

FEMO: A Platform for Free-weight Exercise Monitoring with RFIDs

Han Ding*, Longfei Shangguan†, Zheng Yang‡, Jinsong Han*,
Zimu Zhou‡, Panlong Yang§, Wei Xi*, Jizhong Zhao*,
*Xi'an Jiaotong University

†Hong Kong University of Science and Technology

‡Tsinghua University

§PLA University of Science and Technology

{dinghanxjtu,shanggdtk,hmilyyz,zhouzimu.hk,panlongyang,weixi.cs}@gmail.com,
{hanjinsong,zjz}@mail.xjtu.edu.cn

ABSTRACT

Regular free-weight exercise helps to strengthen the body's natural movements and stabilize muscles that are important to strength, balance, and posture of human beings. Prior works have exploited wearable sensors or RF signal changes (*e.g.*, WiFi and Blue-tooth) for activity sensing, recognition and counting *etc.*. However, none of them have incorporate three key factors necessary for a practical free-weight exercise monitoring system: recognizing free-weight activities on site, assessing their qualities, and providing useful feedbacks to the bodybuilder promptly. Our FEMO system responds to these demands, providing an integrated free-weight exercise monitoring service that incorporates all the essential functionalities mentioned above. FEMO achieves this by attaching passive RFID tags on the dumbbells and leveraging the Doppler shift profile of the reflected backscatter signals for on-site free-weight activity recognition and assessment. The rationale behind FEMO is 1): since each free-weight activity owns unique arm motions, the corresponding Doppler shift profile should be distinguishable to each other and serves as a reliable signature for each activity. 2): the Doppler profile of each activity has a strong spatial-temporal correlation that implicitly reflects the quality of each performed activity. We implement FEMO with COTS RFID devices and conduct a two-week experiment. The preliminary result from 15 volunteers demonstrates that FEMO can be applied to a variety of free-weight activities and users, and provide valuable feedbacks for activity alignment.

Keywords

activity recognition and assessment; RFID

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

SenSys '15, November 1–4, 2015, Seoul, South Korea..

© 2015 ACM. ISBN 978-1-4503-3631-4/15/11 ...\$15.00.

DOI: <http://dx.doi.org/10.1145/2809695.2809708>.

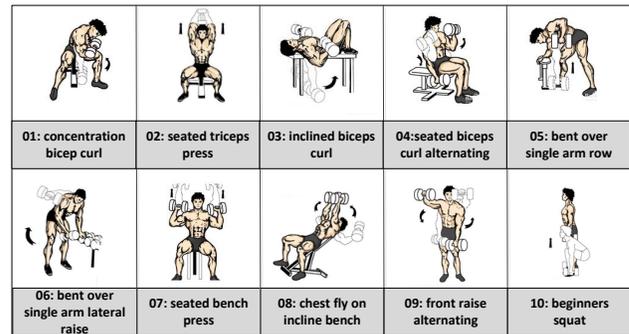


Figure 1: Sketches of the tested free-weight activities

1. INTRODUCTION

Free-weight exercise is indispensable in a balanced exercise program and provides numerous health benefits. It helps to stabilize bones and muscles that are relevant to strength, balance and posture, and contributes to weight loss and overall health [20]. Some researches recommend free-weight training twice a week for adults [2], even for those who walk or run regularly. Moreover, some individuals may prefer free-weight exercises to aerobic exercises as the main fitness activities simply for lifestyle or convenience reasons.

Monitoring and evaluation supports are crucial for free-weight exercises. Compared with aerobic exercises *e.g.* running, people are more vulnerable to inefficient training or even accidental injuries in free-weight training. Stretching or warping muscles improperly, *e.g.*, to the wrong direction or at a high speed, can lead to strains and tears. Timely guidance is also important for exercise safety and quality since people might forget their progress, skip essential steps, or miscout a sequence.

Despite academic and commercial success in aerobic exercise monitoring [10, 11, 12, 16, 43], there is a void in automatic free-weight exercise monitoring and evaluation. A personal trainer is still by far the most common solution to free-weight training monitoring, which incurs high recruitment costs [1]. Unlike aerobic exercises where speed, distance and terrain are essential, the quality of free-weight training is mostly defined by repetitions, durations and activity sets. Existing schemes on general activity recognition [8, 25] and aerobic exercise tracking [12, 16] fail to cap-

ture such high-fidelity information for free-weight exercises. Some pioneer work explored inertial sensors *e.g.* accelerometer and gyroscope, for free-weight training monitoring [9, 28], yet these techniques cannot reliably handle the variety of arm motion patterns and diverse training paces. Furthermore, they also require body-worn sensors to function, which poses inconvenience and might cause unwanted motion changes during training. A promising alternative is to leverage wireless signals for device-free activity sensing. Recent research has explored the feasibility of WiFi signals for gesture and activity recognition [27, 30, 41], yet either involves customized hardware (*e.g.*, software radios [30] and directional antennas [27]) or targets at location-aware activities [41], thus unsuitable for ubiquitous free-weight exercise tracking.

In this paper, we design FEMO, an automatic, non-invasive and light-weight Free-weight Exercise Monitoring system. We enable non-invasive free-weight exercise monitoring by attaching passive Radio Frequency Identification (RFID) tags on assisted instruments (*e.g.* dumbbells) during training, and analyzing signals backscattered from tags during exercises. Attaching passive RFID tags on instruments like a dumbbell poses minimum overhead due to their negligible weight and size. FEMO works by analyzing the Doppler shifts extracted from the backscattered signals. It automatically recognizes, counts, and assesses the exercises on-site and in real time. The detailed assessment feedbacks are also displayed on FEMO’s UI module to assist activity rectification.

The design of FEMO involves the following challenges:

(1) *How to detect and extract accurate Doppler shifts from backscatter signals?* Doppler shifts are accessible from commercial RFID readers via a standard API. However, the raw Doppler measurements are too noisy to precisely portray the tag movements (*i.e.* dumbbell trails).

We address this problem by transforming the received phases from backscatter signals into the corresponding Doppler shifts (§3.1). FEMO tracks this Doppler stream and segments the Doppler shifts of each activity performed even at diverse paces (§3.2).

(2) *How to recognize free-weight exercises on-site?* Traditional activity recognition schemes rely on sophisticated feature selection and complex classification techniques for accurate recognition, which incur a large computational latency and are sensitive to training data.

FEMO addresses this problem by analyzing the temporal patterns of RF signals affected by body movements. Our observation is that each free-weight activity is a unique combination of basic arm motions. These combined arm motions exhibit unique yet stable Doppler shift profiles in the temporal domain, producing light-weight and robust features for activity recognition. We detail this idea and optimize the recognition process in §3.3.

(3) *How to assess user performances?* Since FEMO aims to provide useful feedbacks to users as guidance for improper activity rectification, it is important to quantify the quality of each performed activity.

In FEMO, we define a set of metrics evaluating the quality of activities during free-weight exercises from both the local and global views. An assessment framework is proposed to evaluate both the activity details and the activity consistency within each training group.

We implement FEMO as a framework consisting of four core modules: Doppler value pre-processing, activity segmentation, activity recognition and activity assessment. We prototype FEMO on commodity hardware including a Commercial Off-The-Shelf (COTS) RFID reader, a directional antenna, and two passive RFID tags attached to the dumbbells. It runs on a central server and processes the tag reading stream in pipeline. It also provides an interface to other applications such as activity counting and training process tracker, where they can obtain the current activity primitives from FEMO via programming interfaces. We evaluate the performance of FEMO on free-weight training data collected in two weeks from 15 volunteers. The data cover 1,534 minutes of exercises, with 4,500 repetitions of 10 representative exercises. Results demonstrate that FEMO can be applied to various representative free-weight activities, and provide valuable feedbacks to users, especially beginners.

We summarize the contributions of this paper as follows:

- We introduce the first passive RFID-based system for free-weight exercise monitoring. This system enables on-site activity recognition and assessment, and provides rich feedbacks to the user for activity rectification. Different from prior works that only coarsely recognize to which type the activity is belonging, our FEMO system can provide fine-grained measurements on the activity by using the subtle Doppler values retrieved from COTS passive RFID devices.
- We present a set of algorithms to effectively extract free-weight exercise information from backscatter signals. Specifically, our algorithms enable 1) extracting minute Doppler shifting from noisy tag readings reported by commercial RFID readers; 2) segmenting the Doppler streaming such that each segment contains an intact activity; 3) assessing the exercise performance and providing useful feedbacks to the bodybuilder.
- Last but not the least, we establish a proof-of-concept prototype and conduct two-weeks experiments. Results demonstrate the effectiveness of our system. It can detect ten typical free-weight activities with an average accuracy of 90.4%, and facilitate rectifying irregular exercise behaviors, especially for fitness novices.

2. OVERVIEW

This section briefly introduces the taxonomy of the targeted free-weight activities and the overall work-flow of FEMO.

2.1 Taxonomy of Free-weight Activities

In this paper, we focus on ten common and representative free-weight activities, which can train different parts of the muscle groups. We choose these popular activities based on a two-week questionnaire investigation on fitness enthusiasts in our university. Similar to [9], we categorize these activities into different groups based on the muscle groups desired to be trained. Figure 1 illustrates ten free-weight activities of interests, which will be used to demonstrate our work throughout this paper.

2.2 FEMO Work-flow

Figure 2 presents the work-flow of FEMO. It contains four major steps: preprocessing, activity segmenting, activity recognition, and activity assessment.

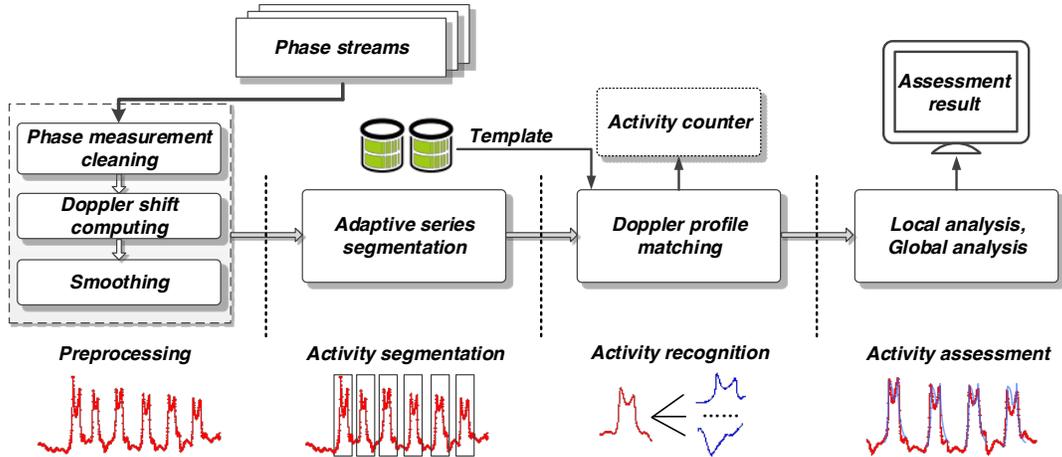


Figure 2: An overview of FEMO’s work-flow

The first step is preprocessing, where the system purifies the raw phase readings and computes the Doppler shifts. In this step, FEMO first mitigates the noisy phase readings introduced by the hardware heterogeneity of the reader and the inconsistency of tag orientations. FEMO then computes the Doppler values and finally employs a moving-average filter to smooth the Doppler values.

The second step is to segment the Doppler stream, such that each segment contains a single free-weight activity. A main challenge lies in the heterogeneity of Doppler profiles with respect to different activities. We find that the state-of-the-art, *e.g.*, threshold-based filters [9, 10] or peak detection schemes [14], fails to segment the activity precisely. In FEMO, we exploit the stochastic characteristics of Doppler values and design a KL-divergence based segmentation algorithm. This algorithm works efficiently and adapts to various free-weight activities and users.

The third step is to recognize each activity from the Doppler segment. Towards accurate and prompt recognition, we build up a body movement model and observe that each activity has a unique yet stable combination of arm motion trails. Based on this observation, we employ the Doppler profile as the feature of each activity and design a fingerprint based activity recognition scheme. To improve the recognition efficiency, we leverage the motion order of arms to prune the unqualified matching candidates in advance.

The final step is to measure the quality of each activity and provide valuable feedbacks to users. In FEMO, we assess each activity from both the local and global views. Local analysis concentrates on the activity details by comparing the proposed features against the standard ones. The global analysis focuses on the consistency of each group of activities by measuring the smoothness and continuity of activities within each group. The assessment results are displayed on the end-user interface to assist activity rectification.

3. SYSTEM DESIGN

This section details the FEMO design and highlights the challenges, key observations, and core techniques behind the activity segmentation, recognition and assessment of FEMO.

3.1 Data Acquisition and Preprocessing

Standard Doppler API: Commercial RFID readers offer a standard API to retrieve Doppler shifts. However, ac-

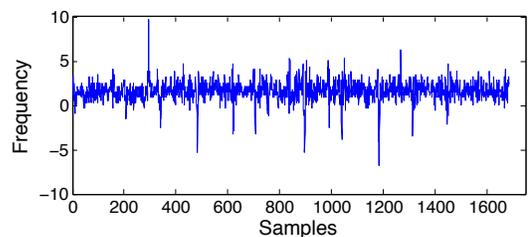


Figure 3: Doppler phase shifting acquired from the API

cording to our initial experiments, we find that the reported Doppler values fail to precisely depict the free-weight activity. As an example, we perform the *Bent over single arm lateral raise* ten times in front of an Impinj R420 reader. Figure 3 plots the Doppler values reported by the standard API. Ideally, the Doppler values should be close to zero when the user keeps still, yet increase positively (negatively) as the dumbbell moves towards (away from) the reader. Since the activity is repeated ten times, the Doppler shifts should show a consistent pattern. However, from Figure 3, we find that the reported Doppler values change irregularly over the whole process. The results indicate that neither the tag’s approaching nor the opposite movement to the reader results in a clear positive or negative peak. Even worse, it is difficult to recognize the repetitive pattern from the collected Doppler readings. The technical report released by Impinj states that the reader estimates the Doppler shift of a tag through the time duration and phase difference of a single tag packet. Thus, the inaccurate Doppler may be related to both the time disalignment and phase noise. Hence we seek other methods to acquire precise Doppler values.

RF phase measurement: Instead of directly using the Doppler shifts from the API, we deduce the Doppler shifts from the phase measurement reported by the commercial reader. The phase value of an RF signal θ describes the offset of the received signal from the original one before transmission, ranging from 0 to 2π .

$$\begin{cases} \theta = (2\pi \frac{2l}{\lambda} + \mu) \bmod 2\pi \\ \mu = \theta_{Tx} + \theta_{Rx} + \theta_{TAG} \end{cases} \quad (1)$$

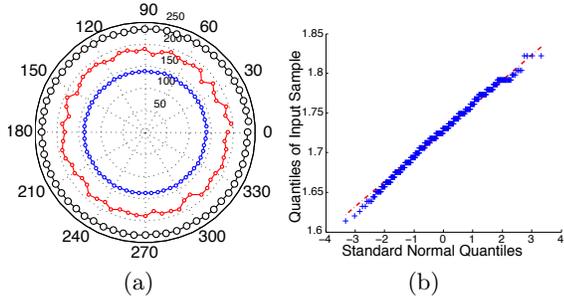


Figure 4: (a) Phase measurements under different tag orientations (b) QQ Plot of sample data versus standard normal distribution

where l is the distance between the antenna and the tag, λ is the wavelength, and μ is the system noise. Due to the imperfect manufacture, the reader's transmit circuits (θ_{Tx}), the receiver circuits (θ_{Rx}), and the tag's reflection characteristics (θ_{TAG}) will introduce additional phase rotations, *i.e.*, noises, to the phase measurement.

3.1.1 Preprocessing

Phase measurement smoothing: To deduce precise Doppler shifts, it is crucial to minimize the phase noise. We change the tag orientation with respect to the antenna with a step of 6° and examine how sensitive the phase measurement is to tag orientations. The antenna is set to be 2m away from the tag. Figure 4(a) shows the raw phase measurements when the tag rotates 2π (marked as red points). We find that due to the inherent circuit noise, the phase values fluctuate continuously and randomly. We further test the distribution of these measurements against the standard Gaussian distribution (shown in Figure 4(b)). The linearity of the points on the Q-Q plot suggests that the data are normally distributed, with a standard deviation of 0.0332 radian. Based on this observation, we model the phase measurement θ as a Gaussian random variable $\mathcal{N}(\mu, 0.0332)$. We then utilize the standard Kalman Filter [17] to smooth the phase values. We test various lookback window sizes, and empirically set it as 10 which optimizes the smoothing performance.

Figure 5 shows the phase measurement before/after smoothing in both the static and movement cases. The difference indicates that the Kalman filter effectively enhances the stability of phase measurements of static tags. For the moving-tag case, we can see that the phase values change steadily after filtering. Yet they still retain a clear profile of each free-weight activity.

Deducing Doppler shifts: Doppler shifts are generated due to the relative movement between a sender and a receiver, *e.g.*, the stationary reader and the moving tag. Suppose at time t_i and t_{i+1} , the reader receives two consecutive signals from the moving tag, with the phase readings θ_i and θ_{i+1} . Let v be the tag's moving speed within the period of $[t_i, t_{i+1}]$. v can be regarded as a constant due to the short interval between two consecutive tag readings. Hence the distance (d) that the tag moves equals to $v \cdot (t_{i+1} - t_i)$. On the other hand, we know that the signal in the backscatter communication traverses 2 times of d . Thus we have:

$$2v \cdot (t_{i+1} - t_i) = \lambda \cdot \left(\frac{\theta_{i+1} - \theta_i}{2\pi} \right) \quad (2)$$

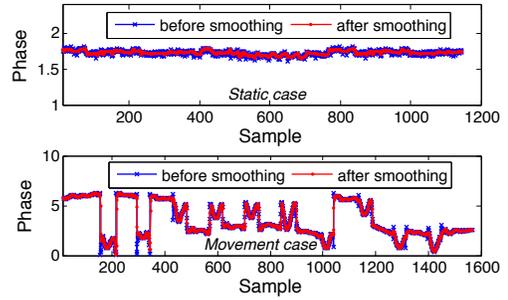


Figure 5: Phase measurements before/after smoothing under static/movement cases

The Doppler shift can be further expressed as:

$$f = \frac{v}{\lambda} = \frac{\theta_{i+1} - \theta_i}{4\pi \cdot (t_{i+1} - t_i)} \quad (3)$$

Figure 6(b) shows the phase-deduced Doppler shifts. The measurements are collected when a volunteer performs *Bent-over lateral raise* ten times. Compared with the noisy Doppler values reported by the API (Figure 6(a)), the deduced Doppler values clearly show ten repetitive patterns.

Despite the high resolution, from Figure 6(b) we find that the deduced doppler value still fluctuates over time. This is because the time interval between any two consecutive readings may vary due to the random access mechanism of ALOHA protocol [32]. Such non-uniform time intervals lead to Doppler fluctuations and jitters that overwhelm the original appearance of each activity. Therefore, after computing the Doppler value, we apply a moving average filter over the last n readings ($n=10$ in FEMO) to smooth the Doppler values. Figure 6(c) shows the final results.

3.2 Activity Segmentation

The activity segmentation module identifies Doppler segments that are likely to contain a complete free-weight activity. We define each segment as $\kappa_i = (t_s : t_e)$, with start time t_s and end time t_e within the Doppler stream. The segmentation yields a set of segments K , with each containing a free-weight activity: $K = \{\kappa_1, \kappa_2, \dots, \kappa_m\}$.

There has been extensive efforts on activity segmentation [8]. Most of them assume that each activity will exhibit a clear peak in the received signal stream. Thus by comparing each signal strength with a static threshold, the activity segments will be located accordingly. However, such a solution is unsuitable for FEMO as a free-weight activity usually contains multiple peaks within an activity period. The one-peak detection scheme may split one activity into multiple segments.

3.2.1 Key observations

In FEMO, the activity segmentation scheme is based on the fact that people tend to take a short rest after each activity to control the training pace. We term such a short rest as a *resting interval*. The resting intervals have small Doppler values, which naturally separate the activities. Thus by extracting the start and end times of each resting interval, we can acquire each activity segment accordingly. Nevertheless, a simple threshold still fails since people may adjust poses slightly within the resting intervals, leading to sharp Doppler peaks in the resting intervals. Such Doppler peaks can exceed the threshold and incur incorrect segments.

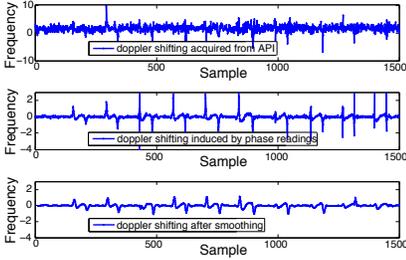


Figure 6: Doppler shift of 10 consecutive *Bent-over lateral raise*

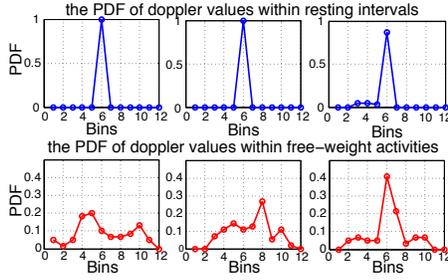


Figure 7: PDF of doppler values within different windows

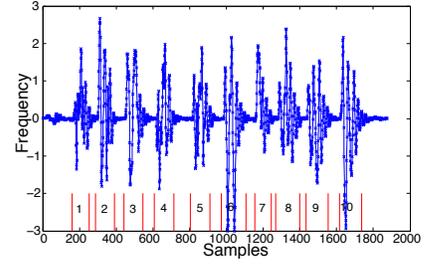


Figure 8: Segmentation result of 10 activity performing

Our segmentation scheme leverages two insights:

- The sharp Doppler values usually take a small portion of the whole data within the resting interval.
- Except for the sharp Doppler values incurred by the pose adjustment, the remaining Doppler readings within the resting interval are relatively small and stable.

Thus if we split a resting interval into multiple consecutive windows, the distribution of Doppler values within each window should be similar. Conversely, the Doppler values outside the resting interval corresponds to the free-weight activity. These values change rapidly and show a completely different distribution from those in the resting interval.

3.2.2 Segmentation scheme

The above analysis leads us to an adaptive segmentation scheme based on the KL divergence [34]. Denote the Doppler stream as $S = (s_i) \in \mathbb{R}^{1 \times N}$, where N is the number of discrete time points t_1, \dots, t_N at which the Doppler values are sampled. For each w consecutive Doppler values, we group them into a window. Within each window, we further categorize the Doppler values into multiple bins. The bin size is empirically set as 0.35. Then we can get the discrete probability distribution function (PDF) of Doppler values within each window. Given two consecutive windows w_i and w_j , let P and Q be their PDF, respectively. The KL divergence of Q from P is defined as:

$$D_{KL}(P||Q) = \sum_i P(i) \cdot \ln \frac{P(i)}{Q(i)} \quad (4)$$

The KL divergence measures the information loss when Q is used to approximate P . In FEMO, there are three cases:

1. both windows are within the resting interval;
2. both windows are within the activity period;
3. one is within the resting interval and another is within the activity period;

In the first case, $D_{KL}(P||Q)$ will be close to zero due to the similar probability distributions within these two windows (Figure 7(a)). In the latter two cases, $D_{KL}(P||Q)$ will be significantly larger than zero as the distribution varies sharply among these two windows (Figure 7(b)). Hence by checking $D_{KL}(P||Q)$, we can ascertain whether the current window is within the resting interval or not. After finding all windows within the resting interval, we can extract the

activity segment accordingly. Figure 8 shows the segmentation result over a Doppler stream, we see that all of these ten activities are correctly identified. We also notice the Doppler profile of each activity is completely contained in the corresponding segment, indicating accurate and robust segmentation.

3.3 Activity Recognition

The activity recognition module aims to identify the free-weight activity within each segment. Many previous works focus on building a robust activity recognition system [9, 10, 29, 30, 41]. The main principle is to extract unique features for the activity from the input signals, and train a classifier to distinguish each unlabelled activity. However, this method has two major drawbacks. First, the system performance is sensitive to the training set. If the training set is small or biased, the classifier will suffer from low recognition accuracy. Although a larger training set may lead to a more accurate classifier, it will incur a higher overhead (*e.g.*, time and cost for collecting the ground-truth) on system realization and deployment. Second, such a method suffers from higher latency. To get a better performance, researchers leverage advanced graphic model (*e.g.*, Hidden Markov Model (HMM) and Conditional Random Field (CRF)) to explore the transition between consecutive activities. These model based enhancements, however, usually incur much higher computational overhead as they require a large amount of training data to determine the parameters of their models before deployment. Moreover, they will involve the states of previous activities when recognizing the current activity, which is also very computationally costly. We are motivated to design our activity recognition scheme balancing accuracy and overhead.

3.3.1 Fingerprint Matching

We find that each free-weight activity can be considered as a combination of different arm motions. For various free-weight activities, each arm motion presents a unique motion trail, generating unique Doppler profiles. Figure 9 demonstrates the effectiveness of Doppler profile. The two upper subfigures verify the stability of Doppler profiles while the uppermost and the lowermost ones confirm that the Doppler profiles disperse in different free-weight activities. Hence, the Doppler profile can discriminate the free-weight activities and act as a reliable signature.

FEMO compares the profile in each segment against the standard to identify free-weight activities. Hence we need to evaluate the similarity between two Doppler profiles. We argue that the Euclidean-distance metric is unsuitable since activity segments may vary in length due to personal

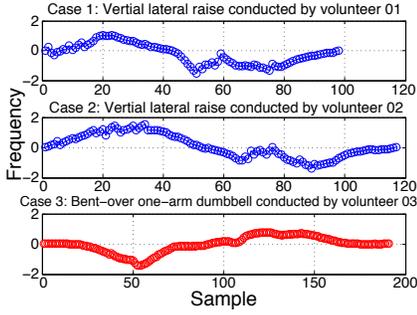


Figure 9: Doppler profiles of gym activities

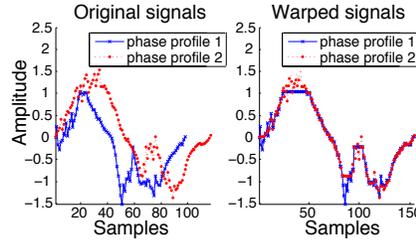


Figure 10: Original and warped signals

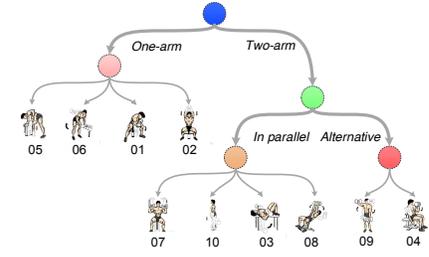


Figure 11: The pipeline of activity recognition

preferences, physical characters (weight, height) and other reasons.

FEMO uses *Dynamic Time Warping* (DTW) [35] to compute the similarity between two Doppler profiles. The benefits are twofold. On one hand, DTW compares two profiles with different lengths. On the other hand, DTW automatically compresses or stretches a sequence to minimize the distance between two sequences, thus focusing on the shape similarity rather than the absolute values. Figure 10 shows the Doppler profiles of bent-over one-arm dumbbell that are aligned with DTW. We observe that Doppler profile 1 is stretched and shifted to match with Doppler profile 2.

3.3.2 Hierarchical activity recognition framework

Directly using DTW for activity recognition is costly. The complexity of DTW is $O(mn)$, indicating a large overhead for long Doppler profiles, especially when we have to compare one Doppler profile against all candidates in the database. To reduce the computational overhead, we design a hierarchical activity recognition framework based on two observations on the arm motion pattern:

1. The free-weight activity involves either a single arm motion or two arm motions.
2. For the two arm motion activities, they are either conducted in parallel or alternatively.

Figure 11 shows the decision-tree based recognition scheme. At the first level, FEMO classifies the candidate by detecting whether the current activity is a single arm activity or not. This process can be achieved by checking whether the attached two tags on the dumbbells are detected together. If yes, FEMO further classifies the candidate by checking whether the current two-arm activity is performed alternatively or not. This process can be done by examining the concurrency of activities within these two doppler streams. Specifically, let t_o and t_l be the overlapping duration and the longest time duration of two activity segments, respectively. If $t_o \geq \beta \cdot t_l$, FEMO ascertains these two motions are conducted in parallel. With above pruning process, we can shrink the candidate group to no more than half its original size.

In our implementation, we set β as $\frac{2}{3}$ by default. Identifying the tag ID takes $O(1)$ time, while judging the concurrency of two activities is achievable in $O(1)$ time. These two simple operations help to prune the unqualified profiles in an early stage of activity recognition, thereby improving the computational efficiency. After the activity recognition, all activities will be labelled and stored for activity assessment.

3.4 Activity Assessment

The activity assessment aims to characterize the *quality* of the exercise and provides feedback to users. We first measure the quality for each individual activity by comparing its characteristics against the standard ones. Note that free-weight activities are often grouped where each group contains a set of repetitive activities. Activity consistency within an activity group is also essential to the gym training [18]. Hence we assess the quality of activities from two perspectives, *i.e.*, the *local view* and *global view*, to reflect both the offset of each individual activity from the standard and the inconsistency of a activity group.

3.4.1 Local Analysis

Local analysis evaluates the quality of each activity by concentrating on its *duration* and *intensity*, which are two general criteria for evaluating the free-weight activities.

Duration: It measures how long an activity is performed by the user. The duration of an activity is critical to the free-weight training [3]. A longer duration indicates a slower arm motion, which potentially corresponds to an ineffective muscle workout. If the duration is too short, the muscle will be stretched or warped fiercely, leading to an excessive muscle workout. Either case degrades the effectiveness of gym training. In order to improve the training efficiency, FEMO measures the difference of the durations between each conducted activity and the standard. Let d_i be the duration of activity i , d_s be the duration of the corresponding standard activity. FEMO computes their difference $d_i - d_s$ and reports the result to the user at the end of each activity group.

Intensity: Intensity is another important metric to evaluate the quality of free-weight activities. It reflects the energy of the arm motions expended by an activity [39]. It is understandable that a high quality activity should show a similar energy trend to the standard. Adapting this idea to the doppler domain, it then becomes much clear that we measure the similarity between their doppler segments. Specifically, Let $A = \{a_1, a_2, \dots, a_m\} \in \mathbb{R}^{1 \times m}$ be the doppler segment of the desired activity, $B = \{b_1, b_2, \dots, b_n\} \in \mathbb{R}^{1 \times n}$ be the corresponding standard activity. In FEMO, we first align these two doppler segments by DTW to address the potential inconsistency of segment length. After that, we compute their similarity using Euclidean distance metric. A shorter distance between two doppler segments indicates a higher similarity, and hence a higher performance of this activity.

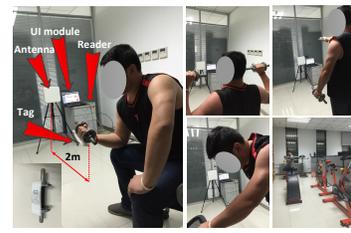


Figure 12: Training tracker service Figure 13: Performance assessment Figure 14: Experiment scenario

3.4.2 Global Analysis

Global analysis aims at monitoring how well each group of activities are performed, with an emphasis on unveiling the abnormal activity pattern and irregular resting intervals within each activity group. This part concentrates on two kinds of characteristics: *smoothness* and *continuity*.

Smoothness: Smoothness reflects how similar each activity is to the remaining activities within an activity group. Dumbbells also require more balance and more muscular control than others, such as the training with barbells or machines, and balance is crucial for optimal performance [6]. Smoothness can well reflect the balance by measuring the similarity of exercises in a group. A larger similarity indicates more regular arm actions, corresponding to effective muscle trainings. To evaluate the smoothness of an activity, we employ the discrete PDFs of all the Doppler values within an activity segment as the proposed features. Specifically, let $PDF(A) = \{p_i\}_{i=1}^m$ be the PDF of the activity segment A , where p_i represents the i^{th} bin value. Since the Doppler value induced is at the granularity of 0.3 Hz, the number of bins to calculate the discrete probability distribution function is also set to the granularity of 0.3 Hz. In our measurements, we find that the fluctuation range of most Doppler values is around 6 Hz, we therefore set the number of bins to 20. To compare the similarity between two activities, we employ the Earth Mover’s Distance (EMD) [33]. EMD measures the dissimilarity between two discrete probability distribution function $PDF(A) = \{p_i\}_{i=1}^m$ and $PDF(B) = \{q_j\}_{j=1}^n$. It is the minimal effort required to transform one histogram into another.

Continuity: Continuity depicts the consistency of resting intervals within an activity group. For efficient training, the users should pace themselves throughout the exercises [7]. A higher consistency of resting intervals indicates that the user has a good motion pace control, *i.e.*, a regular muscle stretching/warping pace. Ideally, within an activity group, the resting intervals should be consistent with each other. However, even a professional gymnasium trainer may fail to keep strictly consistent resting intervals. Thus we model the resting interval of the standard activity group as a standard normal random variable. To evaluate the continuity of activities performed within an activity group, we investigate the statistical characteristic of resting intervals and employ kurtosis as a metric. The coefficient of kurtosis is a measurement on the degree of peakedness in a variable distribution. Specifically, let $R = \{r_i\}_{i=1}^m$ be the vector of resting intervals within an activity group. The kurtosis can be computed as follows:

$$\beta_2 = \frac{\sum_{i=1}^m (r_i - \mu)^4}{(\sum_{i=1}^m (r_i - \mu)^2)^2} - 3 = \frac{\mu^4}{\sigma^4} - 3 \quad (5)$$

where μ and δ are the mean value and standard deviation of the resting interval vector. As larger β_2 value indicates a concentrated distribution of resting intervals, therefore a better continuity of activities that the user performs.

4. SYSTEM IMPLEMENTATION

This section presents both the hardware composition and software realization of FEMO.

Hardware: We implement a prototype of FEMO on COTS UHF RFID devices, including an ImpinJ reader Model R420, a Laird antenna model A9028R0NF (with a gain of 8dbi), and a set of passive RFID tags. As the metal dumbbell will block the magnetic waves, we place the tag on a plastic form, which is further attached to the dumbbell. The reader is connected to a backend PC via an Ethernet cable and continuously reports the signal features backscattered from tags, including RSSIs, phase angles, and Doppler shifts. We time-stamp each tag reading by using the reader’s local clock in order to eliminate the influence of network latency.

Software: The software of FEMO is fully implemented in C#. It comprises of three components: data collection module, data analysis module and UI module. The data collection module is integrated with the Octane SDK, an extension of the LLRP Toolkit, which supports continuous tag interrogation at a rate of 340 readings/s. The data analysis module is responsible for recognizing and assessing the quality of each performed activity. The assessment results are displayed on a web-based UI module. The software runs on a Lenovo PC with an Intel Core i7-4600U 2.10GHz CPU and 8GB RAM.

UI module: FEMO currently provides two services to the bodybuilder: training tracker service and activity performance assessment service. Figure 12 shows the UI page of the training tracker service. This service aims to provide daily training statistics to the bodybuilder, including the accumulative training summary and daily activity summary. The former one records the training frequency, workout and duration of the bodybuilder. The latter one reports the amount of activities and average duration of each activity. In this way, FEMO provides the raw training data to the bodybuilder. Figure 13 illustrates the UI page of the activity performance assessment service. This service offers the quantitative performance assessment on each activity in the daily training. The user can visualize his doppler profiles of the current activity group against the standard template for evaluating his activity performance. The service also offers quantitative reports on the quality of the selected activity group against the standard template, such as the smoothness of each activity and duration/intensity differences from standard template, within each activity group. Based on the

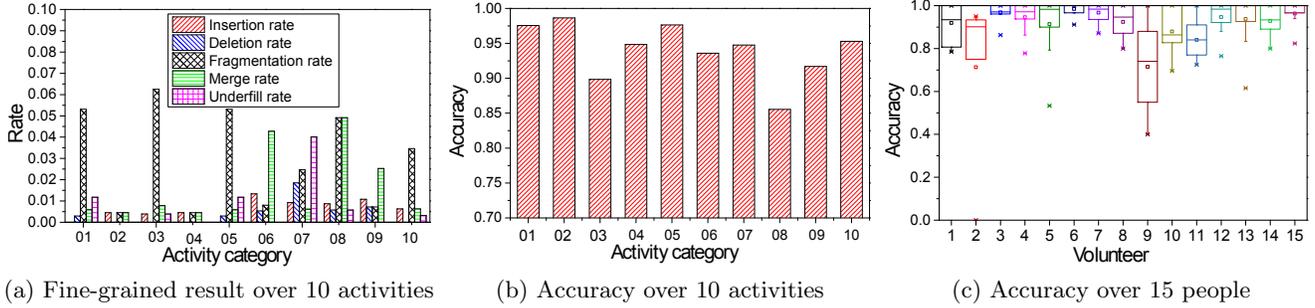


Figure 15: Evaluation result of activity segmentation scheme

comparison, the user obtains a thorough summary on the continuity of activity groups. Although our current FEMO prototype only has fundamental functions, *e.g.*, the visualized quality report for daily activities and logging data along the overall training process, we plan to integrate more advanced functionalities in our future work, such as the customized training reminder and intelligent training advisor.

5. SYSTEM EVALUATION

In this section, we conduct extensive experiments and evaluate the performance of FEMO in terms of accuracy, effectiveness, and overhead.

5.1 Experiment setups

The experiment scenario is shown in Figure 14. We attach two Impinj H47 passive tags on a pair of dumbbells. Each dumbbell weighs 2.5kg. Note that the H47 passive tag is a non-metal mount tag hence does not work on metal surfaces, we thus mount it on a foam plastics that is attached on the dumbbell. The foam plastics sufficiently isolates the tag from the metal. To conduct a comprehensive evaluation, we design a training workout with the ten free-weight activities. In this workout, each activity is required to be performed with three groups of ten repetitions. Then we recruit 15 volunteers (vary in age, gender, height, and weight) to follow this workout and track their training process during two weeks. The total duration of the training is 1,534 minutes, with over 4,500 repetitions in total. The 15 volunteers are diverse in weight, height and exercise frequency. Among them, some ones are our acquaintances, and others are not. To obtain standard templates, we recruit one gym trainer who has over 5 years experience and let him perform this workout under the same settings as the other volunteers. In addition, to get the mostly effective assessment of his training, we recommend the trainer to stand at a constant place during his free-weight exercises.

5.2 Activity segmentation

5.2.1 Evaluation metric

We evaluate the activity segmentation scheme based on six metrics [8]: insertion rate, deletion rate, fragmentation rate, merge rate, underfill rate and accuracy. The former five metrics are used to examine the segmentation robustness while the last is to examine the overall segmentation accuracy. The detailed explanation of these metrics are as follows:

- **Insertion rate:** The proportion of cases that FEMO detects an activity within the resting interval. It ex-

amines how resilient FEMO is to noisy Doppler peaks within resting intervals.

- **Deletion rate:** The proportion of cases that FEMO misses one activity. It examines how sensitive FEMO is to weak Doppler changes incurred by gym activity.
- **Fragmentation rate:** The proportion of cases that FEMO splits a single activity into multiple ones. It evaluates the ability of FEMO in processing complicated or incoherent activities.
- **Merge rate:** The proportion of cases that FEMO merges multiple activities into a single one. It evaluates FEMO’s ability in segmenting the activity with high training pace (*i.e.*, with tiny resting intervals).
- **Underfill rate:** The proportion of cases that the segmented activity is incomplete. It examines whether our method is capable to accurately while completely excavate the entire doppler profile for an activity.
- **Accuracy:** $\frac{\# \text{ of correctly detected activity}}{\# \text{ of activities that are performed}}$. It examines the overall performance of FEMO on activity segmentation.

5.2.2 Evaluation result

Fine-grained segmentation performance: Figure 15(a) shows the fine-grained performance of the segmentation scheme. In particular, the insertion rate is zero for activity 01 and 05. This value then increases steadily to around 0.005 for activities 02, 03, 04 and 10. Finally, it almost reaches 0.015 for activities 06, 07, and 08, which occupies ignorable portion of all the segments. For all activities, the insertion rate is extremely low. *This result indicates that our method is resilient to those doppler peaks within the resting interval.*

After checking the deletion rate, we find that it is also extremely small (below 0.01) for all the desired activities except activity 07. This is because that people usually raise the dumbbells up slowly when performing activity 07, leading to minor Doppler values. Such minor changes, in some cases, may be incorrectly put in resting intervals. Nevertheless, FEMO still controls the deletion rate below 0.02 for activity 07. *This result clearly demonstrates that FEMO is sensitive to Doppler changes incurred by the gym activity.*

As the bar chart in Figure 15(a) shows, FEMO achieves diverse fragmentation rate for different activities. The fragmentation rate is relatively small for activities 02, 04, 06 and 09, *i.e.* below 0.01 on average. It then triples for activities 07 and 10, and further quintuples for activities 01, 03, 05 and 08. This is because activities 01, 03, 05 and

08 contain a reciprocating motion, and people tend to keep a stable posture for a while within these activities. Despite the high disparity, we observe that the overall fragmentation rate for ten activities is below 0.065. *This result shows that the probability of segmenting one activity to multiple ones is very small.*

In Figure 15(a), we also notice a gap between the maximum and minimum merge rates. This is due to the diverse training pace on different gym activities. For example, in our experiment, we find that the resting interval is relative longer for activity 01, 02 and 03, which results in a lower merge rate. While for activity 06, 08 and 09, people tend to take a short rest after performing the activity. Such a short rest leads to activity omissions, yielding a relative larger merge rate. Although the gap exists, it is still relative small (approximate to 0.05) and FEMO can achieve a merge rate of below 0.065 in the worst case. *Therefore, we can conclude that our method scales well to different training paces.*

Figure 15(a) also shows the underfill rate of FEMO. We find that FEMO can precisely capture the entire Doppler profile of activity 02, 04, 06 and 09. Then the underfill rate increases slightly for more complex free-weight activities (*e.g.*, activity 01, 03, 05 and 08). The highest underfill rate is 0.04 for activity 08, indicating FEMO can precisely capture the entire doppler profile of this activity with a success rate of 0.96. *This result clearly states that our method can accurately excavate the entire doppler profile for each activity.*

Accuracy w.r.t Activity diversity: Figure 15(b) plots the result of overall segmentation accuracy with respect to different activities. The performance of the segmentation result can be categorized into three groups. The first group contains activities 01, 02, 05, and 10, where FEMO achieves a segmentation accuracy over 0.95. In the second group, FEMO achieves a segmentation accuracy between 0.9 and 0.95. This group contains activities 04, 06, 07, and 09. While the last group contains activities 03 and 08. For this group, FEMO achieves an accuracy between 0.85 and 0.9. The first two categories together cover eight out of ten activities, indicating that FEMO can achieve a high accuracy (*i.e.*, > 0.9) for the major activities. As for the last group, although the segmentation accuracy reduces, it is still above 0.85. *Therefore, we can conclude that FEMO is robust to activity diversity and can achieve desirable segmentation accuracy.*

Accuracy w.r.t Human diversity: In this experiment, we examine the effect of human diversity on the segmentation accuracy. For each volunteer, we compute the segmentation accuracy on different activities and get the overall accuracy distribution. The result is shown in Figure 15(c). The overall segmentation accuracy for the 15 volunteers maintains in a high level. Specifically, the median accuracy is above 0.9 for 12 volunteers. For the remaining 3 volunteers (volunteer 9, 10, and 11), we can see that FEMO achieves a relative inferior performance, with an accuracy of 0.83 on average. Interestingly, based on the physical characteristic records of volunteers, we find that all the three volunteers seldom go to the gym. As a result, the lack of exercise of these volunteers may lead to non-standard behaviors, thus lowering the segmentation accuracy of FEMO. We further investigate the fine-grained segmentation accuracy distribution of each volunteer. We can see that as shown in Figure 15(c), those who take exercise regularly (*e.g.*, volunteer 3, 4, 6 and 12), the box is relatively short, suggesting

| Actual label | Predicted label | | | | | | | | | |
|--------------|-----------------|------|------|------|------|------|------|------|------|------|
| | A 01 | A 02 | A 03 | A 04 | A 05 | A 06 | A 07 | A 08 | A 09 | A 10 |
| A 01 | 0.89 | 0.03 | 0 | 0 | 0.03 | 0.05 | 0 | 0 | 0 | 0 |
| A 02 | 0.01 | 0.89 | 0 | 0 | 0.05 | 0.05 | 0 | 0 | 0 | 0 |
| A 03 | 0 | 0 | 0.88 | 0 | 0 | 0 | 0.07 | 0.04 | 0 | 0.01 |
| A 04 | 0 | 0 | 0 | 0.93 | 0 | 0 | 0 | 0 | 0.07 | 0 |
| A 05 | 0.02 | 0.02 | 0 | 0 | 0.94 | 0.02 | 0 | 0 | 0 | 0 |
| A 06 | 0.06 | 0.02 | 0 | 0 | 0.03 | 0.90 | 0 | 0 | 0 | 0 |
| A 07 | 0 | 0 | 0.04 | 0 | 0 | 0 | 0.89 | 0.03 | 0 | 0.04 |
| A 08 | 0 | 0 | 0.03 | 0 | 0 | 0 | 0.03 | 0.93 | 0 | 0.01 |
| A 09 | 0 | 0 | 0 | 0.16 | 0 | 0 | 0 | 0 | 0.84 | 0 |
| A 10 | 0 | 0 | 0.01 | 0 | 0 | 0 | 0.02 | 0.02 | 0 | 0.95 |

Figure 16: Confusion matrix of activity recognition

that the segmentation accuracy of overall activities has a high level similarity with each other. While for those sedentary group (*e.g.*, volunteer 9, 10 and 11), we can see that the box of accuracy distribution is comparatively tall, indicating that FEMO holds different accuracy performance on different kinds of activities.

5.3 Evaluation on activity recognition

5.3.1 Evaluation metric

It is possible that the activity recognition system can miss, confuse, or falsely detect activities that did not occur. Thus, in evaluating our activity recognition scheme, we employ the following three metrics:

Precision P : $\frac{TP}{TP+FP}$, where TP and FP represent the true positives and the false positives. Precision is the fraction of correctly recognized activities that are relevant to all the recognized activities.

Recall R : $\frac{TP}{TP+FN}$, where FN is the false negatives. Recall is the fraction of the correctly recognized activities that are relevant to all this kind of activities.

False Positive rate FPR : The proportion of cases that FEMO mistakes an activity for other activities.

5.3.2 Evaluation result

Overall Accuracy on different activities: Figure 16 shows the confusion matrix for the 10 free-weight activities across 15 volunteers. Each row denotes the actual activity performed and each column represents the activity recognized by FEMO. Each element in the matrix represents the fraction of activities in the row that were regarded as the activity in the column. As is shown, the average accuracy is 0.90 with a standard deviation of 0.03 for 10 gym activities. This shows that we can extract rich information about free-weight activities from the Doppler profiles. *The result clearly shows that FEMO achieves a high and stable activity recognition performance, due to its efficient Doppler profile extraction scheme and robust profile matching algorithm.*

Examine the fine-grained performance: In this experiment, we examine the the precision, recall and false positive rate of the recognition performance. The result is shown in Figure 17. FEMO achieves an average precision of 0.90 with a standard deviation of 0.04. This result demonstrates that FEMO substantially returns more actual labels to the activity than error labels. As Figure 17 shows, although the recall fluctuates over different activities, it still maintains a high level for all of these 10 activities, achieving a mean value of 0.91 with a standard deviation of 0.03. This result

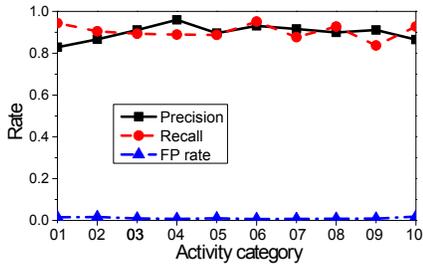


Figure 17: Precision, Recall and FP

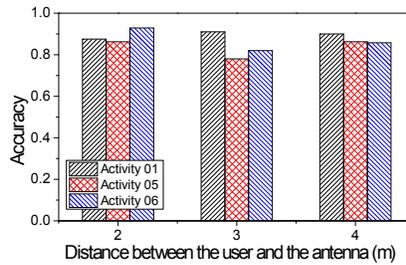


Figure 18: Accuracy over different antenna-to-user distances

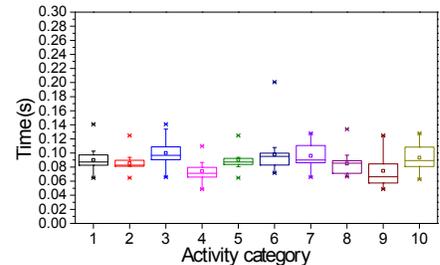


Figure 19: Computational latency

demonstrates that FEMO can recognize all of these 10 activities with high accuracy. *In a nutshell, the high precision and recall achieved by FEMO manifests that our recognition scheme scales well to different activities.*

We then examine the *FPR* of the recognition result. A higher *FPR* implies a great portion of other activities that are mistakenly classified as the targeting activity. As a result, people may ignore the system’s notifications and eventually abandon the system. Figure 17 shows the *FPR* cross 10 activities. We observe that FEMO achieves an average FP rate of 0.011 with a standard deviation of 0.004. Such a low FP rate suggests that FEMO rarely takes other activities as the targeted activity. Therefore, we can trust the activity recognition result with high confidence.

Impact of antenna-to-user distance: we further examine the impact of antenna-to-user distance on the activity recognition. In this trail of experiments, we ask three volunteers to perform activities under different antenna-to-user distance settings. Each activity is performed 30 times by each volunteer. For each activity, we average the recognition accuracy of these volunteers. We randomly pick out three activities and show their recognition accuracy in Figure 18. As the result indicates, when the user is with close proximity to the antenna, the recognition accuracy of these three activities all maintain in a high level. As we expand the user-to-antenna distance, the recognition accuracy changes moderately. *This result clearly demonstrates that our recognition scheme is insensitive to the user-to-antenna distance.*

5.4 Evaluation on activity assessment

FEMO can show the user the detailed assessment information about *Duration*, *Intensity*, *Smoothness*, and *Continuity* of his/her exercises. Then, to evaluate the activity assessment service, we survey the 15 volunteers to get qualitative feedback on the effectiveness and practicability of our system objectively based on their true experience. This includes the user satisfaction and attractive feature. For each aspect, the volunteer is required to give a score between 1 (lowest) to 5 (highest). And we treat these scores as the evaluation for our activity assessment module.

Overall feedbacks: Table 1 shows the overall feedback on FEMO. Participants consider that FEMO could help to reach the training goal with an average satisfaction score of 4.5, the standard deviation is 0.6. The potential to motivate regular training achieves a score of 4.7 on average, with $\text{std}=0.4$. Besides, participants also show high interest (4.6) to use FEMO for monitoring their free-weight training process. Overall, the feedback confirms that FEMO can help to reach training goals faster, motivate the training, and attract participants for long-term use.

| Feedback item | Rating | std |
|--|--------|-----|
| could help to reach the training goal | 4.5 | 0.6 |
| could help to rectify the irregular motion | 4.7 | 0.4 |
| would continue using it | 4.6 | 0.6 |

Table 1: feedback on FEMO after two weeks training

Attractive services: Table 2 shows the ratings on different services provided by FEMO. In the result, the most attractive services that the participants agree with is the accumulative training summary and the local assessment. These two services are scored 4.7. The global assessment is also attractive, achieving a score of 4.6. Although the daily activity summary is less attractive compared with the others, it still achieves a pretty high score (4.3 with $\text{std}=0.4$). Therefore, from above results we believe that FEMO really provides valuable feedback to the user.

5.5 System Overhead

Recognition Latency: FEMO should provide timely activity recognition such that the statistical information of the training process can be displayed promptly. We measure the recognition latency of each activity as the duration from the time point that bodybuilder finishes this activity to the time point the recognition result is shown in the UI module. For each kind of activities, we randomly choose 20 repetitions from 15 volunteers and record their recognition latency. The distribution of the recognition latency for each kind of activities is shown in Figure 19. It shows that FEMO achieves a recognition latency of 0.09s on average for these 10 activities. *Therefore, we can conclude that FEMO can provide on-site activity recognition result.*

Storage overhead: the RFID reader in FEMO periodically interrogates with tags to sense the body motion. The acquired data is thereby accumulated gradually. To examine the storage overhead of FEMO, we give a snapshot of storage augmenting when a volunteer performs activity 03. The result is shown in Figure 20. For comparison, we also plot the accumulated raw data within the same period acquired from the reader before the FEMO processing. We observe that the storage overhead of the raw data stream increases linearly with the time. It rapidly accumulates to over 7.5Mb at the end of the 30th second. Since conducting a workout commonly takes over one hour, recording this whole training process may incur about 1Gb storage overhead. In contrast, we can see that FEMO incurs significantly lower storage overhead, which accumulates to 0.33Mb at the end of the 30th second. From a long-term view, FEMO would accumulate about 40Mb of a one-hour training procedure,

| Service item | Rating | std |
|-------------------------------|--------|-----|
| accumulative training summary | 4.7 | 0.3 |
| daily activity summary | 4.3 | 0.4 |
| global assessment | 4.6 | 0.5 |
| local assessment | 4.7 | 0.4 |

Table 2: Service assessment after two weeks training

which is negligible compared to one-hour video based recording. Therefore, we can conclude that as a long-term running system, FEMO is much storage-efficient.

6. CASE STUDY

We deploy our FEMO system into a small fitness room in our lab. As the bottom-right figure shown in Figure 14, there are other equipments coexisting, such as treadmills and pedalling machines. We conduct experiments in this scenario and test the robustness of FEMO.

In this trail of experiments, we ask three volunteers to perform activities under various conditions. The first group of experiments (Condition #1) are conducted in normal environment, *e.g.*, relatively constant surroundings. In the second group of experiments (Condition #2), the volunteers perform their dumbbell activities with some other trainers running on the treadmills concurrently. During the third group of experiments (Condition #3), a disturber walks across the line of sight between the reader antenna and the volunteer frequently. We pick out three one-arm activities as the examples. Every volunteer performs each activity 30 times under each condition. The segmentation and the recognition accuracy are shown in Figure 21 and Figure 22. From the results we can see that the segmentation accuracy scales well under various conditions. On the other hand, the recognition accuracy has a significant reduction under condition #3. It tells that when the line of sight from the reader antenna to the user is frequently blocked, the Doppler profile will change and become hardly recognizable. However, both segmentation and recognition accuracy (say 0.93 and 0.87) under condition #2 are slightly influenced, which supports the deployment of FEMO in a real gym in the future.

7. DISCUSSION AND FUTURE WORKS

7.1 Discussion

This section discusses limitations and practical deployment issues.

(1) **Impact of Item Placement:** In a real gym, there exists various objects or fitness equipments, and they are considered as obstacles to FEMO. These obstacles (some of them are in the metal material) have slight impact on the accuracy of FEMO, as long as they do not exactly block the line of sight from the reader antenna to the user (tag). In common scenarios, the obstacles may reflect or absorb some signals, resulting in the reduction of signal energy. But they have negligible inference to the Doppler profile as well as our solutions.

(2) **Deployment Cost for Full Coverage:** The current Impinj reader we adopted is not particularly cheap. However, passive RFID technology is promising to be widely spread in the near future. And Impinj reader is a mainstream COTS device, which supports the extension and application of FEMO.

(3) **Scaling to Reader Diversity:** The impinj reader offers a standard API to retrieve Doppler values, but they are

not accurate. If we use other devices that can report accurate Doppler shift, FEMO can skip the signal preprocessing and utilize subsequent modules to monitor free-weight activities directly. Thus, FEMO works well with current readers that report low-quality Doppler shift readings, and will perform much better and more efficient if using those readers with high-quality Doppler shift reporting.

(4) **Handling Unknown Activities:** In this paper, we evaluate 10 popular free-weight activities. It is possible for users to perform incomplete actions during the exercises due to *e.g.* fatigue, which may lead to unknown Doppler profiles. In the current prototype, we filter out these unknown Doppler profiles by comparing their DTW distance with a pre-calibrated threshold. Similarly, other activities, such as picking up or putting down dumbbells, can also be filtered. Only those qualified profiles (*i.e.*, DTW distance below the threshold) are counted.

(5) **All-time Accessibility:** In FEMO, we require the user to do exercise within the interrogation zone of the antenna so that the tag can be interrogated interruptedly during the exercise. It is possible to relax this constraint on user location via boosting the transmitting power and optimizing the antenna position. It is also viable to extend the coverage of the reader by using multiple antennas to ensure all-time accessibility of tags even if the user is out of the interrogation zone of one or several antennas.

(6) **Extending to Multiple Users:** In the current prototype, we only evaluate FEMO in single-user cases. We envision FEMO can be extended to multi-user case via multiple antennas or distributed MIMO technology, *e.g.* each antenna separately monitors one user. As part of our on-going studies, we also leverage multiple antennas to monitor synchronized sports like group dancing training and computer-supported cooperative working.

7.2 Future Works

(1) **Scaling to non-stop activities:** We admit that our proposed segmentation algorithm is based on the fact that people tend to take a short rest after each activity to control the training pace. When adopting the algorithm to deal with non-stop activities, the segmentation accuracy will degrade to a certain extent. We will design a more robust algorithm in our future work.

(2) **Quantitative metric for activity assessment:** In the current prototype, we take the data collected from one gym trainer with 5-year experience as the standard templates. To further improve the effectiveness and availability of FEMO and provide more accurate assessment result, we would extend the standard templates to suit different users, *e.g.*, recruiting more professional trainers or coaches (varying in gender, height, weight, *etc.*) to provide the standard activity templates. Thus, FEMO can choose appropriate templates for the user based on his/her personal information when he/she first uses FEMO. Also, to improve the design of assessment criteria so as to further improve the availability of FEMO, we would consult these coaches to give more professional advices or acquire qualitative feedbacks from focus groups.

8. RELATED WORK

The design of FEMO is closely related to the following research categories.

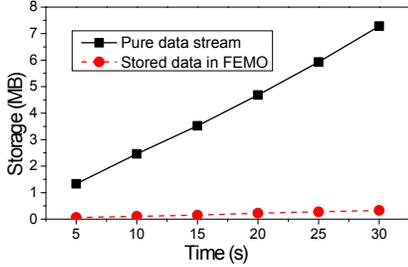


Figure 20: Storage overhead

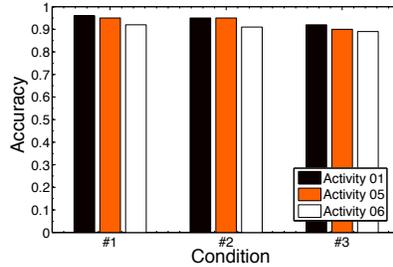


Figure 21: Segmentation accuracy over various conditions

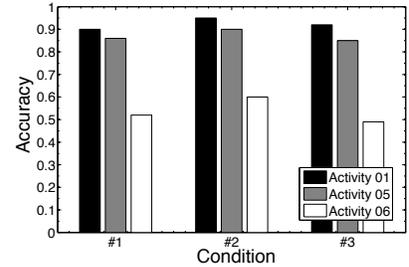


Figure 22: Recognition accuracy over various conditions

Activity Recognition: There is a large body of works on human activity recognition. Based on different processing patterns, the works in this domain can be broadly divided into two categories. The first category of works leverage dedicated sensors, *e.g.*, the gyroscope and accelerometer, for human activity recognition. These works build upon the fact that the sensor can provide rich information reflecting the gesture recognition. UbiFit Garden [10] infers the body movement via a on-body sensing module, and displays the result on the mobile terminal to encourage individuals’ training enthusiasm. RistQ [29] leverages the accelerations from a wrist strap to detect and recognize smoking gestures. Chang *et al.* [9] also use accelerometer sensors embedded in a glove to recognize and track the free-weight exercises in the gym. Although these proposals have demonstrated an inspiring power in activity recognition, the requirement of wearing some dedicated sensors is usually cumbersome or unsuitable for gym activity recognition. Besides, the performance of these sensor based schemes is sensitive to the hardware characteristics (*e.g.*, the sampling rate or computational capacity), which limits the wide adoption in practice.

Another brunch of activity recognition solutions exploit the change of wireless signals incurred by the human’s actions. RF-IDraw [40] can infer a human’s writing by tracking a passive RFID tag attached to his/her pen. E-eyes [41] leverages WiFi signals to recognize the in-home activity of human beings. In particular, there are many prior works detecting movements by resorting to Doppler effect [5, 19, 37, 38]. Their common insight is that the Doppler shift generated from the non-rigid-body motions of humans contains valuable information related to the human movement. The authors in [5, 37] extract certain features from the Doppler shift to distinguish among humans, animals, and vehicles. The work proposed in [19] can classify seven human activities (such as the running, walking, etc.) by extracting Doppler features from the signals collected by a 2.4GHz Doppler radar. By deploying Doppler sensors on the wall or tables, the work proposed in [26] can sense daily human activities using the Doppler effect in 24GHz microwaves. WiSee [30] exploits the Doppler effect of WiFi signals caused by the body reflection to infer nine typical human motions. Similar to our work, these approaches are device-free such that users are released from wearing or carrying any devices. However, they either adopt dedicated hardware (*e.g.*, Software defined radio, Doppler radar, *etc.*) or require complicated signal processing procedure, introducing a high deployment cost or huge computational overhead. Moreover, they have to rely on high-frequency sig-

nals (such as 2.4GHz, 24GHz), in which cases the Doppler change induced by human motions are obvious and easy to track. Our work releases this constraint in that we realize the accurate activity recognition and assessment by using a much lower frequency spectrum, *i.e.*, 860MHz ~ 960MHz, which is adopted by COTS passive RFID devices. Within this spectrum, the Doppler changes are subtle and vulnerable to noises [22]. In addition, we enable the on-site activity recognition and assessment by adopting several efficient techniques, *e.g.*, DTW, and provide rich feedbacks to the user for assessments.

Context-awareness Sensing: Recent advances in light-weighted sensors [23] and signal processing techniques have witnessed the prosperity of context-awareness sensing. CrossNavi [36] enables the crossroad navigation for the blind with the calibration of built-in sensors on commodity phones. SurroundSense [4] exploits built-in sensors on smartphones to characterize ambient features for logical localization. Similarly, SoundSense [24] analyzes the characteristics of sound events via a mobile phone to recognize the context around the user. The sound characteristics are further used for monitoring the face-to-face interaction [21], non-speech body sound recognition [31], speaker counting [42], and sleep quality monitoring [15, 13]. These works can provide partial indication on human environment, thereby enhancing the activity recognition performance.

9. CONCLUSION

In this paper, we present the design, implementation and evaluation of FEMO, a passive RFID based free-weight activity monitoring system. FEMO attaches passive RFID tags to the training devices, *i.e.*, dumbbell in this work, and leverages the backscattered signal for on-site activity recognition and assessment. The result of extensive experiments collected from 15 volunteers demonstrates that FEMO can be applied to a variety of free-weight activities, providing valuable feedbacks for users’ activity rectification.

10. ACKNOWLEDGEMENTS

This work was supported in part by National Basic Research Program of China (973 Program) under Grant No. 2015CB351705, NSFC under Grant No. 61572396, 61373175, and 61402359; the Natural Science Basic Research Plan in Shaanxi Province of China under Grant No. 2014JQ832; the Specialized Research Fund for the Doctoral Program of Higher Education under Grant No. 20130201120016. We thank the valuable comments from our shepherd Dr. Tian He and anonymous reviewers.

11. REFERENCES

- [1] Are Personal Trainers Worth the Price. <http://www.telegraph.co.uk/health/dietandfitness.html>.
- [2] Centers for Disease Control and Prevention. Physical activity recommendations for adults. cdc.gov/physicalactivity/everyone/guidelines/adults.html.
- [3] Learn How to Slim and Strengthen Your Midsection with the Best Core Exercises. <http://www.askthetrainer.com/best-core-exercises/>.
- [4] M. Azizyan, I. Constandache, and R. Roy Choudhury. SurroundSense: Mobile Phone Localization via Ambience Fingerprinting. In *Proceedings of ACM MobiSys*, 2009.
- [5] I. Bilik, J. Tabrikian, and A. Cohen. GMM-based Target Classification for Ground Surveillance Doppler Radar. *IEEE Transactions on Aerospace and Electronic Systems*, 42(1):267–278, 2006.
- [6] Dumbbells, Barbells, and Kettlebells. <http://www.bodybuilding.com/fun/grapgy9.htm>.
- [7] A Simple Protocol for Testing Your Work Capacity. <http://breakingmuscle.com/strength-conditioning/a-simple-protocol-for-testing-your-work-capacity>.
- [8] A. Bulling, U. Blanke, and B. Schiele. A Tutorial on Human Activity Recognition using Body-worn Inertial Sensors. *ACM Computing Surveys*, 46(3):33, 2014.
- [9] K.-H. Chang, M. Y. Chen, and J. Canny. Tracking Free-weight Exercises. In *Proceedings of ACM UbiComp*, 2007.
- [10] S. Consolvo, D. W. McDonald, T. Toscos, M. Y. Chen, J. Froehlich, B. Harrison, P. Klasnja, A. LaMarca, L. LeGrand, R. Libby, et al. Activity Sensing in the Wild: A Field Trial of Ubitfit Garden. In *Proceedings of ACM CHI*, 2008.
- [11] S. B. Eisenman, E. Miluzzo, N. D. Lane, R. A. Peterson, G.-S. Ahn, and A. T. Campbell. Bikenet: A mobile sensing system for cyclist experience mapping. *ACM Transactions on Sensor Networks*, 6(1):6, 2009.
- [12] Fitbit. <https://www.fitbit.com/>.
- [13] W. Gu, Z. Yang, L. Shangguan, W. Sun, K. Jin, and Y. Liu. Intelligent Sleep Stage Mining Service with Smartphones. In *Proceedings of ACM UbiComp*, 2014.
- [14] N. Y. Hammerla, T. Plötz, P. Andras, and P. Olivier. Assessing Motor Performance with PCA. In *Proceedings of IWFAR*, 2011.
- [15] T. Hao, G. Xing, and G. Zhou. iSleep: Unobtrusive Sleep Quality Monitoring using Smartphones. In *Proceedings of ACM MobiSys*, 2013.
- [16] iOS 8 HealthKit. <http://www.apple.com/ios/whats-new/health/>.
- [17] R. E. Kalman. A new Approach to Linear Filtering and Prediction Problems. *Journal of Fluids Engineering*, 82(1):35–45, 1960.
- [18] C. Kennedy-Armbruster and M. Yoke. *Methods of Group Exercise Instruction*. Human Kinetics, 2014.
- [19] Y. Kim and H. Ling. Human Activity Classification Based on Micro-Doppler Signatures Using a Support Vector Machine. *IEEE Transactions on Geoscience and Remote Sensing*, 47(5):1328–1337, 2009.
- [20] J. Kruger, H. M. Blanck, and C. Gillespie. Dietary and physical activity behaviors among adults successful at weight loss maintenance. *International Journal of Behavioral Nutrition and Physical Activity*, 3(1):17, 2006.
- [21] Y. Lee, C. Min, C. Hwang, J. Lee, I. Hwang, Y. Ju, C. Yoo, M. Moon, U. Lee, and J. Song. Sociophone: Everyday Face-to-face Interaction Monitoring Platform using Multi-phone Sensor Fusion. In *Proceeding of ACM MobiSys*, 2013.
- [22] H. Li, Y. Can, and P. S. Alanson. IDSense: A Human Object Interaction Detection System Based on Passive UHF RFID. In *Proceedings of ACM CHI*, 2015.
- [23] Z. Li, W. Chen, C. Li, M. Li, X.-y. Li, and Y. Liu. FLIGHT: Clock Calibration Using Fluorescent Lighting. In *Proceedings of ACM MobiCom*, 2012.
- [24] H. Lu, W. Pan, N. D. Lane, T. Choudhury, and A. T. Campbell. SoundSense: Scalable Sound Sensing for People-centric Applications on Mobile Phones. In *Proceedings of ACM MobiSys*, 2009.
- [25] H. Lu, J. Yang, Z. Liu, N. D. Lane, T. Choudhury, and A. T. Campbell. The Jigsaw Continuous Sensing Engine for Mobile Phone Applications. In *Proceedings of ACM SenSys*, 2010.
- [26] S. Masatoshi and M. Kurato. Activity Recognition Using Radio Doppler Effect for Human Monitoring Service. *Journal of Information Processing*, 20(2):396–405, 2012.
- [27] P. Melgarejo, X. Zhang, P. Ramanathan, and D. Chu. Leveraging Directional Antenna Capabilities for Fine-grained Gesture Recognition. In *Proceedings of ACM UbiComp*, 2014.
- [28] D. Morris, T. S. Saponas, A. Guillory, and I. Kelner. RecoFit: Using a Wearable Sensor to Find, Recognize, and Count Repetitive Exercises. In *Proceedings of ACM CHI*, 2014.
- [29] A. Parate, M.-C. Chiu, C. Chadowitz, D. Ganesan, and E. Kalogerakis. RisQ: Recognizing Smoking Gestures with Inertial Sensors on A Wristband. In *Proceedings of ACM MobiSys*, 2014.
- [30] Q. Pu, S. Gupta, S. Gollakota, and S. Patel. Whole-home Gesture Recognition using Wireless Signals. In *Proceedings of ACM MobiCom*, 2013.
- [31] T. Rahman, A. T. Adams, M. Zhang, E. Cherry, B. Zhou, H. Peng, and T. Choudhury. BodyBeat: A Mobile System for Sensing Non-speech Body Sounds. In *Proceedings of ACM MobiSys*, 2014.
- [32] L. G. Roberts. ALOHA Packet System with and Without Slots and Capture. *ACM SIGCOMM Computer Communications Review*, 5(2):28–42, 1975.
- [33] Y. Rubner, C. Tomasi, and L. J. Guibas. The Earth Mover’s Distance as a Metric for Image Retrieval. *International Journal of Computer Vision*, 40(2):99–121, 2000.
- [34] R. A. L. S. Kullback. On Information and Sufficiency. *Annals of Mathematical Statistics*, 1951.
- [35] S. Salvador and P. Chan. Toward Accurate Dynamic Time Warping in Linear Time and Space. *Intelligent Data Analysis*, 11(5):561–580, 2007.
- [36] L. Shangguan, Z. Yang, Z. Zhou, X. Zheng, C. Wu, and Y. Liu. CrossNavi: Enabling Real-time Crossroad

- Navigation for the Blind with Commodity Phones. In *Proceedings of ACM UbiComp*, 2014.
- [37] A. Stove and S. Sykes. A Doppler-based Automatic Target Classifier for a Battlefield Surveillance Radar. In *Proceedings of IEEE RADAR*, 2002.
- [38] D. A. Tesch, E. L. Berz, and F. P. Hessel. RFID Indoor Localization based on Doppler Effect. In *Proceedings of IEEE ISQED*, 2015.
- [39] How to measure exercise intensity.
http://www.weightwatchers.com/util/art/index_art.aspx?tabnum=1&art_id=20971.
- [40] J. Wang, D. Vasisht, and D. Katabi. RF-IDraw: Virtual Touch Screen in the Air Using RF Signals. In *Proceedings of ACM SIGCOMM*, 2014.
- [41] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu. E-eyes: Device-free Location-oriented Activity Identification using Fine-grained WiFi Signatures. In *Proceedings of ACM MobiCom*, 2014.
- [42] C. Xu, S. Li, G. Liu, Y. Zhang, E. Miluzzo, Y.-F. Chen, J. Li, and B. Finner. Crowd++: Unsupervised Speaker Count with Smartphones. In *Proceedings of ACM UbiComp*, 2013.
- [43] A. Zhan, M. Chang, Y. Chen, and A. Terzis. Accurate caloric expenditure of bicyclists using cellphones. In *Proceedings of ACM SenSys*, 2012.