2024 23rd ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN) | 979-8-3503-6201-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/IPSN61024.2024.00035

Demo Abstract: *MARS* -An mmWave-based Multi-user Activity Tracking Solution

Argha Sen
IIT Kharagpur, India
arghasen10@gmail.com

Anirban Das*
NIIT University, India
anirbanfuture@gmail.com

Swadhin Pradhan
Cisco Systems, USA
swadhinjeet88@gmail.com

Sandip Chakraborty
IIT Kharagpur, India
sandipc@cse.iitkgp.ac.in

ABSTRACT

Developing robust wireless sensing mechanisms for continuously monitoring human activities and presence is crucial for creating pervasive interactive intelligent spaces. The existing literature lacks solutions that continuously monitor multiple users' activities without prior knowledge of the environment. This requires simultaneous localization and tracking of multiple subjects and identifying their activities at various scales, including macro-scale activities like walking and squats and micro-scale activities like typing or sitting. In this demo, we present *MARS* , a holistic system using a single off-the-shelf mmWave radar. *MARS* employs an intelligent model to sense both macro and micro activities and uses a dynamic spatial time-sharing approach to sense different subjects simultaneously. Our thorough evaluation demonstrates that *MARS* can continuously infer activities with over 93% accuracy and an average response time of approximately 2 seconds, even with five subjects performing 19 different activities.

KEYWORDS

mmWave, FMCW Radar, Multi-user Activity Recognition

1 INTRODUCTION

Living in an intuitively interactive space where interactions are natural and seamless has long been a vision, yet its realization remains elusive. To achieve this, we believe there is a need for multi-user continuous room-scale activity tracking through passive sensing. This demo aims to create an activity-sensing system to make indoor living spaces truly intelligent.

Key features of such a system include monitoring multiple subjects, track different activities over time, support both macro-scale (significant body movements) and micro-scale (minor body movements) activities, provide real-time activity inference, and enable continuous subject tracking. Existing works [2, 4] have made strides in these areas but often focus on a subset of these objectives.

To address these challenges, we propose *MARS* , an mmWave-based Multi-user Activity Tracking solution via Room-Scale Sensing system [3] that can track multiple users' activities in real time. We employ a novel technique that utilizes a single rotating mmWave radar to achieve continuous multi-user tracking and expand the Field-of-View (FoV). This innovative technique overcomes challenges such as occlusions without the need for multiple radars, simplifying the system, and avoiding complex interference patterns. *MARS* also uses differentiated stacking of range-doppler frames and opportunistic radar configurations to simultaneously detect macro and micro activities. Stacking the range-doppler frames enables the capture of spatiotemporal features of the received mmWave signals

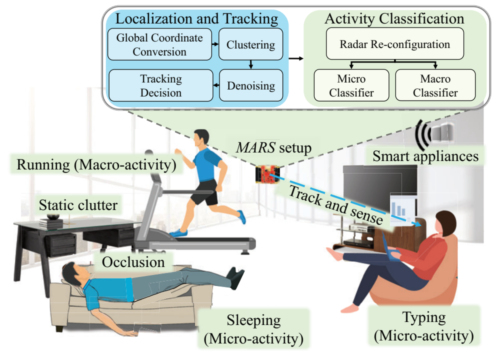


Figure 1: Overview of *MARS* .

per activity. Our system can monitor the highest number of human activities in the mmWave domain, making it a practical solution for real-world deployment.

We extensively evaluate and demonstrate the superiority of *MARS* over several baselines [1, 2, 4] in achieving high classification accuracy with low latency. Our work contributes to a practical and efficient system for continuous multi-user activity monitoring, bringing us closer to the vision of truly intelligent living spaces. A video demonstration of *MARS* is available online[1].

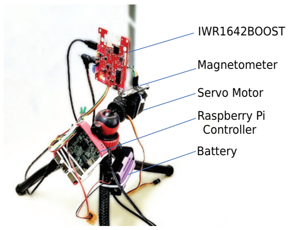
2 METHODOLOGY

MARS is a system designed to track multiple subjects performing various activities using a single mmWave radar. *MARS* solves two sub-problems: (i) localization and tracking of subjects; and (ii) monitoring the activity of individual subjects. To achieve localization and tracking, *MARS* utilizes point cloud data for subject detection and localization. It removes zero-valued Doppler bins to segregate static objects and generate a clutter-free point cloud, while a magnetometer maintains a global reference coordinate system for the radar, enabling tracking of subjects while the radar is rotating. Having a global context of the environment enables *MARS* to keep track of all the subjects even outside the radar's field of view. The clustering of point cloud data is performed using DBSCAN, with each cluster assigned a unique ID for subject tracking. A Kalman filter is applied to individual point cloud queues for each subject to estimate subjects' motion states and denoise random points due to occlusions or blind spots. Continuous tracking of subjects is ensured by using servo motors to rotate the radar for a 360°field of view. Activity classification is achieved using a Random Forest Classifier to predict subjects' activity scale (macro or micro) based

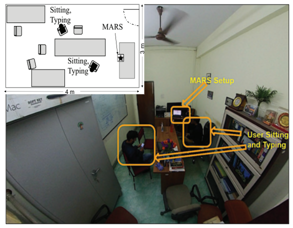
*The author was affiliated to IIT Kharagpur during this project.

[1]https://www.youtube.com/watch?v=Dxg98HU8yts

979-8-3503-6201-5/24/$31.00 ©2024 IEEE
DOI 10.1109/IPSN61024.2024.00035

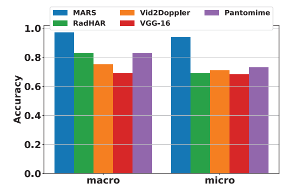


(a)

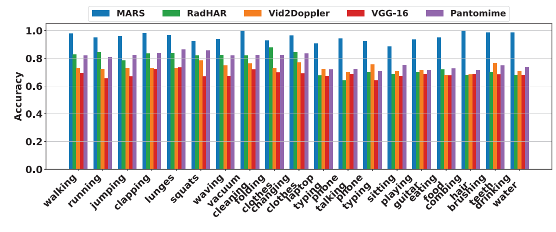


(b)

Figure 2: *MARS* setup and data collection in a room.



(a)



(b)

Figure 3: (a) Overall accuracy of *MARS* , (b) Accuracy across different activities.

on point cloud data. For this purpose, we capture the mean, standard deviation, kurtosis, and skewness in the denoised point cloud queue for each cluster for a time window of 1 seconds.

Different radar configurations are employed for macro and micro activities, using low and high Doppler resolutions, respectively. Activity signatures for individual subjects are segregated, and frames of range-doppler data are stacked to capture temporal features. A 2D Convolutional Neural Network (2D-CNN) is used for activity classification, with different configurations for macro and micro activities. Convolution 2D operation considers the dependency of neighboring spatial values and the temporal relationship of past t (t = Frames Per Second (FPS)) frames.

MARS is built on a Commercial off-the-shelf (COTS) millimeter-wave radar, `IWR1642BOOST`, tested in three rooms (R1, R2, R3) of different sizes. Ground truth activity is annotated from video captured by a USB camera. The system consists of a front-end radar and a backend `Raspberry Pi-4` for data processing. We train the classifiers on an iMac-M1 and deploy them on the Raspberry Pi-4 for live inference. The `IWR1642BOOST` radar uses two transmitters and four receivers with frequencies of 77-81 GHz. Different radar configurations are used for localization (FPS: 30, Doppler Resolution (D_r): 0.13 m/s), macro (FPS: 5, D_r: 0.13 m/s), and micro (FPS: 2, D_r: 0.01 m/s) activity classification. To achieve a 360° field-of-view, the radar is mounted on a `TowerPro MG995` servo motor. A `GY-273` magnetometer sensor is used for global coordinate transformation.

Data is collected from 7 subjects for 44 hours across 19 activities. Around 30 hours of the dataset are used for training. Ground truth for localization is generated by manual marking. Data collection includes controlled, semi-controlled, and uncontrolled scenarios. *MARS* is compared against RadHAR [4], Vid2Doppler [1], VGG-16, and Pantomime [2]. Evaluation metrics are based on the accuracy of classifying the activities.

3 EVALUATION

We evaluate the performance of the macro and the micro activity classifiers w.r.t. the baseline in terms of accuracy. As shown in Figure 3a, the accuracy for *MARS* is 97% in the case of macro activities and 93% in the case of micro activities. The lower accuracy for RadHAR is primarily because it relies on the point cloud dataset for the voxel formation and generates sparse point clouds in case of micro activities. Lower accuracy for Pantomime is due to its sparse point clouds, which can easily overlap with other subjects, and its radial velocity resolution of 0.87 m/s (see [2, Sec 5.1]), which is almost 9× our velocity resolution for macro configuration (0.13m/s). For Vid2Doppler, the poor accuracy is primarily because it takes only 32 Doppler bins, which are unsuitable for micro activity monitoring, and the model feature extraction part is pre-trained on macro activity datasets. As the body movements in the case of macro activities are significant, thus the classifier can segregate individual classes with excellent accuracy (close to ≈ 0.97). In the case of the micro activities, the body movements are less significant, but with the proposed classification pipeline with a higher Doppler resolution, we can achieve an accuracy of 0.93. Among the micro activities, laptop typing, eating food, and playing guitar involve higher body movements, and thus for these particular activities, we observe higher accuracy (see Figure 3b). Activities such as sitting, typing, and talking on a phone are carried out while subjects sit on a chair. Thus, the Doppler shift for these activities is very low. When the subject talks on a phone, the overlap with the *sitting* class is more significant (≈ 10%). In Figure 3b, we show activity wise accuracy of *MARS* w.r.t. the baselines. Although *MARS* supports a higher number of activities compared to the baselines (19 in *MARS* versus 5 and 12, respectively for [4] and [1]), the classification accuracy of the baselines significantly drops in comparison to *MARS* .

REFERENCES

[1] Karan Ahuja, Yue Jiang, Mayank Goel, and Chris Harrison. 2021. Vid2Doppler: Synthesizing Doppler radar data from videos for training privacy-preserving activity recognition. In *ACM CHI*. 1–10.

[2] Sameera Palipana, Dariush Salami, Luis A Leiva, and Stephan Sigg. 2021. Pantomime: Mid-air gesture recognition with sparse millimeter-wave radar point clouds. *ACM IMWUT* 5, 1 (2021), 1–27.

[3] Argha Sen, Anirban Das, Swadhin Pradhan, and Sandip Chakraborty. 2024. Continuous Multi-user Activity Tracking via Room-Scale mmWave Sensing. *ACM/IEEE IPSN*.

[4] Akash Deep Singh, Sandeep Singh Sandha, Luis Garcia, and Mani Srivastava. 2019. Radhar: Human activity recognition from point clouds generated through a millimeter-wave radar. In *ACM mmNets*. ACM, 51–56.