Passive RFID for Object and Use Detection during Trauma Resuscitation

Siddika Parlak, Ivan Marsic, Aleksandra Sarcevic, Waheed U. Bajwa, Lauren J. Waterhouse, and Randall S. Burd

Abstract—We evaluated passive radio-frequency identification (RFID) technology for detecting the use of objects and related activities during trauma resuscitation. Our system consists of RFID tags and antennas, optimally placed for object detection, as well as algorithms for processing RFID data to infer object use. To evaluate our approach, we tagged 81 objects in the resuscitation room and recorded RFID signal strength during 32 simulated resuscitations performed by trauma teams. We then analyzed RFID data to identify cues for recognizing resuscitation activities. Using these cues, we extracted descriptive features and applied machine-learning techniques to monitor interactions with objects. Our results show that an instance of a used object can be detected with accuracy rates greater than 90 percent in a crowded and fast-paced medical setting using off-the-shelf RFID equipment, and the time and duration of use can be identified with up to 83 percent accuracy. We conclude with insights into the limitations of passive RFID and areas in which RFID needs to be complemented with other sensing technologies.

Index Terms—Machine learning, medical information systems, medicine and science, sensors, RFID, object-based sensing, activity recognition, emergency medicine, trauma resuscitation

1 Introduction

Time- and safety-critical settings require efficient and error-free task performance. Trauma resuscitation—the initial management of critically injured patients in the trauma bay, a dedicated room in the emergency department—is an example of such setting. The resuscitation environment is dynamic and often chaotic, which can contribute to medical errors and miscomunication. Team members may forget the parameters of past tasks (e.g., given amount of intravenous (IV) fluid) or miss important steps, leading to inefficiencies and adverse outcomes. The benefits of computerized support for trauma teams have been shown using an expert system that tracked and validated the resuscitation progress based on manual data input [7]. This system, however, had limited usability due to the need for manual data entry and difficulty of capturing information from multiple sources.

Automating context awareness requires selection and deployment of appropriate sensors, as well as processing of the acquired data to infer context, such as activities or object use. Passive RFID technology offers an unintrusive, low-cost and privacy-preserving sensing solution. Unlike accelerometers [4] and active RFID tags [12], passive RFID tags do not require maintenance because they operate without batteries. Passive RFID tags are also smaller (convenient for small medical objects) and cheaper (disposable and usable at the item level). Although computer vision has similar advantages, its use may be limited by privacy concerns. Cameras capture at least a temporary record of people, whereas RFID data contain little or no personal information. Moreover, RFID is better at detecting small or randomly oriented objects, and at tolerating occlusions. Despite these advantages, long-range passive RFID has received limited attention in the activity recognition community due to performance issues. Near-field RFID readers have been used for achieving robust tag detection [5], [10], [18], [19]. These readers, however, are intrusive and require users to remember to wear them.

Our long-term goal is to develop a context-aware system that automatically recognizes human activities in real time and provides feedback to improve the efficiency and effectiveness of time-critical medical work, such as trauma resuscitation. Building such a system requires a combination of approaches and technologies, including RFID, speech recognition, computer vision, and other sensors. In this paper, we focus on long-range passive RFID technology and examine its feasibility for detecting used objects and associated activities during trauma resuscitation. We monitored object use by processing received signal strength data from tagged objects with machine-learning techniques. To evaluate the efficacy of this approach, we conducted a study in an actual trauma bay equipped with six RFID readers and 81 medical objects tagged with off-the-shelf RFID tags.

Using experimental data from 32 resuscitations performed by medical teams on a patient mannequin, we found that our method could detect specific objects in a crowded clinical setting with accuracy rates greater than

- S. Parlak is with Qualcomm Technologies Inc., Santa Clara, CA 95051. E-mail: siddika.parlak@gmail.com.
- I. Marsic and W.U. Bajwa are with the Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ 08854. E-mail: {marsic, waheed.bajwa}@rutgers.edu.
- A. Sarcevic is with the College of Computing and Informatics at Drexel University, Philadelphia, PA 19104. E-mail: aleksarc@drexel.edu.
- L. J. Waterhouse and R. S. Burd are with the Division of Trauma and Burn Surgery, Children's National Medical Center, Washington, DC 20010. E-mail: {lwaterho, rburd}@childrensnational.org.

Manuscript received 17 June 2013; revised 6 Apr. 2014; accepted 21 Apr. 2015. Date of publication 21 May 2015; date of current version 2 Mar. 2016. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below. Digital Object Identifier no. 10.1109/TMC.2015.2436393

90 percent. Although the use of some objects could not be detected with high accuracy, the 95 percent confidence intervals of all results were within 5 percent of the average performance for objects detected with high and low accuracy, indicating consistency across different resuscitations and teams.

We make three contributions in this work. First, we provide *a method for detecting the use of objects in complex activities based on passive RFID*. Our RFID setup was optimized to capture the change in object state instead of just high read rates. Our data processing algorithms are based on machine learning and postprocessing for improved performance. Second, we provide *evidence that passive RFID can be used for object use detection with up to 90 percent accuracy, as well as for activity recognition*. Experimental data was collected in an actual trauma bay. Medical personnel were unaware of our data collection and worked in a natural way. We used off-the-shelf RFID equipment, showing a promise for future use of passive RFID in similar patient-care settings. And third, we offer *insights into the strengths and limitations of passive RFID*. Our results showed that RFID consistently performed well for some objects and less well for others, thereby identifying objects that will require other sensors to complement RFID.

2 Related Work

Sensor-based methods are becoming widespread for human activity recognition. While simple activities can be recognized using on-body sensors [4], complex activities require additional cues such as body location [5], [12], [14], speech [14], or objects in use [5], [6], [10], [19], [28].

Objects in use have been found valuable for identifying tasks during surgery and other medical events [1], [5], [10], [19], daily-living activities (e.g., making coffee) [6], [18], and car manufacturing activities (e.g., closing the engine hood) [28]. Agarwal et al. used RFID to track people and medications to infer events during surgery [1]. Ohashi et al. tracked medications and blood administration by equipping carts with RFID readers (read range $\leq$ 10 cm), and placing tagged objects on the cart for detection [19]. Although feasible for bedside patient care, this approach cannot be applied in the operating room, where several people work in a larger space. To remove this restriction, wearable RFID readers have been used for tracking nursing tasks [10] and surgery phases [5].

Near-field RFID technologies (e.g., wearable readers) can achieve high accuracy of interaction detection, but they have three limitations. First, they require human participation, which is intrusive in real-world applications [5]. Even in a relaxed home setting, participants forgot to wear the readers or grasped objects with a non-equipped hand [18]. Second, near-field readers are not feasible for long-term experiments in a clinical setting as they may hinder patient care. This limitation affects our work because we are continuously running experiments and collecting data in the actual trauma bay, rather than arranging only a few experiments. To ensure minimal intrusion, we designed our tagging approach and RFID antennas setup in the trauma bay using the deployment and evaluation methods developed in our earlier work [21]. Finally, near-field readers provide binary detection information, rather than signal strength values. Although received signal strength indication (RSSI) tends to be noisy for passive RFID, it contains rich information that can be extracted using multiple, spatially distributed readers and data processing techniques.

We used six ceiling-mounted antennas, with read ranges of 3 to 4 m, and demonstrated that tag mobility (due to object use) can be inferred from the RSSI. Long-range tag motion detection has been studied with fixed RFID readers and algorithms to detect fluctuations in RSSI [11], [27]. Our work differs in both experimental and algorithmic approaches in three distinct ways. First, we conducted continuous, long-term experiments in a time-critical, crowded setting, unlike home [11] or office settings [27]. Second, medical personnel participating in our experiments were unaware of RFID data collection. Finally, we applied machine-learning techniques, rather than simple logical rules to address data processing challenges in a realistic setting characterized by many objects and usage patterns.

Many prior studies have leveraged machine learning for activity recognition in various problem domains. For example, Hidden Markov Models (HMMs) were used to recognize daily-living activities (e.g., cereal making and teeth brushing [6]), physical activities (e.g., standing, walking, and running [15]), and office activities (e.g., phone conversations and face-to-face interactions [20]). Similarly, previous work has made use of conditional random fields for recognition of GPS-based activities [17] and home activities [30], and also leveraged dynamic Bayesian networks for recognizing daily activities [25]. Finally, support vector machines have been used for activitiy monitoring of the elderly [8], Adaboost and HMM for recognizing physical activities [16], and rule- and tree-based classifiers for activitiy recognition [5], [13]. To our knowledge, however, no study has used machine learning techniques for object use detection based on RSSI data from passive RFID. In particular, the noisiness of the RFID data in our problem domain required careful feature extraction and postprocessing, in conjunction with traditional machine learning methods for classification.

Our early work focused on the components needed to build the overall system for object use detection in the trauma bay [21], [22], [23]. In contrast, this paper focuses on the details and evaluation of the complete system for object use detection in a realistic setting. Specifically, we started with mock resuscitations in a laboratory setting with teams of two experimenters (non-medical students) and nine objects [23]. We built a system for detecting the use of these objects based on their motion and location, which achieved an average precision rate of 63.8 percent and a recall rate of 90.6 percent [23]. We then used these data to experiment with various feature sets and classifiers for detecting whether an object is in motion from passive RFID data [22]. While this previous work provided insights into object use detection from RSSI data [22], [23], their setups were simpler than the realistic experimental setup used in this paper. Finally, the same laboratory setting was used to analyze the performance of RFID antennas and tags placement [21]. This paper leverages the lessons learned from our early studies for the following: optimal placement of RFID antennas in an actual trauma bay; deployment of a comprehensive RFID-based object use detection system, including placement of tags on numerous medical objects inside the

TABLE 1
List of Tagged Objects, Number of Tags, and Activity Involving the Object

Object (# of tags)	Activity
Cervical collar (2)	Neck immobilization
Stethoscope (2)	Chest auscultation
Thermometer (1)	Temperature measurement
Laryngoscope (1)	Intubation
CO_2 indicator (2)	Intubation
Endotracheal (ET) tube (4)	Intubation
IV fluid bag (8)	Fluid administration
IV catheter (31)	IV line placement
IV start kit (16)	IV line placement
IV tubing (4)	Fluid administration
Bag valve mask (2)	Ventilation
Rapid infuser tubing (1)	Rapid fluid infusion
Otoscope (1)	Ear assessment
Ophthalmoscope (1)	Eye assessment
Broselow tape (1)	Patient weight estimation
Foley catheter (1)	Urine assessment
Orogastric tube (1)	Gastric decompression
Blood pressure (BP) cuff (1)	BP measurement
Intraossesous access drill (1)	IO catheter placement
Team role tags (8)	n/a

Manually annotated object types during video review are highlighted in gray.

actual trauma bay; extensive collection of RSSI data during realistic resuscitation simulations in the trauma bay; and evaluation of our approach for object use detection in the trauma bay using this large collection of realistic RSSI data.

3 Study Setup

We next describe our RFID deployment at the research site, data collection methods, and the recorded dataset.

3.1 Identifying Activities and Objects for Tracking

Using a hierarchical task analysis performed by medical experts on our research team, we created a list of activities that were judged critical for the performance of trauma resuscitation (Table 1). If omitted or performed incorrectly, these activities could lead to adverse patient outcomes. We excluded low-level subtasks such as opening the patient's mouth or removing a stylet from an endotracheal tube because their detection is challenging and noncritical for achieving our goals [24].

3.2 RFID Tag Type and Placement

RFID tag type and position were determined based on the composition and size of objects, as well as their usage patterns. We evaluated several tagging configurations for each of these factors and selected the configuration for each object that yielded the highest RFID read rates [21].

1) *Tag type*: We selected the tag type based on object material. Passive RFID tags perform poorly when attached to metallic surfaces or liquid containers. We tagged metallic items (e.g., laryngoscope) with rigid on-metal tags. Although these special tags are expensive, they are feasible for reusable items. Liquid containers (e.g., IV fluid bags) and objects in aluminum packaging (e.g., CO_2 detector) were tagged with regular tags, keeping the tag contact with liquid or aluminum minimal.
2) *Number of tags and their placement*: For each object, we identified surfaces available for tagging and selected the largest fitting tag. The surface availability depended on (a) *object protection*: for sterile objects, only the wrapping could be tagged; (b) *shape constraints*: flat surfaces were preferred as tag folding degrades performance; (c) *smoothness*: tags adhered better to smooth surfaces; (d) *size*: most objects were tagged with two RFID tags for robust detection so that if one tag was not readable, the other might still be readable (single tag was used on small objects to avoid signal distortion [11]); and (e) *duration of contact:* if an object was held longer than 10 seconds, we placed one tag at a point of contact with provider's hand or patient's body, and the other tag at the location where it would remain exposed ("tandem tagging" [24]). Weaker or no signal from the first tag and strong signal from the second tag indicated that the object was in use.

Fig. 1. Tagged objects (tags circled). Digital thermometer (upper left); fluid bag (lower left); and bag valve mask or BVM (right).

We used 109 passive RFID tags for tagging 81 objects of 19 types (Table 1). Examples of tagged objects include a thermometer, intravenous fluid bag, and bag valve mask (Fig. 1). During a pilot resuscitation, we noted that some tags were not detectable (e.g., on the stethoscope or thermometer probe). These tags were small and folded around the object shape. In these cases, we replaced the initial tag with a larger one if it did not interfere with the object's use. If a larger tag was not feasible, we kept the small tag but relocated it to improve detection rates.

3.3 RFID Antenna Placement

Our goal was to ensure coverage of the entire trauma bay and maximize object use detection with a minimum number of antennas due to the lack of space, possible interference with medical equipment and cost. To determine the optimal antenna placement, we studied providers' locations and their interactions with medical objects. Using this analysis, we divided the workspace into five zones: patient-bed zone,

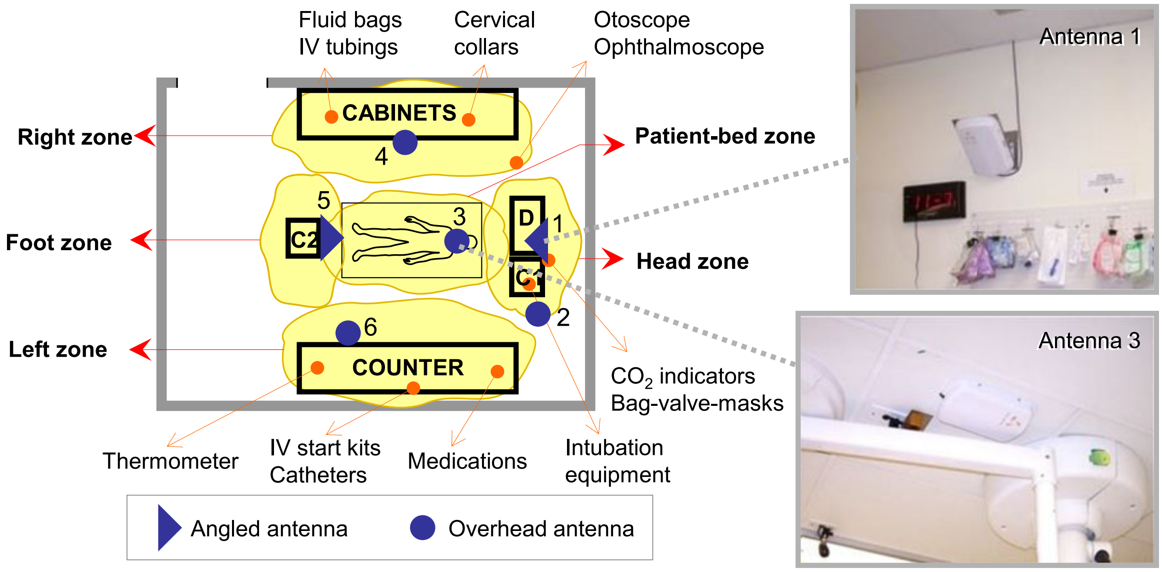

Fig. 2. Environmental setting of the trauma bay: zones where medical objects appear during resuscitation, storage locations for supplies and equipment, and antenna positions with shaded areas of antenna coverage. Actual deployment of antennas #1 and #3 is shown on the right.

right and left zones, and foot and head zones (Fig. 2). When in use, objects appear in the patient-bed zone; when stored or left idle, objects appear in the left, right, and head zones. The foot zone is rarely used for storing or using objects, so we ignored this area.

We evaluated several antenna setups to achieve high readout rates and maximize deviation in the RSSI signal when objects change status [21]. Our configuration chosen for the experiments included: *(1) one ceiling-mounted antenna above each zone*, facing the floor to reduce the effects of human presence and movement on RFID tracking (filled circles in Fig. 2 and shaded coverage areas); and *(2) two additional antennas mounted in the patient-bed zone* to improve signal detection rates, as well as the accuracy of localization and movement detection (filled triangles in Fig. 2). These antennas were angled to face the bed and were >2 meters above the floor to avoid work obstruction, and to reduce interference caused by human occlusion and movement. The actual deployment of antennas #1 and #3 is shown in Fig. 2 (right).

We used RFID readers from Alien Technology (ALR-9900+) [2]. We installed two readers in the trauma bay, hidden in a space above the ceiling. Antennas #1, #2 and #3 were connected to reader 1, and antennas #4, #5 and #6 were connected to reader 2. These associations allowed a pair of antennas to be simultaneously active: the pairs 1-4, 2-5, and 3-6 were activated sequentially. To reduce interference, antennas scanning the patient bed zone were never active at the same time.

3.4 RFID Data Collection

RFID data were collected during 32 simulated resuscitations, each about 20 min long. All teams performed four resuscitation scenarios with injuries requiring different treatments, including endotracheal intubation, administration of fluids and medications, temperature control and chest-tube insertion. The use of tagged objects of each type depended on the scenario. For example, we tagged two cervical collars of different sizes because scenarios involved different patient types (i.e., infant and child).

Before resuscitation, the tagged objects were stored in cabinets. A member of the research team initiated RFID data collection when the patient simulator arrived in the room. RFID readers operated autonomously and were continuously scanning the environment until stopped.

3.4.1 Ground Truth Data for Algorithm Evaluation

Object interactions were manually annotated using video review of each resuscitation. Annotation included the event ID, object ID (e.g., large collar, stethoscope 2), interaction start and end times, and activity involving the object, if any. Of the 81 tagged objects, we annotated interactions with 73 objects of 11 different types across all 32 resuscitations (Table 1). The remaining objects either had low data rates, preventing algorithm training (e.g., otoscope and ophthalmoscope), or were not used in a significant number of resuscitations (e.g., Broselow tape, Foley catheter, intraosseous access drill). We decided not to annotate interactions with the blood pressure cuff because we could detect its use from the vital signs monitor.

3.5 RFID Data Analysis

Average read rates indicated which objects may or may not be easily detected. Objects with low rates were problematic because of insufficient data for processing. We calculated the average read rate for an object by normalizing the total number of readings with the number of resuscitations and the number of objects of that type.

The average read rates varied among objects (Fig. 3) for several reasons:

1) *Irregular object shapes required tag folding or a smaller tag, weakening the ability of the reader to detect the tag.* Examples include the otoscope, ophthalmoscope, and stethoscope ("**IS**" labeled columns in Fig. 3).
2) *Some objects were stored in locations with weak antenna coverage.* Examples include the otoscope, ophthalmoscope and CO_2 detector. The otoscope

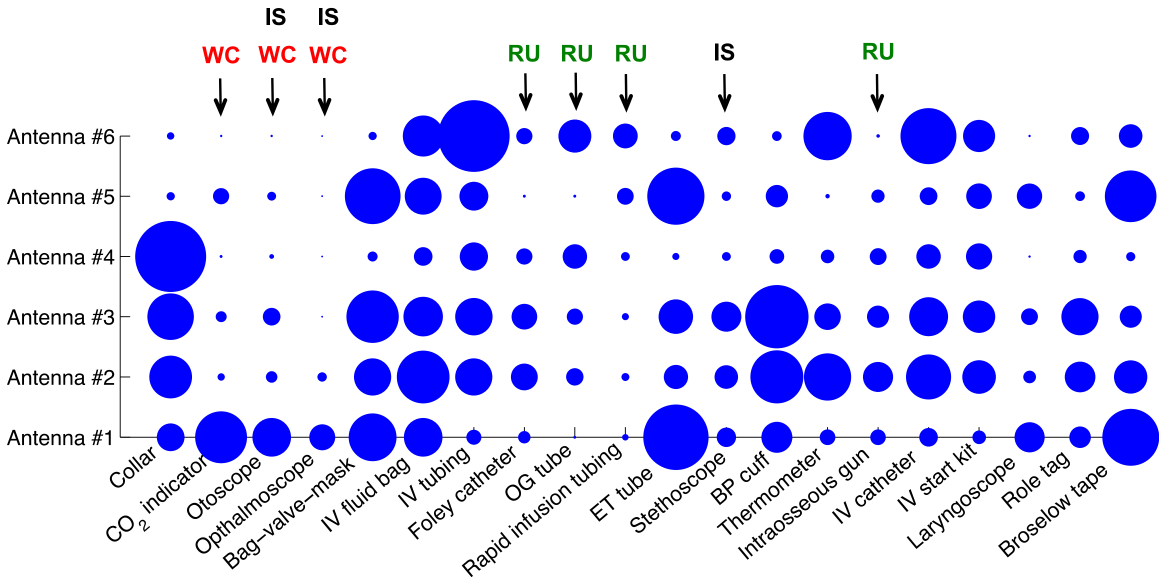

Fig. 3. Average read rates for each object and antenna (averaged over 32 resuscitations and number of objects of the same type). Larger bubbles indicate higher read rates. (None: 0 readings/resuscitation; Largest: 1,388 readings/resuscitation). Arrows indicate the objects with low number of readings, due to irregular shape, weak coverage (WC), or rare use (RU).

and ophthalmoscope were mounted on a movable mechanical arm, usually located between the head and right zones (Fig. 2). The CO_2 detector was detectable only intermittently while stored, hanging on the wall ("**WC**" labeled columns in Fig. 3).

3) Some objects were used rarely. Objects stored in cabinets could only be detected when taken out. Their read rates were high after they appeared or when used, but their total read rates were low. Examples are the orogastric tube, intraosseous access drill and tubing for rapid fluid infusion ("**RU**" labeled columns in Fig. 3).

Read rates of antennas varied by object trajectory, which in turn depended on where the object was stored and used (Fig. 3). For example, when not in use, the cervical collar was stored in the right zone, with antenna #4 having the best reception; when the collar was brought to the patient bed area, reception from antennas #1, #2 and #3 increased; when the collar was placed around the patient's neck, antenna #5 was the most distant and had the lowest reception. This change in use status caused a large fluctuation in the RSSI of the collar, which facilitated interaction detection. Detecting object interaction by using data from a single antenna (e.g., CO_2 detector and otoscope) or continually from the same set of antennas (e.g., stethoscopes are carried around the neck and mostly stay in the patient-bed zone) was challenging. Interaction detection was easier for objects moving across zones because the reception of different antennas depended on object location (e.g., cervical collar, IV fluid bag, bag-valve mask, IV tubing, IV catheter, and thermometer).

4 Method for Detecting Objects in Use

Medical objects differ based on their storage location, usage pattern and interaction style. Manually defined rules for use detection are therefore not feasible. We formulated object use detection as a binary classification problem and developed a machine-learning-based strategy with three steps: feature extraction from RSSI data, classification and postprocessing.

4.1 Feature Extraction

For each object, we segmented RSSI data into fixed-size overlapping windows and extracted the relevant features from each window (Fig. 4). To determine the relevant features, we analyzed interactions with objects and identified three main cues that indicate objects use [24]:

1) *Zone-based location:* Objects are fetched from their storage and moved to a place where they are used. We used RSSI from different antennas to represent the zone-based location (Table 2).
2) *Motion:* Movement of a used object (and its attached tag) causes fluctuations in RSSI, which can be detected by quantifying the variability in the current window. Our previous work showed the feasibility of this approach in a crowded laboratory setting [22]. To quantify this variability, we divided a time window into left (L) and right (R) sub-windows, and computed statistics to quantify the dissimilarity between them (Table 2). Because moving objects are more likely to be detected by several antennas, we also counted the number of antennas that detected the object at least once in the current window.
3) *Contact:* Contact between an object and a provider or patient means that the object is likely in use. For objects with tandem tagging, we expected strong signal from both tags when the object was idle. When in use, the tag in contact with a provider or patient emitted weaker signal or no signal at all. We detected contact based on the percentage RSSI contributed by each tag on the object (Table 2).

A cue may show different characteristics depending on the object type. For example, when in use, a cervical collar is located at the head of the bed, while an IV fluid bag is located at the left side. These object locations generate

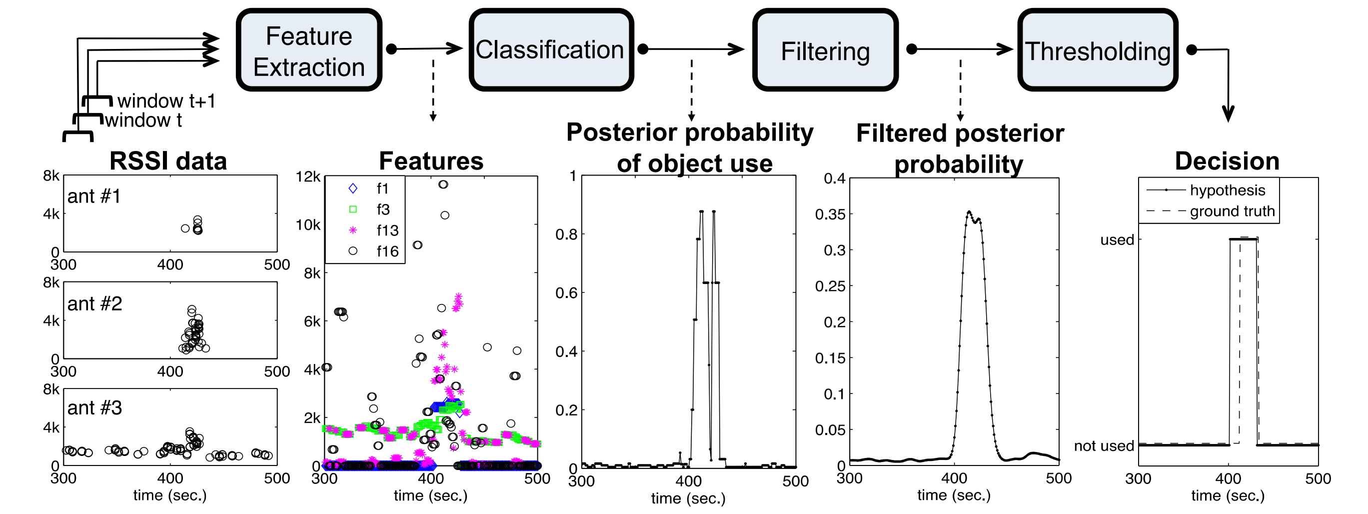

Fig. 4. A schematic illustrating our machine learning-based approach for detecting the use of objects.

different RSSI patterns, although they both mean *"in use."* Each cue then must be interpreted separately for each object type. To uncover these cues, we extracted 18 features from the RSSI data and concatenated them to a feature vector. To handle the feature variance across object types, we trained a separate classifier for each object type.

4.2 Classifier Training

We used a supervised approach for training a classifier. Our dataset included RSSI data from tags and manually generated annotations. Features extracted from RSSI and annotations served as input to a learning algorithm that outputted a classifier. In the testing phase, the features were mapped to binary labels using the classifier (Fig. 4).

By visualizing the distribution of features for our dataset, we found they could not be modeled with a simple and well-known probability distribution. Instead, we used algorithms that directly learn the discriminator via discriminative classification methods [3]. We evaluated four rule-based and tree-based classifiers that have performed well in similar detection tasks [5], [13]: decision trees, random forests, boosting, and JRip. To train and test these classifiers, we used the Weka data mining software [8].

TABLE 2
Cues for Detecting Object Use and Related Features Extracted from RSSI to Capture Those Cues

Cue	Feature
Location	Average RSSI (from each antenna)
Motion	Difference of average RSSI between *L*eft and *R*ight sub-window (from each antenna)
	Total difference between L and R
	Spearman Rank Correlation Coefficient between L and R
	Number of antennas that are common in L and R
	Mahalanobis distance between L and R
	Number of visible antennas
Contact	Percentage RSSI contributed by each tag on the object

4.3 Postprocessing

A classifier generates hypothesized labels as outputs, but it does not provide information about confidence of the hypotheses. We obtained confidence estimates (posterior probabilities) of the binary labels by fitting a logistic regression model into the classification output [3]. We also processed the posterior probability sequence with (Fig. 4):

1) *Smoothing*: The sequence was filtered with a smoothing Gaussian filter to eliminate sudden jumps in the posterior probability sequence.
2) *Thresholding*: An instance of object use was declared if the smoothed probability values exceeded a threshold. A precision-recall curve was obtained by adjusting this threshold.
3) *Merging adjacent instances of use*: If two use instances were $<$ 30 sec apart, we merged them to a single event. The rationale was that proximate uses of the same object most likely represent one activity.
4) *Eliminating short interactions*: We removed use instances shorter than a specified time interval derived from annotated data, depending on the object type. For example, annotations showed that the mean use time for thermometer was 23 sec and the minimum use time was 12 sec. As a result, we set a 10 sec threshold for the thermometer.

5 Evaluation

We next describe evaluation of the object use detection performance in relation to two aspects:

- *Object instance detection:* Identifying which instance of a given object type was used during an event (e.g., IV catheters #5 and #6 were used in event #17). By knowing object instances we can obtain the count and parameters of used objects; this is important contextual information that teams may overlook or forget.
- *Time of object use:* Detecting the exact time interval of object use (e.g., IV catheter #6 was used between 126-149 seconds in event #17). By finding the time of object use, we can track and analyze team activities.

5.1 Evaluation Method and Metrics

The performance of object-instance and use detection was evaluated by five-fold cross-validation. Each resuscitation event produced a separate data sequence. We divided the set of these sequences into five subsets to apply five-fold cross-validation. The sequences were not segmented into pieces and each sequence appeared unbroken in the training or testing set.

We evaluated use detection performance with three sets of metrics. The first set was *precision* and *recall*, widely used in detection problems [26]. We also used *F-measure* (the harmonic mean of precision and recall) as a combined measure of performance. Precision, recall and F-measure vary in the range of [0, 1]. High precision rates mean that most detections are correct (few false alarms), whereas high recall rates mean that most true instances are detected (few misses). Although precision, recall and F-measure show the number of false positives and false negatives, they do not specify the type of misses and false alarms. For example, a false alarm may be due to a late detection or complete miss, and one of these errors may be more serious than the other. To address this drawback, we used a second set of metrics defined by Ward et al. [31]. False negatives were categorized into three types:

1) *Underfill:* Predicted segment matches a ground truth segment but partially misses at the start or end (underfill_start, underfill_end).
2) *Fragmentation:* Two or more predicted segments match a ground truth segment.
3) *Deletion:* A ground truth segment is not matched.

False positives were also categorized into three types:

1) *Overfill*: Predicted segment matches a ground truth segment with a spill at the start or end (overfill_start, overfill_end).
2) *Merge*: Predicted interval includes two or more ground truth segments.
3) *Insertion*: Predicted interval does not match a ground truth segment.

Another flaw of precision and recall metrics is that they do not account for true negatives, especially when the data is skewed, potentially biasing the calculations. Our third set of metrics included unbiased measures of performance: *informedness*, *markedness* and *Matthews correlation coefficient (MCC)* [26]. MCC is considered one of the best representations of a confusion matrix because it incorporates all of its elements. Informedness and markedness vary in the range of [0, 1]. Because maximum value for inverse precision and inverse recall is 100 percent, informedness and markedness are often smaller than precision and recall. MCC measures the goodness of a hypothesis by computing its correlation with the ground truth: the greater the correlation, the more accurate the hypothesis. Being a correlation-based metric, MCC varies in the range $[-1, 1]$. MCC values close to 1 indicate high correlation between the ground truth and hypothesis, and hence good use-detection performance. An MCC of zero means that the hypothesis and ground truth are not correlated, indicating poor performance. Negative MCC values represent inverse correlation between two signals. When evaluating detection performance, however, inverse correlation is not relevant because negative performance scores are not valid. We thus observed negative MCCs only when they were close to zero.

5.2 Experimental Results

We evaluated the use detection performance on 72 objects of 11 commonly used types: two cervical collars, two stethoscopes (one tagged with two tags; one tagged with four tags), thermometer, laryngoscope, two CO_2 indicators, four ET tubes, two bag-valve-masks, eight IV fluid bags, 31 IV catheters, 16 IV start kits and four sets of IV tubing. The two stethoscopes were evaluated separately (Table 1).

5.2.1 Characterizing Passive RFID Performance for Detecting Object Parameters

Certain types of objects are available in different sizes to accommodate different patients and needs, e.g., cervical collars, endotracheal (ET) tubes and IV catheters. IV fluid bags may also contain fluids with different compositions. The parameters for these objects (e.g., size, volume) are determined based on the patient's age, weight or current condition. Using objects with improper parameters can be a medical error and may lead to adverse outcomes.

Objects of the same type but with different parameters often have similar shape or packaging. It is difficult to identify the parameters through analysis of videos. In the resuscitations that we videotaped, only parameters of the cervical collars and CO_2 indicators were visible because collars had different sizes and CO_2 indicators were in packages of different colors. The contents of fluid bags and size of IV catheters and endotracheal tubes, on the other hand, were not distinguishable from the video.

Through human and machine vision, the object type is first recognized (e.g., a cervical collar), and then its parameters are found (e.g., a *small* cervical collar). Using identification technologies, on the other hand, both object type and parameters can be found simultaneously. When an RFID tag is detected as *in use*, the object parameters are retrieved from a database using the tag ID. Accurate knowledge of objects and their time of use perfectly convey the parameter information for the used objects. In cases during which only use time is inaccurate (detected time interval does not exactly match the true time interval), RFID still provides the parameters of objects that were used in a particular resuscitation.

In this experiment, we evaluated long-range passive RFID technology for identifying specific instances of object type used during resuscitations. Once the exact instance was detected, identifying its parameters was straightforward. We first ran our use detection algorithm for each resuscitation. If an object was detected as being used at least once, we assumed a positive detection for the object in that resuscitation. Similarly, if ground truth was positive at least once throughout the resuscitation, we assumed a positive ground truth for the object in that resuscitation. Here we report results for eight out of 11 types of objects because multiple instances were tagged for these eight types: cervical collar (2), CO_2 indicator (2), endotracheal tube (4), bag-valve-mask (BMV) (2), IV fluid bag (8), IV start kit (16), IV catheter (31) and IV tubing (4).

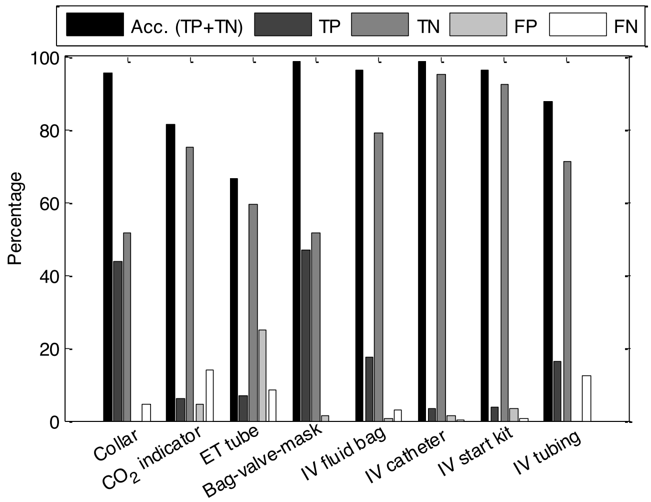

Fig. 5. Performance scores for object parameter detection. (Acc.: accuracy, TP: true positive, TN: true negative, FP: false positive, and FN: false negative).

For all types, except the ET tube and CO_2 indicator, we were able to detect the exact object instance, and hence its parameters, with an accuracy of > 90 percent, where accuracy implies the sum of true positive (TP) and true negative (TN) rates (Fig. 5, first bar in each group). For some objects (e.g., collar, BVM), TP and TN rates contributed almost equally. For others (e.g., IV start kit, catheter, fluid bag), TN rate was significantly greater than TP rate. The reason is that object types with many instances generated high TN rates. For example, we tagged eight IV fluid bags and two collars. When a team used one instance of each type, the used instances (one collar and one fluid bag) represented ground truth positives, whereas unused instances (one collar and seven fluid bags) represented ground truth negatives. Because fluid bags had more unused instances, these objects had a greater TN rate.

The use of CO_2 indicators was missed primarily due to low read rates (Section 3.5). False positive (FP) rate was high for endotracheal tubes because they were in antenna's view even when not in use. Locations of used and unused tubes were in close proximity and could not be distinguished using RFID. In adition, team members interacted with almost all tubes when searching for the appropriate one, triggering more false alarms.

5.2.2 Characterizing Passive RFID Performance for Detecting Time of Object Use

To detect the time and duration of object use, we extracted features from the RSSI data using a 12-second sliding window, and then trained a separate Logitboost classifier for each object type (a total of 12 classifiers). Our findings from this experiment are as follows:

*1) Use detection performance depended on the object type and its storage.*We obtained the best MCC scores for cervical collars (82.4 percent), bag-valve mask (84.2 percent), fluid bags (53.4 percent), and thermometer (40.6 percent) (Table 3). Performance on some objects was low because of an unbalanced training dataset, acquired from actual work where some objects are used frequently and others rarely. All objects, except the bag-valve mask, were relocated from left or right zones to the patient-bed zone when needed. Relocations from other zones to the patient-bed zone or within patient-bed zone were more difficult to detect because the displacement was relatively small. The bag-valve mask was stored in the head zone, hung on the wall above the patient bed. When needed, this object was relocated to the patient-bed zone and to a lower height, which increased the fluctuations in the RSSI signal and facilitated use detection. The MCC for thermometer was lower than for collars or fluid bags for several reasons. First, the thermometer was often brought to the patient bed long before its use. When it was actually needed, a nurse relocated it closer to the patient, but this move was not always noticeable in the RSSI signal. Second, collars, bag-valve mask and fluid bags remained in their location of use for long periods (e.g., once the collar was placed around the patient's neck, it stayed there throughout the resuscitation). In contrast, the thermometer was in use for a much shorter time. As the use time became shorter, it was more likely that the use moment was missed or confused with accidental movement.

TABLE 3
Precision (P), Recall (R), F-measure (F), Markedness (M), Informedness (I) and Matthews Correlation Coefficient for Several Objects

Tagged obj.	P	R	F	M	I	MCC
Collar	82.2	92.5	87.0	79.4	85.5	82.4 ± 1.3
Stethosc. 4 tags	30.4	24.7	27.6	21.2	17.6	19.5 ± 4.7
Stethosc. 2 tags	9.3	9.3	9.6	4.0	4.3	4.2 ± 3.6
Thermometer	24.8	72.6	37.1	24.1	67.0	40.6 ± 4.1
Laryngoscope	7.3	13.2	9.4	1.9	3.3	2.5 ± 2.2
CO_2 indicator	7.0	6.3	6.1	6.4	5.4	5.6 ± 5.7
ET tube	5.3	20.0	8.3	4.7	17.5	9.1 ± 6.3
BVM	92.0	88.1	90.1	85.0	83.5	84.2 ± 1.0
Fluid bag	73.7	44.7	55.2	67.5	42.7	53.4 ± 4.7
IV catheter	10.0	70.0	17.4	9.9	65.3	25.4 ± 5.0
IV start kit	4.9	22.3	8.0	4.8	21.8	10.2 ± 2.3
IV tubing	26.5	20.3	23.0	24.8	19.1	21.8 ± 3.1

Confidence interval for MCC at 95 percent confidence level is also shown.

For the thermometer and fluid bags, almost half of the misses were deletions and the remaining half were underfills and fragmentations (Fig. 6). The average length of underfill was 18 sec, which is short compared to the overall usage duration for these objects. This order-of-seconds lag in use detection (due to underfill) is not likely to be clinically relevant given the much longer period of use for these objects. Considering that human-generated annotations of start and end use times are error prone, as is video-to-RFID synchronization (see Section 6.5), the actual underfill interval may even be shorter. For the IV fluid bags, fragmentation rate was high because this object remained in use for a long time (Fig. 6). We believe that fragmentations are not serious errors for detection of this object and can be ignored because interruptions are unlikely once fluid administration has started.

We observed the lowest MCC for the stethoscope tagged with two tags (Table 3). Hypothesized labels for this object included a significant number of insertions and deletions, which constituted most false negatives and false positives (Fig. 6). For most errors, the predicted time interval did not overlap with the ground truth. Although using two

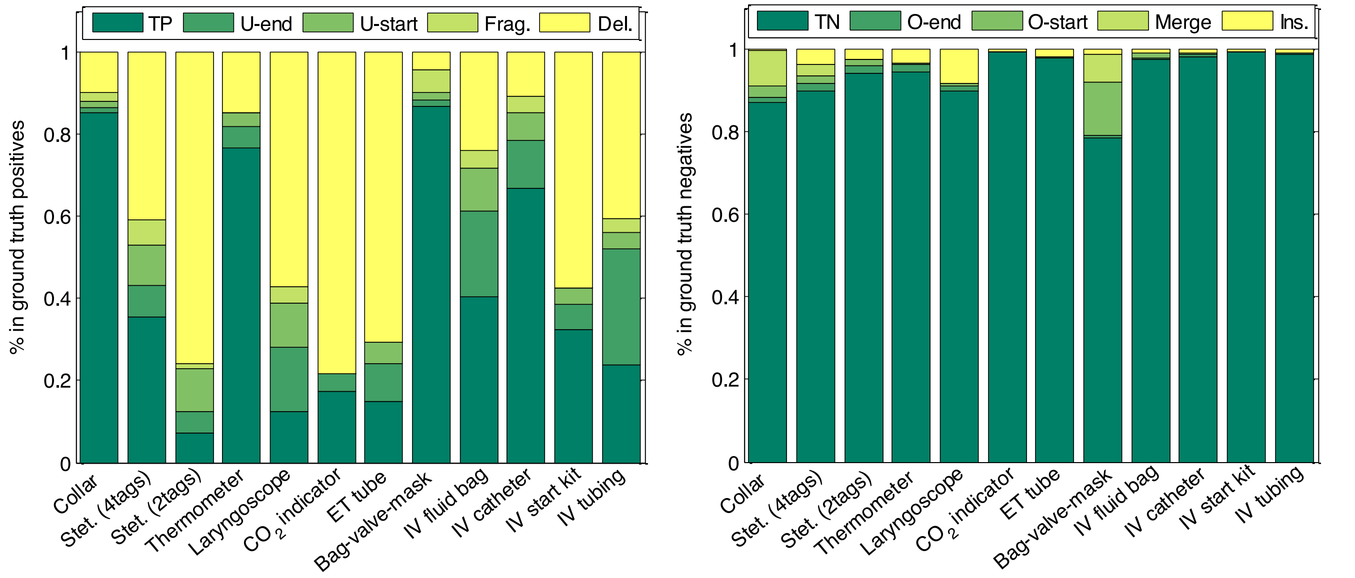

Fig. 6. Distribution of errors in object use detection. (TP: true positive, U-end: underfill-end, U-start: underfill-start, Frag.: fragmentation, Del.: deletion, TN: true negative, O-end: overfill-end, O-start: overfill-start, and Ins.: insertion).

additional tags provided improvements, MCC remained at about 20 percent. The stethoscope has challenging features for RFID-based detection. First, being carried by a provider at all times, it can be considered a personal object. Because providers are mobile and gather around the patient bed, stethoscopes are always in motion and close to the patient. Second, it is often in contact with human body. When the stethoscope was carried around the neck or held in hand, the tag was not detectable. When it was in use, the tagged part was not in contact with human body, but was instead occluded by the user leaning over the patient to listen to breath sounds. The data rate for this object was much lower than for other objects. Instead of disappearing only when covered by hands during use, one of the tandem tags often disappeared due to occlusion. Finally, we also observed variability in providers' usage patterns. When idle, some providers carried the stethoscope around their neck and others over their shoulder, with the stem part on the back. These factors made it challenging to detect the actual usage of the stethoscope.

Time of use for intubation equipment (laryngoscope, ET tube, and CO_2 indicator) could not be detected with high accuracy because these objects were stored in the head zone, which is closer to the patient bed compared to left and right zones. When a team decided to intubate the patient, these objects were brought near the patient's head for easy access, making it difficult to accurately detect the exact time of use. Similarly, IV catheters and IV start kits were usually prepared in advance and brought to the patient bed long before the actual use (Fig. 8, IV catheter). MCC was therefore greater for IV tubing because this object was brought at the time of use (Fig. 8).

We concluded that inter-zone relocations (especially from left and right zones to the patient-bed zone) were detectable using passive RFID. If object was fetched from its storage long before usage and then only relocated within the zone, the exact time of use could not be detected reliably using our current deployment of passive RFID technology.

2) *Greater data rates did not always lead to improved detection of object use*. Although low data rates affected the use detection performance, high data rates did not always improve use detection scores either. In our dataset, ET tubes generated high amounts of RFID data (Fig. 3), but their use instances could not be detected with high accuracy (Fig. 6, Table 3). This finding is important because a common criterion for the success of RFID systems is the read rate. High read rates, however, do not guarantee that a change will be detectable. A change in RSSI must be detected to decide whether an object is in use. When deploying RFID tags and antennas for use detection, we must first ensure that every object is detectable with sufficient data rates. As a second priority, the deployment strategy should maximize the change in signal pattern due to a change in object's state, instead of maximizing data rates.

3) *Confidence intervals of results were at most* $\pm$ *6.3 percent*. We ran the cross-validation ten times, calculated the performance metrics for each run, and found the averages to produce results (Table 3). Because training and test sets are determined randomly in each run, they are expected to produce different results. Reliability and repeatability of results, however, can be claimed only when the variance across different runs is low.

For analyzing the variation of results across different runs, we calculated the confidence interval at 95 percent for each object, reported for MCC in Table 3. The confidence interval was $\leq \pm 6.3$ percent, and much smaller for most objects. The consistent performance at both high- and low-performance ends provides an insight into the strengths and limitations of passive RFID, and aids in our understanding of how other sensory modalities may complement RFID. For object types with high use detection, passive RFID provided adequate performance. For challenging object types, such as stethoscope, other sensor modalities should be considered, such as active RFID tags or accelerometers. A key advantage of passive tags—the battery-free operation—may not be important for the stethoscope. Unlike other objects in the trauma bay, the stethoscope is a

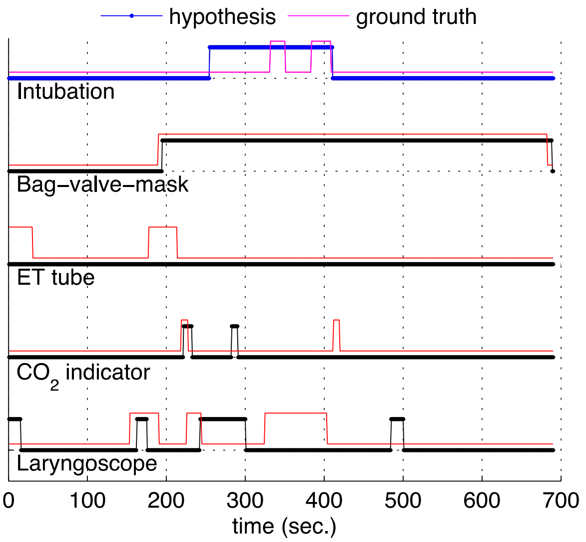

Fig. 7. Hypothesis and ground truth for detecting the intubation activity and associated used objects.

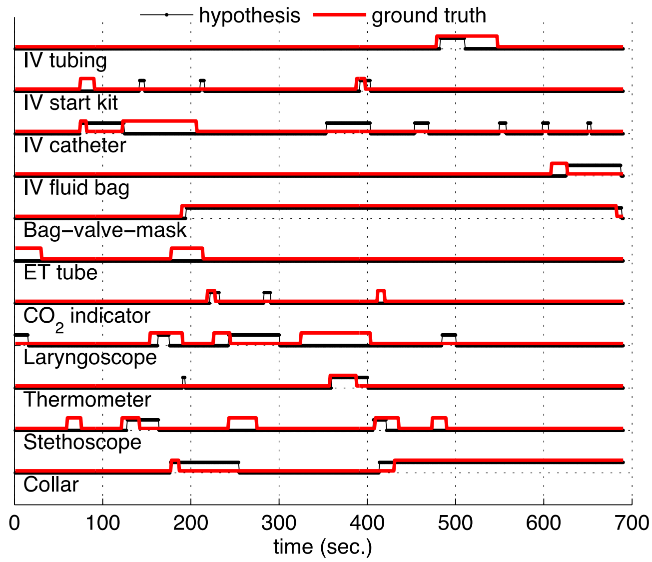

Fig. 8. Hypothesis and ground truth for detecting the object use in Resuscitation #17.

personal object, similar to a pager. It is reasonable to expect the owner to monitor and replace the batteries required for active RFID tags or accelerometers.

4) *Different metrics emphasized different aspects of performance evaluation*. We report results for three sets of metrics:

a) Set 1: Precision, recall, F-measure (Table 3).
b) Set 2: Informedness, markedness and Matthews correlation coefficient (Table 3).
c) Set 3: Distribution of correct and erroneous inferred labels, and type of errors (Fig. 6).

Because informedness and markedness are often smaller than precision and recall [26], we expected that the first set of metrics would yield higher values. Object types with few instances (e.g., collar, thermometer) met our expectation. Objects with many instances of the same type (e.g., IV catheter, ET tube), however, generated high TN rates, which compensated for the difference between the first and second sets of metrics. For detection problems with high number of ground-truth negatives, we conclude that informedness-markedness-MCC family of metrics provides a more objective evaluation by taking true negatives into account.

The third set of metrics (Fig. 6) highlighted the proportion of different kinds of errors. Because they are normalized by the number of ground truth negatives or positives, it was not possible to compare them directly with the first and second sets of metrics. For example, thermometer and IV catheter showed similar distribution of error types (Fig. 6). However, MCC for thermometer is 40.6 percent and MCC for IV catheter is 25.4 percent (Table 3).

6 Recognizing Activities from Used Objects

Although recognizing trauma team activities from used objects is outside the scope of this paper and part of our future work, we wanted to demonstrate how object use detection could lead to activity recognition.

For recognizing activities, we first ran the object-use detection algorithm to obtain decisions about object use. We then defined simple rules to infer the activity from the used object. For activities performed with a single object (e.g., neck immobilization, temperature measurement), the time of activity corresponded to the time of use for the associated object (i.e., temperature measurement was performed when the thermometer was used). The activity recognition performance would be close to the performance of detecting the associated object's use (Table 3).

For activities performed with multiple objects (e.g., intubation), the "activity" is usually defined conceptually and cannot be observed directly from sensory data. For example, the intubation activity roughly corresponds to the moment when the laryngoscope is used to hold the tongue back and the ET tube is inserted into the trachea (see the ground truth in the top graph of Fig. 7). Note that in Fig. 7 the interaction with the ET tube during the "intubation activity" period is not recorded although it took place. Because an RFID tag was placed on the ET tube packaging rather than the tube itself (required for sterility), the ground truth for ET tube only shows interactions with the ET-tube packaging. When the ET tube is removed from the packaging while the team prepares for intubation, the packaging is discarded along with the tag.

We defined simple rules to recognize the intubation activity, such that (*i*) the use of at least two intubation-related objects must be detected and (*ii*) their time of use must be sufficiently close. Despite high false alarm rate for the laryngoscope, and false alarms and misses for ET tube and CO_2 indicator, the intubation activity could be reliably detected. In this example, use detection occured >1 minute before the intubation actually started (overfill-start). If the use of laryngoscope could be detected more accurately, intubation could also be better localized in time. Errors in object use detection, therefore, cause a noisy input to activity recognition. In addition, contribution of each object to the activity is different (e.g., bag-valve mask is used throughout the intubation activity, but CO_2 indicator is used only for a short time). For these reasons, more complex and probabilistic rules are required for activity recognition, which will be addressed in our future work.

7 Discussion

7.1 Detecting Object Parameters with RFID

Using long-range passive RFID, we were able to identify the exact instance of a used object type (e.g., IV fluid bags #2 and #3 were used among all seven IV fluid bags) with accuracy rates $>$ 90 percent. By identifying the instance, we were also able to obtain the count and parameters of used objects, which are important for the resuscitation workflow. For example, the number of used IV fluid bags indicates the volume of administered IV fluid to the patient.

Object instance detection is similar to that of medication and blood tracking, which has been studied by others [14], [19]. These prior studies, however, were based on short-range technologies, such as barcode readers or low-frequency RFID. Short-range technologies require user participation in the sensing process. This requirement hinders work activities and may be forgotten or ignored. We showed that specific objects could be tracked using long-range passive RFID without the need for user participation in the sensing process. Although the current $>$ 90 percent accuracy rate for our method may appear inadequate for real-world scenarios, we believe that the accuracy will improve with technology advances, and the overall performance of competitive methods may be worse when accounted for unreliable user participation.

7.2 Detecting Time of Object Use with RFID

Our experimental results showed that detecting time of object use depended on three main factors:

1) *Storage location relative to use location*: Objects with best detection rates included cervical collars, IV fluid bags, IV tubing sets, and the thermometer. These objects are stored in the left or right zones, which are sufficiently distant from the patient bed. Relocation from their storage to usage zone was clearer in the RSSI sequence, compared to relocation from the head zone to the patient bed.
2) *Duration of use*: Cervical collars and fluid bags are in use for an extended time, leading to fewer chances for misses or confusion with other disturbances. As the duration of interaction with the tag increased, detection rates improved.
3) *Time gap between relocation and use*: Some objects are prepared and brought to the patient bed long before use, even before patient arrival. Examples include IV start kits and IV catheters in almost all resuscitations, and occasionally cervical collars (Fig. 8). These objects are small and do not occupy much space on the patient bed, but make detection of the exact use time more difficult.

7.3 Scalability of RFID-Based Methods in Medical Settings

RFID technology deployment in a hospital setting requires mounting multiple antennas, coordinating RFID readers and tagging objects. Through this research, we have designed and evaluated approaches for placing RFID antennas, readers, and tags in the trauma bay. Because the arrangement of equipment, patient bed and medical personnel across trauma centers is similar, we believe that placement of multiple antennas and readers in any trauma bay can follow our approach of dividing the room into zones, with some modifications based on the room parameters [21], [24]. Once this process is completed, placing antennas and readers is a one-time operation. For our study, the hospital staff mounted the antennas based on instructions from our research team and the process took only few hours without disturbing the day-to-day activities. Post-installation maintenance of these antennas requires considerably less effort than that of most other medical equipment in the hospital.

Although tagging every medical object requires effort, our study provided guidance for placing tags based on the size and shape of an object, potentially decreasing the time it takes to place tags. In addition, the cost of RFID tags is declining, making it more likely that RFID tags will soon become as integral a part of medical objects as the barcodes are today. Finally, passive RFID is already widely used in supply-chain management, where manufacturers place passive RFID tags on countless objects for inventory control [28], [32]. Once placed on objects, passive RFID tags are far less intrusive compared to other sensors. As we described before, participants in our study were not aware of tags on medical objects because the tags did not obstruct interaction with the objects. This nonintrusiveness during work has a significant advantage that can compensate for the effort of placing tags on objects before or after resuscitation.

Finally, the trained machine learning models for object use detection depend on the room layout. The layout of resuscitation areas, however, is similar at most trauma centers. This similarity suggests that a model trained for one room can be used as a baseline for another room. Differences can still be expected in use-activity signatures from one room to another. These differences can be resolved by parameter calibration or model adaptation (similar to adapting a speech recognition system, such as Siri, to an unknown user). In particular, data required for such adaptations might be collected by choreographing a series of activities including multiple objects, either during the installation phase of the system or on a regular basis post installation.

7.4 Limitations of RFID and Potential Sensors to Complement RFID

Our long-term goal is to develop a context-aware system that automatically acquires information about human activities and provides real-time feedback to improve the efficiency of the trauma resuscitation process. These types of systems must be error-free to provide benefits in a realistic scenario. Building such a system using only one type of sensors (e.g., RFID) is not feasible because no single modality can be expected to satisfy all requirements independently. Instead, it is required to have a combination of different approaches and technologies, such as RFID, computer vision, and other sensors, as well as a detailed plan on how to fuse each one of these sensor outputs for an optimum decision making. In this paper, we took an initial step toward this goal and analyzed how passive RFID can play a role in such a multi-modal system by identifying the strengths and limitations of this technology for detecting objects in use and related activities. We found that passive

RFID performance scores are promising and that complemeting this technology with other modalities is likely to lead to acceptable error rates. Real life deployment of our system therefore requires integration with other sensing technologies so that all medical objects and tasks can be detected with sufficient accuracy.

A major limitation of sensor-based (including RFID-based) activity tracking is the inability to recognize activities performed without instruments, such as manual palpations, pulse assessment or verbal statements. These actions are often a part of complex activity, and are thus important to detect. Vision or speech-based technologies could be used for detecting these kinds of activities.

Sterile items in the trauma bay, such as ET tubes and CO_2 indicators, are stored in sterile wrappings. A sensor can only be attached to the wrapping, which prevents tracking after wrapping is discarded. Upon wrapping removal, the item may be used with a delay or not at all. Using active sensors for these objects is not feasible because the sensor would be discarded along with the wrapping. Although machine vision is an option, discerning wrapped objects is challenging because most types of wrapped objects have similar packaging. Vision may be combined to pick up from where RFID left off and continue object tracking after its wrapping is removed. In addition, if multiple objects are used for the same activity, it is possible to improve the recognition of the overall activity based on use detection of those objects (Section 6).

Finally, we faced challenges when attaching RFID tags to objects with irregular shapes (ISs) and uneven surfaces (e.g., otoscope, stethoscope and laryngoscope). As a result, the reception from these objects was relatively low, as were the use detection rates. These objects were mostly personal or relatively expensive (i.e., not disposable). Using active RFID tags or accelerometers may be a feasible solution for detecting the use of such objects.

7.5 Practical Observations

7.5.1 Errors in Annotations Indicated by RFID

We observed two types of errors in the ground-truth annotations, each serving as a realistic comparison between human visual processing and RFID-based data mining. The first error was annotating the wrong instance of an object. We tagged two cervical collars, one small and one large. In two out of 32 recordings, the large collar was annotated as used although the small collar was actually used. Based on our discussion with the annotator, the trauma team should have used the large collar in these scenarios based on the patient age. The RFID detection algorithm caught a team error that was not noticed by the annotator. This finding highlights the strength of RFID in identifying the parameters of objects (e.g., volume, size), similar to medication and fluid tracking applications using near-field technologies. Our experiments showed that passive RFID technology with long-range readers could be used for identifying the parameters of used objects.

The second error was annotating objects that were not part of the experiment. An untagged object was mistakenly used instead of a tagged object of the same type in nine out of 32 resuscitations. Because the tag is small, it was difficult to see in the video whether the used object was tagged. These instances were incorrectly annotated and represented erroneous ground truth (detected and fixed before we ran our experiments reported in this paper). In contrast, RFID data showed that the tagged objects were in storage. This observation indicates the strength of RFID in detecting the location or use of small objects.

7.5.2 Video Review and Annotation Challenges

We annotated the ground truth of object interactions for our system training and evaluation. This effort pointed to several practical challenges stressing the need for an automated system for detecting and analyzing trauma team errors. First, annotating object interactions by reviewing videotaped resuscitations was a laborious task—it took approximately 40 minutes to annotate a 10-minute video for an object that was used or relocated frequently. The time required for annotation varied widely based on how often the object was moved (stethoscope versus cervical collar) and how difficult it was to see it in the video (transparent ET tubes versus thermometer). In addition, identifying object parameters (and individual object instances) from video recordings was often challenging. In our dataset, only the parameters of cervical collars and CO_2 indicators were clearly visible in videos because collars had different sizes and different CO_2 indicators were in packages of different colors. The contents of IV fluid bags and size of IV catheters and ET tubes were not discernable from the videos unless verbally reported by the providers. We were still able to identify these parameters by comparing the RFID data and annotations.

Second, videotaping and RFID data recording had to be started and stopped independently, but this did not always happen simultaneously. We observed an offset of up to 280 sec between video-based annotations and RFID data for each resuscitation. Also, in several cases the video recording and the RFID data recording for the same resuscitation were not of the same length. We cropped the longer one from the end to make them equal.

8 Conclusion

Detecting object use is necessary for recognizing complex medical activities and establishing situation awareness in dynamic medical settings. We developed a passive RFID-based system for non-intrusive detection of used medical tools during trauma resuscitation. Our sytem consisted of optimally placed RFID tags and antennas, as well as a method for processing the RSSI data to infer object usage based on machine learning.

We deployed our system in an actual trauma bay and recorded radio signals during 32 simulated resuscitaions performed by trauma teams. Using long-range passive RFID, we were able to identify the parameters of used objects with high accuracy ($>$ 90 percent). Our goal was similar to medication or blood tracking, where barcodes or other near field technologies have been previously used. We showed that specific instances of objects could be tracked using long-range passive RFID, without the need for human cooperation in the sensing process.

Performance of detecting the time of object use depended on several factors, such as storage location, duration of use and the time between relocation and use. Passive RFID-based tracking yielded better results for objects that are stored sufficiently far away from the usage zone, used right after being relocated, and stayed in use for long times. We found that although low data rates affected the use detection performance, high data rates did not guarantee better use detection. We proposed complementary sensing methods to track objects for which RFID did not perform well. Finally, we illustrated by example how it is possible to recognize activity even when the use detection rates of individual objects are not high.

Acknowledgments

This research was supported by the US National Science Foundation (NSF) grant IIS-08-03732. The authors thank the medical personnel at Children's National Medical Center for participating in this study, as well as the hospital's Biomedical Engineering Department for their assistance in installing the RFID equipment.

References

[1] S. Agarwal, A. Joshi, T. Finin, Y. Yesha, and T. Ganous, "A pervasive computing system for the operating room of the future," *Mobile Netw. Appl.*, vol. 12, no. 2/3, pp. 215–228, 2007.
[2] Alien Technology 〉 Products 〉 RFID Readers. (2012). [Online]. Available: http://www.alientechnology.com/readers/alr9900.php
[3] E. Alpaydin, *Introduction to Machine Learning*. Cambridge, MA, USA: MIT Press, 2005.
[4] L. Bao and S. Intille, "Activity recognition from user-annotated acceleration data," *Pervasive*, vol. 3001, pp. 1–17, 2004.
[5] J. Bardram, A. Doryab, R. Jensen, P. Lange, K. Nielsen, and S. Petersen, "Phase recognition during surgical procedures using embedded and body-worn sensors," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, 2011, pp. 45–53.
[6] M. Buettner, R. Prasad, M. Philipose, and D. Wetherall, "Recognizing daily activities with RFID-based sensors," in *Proc. 11th Int. Conf. Ubiquitous Comput.*, 2009, pp. 51–60.
[7] M. Fitzgerald, P. Cameron, C. Mackenzie, and N. Farrow, "Trauma resuscitation errors and computer-assisted decision support," *Archives Surgery*, vol. 146, no. 2, pp. 218–225, 2011.
[8] A. Fleury, M. Vacher, and N. Noury, "SVM-based multimodal classification of activities of daily living in health smart homes: Sensors: Algorithms and first experimental results," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 274–283, Mar. 2010.
[9] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software," *SIGKDD Explorations*, vol. 11, no. 1, pp. 10–18, 2009.
[10] T. Inomata, F. Naya, N. Kuwahara, F. Hattori, and K. Kogure, "Activity recognition from interactions with objects using dynamic Bayesian network," in *Proc. 3rd ACM Int. Workshop Context-Awareness Self-Managing Syst.*, 2009, pp. 39–42.
[11] B. Jiang, K. 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, Feb. 2006.
[12] T. Kannampallil, Z. Li, M. Zhang, T. Cohen, D. Robinson, A. Franklin, J. Zhang, and V. Patel, "Making sense: Sensor-based investigation of clinician activities in complex critical care environments," *J. Biomed. Informat.*, vol. 44, no. 3, pp. 441–454, 2011.
[13] K. Kleisouris, B. Firner, R. Howard, Y. Zhang, and R. Martin, "Detecting intra-room mobility with signal strength descriptors," in *Proc. 11th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2010, pp. 71–80.
[14] M. Kranzfelder, A. Schneider, S. Gillen, and H. Feussner, "New technologies for information retrieval to achieve situation awareness and higher patient safety in the surgical operating room," *Surgical Endoscopy*, vol. 25, pp. 696–705, 2011.
[15] Y. Lee and S. Cho, "Activity recognition using hierarchical hidden markov models on a smartphone with 3D accelerometer," *LNCS Hybrid Artif. Intell. Syst.*, vol. 6678, pp. 460–467, 2011.
[16] J. Lester, T. Choudhury, and G. Borriello, "A practical approach to recognizing physical activities," *Pervasive Comput.*, vol. 3968, pp. 1–16, 2006.
[17] L. Liao, D. Fox, and H. Kautz, "Hierarchical conditional random fields for GPS-based activity recognition," in *Proc. 12th Int. Symp. Robot. Res.*, 2007, vol. 28, pp. 487–506.
[18] B. Logan, J. Healey, M. Philipose, E. Tapia, and S. Intille, "A long-term evaluation of sensing modalities for activity recognition," in *Proc. 9th Int. Conf. Ubiquitous Comput.*, 2007, pp. 483–500.
[19] K. Ohashi, S. Ota, L. Ohno-Machado, and H. Tanaka, "Smart medical environment at the point of care: Auto-tracking clinical interventions at the bedside using RFID technology," *Comput. Biol. Med.*, vol. 40, no. 6, pp. 545–554, 2010.
[20] N. Oliver and E. Horvitz, "Layered representations for human activity recognition," in *Proc. IEEE 4th Int. Conf. Multimodal Interfaces*, 2002, pp. 3–8.
[21] S. Parlak, S. Ayyer, Y. Liu and I. Marsic, "Design and evaluation of RFID deployments in a trauma resuscitation bay," *IEEE J. Biomed. Health Informatics*, vol. 18, no. 3, pp. 1091–1097, May 2014.
[22] S. Parlak and I. Marsic, "Detecting object motion using passive RFID: A trauma resuscitation case study," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 9, pp. 2430–2437, Sep. 2013.
[23] S. Parlak, I. Marsic, and R.S. Burd, "Activity recognition for emergency care using RFID," in *Proc. ACM 6th Int. Conf. Body Area Netw. Conf.*, Beijing, China, Nov. 2011, pp. 40–46.
[24] S. Parlak, A. Sarcevic, I. Marsic, and R.S. Burd, "Introducing RFID technology in dynamic and time-critical medical settings: Requirements and challenges," *J. Biomed. Informat.*, vol. 45, no. 5, pp. 958–974, Oct. 2012.
[25] D.J. Patterson, D. Fox, H. Kautz, M. Philipose, "Fine-grained activity recognition by aggregating abstract object usage," in *Proc. 9th IEEE Int. Symp. Wearable Comput.*, Oct. 2005, pp. 44–51.
[26] D. Powers, "Evaluation: From precision, recall and f-measure to ROC, informedness, markedness and correlation," *J. Mach. Learn. Technol.*, vol. 2, no. 1, pp. 37–63, 2011.
[27] L. Ravindranath, V. Padmanabhan, and P. Agrawal, "SixthSense: RFID-based enterprise intelligence," in *Proc. ACM MobiSys Conf.*, 2008, pp. 253–266.
[28] A. Sarac, N. Absi, S. Dausere-Peres, "A literature review on the impact of RFID technologies on supply chain management," *Int. J. Prod. Econ.*, vol. 128, no. 1, pp. 77–95, 2010.
[29] T. Stiefmeier, D. Roggen, G. Troster, G. Ogris, and P. Lukowicz, "Wearable activity tracking in car manufacturing," *IEEE Pervasive Comput.*, vol. 7, no. 2, pp. 42–50, Apr.–Jun. 2008.
[30] T. Van Kasteren, A. Noulas, G. Englebienne, and B. Krose, "Accurate activity recognition in a home setting," in *Proc. 10th Int. Conf. Ubiquitous Comput.*, 2008, pp. 1–9.
[31] J. Ward, P. Lukowicz, and H. Gellersen, "Performance metrics for activity recognition," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 1, pp. 6:1–6:23, 2011.
[32] R. Weinstein, "RFID: a technical overview and its application to the enterprise," *IT Prof.*, vol. 7, no. 3, pp. 27–33, 2005.

Siddika Parlak received the BS and MS degrees in electrical engineering from Bogazici University, Istanbul, Turkey, in 2006 and 2008, respectively. She received the PhD degree in electrical and computer engineering from Rutgers University, New Brunswick, NJ, in 2013. She is currently a Senior Engineer at Qualcomm Technologies Incorporated. Her research interests include machine learning, sensors and context-aware systems.

Ivan Marsic received the DiplIng and MS degrees in computer engineering from the University of Zagreb, Croatia, and the PhD degree in biomedical engineering from Rutgers University, NJ, in 1994. He is a professor of electrical and computer engineering at Rutgers University. His research includes software engineering and sensor networks for healthcare applications.

Aleksandra Sarcevic received the MS and PhD degrees from the School of Communication and Information, Rutgers University, in 2005 and 2009, respectively. She is an assistant professor of informatics at Drexel University. Her research interests include human-computer interaction and computer-supported cooperative work.

Waheed U. Bajwa received the BE degree from the National University of Sciences and Technology, Pakistan in 2001, and the MS and PhD degrees from the University of Wisconsin-Madison in 2005 and 2009, respectively. He is an assistant professor of electrical and computer engineering at Rutgers University. His research includes high-dimensional inference and inverse problems, compressed sensing, statistical signal processing, and wireless communications.

Lauren J. Waterhouse received the BS degree in public health from George Washington University in 2007. She is a research coordinator in the Division of Trauma and Burn Surgery, Children's National Medical Center. Her research includes team coordination and protocol adherance in pediatric trauma resuscitation.

Randall S. Burd received the BA degree from Dartmouth University in 1983, the MD degree from Columbia University in 1987, and the PhD degree from the University of Minnesota in 1994, where he also completed his residency in surgery. He is a pediatric surgeon and the chief in the Division of Trauma and Burn Surgery, Children's National Medical Center. His research focuses on developing novel strategies for improving the early care of injured children.