Wi-Fi CSI-Based Activity Recognition with Adaptive Sampling Rate Selection

Yuka Tanno(✉), Takuya Maekawa, and Takahiro Hara

Osaka University, Suita, Japan
{tanno.yuka,maekawa,hara}@ist.osaka-u.ac.jp

Abstract. Activity recognition methods using Wi-Fi Channel State Information (CSI) have been actively studied in the mobile and ubiquitous computing community. Many prior studies on CSI-based context recognition systems employ CSI data collected at a high and constant sampling rate, resulting in always high computation costs for context recognition. In this study, we propose a CSI-based activity recognition method that adaptively adjusts the sampling rate using reinforcement learning. In the proposed method, the "action" in the reinforcement learning is defined as the selection of a sampling rate of CSI, and the "state" is defined as an intermediate output of a neural network for activity recognition in the environment, which is expected to include information describing the complexity of the current activity. Moreover, we design an activity recognition model that can accept CSI inputs collected at an arbitrary sampling rate in principle, and extract sampling-rate-independent intermediate representations in its intermediate layers, enabling the reinforcement learning agent to switch to an appropriate sampling rate regardless of the current sampling rate. We evaluated the proposed approach using data collected in real environments.

Keywords: Wi-Fi CSI · Activity recognition · Reinforcement learning

1 Introduction

With the development of sensing technology, research on recognizing human activity by analyzing real-world sensor data has been actively studied. Activity recognition is a technology that estimates an activity class of a target person based on real-world sensor data, and can be applied to building various real-world applications such as home automation and surveillance of elderly people living alone. In addition to conventional methods using video data from cameras and/or accelerometer data from wearable devices [5,6], activity recognition methods using Wi-Fi Channel State Information (CSI) have been attracting attention in recent years [8]. Wi-Fi CSI contains information about propagation paths of Wi-Fi obtained from the physical layer between the transmitter and receiver that perform wireless communication, and is composed of the amplitude and phase of the complex channel information for each subcarrier between

T. Hara and H. Yamaguchi (Eds.): MobiQuitous 2021, LNICST 419, pp. 507–512, 2022.
https://doi.org/10.1007/978-3-030-94822-1_31

each transmit-receive antenna pair. The methods using Wi-Fi have the following advantages over the conventional methods. (i) They do not suffer from privacy problems caused by methods using camera images. (ii) They can recognize a target's activity even when the target is at the camera's blind spot. (iii) No wearable device is required. (iv) Commercially available Wi-Fi access points can be used to construct a recognition system.

Because of the limited transmission range of Wi-Fi, multiple CSI-based activity recognition systems (i.e., transmitter and receiver pairs) are installed in an environment, e.g., office and house. When the system is installed for each room, for example, the amount of CSI data to process is proportional to the number of systems, requiring significant computation costs when we assume large-scale deployment of the systems. Because the computation costs of always-on monitoring systems are crucial for achieving green ICT, an energy-saving approach that maintains precise activity recognition is necessary. Our solution proposed in this study is to adaptively control the sampling rate of the CSI-based activity recognition system according to a state of an environment in which the recognition system is installed. Prior studies on CSI-based context recognition systems employ CSI data collected at a high and constant sampling rate [4,7]. However, when a person in an environment of interest is sleeping, for example, the high sampling rate CSI data is not required to recognize the simple activity.

Therefore, in this study, we propose a CSI-based activity recognition method that adaptively adjusts the sampling rate according to the current state of an environment of interest. The proposed method employs reinforcement learning to adjust the sampling rate of CSI sensing by the recognition system.

2 Proposed CSI-Based Activity Recognition Method

2.1 Overview

In this study, we assume an indoor environment where a transmitter and a receiver of Wi-Fi are installed. The transmitter transmits packets and the receiver receives them and obtains CSI for each packet. The frequency of the transmitter's packet transmission corresponds to the sampling rate of the CSI in this study. The proposed method estimates an activity class of a person in the environment for each time window of data, and at the same time the method determines the sampling rate to be used for the next time window. The sampling rate of the CSI is either one of three: HIGH, MIDDLE, and LOW, 25 Hz for HIGH, 5 Hz for MIDDLE, 1 Hz for LOW in our implementation.

The proposed method consists of an activity recognition model and a sampling rate selection module, as shown in Fig. 1.

2.2 Activity Recognition Model

Here we introduce the activity recognition model used in the proposed method shown in Fig. 1. The activity recognition model is a neural network trained to

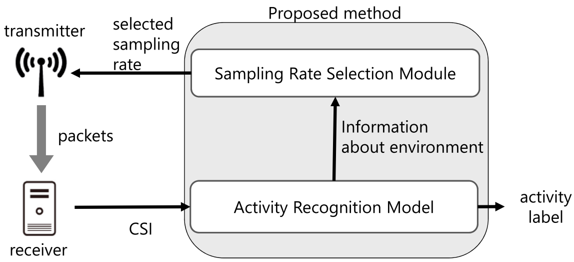


Fig. 1. Overview of proposed method

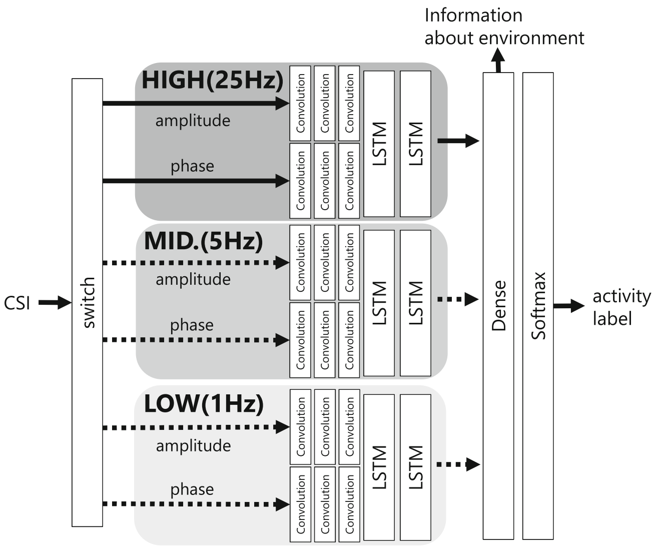
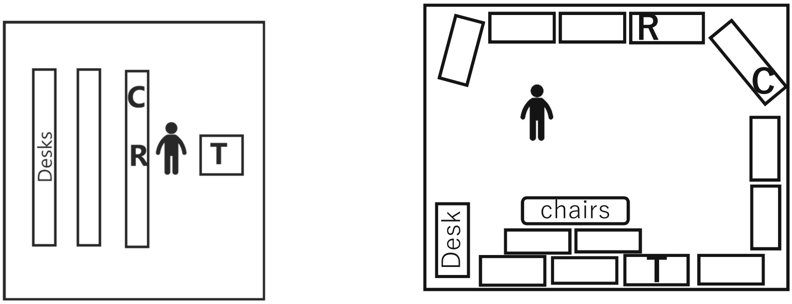


Fig. 2. Activity recognition model of proposed method

estimate an activity label of input CSI data. In the proposed method, the intermediate output of the activity recognition model is used as the input of the sampling rate selection module. The intermediate output of the activity recognition model can be regarded as features extracted from CSI data for activity recognition that contain information about the complexity of the current activity, and is useful for selecting the sampling rate. However, because the proposed method assumes CSI acquired at different sampling rates as an input, it is necessary to prepare a different activity recognition model for each sampling rate in general. Because the intermediate output of each model is considered to have different meanings, the intermediate outputs of the models cannot be simply used as an input of the sampling rate selection module. To address this issue, we design an activity recognition model that can process CSI data collected at different sampling rates as shown in Fig. 2.



(a) Dataset 1 (b) Dataset 2

Fig. 3. Experimental environments. T is the transmitter, R is the receiver, and C is the camera for acquiring ground truth labels.

2.3 Sampling Rate Selection Module

The sampling rate selection module is implemented based on a reinforcement learning model, Deep Q Network (DQN) [3]. In this study, the "action" in reinforcement learning is selection of a CSI sampling rate (HIGH, MIDDLE, or LOW), and the "state" is the information obtained from the observed CSI (intermediate output of the activity recognition model). In order to achieve recognition at a low sampling rate, a higher "reward" is given when the reinforcement learning agent can correctly recognize an activity at a lower sampling rate.

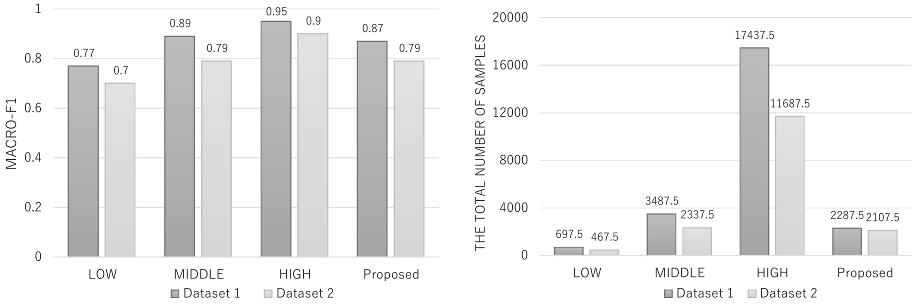
Note that, to efficiently train the neural network, we modify the original DQN in several ways according to [2].

3 Evaluation Experiment

3.1 Dataset

The datasets for the evaluation experiment were collected in two different environments; Dataset 1 and Dataset 2. Dataset 1 consists of 12 sessions of labeled CSI data acquired in the environment depicted in Fig. 3(a). In each session, a participant performed a set of seven activities: walk, tooth brushing, using vacuum cleaner, abdominal exercises, squat, jumping rope, and side-to-side jumping, in a random order. Dataset 2 consists of 12 sessions of labeled CSI data acquired in the environment shown in Fig. 3(b). In each session, a participant performed a set of five activities; walk, sleeping on chairs, using a laptop, squat, and juggling with a ball, in a random order. The duration of each activity is about 20 s.

The transmitter and receiver were PCs equipped with Intel 5300 NIC Wi-Fi modules. We installed the CSI tool published by Halperin et al. [1] on each of them and used it to acquire CSI data. The transmitter has three antennas, the receiver has one antenna, and the number of subcarriers is 30.



(a) F1 scores

(b) The total number of processed CSI samples per test session

Fig. 4. Results of the comparison methods and proposed method

3.2 Evaluation Method

To verify the effectiveness of the proposed method, we prepared the following comparison methods.

- HIGH: A method that constantly 25 Hz as the sampling rate
- MIDDLE: A method that constantly 5 Hz as the sampling rate
- LOW: A method that constantly 1 Hz as the sampling rate

In the experiment, we used 9-session data as training data for activity recognition, 1-session data as training data for reinforcement learning, and the remaining sessions as test data. For reinforcement learning, the 1-session training data was used by randomly changing the order of activities per episode.

Since the above methods output an activity estimate for each window, the macro-averaged F-measure (macro-f1 score) calculated with the estimates is used as an evaluation metric for recognition performance. In addition, the total number of CSI samples processed for recognition in the test data is used as a metric to evaluate the computational cost of recognition.

3.3 Results

Figure 4(a) shows the average F-measures of the comparison methods and the proposed method for Datasets 1 and 2. The F-measures of the proposed method were 0.87 for Dataset 1 and 0.79 for Dataset 2, which are almost identical to those of MIDDLE. The average numbers of total samples used in the processing of the comparison methods and the proposed method are shown in Fig. 4(b). The total numbers of samples of the proposed method were 2287.5 for Dataset 1 and 2107.5 for Dataset 2.

As shown the results of Dataset 1 in Fig. 4(a) and 4(b), the proposed method results in a reduction of the F-measure by only 2% compared to MIDDLE, while the number of samples processed was reduced by 35%. For dataset 2, the F-measure of the proposed method was identical to the F-measure of MIDDLE. In contrast, the number of samples processed was reduced by 10%. Therefore, we can say that the proposed method could significantly reduce the computational cost while maintaining the recognition accuracy.

4 Conclusion

In this study, we proposed a method to perform activity recognition with a lower sampling rate while maintaining high accuracy by adaptively adjusting the sampling rate according to the current state, and showed its effectiveness through evaluation experiments in real environments. As a part of our future study, we plan to construct a reinforcement learning agent that can work in any environments without the necessity of environment-dependent training.

Acknowledgement. This study is partially supported by JSPS JP16H06539 and JP21H03428.

References

1. Halperin, D., Hu, W., Sheth, A., Wetherall, D.: Tool release: gathering 802.11n traces with channel state information. Comput. Commun. Rev. **41**(1), 53 (2011). https://doi.org/10.1145/1925861.1925870
2. Hessel, M., et al.: Rainbow: combining improvements in deep reinforcement learning. CoRR abs/1710.02298 (2017). http://arxiv.org/abs/1710.02298
3. Mnih, V., et al.: Playing atari with deep reinforcement learning. Computing Research Repository, CoRR abs/1312.5602 (2013). http://arxiv.org/abs/1312.5602
4. Ohara, K., Maekawa, T., Matsushita, Y.: Detecting state changes of indoor everyday objects using wi-fi channel state information. In: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, IMWUT, vol. 1, no. 3, pp. 88:1–88:28 (2017)
5. Ravi, N., Dandekar, N., Mysore, P., Littman, M.L.: Activity recognition from accelerometer data. In: The Twentieth National Conference on Artificial Intelligence and the Seventeenth Innovative Applications of Artificial Intelligence Conference, pp. 1541–1546 (2005)
6. Ulutan, O., Rallapalli, S., Srivatsa, M., Manjunath, B.S.: Actor conditioned attention maps for video action detection. Computing Research Repository, CoRR abs/1812.11631 (2018). http://arxiv.org/abs/1812.11631
7. Wang, W., Liu, A.X., Shahzad, M., Ling, K., Lu, S.: Device-free human activity recognition using commercial WiFi devices. IEEE J. Sel. Areas Commun. **35**(5), 1118–1131 (2017). https://doi.org/10.1109/JSAC.2017.2679658
8. Wang, Y., Liu, J., Chen, Y., Gruteser, M., Yang, J., Liu, H.: E-eyes: device-free location-oriented activity identification using fine-grained WiFi signatures. In: The 20th Annual International Conference on Mobile Computing and Networking, MobiCom 2014, pp. 617–628 (2014). https://doi.org/10.1145/2639108.2639143