RAM: Radar-based Activity Monitor

Md Abdullah Al Hafiz Khan*, Ruthvik Kukkapalli†, Piyush Waradpande†,
Sekar Kulandaivel†, Nilanjan Banerjee†, Nirmalya Roy*, Ryan Robucci†
*Department of Information Systems
†Department of Computer Science and Electrical Engineering
University of Maryland Baltimore County
Email: mdkhan1@umbc.edu, sb20615@umbc.edu, pi5@umbc.edu,
sekark1@umbc.edu, nilanb@umbc.edu, nroy@umbc.edu, robucci@umbc.edu

***Abstract*—Activity recognition has applications in a variety of human-in-the-loop settings such as smart home health monitoring, green building energy and occupancy management, intelligent transportation, and participatory sensing. While fine-grained activity recognition systems and approaches help enable a multitude of novel applications, discovering them with non-intrusive ambient sensor systems pose challenging design, as well as data processing, mining, and activity recognition issues. In this paper, we develop a low-cost heterogeneous Radar based Activity Monitoring (RAM) system for recognizing fine-grained activities. We exploit the feasibility of using an array of heterogeneous micro-doppler radars to recognize low-level activities. We prototype a short-range and a long-range radar system and evaluate the feasibility of using the system for fine-grained activity recognition. In our evaluation, using real data traces, we show that our system can detect fine-grained user activities with 92.84% accuracy.**

I. Introduction

Activity recognition is an important enabler of the successful implementation and deployment of several applications including home energy management [1], assistive care for the elderly [2], and recommendation and alert systems for individuals suffering from Alzheimer's and Dementia [3]. For instance, activity recognition can augment home energy management systems by providing the underlying contexts associated with energy usage [4]. The inferred activities can provide better understanding of why energy consumption was high at certain times of the day, and can form the basis of directed recommendations for energy usage reduction. At assistive care facilities for the elderly, realtime activity monitoring systems can help provide just-in-time intervention in the event of a fall or medical emergency. Inferring activities can also form the basis of other health applications such as remote physiotherapy, diet and obesity management, as well as functional, cognitive and behavioral health assessment [5].

Accurately inferring a user's activity is a non-trivial and challenging problem. Prior research has focused on two distinct classes of systems to infer user activities in a home setting. The first set of systems focus on wearable sensors such as smart bands, smart watches, wireless neckless, and sensors built into smartphones [6] for activity recognition. Such systems have to be worn by the user at all times which can be tedious, especially for individuals with Alzheimer's and Dementia. An alternative

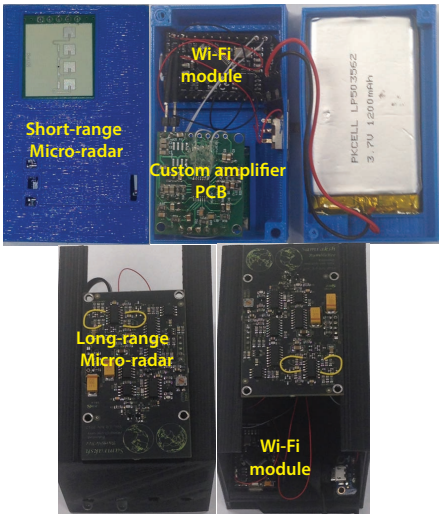


Fig. 1: *The figure illustrates the* ***RAM*** *prototype. It comprises two types of radar sensors: (1) a micro-doppler radar tuned for short ranges while sensitive to fine-grained movements; (2) longer-range radar that is sensitive to coarse-grained movements.* ***RAM*** *uses a custom-designed amplifier PCB for the short-range measurements and is enclosed in a custom 3D-printed enclosure making it easy to deploy. The radars use a Wi-Fi module and a micro-controller for data transfer.*

to wearable sensors is to use ambient sensors deployed in the environment to infer activities. These include camera systems, RFID-based systems [7], and Wi-Fi based activity recognition systems [8]. These systems, however, are suited for fairly coarse-grained activity recognition. For instance, most systems can distinguish walking from running and crawling. While coarse-grained activity recognition can be useful for certain applications, they cannot be used for fine-grained behavioral assessment. The ability to distinguish fine-grained activities such as bathroom activities (brushing and washing hands) or cooking activities (chopping and stirring) can help with applications related to obesity management [9], cognitive health assessment [7] or determining compliance with medication intake and physiotherapy performance.

To address the above challenges, in this paper, we present a **R**adar-based **A**ctivity **M**onitoring system, **RAM**, that uses an array of heterogeneous micro-doppler radars to determine fine-grained user activities. Figure 1 illustrates the prototype wireless battery-powered micro-

978-1-4673-9953-1/16/$31.00 ©2016 IEEE

radar based on the Bumblebee used in **RAM** and it also shows the short range radar. The radars are strategically placed to recognize movements in certain portions of the human body. Aggregating data from multiple radars can then be used to infer fine-grained activities such as brushing, sweeping, stirring, and chopping. In **RAM**, we design a signal processing stack that takes data from multiple radar sensors, applies appropriate low-pass filters, extracts features, and then uses a machine learning algorithm to distinguish activities at different locations in the house such as bathroom, kitchen, and bedroom. These activities can then be used for applications described above.

The design, implementation, and evaluation of **RAM** presents the following novel research contributions.

- **Multi-radar System for Activity Recognition:** We fuse data from an array of micro-doppler radars to determine activities. Each radar measures human body movements from a different vantage point and combining data from multiple radars provides higher accuracy of detection over using a single radar. **RAM** uses a signal processing stack that uses intelligent filtering of the data, feature extraction, and classification using a random forest classifier to determine activities.
- **Functional Prototype Implementation:** We have designed, implemented, and evaluated a realtime activity recognition system using four micro-doppler radars. The data from the radars is collected wirelessly (over a Wi-Fi connection) to a central computer where our signal processing stack uses the data to infer the activities. We show that **RAM** can recognize activities in real-time with an accuracy of 92.84%.

II. Related Work

RAM builds on previous work on activity recognition sensors, the underlying signal processing algorithms, and applications of activity recognition and discovery to healthcare and home energy management. Here we compare and contrast **RAM** with the most relevant literature.

A. Activity Recognition

A significant amount of research has been performed on inferring activities of humans in a home and caregiver setting. The sensors used include those deployed in the environment such as cameras [10], RFID tags [11], and RF-based systems such as radars and Wi-Fi [12], [13]. However, most of these systems focus on recognition of coarse-grained activities such as walking, running, and crawling. Moreover, these systems are evaluated in a constrained environment like a treadmill [14]. **RAM** on the other hand focuses on recognition of fine-grained quotidian activities related to cooking, cleaning, entertainment, or bathroom activities. It aggregates data from an array of micro-doppler radars to infer subtle differences between various activities performed in the home. Researchers have also proposed the use of wearable sensors for activity recognition [15]. These include wrist-worn bands [16], sensors built into smartphones [17], and textile sensors [18]. While wearable sensors may be accurate in determining user activities, they require the user to carry the sensor at all times, which is cumbersome for users who suffer from Alzheimer's or Dementia. **RAM** uses portable proximity-based micro-doppler radars that are deployed in the environment and can determine user activities with minimal intrusiveness.

B. Signal Processing

Activity recognition systems use a wide range of signal processing algorithms that convert data from the sensors into activity labels. For instance, Otero et al. [19] use spectrogram analysis to design a simple binary classifier to determine whether a user is walking. Kim et al. [20] extract features from Doppler radar's spectrogram and classified seven different activities using a support vector machine. Bilik et al. [21], and Smith et al. [22] explored other micro-doppler radar features to distinguish movements in humans, animals, and vehicles. In our work, we use different features from the $I(t)$ and $Q(t)$ channels. These features are computationally cheap to extract but show more than 90% accuracy in recognizing human activities. Other signal processing algorithms such as Hidden Markov Models are used in activity classifications when wearable inertial and textile sensors are used to determine user activities [23].

C. Applications of Activity Recognition

Activities recognized in a home environment can have a plethora of applications. In the home energy management setting, activities can form the basis of context to determine how and why the home consumes the energy it does [25]. Activity inferencing can be used to determine anomalies in behavior for users suffering from Dementia or Alzheimer's [26], and can be a good predictor of falls in elderly assistive care facilities. Radar-based activity recognition has been used for other applications such as gait analysis [27]. This work is complementary to these applications. Our goal is to design a robust, non-contact activity recognition system in a home environment that can consequently be used for various applications.

III. Multiple Radars for Fine-grained Activity Recognition

Our work focuses on classifying activities described in Table I. For instance, in the bathroom we need to distinguish whether the user is brushing his teeth, washing his hands, or walking. Inferring these fine-grained activities has several applications. For example, understanding whether the user is brushing his teeth regularly or whether he is washing his hands can help determine his compliance to regular hygiene. To distinguish these activities accurately, however, it is important that fine-grained motion of different parts of the human body be captured accurately. For example in order to distinguish stirring from chopping it is important to determine the differences in the subtle movement of the hand for these

TABLE I: *The table describes the set of activities that* ***RAM*** *recognizes. The table is divided into the high-level activity (based on room) and the corresponding low level activities in each category. For example, dancing, standing, and sitting comprise of the Entertainment activity that may be performed in a living room.*

High-level Activity	Low-level Activities
Entertainment	Sitting Dancing Standing
Bathroom Activities	Brushing Walking Washing Hands
Cleaning	Spraying Wiping
Cooking	Stirring Chopping Walking

two activities. Distinguishing walking from dancing requires capturing the difference in movement of the upper and lower extremity of the human body. Cameras such as Kinect can be used to distinguish these movements, however, they are expensive ($200/unit), require wall power to operate, and impose privacy concerns if they are deployed in a home setting.

In this paper we use two types of radars: (i) a high-frequency radar tuned to operate at a short range; and (ii) a low-frequency long-range radar. The short-range radar can differentiate between subtle movements while the long-range radar covers a large area. By combining an array of these two types of radars, **RAM** can provide inexpensive coverage and help infer fine-grained activities. Both custom radars if designed from scratch will cost less than $20/unit. In the rest of the section, we provide an overview of our radars and then motivate the need for an array of heterogenous radars.

A. Micro-doppler Radars

RAM uses a 5.8-GHz Bumblebee radar and a 25-GHz KLC-2 short-range radar. The Bumblebee radar is an off-the-shelf pulse-capable radar [28] and the other is a continuous-mode radar used in a modified custom in-house hardware for measuring dominant-signal phase changes [29]. Doppler radars work on the principle of effect [30]. When the radar emits a signal at frequency (f_t) and reflected signal from the object is received by the radar at a shifted frequency (f_r). The received wave is demodulated in the $I(t)$ and $Q(t)$ channels. If the object is moving at a velocity v relative to the radar, then $f_r = f_t \cdot (\frac{c-v}{c+v})$, and the frequency shift $f_d = f_t - f_r \approx f_t \cdot \frac{2v}{c}$ since $v \ll c$.

Formally, if the radar transmits a wave described by $S(t) = A_T \cos(2\pi f_t t)$, where A_t represents the amplitude of the transmitted signal, the received wave can be represented by $R(t) = A_R \cos(2\pi(f_t + f_d)t + \phi)$, where A_R is the amplitude, ϕ is the initial phase that represents the initial distance between the radar and the object [29] and f_d is the doppler frequency shift. The received signal is demodulated in the I and Q channel. The I (real part of the wave) can be represented as:

$$\begin{aligned} I'(t) = R(t)S(t) &= A_T A_R \cos(2\pi(f_t + f_d)t + \phi)\cos(2\pi f_t t) \\ &= \frac{A_T A_R}{2}\{\cos(2\pi(2f_t + f_d)t + \phi) + \cos(2\pi f_d t + \phi)\} \end{aligned} \tag{1}$$

Using a low-pass filter the high frequency component, $(2\pi(2f_t + f_d)t + \phi)$ is eliminated and the I component of the signal can be written as:

$$I(t) = \frac{A_T A_R}{2}\cos(2\pi f_d t + \phi) \tag{2}$$

Similarly, the Q component of the wave can be written as follows:

$$\begin{aligned} Q(t) &= \frac{A_T A_R}{2}\cos\left\{(2\pi f_d t + \phi) - \frac{\pi}{2}\right\} \\ &= \frac{A_T A_R}{2}\sin(2\pi f_d t + \phi) \end{aligned} \tag{3}$$

From Eq. 2 and 3 we can calculate phase of the wave as follows:

$$\theta = \arctan\left\{\frac{Q(t)}{I(t)}\right\} = 2\pi f_d t + \phi = 4\pi\frac{vt}{\lambda} \tag{4}$$

The rate of change of θ is proportional to the velocity of the object relative to the radar. Note that (4) returns a value in the interval $(\frac{\pi}{2}, -\frac{\pi}{2})$ that leads to discontinuities in the phase calculations. Phase unwrapping is the process of removing this discontinuity by adding or subtracting an integer number of rotations to each sample point so that the unwrapped phase value corresponds to the measured phase. The phase unwrapped equation can be represented as follows:

$$\theta_{u,i} = \theta_{u,i-1} + mod(\theta_{w,i} - \theta_{w,i-1} - \pi, 2\pi) + \pi \tag{5}$$

where $\theta_{u,i}$ and $\theta_{w,i}$ denote respectively the unwrapped phase and the measured phase at i^{th} time. To begin unwrapping, we assume $\theta_{u,0} = \theta_{w,0}$.

Θ, the derivative of θ, and the energy in the I and Q components of the signal are good distinguishing features for activities. We use these features in our signal processing algorithm (described in Section II-B) to classify the activities. To distinguish change in velocity at short distances between the radar and the object, the amplitudes of the I and Q components of the signal needs to be amplified. We, therefore, have designed a custom amplifier board (see Figure 1) for our short-range radar. Our actual physical system will observe multiple moving objects, which results in a composite complex signal. We apply a simplifying assumption for discussion that a dominant energy in a narrow band generates the majority of the signal. Thus after extracting the phase and applying a smoothing (low-pass) filter we create a transient phase signal representing an effective movement that is used for analysis.

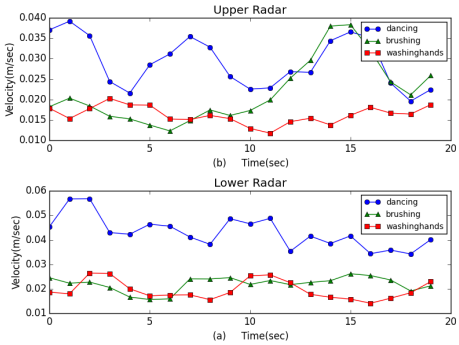


Fig. 2: *The figure shows the efficacy of using multiple radars for activity recognition. Radar Upper was placed pointing to the top portion of the user's body and Radar Lower was placed pointing to the lower portion of the human body. The figure plots the velocity calculated from the data from the radars when the user was performing the dancing, brushing, and washing hands activities. The figure shows that combining data from the two radars can help distinguish the three activities.*

B. Case for an Array of Heterogeneous Radars

RAM uses an array of heterogeneous radars to infer activities accurately. In this section, we motivate the need for an array of heterogeneous radars.

Determining differences in activities: Activities such as walking, brushing, cleaning, and stirring involve movements in more than one part of the human body. For example, walking involves moving both legs and the hands, and brushing involves moving hands from left to right (or vice-versa) and movement of the head. A radar uses a directional antenna (with a cone of view of 30°), hence placing multiple radars positioned at different parts of the human body can help distinguish activities accurately. Figure 2 illustrates an example with two radars. The first radar (Upper Radar) is placed facing the upper part of the body, and the second radar (Lower Radar) faces the lower part of the body. Figure 2 plots the velocity (differential of θ described in (4)) for both radars over time as these activities are performed. Upper and Lower radars exhibit higher velocity for dancing since it involves movement of the upper and lower part of the body. On the other hand, washing hands and brushing activities show lower velocity with respect to the upper and lower radar compared to dancing. Moreover, the velocity is higher for brushing compared to washing hands since it involves faster movement of the hands. Therefore, using an array of heterogeneous radars strategically placed can help distinguish activities accurately.

Distinguishing fine-grained movements and maximizing coverage: Figure 3 (a) and (b) show the precision and recall of classifying washing hands and brushing as a function of the distance of the user from the radars. For this experiment we used the short-range radar (KLC-2) and the long-range radar (Bumblebee). The figure shows that short-range radar has higher precision and recall for the activities at ranges lower than 0.5 m, while the long-range radar shows higher precision and recall values at distances greater than 0.5 m. The results demonstrate the importance of using a heterogeneous mix of radars depending on room size and coverage. For instance, in a bathroom which is usually a smaller in size than a large room, the short range radars can provide higher accuracy, while in a larger living room the long-range radar can be deployed for better coverage and accuracy.

IV. **RAM** Software Architecture

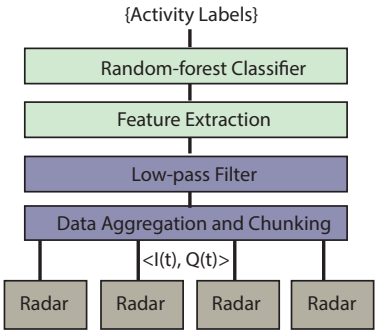
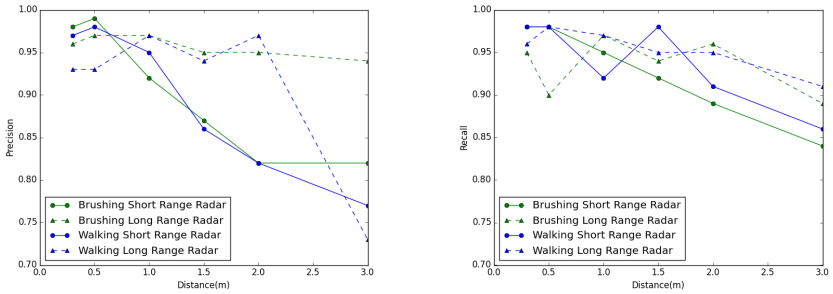


Fig. 4: *The figure illustrates the overall software architecture for **RAM**. Data from an array of micro-radars is processed, filtered, and then chunked into frames. Consequently, features are extracted from the data chunks which are then used for activity classification using a random forest classifier.*

RAM comprises of an array of micro-doppler radars. These radars attach to a custom Wi-Fi module. Data is aggregated over Wi-Fi on a central computer. The signal processing algorithm for **RAM** runs on the computer. The overall architecture of the signal processing architecture is illustrated in Figure 4. Data from the array of micro-radars is aggregated and chunked into frames of sufficient length to identify an activity. We evaluate the sensitivity of **RAM** to the frame size in the evaluation section. **RAM** then extracts features from the data chunk and uses a random forest classifier to classify the activity. In the system we assume that the location of the radar is known a priori. The system, therefore, distinguishes activities at a specific location. For example, in the bathroom the system distinguishes brushing, walking, and washing hands. The set of activities specific to a location is described in Table I. We now describe each functional component of the **RAM** software architecture in detail.

A. Data Chunking and Signal Filtering

The reflected signal from the radars is demodulated in the I and Q components. The signals however are noisy and include artifacts due to other movements in the environment. It is important to filter these artifacts to meaningfully extract features from the data required for our activity recognition. We have therefore designed a simple low-pass filter to remove high frequency noise. To determine the band for the filter, we performed a simple calculation based on the velocity of movement of the human body while performing activities. We note that while performing activities, the human body moves with speeds in the range 0–2 m/s [31]. For the 25-GHz radar, applying (4), the received frequency will lie in



(a) (b)

Fig. 3: *The figure shows the accuracy of classifying brushing, and walking as a function of the distance of the radar from the user. The KLC-2 short-range radar has higher accuracy at short distances while the Bumblebee long-range radar has a higher accuracy at longer distances. Therefore, a heterogeneous mix of these radars can help with higher accuracy and coverage in houses that have rooms of varying sizes.*

the interval 0–30 Hz. We, however, found by applying a FFT on the data that most of the high-energy frequency components lie in between 0–10 Hz. We therefore apply a low-pass filter with a maximum frequency of 10 Hz.

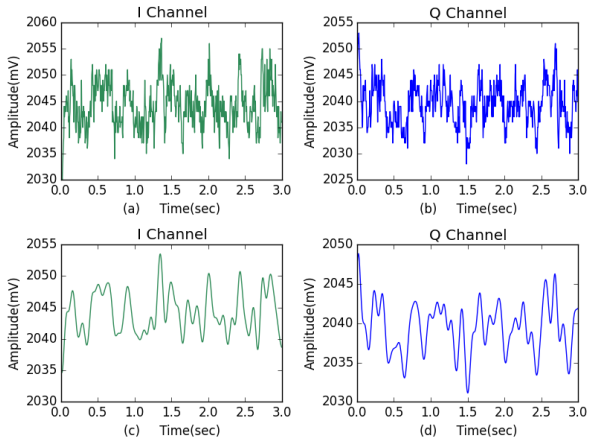


Fig. 5: *(a), (b) Original I and Q components of the received signal. (c), (d) Filtered I and Q signal of the received signal*

Figure 5 (a) and (b) illustrate the unfiltered I and Q channel waveforms obtained from the radar and Figure 5 (c) and (d) represent the filtered counterparts after applying a low pass filter. The filtered data is smooth and can be used for feature extraction.

B. Feature Extraction and Classification

The filtered I and Q components of the received signal is chunked into *frames*. Ideally, each frame should have sufficient data that corresponds to a single occurrence of an activity. We evaluate the accuracy of **RAM** as a function of the frame length in the evaluation section. For each frame, **RAM** extracts the following features. (i) **Phase (Θ) of the signal**: The phase provides a measure of the amount of displacement of a body part when performing an activity. For instance, the displacement of the hand inherent in brushing is higher than washing hands. Figure 6 illustrates how the phase can be used to distinguish dancing (requires high displacement) from brushing (medium displacement) and washing hands (low displacement). The radar pointing to the upper part of the body (Upper Radar) shows higher displacement for brushing than washing hands, while the radar pointing to the lower part of the body (Lower Radar) shows higher displacement for the washing hands over brushing. (ii) **Velocity:** The rate of change of θ quantifies the velocity of movement of the object relative to the radar. The velocity can help distinguish activities such as walking from running. Figure 2 illustrates an example of how velocity can be used for distinguishing the three activities. (iii) **Raw Features from I and Q channel data**: We extract a set of features from the filtered I and Q channel data such as average, standard deviation, range (difference between the maximum and minimum values) and total energy $|I|^2 + |Q|^2$ in a frame. These metrics quantify the direction of movement and the randomness associated with the fine-grained movements of an activity.

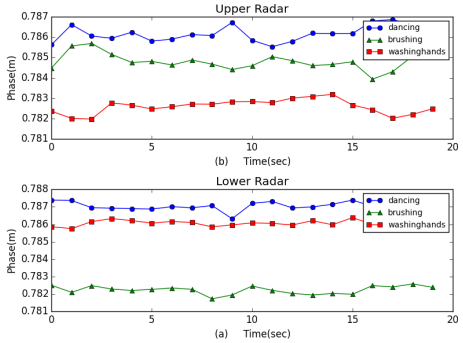


Fig. 6: *The figure plots the phase as a function of time for three activities and two radars. The top figure corresponds to the radar pointing to the upper portion of the human body and the bottom figure corresponds to the radar pointing to the lower portion of the body. The phase is a measure of the effective displacement of the body relative to the radar and can be used to distinguish the movements.*

Feature Selection: We also perform feature selection to reduce the dimension of the feature set. Too many irrelevant features may decrease accuracy of classification and can cause over-fitting. Reducing the number of features can improve accuracy and also reduce training

time. There are two approaches that can be applied for feature selection: feature ranking and subset selection. In feature ranking, each feature is assigned a weight based on some criteria such as predefined threshold. In feature subset selection, the algorithm searches for an optimal subset of features based on a cost function. In this work, we use linear support vector machine to discard lower-weighted features and keep higher-weighted features. Linear Support Vector Classifier (LSVC) [32] generates a hyperplane that has maximum separation for two classes of data. For instance, if we consider a set of instance-label pairs (x_i, y_i) where $x_i \epsilon R^n$, $y_i \epsilon \{1, -1\}$ and $i = 1, ..., l$, a SVC problem can be formulated as follows:

$$\min_{w,b} \frac{1}{2} w^T w + C \sum_{i=1}^{l} L(w, b; x_i, y_i) \tag{6}$$

where $L(w, b; x_i, y_i) = \max(1 - y_i(w^T \phi(x_i) + b), 0)$ represents $L1$ loss function and $C \geq 0.01$ is the penalty parameter. ϕ is a function that maps training instances (x_i) to the higher dimension. For our feature selection algorithm, we calculate a confidence score that is equal to the signed distance of a sample from the hyperplane. Based on this score we rank our features and discard features that have confidence scores of less than a predefined threshold.

Random Forest Classifier: To infer an activity from the radar data we use a Random Forest (RF) classifier because it ables to handle large activity data sets. The classifier is trained on multiple aggregated streams of features extracted from the array of radars. A random forest classifier is a collection of k decision trees. Each decision tree is trained using a vector of inputs $\Theta_1, \ldots, \Theta_{k-1}, \Theta_k$ taken from the training set. During the testing phase, each tree classifies the test input and the mode of the classifications is taken as the final activity label.

V. **RAM** Hardware Prototype

We have designed and implemented a fully functional prototype of **RAM** using two radars: the Bumblebee [28] and the KLC-2 radar [33]. Our prototype is illustrated in Figure 1 and includes a Spark core micro-controller and Wi-Fi module that reads data from the radars directly using an analog-to-digital converter. We enclose the entire system using a custom-designed 3D-printed enclosure for portability and deployment ease. We have written custom software that runs on the micro-controller for collecting and disseminating the radar data, and software on the central computer that implements the signal processing stack. We briefly describe the characteristics of the two radars and the wireless module below.

Bumblebee. Bumblebee is a low-power, mote-scale Pulse Doppler Radar that detects movements of a target within a range of 10 m [28]. The Bumblebee radar is capable of detecting radial velocities up to 2.6 m/s at a distance of up to 10 m, which is well-suited to capture various human activities inside a smart home environment. It operates at a 5.8-GHz center frequency and has an on-board internal antenna with a 60° conical coverage. The board operates at a voltage ranging from 3.6 to 12 V and outputs a voltage ranging from 0 to 3.2 V with a 1.65-V DC offset for the I and Q channels.

RFBeam KLC-2. RFBeam KLC-2 is a short range K-band development module which operates at 25-GHz. The KLC-2 module also provides I-Q channel output similar to that of the Bumblebee. For determining subtle movements of human body such as those at the upper/lower extremities, hands, legs or neck movements at short distances, we have designed a custom amplifier board with a gain of 94 dB to amplify the I and Q channel inputs.

Spark Core. To collect data from the radars and transmit it wirelessly to the central computer, we integrated a Wi-Fi module and a micro-controller module. The wireless and micro-controller module consist of a Texas Instruments CC3000 radio and STM32F103 72-MHz ARM Cortex M3 processor. The entire unit, known as Spark Core [34] runs at an input voltage ranging from 3.6 to 6 V and has a 3.3-V logic level, which allows its 8-channel 12-bit anoalog-to-digital converter (ADC) to cover the output range of the radars.

VI. System Evaluation

We evaluate **RAM** while focusing on the following key questions: (i) What is the accuracy of activity classification for **RAM**? (ii) How does the accuracy of classification vary as a function of frame length and overlapped window between frames? (iii) How is the accuracy dependent on the distance of the human from the radar?

A. Experimental Setup

We recruited 5 subjects in the age group of 18 to 50 years to perform the activities. The height and weight of the subjects ranges from 5'6" to 5'11" and 155–180 lbs respectively. All subjects are asked to perform various activities presented in Table I. We placed four **RAM** devices at approximately 3 m from the subjects. All the four **RAM** radar systems were placed in front of the subject with two modules at a 15° angle while the other two at a 45° angle to the right and left side of the subject. The radars were numbered 1 through 4 starting from the subject's left to the subject's right side. The radars 1 and 4 were placed at the waist height level ($\approx$ 3 feet) and the radars 2 and 3 are placed at the eye-level ($\approx$ 5 feet) to detect both the lower and upper extremities body movements of the subjects. For some of our experiments, we varied the distance between the radar and the subject while they performed the activities. Subjects performed their activities in a natural way without following any specific order or sequence. The radar data was recorded at a sampling frequency of 200 Hz. We first collected radar data for the training set and subsequently for the testing set across multiple individuals. We used the Kinect to capture ground truth video data. To record the activities each experiment was repeated multiple times for a duration ranging from 60 to 300 sec per subject. Each radar's signal values consisting of I and Q

channel data were transferred wirelessly to a computer for filtering and processing. Combining all data sets together across the 5 human subjects, we collected a data spanning 24 480 sec (≈ 7 hrs) in total. This corresponds to 4 896 000 values received from each of the radar sensor with time intervals of 5 ms.

TABLE II: *Activity Recognition Accuracy for ESR (%)*

Radar 1	Radar 2	Radar 3	Radar 4	Sub-activities
91.3	60.87	86.96	69.57	Spraying Wiping
73.68	85.96	77.19	98.25	Brushing Washing Hand Walking
68.42	84.21	78.95	84.21	Chopping Stirring Walking
92.98	100.0	91.23	80.7	standing dancing sitting

B. System Configurations

We evaluate the accuracy of the activity recognition system using three configurations of **RAM**.

Employing Single RAM (ESR): In this setting, we consider the data from an array of 4 **RAM**s individually and feed this data into our signal processing stack, feature extraction and classification pipeline. For example, in case of kitchen activities such as chopping, stirring and walking we use the data from all four RAMs separately to classify the activities.

Fusing Multiple RAM of Same Type (FMR-ST): In this setting, we aggregate data from multiple **RAM** devices of the same type (long-range). The data is time-synchronized. We selectively choose the **RAM** device based on the span of the activity (entire body or upper body movements) and the **RAM** system coverage and location. For example, in case of brushing the radars at the height of upper body extremity provide better signals than radar pointing at the lower body extremity.

Fusing Multiple RAM of Different Types (FMR-DT): In this setting, we fuse not only multiple **RAM** of same type (long-range) but different types of **RAM** devices, i.e., short and long range based on the activity and required sensor coverage. The small-range radar is tuned to detect movements at a shorter distance while long-range radar is suited to detect movements relatively at a larger distance.

C. Activity Classification Accuracy

In this section, we compare the accuracy results for the three **RAM** configurations described above. Overall the ESR strategy resulted in low accuracy as shown in Table II. The average classification accuracy was 78.95% for bathroom activities and 77.17% for cleaning activities. We note that the performance of radar 1 deteriorates for detecting short-range upper extremity body movements such as brushing (≈ 73.68%) compared to lower extremity movements such as spraying and wiping (≈ 91.3%).

TABLE III: *This table illustrates the precision and recall for activity classification when combining the short-range and long-range radars.*

Recall	Precision	Sub-activities	High-level Activity
0.76	0.62	Sitting	Entertainment
0.89	0.97	Dancing	
0.61	0.72	Standing	
0.61	0.55	Brushing	Bathroom Activities
0.55	0.88	Walking	
0.95	0.75	Washing Hands	
0.77	1.	Spraying	Cleaning
1.	0.81	Wiping	
0.97	0.69	Stirring	Cooking
0.58	0.49	Chopping	
0.39	1.	Walkin	

This attests that selecting radar sensors strategically based on the coverage, position, activity performance speed, body area movements may help improve the accuracy of activity recognition.

Table IV shows the multi-radar fusion results. Table V and Table VI present the precision and recall values for the same classifications. Combining radar 1 and radar 2 data is termed *Fusion 1* and combining radar 3 and radar 4 is termed *Fusion 2*. We time-synchronized multiple **RAM** data streams to form a single data source and fed that to the random forest classifier. We train our model using the aggregated datasets with the ground truth from the Kinect and performed cross-validation. From Table II, we note that the radar 1 gives 68.42% accuracy and radar 2 gives 84.21% accuracy but fusing data from both helps improve the accuracy of detecting chopping, stirring and walking activities to 98.25%. 88.3% accuracy is achieved for detecting spraying and wiping for *Fusion 1* due to the poor accuracy of radar 2 (≈ 60.87%) in detecting these activities, whereas for *Fusion 2* based on a different set of radars, the multi-sensor **RAM** system provide an accuracy of 84.04%.

Table III shows results for the FMR-DT setup. The experiments were performed when the subject was at a distance of 1m from two radars (short-range and long-range). The table shows that combining these two radars does not provide high classification accuracy for most activities. In fact, it is worth noting (see Figure 3) that these radars individually perform well for activity recognition depending on the distance between the human and the radar. The results shows that the radars must be strategically placed depending on the size of the room. For example, in bathrooms that are smaller in size, the short-range radar must be used while the living area an array of Bumblebee radars are likely to perform better.

D. Effect of frame-length and overlap with multi-radar fusion

Varying the size of the frame for chunking the radar data and overlap during the classification process helps improve the accuracy of recognizing different activities. The frame length determines the chunk of data that corresponds to a single activity while the overlap is the overlap between frames when generating the data chunks. We performed a set of experiments to determine the optimal combination of frame length and overlap.

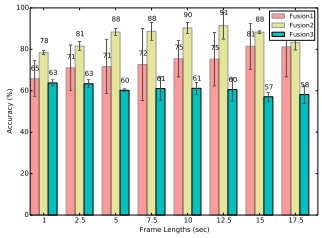


Fig. 7: *Accuracy with different frame-lengths for Bathroom activities.*

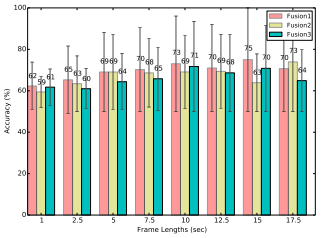


Fig. 8: *Accuracy with different frame-lengths for Cleaning activities.*

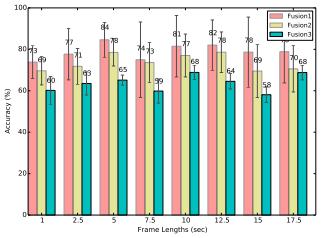


Fig. 9: *Accuracy with different frame-lengths for Cooking activities.*

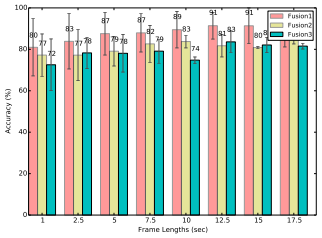


Fig. 10: *Accuracy with diferent frame-lengths for Entertainment activities.*

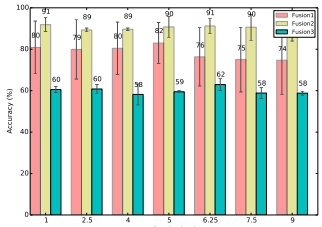


Fig. 11: *Accuracy with different window overlap for Bathroom activities.*

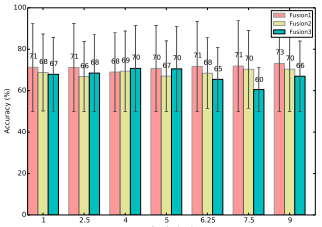


Fig. 12: *Accuracy with different overlap for Cleaning activities.*

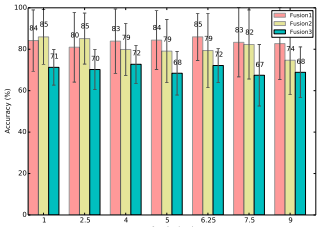


Fig. 13: *Accuracy with different overlap for Cooking activities.*

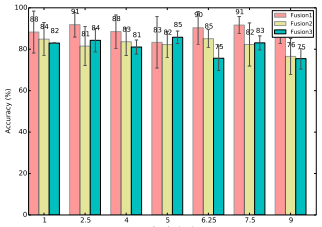


Fig. 14: *Accuracy with different overlap for Entertainment activities.*

TABLE IV: *Activity Recognition Accuracy (%) for Multi-Radar Data Fusion*

Fusion 1 (radar 1+2)	Fusion 2 (radar 3+4)	Fusion 3 (radar 1+2+3+4)	Sub-activities
88.3	84.04	87.37	Spraying Wiping
89.47	98.25	61.84	Brushing Washing Hand Walking
98.25	95.61	78.95	Chopping Stirring Walking
98.25	92.11	88.16	standing dancing sitting

TABLE V: *Precision values classifying activities when using multi-sensor fusion (of same type of radars).*

Fusion 1	Fusion 2	Fusion 3	Sub-activities	Activity
0.97 0.82	0.94 0.78	0.92 0.84	Spraying Wiping	Cleaning
1 1 0.95	0.89 1 0.87	0.9 1 0.77	Standing Dancing Sitting	Entertainment
0.82 0.86 1	0.97 0.97 1	0.43 0.42 1	Brushing Washing Hands Walking	Bathroom
0.95 0.97 1.00	0.88 1.00 1.00	0.95 1.00 0.95	Chopping Stirring Walking	Cooking

Specifically, we performed the experiments with the following **RAM** configurations: *Fusion 1* (radar 1 + radar 2), *Fusion 2* (radar 3 + radar 4), and *Fusion 3* (radar 1 + radar 2 + radar 3 + radar 4). Figures 7, 8, 9, and 10 showcase the activity recognition accuracy in case of *Fusion 1, Fusion 2 and Fusion 3* for bathroom, cleaning, cooking and entertainment activities respectively for frame length varying from 1 to 17.5 sec with an increment of 2.5 sec at each step. *Fusion 2*, a combination of radar sensors 3

TABLE VI: *Recall values for classifying activities when using multi-sensor fusion (of same type of radars).*

Fusion 1	Fusion 2	Fusion 3	Sub-activities	Activity
0.97 0.82	0.72 0.96	0.82 0.93	Spraying Wiping	Cleaning
0.95 1 1	0.87 1 0.89	0.72 1 0.92	Standing Dancing Sitting	Entertainment
0.87 0.82 1	0.97 1 0.97	0.46 0.39 1	Brushing Washing Hands Walking	Bathroom
0.97 0.97 1	1 0.87 1	0.88 0.49 1	Chopping Stirring Walking	Cooking

and 4 outperforms bathroom activity detection with an average accuracy of 91% for a frame length of 12.5 sec as shown in Figure 7 and overlap of 6.5 sec as shown in Figure 11. Similarly in case of recognizing cleaning, cooking and entertainment activities, *Fusion 1* outperforms all other combination of radar sensors. This set of results attest that while multiple radar fusion helps improve the accuracy of activity recognition it does depend upon the radar sensor position and specific activity under consideration. *Fusion 3* which is a combination of all 4 **RAM** performs poorly comparative to other multi-radar fusion options i.e., *Fusion 1 and Fusion 2*.

Figures 11, 12, 13, and 14 depict the activity recognition accuracy in case of *Fusion 1, Fusion 2 and Fusion 3* for bathroom, cleaning, cooking and entertainment activities respectively with overlap varying from 1 to 9 sec. *Fusion 2* outperforms the bathroom activity recognition accuracy with 91% accuracy for a overlap of 6.25 seconds. *Fusion 1* outperforms in all other cases, recognizing better the cleaning, cooking and entertainment activities as shown in Figures 12, 13, and 14 conforming the previous set of results obtained using varying frame length. We note that an appropriate combinations of frame length and overlap while boost the performance of the random for-

est classifier to detect a variety of fine-grained activities, an intelligent exploration of multiple RAMs fusion also helps increase the accuracy.

VII. Conclusion

In this work, we develop **RAM**, a non-intrusive low-cost heterogeneous ambient radar sensor system for activity recognition. **RAM** uses an array of micro-doppler radar sensors to detect hand, leg, upper extremity and lower extremity body-movements to recognize fine-grained activities of daily living (ADLs). We first investigate micro-Doppler radar signatures for recognizing finer body movements and develop signal processing techniques to relate different body movements with Doppler signal variations. We use those movements to recognize an ongoing high-level ADL episode, for example cooking, and cleaning. We have designed, implemented, and evaluated **RAM** in a real setting and show that our system can achieve better than 90% classification accuracy.

VIII. Acknowledgment

This material is based upon work supported by the National Science Foundation under awards CPS-1544687, CNS-1305099 and IIS-1406626, CNS-1308723, CNS-1314024, and the Microsoft SEIF Awards. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF or Microsoft.

References

[1] Antimo Barbato, Luca Borsani, Antonio Capone, and Stefano Melzi. Home energy saving through a user profiling system based on wireless sensors. In *Proc. of ACM SenSys (2009)*.

[2] Jürgen Nehmer, Martin Becker, Arthur Karshmer, and Rosemarie Lamm. Living assistance systems: An ambient intelligence approach. In *Proc. of ICSE (2006)*.

[3] Chung-Chih Lin, Ming-Jang Chiu, Chun-Chieh Hsiao, Ren-Guey Lee, and Yuh-Show Tsai. Wireless health care service system for elderly with dementia. *IEEE Transactions on Information Technology in Biomedicine 10,4 (2006)*.

[4] Yuvraj Agarwal, Bharathan Balaji, Rajesh Gupta, Jacob Lyles, Michael Wei, and Thomas Weng. Occupancy-driven energy management for smart building automation. In *Proc. of the ACM SenSys (2010)*.

[5] Upkar Varshney. Pervasive healthcare and wireless health monitoring. *In Mobile Networks and Applications 12, 3 (2007)*.

[6] P. Leijdekkers and V. Gay. Personal heart monitoring and rehabilitation system using smart phones. In *Proc. of ICMB (2006)*.

[7] Michael Buettner, Richa Prasad, Matthai Philipose, and David Wetherall. Recognizing daily activities with rfid-based sensors. In *Proc. of ACM UbiComp (2009)*.

[8] Sung-Nien Yu and Jen-Chieh Cheng. A wireless physiological signal monitoring system with integrated bluetooth and wifi technologies. In *Proc. of IEEE EMBS (2005)*.

[9] Randall S Stafford, John H Farhat, Bismruta Misra, and David A Schoenfeld. National patterns of physician activities related to obesity management. *The Journal of Archives of Family Medicine 9, 7 (2000)*.

[10] JK Aggarwal and Lu Xia. Human activity recognition from 3d data: A review. *Elsevier Pattern Recognition Letters (2014)*.

[11] Dany Fortin-Simard, J Bilodeau, Kevin Bouchard, Sebastien Gaboury, Bruno Bouchard, and Abdenour Bouzouane. Exploiting passive rfid technology for activity recognition in smart homes.

[12] Tuan-Jie Li, Meng-Meng Ge, and Gao-Wei Yuan. Human activity recognition using uwb radar and cameras on a mobile robot. In *Proc. of IEEE ICIEA (2009)*.

[13] Liang Liu, Mihail Popescu, Marilyn Rantz, and Marjorie Skubic. Fall detection using doppler radar and classifier fusion. In *Proc. of IEEE-EMBS BHI (2012)*.

[14] Philip van Dorp and Frans CA Groen. Human motion estimation with multiple frequency modulated continuous wave radars. *IET radar, sonar & navigation 4, 3 (2010)*.

[15] Andreas Bulling, Ulf Blanke, and Bernt Schiele. A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys (CSUR) 46, 3 (2014)*.

[16] Takuya Maekawa, Yutaka Yanagisawa, Yasue Kishino, Katsuhiko Ishiguro, Koji Kamei, Yasushi Sakurai, and Takeshi Okadome. Object-based activity recognition with heterogeneous sensors on wrist. In *Pervasive Computing (2010)*.

[17] Young-Seol Lee and Sung-Bae Cho. Activity recognition using hierarchical hidden markov models on a smartphone with 3d accelerometer. In *Hybrid Artificial Intelligent Systems (2011)*.

[18] Jan Meyer, Paul Lukowicz, and Gerhard Tröster. Textile pressure sensor for muscle activity and motion detection. In *Proc. of IEEE Wearable Computers (2006)*.

[19] Michael Otero. Application of a continuous wave radar for human gait recognition. In *Defense and Security*. International Society for Optics and Photonics, 2005.

[20] Youngwook Kim and Hao Ling. Human activity classification based on micro-doppler signatures using a support vector machine. *IEEE transaction on Geoscience and Remote Sensing 47, 5 (2009)*.

[21] Igal Bilik, Joseph Tabrikian, and Arnon Cohen. Gmm-based target classification for ground surveillance doppler radar. *IEEE transaction on Aerospace and Electronic Systems 42, 1 (2006)*.

[22] Graeme E Smith, Karl Woodbridge, and Chris Baker. Template based micro-doppler signature classification. In *IET seminar on High Resolution Imaging and Target Classification (2006)*.

[23] Chun Zhu and Weihua Sheng. Human daily activity recognition in robot-assisted living using multi-sensor fusion. In *Proc. of IEEE ICRA (2009)*.

[24] Piero Zappi, Thomas Stiefmeier, Elisabetta Farella, Daniel Roggen, Luca Benini, and Gerhard Tröster. Activity recognition from on-body sensors by classifier fusion: sensor scalability and robustness. In *Proc. of IEEE ISSNIP (2007)*.

[25] Takashi Matsuyama. i-energy: Smart demand-side energy management. In *Smart Grid Applications and Developments (2014)*. Springer.

[26] Ahmad Lotfi, Caroline Langensiepen, Sawsan M Mahmoud, and M Javad Akhlaghinia. Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behaviour. *Journal of Ambient Intelligence & Humanized Computing 3, 3 (2012)*.

[27] Jonathan L Geisheimer, William S Marshall, and Eugene Greneker. A continuous-wave (cw) radar for gait analysis. In *Proc. of IEEE Signals, Systems and Computers (2001)*.

[28] Bumblebee. https://samraksh.com/products/sensors/32-product-pages/products-sensors/71-bumblebee-radar.

[29] Zheng Li, Ryan Robucci, Nilanjan Banerjee, and Chintan Patel. Tongue-n-cheek: non-contact tongue gesture recognition. In *Proc. of ACM IPSN (2015)*.

[30] Victor C Chen, Fayin Li, Shen-Shyang Ho, and Harry Wechsler. Micro-doppler effect in radar: phenomenon, model, and simulation study. *IEEE Transactions on Aerospace and Electronic Systems 42, 1 (2006)*.

[31] Bahri Cagliyan, Cesur Karabacak, and Sevgi Zubeyde Gurbuz. Human activity recognition using a low cost, cots radar network. In *Proc. of IEEE Radar (2014)*.

[32] Yin-Wen Chang and Chih-Jen Lin. Feature ranking using linear svm. *Causation and Prediction Challenge Challenges in Machine Learning 2, 47 (2008)*.

[33] Klc2 radar. http://www.rfbeam.ch/.

[34] Spark core. https://www.particle.io/.