RF-ARP: RFID-based Activity Recognition and Prediction in Smart Home

Yegang DU*, Yuto LIM* and Yasuo TAN*
*School of Information Science, Japan Advanced Institute of Science and Technology
1-1 Asahidai, Nomi, Ishikawa, Japan
Email:{duyegang, ylim, ytan}@jaist.ac.jp

***Abstract*—Smart Home is generally considered to be the final solution for human living problem, especially for health care of the elderly and disabled, power saving, etc. Human activity recognition in smart home is the key to achieve home automation, which enables smart services automatically run according to human mind. Recent researches have made several progresses in this field, however most of them can only recognize default activities which is probably not needed by smart home services. In addition, low scalability makes such researches infeasible out of laboratory. In this work, we unwrap this issue and propose a novel framework to not only recognize human activity, but also predict it. The framework contains three stages: recognition after the activity; recognition in progress and activity prediction in advance. With the help of RFID tags, the hardware cost of our framework is low enough to popularize. And the experiment result shows that our framework can realize good performance in activity recognition and prediction with high scalability.**

***Index Terms*—Human Activity Recognition, Object Usage Sensing, Activity Prediction, RFID, Smart Home**

I. Introduction

Over the last few years, the Internet of Things (IoT) has been greatly developed with the help of mobile computing, edge computing and cloud computing. One of the most representative applications is smart home, which has good prospects in the future. In a typical smart home system, all the resources and devices can be controlled by the platform. The long term goal of smart home is to achieve automatically control according to both environment and inhabitant. Thanks to the advances of sensing technology, it is not difficult to obtain the environment data such as: illumination, temperature, humidity. However, there is still no practical solution to recognize human activity at home to enable scenario-based smart services [1]. To make the smart home platform know more about their host, human activity recognition (HAR) becomes an urgent challenge to the researchers.

Rather than online consumer activity, activity of daily living (ADL) usually can not produce any data for the computer system. Thus, this causes the gap between inhabitant and smart home system. To bridge this gap, existing work has shown us a bright direction. The interaction between human and devices could be the channel to recognize the activity. Wearable devices have already been widely used in recent years. Such devices equipped with different kinds of sensors and microprocessor are put on human body to monitor the state of human [2]. Represented by Apple Watch, smart watch and smart wristband have the ability to recognize simple activities like waving hands, sitting still, etc [3], [4], [5]. However, such activities do not contain semantic meaning, which is more appropriate to be called gesture recognition or action recognition [6]. These activities can not be directly used by the smart home system to provide scenario-based services. Another way to detect the interaction between human and devices is to attach sensors to ubiquitous objects used by human [7]. But, such smart electronic sensors rely on battery, so that the size of these smart sensors is not small enough to be attached to all the devices. Not to mention the cost of both maintain and the price of themselves. It is worth noting that passive RFID tags seem to have the ability to take place of such sensors. And several works have been proposed to prove that passive RFID tag is a good way to detect object usage [8], [9]. On the basis of these researches, our work gets even further results of activity recognition based on object usage.

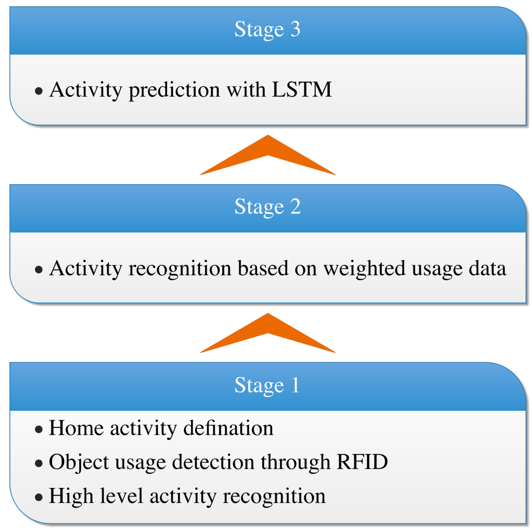


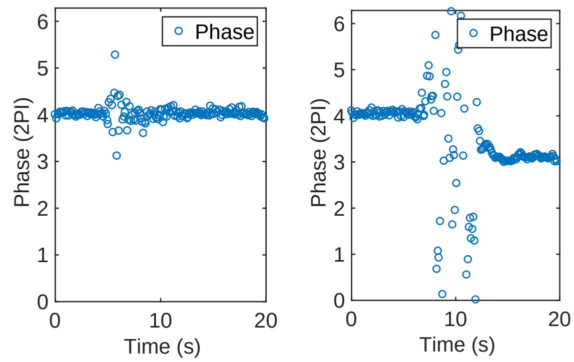
Fig. 1. Three-stage framework to recognize and predict human activity in smart home.

In this paper, we deeply analyse the characteristic of human activity in home environment firstly. This further clarified the goal of HAR in smart home, that is to provide scenario knowledge to the smart home platform to reach human centered automatic service. Then, we propose RF-ARP framework to recognize and predict the human activity in smart home, as shown in Fig. 1. Different from address resolution protocol

(ARP) that translates IP address into MAC address, our RF-ARP translates wireless RF signal to the human activity. The framework mainly contains three stages: recognition after the activity; recognition in progress and activity prediction in advance. In the first stage, we utilize passive RFID tags to detect the interaction between human and device and recognize the high level activity by combining those low level activities together. In this stage, we can make a record of what the inhabitant have done. And in the second stage, we weight the device by term frequencyinverse document frequency (tf-idf) to make sure the significance of each device to each activity. In this way, we will not have to give the recognition result after the activity has completed. While in the third stage, we already have the log of activity. Thus, we could use long short-term memory (LSTM) network to model the ADL of the inhabitant. So that the proposed framework will be able to predict the next activity that perhaps happens after the current activity. We finally test our framework with off-the-shelf equipment and open source database, and the effectiveness and efficiency of RF-ARP is proved.

To build such a system, there are several challenges we have to face. The first one is object usage detection without wearable devices. Existing work usually uses wearable RFID reader to detect object usage according to the distance between object and the hand of human [8]. While in our work, we use fixed long distance antenna to cover as more region as possible leading to the invalid of previous method. So we propose a way to detect object usage by phase which is a physical feature of RFID signal. The second challenge is concurrent activities recognition. We come up with a task-oriented generative approach rather than discriminative approach, so that the recognized result could be more than one activity. In addition, traditional machine learning methods rely on training data, which causes the so-called 'cold start' problem. Our approach utilizes the prior knowledge to define the activity, and then the training data is not required in the first stage. Thus, the upper stages could be in motion after stage 1 producing enough data.

Compared with existing work on activity recognition in smart home, our framework has multiple advantages on different aspects. First of all, scalability is the most important strength of our approach, which is reflected on two aspects. One is that we allow the smart home platform to define all kinds of activities as it needs. The other is that our approach cloud works in different houses, even when the objects and devices are different from each other. Besides, our activity prediction stage is going further than current HAR in smart home. This may largely promote the fully automatic smart home in the near future. The next strength is that our three-stage framework unwraps the task to recognize high level activity from wireless signal data. This brings huge flexibility, because every stage can be optimized independently or even replaced by other algorithms. For example, we can substitute LSTM in stage three with any other time series data mining algorithm, due to the labeled data provided by stage 1 can be used to train different models. Last but not the least, both the cost of RFID tags and computational complexity are low



(a) Passing by. (b) Picking up.

Fig. 2. Interactions and corresponding phase changes.

enough to implement our framework to current houses without much effort.

II. Design of RF-ARP Framework

In this section, we introduce in detail of our framework. The basic idea is to combine the devices with inherent semantic meaning and the interaction between human and devices to detect the object usage. And then we infer the higher level activity according the definition of each activity. With the log of recognized activities, we weight the devices with tf-idf to recognize the activity while it is in progress. And finally, we utilize the LSTM to model the habit of ADL, and predict the activity that the inhabitant might do later.

A. Object Usage Detection

Passive RFID tags are small and cheap enough to be widely used in logistics, warehousing, etc [10], [11]. And researchers have found that RFID tags can be used to detect the object usage by attaching them on everyday objects [12]. Several years before, Philipose et al. have used glove equipped with near-field RFID reader to do such work by the state of readable or not of tags [13]. While in our work, we chose long distance antenna and UHF reader to do the same task. The reason is that fixed long distance antenna can cover large area and scan all of the tags in the area almost at the same time. And some objects are not used by hand, such as chairs, beds, etc. Besides, we do not require the inhabitant to wear any devices, so that we also reduce the inconvenience to the inhabitant.

Phase is a back-scattered RF channel parameter, which can be continuously read by UHF RFID reader. In our work, we use Impinj R420 as the reader to obtain the phase value. According to our previous research, phase is sensitive to the interaction between human and RFID tags. As shown in Fig. 2, different interactions will change the phase value in varying degrees. In Fig. 2(b), there is a human passing by the object with a RFID tag. While in Fig. 2(a), the man walks to the object and picks it up, then puts it back and walks away. This inspires us to use the dispersion of the phase data in a sliding window to distinguish interactions.

TABLE I
OBJECT USAGE DETECTION THROUGH INTERACTION.

Usage	Tag state	Interaction	Objects
1	Covered	Sitting, lying, blocking	Chair, bed, sofa, switch, etc
	Picked up	Picking up	Knife, toothbrush, chopsticks, etc
0	Interfered	Passing by	All
	Still	Absence	All

Apart from the above interactions, people also interact with big furniture without moving them, like bed, sofa, etc. In these cases, the object usage can be detected by the phase all the same. UHF RFID readers are able to scan the tags several times in one second. In our verification experiment, the average sampling rate for each tag is 12 times per second. Although the reader can keep receiving the back-scattered signal when the tag is interfered, it can not see the tag while the tag is completely blocked. This enables us to use a simple way to detect such interactions, as shown by Eq. (1).

$$covered = if((t_0 - t_1) > T) \tag{1}$$

In the above equation, t_0 represent the current timestamp and t_1 represents the timestamp of previous round scanning. And T is the threshold for the tag state. In this work, we set T as one second to ensure enough sensitivity to detect short term interaction. For some specific objects like bed, we can increase T to detect the right interaction. Note that, when the tag works as a switch, one more step is needed to translate the interaction to equip state. The detail is introduced in our previous work [14].

So far we have introduced the way to detect object usage, and Table. I shows the way to determine the usage state. When the tagged object is covered or picked up, it means the object is being used and the usage state is set to '1'. Otherwise, when the tagged object is interfered or still, it means the object is not being used and the usage state is set to '0'. As depicted in Fig. 3, the matrix with white background color is the example of object usage array.

B. High Level Activity Recognition

To make sure the result of HAR is required by smart home platform, the best solution is to grant smart home platform the authority to define what it needs to know. Thus, we make a rule to enable such authority. The definition of an activity only includes the objects that will be used only in the activity. For example, television has the semantic meaning that greatly indicate the activity of 'watching TV'.

As shown in Fig. 3, after the usage states of objects have been detected, we build two queues to store the usage states. 'On queue' contains the object ID and the timestamp that the object starts to be used. While the 'Off queue' contains the object ID and the timestamp that the object ends using. In this stage, the length of the queues is set as fixed one day. It means

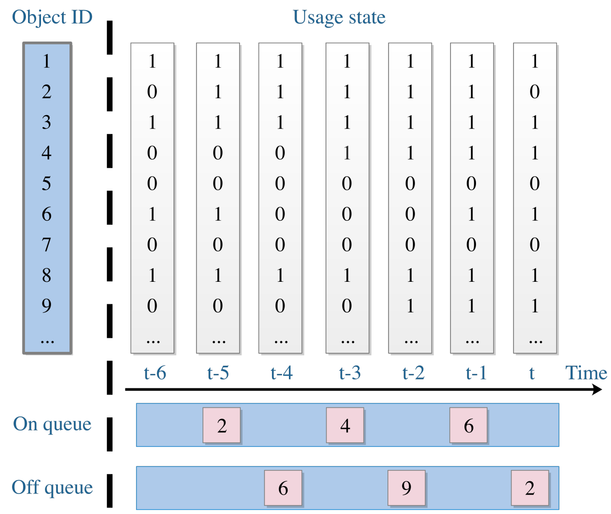
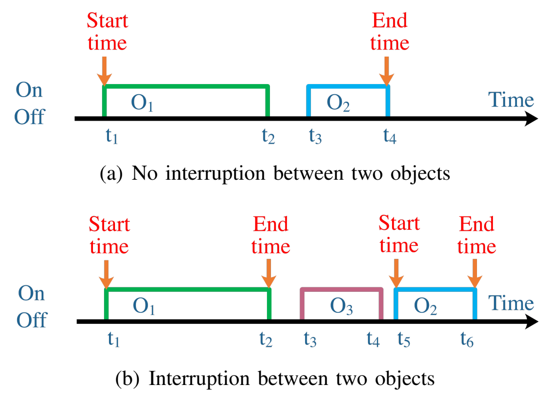


Fig. 3. The usage state vector changes with time, and we can generate 'On queue' and 'Off queue' respectively.



(a) No interruption between two objects

(b) Interruption between two objects

Fig. 4. We propose two strategies to determine the start time and end time. O_1 and O_2 are objects that belong to one activity, and O_3 belongs to some other activity.

that the task of this stage is to make a record of activities that have been done in the last one day.

Here, we can also take 'watching TV' as an example and object No. 2 represents television. At $t-5$, the television is turned on , and it is turned off at t. So, we can say that the activity of 'watching TV' has been acted, besides the start time is $t-5$ and the end time is t. However, this is merely the single object activity. And generally speaking, high level activity tends to include more than one objects. So, we set the start time as when the first object is used, and the end time is the last object ends using. However, when we do so, we can not get the right end time. Because, the activity may be interrupted by other activities. In this case, we are confused about when the activity is end.

To overcome the above problem, we treat the object in the definition individually as a subset of the whole objects set. And two strategies are proposed to determine the start time and end time, as shown in Fig. 4. In Fig. 4(a), there is no 'On' action between t_1 and t_3. It means that the activity is ongoing continuously, although the usage of O_1 ends before the usage of O_2 starts. So, we can merge the two subsets to a new subset. The start time is still the timestamp of the first object in the 'On queue'. And the end time is set to the timestamp of the last object in the 'Off queue'. When the next object belonging to the same activity start to be used, we check the interruption as the same. If there is still no interruption, then we keep merging the subset to the former subset and reset the end time. While when there is interruption, we use the strategy shown in Fig. 4(b). We can see that between t_1 and t_5, O_3 has been used at t_3. In this case, the activity is interrupted by other activity. We cut the relationship between the current subset and the former one that belongs to the same activity. The former subset has finished at t_2, so the end time is set to t_2. And the current subset becomes the initial subset, and the start time is t_5. While the end time is t_6 and may be rewritten by later subset.

C. Recognition in progress

In the first stage of our framework, we define the activity with the objects that only used in this activity. Even though it can record the log of activities that have been done efficiently, the precision of start time and end time is not so satisfactory. Besides, the definition of activity has limitations. There are more objects can not be involved in the definition, because they might be used by more than one activities. Thus, we come up with an approach to extend the definition of activity and recognize the activity with the new definition.

In the past research, tf-idf is commonly used in natural language processing and information retrieval [15], [16]. It is a numerical statistic that is intended to reflect how important a word is to a document in a collection or corpus [17]. Due to the ability to reflect the importance of element in frequency, we utilize the tf-idf to weight the objects in definition of activities. Before we calculate the tf-idf value of each object, we have to generate the training data. In the first stage, the interval between two activities is big, because of the too strict definition. So, we extend the start time of one activity to the end time of the former one and extend the end time of the activity to the start time of the last one. The objects used to exist between two activities are involved in two definitions of activities on both sides. In this way, the definition of activity is extended greatly.

Without loss of generality, we represent the set of objects O as:

$O = \{o_1, o_2, ... , o_n\}$

where n is the amount of objects. The set of activities A is denoted as:

$A = \{a_1, a_2, ... , a_m\}$

where m is the amount of activities. Also, we define the set of activities in the process $P \subseteq A$.

In the log data, we count the frequency of every objects o_i ($i \in n$) that have been used in a specific activity a_j and note it as g_i^j. Note that, we only care about that if the object has been used in one round of the activity, ignoring the number of times. If the object is used more than once in one round of activity, we still count one in one round. And to reduce the noise caused by the definition extension, we the equation below as a high-pass filter:

$$g_i^j = \begin{cases} 0, & g_i^j < z * max(g_i^j) \\ g_i^j, & otherwise \end{cases} \tag{2}$$

where z is a threshold set as 0.5 to control the filter in our work. If the higher z we set, the more strict definition we get. Particularly, if we set z to 1, the activity definition will be the same in stage 1.

The term frequency tf_i^j can be calculated by the equation below:

$$tf_i^j = \frac{g_i^j}{\sum_{i=1}^{n} g_i^j} \tag{3}$$

And the inverse document frequency idf_i^j can be calculated by Eq. (4) and Eq. (5)

$$idf_i^j = log(\frac{m}{\sum_{j=1}^{m} f_{i,j}(T)}) \tag{4}$$

$$f_{i,j}(T) = \begin{cases} 0, & g_i^j = 0 \\ 1, & otherwise \end{cases} \tag{5}$$

After we get both tf_i^j and idf_i^j, the $tf - idf_i^j$ then can be calculated as follows:

$$tf - idf_i^j = tf_i^j * idf_i^j \tag{6}$$

With the training data, we can finally generate a weight matrix that illustrates the importance of each object to different activities. Then, we utilize this matrix to realize online recognition.

Similar to the first stage, when a new object usage is detected, the object ID i ($i \in n$) is put into 'On queue'. Then we start to check the weight matrix to find the maximum $tf - idf_i^j$ and the corresponding j. In other words, the weight matrix tells the most possible activity, since this object is most representative to that activity. Then, we need to check the set of activities in the process P to verify if this activity is in P or not. If it is in the set P, we do not need to change anything and keep waiting the next object usage state change. If the activity is not in the set P, we then add the activity ID in P and note the start time of this activity with current timestamp.

When a new object is detected to end the usage, we also need to check the weight matrix with the object ID i. We retrieve the activities whose $tf - idf_i^j$ are not 0, then pull out all the relevant activities from the set P and set their end times respectively. In addition, if the next activity is a part of the activities that have just pulled out, we merge them together as the first stage does and set their end time to empty.

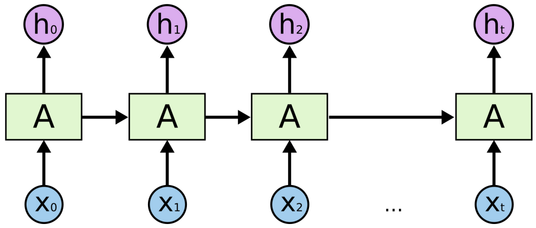


Fig. 5. Activity sequence and RNNs model.

D. Activity Prediction with LSTM

In this research, we treat human activity prediction as a time sequence predicting problem. We believe the inhabitants act different activities in a relative fixed pattern. For example, there is a user who always goes to watch TV after having dinner according to the activity log. If the user is detected to have dinner currently, then the next activity the user will act is most probably watching TV. Such problem to predict next state based on current state can be solved by classical machine learning method. Nevertheless, the next activity has relation with not only the current activity but also the previous ones. So we introduce deep learning to this problem. RNN performs well on spatial temporal predicting problems, such as location prediction [18]. LSTM networks are a special kind of RNN, which are proved to be more efficient [19]. LSTM networks have the ability to memorize both long and short term knowledge, which tally with human mind.

As depicted in Fig.5, this is a spread LSTM networks. X_0 to X_t in this paper represents the activity log, and h_0 to h_t represent the prediction result which is the next activity. We can see that when the time stamp is t, the input of the model is the current activity X_t and the past knowledge remembered from $t-1$ to $t-n$. It means the model can predict the next activity using not only the current activity but also the past several activities. This just accords with our assumption that the activity does not happen randomly and the motivation of next activity is what the human has done. In this way, the prediction accuracy is higher than the classical machine learning approach, because more knowledge is taken into consideration to model the habit of inhabitant.

In addition, we also apply the method in stage 2 to the process of prediction. Besides modeling the activity habit, we also utilize LSTM to model the object usage habit. Therefore, the next object that might be used can also be predicted by LSTM. Then, we find out the relevant activities of the object. Finally, we find the intersection of two prediction results to further improve the performance of prediction. Knowing the current activity and the current objects in use, we will be able to predict the next activity with a relative high accuracy.

III. Experiment and Evaluation

In this section, we show the performance of the three-stage framework to recognize and predict the activity of inhabitant. Since we unwrap the HAR task into three stages, we test performance of three stages respectively. To evaluate the part of activity recognition, we conduct the experiment on open source dataset. The dataset generated by Ordonez [20] includes 11 ADLs performed by the users on a daily basis in their own house for 35 days. We chose this dataset because most sensors in this dataset can be replaced by RFID tags to represent the usage as the same.

A. Stage 1

We attach two RFID tags to two commonly used objects: chair and toothbrush. And then, we ask the volunteer to do specified interaction with the objects 50 times respectively and note the tag states and corresponding usage states. Since 'interfered' and 'still' both represent no usage, we treat them as one interaction. In Table. II, we set that:

- TP represents that usage has been right detected;
- TN represents that interference has been right detected;
- FP represents that interference is detected as usage by mistake;
- FN represents that usage is detected as interference by mistake.

TABLE II
The result of object usage detection.

Objects	TP	TN	FP	FN
Chair	50	49	1	0
Toothbrush	49	47	3	1

Then, the average accuracy can be calculated as 97.5%. And the precision and recall are 96.1% and 99%. The performance is good enough to prove that RFID tags can be used to detect the object usage.

B. Stage 2

In the training part, we first classify the object usage data according to their corresponding activity ID, and each activity will contain several objects. Then, we calculate the weight of objects in specific activity to generate the weight matrix. In the testing part, we use the proposed approach to produce the log of activities. After that, we compare the log with the ground truth log of activities. If the label in ground truth log matches the recognized log, we take a count of TP for this labeled activity ID. If the ground truth activity a_p is recognized as other activity a_q, we take a count of $F_{p,q}$.

Here, for easy understanding, we use the verify matrix to represent the recognition result in this stage, as shown in Table. III. As shown in the table, FN is the sum of the row apart from the TP in this row, and FP is the sum of the column apart from the TP in this column. FN represents the false negative to the activity, and FP represents the false positive to the activity. Then, the precision and recall can be calculated by Eq. (7) and Eq. (8).

$$precision = \frac{1}{m}\sum_{j=1}^{m}\frac{TP_j}{TP_j + FP_j} \tag{7}$$

TABLE III
AN EXAMPLE OF THE VERIFY MATRIX OF TP, FN AND FP.

Activity ID	1	2	3	FN
1	TP_1	$F_{1,2}$	$F_{1,3}$	FN_1
2	$F_{2,1}$	TP_2	$F_{2,3}$	FN_2
3	$F_{3,1}$	$F_{3,2}$	TP_3	FN_3
FP	FP_1	FP_2	FP_3	-

$$recall = \frac{1}{m}\sum_{j=1}^{m}\frac{TP_j}{TP_j + FN_j} \quad (8)$$

According to the calculation, the average precision of our framework in stage 2 is 85.7%, and the average recall is 87.3%. In the experiment, we find that the main reason causes the false recognition is the temporal-sensitive activities. In the dataset, 'breakfast', 'lunch' and 'dinner' are treated as three different activities. Even though they are different in temporal space, the main objects related with them are similar. While our proposed framework does not take temporal knowledge into consideration, and this makes it hardly to distinguish those activities.

C. Stage 3

In the experiment, we build a typical LSTM model on TensorFlow-GPU with Keras as the high level API. The training epoch is set to 10000 to make sure the model is well trained. There are 4 layers in the LSTM model: one input layer, two hidden layer and one out layer. The loss is set $ascategoricalcrossentropy$, and the optimizer is $adam$. The $timestep$ and $neurons$ in hidden layers are hyperparameters. And we adjust the hyper parameters to make the model reaches the top performance.

We find that the test accuracy reaches to the top while the $timestep$ equals to 3. This means the LSTM model can utilize the past 3 activities to predict the next activity and get the highest accuracy, which accords with our assumption. Also, the accuracy starts to decrease after that. It means that the pattern of activities can not be too large, otherwise, too much noise will be used.

We also compare our model with the classical Naive Bayes method as shown in Fig. 6. Because the Naive Bayes method only uses the current one activity to predict next activity. We can see that our solution gets much higher accuracy than Naive Bayes. The top two prediction accuracy reaches to 65.2%. Moreover, when we apply the method in stage 2 to the process of prediction, the accuracy will be as high is 78.3%

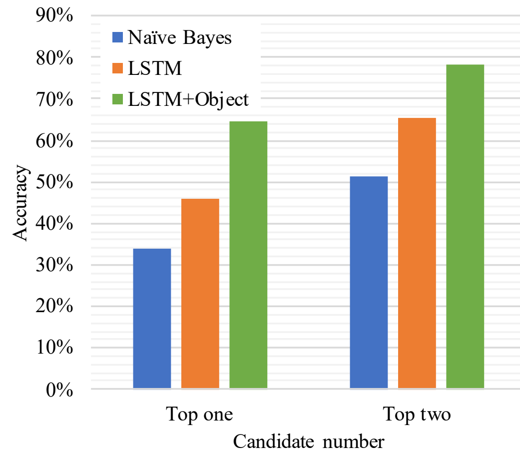


Fig. 6. Accuracy of Naive Bayes and LSTM solution.

IV. CONCLUSION AND DISCUSSION

In this work, we present RF-ARP which is a three stages framework to deal with the issue of HAR in smart home. According to the usage of objects, our framework can infer the high level activity and further predict the next possible activity. Without any requirement to the inhabitant, the proposed framework can be widely promoted to different house at a relative low cost of both money and energy. The framework is evaluated on an open source dataset of ADL. The recognition precision can reach 85.7% and the prediction accuracy is 78.3% in the condition of two candidates. Compared with existing work, our framework performs better.

ACKNOWLEDGMENT

This research is partly supported by the scholarship from China Scholarship Council (CSC) under the Grant CSC NO.201608050081.

REFERENCES

[1] S. Feng, P. Setoodeh, and S. Haykin, "Smart home: Cognitive interactive people-centric internet of things," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 34–39, 2017.

[2] F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou, and Y. Amirat, "Physical human activity recognition using wearable sensors," *Sensors*, vol. 15, no. 12, pp. 31 314–31 338, 2015.

[3] Y. Zou, J. Xiao, J. Han, K. Wu, Y. Li, and L. M. Ni, "Grfid: A device-free rfid-based gesture recognition system," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 381–393, 2016.

[4] S. Pradhan, E. Chai, K. Sundaresan, L. Qiu, M. A. Khojastepour, and S. Rangarajan, "Rio: A pervasive rfid-based touch gesture interface," in *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*. ACM, 2017, pp. 261–274.

[5] W. Jiang and Z. Yin, "Human activity recognition using wearable sensors by deep convolutional neural networks," in *Proceedings of the 23rd ACM international conference on Multimedia*. Acm, 2015, pp. 1307–1310.

[6] M. Ziaeefard and R. Bergevin, "Semantic human activity recognition: A literature review," *Pattern Recognition*, vol. 48, no. 8, pp. 2329–2345, 2015.

[7] 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 *International Conference on Pervasive Computing*. Springer, 2006, pp. 117–134.

[8] T. Gu, S. Chen, X. Tao, and J. Lu, "An unsupervised approach to activity recognition and segmentation based on object-use fingerprints," *Data & Knowledge Engineering*, vol. 69, no. 6, pp. 533–544, 2010.

[9] P. Palmes, H. K. Pung, T. Gu, W. Xue, and S. Chen, "Object relevance weight pattern mining for activity recognition and segmentation," *Pervasive and Mobile Computing*, vol. 6, no. 1, pp. 43–57, 2010.

[10] X. Liu, J. Cao, Y. Yang, W. Qu, X. Zhao, K. Li, and D. Yao, "Fast rfid sensory data collection: Trade-off between computation and communication costs," *IEEE/ACM Transactions on Networking*, vol. 27, no. 3, pp. 1179–1191, 2019.

[11] X. Liu, X. Xie, X. Zhao, K. Wang, K. Li, A. X. Liu, S. Guo, and J. Wu, "Fast identification of blocked rfid tags," *IEEE Transactions on Mobile Computing*, vol. 17, no. 9, pp. 2041–2054, 2018.

[12] X. Liu, X. Xie, S. Wang, J. Liu, D. Yao, J. Cao, and K. Li, "Efficient range queries for large-scale sensor-augmented rfid systems," *IEEE/ACM Transactions on Networking*, 2019 in press.

[13] M. Philipose, "Large-scale human activity recognition using ultra-dense sensing," *The Bridge, National Academy of Engineering*, vol. 35, no. 4, 2005.

[14] Y. Du, Y. Tan, and Y. Lim, "Rf-switch: A novel wireless controller in smart home," in *2018 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*. IEEE, 2018, pp. 1–2.

[15] Y. Du, Z. Li, M. Stojmenovic, W. Qu, and H. Qi, "A low overhead progressive transmission for visual descriptor based on image saliency." *Journal of Multiple-Valued Logic & Soft Computing*, vol. 25, no. 1, 2015.

[16] Y. Du, Z. Li, W. Qu, S. Miao, S. Wang, and H. Qi, "Mvss: Mobile visual search based on saliency," in *2013 IEEE 10th International Conference on High Performance Computing and Communications & 2013 IEEE International Conference on Embedded and Ubiquitous Computing*. IEEE, 2013, pp. 922–928.

[17] A. Rajaraman and J. D. Ullman, *Mining of massive datasets*. Cambridge University Press, 2011.

[18] Q. Liu, S. Wu, L. Wang, and T. Tan, "Predicting the next location: A recurrent model with spatial and temporal contexts." in *AAAI*, 2016, pp. 194–200.

[19] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[20] F. Ordóñez, P. de Toledo, A. Sanchis *et al.*, "Activity recognition using hybrid generative/discriminative models on home environments using binary sensors," *Sensors*, vol. 13, no. 5, pp. 5460–5477, 2013.