When RFID Meets Deep Learning: Exploring Cognitive Intelligence for Activity Identification

Xiaoyi Fan, Fangxin Wang, Feng Wang, Wei Gong, and Jiangchuan Liu

Abstract

Cognitive communication and computing have seen deep penetration in many networking areas in the past decades. With the recent advances in big data analysis and deep learning, we have seen great potential toward exploring cognitive intelligence for a wide range of applications. A notable example therein is human activity recognition, especially through RFID. Existing RFID activity identification solutions are mostly designed for static or slowly moving targets, rendering them far from satisfactory. More importantly, we observe that they suffer serious performance degradation in typical indoor environments with multipath interference. In this article, we argue that the recent advance of deep learning brings new cognitive intelligence for human activity identification. We first review the literature and research challenges of multipath effects in indoor environments. Then we introduce an advanced RFID activity identification framework, DeepTag, which uses a deep-learning-based approach for activity identification in multipath-rich environments. DeepTag gathers massive phase information from multiple tags, and preprocesses them to extract such key features as pseudospectrum and periodogram. We feed the preprocessed signal power and angle information into a deep learning architecture that combines a convolutional neural network and long short-term memory (LSTM) network. Our DeepTag framework can well adapt to both tag-attached and tag-free activity identification scenarios. Our extensive experiments further demonstrate its superiority in activity identification in multipath-rich environments.

Introduction

In the past decade, cognition has received significant attention in modern communication, networking, and computing systems. State-of-the-art solutions (e.g., cognitive radio [1, 2]) have mostly focused on the use of cognition to improve the utilization of such system resources as wireless spectrum. With the recent advances in big data analysis and deep learning, we have seen great potential toward exploring cognitive intelligence from these resources for a wide range of applications. A notable example therein is human activity recognition.

Traditional solutions for human activity recognition rely on sensor- or device-based approaches [3], but the required sensors/wireless devices are often not negligible in both size and weight, which restricts the application scenario. Radio frequency identification (RFID) is a promising technology that can overcome those difficulties, with the advantages of low cost, small form size, and the batteryless feature. Basically, an RFID system consists of a reader and many tags, where tags can be activated and powered by the signal from the reader and also send signals back to the reader without extra batteries. One single RFID reader can operate thousands of tags at a time. For example, IKEA Canada has completed a solution that enables shoppers to purchase merchandise with the tap of a spoon that has a built-in tag, freeing shoppers from having to push carts or carry baskets around the store (IKEA Canada Engages Customers With RFID at Pop-up Store; HTTP://www.rfidjournal.com/articles/view?14719). Disney has built an RFID gaming system that can sense when the player is moving or touching objects attached with tags in near real time.

Even though RFID reveals many benefits, the information offered by today's RFID tags is still quite limited, and the typical raw data, namely, received signal strength indicator (RSSI) and *phase angle*, mostly target stationary reading scenarios. As such, existing RFID-based activity identification solutions are far from satisfactory. It is well known that RSSI readings almost have no change with small human activities [3], such as shaking hands; and the phase angle, although sensitive to activities, is hardly a reliable indicator. Moreover, real-world multipath-rich environments bring more challenges for current RFID-based activity identification approaches to be applied in reality. For example, a person is often occluded by furniture and other persons, resulting in the signals of tags possibly being deflected and taking multiple paths to arrive at the RFID reader. Therefore, the received raw signals are not accurate enough to directly reflect the corresponding activity.

However, we argue that multipath indeed brings rich information that can be explored to identify human activities. Both the backscattered signal power and angle are highly related to human activities, impacting multiple paths at

Xiaoyi Fan is with the University of British Columbia and also with Simon Fraser University; Fangxin Wang (corresponding author) and Jiangchuan Liu are with Simon Fraser University; Feng Wang is with the University of Mississippi; Wei Gong is with the University of Science and Technology of China and also with Simon Fraser University.

Digital Object Identifier: 10.1109/MWC.2019.1800405

1536-1284/18/$25.00 © 2018 IEEE

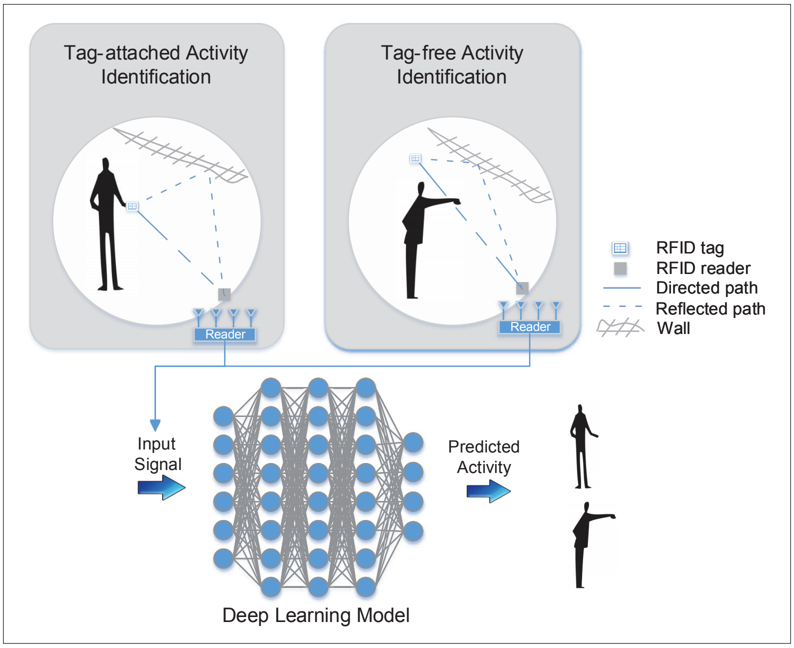

FIGURE 1. Two general approaches for activity identification with RFIDs.

different levels. If we can capture these changing features of the path, activities could be identified with high sensitivity. However, identifying the relevant features can be very time-consuming and complicated, and hence so is defining the rules for accurate classification of activities. Different from conventional solutions that directly rely on unreliable raw data, we develop a novel mechanism that jointly considers both pseudospectrum [4] and periodogram [5] to take in raw multipath signal mixtures and output carefully de-coupled angles and powers for different paths. We gather massive angle and power information from multiple tags, and preprocess them to extract key features. Yet another challenge is that due to the received signals being a dynamic mixture of multi-path, the features of RFID-based activities are hard to pre-select manually, and the rules for making correct estimations are hard to pre-defined as well.

In this article, we note that the recent advance of artificial intelligence brings new possibilities for the cognitive ability of RFID technology, empowering RFID communications with deep intelligence for accurate activity identification. Specifically, deep learning, as a new emerging generation of machine learning, can well accommodate this aforementioned challenge, and thus shed new light on the problem of activity recognition in a multipath-rich environment. To this end, we present a deep learning architecture, namely, DeepTag, that can take advantage of the RFID signal information from the pre-processing scheme and jointly use a convolutional neural network (CNN) and a long short-term memory (LSTM) network to solve the activity identification problem. We highlight that our DeepTag approach can well adapt to both *tag-attached* and *tag-free* activity identification scenarios, where the former directly attaches RFID tags to objects (e.g., the human body), and the latter only puts tags on fixed positions in the environment and thus makes objects tag-free. We conduct extensive experiments to evaluate our DeepTag in multipath-rich environments and report significant performance gains over a number of state-of-the-art feature-based approaches. It is also worth noting that our DeepTag is readily deployable using off-the-shelf RFID readers (e.g., a single UHF reader with a limited number of antennas) and allows reusing existing RFID readers for indoor activity identification.

The rest of this article is organized as follows. In the following section we present a brief literature review on activity identification with RFIDs. Then we discuss the multipath effects and de-coupling multipath signals. Following that, we present the design of our deep learning architecture for activity identification. Then we discuss the framework implementation details, with its performance being evaluated. We conclude this article in the final section.

Activity Identification with RFIDs

Figure 1 illustrates the two general approaches for activity identification with RFIDs. In tag-attached approaches, an RFID tag is attached to the human body, and the activities are then captured by a tag reader [6, 7]. Recently, tag-free approaches have also been suggested [8]. Instead of attaching tags to human bodies, which can be inconvenient and considered intrusive, multiple stationary tags are deployed in the environment as references, whose readings are expected to be affected by human activities in close proximity. Through analyzing the backscattered signals from the reference tags, the activities can then be identified. We next briefly review both approaches in the literature, as summarized in Table 1.

Tag-Attached Approach

This approach is the most straightforward, as RFID tags are suitable to be attached to objects due to their low cost and small size, and the batteryless feature. In activity identification, most solutions using this approach exploited the change of wireless signals incurred by human actions, and accurate localization techniques were often used to achieve the goal of activity identification.

Early RF-based localization primarily relied on RSSI information [3]. The major limitation of RSSI-based solutions is unreliability, since RSSI is insensitive to small body movement, and thus it is difficult to achieve high-precision identification. Recently, phase-based localization techniques have successfully achieved centimeter accuracy. For example, Tagoram [9] leveraged tag mobility to construct a virtual antenna array and built a differential augmented hologram using the phase values collected from the antennas. RF-IDraw [6] achieved good tracking accuracy with eight antennas connected to two RFID readers. FEMO [7] used the frequency shifts of movements to determine what exercise a user is performing. While such advanced solutions as RF-IDraw [6], Tagoram [9], and FEMO [7] achieve high accuracy through exploring antenna arrays, their performance degrades heavily for indoor environments, where multi-path reflections are prevalent and strong.

Solutions	Categories	Approach	Data	Technical Improvement
Tagoram [9]	Localization-based identification	Tag-attached	RF phase	Leverages the tag mobility to construct a virtual antenna array
RF-IDraw [6]	Localization-based identification	Tag-attached	RF phase	Employs multiple antennas (8) to eliminate this ambiguity
D-Watch [10]	Localization-based identification	Tag-free	RF phase	Utilizes both the direct path and the reflection paths to track targets
RFIPad [11]	Localization-based identification	Tag-free	RSSI & RF phase	Transforms a tag plane into a virtual touch screen
Twins [12]	Localization-based identification	Tag-free	Mutual inductance	Utilizes the coupling effect among passive tags
FEMO [7]	Direct activity identification	Tag-attached	RF phase	Uses the frequency shifts of the movements
Li et al., [13]	Direct activity identification	Tag-attached	RSSI	Presents a deep learning architecture
i2tag [14]	Direct activity identification	Tag-attached	RSSI & RF phase	Quantifies different levels of mobility and utilizes supervised learning framework
TASA [8]	Direct activity identification	Tag-free	RSSI	Deploys stationary tags as references and uses location-based activity identification
APID [15]	Direct activity identification	Ttag-free	Signal energy	Uses energy changes of backscatter signals

TABLE 1. Recent research on activity identification with RFIDs.

Tag-Free Approach

The tag-attached approach requires the target to be attached to a tag capable of emitting or reflecting RF signals. This, however, makes the approach not applicable in some scenarios. For example, in intruder detection, the targets will deliberately discard any device that can be tracked. In elder care, older people are usually reluctant to hold mobile devices, wear wearables or be attached with RFID tags. These real-life scenarios motivate the need for tag-free activity identification, which does not require any device to be attached to the target. Thus, tag-free activity identification has attracted extensive research interest recently.

In a tag-free configuration with stationary RFID tags being deployed in the environment as references (e.g., on walls or furniture), the communication link established with fixed readers can be disturbed by human activities in close proximity, hence changing RSSI or phase readings as well. Toward this direction, TASA [8] was proposed to rely on RSSI fingerprints, where reference tags are deployed in a regular way on a monitoring region and training data are generated during the training phase by requiring a person to act in different locations. In the testing phase, the resulting RSSI is mapped to the closest fingerprint to identify the status of the person. Such fingerprint-based methods, however, need a large amount of human effort to acquire and update the fingerprint database. Changes in the environment, such as the movements of furniture, will change the fingerprints, causing mismatches between the database and the new measurements.

Later, angle of arrival (AoA)-based schemes became popular with the opportunity of multiple antennas attached to a single RFID reader. The AoA of an RF source is computed by comparing the phases of the received signals at antennas. AoA estimation is widely used in RF-based positioning given the different propagation distances to different antennas, and serves as a foundation for activity identification [6]. D-Watch [10] efficiently utilized both the direct path and reflection paths to identify the angle information of the target. Twins [12] leveraged the coupling effect caused by interference among passive tags to detect a single moving subject. APID [15] was proposed to detect arm reaching by analyzing backscatter signals from a passive RFID tag. RFIPad [11] transformed a tag plane into a virtual touch screen by analyzing the induced disturbance of RF signals. However, AoA-based schemes may still suffer similar performance degradation when facing the challenges of the multipath-rich environment [14], as is further summarized in the next subsection.

Summary

It is easy to see that although RFID-based activity identification solutions can be categorized into tag-attached and tag-free approaches, the core technologies and the corresponding challenges are actually similar. For example, in real-world indoor scenarios, the received raw RFID signals are the dynamic mixture of many signals from multiple paths, which, as indicated in previous research work [13], may not be accurate enough to be immediately applied to activity identification. Recently, learning-based techniques have become a very active research area for general activity understanding. Toward this direction, i2tag [14] employed a supervised learning framework based on a fine-grained mobility profile, which can quantify different levels of mobility. Li *et al*. [13] proposed to directly apply a deep learning approach on collected coarse-grained RSSI readings to range about the tag for activity identification. Our DeepTag complements these works well by demonstrating the necessity and benefits of appropriate data preprocessing on mixed multipath signals and further proposing a deep learning architecture that can take full advantage of this and maximize the performance gain. In particular, we therefore propose to gather massive angle and power information from multiple tags, and preprocess them for key feature extractions. As the features of RFID-based activities are hard to pre-select manually, and the rules for making correct estimations are hard to pre-define, we further propose a deep learning architecture to handle such dynamics well and provide an activity identification framework in a multipath-rich environment.

Multipath Preprocessing

In practice, AoA estimation may not work well because of the multipath effect, which we address in this subsection. It is known that the estimation

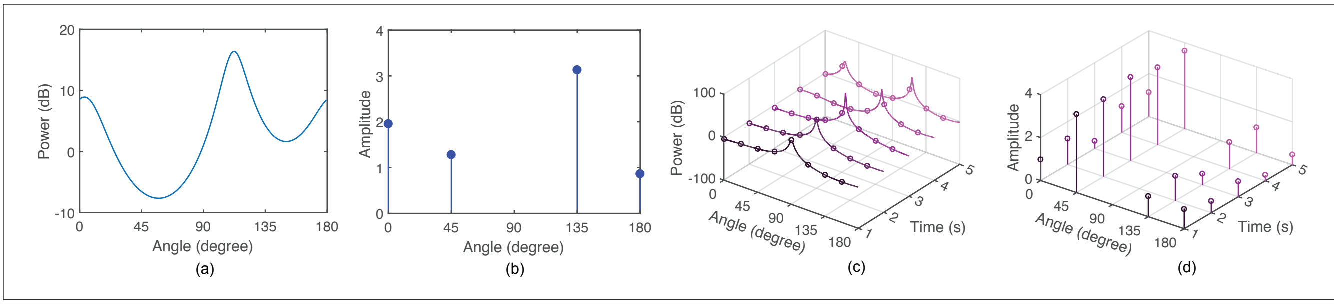

FIGURE 2. Illustration of data preprocessing and spectrum frame design: a) pseudospectrum estimation; b) periodogram estimation; c) pseudospectrum frames; d) periodogram frames.

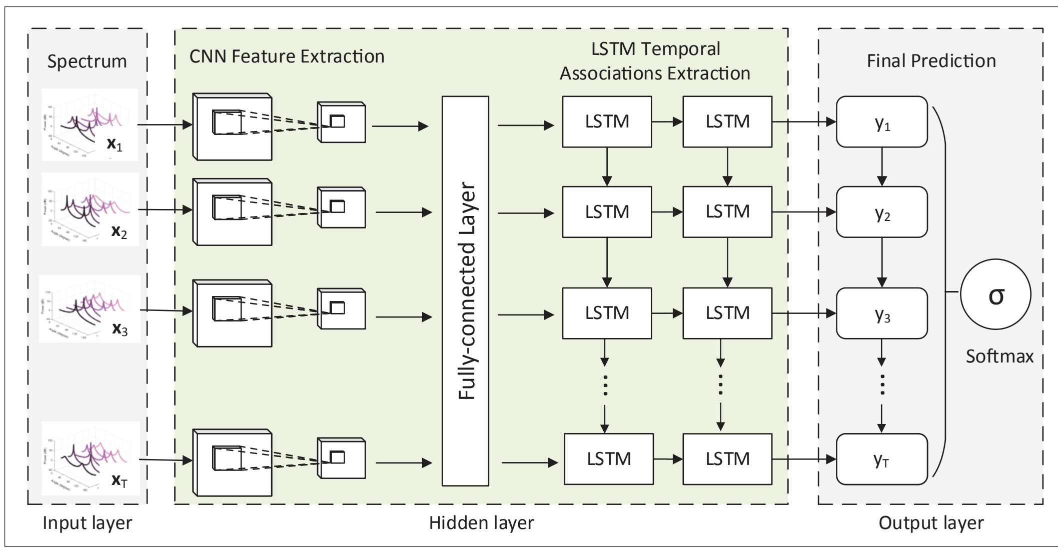

FIGURE 3. DeepTag network architecture.

is quite challenging in a multi-path indoor environment. To this end, we adopt the pseudospectrum [4] and periodogram [5] estimation to de-couple the signal multipaths.

Our pseudospectrum estimation design is mainly based on the Multiple SIgnal Classification (MUSIC) algorithm [4], which is one of the high resolution subspace AoA algorithms and was originally used to estimate the number of received signals by calculating the directions of arriving signals as well as their peak amplitudes. Figure 2a shows an example where the higher peaks are of great power, and each corresponds to an estimated AoA.

We introduce the periodogram [5] to strengthen pseudospectrum estimates by taking the accurate power information into consideration. In our system, we use fast Fourier transform to estimate the power distribution. According to Parseval's theorem, the Fourier transform is unitary, that is, the sum (or integral) of the square of a function is equal to the sum (or integral) of the square of its transform. Figure 2b shows an illustration, where we have four antennas connected to the RFID reader and thus can get four values in the periodogram for the power density distribution.

The multiple signals may also twist with each other and sometimes hide behind noises, so the relationships to human activities cannot easily be identified. All these call for solutions to dynamically identify and extract intrinsic features from the massive spectrum data with high accuracy. We accordingly introduce a deep learning design, which not only is effective in uncovering features for common activities, but also can scale up to identify more complex activities.

Deep Learning Design for Activity Identification

This section describes the main components of our DeepTag design. As illustrated in Fig. 3, our deep learning design takes the results from data preprocessing (i.e., periodogram and pseudospectrum) as inputs into our DeepTag network. We use an integrated design of a CNN and an LSTM network. The CNN has seen great success in the computer vision community, and the unique convolutional calculation is powerful in extracting the implicit spatial relationships in a single spectrum frame. LSTM is widely used in the speech recognition field, and it is able to learn dynamic temporal relationships from a sequence of spectrum frames. The output is the classification of object activities using a softmax layer. We next discuss each layer one by one.

The deep learning architecture starts with the design of our spectrum frames. The preprocessing stage outputs the spectrum for each tag, where we utilize the spectrum of all tags to build the spectrum frame. Specifically, we provide the following as input to the architecture:

- Pseudospectrum frames for AoA (as illustrated in Fig. 2c)
- Periodogram frames for power spectral density (as illustrated in Fig. 2d)

By combining these two types of information,

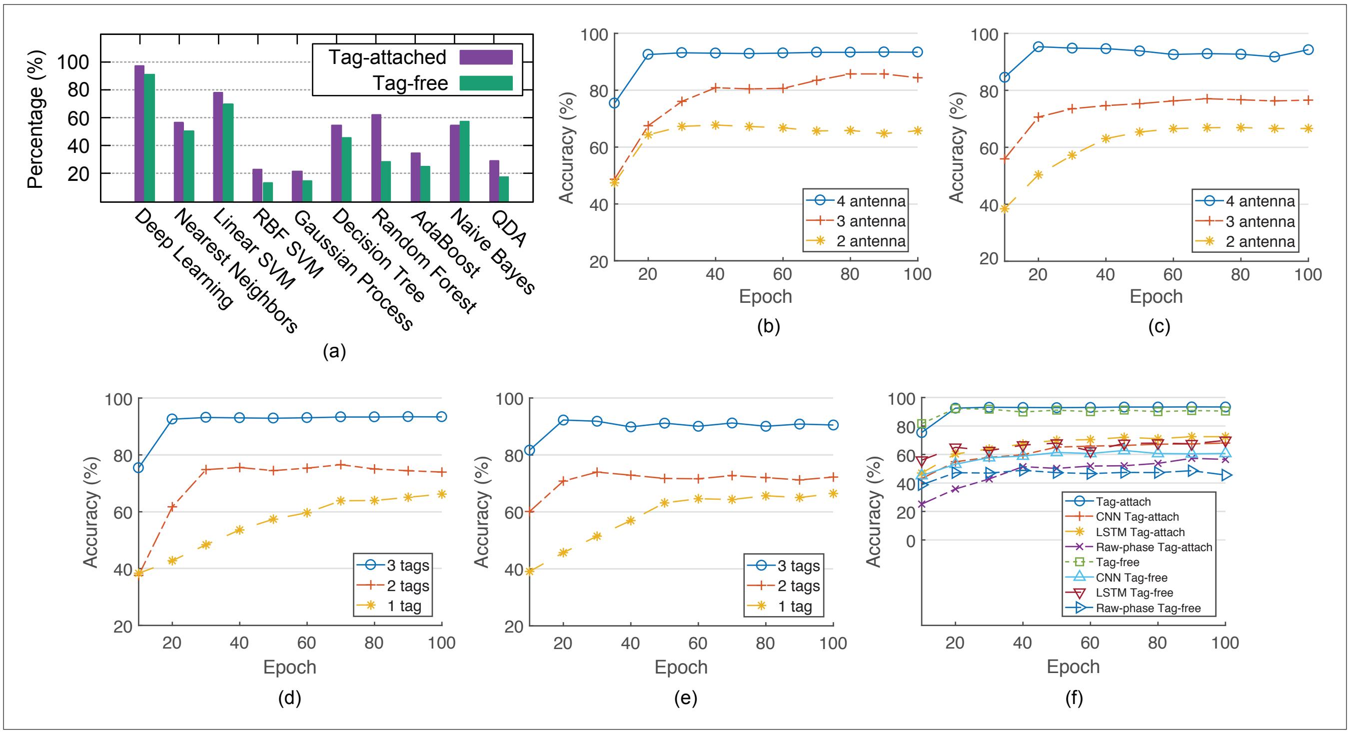

FIGURE 4. The performance evaluation of DeepTag when comparing with other classification approaches, varying the number of antennas, varying the number of tags, and varying learning networks: a) overall performance of DeepTag system; b) impact of number of antennas (tag-attached); c) impact of number of antennas (tag-free); d) impact of number of tags (tag-attached); e) impact of number of tags (tag-free); f) impact of different learning networks.

the architecture can take into account both the angle and power information of signals. The input layer then takes all the spectrum outputs from our preprocessing stage and builds the corresponding spectrum frame, where a series of spectrum frames along the time will further serve as the initial input for the hidden layer.

The hidden layer integrates a CNN structure and an LSTM structure. We construct a CNN to take the spectrum frames as input and provide the output to be fed into the LSTM structure, where the extracted lower dimension features then form the input as a sequence to the LSTM structure. In this work, we use the fully connected layer to merge all the inputs, where these features are outputs of rectified linear units. In our design, a stacked LSTM first encodes the frames one by one from the output of the CNN. LSTM is a subnet that allows the context information to easily be memorized for long periods of time in sequential data. The LSTM cells are then grouped and organized into a deep LSTM architecture. Inside the architecture, the output from one LSTM layer will be the input for the next LSTM layer. We use two stacked LSTM layers, each with 32 memory cells. Following the LSTM layers, a softmax classifier at the output layer is used to make a prediction at every spectrum frame, where the outputs from the last hidden layer are normalized with the softmax function.

System Implementation

Our design can be fully implemented based on a commercial reader and requires no modifications on tags. In this section, we further describe the key implementation details that are not covered in the previous sections.

Our prototype implementation uses an Impinj Speedway R420 reader (https://support.impinj.com/) without any hardware or firmware modification, which has been extensively used in the research community. The Impinj Speedway R420 reader has four antenna ports and is compatible with the EPC Gen2 standard, where the antennas work in a time-division multiplexing mode. The number of RF ports in the reader limits the scale of our antenna array, and we can increase the antenna number by Impinj antenna hubs. We set the typical wavelength λ to 0.32 m. We use Impinj tags, which are one of the cheapest tags available on the market and cost US$0.05 each.

The system employs a typical client-server architecture. The processes adopt Octane Java SDK with LLRP protocol to communicate with the reader, collect the readings, and upload them to the backend module. We utilize the multiple threads method, where a loop is used to execute the tag reading operation and immediately returns a sequence of RFID readings to the calling thread. The calling thread then uploads the tag readings to the server. The backend module on the server accepts the streaming of tag readings, where the server also stores the training data in the database and executes our algorithms to identify the activity. CNN and LSTM classifiers are implemented in Keras with a Tensorflow backend on Dual NVIDIA GeForce GTX 1080 Ti GPUs, and the multiclass classifiers based on machine learning tools are implemented based on the Scikit-learn library (http://scikit-learn.org/).

Evaluation

In the evaluation, we invite 10 volunteers and use three tags for each volunteer.[1] To conduct a comprehensive evaluation, we test seven scenarios, that is, standing, sitting, waving, bowing, walking,

[1] Note that these volunteers are varied in age, gender, height, and weight.

DeepTag de-couples the multipaths using the array of antennas; therefore, the number of antennas limits the number of multipaths that can be detected by our pre-processing scheme. With the information of more signal paths, DeepTag achieves a higher multipath density in the area and improves the activity identification accuracy.

running, and working. They are used to evaluate the robustness of DeepTag in tag-attached and tag-free scenarios for activity identification, respectively. In the tag-attached scenario, tags are attached to each person; in the tag-free scenario, the volunteers stand between the tags and the antenna array in our experiments. To evaluate the prediction quality, we run an experiment on the real-world data with about 6000 activity examples. With the training configurations mentioned previously, the total training time for the combined CNN and LSTM architecture is about 10 minutes, and our model can achieve stable accuracy with 500 iterations, which is not a large overhead. Besides, our system leverages a backend module to process the streaming data for inference to achieve real-time identification, which only depends on the sliding window size as the inference time is negligible.

Figure 4a shows the performance of our DeepTag compared to eight mainstream classifiers, including *k*-Nearest Neighbors, one-vs,-all linear support vector machine (SVM), one-vs.-all RBF SVM, Gaussian process, decision tree, Random Forest, AdaBoost, Bayesian net, and quadratic discriminant analysis (QDA), where all the classification methods take the same spectrum frame data as input. DeepTag achieves 94 percent accuracy of activity identification on average. In the tag-attached scenario, we can see that our DeepTag performs the best among all approaches with an accuracy of 97 percent on average, which is 20 percent better than the runner-up (SVM). In contrast, the classical machine learning methods (e.g., linear SVM) only have an accuracy of 77 percent or lower, reaffirming the effectiveness of our deep learning scheme. The underlying reason is that the deep learning architectures including CNN and LSTM are more capable of capturing the intrinsic but hidden features related to the activities from complicated signals, and distinguish them from that unrelated information. The experiments in a tag-free scenario show that the performance of other classifiers degrades heavily, and DeepTag still maintains a high accuracy of 91 percent on average. In the next experiments, we use different experiment parameters to investigate different tag number, antenna number, and deep learning architecture in tag-attached and tag-free scenarios, respectively.

DeepTag de-couples the multipaths using the array of antennas; therefore, the number of antennas limits the number of multipaths that can be detected by our pre-processing scheme. With the information of more signal paths, DeepTag achieves a higher multipath density in the area and improves the activity identification accuracy. We thus investigate the impact of the number of antennas, as shown in Figs. 4b and 4c for the tag-attached and tag-free scenarios. We can see that when the number of antennas increases from 2 to 4, more angle and power information of multipath can be detected, and thus DeepTag can achieve higher accuracy of activity identification.

With more tags, more signals will be produced to provide more information for activity identification. In the experiments, we vary the number of tags from 1 to 3 and investigate its impact. The results are shown in Figs. 4d and 4e, with respect to the tag-attached and tag-free scenarios. We can see that more tags are helpful to provide more information and improve the activity identification accuracy. Since the amount of multipath that our data pre-processing scheme can detect for each tag is limited by the number of antennas on the reader, the number of tags actually is the most effective and cheapest method to increase the path diversity in the environment.

We compare the results of our DeepTag with various combinations of deep learning architectures in both tag-attached and tag-free scenarios, as shown in Fig. 4f. For example, "CNN Tag-attach" indicates using a combination of the tag-attached sensing approach and the only CNN-based learning model; LSTM Tag-free indicates using a combination of the Tag-free approach and only the LSTM-based learning model; and Tag-free uses the CNN and LSTM learning architecture. The rule applies the same as for other definitions. First, we compare the performance of DeepTag with CNN networks, and both of them integrate the preprocessing scheme of DeepTag. DeepTag can achieve a 34 percent higher accuracy on average than CNN networks, which demonstrates that the LSTM architecture is necessary for activity identification. Then we evaluate DeepTag against LSTM networks. It clearly shows that DeepTag can achieve 22 percent higher accuracy than the LSTM networks on average, illustrating that CNN can efficiently extract the features for activity identification. Compared to the deep learning architecture that takes the raw phase data as inputs, DeepTag can also achieve an accuracy gain of 45 percent. In summary, the benefits of DeepTag come from both the preprocessing scheme and the deep learning architecture consisting of both LSTM and CNN, which work jointly to harvest the rich phase information for activity identification in a multipath environment.

Conclusion

RFID-based human activity identification is one of the most popular applications for cognitive wireless communication and computing. In this article, we further explore this topic and introduce DeepTag, a deep-learning-based RFID activity identification framework to tackle the challenges brought by real-world multipath-rich environments. In particular, we adopt the pseudospectrum and periodogram estimation to de-couple the signal multipaths for extracting the activity key features.

We further develop a deep learning architecture that combines the convolutional neural network and long short-term memory network to dynamically identify and extract intrinsic features from the massive multipath spectrum data for high-accuracy activity identification. Through extensive evaluations by our real-world prototype using the off-the-shelf Impinj reader and tags, we show the superiority of DeepTag on activity identification in multipath-rich environments for both tag-attached and tag-free scenarios.

Acknowledgment

This work is supported by a Canada Technology Demonstration Program (TDP) grant and a Canada NSERC Discovery Grant. Feng's research is partly supported by an NSF I/UCRC Grant (1822104).

References

[1] A. A. Khan, M. H. Rehmani, and A. Rachedi, "Cognitive-Radio-Based Internet of Things: Applications, Architectures, Spectrum Related Functionalities, and Future Research Directions," *IEEE Wireless Commun.*, vol. 24, no. 3, June 2017, pp. 17–25.

[2] M. Jia *et al.*, "Broadband Hybrid Satellite-Terrestrial Communication Systems Based on Cognitive Radio Toward 5G," *IEEE Wireless Commun.*, vol. 23, no. 6, Dec. 2016, pp. 96–106.

[3] A. Rai *et al.*, "Zee: Zero-Effort Crowdsourcing for Indoor Localization," *Proc. ACM MobiCom*, 2012.

[4] R. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," *IEEE Trans. Antennas and Propagation*, vol. 34, no. 3, 1986, pp. 276–80.

[5] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*, Prentice-Hall, 1997.

[6] J. Wang, D. Vasisht, and D. Katabi, "Rf-Idraw: Virtual Touch Screen in the Air Using RF Signals," *Proc. ACM SIGCOMM*, 2015.

[7] H. Ding *et al.*, "FEMO: A Platform for Free-Weight Exercise Monitoring with RFIDs," *Proc. ACM Sensys*, 2015.

[8] D. Zhang *et al.*, "TASA: Tag-Free Activity Sensing Using RFID Tag Arrays," *IEEE Trans. Parallel and Distributed Systems*, vol. 22, no. 4, 2011, pp. 558–70.

[9] L. Yang *et al.*, "Tagoram: Realtime Tracking of Mobile RFID Tags to High Precision Using COTS Devices," *Proc. ACM MobiCom*, 2014.

[10] J. Wang *et al.*, "D-Watch: Embracing 'bad' Multipaths for Device-Free Localization with COTS RFID Devices," *Proc. ACM CoNEXT*, 2016.

[11] H. Ding *et al.*, "RFIpad: Enabling Cost-Efficient and Device-Free In-Air Handwriting Using Passive Tags," *Proc. IEEE ICDCS*, 2017.

[12] J. Han *et al.*, "Twins: Device-Free Object Tracking Using Passive Tags," *IEEE/ACM Trans. Networking*, vol. 24, no. 3, 2016, pp. 1605–17.

[13] X. Li *et al.*, "Deep Learning for RFID-Based Activity Recognition," *Proc. ACM Sensys*, 2016.

[14] X. Fan, W. Gong, and J. Liu, "i2tag: RFID Mobility and Activity Identification Through Intelligent Profiling," *ACM Trans. Intelligent Systems and Technology*, vol. 9, no. 1, 2017, p. 5.

[15] H. Ding *et al.*, "Device-Free Detection of Approach and Departure Behaviors Using Backscatter Communication," *Proc. ACM Ubicomp*, 2016.

Biographies

Xiaoyi Fan [S'14] received his B.E. degree from Beijing University of Posts and Telecommunications, China, in 2013, his M.Sc. degree from Simon Fraser University, Canada, in 2015, and his Ph.D. degree from Simon Fraser University in 2018. He is now a postdoctoral research fellow in the Department of Electrical and Computer Engineering at the University of British Columbia, Canada. His areas of interest are the Internet of Things, big data, and deep learning.

Fangxin Wang [S'15] received his B.E. degree from Beijing University of Post and Telecommunications in 2013 and his M.E. degree from Tsinghua University, China, in 2016. He is currently pursuing a Ph.D. degree at Simon Fraser University. His research interests include the Internet of Things, wireless networks, edge computing, and deep learning.

Feng Wang [S'07, M'13, SM'18] received both his Bachelor's degree and Master's degree in computer science and technology from Tsinghua University in 2002 and 2005, respectively. He received his PhD degree in computing science from Simon Fraser University in 2012. He is currently an associate professor in the Department of Computer and Information Science at the University of Mississippi. He is a recipient of the IEEE ICME Quality Reviewer Award (2011) and the ACM BuildSys Best Paper Award (2018). He is a Technical Committee Member of Elsevier *Computer Communications*. He served as Program Vice Chair for the International Conference on Internet of Vehicles 2014, and as TPC Co-Chair for IEEE CloudCom 2017 for the Internet of Things and Mobile on Cloud track. He also serves as a TPC member for various international conferences such as IEEE INFOCOM, ICPP, IEEE/ACM IWQoS, ACM Multimedia, IEEE ICC, IEEE GLOBECOM, and IEEE ICME.

Wei Gong [M'14] received his Ph.D. degree from the Department of Computer Science and Technology at Tsinghua University. Currently, he is a professor with the School of Computer Science and Technology at USTC. His research interests include backscatter communication, distributed computing, and Internet of Things applications.

Jiangchuan Liu [S'01, M'03, SM'08, F'17] is a University Professor in the School of Computing Science, Simon Fraser University. He is an NSERC E.W.R. Steacie Memorial Fellow. He received his B.Eng. degree (cum laude) from Tsinghua University in 1999, and his Ph.D. degree from the Hong Kong University of Science and Technology in 2003. He has served on the Editorial Boards of *IEEE/ACM Transactions on Networking*, *IEEE Transactions on Big Data*, *IEEE Transactions on Multimedia*, *IEEE Communications Surveys & Tutorials*, and the *IEEE Internet of Things Journal*. He is a Steering Committee member of *IEEE Transactions on Mobile Computing*.

Through extensive evaluations by our real-world prototype using the off-the-shelf Impinj reader and tags, we show the superiority of DeepTag on activity identification in the multipath-rich environments for both tag-attached and tag-free scenarios.