Detecting Object Motion Using Passive RFID: A Trauma Resuscitation Case Study

Siddika Parlak, *Member, IEEE,* and Ivan Marsic, *Member, IEEE*

***Abstract*—We studied object motion detection in an indoor environment using RFID technology. Unlike prior work, we focus on dynamic scenarios, such as emergency medical situations, subject to signal interference by people and many RFID tags. We build a realistic trauma resuscitation setting and record a dataset of around 14 000 detection instances. We find that factors affecting radio signal, such as tag motion, have different statistical fingerprints, making them discernible using statistical methods. Our method for object motion detection extracts descriptive features of the received signal strength and classifies them using machine-learning techniques. We report experimental results obtained with several statistical features and classifiers, and provide guidelines for feature and classifier selection in different environments. Experimental results show that object motion could be detected with an accuracy of 80% in complex scenarios and 90% on average. The motion type, on the other hand, could not be identified with such high accuracy using currently available passive RFID technology.**

***Index Terms*—Machine learning, motion detection, radio signal strength, RFID, trauma resuscitation.**

I. Introduction

THE MOBILITY of a user or an object provides valuable information for context-aware applications, such as human activity recognition by monitoring the use of objects [1]. People or object motion can be captured by cameras [2], [3], accelerometers [4], [5], passive radio-frequency identification (RFID) tags [6]–[8], active RFID tags [9], and WLAN sensors [10]. The sensor type is decided by application constraints and requirements. Active sensors (accelerometers) are favored for reliable and precise sensing. However, they are inconvenient for tagging small objects and costly for large-scale use. Passive sensors (barcodes and passive RFID tags) are preferred when many small and inexpensive items are tracked. An example application is recognizing daily activities by RFID tagging of everyday objects (milk bottle and mug) [1].

Passive RFID technology is a nonintrusive, low-cost, and privacy-preserving solution for sensing small and inexpensive objects in a dynamic environment. Relative to barcodes, RFID provides faster and simultaneous scanning of multiple items, longer read-range, and non-line-of-sight operation. Unlike active sensors, passive RFID tags do not require maintenance as they operate without batteries. Computer vision offers most of these advantages, but it also raises privacy concerns because of a permanent visual record of people and their activities. RFID is also better at detecting small and randomly oriented objects, and less sensitive to occlusions, both of which pose challenges for vision [11]. Despite its advantages, long-range passive RFID has received limited attention for activity recognition due to performance issues. Near-field RFID readers (fixed or worn, e.g., on a bracelet) have been preferred [1], [12]. They, however, may be intrusive, as they require users to remember to wear them or to swipe the objects near readers.

We explored using passive ultra-high frequency (UHF) RFID for long-range (2–3 m) detection of object motion within a room. We define object motion as any human interaction that causes a change in object's orientation and location, as well as occlusions with hand or body. These changes affect the energy reflected by RFID tags and result in frequent changes in received signal strength indication (RSSI). These signal changes have different statistical characteristics compared to the fluctuations caused by changes in the environment (e.g., human movement near the tag). Our method for motion detection processes the RSSI sequence by machine-learning techniques to detect fluctuations caused by tag motion.

Our target application is activity recognition during trauma resuscitation, the initial treatment of a severely injured patient in the emergency department. It is a challenging setting due to human presence and a large number and variety of objects. It requires careful placement of RFID antennas and tags [13], and the use of statistical methods for data processing (our focus here). We extract features from RSSI and classify them as *moving* or *still* using binary classifiers. Our results show that the advantages of scenario-dependent features and classifiers, and are applicable to other context-aware applications. We also studied recognition of motion types as *linear* versus *random*. We found that the identifying motion type is more challenging than detecting object motion with current passive RFID technology.

A. Related Work in Motion Detection and RFID

In the field of instrumentation and measurement several mobility sensors types have been used—accelerometers [4], [5], cameras [2], [3], and RFID tags [7]. Hache *et al.* [4] distinguished mobile from stationary states with 97.4% accuracy by applying thresholds on sensor data. Mobility type (walking and running) can also be found using thresholds [4] or machine-learning [5]. Accelerometers are reliable and precise, but our domain requires either low-cost sensors or

Manuscript received June 15, 2012; revised October 07, 2012; accepted December 30, 2012. Date of publication June 21, 2013; date of current version August 7, 2013. This paper was supported by the NSF under Grant 0803732. The Associate Editor coordinating the review process was Dr. Jesús Ureña.

S. Parlak is with Qualcomm Technologies Inc., Santa Clara, CA 95051 USA (e-mail: parlak@qti.qualcomm.com).

I. Marsic is with the Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ 08854 USA (e-mail: marsic@ece.rutgers.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIM.2013.2258772

TABLE I
CHARACTERISTICS OF THE TRAUMA RESUSCITATION DOMAIN

No.	Characteristics	Requirements/challenges	Value
1.	Number of people in experimental area	Human body occludes and absorbs radio signals. Effects become more severe as the number of people increases	5–20
2.	Number of tagged objects in view	Due to the collision avoidance mechanism (Slotted Aloha Protocol), number of readings from a tag decreases with the increasing number of tags in view	10–100
3.	Number of concurrently moving objects	Nearby concurrently moving tags may affect each other's signal strength and interfere	5–10
4.	Duration of interaction	A quick interaction may not be captured as its effect on the signal strength is smoothed while windowing (Section IV-A)	Seconds to minutes
5.	Speed of object movement	Very slow motion may be perceived as if the object is still. Swift movements may go undetected as the change in signal strength is smoothed while windowing	Slow to moderate (0.5–1 m/s)
6.	Speed of human movement	Fast human movement causes frequently changing environmental characteristics, causing fluctuations even for a still tag	Moderate (∼1–3 m/s)
7.	Frequency of human movement	Frequent human movement causes frequently changing environmental characteristics and more occlusions, causing fluctuations even for a still tag	Many times per minute
8.	Locale of object interactions	Regions with significant object concentration must be in coverage of RFID antennas	Scattered, near patient bed
9.	Tolerance to detection latency	Less tolerance to latency restricts the use of postprocessing techniques	Few seconds

cameras to track a large number of small, inexpensive, and disposable objects. Cameras provide rich contextual information and statistical methods for motion tracking [2], [3], but raise privacy concerns, especially in medical domains. Hence, we opted for passive RFID.

Passive RFID has been used for object localization [14], [15], identification [16], and motion detection [7]. Methods in [14] and [15], although related, do not apply to our study as our goal is detecting object motion. Also, we do not derive motion status from location since our definition of object motion includes both orientation changes and small but frequent location changes. A work closest to ours [7] used a rule-based method to detect movement of RFID-tagged objects. Although they achieved an accuracy of 94%, the algorithm worked poorly (40%–65%) in scenarios with human movement. Our approach achieved 95% accuracy in the baseline scenario, and 80% accuracy in most challenging scenarios with human movement and multiple tags. Machine learning not only provided superior performance but also obviated the need for defining the rules and associated thresholds.

II. APPLICATION DOMAIN: TRAUMA RESUSCITATION

Given the safety- and time-critical nature of trauma resuscitation, maintaining situation awareness is important for patient care. This process requires interaction with medical objects, most of which are uniquely associated with tasks. Object-use information has been applied for tracking hospital activities [17] and classifying surgery phases [18].

Prior work on long-range passive RFID-based motion detection focused on less dynamic domains, such as smart homes [6], [8] and offices, with few tagged objects and one or no persons present. As a result, in [19] the states were clearly discernible in the raw RSSI data. In contrast, trauma resuscitation poses challenges for RFID-based motion detection (Table I). Resuscitations require the collaborative work of medical teams, often with 15–20 members using tens to hundreds of tagged objects. Because tasks are performed in parallel, multiple objects might be in use simultaneously, and the speed and frequency of human movement are high.

TABLE II
INTERACTION STATISTICS FOR VARIOUS OBJECTS AND INTERACTION TYPES (ACROSS 32 TRAUMA RESUSCITATIONS)

Object	Interaction type	Interactions per resuscitation	Avg. dur. (s)	Min. dur. (s)
CO_2 detector	Relocation	0.5	8.9	2
	use	n/a	n/a	n/a
Laryngoscope	Relocation	1.2	12.4	1
	use	0.9	28.1	8
Thermometer	Relocation	0.4	4.6	2
	use	0.8	23.2	12
Fluid bag	Relocation	0.8	19.7	2
	use	1.5	21.3	2
Cervical collar	Relocation	0.6	18.8	4
	use	0.9	32.6	7

To design our experiments, we observed object motion in actual resuscitations. The average duration of object relocation depends on the distance of storage and usage locations (Table II). For example, CO_2 detector and laryngoscope are stored on carts near the patient bed. Fluid bags and cervical collars are stored away, in, or atop a cabinet. Duration of use, on the other hand, depends on many factors, including the task and number of attempts to accomplish the task. We observed that the most objects on average are used $\geq$20 s (Table II). Also, minimum duration of use is often longer than that of relocation. We use this insight to filter out short accidental interactions. Object relocation does not always imply usage, and number of interactions involving only relocation may be as high as that involving usage (Table II). Motion type provides additional contextual information to reduce false alarms. We identified two motion types: linear (relocation) and random (during use).

Duration of object use cannot be determined with sensor-based methods for wrapped items, such as tubes and CO_2 detectors, because we can tag only their wrapping. Tags on these objects are in view only briefly (<10 s), during unwrapping, after which the wrapping along with the tag is discarded. Although the act of unwrapping can be interpreted as usage, it is swift and may be confused with noise. The use of wrapped objects can be detected by other cues, such as

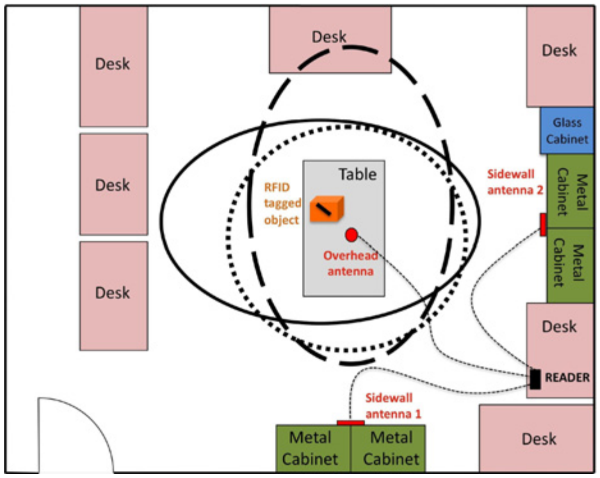

Fig. 1. Experimental layout. Coverage zones of antennas: sidewall antenna (1): dashed line; sidewall antenna (2): solid line; overhead antenna.

location (e.g., objects on patient bed are likely to be in use) or relationship information (e.g., if laryngoscope is in use, then CO_2 detector is likely in use) [13]. This paper focuses on detecting motion of nonwrapped objects, which are used for longer times, thus providing reliable motion detection.

III. Experimental Setup

Our experimental setup and scenarios were designed to evaluate the effects of: i) human presence and movement; ii) multiple tags in view; and iii) concurrent and nearby tag movement. These factors are variable in trauma resuscitation (Table I) and likely to affect the radio signal. Our experiments varied these parameters and kept others (items 4–9 in Table I) at fixed values, specified by the usual conduct of resuscitation.

A. Environmental Setting and RFID Equipment

We designed our laboratory setting similar to a resuscitation room: a patient bed at the center, side furniture, and space in between (Fig. 1). This is a challenging environment for RFID because the furniture causes multipath fading and distortion of the radio signal. A tagged object was used at the central table and in the surrounding space. We focused on this area because most interactions with objects occur around the patient (Table I). We assumed complete antenna coverage (see [13]) and focused on evaluating our motion detection algorithms.

We used off-the-shelf equipment from Alien Technology: an RFID reader (ALR-9900), three circularly polarized antennas (ALR-9611-CR), and passive tags (Squiggle ALN-9540). Since such antennas allow tag detection in two orientations, two perpendicular antennas usually suffice to detect tags in all orientations. A third perpendicular antenna (Fig. 1) created redundancy to mitigate the adverse effects on radio signals (e.g., occlusion and multipath). One antenna was ceiling-mounted (2.7 m above floor) facing the center of a plastic-top table sized $0.76 \times 1.27 \times 0.76$ m. To reduce human occlusion the other two antennas were sidewall mounted, 2-m above the floor facing the table at an angle of 60deg to the floor. Since workers mostly gather around the patient, ceiling-mounted antenna was less likely occluded and thus critical for our design.

The reader operated autonomously, collecting data for a given time with no control inputs. Regardless of the number of present tags, the reader scanned for multiple tags instead of fast searching for a target tag. For scalability to more readers, we used a single reader working in dense reader mode (DRM) to prevent readers' interference. DRM performs best when tag-to-reader distance is $\geq$1.5 m [20]. Radio signal (1 W) was emitted in a round robin way by one antenna at a time (0.5 s).

B. Data Collection

Our dataset consisted of 240 RSSI recording sessions. Each session lasted 60 s yielding 14 400 instances of motion detection. Based on our target application, each session included linear, random, and no movement, simulated by interacting with an RFID-tagged cardboard box (19 cm $\times$ 10 cm $\times$ 7 cm) in three ways: i) holding the box and walking at $\approx$1 m/s (object moving linearly); ii) standing and wiggling the box, rotating in 3-D and occluding by hand (object moving randomly); and iii) not interacting at all (object standing still). We refer to the tag on the box as "target tag" to distinguish it from other tags in the experimental area.

Random movement for 20 s simulated average use duration (Table II). Relocation duration (linear motion) is usually less than both use (random motion) and nonuse (still); however, we set relocation and still durations to 20 s to generate equal amounts of data for each motion type and analyze all motion types under equal conditions. The three 20-s interactions occurred contiguously in varying order for sessions of 60 s.

Different scenarios introduced challenges during actual resuscitations and thus affecting the radio signal behavior.

Scenario #1—Baseline: One tag in view and no movement in the environment (60 sessions).

Scenario #2—Human movement: One tag in view.

Scenario #2a: One person moving (30 sessions).

Scenario #2b: Two people moving (30 sessions).

The experimenter(s) continuously walked around the table and from the surrounding furniture to the table, at about 1 m/s, simulating human motion observed during resuscitations.

Scenario #3—Multiple tags: Ten tags (including target tag) uniformly scattered on the table.

Scenario #3a: No people movement (30 sessions).

Scenario #3b: One person moving, not touching tags (30 sessions).

In Scenario #3a, experimenters in the room stood still.

Scenario #4—Concurrent and nearby tag movement: Two tags attached to different cardboard boxes.

Scenario #4a: Two experimenters holding tagged objects walked around the table and from periphery to the table at $\approx$1 m/s (30 sessions). Their movement was not synchronized.

Scenario #4b: The target tag stationary; the experimenter, standing still, randomly moved another tag while maintaining the average tags distance at $\approx$15 cm. The goal was to study a nearby tag movement with a still target tag (30 sessions).

IV. Method

We saw motion detection as a binary classification problem where raw RFID data were represented with feature vectors

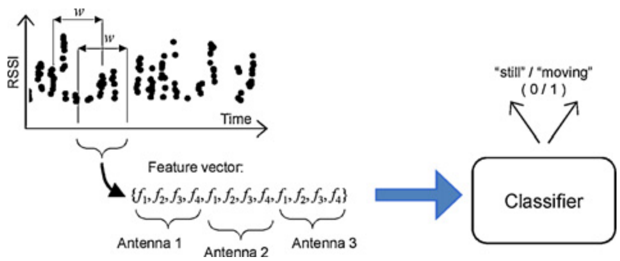

Fig. 2. Diagram illustrating the data processing in our system.

and object's mobility status was represented with class labels. We used a sliding-window to map the RSSI data to features. At each time, a feature vector was extracted from the data in the current window. A classifier then assigned each feature either *moving* or *still* label (Fig. 2). We also filtered the label sequence generated by nontemporal classifiers to remove spurious transitions. We next describe these steps in detail.

A. Windowing

We experimented with fixed windows of 1.5, 2, 3, 5, and 10 s. We adjusted the shift size to half of the window length [12], [21]. As the window shifted by a half-window length, one classification decision was output for each window (Fig. 2). A 1.5-s window was the shortest because our setup had three antennas, each with a scanning time of 0.5 s. A 10-s window is the longest because the latency and smoothing of a longer window would not be appropriate for our domain. The average duration of object use in trauma resuscitation is relatively short (20 s, Table II). To detect such interactions and be able to generalize our results, we limited the window to 10 s.

B. Interpolating the RSSI Signal

In our setup, tag reads occurred at irregular intervals since:

1) the autonomous mode setting caused irregular data arrivals. Data rates varied from 1 to 33 readouts/s, with a median rate of 23 readouts/s (Table III);
2) antennas were activated in a round-robin way to avoid mutual interference. Readouts arrived only from the antenna through which the reader was transmitting;
3) multiple tags competed for wireless channel (ALOHA protocol), reducing data rates per tag (for two tags 90th percentile dropped from 31 to 24 readouts/s). More tags significantly reduced data rates (Table III).

Before feature extraction, we preprocessed RSSI data within the window to fill in missing samples. The Alien RFID reader reports unitless RSSI values in the range 500–50 000, up to 1 decimal point, which were used without quantization. We first removed the gaps by concatenating the intervals when the antenna was active; then linearly interpolated the RSSI values within the window. The number of interpolation points was determined by the window size so that the resulting data rate per second was the same for all window sizes in all sessions.

C. Feature Extraction

By visualizing the RSSI distribution in different scenarios (Fig. 3), we observed that the variance was significantly higher for a moving tag (first row) than for a still tag (second row). Human movement caused the RSSI to deviate more for both states, but less for a still tag. With multiple tags, the variance

TABLE III
DATA RATE STATISTICS BASED ON THE NUMBER OF TAGS IN THE ENVIRONMENT (READOUTS PER SECOND)

# Tags	Min.	10th perctl.	Median	90th perctl.	Max.
1	1	16	23	31	33
2	4	18	23	24	32
10	0	5	7	9	11

became smaller because of more RSSI interpolation. Similar to human movement in Scenarios #2a and #2b, concurrent nearby tag movement in Scenarios #4a and #4b caused more scattered RSSI values, but the variance was notably higher for a moving tag. These observations indicate that the standard deviation is a useful feature for detecting object motion.

We also analyzed temporal behavior of RSSI (Fig. 3, bottom 2 rows). In all scenarios, difference between consecutive RSSI readings was much smaller for a still tag. Compared to the deviation of RSSI (Fig. 3, top rows), difference of RSSI values provided better separation among moving and still states. Based on these observations, we defined *standard deviation* (f_1) and *difference* (f_2) as our baseline features. We defined two more features to investigate the effects of an enhanced feature set.

1) *Delta Mean* (f_3): It is the amount of change between the RSSI averages of two successive windows. It should be high when the interaction status changes.
2) *Trend* (f_4): It is the slope of the line fitted to the RSSI series in the current window. It captures the RSSI tendency and should be high for the linear motion type.

The features f_1 to f_4 were computed for each antenna separately and concatenated (instead of averaging) to obtain the final feature vector. Our choice of concatenation over averaging was based on our preliminary experiments. Change in tag orientation may increase reception for one antenna, and decrease that for the other, leaving average reception the same but missing the orientation change. To preserve orientation change information, which implies object motion, we concatenated the feature values from different antennas.

D. Classification

The above analysis showed that the *moving* and *still* readouts are overlapping in the feature space, especially in scenarios with human presence and movement. The time correlation of the data can be exploited to obtain more accurate classification results. We incorporated this correlation in three ways: i) by applying nontemporal classification and smoothing the output labels with a Hidden Markov Model (HMM); ii) by applying temporal classification with generative models; and iii) by using temporal classification with discriminative models. Discriminative temporal models have been used for RFID-based activity recognition [22], but not for long-range passive RFID. The details of the three classification methods follow.

1) *Nontemporal Classification:* We experimented with nontemporal classifiers in WEKA [23].

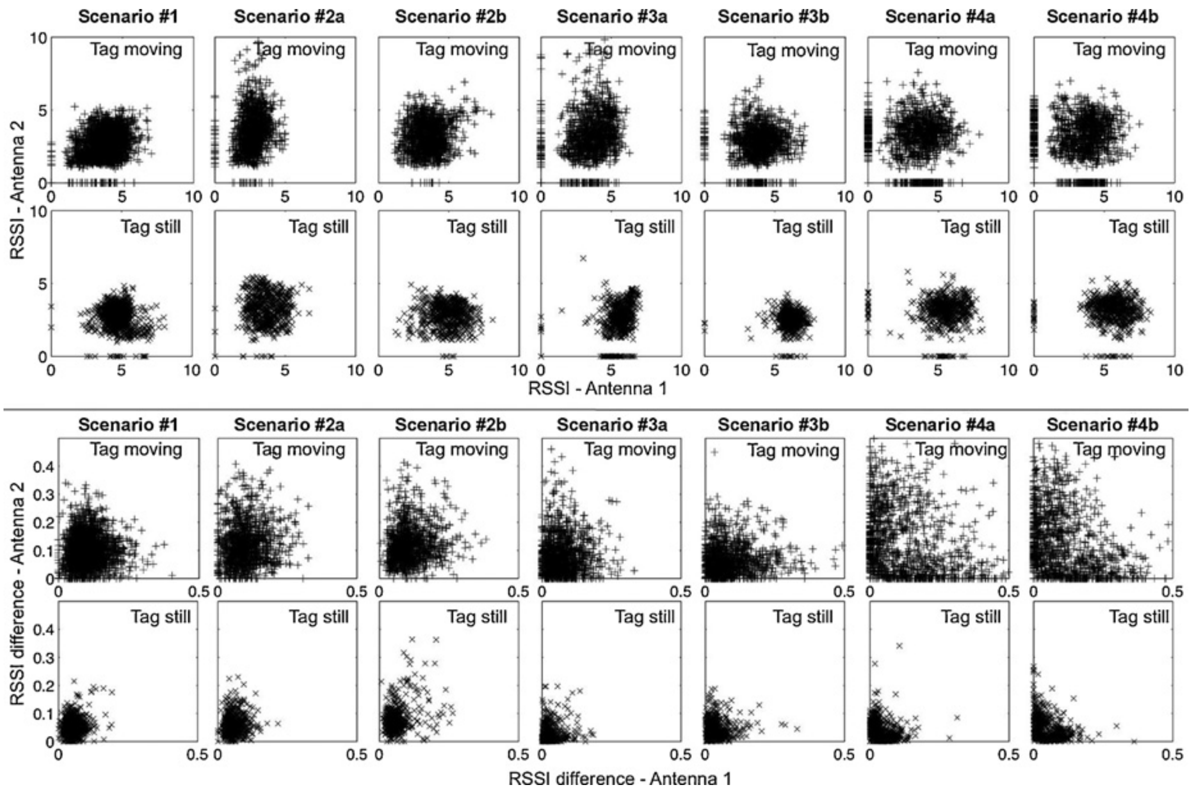

Fig. 3. RSSI values (top 2 rows) and differences of consecutive RSSI readings (bottom 2 rows) for a moving and still tag. Shown are data from Antenna 1 (x-axis) and Antenna 2 (y-axis) (Fig. 1). Each chart has $\geq$600 datapoints. Scenario is specified at the top (#1: 1 tag, no people present; #2a: 1 tag, 1 person moving; #2b: 1 tag, 2 people moving; #3a: 10 tags, no people present; #3b: 10 tags and 1 person moving; #4a: 2 people concurrently moving and carrying tags; #4b: 2 tags and no people movement (nontarget tag moved randomly in place).

Decision Trees (DT): A tree of decision nodes based on the C4.5 algorithm [24]. This classifier has been used in similar tasks and yielded satisfactory results [12], [21].

Support Vector Machines (SVM): Trained to maximize the margin between different classes [24]. SVMs are known for their low generalization error.

Random Forests: An ensemble classifier (i.e., collection of classifiers), with many DTs, each voting for a class [24]. RFID signals are sensitive to environmental effects and building an accurate statistical model requires large volume of data from diverse scenarios. Ensemble classifiers integrate several models to lessen overfitting and improve classification.

Boosting: An ensemble classifier, combining weak learners to obtain a strong learner. We used the logit boost algorithm [24].

The classifier output was postprocessed using an HMM to remove spurious transitions and smoothen the label sequence. We chose an HMM over approaches, such as median filtering as HMM incorporates the domain knowledge. For example, during resuscitations some objects are used only briefly and else stay still, e.g., laryngoscope is typically used for 28 s in a 20- to 30-min long resuscitation (Table II). Expected transition frequency between *moving* and *still* states and the frequency of self-transitions, can be incorporated into the HMM state transition matrix. Since *moving* and *still* classes in our dataset were about equal size, we assumed an equal self-transition probability for both (0.99). Probability of confusion of the two states was estimated by classification of training data and integrated into the HMM as observation probabilities. We compared predicted labels to manually obtained ground truth.

2) *Temporal Classification With Generative Models:* Deciding motion status over time is a sequential learning problem. An HMM is a generative model that incorporates temporal or spatial information [24]. We built an HMM where the actual motion states form the hidden state set (*still* or *moving*) and the predicted states form the observation symbols set (named the same: *still* or *moving*). Transition probabilities were estimated from the training data. Observation probabilities were modeled with a Gaussian density because histograms of extracted features showed Gaussian distribution (also visible in Fig. 3). Parameters were estimated from the training data using the Maximum Likelihood principle [25].

3) *Temporal Classification With Discriminative Models:* Classification using generative models requires fitting a distribution to observations. Given the dynamic nature of resuscitation with many variable parameters, estimating an accurate model is difficult. We therefore selected a model that does not require fitting a distribution to the observed data. Conditional random fields (CRF) are discriminative models, specifying the probability of label sequences conditionally, based on the observations rather than joint probabilities [26]. We used the hCRF library to train the CRF-based model.

V. Experimental Results

We evaluated motion detection using six-fold cross validation, where the dataset is randomly partitioned in six subsamples, and one subsample is retained as the validation set. This process is repeated six times, such that each subsample is used once as validation data. Because readouts across a session were

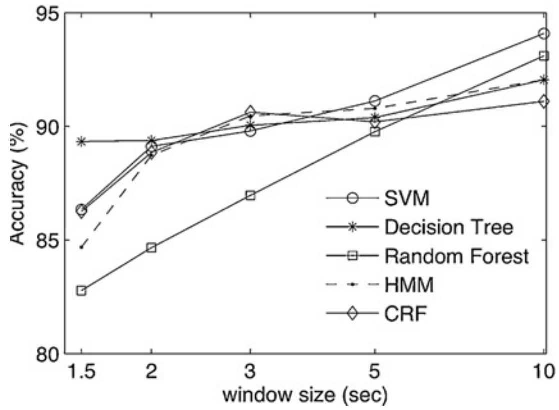

Fig. 4. Comparison of moving versus not moving classification accuracy for different classifiers and window sizes.

correlated, a complete session must be used either in training or in testing. Including a part of a session in the training and the remainder in the testing set would have artificially increased accuracy.

We started with baseline features [standard deviation (f_1) and difference (f_2)] and later added other features (f_3 and f_4). Our evaluation metric is the percentage of accurately classified labels in a session. The performance score is the average accuracy over all sessions in the testing set.

A. Effects of Classifier and Window Size

Here, we evaluated motion detection performance of several classifiers and the effects of window size on classification accuracy. A 10-s window yielded the highest accuracy for all classifiers (Fig. 4). DT outperformed other classifiers for the smallest window size (1.5 s). SVM outperformed others as the window size increased. Random Forests were no better than DT although they are a boosted version of DT. Random Forests are known to overfit for noisy datasets [27]. Our results confirmed that these prior findings as we observed significant improvement with increased window size, which yields less noisy features (Fig. 4). Our subsequent experiments showed that the overfitting could be prevented by limiting the growth of trees in the Random Forest. By limiting the growth, we obtained improved scores for Random Forest, which were even better than the scores obtained using DT.

Our results showed that the advantage of DT for small windows, e.g., for detecting motion of objects that are used only briefly, such as wrapped items or fluid bags (Table II). Our results also suggest using SVM for longer windows, to detect motion of long-used objects, e.g., laryngoscope (Table II). Since each RFID tag provides its object identity, it is possible to set the window size adaptively at runtime based on object type.

Temporal classifiers (HMM and CRF) slightly outperformed for a 3-s window, but the difference was not statistically significant. Nontemporal classifiers better captured the change in RSSI due to object motion and were less sensitive to environment effects. Integrating temporal information by smoothing the classifier output seems to suffice (Section IV-D).

B. Motion Detection Performance in Different Scenarios

We analyzed how motion detection performance varied from ideal to challenging scenarios seen in dynamic settings, such as trauma resuscitation. All classifiers provided an accuracy >90% in the ideal scenario (Scenario #1; Fig. 5). The accuracy decreased with human movement (Scenarios #2a and #2b). The reduction from Scenario #1 to #2a was higher than that from Scenario #2a to #2b. With each new person, the accuracy reduction decreased. The HMM-based classifier yielded notably higher scores in the human movement Scenarios #2a and #2b. We infer that the observation model has not dramatically changed in the presence of human movement. HMMs may be useful in scenarios with few tags and significant human movement.

Although multiple tags (Scenario #3a) decreased the accuracy, the combined effect of multiple tags and human movement was harsher (Scenario #3b). Concurrently moving tags (Scenario #4a) and movement of a nearby tag (Scenario #4b) did not decrease motion detection performance, and even provided an improvement. The nearby tag caused more scattered feature values when the target tag was moving, but had no significant effects when the target tag was still (Fig. 3).

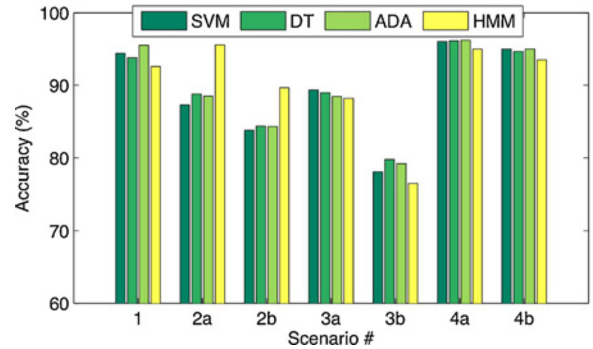

Fig. 5. Motion detection performance of classifiers in different scenarios.

TABLE IV
MOTION DETECTION ACCURACY (%) DEPENDING ON RSSI INTERPOLATION. INCLUDES ALL DATA COLLECTED IN ALL SCENARIOS

	No interpolation	With interpolation
SVM	87.0	90.2
HMM	70.9	90.5

C. Effect of RSSI Data Interpolation

Performance of all classifiers significantly improved with data interpolation (Table IV). Because the difference features (f_2) for the same motion type in different scenarios varied, they could not be represented by a single parameter set. Due to this reason, decrease in accuracy was higher for HMMs (Table IV); HMM is a parametric model and a single Gaussian was not sufficient for all scenarios. Although a more complicated model (e.g., Gaussian mixture) could yield better results, its training requires more data to estimate the increased number of parameters, as well as knowing the number of mixtures. Interpolation provided a simple solution to decrease the dependency of feature values on the number of visible tags.

D. Effect of Individual Features

Here, we analyzed the efficiency of features. We included delta mean (f_3) and trend (f_4) features, in addition to our

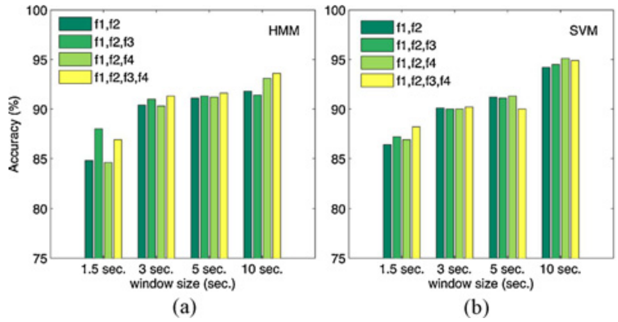

Fig. 6. Motion detection performance of different feature sets and window sizes, using an (a) HMM-based and (b) SVM-based classifier.

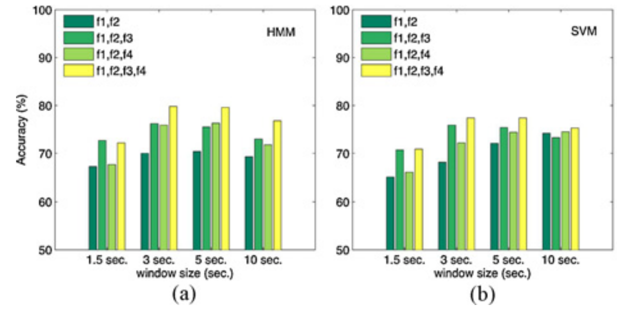

Fig. 7. Motion *type* classification performance of different feature sets and window sizes, using an (a) HMM-based and (b) SVM-based classifier.

baseline feature set of standard deviation (f_1) and difference (f_2), and experimented with several feature subsets. We present the results of using HMM- and SVM-based classifiers as representatives of temporal and nontemporal classifiers.

A richer set of features yielded better scores for HMM [Fig. 6(a)]. The trend feature (f_4) degraded performance for short windows, but slightly improved it for longer ones. Estimation of f_4 requires longer RSSI observation than for other features, making it suitable for longer movements. For SVM, a richer feature set improved the performance only for a 1.5-s window, since a short window contains less data and yields poor feature estimation. In this case, two additional features compensated for the noise in feature calculation. As the window size grew, added features did not contribute significantly since the two baseline features were already reliable [Fig. 6(b)]. We conclude that the additional features improved motion detection accuracy for short windows and generative classifiers, such as HMMs. Else, the baseline feature set performed equally well.

E. Motion Type Recognition

Here, we classified object motion as “still,” “moving linearly,” or “moving randomly,” instead of “still” versus “moving.” Motion type helps reduce false alarms during activity recognition, since random motion is a strong indicator of object use (Section II). Consider a scenario where a nurse fetches a thermometer to measure patient’s temperature. The patient has an oxygen mask on his face, preventing temperature measurement, so she leaves the thermometer on the bed. Later, she takes the thermometer and measures the temperature. If the algorithm declared that any interaction indicated usage, the first interaction in this example is a false alarm, which is avoidable if the algorithm discerned relocation (linear motion) from usage (random motion).

The best motion recognition scores were obtained for windows of length 3 and 5 s [Fig. 7(a)]. As the window size grew, linear and random movements tended to create statistically similar RSSI sequences, making their discrimination challenging. For different applications, the optimum window size can be adjusted depending on the speed of movement and distance traveled, which can be determined by analyzing the application domain.

HMM achieved accuracy rates of up to 80% using the enhanced feature set [$f_1, \ldots, f_4$; Fig. 7(a)]. Compared to binary classification [Fig. 7(a)], the enhanced feature set yielded greater accuracy. Recognizing the type of motion as random or linear requires defining a rich set of features.

We also experimented with cascaded classification, where one classifier discriminated between moving and still and the second determined motion type as linear and random. For DT and boosting, cascaded classification yielded better scores that were close to that of SVMs [Fig. 7(b)]. However, cascading did not provide any gain for SVM, which is originally a binary classifier and encapsulates multiple classifiers for performing multiclass classification. Our cascading strategy was then implicitly embedded in the multiclass SVM.

VI. Conclusion

Passive RFID offered a nonintrusive, cost-effective, and privacy-preserving sensing solution for dynamic settings with people and many tagged objects. Our motion detection algorithms were based on statistical machine learning to process noisy RSSI data, rather than the rule-based approaches used in prior work. Our experiments showed that the feasibility of using RFID for object motion detection. Our observations follow.

A. Effect of Human Interference

It was difficult to model RF propagation in cluttered indoor environments, and more so in dynamic settings with human motion. In our experiments, machine-learning tools were able to learn statistical “fingerprints” of object motion and discern them from environmental effects. Although motion detection was affected by human presence, the accuracy remained high.

B. Effects of Multiple Objects (Tags)

Objects for tagging differ in material, size, shape, packaging, and usage style. These features must be considered both when placing the tags [13] and when developing the algorithms (this paper). Use duration for medical objects varied from very short to relatively long, with an average about 20 s (Table II). Objects that were used briefly (e.g., fluid bags) required shorter windows for detection. Our results showed that some classifiers perform better for short windows, suggesting that they will be applied for briefly used objects. This was easily done because object identity was known from the tag ID. Another parameter that depends on the window size was the optimum feature set. A short window yields poor estimate of some features (e.g., trend) that can be remedied with a more diverse features. For long windows, adding new features did not resulted in a visible improvement.

At least one tag must be attached to each object, and most objects required multiple tags [13]. The high degree of tagging led to many tags in the antenna view at any time, reduced the read rate from any single tag. Moreover, temporal features (e.g., difference of RSSI readings) depend on the data arrival rate, which in turn depends on the number of tags in view. We showed that the dissimilarity of features from different conditions could be reduced by preprocessing (interpolation).

C. Other Observations

Because of temporal continuity of RSSI data, we expected that the temporal classifiers would perform better. We did not observed any superiority of temporal classifiers, except in the scenarios with few tags and significant human movement.

Although $>90\%$ accuracy rates were achievable for motion detection, lower scores resulted for motion type recognition. It may be more reliable to discern linear and random motion types by coarsely estimating the initial and final position of an object, instead of detecting the relocation from the RSSI pattern. Also, object relations could be exploited for tasks with multiple objects. If the use of one object can be identified reliably, another object under the same task may also be in use.

D. Lessons Learned for Detecting Motion in Dynamic Settings

Object use detection was the key for recognizing human activities that involved objects—object motion was a key indicator of use. With long-range passive RFID, we achieved a motion detection accuracy of 90% on average, and $\approx$80% under challenging scenarios. These scores were promising for a passive technology and even comparable to some active sensors (active RFID tags [9], Wi-Fi [10], [27]). However, the performance adequacy depends on the application requirements. Life-critical medical tasks, e.g., require high detection rates and other sensors were needed to complement passive RFID. Either individually or combined with other sensors, passive RFID was promising and worth pursuing in dynamic work settings, where: 1) distractions are not tolerated; 2) privacy is a concern; 3) tracked objects are small in size; and 4) tracked objects are large in number and sometimes disposable, necessitating low-cost sensors.

REFERENCES

[1] M. Philipose, K. Fishkin, M. Perkowitz, D. Patterson, D. Fox, H. Kautz, and D. Hahnel, "Inferring activities from interactions with objects," *IEEE Pervasive Comput.*, vol. 3, no. 4, pp. 50–57, 2004.

[2] P. Vadakkepat and L. Jing, "Improved particle filter in sensor fusion for tracking randomly moving object," *IEEE Trans. Instrum. Meas.*, vol. 55, no. 5, pp. 1823–1832, 2006.

[3] N. H. Dardas and N. D. Georganas, "Real-time hand gesture detection and recognition using bag-of-features and support vector machine techniques," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 11, pp. 3592–3607, 2011.

[4] G. Hache, E. D. Lemaire, and N. Baddour, "Wearable mobility monitoring using a multimedia smartphone platform," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 9, pp. 3153–3161, 2011.

[5] M. Ermes, J. Parkka, J. Mantyjarvi, and I. Korhonen, "Detection of daily activities and sports with wearable sensors in controlled & uncontrolled conditions," *IEEE Trans. Inform. Tech. Biomed.*, vol. 12, no. 1, pp. 20–26, 2008.

[6] K. Fishkin, B. Jiang, M. Philipose, and S. Roy, "I sense a disturbance in the force: Unobtrusive detection of interactions with RFID-tagged objects," in *Proc. UbiComp*, 2004, pp. 268–282.

[7] B. Jiang, K. P. Fishkin, S. Roy, and M. Philipose, "Unobtrusive long-range detection of passive RFID tag motion," *IEEE Trans. Instrum. Meas.*, vol. 55, no. 1, pp. 187–196, 2006.

[8] L. Bao and S. Intille, "Activity recognition from user-annotated acceleration data," in *Proc. Pervasive Computing*, vol. 3001. 2004, pp. 1–17.

[9] K. Kleisouris, B. Firner, R. Howard, Y. Zhang, and R. P. Martin, "Detecting intra-room mobility with signal strength descriptors," in *Proc. ACM MobiHoc*, 2010, pp. 71–80.

[10] J. Krumm and E. Horvitz, "Locadio: Inferring motion and location from Wi-Fi signal strengths," in *Proc. MobiQuitous*, 2004, pp. 4–13.

[11] P. Turaga, R. Chellappa, V. Subrahmanian, and O. Udrea, "Machine recognition of human activities: A survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 11, pp. 1473–1488, 2008.

[12] B. Logan, J. Healey, M. Philipose, E. M. Tapia, and S. Intille, "A long-term evaluation of sensing modalities for activity recognition," in *Proc. UbiComp*, 2007, pp. 483–500.

[13] S. Parlak, A. Sarcevic, I. Marsic, and R. Burd, "Introducing RFID technology in dynamic and time-critical medical settings: Requirements and challenges," *J. Biomed. Inform.*, vol. 45, pp. 958–974, 2012.

[14] A. F. C. Errington, B. L. F. Daku, and A. F. Prugger, "Initial position estimation using RFID tags: A least-squares approach," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 11, pp. 2863–2869, 2010.

[15] C. Alippi, D. Cogliati, and G. Vanini, "A statistical approach to localize passive RFIDs," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2006, pp. 21–24.

[16] M. Jo, H. Y. Youn, S. Cha, and H. Choo, "Mobile RFID tag detection influence factors and prediction of tag detectability," *IEEE Sensors J.*, vol. 9, no. 2, pp. 112–119, 2009.

[17] J. Favela, M. Tentori, L. Castro, V. Gonzalez, E. Moran, and A. Martínez-García, "Activity recognition for context-aware hospital applications: Issues and opportunities for the deployment of pervasive networks," *Mobile Netw. Appl.*, vol. 12, no. 2, pp. 155–171, 2007.

[18] J. E. Bardram, A. Doryab, R. M. Jensen, P. M. Lange, K. L. G. Nielsen, and S. T. Petersen, "Phase recognition during surgical procedures using embedded and body-worn sensors," in *Proc. IEEE PerCom*, 2011, pp. 45–53.

[19] L. Ravindranath, V. Padmanabhan, and P. Agrawal, "Sixthsense: RFID-based enterprise intelligence," in *Proc. MobiSys*, 2008, pp. 253–266.

[20] M. Buettner and D. Wetherall, "An empirical study of UHF RFID performance," in *Proc. ACM MobiCom*, 2008, pp. 223–234.

[21] S. Reddy, J. Burke, D. Estrin, M. Hansen, and M. Srivastava, "Determining transportation mode on mobile phones," in *Proc. IEEE ISWC*, 2008, pp. 25–28.

[22] T. van Kasteren, A. Noulas, G. Englebienne, and B. Kröse, "Accurate activity recognition in a home setting," in *Proc. UbiComp*, 2008, pp. 1–9.

[23] M. Hall, E. Frank, H. Geoffrey, B. Pfahringer, P. Reutemann, and I. H. Witten, "The Weka data mining software," *SIGKDD Explor.*, vol. 11, no. 1, pp. 10–18, 2009.

[24] E. Alpaydin, *Introduction to Machine Learning*. Cambridge, MA, USA: MIT Press, 2010.

[25] S. Parlak and I. Marsic, "Monitoring interactions with RFID tagged objects using RSSI," in *Proc. Mobile Ubiquitous Syst. (Mobiquitous 2010)*, LNICS 73. 2012, pp. 374–380.

[26] J. D. Lafferty, A. McCallum, and F. C. N. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in *Proc. ICML*, 2001, pp. 282–289.

[27] M. Segal, "Machine learning benchmarks and random forest regression," *Center Bioinform. Mol. Biostat.*, 2004.

[28] K. Muthukrishnan, M. Lijding, N. Meratnia, and P. Havinga, "Sensing motion using spectral and spatial analysis of WLAN RSSI," in *Proc. 2nd Eur. Conf. Smart Sens. Context*, 2007, pp. 62–76.

Siddika Parlak (S'08–M'13) received the B.S. and M.S. degrees from the Department of Electrical and Electronics Engineering, Bogazici University, Istanbul, Turkey, in 2006 and 2008, respectively. For the M.S. degree, she worked on speech recognition and spoken information retrieval. She received the Ph.D. degree from the Department of Electrical and Computer Engineering, Rutgers University, New Brunswick, NJ, USA, in 2013.

Her current research interests include speech processing, ubiquitous computing, and context-aware systems.

Ivan Marsic (M'94) received the Dipl.Ing. and M.S. degrees in computer engineering from the University of Zagreb, Croatia, and the Ph.D. degree in biomedical engineering from Rutgers University, New Brunswick, NJ, USA, in 1994.

He is currently a Professor with the Department of Electrical and Computer Engineering, Rutgers University. His current research interests include software engineering and sensor networks for health-care applications.