RadioSense: Exploiting Wireless Communication Patterns for Body Sensor Network Activity Recognition

Xin Qi, Gang Zhou, Yantao Li, Ge Peng
Department of Computer Science
The College of William and Mary
Williamsburg, VA, 23187, USA
{*xqi,gzhou,yantaoli,gpeng*}*@cs.wm.edu*

***Abstract*—Automatically recognizing human activities in a body sensor network (BSN) enables many human-centric applications. Many current works recognize human activities through collecting and analyzing sensor readings from on-body sensor nodes. These sensing-based solutions face a dilemma. On one hand, to guarantee data availability and recognition accuracy, sensing-based solutions have to either utilize a high transmission power or involve a packet retransmission mechanism. On the other hand, enhancing the transmission power increases a sensor node's energy overheads and communication range. The enlarged communication range in consequence increases privacy risks. A packet retransmission mechanism complicates on-body sensor nodes' MAC layer and hence increases energy overheads.**

In contrast to the sensing-based solutions, we build RadioSense, a prototype system that exploits wireless communication patterns for BSN activity recognition. Using RadioSense, we benchmark three system parameters (transmission (TX) power, packet sending rate, and smoothing window size) to design algorithms for system parameter selection. The algorithms aim to balance accuracy, latency, and energy overheads. In addition, we investigate the minimal amount of training data needed for reliable performance. We evaluate our RadioSense system with multiple subjects' data collected over a two-week period and demonstrate that RadioSense achieves reliable performance in terms of accuracy, latency, and battery lifetime.

***Keywords*-Body Sensor Network; On-body Sensor Nodes; Communication Patterns; Activity Recognition**

I. Introduction

As the cost of wireless sensing devices declines, there is a trend to deploy body sensor network (BSN) as the solution for human-centric applications, such as personal health monitoring [1], assisted living [2], and physical fitness assessment [3]. As an important component, activity recognition enables BSN applications to be aware of user's activity. Many sensing-based solutions [4][5][6][7][8] have been proposed for activity recognition in BSN. These efforts collect sensor readings from on-body sensor nodes and utilize pattern recognition algorithms to recognize user's activity at an aggregator, such as a PC or a smartphone.

Although the effectiveness of the sensing-based solutions has been demonstrated in previous works, they face a dilemma. On one hand, packet loss undermines accuracy. Research on BSN link layer behavior [9] has demonstrated that the impermeability of human body causes the packet delivery ratio (PDR) to decrease. Packet loss reduces the availability of sensor readings at the aggregator. The incomplete sensor readings in consequence undermine classifier performance [10]. On the other hand, two strategies could be adapted to enhance the availability of sensor readings at the aggregator, but they have their own limitations. First, on-body sensor nodes may maintain their sending power at a high level (e.g., -15 dBm [9]) in order to achieve a satisfying PDR (e.g., 90% [9]). However, an increased sending power not only increases energy overheads, but also enlarges the communication range. As a result, private information within a BSN (e.g., acoustic sensor readings) could be overheard and analyzed by malicious entities. Although privacy could be enhanced by using a security protocol [11], the extra energy overheads cannot be ignored. Second, on-body sensor nodes may install a MAC protocol that contains a packet retransmission mechanism. This Quality of Service (QoS) component complicates the MAC protocol and increases energy overheads. In addition to this dilemma, sensing-based solutions also require the availability of various sensors [12][13][14], and continuous sensing also accounts for a large portion of energy overheads [15].

In contrast to the sensing-based solutions, we propose RadioSense - a prototype system that exploits wireless communication patterns for BSN activity recognition. Previous work [9] has found that the impermeability of human body causes packet loss and fading of transmitted signals. RadioSense is built upon the fact that *different activities have different patterns of packet loss and fading, which we call communication patterns.* RadioSense collects communication patterns from on-body sensor nodes and utilizes them for activity recognition.

RadioSense is built upon two dedicated on-body sensor nodes, which send simple fake packets to the aggregator at a low rate and power level. The aggregator extracts each node's communication features (such as PDR and the mean of Received Signal Strength Indicator (RSSI) [16] values) from arrival packets within a fixed time window. Features of all on-body sensor nodes form a communication pattern. The aggregator uses the communication pattern within each time window as a signature to identify the corresponding activity. In this paper, we demonstrate that RadioSense has the following features:

- **Low transmission power:** Through experiments, we

1052-8725/12 $26.00 © 2012 IEEE
DOI 10.1109/RTSS.2012.62

demonstrate that communication patterns are most discriminative among different activities when the transmission (TX) power is low (e.g. -28.7 dBm). Thus, in RadioSense, not only does a low TX power result in small energy overheads, but also high recognition accuracy. Another benefit of a low TX power is a small communication range. In experiments, we find that an on-body sensor node's largest communication range is less than 2 meters when the TX power is < -25 dBm. In RadioSense, on-body sensor nodes send fake packets within such a small range. As a result, privacy risks are reduced.

- **Preference for packet loss:** Through experiments, we demonstrate that packet loss is an important feature for achieving high recognition accuracy at low TX powers. Thus, in RadioSense, packet loss boosts accuracy rather than undermining it.
- **MAC layer simplicity:** RadioSense does not depend on a complicated MAC protocol with QoS assurance. A simple MAC layer entails low energy overheads.
- **Device simplicity:** RadioSense does not rely on any special sensors. Only a low power radio (e.g. CC2420 [17]) is required on each on-body sensor node. In contrast to the sensing-based solutions, energy overheads for obtaining and transmitting sensor readings are removed.

The main contributions of this paper are summarized as follows:

- We are among the first to reveal the discriminative capacity of communication patterns. Moreover, we propose RadioSense, a prototype system that collects communication patterns from on-body sensor nodes and exploits them for activity recognition.
- We benchmark three system parameters (TX power, packet sending rate, and smoothing window size) and provide algorithms for system parameter selection during the training phase. The algorithms aim to balance accuracy, latency, and energy overheads. In addition, we also investigate the minimal amount of training data needed for reliable performance.
- We evaluate RadioSense using data collected from three subjects at mutiple places with different background noise levels during a two-week period. The results demonstrate that RadioSense achieves reliable performance in terms of accuracy, latency, and battery lifetime.

The rest of the paper is organized as follows. We first discuss related works. Next, we present the potential of utilizing communication patterns for BSN activity recognition. Then, we propose a detailed design of RadioSense, after which we benchmark three system parameters to balance accuracy, latency, and energy overheads. Next, we demonstrate performance evaluation of RadioSense. Finally, we present conclusions and future work.

II. Related Work

Many works propose methods that utilize sensor readings for activity recognition. For example, PBN [5] is a practical activity recognition system that utilizes sensor readings collected from on-body sensors. CenceMe [18] utilizes sensors on smartphone for activity recognition. Rather than using sensor readings, our approach exploits the discriminative capacity of communication patterns for activity recognition.

There are a few works utilizing communication information for activity recognition. In [5], RSSI values are used as one type of sensor readings for practical activity recognition in BSN. In the paper, the authors use a default TX power level for on-body sensor nodes and illustrate that raw RSSI readings do not have strong discriminative capacity through K-L divergence analysis. However, rather than using raw RSSI readings at a default TX power level of on-body sensor nodes, our work automatically selects the optimal TX power level, at which we experimentally demonstrate the statistical features of RSSI readings are adequate for activity recognition. In [6], RSSI readings are used to differentiate between sitting and standing. In contrast, we design algorithms to automatically select the optimal system parameters (TX power level and packet sending rate) for on-body sensor nodes. With the selected parameters, we further explore the discriminative capacity of RSSI readings and combine RSSI features with packet delivery ratios (PDRs) to recognize more complex activities. In [19], the authors deploy wireless devices at critical spots within a room. The RSSI values collected from these devices are used to recognize both the room state (empty or not) and human activities (sitting, standing and walking). Our work uses more communication features collected from on-body sensor nodes, which work at the optimal system parameters, and aims to handle more complex activities.

In addition to activity recognition, communication information has been exploited for other applications. For example, radio tomography uses the fading of RSSI caused by human body for indoor motion tracking [20]. Radio signals have also been used for localization [21]. In contrast, we utilize different fading patterns caused by different human activities in BSN for activity recognition.

III. Motivation

In this section, we demonstrate the potential of utilizing communication patterns to recognize activities.

A. Experiment Setup

In experiment, we use Tmotes as on-body sensor nodes. Each Tmote is equipped with a CC2420 radio with 31 transmission power levels, which range from 1 (lowest) to 31 (highest). Although Tmote also has several sensors (e.g., light and temperature sensors), we only use its radio.

Figure 1. Setup.

In experiment, a user wears three Tmotes with CSMA enabled as depicted in Figure 1. The one on the right wrist and the one on the right ankle are used as on-body sensor nodes and highlighted with squares. Wrist and ankle are critical places for deploying on-body sensor nodes for activity recognition [5]. The third, which is used as a base station, is highlighted with a cycle. Each of the two on-body sensor nodes sends 4 packets per second at the TX power level 2 (-28.7 dBm). Each packet only contains the id of the node from which it is sent out. The base station receives packets from the two on-body sensor nodes, extracts the node ids, and sends messages containing the ids to a laptop through the USB interface. A software module on the laptop receives messages from the base station and logs the timestamp of each arrival message. Each run of this software module lasts for one minute, during which the user is required to perform one activity. There are three activities (sitting, standing, walking) to perform in total. Sitting is performed with legs under a table.

B. The Potential of Exploiting Communication Patterns

With the logged message arrival times, we plot the message arrival patterns of each activity in one minute in Figure 2. The x-axis represents the time point (in seconds) during this one minute. A spike at time t means that a message arrives at t.

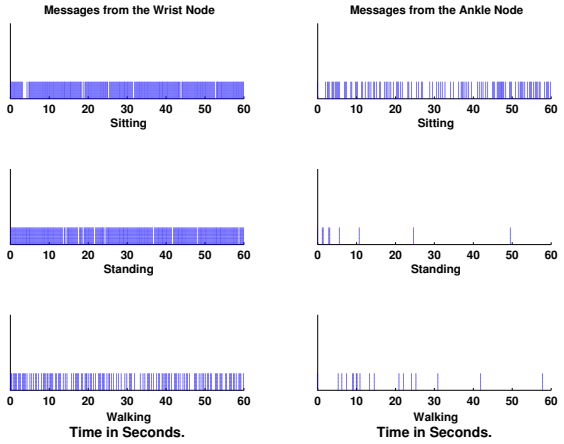

Figure 2. One Minute Message Arrival Patterns at the Base Station.

From Figure 2, we observe that the message arrival pattern of the wrist node is enough for distinguishing walking from the other two activities. In contrast, sitting and standing have very similar message arrival patterns of the wrist node. However, they can be distinguished from each other by the message arrival patterns of the ankle node. Thus, intuitively, these three activities are able to be separated from each other with the message arrival patterns of both the wrist and ankle nodes.

With only three activities being considered, we demonstrate the potential of exploiting communication patterns of the right wrist and ankle nodes for recognizing activities. However, it is just a concept proof. In the rest of this paper, we consider more complex activities.

IV. RadioSense Architecture

In this section, we introduce the architecture of RadioSense, a prototype system that exploits communication patterns for activity recognition. RadioSense mainly contains three components (as depicted in Figure 3), a sensor node working as a base station, a laptop working as an aggregator, and two dedicated on-body sensor nodes that have low power radios, such as CC2420. The base station is placed at the center of user's body. The on-body sensor nodes are placed at user's wrist and ankle, which are the critical positions to deploy sensor nodes for activity recognition [5]. The base station and on-body sensor nodes form a body sensor network. The base station is connected to the aggregator through the USB interface.

To unobtrusively recognize a user's activities, RadioSense must continuously collect communication information, from which communication patterns are extracted. The communication information collected by RadioSense includes the message arrival patterns as depicted in Figure 2 and the RSSI value of each arrival packet.

In RadioSense, each message contains two fields, $nodeid$ and $rssi$. Each on-body sensor node assigns its id number to the $nodeid$ field and sends messages encapsulated in packets periodically to the base station with the optimal parameters (TX power level and packet sending rate) obtained by the Training Module. On receipt of each packet, the RelayApp calls the MAC layer interface to measure the RSSI value of the received packet, extracts the message, and assigns the RSSI value to the $rssi$ field in the message. Then, the RelayApp module transmits the message containing the RSSI value to the aggregator through the USB interface. At the aggregator, the Time Stamper module logs the time and the RSSI value of each arrival message. It also logs the ground truth labels inputted from the GUI module during the training phase. The message arrival times of each on-body sensor node form its message arrival pattern.

A. Features and Classifier

The features used by RadioSense are all communication features extracted from the log, which is generated by the Time Stamper module. The log contains each on-body sensor node's message arrival pattern and the RSSI values of packets at runtime. It also contains the ground truth labels from the GUI module during the training phase. Smoothing window is the runtime time window within which features are extracted and classified. From message arrival patterns, the Feature Extraction module in Figure 3 extracts PDRs of on-body sensor nodes within each smoothing window based

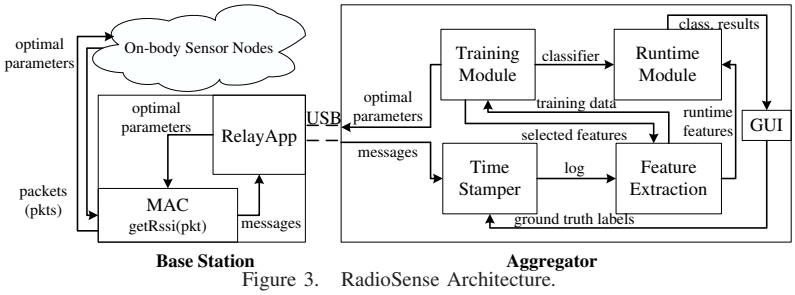

Figure 3. RadioSense Architecture.

on their current packet sending rate. From RSSI values of each sensor node's packets, the Feature Extraction module extracts 18 statistical features (the max, min, max−min, mean, var, median, mean crossing rate, values of the RSSI histgram with 10 bins, and iqr (interquartile range) of RSSI values) within each smoothing window. During the training phase, the Feature Extraction module extracts these features and forwards them with the ground truth labels to the Training Module. The Training Module uses the feature selection algorithm with sequential forward strategy [22] to select the best features and returns them to the Feature Extraction module. In addition, the Training Module trains a classifier with the selected features for the Runtime Module. At runtime, only the selected features are extracted by the Feature Extraction module and used by the Runtime Module to classify activities. The activity classification results are sent to the GUI module for display.

In RadioSense, we choose Support Vector Machine (SVM) combined with RBF kernel to differentiate user's activities. SVM is one of the best classifiers and has been successfully applied in many real-world classification problems, including activity recognition [15] and epileptic seizure detection [23]. In general, SVM has four advantages over other classifiers [24]: (i) the optimization involved in SVM corresponds to a convex optimization problem, whose local solution is also a global solution; (ii) it is able to achieve high accuracy with a relatively small number of training examples; (iii) it scales well with data dimensionality; (iv) it is in a simple form and hence fast to execute in runtime. These four properties make SVM a promising fit for activity recognition.

B. Training Module

In addition to the runtime features and classifier, the Training Module also generates the optimal parameters for on-body sensor nodes with the aim of balancing accuracy, latency, and energy overheads. As demonstrated by [9], on-body sensor nodes' TX power determines their PDRs and hence affects the communication pattern utilized by RadioSense. In addition, on-body sensor nodes' packet sending rate also has effects on the communication pattern. For example, a higher packet sending rate captures more information of the RSSI variation during each activity. Since accuracy depends on how discriminative the communication pattern is, different TX powers and packet sending rates that result in different communication patterns in consequence affect accuracy. Moreover, we do not expect on-body sensor nodes to use a high TX power and packet sending rate, since doing that quickly depletes the battery. Thus, one function of the Training Module is to determine the optimal TX power and packet sending rate for on-body sensor nodes. The optimal parameters should achieve high accuracy and low energy overheads at the same time.

In addition, the smoothing window size is a system parameter that balances accuracy and latency [23][25] at runtime. The Training Module also determines the optimal smoothing window size that achieves high accuracy and fast response at the same time.

V. Benchmarking System Parameters

In this section, we benchmark three system parameters (TX power, packet sending rate, and smoothing window size) with the data collected by RadioSense. According to benchmarking results, we propose algorithms for the Training Module to automatically select the optimal system parameters. During the training phase, the Training Module selects the optimal system parameters in the order of TX power level, packet sending rate, and smoothing window size. In addition, we also investigate the minimal amount of training data needed for reliable performance.

A. Data Collection

To benchmark the three system parameters, we collect data of seven activities (running, sitting, standing, walking, lying, riding, and cleaning) from one subject. Riding is performed on the cycling machine in gym. Unlike [25] that selects the optimal system parameters through benchmarking multiple subjects' data, we only use data from one subject because we experimentally find that accuracy is not a monotonous function of some system parameter, such TX power. Multiple subjects' data may blur the relationships between accuracy and system parameters. With one subject's data, we can observe the clear relationships between accuracy and system parameters and experimentally justify our observations, based on which we design algorithms to select the optimal parameters for individual subject. Compared to the one-for-all optimal parameters obtained in [25], our algorithms results in different optimal system parameters for different subjects.

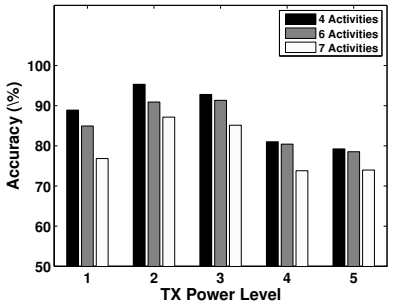

Figure 4. Recognition Accuracy vs. TX Power Level. Smoothing Window Size = 9 seconds, Packet Sending Rate = 4 pkts/s.

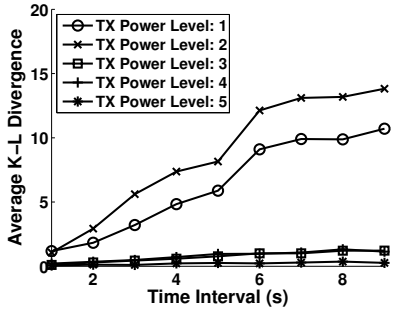

Figure 5. Average KL Divergence of PDR Distributions vs. Different Time Intervals. Packet Sending Rate = 4 pkts/s.

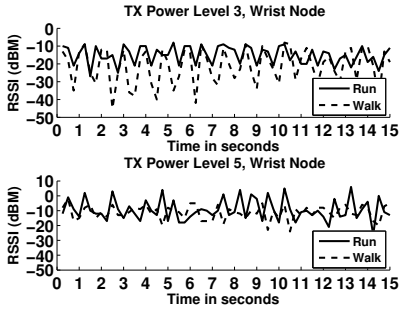

Figure 6. Snapshots of RSSI Values of Wrist Node's Packets. Packet Sending Rate = 4 pkts/s.

The subject wears on-body sensor nodes in the manner illustrated in Figure 1. Data is collected with different combinations of TX power level and packet sending rate. The range of TX power level in data collection is from 1 to 5. The range of packet sending rate in the data collection is from 1 pkt/s to 4 pkts/s. Thus, there are in total 20 parameter combinations. We collect data with 9 parameter combinations. For each parameter combination, both on-body sensor nodes are configured with the corresponding parameters, and each activity is performed for 30 minutes. To introduce the diversity of background noise levels (e.g. interferences and multi-path fading, etc), we collect data from different places, such as lab, classroom, living room, gym, kitchen, and outdoor. Each activity is performed at multiple places with different background noise levels. The data collection process lasts for 15 days in total.

We group the collected data into three activity sets. The first activity set contains four activities: running, sitting, standing, and walking. The second activity set contains two more activities: lying and riding. The third activity set contains one more activity: cleaning. Previous work [26] has identified cleaning as a difficult activity to classify. With these three activity sets, we investigate whether adding new activities into an activity set affects benchmarking results or not. We use the Java interface of libsvm [27] for the SVM classifier training and testing.

B. TX Power Level

Figure 4 depicts how accuracy changes as on-body sensor nodes' TX power level increases. In this figure, the smoothing window size and packet sending rate are set to be 9 seconds and 4 pkts/s, respectively. We follow the routine of 10-folds cross validation to estimate accuracies. In each round of cross validation, data is divided into 10 subsets, 9 of which are used for training and the remaining 1 is used for testing. This process is repeated 10 times and each of the 10 subsets is used exactly once as the testing data. The estimated accuracy is the average accuracy over 10 rounds [24].

We have two observations from Figure 4. First, the accuracy increases at the beginning and then decreases as the TX power level increases. TX power level 2, after which the accuracy decreases, is the optimal TX power level, because it has the lowest TX power that achieves the highest accuracy. Second, the above observation holds for all three activity sets. To better understand the first observation, we tabulate the feature selection results during the training phase in Table I.

TX Power Level	No. of Activities	PDR (Wrist)	PDR (Ankle)	No. of RSSI Features
1	4	√	√	3
	6	√	√	4
	7	√	√	5
2	4	√	√	3
	6	√	√	3
	7	√	√	5
3	4	×	×	4
	6	×	×	5
	7	×	×	5
4	4	×	×	4
	6	×	×	4
	7	×	×	5
5	4	×	×	4
	6	×	×	4
	7	×	×	5

Table I
SELECTED FEATURES AT EACH TX POWER LEVEL. SMOOTHING WINDOW SIZE = 9 SECONDS, PACKET SENDING RATE = 4 PKTS/S.

From table I, we observe that the PDRs of two on-body sensor nodes are selected for activity recognition when the TX power level is low (e.g., 1 or 2). As the TX power level becomes higher, the PDRs are no longer selected. It is because a human body's height is limited and a higher TX power level results in higher PDRs for different activities, but the PDRs are less discriminative when they are high for all activities.

Figure 5 supports the above explanation. In Figure 5, the average KL divergence over all activity pairs and all on-body sensor nodes are demonstrated at different TX power levels and time intervals. KL divergence is a metric that measures how different two distributions are [24]. A higher KL divergence value means the two distributions being compared are more different and vice versa. The x-axis in the figure is the time interval during which the PDR of each activity is calculated. To calculate each average KL divergence value in Figure 5, we first fix the TX power

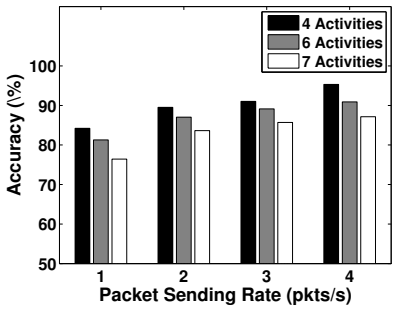

Figure 7. Recognition Accuracy vs. Packet Sending Rate. Smoothing Window Size = 9 seconds, Power Level = 2.

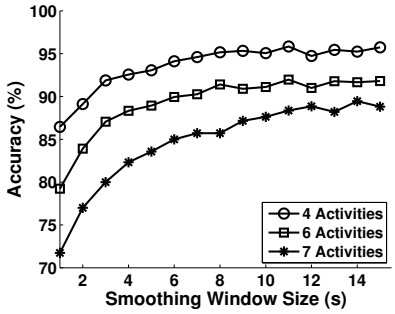

Figure 8. Recognition Accuracy vs. Smoothing Window Size. Power Level = 2, Packet Sending Rate = 4 pkts/s.

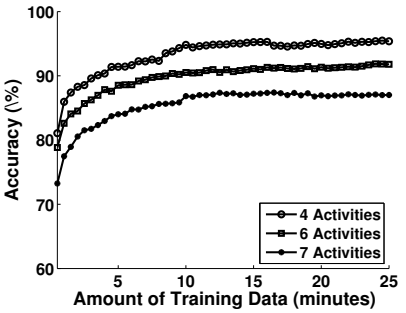

Figure 9. Recognition Accuracy vs. Amount of Training Data.

level and time interval. Second, for each on-body sensor node, the PDR distribution of each activity during 2 minutes is then computed. We empirically choose 2 minutes, the minimal total time that results in clear differences of average KL divergence among TX power levels. Third, for any pair of activities, we calculate the KL divergence of their PDR distributions. Fourth, the average KL divergence for this on-body sensor node is obtained by averaging the KL divergences of all activity pairs. At last, the average KL divergences of all on-body sensor nodes are averaged. From this figure, we observe that TX power level 2 has the largest average KL divergence, which means that the PDRs are most discriminative at this TX power level. In contrast, average KL divergences are much smaller at higher TX power levels. Thus, the PDRs are less discriminative at these power levels.

From table I, we also observe that when the TX power level is higher than 2, the feature selection algorithm only selects RSSI related features. As illustrated by Figure 4, accuracy decreases from TX power level 3. It can be inferred that the discriminative capacity of RSSI features is decreasing as the TX power level increases. Figure 6 supports this inference. Figure 6 illustrates the typical snapshots of RSSI values of the wrist node's packets at TX power levels 3 and 5. RSSI values of 15 seconds during running and walking are depicted. From this figure, we observe that the RSSI values are more discriminative between the two activities when the TX power level is 3. One possible reason for this observation is that the differences of human body's fading effects among different activities become smaller as the TX power level increases.

Among the selected RSSI features listed in Table I, most of them are the features extracted from the RSSI values of the wrist node's arrival packets. It can be inferred that the RSSI values of the wrist node captures more information for recognizing activities, because the wirst node is near the base station and has less packet loss.

From Figure 5 and 4, the average KL divergence is a metric that indicates the discriminative capacity of the PDRs. Moreover, we observe that when the PDRs are most discriminative, the highest accuracy is achieved. Thus, in our design, the Training Module uses the average KL divergence as a metric to select the optimal TX power level. In addition, Figure 5 indicates that the time interval of 8 seconds achieves the maximal average KL divergence difference among different TX power levels.

Therefore, we design the following steps for the Training Module to automatically select the optimal TX power level: (i) It configures the on-body sensor nodes with a relatively high packet sending rate (e.g., 4 pkts/s) because the optimal packet sending rate is determined after the optimal TX power level; (ii) For each TX power level, it collects message arrival patterns for 2 minutes and calculates the average KL divergence by using 8 seconds as the time interval; (iii) Starting from TX power level 1 and following its increasing order, as long as the average KL divergence of the current TX power level is smaller than that of the previous TX power level, the previous TX power level is selected, and the selection process stops.

C. Packet Sending Rate

Figure 7 depicts how accuracy changes as on-body sensor nodes' packet sending rate varies. In this figure, the TX power level and smoothing window size are set to be 2 and 9 seconds, respectively. Each accuracy in this figure is computed from 5-minute data of each activity following the routine of 10-fold cross validation. We empirically choose 5 minutes, the minimal total time that results in stable estimated accuracy during the cross validation process.

From Figure 7, we observe that the accuracy increases as the packet sending rate increases. It is because a higher packet sending rate captures more information of the RSSI variation during each activity. All activity sets share this observation. However, a higher packet sending rate means increased energy overheads. Thus, to balance accuracy and energy overheads, the optimal packet sending rate should be the rate above which there is no obvious accuracy improvement. In Figure 7, compared to the accuracy (for the seven-activity set) achieved at 3 pkts/s, we observe that the accuracy increasing rate becomes slow as the packet sending rate is increased to 4 pkts/s. It is reasonable to estimate that the packet sending rates, which are larger than 4 pkts/s could only achieve limited accuracy improvement.

Therefore, we design the Training Module to automatically select the optimal packet sending rate as follows: (i) the Training Module fixes the TX power level at the optimal one and collects data of each activity for 5 minutes at each packet sending rate; (ii) the Training Module selects the packet sending rate i pkts/s as the optimal one if: (a) the packet sending rate i achieves 90% accuracy with a relatively large smoothing window size (e.g., 9 seconds); (b) or $i >= 4$ pkts/s and the accuracy improvement achieved by the packet sending rate $i + 1$ is less than 2%.

D. Smoothing Window Size

Features extracted from a longer smoothing window are more robust to noise and hence result in higher accuracy. However, the delay of a longer smoothing window is larger compared to a shorter smoothing window. Thus, smoothing window size has to balance accuracy and latency.

Figure 8 depicts how accuracy changes as the smoothing window size varies. To get this figure, we use 10-minute data of each activity and set the TX power level and packet sending rate to be 2 and 4 pkts/s, which are the optimal parameters for the subject. The accuracies are computed following the 10-fold cross validation routine. From Figure 8, we observe that at the beginning the accuracy increases as the smoothing window size increases. It is because features extracted from a longer smoothing window are more robust to noise. The increasing rate becomes slow or near zero after the smoothing window size passes a threshold. It is because there is not much information gain by using a size which is larger than the threshold. In Figure 8, for example, the threshold is 10 seconds for the seven-activity set.

Therefore, we design the Training Module to automatically select the optimal smoothing window size as follows: the Training Module fixes the TX power level and packet sending rate at the optimal ones and uses the total training data (10-minute data of each activity) to select the smoothing window size i as the optimal one if: (a) the smoothing window size i achieves 90% accuracy; (b) or $i >= 10$ seconds (Here, 10-second is an acceptable latency for activity recognition [5].) and the accuracy improvement achieved by using the smoothing window size $i + 1$ is less than 2%.

E. Amount of Training Data

Intuitively, accuracy increases as the amount of training data increases, because the trained classifier can encode more information. However, labeling data is tedious. Thus, we also investigate the minimal amount of training data needed for reliable performance with the data of all three subjects (collected in Section VI). Different from benchmarking the three system parameters (Section V-B$\sim$V-D), we use the data of all subjects to obtain the minimal amount of training data for all subjects. Since it is meaningless to optimize the amount of training data individually for each subject during the training phase, we choose an optimal one for all subjects before the training phase.

Figure 9 demonstrates how accuracy varies with different amount of training data of each activity. The TX power level, packet sending rate, and smoothing window size are set to be the optimal ones for each subject and each activity set. The average accuracies of all three subjects are given for training data of each activity up to 25 minutes.

From Figure 9, we observe that the accuracy increases quickly in the first 10 minutes. After 10 minutes, the accuracy increases slowly. Thus, in the deign, the Training Module collects 10-minute data of each activity for training. With the smoothing window size of 8 seconds (the optimal smoothing window size for 6 activities of subject 1), there are about 75 instances of each activity during 10 minutes.

F. Summary

During the training phase, the Training Module follows the steps below to obtain the optimal system parameters: (i) Starting from TX power level 1, the Training Module collects 2-minute data of each activity to select the optimal TX power level. (ii) Then, starting from packet sending rate 1 pkts/s, the Training Module collects 5-minute data of each activity at the optimal TX power level to select the optimal packet sending rate. (iii) After selecting the optimal parameters for on-body sensor nodes, the Training Module collects 10-minute data of each activity. With the collected data, the Training Module searches the optimal smoothing window size for feature extraction. (iv) With the optimal smoothing window size and training data, the Training Module trains a classifier for the Runtime Module.

VI. Evaluation

In this section, we evaluate RadioSense from three aspects: accuracy, latency, and battery lifetime. In addition, we measure the communication range of RaidoSense and investigate whether RadioSense is able to coexist with other on-body sensor nodes.

A. Experiment Setup

Subject	Gender	Height (m)	Weigh (kg)
1	Male	1.80	85.0
2	Male	1.68	63.0
3	Female	1.56	48.0

Table II
SUBJECT INFORMATION.

In experiment, we implement RadioSense on three Tmotes and a laptop. We collect data of activities (running, sitting, standing, walking, lying, riding, and cleaning) from three subjects. The information of each subject is summarized in Table II.

Each subject wears sensor nodes in the manner depicted in Figure 1. During the training phase, data is first collected for selecting the optimal parameters for on-body sensor nodes. Then, each subject performs each activity for 10 minutes for the Training Module to select the optimal smoothing window

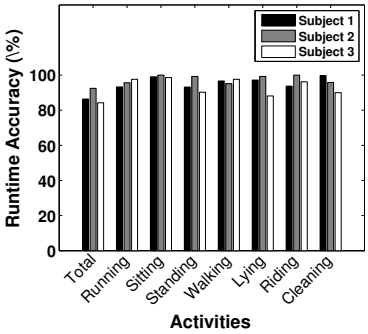

Figure 10. Activity Classification Accuracy.

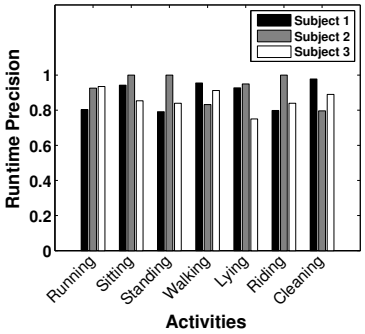

Figure 11. Activity Classification Precision.

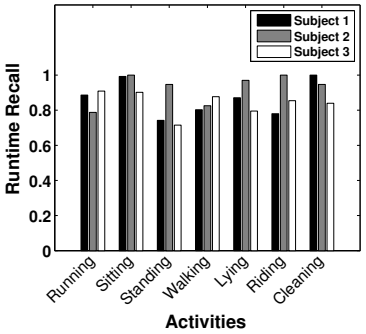

Figure 12. Activity Classification Recall.

size and train a classifier. To test the runtime accuracy, each subject performs each activity for 30 minutes. Each activity is performed in various environments (such as lab, classroom, living room, gym, kitchen, and outdoor) with different background noise levels in both training and testing phases. The whole testing process lasts for 2 weeks.

B. Accuracy and Latency

In this subsection, we first present the optimal system parameters selected by the Training Module for all subjects. Second, we demonstrate the runtime accuracy and latency of each subject. Finally, we demonstrate the accuracies of more activities for one of the subjects.

1) Optimal System Parameters: The first parameter optimized by the Training Module is the TX power level. Table III summarizes the average KL divergence computed by the Training Module for each subject. The time interval for computing PDR is 8 seconds.

	Subject 1	Subject 2	Subject 3
TXPowerLevel 1	9.88	12.86	**12.90**
TXPowerLevel 2	**13.18**	**20.79**	8.80
TXPowerLevel 3	1.21	1.59	1.58

Table III
AVERAGE KL DIVERGENCE OF PDR DISTRIBUTIONS OF EACH SUBJECT.

In Table III, the average KL divergence at TX power level 2 is the largest for subject 1 and 2. Thus, the Training Module selects TX power level 2 for subject 1 and 2. In contrast, for subject 3, the Training Module selects TX power level 1 that has the largest average KL divergence. It is because subject 3 is shorter than the first two subjects and TX power level 2 results in a relatively high PDR during her different activities. Thus, at TX power level 2, the differences among the PDR distributions of her different activities are smaller than those at TX power level 1.

All the optimal system parameters selected by the Training Module for each subject are summarized in Table IV.

Subject	TX Power Level	Packet Sending Rate (pkts/s)	Smoothing Window Size (seconds)
1	2	4	8
2	2	3	6
3	1	4	6

Table IV
OPTIMAL SYSTEM PARAMETERS SELECTED BY THE TRAINING MODULE.

From Table IV we observe that subject 2 has a smaller packet sending rate and smoothing window size. It may be because subject 2 performs each activity more cleanly and hence his data has less noise and is easier to classify.

2) Runtime Accuracy and Latency: In Figure 10, we present the total runtime accuracy of classifying seven activities for all three subjects. In addition to the total runtime accuracy, for each activity, we also plot its runtime accuracy ((true positive + true negative)/(true positive + false positive+ true negative + false negative)) in Figure 10, runtime precision (true positive/(true positive + false positive)) in Figure 11, and runtime recall (true positive/(true positive + false negative)) in Figure 12.

In Figure 10, Subject 1, 2, and 3 have the total runtime accuracies of 86.3%, 92.5%, and 84.2%, respectively. Interestingly, subject 2 performs much better than the other two subjects. Moreover, from Table III, we observe that subject 2 has the largest average KL divergence of PDR distributions among all subjects at their optimal TX power levels. Thus, it can be inferred that the average KL divergence of PDR distributions is a validated metric for indicating accuracy during the TX power level selection. The reason why subject 2 performs much better may be because that subject 2 performs each activity more cleanly.

In addition, we observe that most individual activities have accuracies over 90%. Compared to the most recent sensing-based solution [5], RadioSense achieves comparable accuracies in recognizing the locomotion activities (such as sitting, standing, walking, and lying). Running and riding are not considered in [5]. From Figure 11, we also observe that most individual activities have precisions over 0.8, which indicates the consistency of the accuracy results in Figure 10. Finally, we conclude that 10 minutes training data of each activity is adequate for RadioSense to achieve promising accuracy results.

To achieve the accuracies, precisions, and recalls depicted in Figure 10, 11, 12, the average latencies for subject 1, 2, 3 are 9.16 seconds, 6.46 seconds, and 7.12 seconds, respectively. In contrast to subject 2, subject 1 and 3's latencies are much larger than their smoothing window sizes. It is because the accuracies for subject 1 and 3 are lower than

that of subject 2. One misclassified activity introduces an additional latency of one smoothing window size. The more activities are misclassified, the more latencies are introduced.

3) Evaluation with More Activities: Aside from evaluating the activity set that includes high level activities, we collect data for two more activity sets from subject 1. The activity sets include more specific activities during sitting and cleaning. The first set is the *sitting set*, which includes driving, working, reading, eating, and watching TV. All these activities are performed when subject 1 is sitting. The second set is the *cleaning set*, which includes cleaning table, cleaning floor, cleaning bathtub, and cleaning blackboard. Previous work [26] has identified cleaning as a difficult activity to classify. The system parameters are configured with subject 1's optimal system parameters. Each activity is performed for 10 minutes during training and for 30 minutes during testing. The runtime accuracies of these two activity sets are summarized in Figure 13.

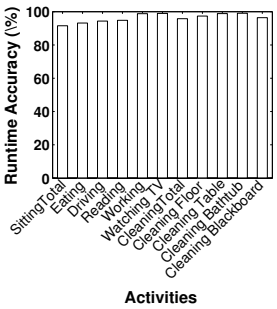

Figure 13. Accuracies for More Activities.

For the sitting set, the total runtime accuracy and latency are 91.5% and 8.35 seconds, respectively. For the cleaning set, the total runtime accuracy and latency are 95.8% and 8.29 seconds, respectively. All the accuracies are summarized in Figure 13.

Interestingly, only the PDR of the ankle node is selected as an important feature for classifying the activities in the sitting set. It is because the communication range of the wrist node well covers the base station even at the lowest TX power level when the subject is sitting. The PDRs of the wrist node during different activities are not significantly different. Moreover, for different activities, the distance between the subject's right hand and the base station is different and the subject's right hand moves differently. Thus, three RSSI features of the wirst node are selected by the feature selection algorithm during the training phase and play an important role in classifying the activities.

C. Battery Lifetime

We use nine Tmotes to measure the average battery lifetimes of on-body sensor nodes in RadioSense system. These Tmotes are divided into three groups, each of which is configured with the optimal parameters obtained by the Training Module for each subject. We run each Tmote with brand new batteries (AA, Alkaline, LR6, 1.5V) and record the battery lifetimes in hours.

The average battery lifetimes of the three groups are 159.3 hours, 168.7 hours, 175.3 hours, respectively. From these results, we conclude that on-body sensor nodes in RadioSense have long battery lifetimes.

D. Communication Range and Privacy

TX Power Level	TX Power (dBm)	Max Comm. Diameter (cm)
1	-33.0	77.2
2	-28.7	108.3
3	-25.0	388.0
7	-15.0	923.2

Table V
THE TX POWER AND THE LARGEST COMMUNICATION DIAMETER OF EACH TX POWER LEVEL.

We measure the communication ranges of TX power levels 1, 2, 3, and 7. In the measurement, a sensor node is defined as being connected with a base station as long as the PDR of this sensor node at the base station is larger than 20%. The measured ranges with the corresponding TX powers [17] are organized in Table V. From this table, we observe that TX power level 1 and 2, which are selected as the optimal TX power levels for three subjects, correspond to very low TX powers.

A lower TX power level not only indicates lower energy overheads, but also generates a smaller communication range. For example, instead of TX power level 7 (to achieve 90% PDR in BSN [9]) whose communication range is 923.2cm, RadioSense selects TX power level 1 and 2 for the three subjects. From Table V, the communication ranges when using these two TX power levels are less than 2 meters. On-body sensor nodes in RadioSense only send fake packets within such small communication ranges. As a result, privacy risks are reduced.

E. Coexistence with other On-body Sensor Nodes

RadioSense utilizes two dedicated on-body sensor nodes (the RadioSense nodes). In this subsection, we investigate whether the RadioSense nodes can coexist with other on-body sensor nodes. In the experiment, subject 1 wears RadioSense nodes and two extra on-body sensor nodes, which are placed on the left wrist and left ankle. These two extra on-body sensor nodes are deployed for general purpose. The general purpose nodes work at TX power level 7 and their packet sending rate is 4 pkts/s. Both the RadioSense nodes and general propuse nodes work on the same communication channel with CSMA enabled and communicate with the base station. At runtime, subject 1 performs each of the seven high level activities for 30 minutes. First, the total runtime accuracy is 90.8%. Compared to the total runtime accuracy (86.3%) without the the general purpose nodes, the accuracy is higher. The reason may be that the general purpose nodes contend with the RadioSense nodes for the communication channel. The contention increases the discriminative capacity of the communication pattern utilized by RadioSense. Second, the average PDRs of the general purpose nodes are 98.0% and 95.6%, which indicate very good link quality [28]. Since the RadioSense nodes work at a low TX power, they have limited effects on the general purpose nodes' communication. Thus, we conclude that the existence of the

RadioSense nodes rarely affects the communication qualities of the general purpose nodes.

VII. Conclusions and Future Work

In this paper, we propose RadioSense, a system that demonstrates the feasibility of utilizing wireless communication patterns for BSN activity recognition. Through our experiments with data gathered from three subjects during a two-week period, we demonstrate that RadioSense achieves promising results in terms of accuracy, latency, and battery lifetime. We also demonstrate that RadioSense is able to coexist with other on-body sensor nodes.

In future, we will continue our work from three aspects. First, we plan to program a smartphone as the base station in RadioSense, which makes RadioSense more practical. Second, we plan to consider communication features (e.g. inter-link reception correlation) between on-body sensor nodes. Intuitively, such communication features include more comprehensive information of activities. Third, we plan to investigate the impacts of different deployment locations of on-body sensor nodes.

References

[1] O. Chipara, C. Lu, and et al., "Reliable clinical monitoring using wireless sensor networks: experiences in a step-down hospital unit," in *Proc. of Sensys '10*.

[2] T. Bernd, B. Andreas, and et al., "Recognition of hearing needs from body and eye movements to improve hearing instruments," in *Proc. of Pervasive '11*.

[3] F. Albinali, S. Intille, and et al., "Using wearable activity type detection to improve physical activity energy expenditure estimation," in *Proc. of Ubicomp '10*.

[4] J. Doppler, G. Holl, and et al., "Variability in foot-worn sensor placement for activity recognition," in *Proc. of ISWC '09*.

[5] M. Keally, G. Zhou, and et al., "Pbn: Towards practical activity recognition using smartphone-based body sensor networks," in *Proc. of Sensys '11*.

[6] M. Quwaider and S. Biswas, "Body posture identification using hidden markov model with a wearable sensor network," in *Proc. of BodyNets '08*.

[7] L. Wang, T. Gu, and et al., "Real-time activity recognition in wireless body sensor networks: From simple gestures to complex activities," in *Proc. of RTCSA '10*.

[8] P. Zappi, C. Lombriser, and et al., "Activity recognition from on-body sensors: accuracy-power trade-off by dynamic sensor selection," in *Proc. of EWSN'08*.

[9] A. Natarajan, B. de Silva, and et al., "Link layer behavior of body area networks at 2.4 ghz," in *Proc. of MobiCom '09*.

[10] H. Sagha, J. del R. Millán, and R. Chavarriaga, "A Probabilistic Approach to Handle Missing Data for Multi-Sensory Activity Recognition," in *Workshop on Context Awareness and Information Processing in Opportunistic Ubiquitous Systems Ubicomp*, 2010.

[11] C. Karlof, N. Sastry, and D. Wagner, "Tinysec: a link layer security architecture for wireless sensor networks," in *Proc. of SenSys '04*.

[12] O. Amft, H. Junker, and G. Troster, "Detection of eating and drinking arm gestures using inertial body-worn sensors," in *Proc. of ISWC '05*.

[13] D. Gordon, H. R. Schmidtke, and et al., "A novel micro-vibration sensor for activity recognition: Potential and limitations," in *Proc. of ISWC '10*.

[14] T. Stiefmeier, G. Ogris, and et al., "Combining motion sensors and ultrasonic hands tracking for continuous activity recognition in a maintenance scenario," in *Proc. of ISWC '06*.

[15] A. Krause, M. Ihmig, and et al., "Trading off prediction accuracy and power consumption for context-aware wearable computing," in *Proc. of ISWC '05*.

[16] G. Andrea, *Wireless Communications*. Cambridge University Press, 2005.

[17] "CC2420 Data Sheet," http://inst.eecs.berkeley.edu/~cs150/Documents/CC2420.pdf.

[18] E. Miluzzo, N. D. Lane, and et al., "Sensing meets mobile social networks: the design, implementation and evaluation of the cenceme application," in *Proc. of SenSys '08*.

[19] M. Quwaider and S. Biswas, "Sensewaves: Radiowaves for context recognition," in *Pervasive '11*.

[20] J. Wilson and N. Patwari, "See-through walls: Motion tracking using variance-based radio tomography networks," *IEEE TMC 2011*.

[21] M. Youssef and A. Agrawala, "The horus location determination system," *Wirel. Netw.*

[22] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, 2003.

[23] A. H. Shoeb and J. V. Guttag, "Application of machine learning to epileptic seizure detection," in *ICML'10*.

[24] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag, 2006.

[25] H. Lu, A. J. B. Brush, B. Priyantha, A. K. Karlson, and J. Liu, "Speakersense: Energy efficient unobtrusive speaker identification on mobile phones," in *Proc. of Pervasive '11*.

[26] J. Lester, T. Choudhury, and et al., "A hybrid discriminative/generative approach for modeling human activities," in *Proc. of IJCAI'05*.

[27] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, 2011.

[28] K. Srinivasan, M. A. Kazandjieva, S. Agarwal, and P. Levis, "The β-factor: measuring wireless link burstiness," in *Proc. of SenSys '08*.