO with an optical sensor).

Fig. 5 shows an overview of our instrumented CoreLab populated with a variety of wireless multimodal sensors to collect data for various ADLs during experiments. In actual deployment situations, we would only utilize those sensors considered helpful or essential after evaluating their usefulness in distinguishing activities of interest.

C. Invariant Feature Evaluation, Sorting, Selection, and Composition

Invariant features extracted from sensors are often commonly used in vision-based activity recognition (e.g., SIFT [41]). However, it is very challenging to generate invariant and highly discriminative features for effectively distinguishing human activities due to the time variance and the activity composition for different residents. This motivates us to establish an efficient mechanism to retrieve invariant features and evaluate their usefulness. As for extracting invariant features, good sensor selection is the most critical procedure to begin with. Since we will deploy only those sensors that are truly useful to detect the activities of interest, domain knowledge is necessary for sensor selection from the outset.

We have divided extracted features from multimodal sensors into explicit and implicit categories based on their requirements in computing resources. Explicit features are those demanding less computing power (such as mean, variance, area under curve, maximum, and minimum) and they are more intuitive to interpret in the time domain. By contrast, implicit features (e.g., fast Fourier transform (FFT) as listed in [42]) are more computation-demanding, nonintuitive and need in-depth understanding for system designers. Our goal is to generate all potentially useful features directly from AICOs; however, current processing power on an embedded processor is not sufficient to enumerate all useful features. Besides, we cannot completely rely on computers to automatically select sensors without help from domain knowledge.

Explicit features can assist humans more intuitively to select necessary sensors. This is why we are currently focusing on exploring explicit features and evaluating their usefulness. Exploration and evaluation of latent and useful implicit features will be performed in the next phase of our "Attentive Home" project.

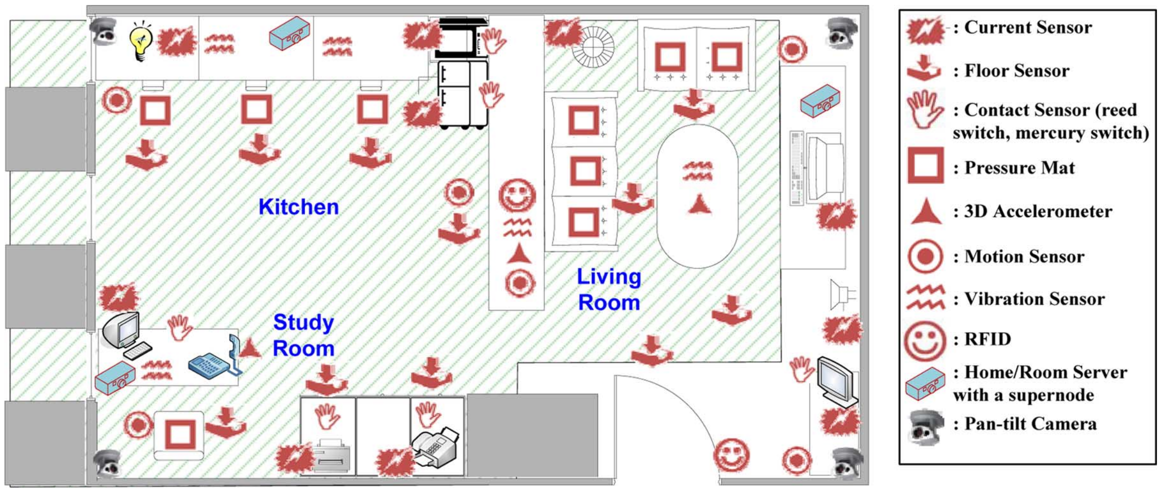


Fig. 5. Deployment overview in the CoreLab at NTU for experiment data collection. Each room has its own room server to serve as a data processing center. Note that the four cameras are only meant for collecting ground-truth for labeling the training data.

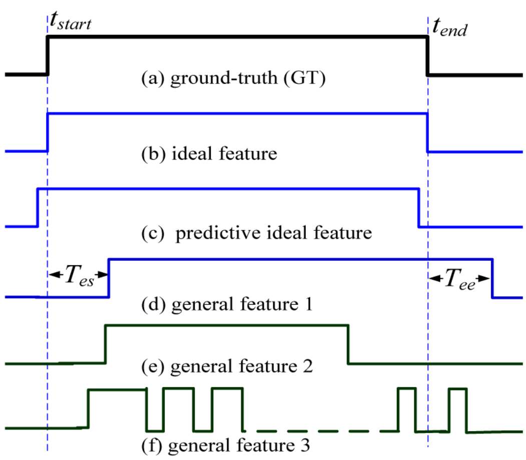


Fig. 6. An illustration of an activity and its simplified (or binarized) corresponding features in the time domain.

At that point, we will discuss how to make good use of both types of features to increase the overall recognition rate.

1) Time Invariant Feature Extraction and Evaluation: Fig. 6 illustrates various types of simplified (or binarized) explicit features in the time domain (each feature can be treated as an energy function). However, the actual explicit features are analog and probably biased. It is very challenging to generate an ideal feature as in Fig. 6(b), which exactly matches the duration of the original ground truth (GT) in Fig. 6(a), let alone a predictive one [Fig. 6(c), [43]]. More realistic features look like Fig. 6(d) with delayed detections of the activity. Fig. 6(e) and (f) illustrate some features that are very common in our ADLs. For example, the features from a mercury switch attached to a remote control to detect "watching TV" would resemble Fig. 6(f), since residents tend to flip through channels from time to time in an unpredictable manner.

All training samples consist of raw data and can be processed to obtain N possible useful features $\boldsymbol{f} = \{f_1, \ldots, f_N\}$ with a total of K activities of interest from $\boldsymbol{A} = \{A_1, \ldots, A_k, \ldots A_K\}$. Given a feature f_i, we propose two major properties to evaluate the usefulness in the time domain and decide a cutoff point as in [29] to eliminate less useful features.

The first property is the time invariance (Ψ) that is a characteristic indicating how a feature can adapt itself to completely fit the time duration to its corresponding GT in the time domain. Ψ is evaluated by cross correlation between the GT and a candidate feature f_i for an activity. The cross correlation originally refers to the measurement of similarity between two signals in the signal-processing domain by comparing an unknown signal to a known one. The higher the Ψ value, the closer a candidate feature to its corresponding GT. Ψ can be evaluated by (1)

$$\psi_{k,i} = \frac{1}{T_k} \max\left(\sum_{t=1}^{T_k - l_k} \mathrm{GT}_k[t + l_k] \cdot f_{k,i}[t]\right) \tag{1}$$

where $\Psi_{k,i}$ is the relative time invariance for feature i based on the GT of activity k. g_k is the GT function for activity k (whose duration is l_k) and $f_{k,i}$ is the unknown normalized feature (or a candidate feature function) to be evaluated. T_k is the maximum cross-correlation value from the training data as a factor for normalization.

The second factor is the detection sensitivity (Δ) which evaluates how small the early activity-start detection time T_{es} and early activity-end detection time T_{ee} can be (as in Fig. 6(d)) as shown in (2). We only evaluate the first T_{es} and the last T_{ee} within the duration of their corresponding GT for simplification

$$\Delta_{k,i} = \alpha \cdot \left[N_{\mu=0,\sigma^2_{T_{\mathrm{es}}}}\left(T^{k,i}_{\mathrm{es}}\right)\right] + (1-\alpha) \cdot \left[N_{\mu=0,\sigma^2_{T_{\mathrm{ee}}}}\left(T^{k,i}_{\mathrm{ee}}\right)\right] \tag{2}$$

where α is the weight. $\Delta_{k,i}$ is approximated with a weighted sum of two Gaussian distributions (N) for $T^{k,i}_{\mathrm{es}}$ and $T^{k,i}_{\mathrm{ee}}$ with their corresponding variances (σ^2); theoretically, both $T^{k,i}_{\mathrm{es}}$ and $T^{k,i}_{\mathrm{ee}}$ will be very close to zero for a time invariant feature. Additionally, the higher the $\Delta_{k,i}$, the more sensitive a candidate feature is for perceiving the beginning and the end of an activity.

2) Feature Sorting and Selection: Finally, we calculate a weighted sum to determine a feature's usefulness as follows:

$$U_{k,i} = \frac{\sum_{p=1}^{M} w_p \cdot R_p}{\sum_{p=1}^{M} R_p} = \frac{(w_\Psi \cdot \Psi_{k,i}) + (w_\Delta \cdot \Delta_{k,i})}{\Psi_{k,i} + \Delta_{k,i}} \tag{3}$$

where there are M confidence parameters (R_p) to be evaluated for each feature $f_{k,i}$ with their corresponding weights w_p (and $\sum_{p=1}^{M} w_p = 1$). In the current phase, there are only two confidence parameters consisting of $\Psi_{k,i}$ and $\Delta_{k,i}$ as defined in (1) and (2). To sum it all up, the larger the time invariance $(\Psi_{k,i})$ and the detection sensitivity $(\Delta_{k,i})$ are, the more useful the feature $(f_{k,i})$ is for recognizing activity A_k.

Instead of utilizing all possible features for an activity A_k, we are more interested in finding a ranking of feature sets $\boldsymbol{f_{k,R}} = \{f_{k,1}, \ldots, f_{k,i}, \ldots, f_{k,C_k}, \ldots f_{k,N}\}$ based on their usefulness. This way we can minimize recognition errors while also mitigating the computational load for the system. That is, those features after f_{k,C_k} (C_k is the cutoff point for activity k) will not provide further improvement for the accuracy during the training phase.

3) Feature Composition: In order to increase the number of time invariant features and enhance overall robustness, we not only measure the similarity between a candidate feature and its GT, but also define two further properties to evaluate the relationship among candidate features.

The first property is the backup (or redundancy) ability to measure the similarity between two candidate features. The backup property for a feature can also be reckoned using a similar convolution equation as (1), with the difference being the GT function is replaced by another feature function. The backup property states the ability of a feature to back up another, and it is a useful property for improving overall robustness.

The second property is the orthogonal ability to estimate how a feature can compensate another one. The orthogonal property (Λ) is inspired by the idea of orthogonal base vectors in linear algebra and its evaluation makes use of an inner product calculation between two feature vectors in the time domain, as shown in (4)

$$\Lambda_{i,j}^{k} = \langle f_{k,i}, f_{k,j} \rangle = \frac{1}{l_k}\left(\sum_{t=1, i\neq j}^{l_k} f_{k,i}[t] \cdot f_{k,j}[t]\right) \tag{4}$$

where $f_{k,i}$ $[t]$ and $f_{k,j}$ $[t]$ are the values for feature i and j at time t and their GT is activity k (whose duration is l_k). The higher the $\Lambda_{i,j}^{k}$ is, the less orthogonal these two features become.

It is very challenging to obtain time invariant features simply from one single sensor, and thus, composite features become desirable. Since an AICO is not restricted to using only one sensor, we can connect multiple sensors with appropriate characteristics to obtain composite features if necessary. The system can systematically select useful features with proper orthogonal and backup properties to generate more time invariant features. Additionally, a composite feature can reckon its usefulness as in (1) and (2) to evaluate its contribution to the outcome.

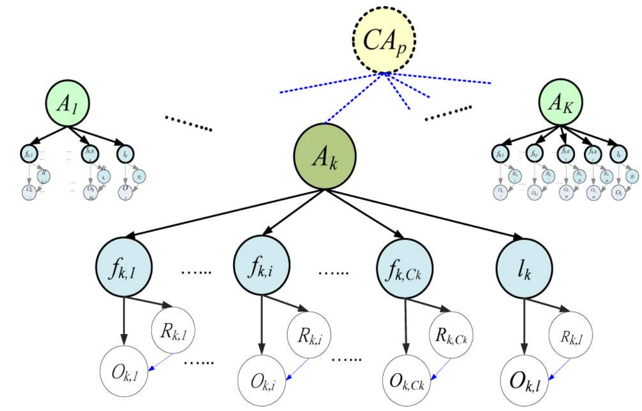


Fig. 7. Multiple enhanced naive Bayesian models for performing data fusion and inferring interleaved activities for a single resident.

By taking appropriate advantage of the properties stated above, we can overcome some critical challenges inherent in activity recognition problems. With time invariant features, we can detect if a feature occurs in a specific period as in [5] to help recognize the activities of interest. As for implicit features, we plan to make use of a similar mechanism as described in [29] to obtain more accurate classification results in future experiments.

D. Feature Fusion and Location-Aware Activity Inference

To effectively make full use of ranking features, a generalized fusion and inference engine for various activities is desirable. Furthermore, in order to increase performance and take into account uncertainties inherent in wireless sensors, we enhance a naive Bayes classifier by incorporating reliability factors to detect potential malfunction so that the system can exclude malfunctioning nodes from continuing to contribute erroneous information, thus increasing overall robustness.

Fig. 7 shows our proposed multiple enhanced Bayesian models for performing feature fusion and activity inference by incorporating ranking features as well as their corresponding reliability factors given the observations from a variety of multimodal sensors. For an activity A_k, the nodes include $f_{k,i}$ (the ith ranking feature), l_k (its location-aware feature), and their corresponding reliability factors ($R_{k,i}$ and $R_{k,l}$) given the current observation vector $\boldsymbol{O}$. According to Bayesian Network theory, the joint probability of interest can be factorized based on the network structure and then expressed by the following equation:

$$p(X_F|O) = \frac{p(x_1, \ldots, x_N, O)}{p(O)} \propto \prod_{i}^{N} p(x_i|\pi_{x_i}) \prod_{j} p(o_j) \tag{5}$$

where $p(\boldsymbol{X}_F|\boldsymbol{O})$ is the conditional probability of querying nodes $\boldsymbol{X}_F$ given observations $\boldsymbol{O}$; π_{x_i} is the parent set of a certain node x_i. Based on the Bayesian Network of Fig. 7, its joint probability of interest can be factorized into (6)

$$\begin{aligned} p(A_k, \boldsymbol{R_k}, R_{k,l}, \boldsymbol{O}) &\propto \beta \cdot \sum_{l_k}\sum_{f_k}\left\{\left[p(l_k|A_k)\cdot\prod_{i=1}^{C_k} p(f_{k,i}|A_k)\right]\right. \\ &\cdot\left[p(R_{k,l}|l_k)\cdot\prod_{i=1}^{C_k} p(R_{k,i}|f_{k,i})\right] \\ &\left.\cdot\left[p(O_{k,l}|l_k)\cdot\prod_{i=1}^{C_k} p(O_{k,i}|f_{k,i})\right]\right\} \end{aligned} \tag{6}$$

where β is a dynamically adjustable value for normalization based on currently selected features. $p(A_k, \boldsymbol{R}_k, R_{k,l}, \boldsymbol{O})$ is the marginal joint probability of interest regarding the unobservable A_k given reliability factors ($\boldsymbol{R}_k$ and $R_{k,l}$) and evidence vector $\boldsymbol{O}$. $\boldsymbol{R}_\mathrm{k}$ denotes the vector composed of the reliability factors from $R_{k,1}$ to R_{k,C_k}. $p(f_{k,i}|A_k)$ and $p(l_k|A_k)$ represent the likelihood that A_k may occur given the feature $f_{k,i}$ and l_k. For now, the prior probabilities $p(A_k)$ for all activities are assumed equal. $R_{k,i}$ and $R_{k,l}$ are reliability factors and can be used to control the contribution of one single feature by using the likelihood $p(R_{k,i}|f_{k,i})$ and $p(R_{k,l}|l_k)$; $p(O_{k,l}|l_k)$ and $p(O_{k,i}|f_{k,i})$ are determined by sensor models, both of which are closely related to whether some specific sensors are triggered to generate corresponding features given the current observations. Obviously, these models with ranking features are more concise than the ones using all possible features, thus greatly reducing the computational burden. Note that we omit all of the time variables in all equations to simplify the expressions.

In order to estimate A_k, each of the individual probabilities in (6) need definitions based on the instrumented environment and required accuracy and robustness. Equation (6) enables us to use a "divide and conquer" approach to estimate $p(A_k|\boldsymbol{O})$ more efficiently. Take $p(R_{k,i}|f_{k,i})$ as an example; different system designers could define it based on their concerns about system robustness such that the resultant system is adaptive to distinct sensors in their instrumented houses.

In order to generalize the definition of reliability factors for $p(R_{k,i}|f_{k,i})$, we adopt a similar approach used in [44], which is dedicated to audio and video sensors, to be applicable for multimodal wireless sensors. The reliability is evaluated based on two common elements, which are not only applicable to our settings. The first element is the natural patterns or the spatio-temporal smoothness when residents perform activities. That is, we utilize recurrent window approach where recent activities will be helpful in predicting the current one. The second element is the persistent tendency of a malfunction to keep contributing erroneous output to the data fusion mechanism even if the output is unstable.

According to the above considerations, the reliability $R_{k,i}$ used to improve the smoothness and performance at time t can be defined as

$$R_{k,i} = \omega \cdot \sum_{t-W_A}^{t} A_k[t] + (1-\omega)\cdot\frac{1}{f_{\mathrm{sample}}}\cdot\sum_{t-W_p}^{t} P_{k,i}[t] \tag{7}$$

where ω is the weight and $\sum A_k[t]$ is the total number of detections of the activity A_k during the interval ranging from time $t - W_A$ to time t. W_A is the window width of the detecting interval. We have discovered from some pilot tests that the number of correctly received packets is closely related to the stability of a remote sensor, and $P_{k,i}[t]$ is the total number of packets (with the frequency f_{sample} as a normalizing constant) correctly received for feature i within the interval W_p looking back from time t. Now, we define $p(R_{k,i}|f_{k,i})$ using a sigmoid function as

$$p(R_{k,i}|f_{k,i}) = \begin{cases} \frac{e^{\lambda\cdot R_{k,i}}}{1+e^{\lambda\cdot R_{k,i}}}, & \text{if } R_{k,i} \geq \tau_R \\ \tau_{\mathrm{fail}}, & \text{otherwise} \end{cases} \tag{8}$$

where λ is a constant to control the slope of the sigmoid function and τ_{fail} and τ_R are predefined constants. Likewise, we can define $p(R_{k,l}|l_k)$ using the concepts as depicted in (7) and (8).

As for $p(l_k|A_k)$, we define it as

$$p(l_k|A_k) = \begin{cases} 1, & \text{if } l_k \subset Z_k \\ N_{\mu=0,\sigma^2_{Z_k}}(d_{l_k,Z_k}), & \text{otherwise} \end{cases} \tag{9}$$

where l_k is the feature representing the current location and $Z_{,k}$ contains the activity zones consisting of the IDs of floor-AICOs for A_k. We model $p(l_k|A_k)$ with a Gaussian function with the Euclidean distance d_{l_k,Z_k} between l_k and a floor-AICO within Z_k as its argument and with $\sigma^2_{Z_k}$ as its variance. The location feature incorporates high-level contextual information or domain knowledge about the AICOs in which the sensors are embedded (e.g., on the door of a microwave in the kitchen).

To estimate A_k and speed up the computation, this work analytically updates $p(R_{k,i}|f_{k,i}), p(R_{k,l}|l_k)$ and $p(l_k|A_k)$ in (6) and then marginalizes all unobservable nodes. Finally, the possible ongoing activities $\hat{A}$ can be derived by

$$\hat{A} = \{A_i | p(A_i|O) \geq A_\tau, A_i \in \{A_1, \ldots, A_K\}\} \tag{10}$$

where A_τ is a threshold to control the how many activities can be detected at the same time. Note that A_k can serve as a component of a more complicated activity (CA_p) for future enhancement, as shown in Fig. 7

IV. EXPERIMENT

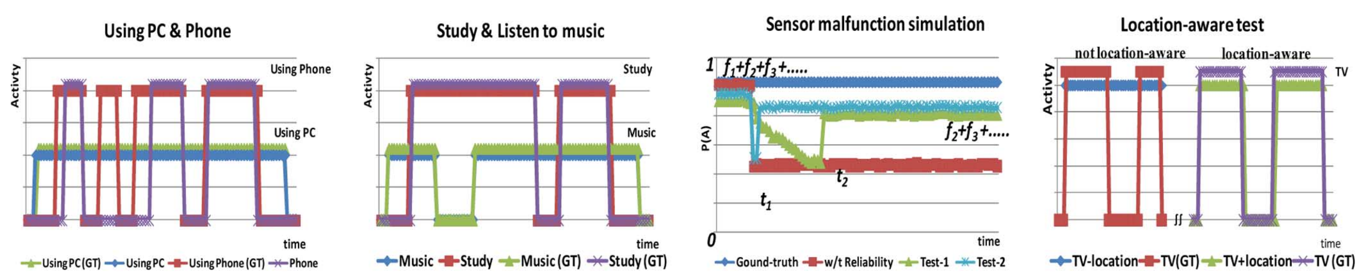
To validate our approach, we have collected training data about activities of interest that commonly occur in a regular home, as listed in Table II. Some of these activities are even interleaved. The dataset was collected across multiple days in the CoreLab, as shown in Fig. 5, from 11 volunteers (ten of whom are not researchers). The volunteers were asked to read the brief instructions in Table II as rough guidance. On average, we have collected varying lengths of data per activity and we have used four cameras deployed on the four corners of the Corelab to collect ground truth for labeling training data afterwards.

The average accuracies (percentage of time that an activity is correctly detected) of the experimental results are shown in Table II, where the comparison between the results with/without the assistance of location context is presented to verify the usefulness of location-awareness. The results with location information outperform the others without location assistance; this

TABLE II
EXPERIMENTAL RESULTS FOR LOCATION-AWARE ACTIVITY RECOGNITION (ELEVEN VOLUNTEERS)

Activities of interest	Deployed sensors[1]	Accuracy without location assistance	Accuracy with location assistance	Top 3 useful sensors [1]	Brief instructions for volunteers
Using PC	A,C,L,M,P, U,V,Z	91%	91%	P,A,M	Use the PC in the study room to surf the Internet (dialing out to a friend is encouraged).
Using phone	A,L,M,P,R, U,V,Z	88%	88%	U,R,A	Answer the phone whenever it rings and dial out at least once to chat.
Studying	A,C,L,M,P, R,Z	72%	92%	P,C,I	Turn on the study light and choose books/magazines to read (turn on the stereo if necessary and volunteers can leave to take a drink).
Listening to the music	A,C,L,M,P, U,V,Z	98%	98%	C,P,A	Turn on the stereo using the remote control and listen to the music (volunteers can walk around if they want).
Watching TV	A,C,L,M,P, U,V,Z	77%	97%	C,P,U	Sit on the couch in the living room and flip through at least three channels using the remote control or buttons on the front panel of the TV (volunteers must leave to take a drink).
Using microwave	A,C,L,M,R, U,V,Z	98%	99%	C,R,A	Put bread into the microwave and heat it.
Using refrigerator	A,C,L,M,R, U,V,Z	96%	98%	R,A,U	Open the refrigerator and take out anything inside then put it back.
Making hot tea	A,C,L,M,R, U,V,Z	89%	91%	C,A,R	Use water boiler or microwave to heat water and prepare a cup of tea, or coffee.
Using the printer	A,C,L,M,P, U,V,Z	90%	92%	C,V,A	Use the PC in the study room to print a document and check the document as well.
Using other appliances with RFIDs	A,C,I,M,L, U,V,Z	67%	68%	C,I,A	Choose one appliance (shredder, toaster, or water boiler, etc.) and plug it into a power-AICO and then operate the appliance for a while.
Walking	L,M,P,Z	96%	96%	L,M,P	Walk through an activity zone.
Sitting	L,M,P,Z	99%	99%	P,L,M	Sit on any chair instrumented with a pressure sensor.

[1]**A**: 3-axis accelerometer, **C**: current flow, **I**:RFID, **L**:Floor sensor, **M**: motion (including infrared proximity sensor), **P**: pressure mat (or pressure blanket), **R**: reed switch, **U**: mercury switch, **V**: vibration, **Z**: bundled sensors (including node ID, signal strength, temperature, humidity, battery voltage, etc.)



(a) (b) (c) (d)

Fig. 8. (a) Continuous recognition result of interleaved activities for using PC and phone. (b) Continuous recognition result of interleaved activities for studying and listening to the music. (c) The simulation of an erroneous feature f_1 and using reliability to control its contribution. (d) Comparison with/without location-aware assistance for "watching TV" activity.

is especially obvious in the testing scenario of "watching TV" and "studying" because volunteers were asked to take some beverages. Table II also lists the ranking of sensors that can generate the most useful features. The majority of the useful features stem from sensors (embedded in AICOs) to detect power usage, contact, pressure, location, and motion. In particular, the power-AICOs are very helpful in distinguishing activities that involve any home appliances such as "watching TV," "listening to the music," "using microwave," and "using other appliances with RFIDs."

Fig. 8(a) illustrates the result of a continuous testing trace and demonstrates the recognition results of two interleaved (and partially concurrent) activities between "using PC" and "using phone." The majority of the data traces were correctly classified by the model, but there were some sparse misclassifications owing to the less informative features from "using phone." Fig. 8(b) shows satisfactory recognition results of two interleaved activities between "studying" and "listening to the music" since their ranking features are very reliable and well selected.

In order to simulate sensor malfunction, we conducted two tests to purposely weaken the most informative feature during the experiment given that the second useful feature is almost as helpful as the first one. Fig. 8(c) shows the Bayesian Network fused results with/without the help from reliability factors. In the case of a device failure (test-1 for battery removal simulation or test-2 for a more realistic simulation where the number of successfully transmitted packets with correct readings was programmed to decrease on purpose) and without any assistance from the reliability factors, the final fused result became less accurate. However, the simulated results show that the reliability factors can automatically keep the engine away from fusing the failed sensor, thus making the overall system less sensitive to device failure and resulting in improved overall robustness. Fig. 8(d) illustrates the comparisons with/without location contextual information for a continuous testing trace.

The addition of location context to the model helps to reduce the classification errors in detecting "watching TV" activity because the volunteers were asked to leave to drink water or take beverages in the middle of the activity. The results shown in Fig. 8 confirm that the model using a Bayesian Network along with the assistance from both reliability factors and location information improves the overall accuracy and robustness.

The experimental results show that composite features do not further improve the accuracy for those activities with very informative ranking features. For instance, the power-AICO can generate very reliable time invariant features for detecting appliance usage. However, composite features do help increase overall robustness as long as we choose at least one feature with good backup property. In addition, the experimental results demonstrate that composite features become useful especially for those activities incapable of extracting reliable time invariant features. For example, an activity such as "making hot tea" involves touching different objects, and the accuracy is improved by at least 20% when incorporating composite features. That motivates us to explore further composite features in future experiments.

V. Discussion

The reason why the accuracy for "using phone" is lower than expectations is that the volunteers sometimes prefer to use the speakerphone rather than the handset, which, we thought, was crucial for the sensors to detect this activity. This unexpected behavior became even more frequent when the volunteers tried to dial out only. That means our current settings for "using phone" should likely be improved by utilizing other sensors (such as current flow sensors) to detect this initially unexpected event of using the speakerphone. Another lower-than-expected recognition accuracy was observed in "using other appliances with RFIDs"; this was partly because our current passive RFID reader sometimes failed to detect the tag attached on the plug of an appliance and volunteers did not get informed if such an event happened.

The importance of location-awareness lies in its abilities to correct the activity estimates which are tightly related to the places where the activities are performed. For example, residents may leave the living room while watching TV, and come back later with a glass of drink. Without location information, the system may not detect this situation, and "watching TV" will be considered persistent over the entire period of "taking a drink." We deployed more sensors to analyze which sensors are crucial for an activity; namely, the actual number of sensors deployed for later usage is often less than the total amount during evaluation. The top three sensors indicated in Table II are meant for this purpose.

One additional observation worth mentioning here is related to interference problems encountered during the experiments. Currently, other wireless signals (probably WiFi) appear to cause some interference from time to time; this caused a decrease in the number of successfully received packets after we switched and tested different channels. However, we found that by preferentially choosing either channels 25 or 26, the results improved, since both of these channels do not overlap with WiFi channels.

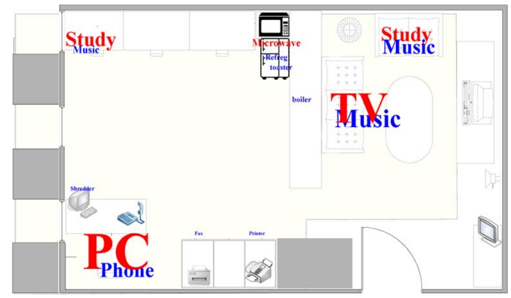


Fig. 9. An activity map for a user in the Attentive Home. The bigger the font size in the map, the more frequent the activity performed in that particular location.

As for interference caused by various other factors (such as different home layout, radio uncertainties or irregularity), we adopted a strategy of simple retransmission of packets because our applications also have the characteristics listed in [3]. These characteristics include low sampling rate, single-hop star topology, unidirectional data flow, tolerable propagation delays, and allowable sporadic packet loss. In addition to randomly delayed retransmission, we utilized a double sampling rate followed by event-driven transmission to ensure satisfactory reception probability of wireless packets.

From questionnaires filled out after the experiments, we found that more than two-thirds of the volunteers had concerns regarding electromagnetic waves generated by wireless nodes. The strategy of event-driven retransmission adopted in this work enables a significant decrease in the amount of wireless transmissions, thus reducing overall exposure to electromagnetic waves and mitigating users' concerns. These concerns do, however, raise the possibility of other interesting research in the future about potential long-term health effects that WSNs might cause.

VI. Applications

Pervasive sensors and actuators in an attentive home can facilitate a multitude of applications ranging from activity recognition, automatic service-provision, energy minimization, health care, to tutors for household tasks (e.g., an LCD to display cooking instructions on a recipe). In particular, they can assist the elderly, thus alleviating some of the burden on their families or caregivers. By understanding more about the activities and significant locations of the residents, our prototype system can better match the expectations of residents in an attentive home. Our experimental results suggest that a number of context-aware applications may soon not only be feasible, but may also, in fact, be quite practical. A feasible wireless sensor network can report up-to-the-minute human/environmental status and provide residents with various summarized information in a more comprehensive and intuitive way.

Fig. 9 illustrates our first version of an "Activity Map" showing activity information inferred from the data of a volunteer. This map utilizes a similar concept to "tag clouds" which are often used for visualizing the popularity of topics in a website. The bigger the font size in the map, the more frequent the activity performed in that particular location. Please refer to [45] for more information.

VII. Conclusion and Future Work

This paper presented a location-aware activity recognition approach utilizing a Bayesian-Network-based fusion engine with inputs from ranking features using assistance from reliability factors reckoned from a variety of wireless sensors to improve overall robustness and performance. This work built on multimodal sensors to collect the most informative features to meet the challenges of recognizing the multifaceted nature of human activities. Our initial work, verified by experiments, has yielded high recognition rates, thus suggesting that this is a feasible approach that may lead to practical ambient intelligence applications such as our first version of the activity map.

Our contribution here is threefold. First, instead of using various simple or dumb sensors directly, we have prototyped AICOs based on proposed design guidelines to facilitate natural interactions with residents. Additionally, the AICOs integrated with our flexible architecture are deployed throughout a living lab rather than directly on the residents, themselves. Second, in order to maximally take advantage of a minimum number of sensors to perform accurate activity recognition, we aim for generating as many time invariant features as possible and evaluate their usefulness for creating a set of ranked distinguishable features. Additionally, feature evaluation, selection, and composition mechanisms can improve overall system performance and robustness. Third, in order to recognize interleaved activities, in the current phase of our work, we have chosen to utilize multiple naive Bayes classifiers and enhanced them by systematically incorporating reliability factors, which serve as confidences to control the contributions from possibly error-prone wireless sensors. Additionally, the results of the location-aware activity recognition are statistically displayed on an activity map in a more pleasant and intuitive way.

The proposed approach has the potential of achieving the goal of implementing a responsive, sensitive, interconnected, contextualized, transparent, and intelligent system that can move from our living lab into the real world. To meet more real human needs and bear with more real-time contextual information in the attentive home, our future work will aim for recognizing multiple residents' concurrent activities by exploring more WSN technologies, extracting more invariant features, and incorporating more complex models (such as HHMM [46] or AHMM [24]) if necessary.

References

[1] Experience and Application Research: Involving Users in the Development of Ambient Intelligence, ISTAG Working Group Final Report - v1 2004, I. A. G. (ISTAG).

[2] J. Rennie and M. Press, "The computer in the 21st century," *Scientific Amer. Special Issue (The Computer in the 21st Century)*, pp. 4–9, 1995.

[3] E. M. Tapia, S. S. Intille, L. Lopez, and K. Larson, "The design of a portable kit of wireless sensors for naturalistic data collection," in *Proc. Pervasive Comput*, 2006, vol. 3968, pp. 117–134.

[4] R. Campbell, J. Al-Muhtadi, P. Naldurg, G. Sampemane, and M. D. Mickunas, "Towards security and privacy for pervasive computing," in *Proc. Int. Symp. Software Security: Theories and Systems: Mext-NSF-JSPS, ISSS*, Tokyo, Japan, Nov. 8–10, 2002, Revised Papers, 2003.

[5] E. M. Tapia, S. S. Intille, and K. Larson, "Activity recognition in the home using simple and ubiquitous sensors," in *Proc. Pervasive Comput.*, 2004, vol. 3001, pp. 158–175.

[6] T. Honkanen, OSGi—Open Service, Gateway Initiative.

[7] K. Koile, K. Tollmar, D. Demirdjian, H. Shrobe, and T. Darrell, "Activity zones for context-aware computing," in *Proc. Int. Conf. Ubiquitous Computing (UbiComp)*, 2003, vol. 2864, pp. 90–106.

[8] D. T. Tran, D. Q. Phung, H. H. Bui, and S. Venkatesh, "Factored state-abstract hidden markov models for activity recognition using pervasive multi-modal sensors," in *Proc. 2005 Int. Conf. Intell. Sensors, Sensor Networks and Information Process. Conf.*, 2005, pp. 331–336, 2005.

[9] S. Dornbush, K. Fisher, K. McKay, A. Prikhodko, and Z. Segall, "XPod: A human activity aware mobile music player," in *Proc. 2nd Int. Conf. Mobile Technol. Appl. Syst.*, 2005, vol. 331, pp. 1–6.

[10] L. Liao, D. Fox, and H. Kautz, "Extracting places and activities from GPS traces using hierarchical conditional random fields," in *Proc. Int. J. Robot. Res.*, 2007, vol. 26, p. 119.

[11] C. D. Kidd, R. Orr, G. D. Abowd, C. G. Atkeson, I. A. Essa, B. MacIntyre, E. Mynatt, T. E. Starner, and W. Newstetter, "The aware home: A living laboratory for ubiquitous computing research," *Cooperative Buildings*, vol. 1670, pp. 191–198, 1999.

[12] M. Philipose, K. P. Fishkin, M. Perkowitz, D. J. Patterson, D. Fox, H. Kautz, and D. Hahnel, "Inferring activities from interactions with objects," *IEEE Pervasive Computing*, vol. 3, pp. 50–57, 2004.

[13] Y. Nakauchi, T. Fukuda, K. Noguchi, and T. Matsubara, "Intelligent kitchen: Cooking support by LCD and mobile robot with IC-labeled objects," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst. IROS*, 2005, pp. 1911–1916.

[14] S. Helal, B. Winkler, C. Lee, Y. Kaddourah, L. Ran, C. Giraldo, S. Kuchibhotla, and W. Mann, "Enabling location-aware pervasive computing applications for the elderly," in *Proc. Pervasive Comput. Commun.*, 2003, pp. 531–546.

[15] D. J. Patterson, L. Liao, D. Fox, and H. Kautz, "Inferring high-level behavior from low-level sensors," in *Proc. 5th Int. Conf. Ubiquitous Computing (Ubicomp)*, pp. 73–89.

[16] E. M. Tapia, N. Marmasse, S. S. Intille, and K. Larson, "MITes: Wireless portable sensors for studying behavior," in *Proc. Extended Abstracts Ubicomp 2004: Ubiquitous Computing*, 2004, pp. 156–157.

[17] M. Bennewitz, W. Burgard, and S. Thrun, "Learning motion patterns of persons for mobile service robots," in *Proc. IEEE Int. Conf. Robot. Autom., ICRA*, 2002, vol. 4, pp. 3601–3606.

[18] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Proc. Pervasive*, 2004, pp. 1–17.

[19] J. Fogarty, C. Au, and S. E. Hudson, "Sensing from the basement: A feasibility study of unobtrusive and low-cost home activity recognition," in *Proc. 19th Annu. ACM Symp. User Interface Softw. Technol.*, 2006, pp. 91–100.

[20] J. Chen, A. H. Kam, J. Zhang, N. Liu, and L. Shue, "Bathroom activity monitoring based on sound," in *Lecture Notes in Computer Science*, 2005, p. 47.

[21] H. Sidenbladh and M. J. Black, "Learning image statistics for Bayesian tracking," in *Proc. Int. Conf. Comput. Vision*, 2001, vol. 2, pp. 709–716.

[22] A. Madabhushi and J. K. Aggarwal, "A Bayesian approach to human activity recognition," in *Proc. 2nd Int. Workshop on Visual Surveillance*, 1999, pp. 25–30.

[23] K. M. Kitani, Y. Sato, and A. Sugimoto*, Deleted Interpolation Using a Hierarchical Bayesian Grammar Network for Recognizing Human Activity*. Piscataway, NJ: IEEE Press, 2005, vol. 246.

[24] H. H. Bui, S. Venkatesh, and G. A. W. West, "Policy recognition in the abstract hidden markov model," *J. Artif. Intell. Res.*, vol. 17, pp. 451–499, 2002.

[25] M. Piccardi and O. Perez, "Hidden Markov models with kernel density estimation of emission probabilities and their use in activity recognition," in *Proc. IEEE Conf. Comput. Vision, Pattern Recogn., CVPR*, 2007, pp. 1–8.

[26] S. Lu, J. Zhang, and D. D. Feng, "Detecting unattended packages through human activity recognition and object association," *Pattern Recognition*, vol. 40, pp. 2173–2184, 2007.

[27] S. S. Intille, K. Larson, E. M. Tapia, J. Beaudin, P. Kaushik, J. Nawyn, and R. Rockinson, "Using a live-in laboratory for ubiquitous computing research," in *Proc. Pervasive*, Berlin, Heidelberg, Germany, 2006, Springer-Verlag.

[28] D. H. Wilson and C. Atkeson, "Simultaneous tracking and activity recognition (STAR) using many anonymous, binary sensors," in *Proc. Pervasive*, 2005, pp. 62–79.

[29] J. Lester, T. Choudhury, N. Kern, G. Borriello, and B. Hannaford, "A hybrid discriminative-generative approach for modeling human activities," in *Proc. Int. Joint Conf. Artif. Intell., IJCAI*, 2005, pp. 766–772.

[30] D. J. Patterson, D. Fox, H. Kautz, and M. Philipose, "Fine-grained activity recognition by aggregating abstract object usage," in *Proc. 9th IEEE Int. Symp. Wearable Comput.*, 2005, pp. 44–51.

[31] H. L. Chieu, W. S. Lee, and L. P. Kaelbling, "Activity recognition from physiological data using conditional random fields," in *Proc. SMA Symp.*, 2006.

[32] W. Shuangquan, Y. Jie, C. Ningjiang, C. Xin, and Z. Qinfeng, "Human activity recognition with user-free accelerometers in the sensor networks," in *Proc. Int. Conf. Neural Netw. Brain, ICNN&B*, 2005, pp. 1212–1217.

[33] J. Beaudin, S. Intille, and E. M. Tapia, "Lessons learned using ubiquitous sensors for data collection in real homes," in *Proc. Conf. Human Factors in Comput. Syst.*, 2004, pp. 1359–1362.

[34] H. Ishii, "Bottles: A transparent interface as a tribute to Mark Weiser," *IEICE Trans. Inf. Syst.*, vol. 87, pp. 1299–1311.

[35] J. H. Friedman, "On bias, variance, 0/1—Loss, and the curse-of-dimensionality," *Data Mining and Knowledge Discovery*, vol. 1, pp. 55–77, 1997.

[36] E. M. Tapia, S. S. Intille, and K. Larson, "Activity recognition in the home using simple and ubiquitous sensors," in *Proc. Pervasive Comput.*, 2004, vol. 3001, pp. 158–175.

[37] J. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in *Proc. 18th Int. Conf. Mach. Learn.*, 2001, pp. 282–289.

[38] Y. Chuang-wen, C. Yi-Chao, C. Ji-Rung, P. Huang, C. Hao-hua, and L. Seng-Yong, "Sensor-enhanced mobility prediction for energy-efficient localization," in *Proc. 3rd Annu. IEEE Commun. Soc. Conf. Sensor, Mesh and Ad Hoc Commun. Networks, SECON*, Sep. 2006, pp. 565–574.

[39] Tmote-Sky. [Online]. Available: http://www.sentilla.com/moteiv-transition.html

[40] Phidgets. [Online]. Available: http://www.phidgets.com/

[41] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. 7th IEEE Int. Conf. Comput. Vision*, 1999, vol. 2, pp. 1150–1157.

[42] J. Lester, T. Choudhury, and G. Borriello, "A practical approach to recognizing physical activities," in *Proc. Pervasive*, 2006, vol. 6, pp. 1–16.

[43] R. Aipperspach, E. Cohen, and J. Canny, "Modeling human behavior from simple sensors in the home," in *Proc. IEEE Conf. Pervasive Comput.*, 2006, vol. 3968, pp. 337–348.

[44] D. Lo, R. A. Goubran, and R. M. Dansereau, "Robust joint audio-video talker localization in video conferencing using reliability information-II: Bayesian network fusion," *IEEE Trans. Instrument. Measure.*, vol. 54, pp. 1541–1547, 2005.

[45] C. H. Lu, Y. C. Ho, and L. C. Fu, "Creating robust activity maps using wireless sensor network in a smart home," in *Proc. IEEE Int. Conf. Autom. Sci. Eng.*, 2007, pp. 741–746.

[46] H. H. Bui, D. Q. Phung, and S. Venkatesh, "Hierarchical hidden Markov models with general state hierarchy," in *Proc. 19th Nat. Conf. Artif. Intell.*, pp. 324–329.

Ching-Hu Lu (M'06) received the B.S. and M.S. degrees in the electrical engineering from the National Taiwan University of Science and Technology, Taiwan, R.O.C., in 1993 and 1995. He is currently working towards the Ph.D. degree at the Department of Computer Science and Information Engineering at the National Taiwan University, Taipei, Taiwan

He also has more than six years of experience working in the IC industry in Taiwan. His research interests include smart environments, intelligent spaces, wireless sensor network, context-aware technologies, factory automation, and topics related to them.

Li-Chen Fu (M'84–SM'94–F'04) received the B.S. degree from the National Taiwan University, Taipei, Taiwan, in 1981, and the M.S. and Ph.D. degrees from the University of California, Berkeley, in 1985 and 1987, respectively.

Since 1987, he has been a Faculty Member with the Department of Electrical Engineering and the Department of Computer Science and Information Engineering, National Taiwan University, where he currently is a Professor. His research interests include robotics, flexible manufacturing systems scheduling, shop floor control, home automation, visual detection and tracking, e-commerce, and control theory and applications.

Dr. Fu is a Senior Member of the IEEE Robotics and Automation Society and the IEEE Automatic Control Society.