Activity-aware ECG-based Patient Authentication for Remote Health Monitoring

Janani Sriram
Department of Computer Science
Dartmouth College, Hanover, NH, USA
janani.sriram@dartmouth.edu

Minho Shin
Institute for Security, Technology, and Society
Dartmouth College, Hanover, NH, USA
mhshin@cs.dartmouth.edu

Tanzeem Choudhury
Department of Computer Science
Dartmouth College, Hanover, NH, USA
tanzeem.choudhury@dartmouth.edu

David Kotz
Institute for Security, Technology, and Society
Dartmouth College, Hanover, NH, USA
kotz@cs.dartmouth.edu

ABSTRACT
Mobile medical sensors promise to provide an efficient, accurate, and economic way to monitor patients’ health outside the hospital. Patient authentication is a necessary security requirement in remote health monitoring scenarios. The monitoring system needs to make sure that the data is coming from the right person before any medical or financial decisions are made based on the data. Credential-based authentication methods (e.g., passwords, certificates) are not well-suited for remote healthcare as patients could hand over credentials to someone else. Furthermore, one-time authentication using credentials or trait-based biometrics (e.g., face, fingerprints, iris) do not cover the entire monitoring period and may lead to unauthorized post-authentication use. Recent studies have shown that the human electrocardiogram (ECG) exhibits unique patterns that can be used to discriminate individuals. However, perturbation of the ECG signal due to physical activity is a major obstacle in applying the technology in real-world situations. In this paper, we present a novel ECG and accelerometer-based system that can authenticate individuals in an ongoing manner under various activity conditions. We describe the probabilistic authentication system we have developed and present experimental results from 17 individuals.

Categories and Subject Descriptors

General Terms
Experimentation, Security

Keywords
Biometrics, ECG, Mobile Computing, Security

1. INTRODUCTION
As traditional healthcare systems struggle to cope with increasing demand and rising costs, many developed countries have been seeking new models of healthcare. Mobile computing and medical sensor technology offer a new paradigm for healthcare, namely remote healthcare [25, 19], that can reduce the cost [12] and improve the quality of healthcare services. Applications include long-term care for patients with chronic disease [22, 20], assistive technology for the elderly [21], risk management for people in rehabilitation [23], lifestyle coaching for people seeking to change unhealthy behavior [8], and fitness monitoring of athletes [18].

A remote health system relies on the ability of the system to correctly authenticate a patient, i.e., recognize whether the sensor data belongs to the right person, so that healthcare professionals can provide appropriate health services. Incorrect financial and medical decision could be made if someone else wears the patient’s sensors, with or without the patient’s permission. Existing authentication methods are not adequate for remote healthcare scenarios. If the authentication is based on passwords, smart cards, or secret keys, any person who has access to the credential can impersonate the patient.

Unlike traditional authentication methods, biometric authentication can discriminate between individuals based on their non-transferable traits such as face, voice, iris scan, or fingerprints. An authentication system usually operates in one of two different modes—identification and verification. During identification the
specifically handle the variability that arises from physical activity.

A one-time authentication has the risk of unauthorized post-recognition use, since someone else can wear the sensors after a successful authentication. Thus, it is desirable to have on-going authentication systems that confirms or rejects patient identity continuously.

The human ECG, an electrical signal that is associated with the electrical activity of the heart (Figure 1), offers several benefits as a biometric: it is universal, continuous, and difficult to falsify. The ECG signal from different individuals conforms to a fundamental morphology but also exhibits several personalized traits, such as relative timing of the various peaks, beat geometry, and responses to stress and activity. There are several factors that influence ECG signal: physiology and geometry of the heart, body build, gender, and age, which cause difficulties in accurate diagnosis [13, 15]. This inter-individual variability is potentially beneficial for discriminating individuals [36].

On the other hand, slow-changing factors, such as body habitus and age, can introduce long-term intra-subject variability, preventing permanent and unique identification of a person. Additionally, physiological responses to stimuli (such as stress and activity), and signal artifacts due to movement can result in large variability in a given individual’s ECG signal. Intra-subject variability may also be caused by electrode placement [1], pharmaceutical drugs [35], and physical activities [32, 14]. Therefore, if ECG information is to be used for authentication in real situations it is necessary to account for the sources of significant intra-individual variability.

In this paper, we introduce an activity-aware biometric authentication system that combines ECG information with accelerometer data to handle the variability that arises from physical activity. Specifically,

- We develop a probabilistic ECG-based biometric authentication scheme that confirms a patient’s identity in an ongoing manner across various activities.
- We implement a mobile prototype portable device using wearable ECG sensors.
- We demonstrate the feasibility of our patient-authentication scheme with 17 human subjects.

2. RELATED WORK

The validity of using the ECG in biometric recognition has been demonstrated in various related works, but most of the prior research considers ECG data collected from subjects at rest. For on-going recognition, we would like to explore the impact of activity on the performance of the ECG based biometric system. Furthermore, we are interested in using a simple wearable ECG acquisition system and demonstrate the practical feasibility of the approach using a mobile prototype.

There are primarily two classes of feature extraction approaches for the ECG that are seen in related work - fiducial and non-fiducial. Fiducial points are points of interest within the heartbeat, such as local maxima or minima. Typical approaches extract different time and amplitude features using these points of reference [4, 16, 30]. Non-fiducial approaches aim to extract discriminative information from the ECG trace without having to localize fiducial markers. Thus, the fiducial features capture information local to a single heartbeat and the non-fiducial features capture global patterns (not restricted to a heart beat) in the ECG trace.

Biel et al. [4] used 12-lead ECG features to identify 20 subjects at rest. They use proprietary equipment to extract automatically extract 30 different fiducial features from each of the 12 leads, which is expensive to compute and more than is likely needed for biometric identification or verification. Furthermore, the 12-lead ECG system requires meticulous placement of 10 electrodes on each person, which can impede its usability.

Shen et al. [30] used 7 fiducial ECG-features from the most invariant part of the heartbeat: the QRS complex and T-wave. They then classify the features with a combination of template matching and decision-based neural network classifiers. Using one-lead ECG data, they achieved an identification accuracy of 100% for 20 individuals selected from the MIT/BIH database. The study demonstrates the potential of the ECG in biometric identification but does not investigate the impact of physical activity.

Israel et al. [16] developed an identification method based on fiducial features and studied the impact of seven states of mental stress on 29 subjects. Their experiments show that normalized temporal distances between fiducial points of ECG signal are invariant to anxiety state and can be used for biometric recognition under different stress levels. However, they do not consider the impact of physical activities.

The studies mentioned above and many other ECG-based identification and verification methods [2, 17] use the fiducial points from different parts of the ECG signal to extract features. However, detection of fiducial points is prone to error since there is no universally acknowledged rule for defining where the wave boundaries or fiducial points, lie [27].

Recent studies [7, 9, 27, 34] show that non-fiducial approaches can also successfully identify individuals from their ECG signal. Plataniotis et al. [27] and Wang et al. [34] have proposed an approach that uses autocorrelation analysis (AC) coupled with Discrete Cosine Transform (DCT) and does not require segmentation of the ECG trace into heart beats. On the other hand, Chiu et al. [9] used a discrete wavelet transform of ECG signal for feature extraction and Chan et al. [7] used a distance measure based on a wavelet transform. These studies also fail to consider the impact of activity that contributes to physiological variation (such as increased subject heart rate) and noise within the spectrum of the ECG signal.

The motion of the person can generate noise in the ECG signal because of unstable contact of an electrode with skin. Tong et al. [33] and Raya and Sison [28] each proposed a noise-cancelling technique to compensate for motion artifacts in ECG data, by using additional sensors: an accelerometer and an anisotropic magnetore-
sitive (AMR) sensor. In this paper, we have attempted to use the accelerometer information to model the variability rather than to correct variability with potential for loss of discriminatory information. During feature extraction we focus our efforts on the more robust part of the ECG signal - the QRS complex.

Even when accurate ECG measurements are available, the characteristics of the ECG signal exhibit physiological changes in response to activity [32, 14]. A few studies [17, 11] chose to complement the ECG signal with other biometric or non-biometric sensors. For example, Israel et al. [17] combined face recognition and an ECG signal to overcome limitations of the ECG signal as a biometrics. Damousis et al. [11] takes a broader approach to build a reliable authentication and monitoring system using trait-based biometrics (such as face or voice), physiology-based biometrics (such as ECG or electroencephalogram), and behavioral data (such as gait). In this paper, we handle activity-induced ECG variation by extracting a set of accelerometer features that characterize different physical activities along with fiducial and non-fiducial ECG features. We believe that such a multimodal feature set can provide the classifier with the necessary auxiliary information to perform accurate authentication. Table 1 summarises the comparison of the above-noted papers with our scheme. It should be noted that we cannot realistically compare our accuracy with different schemes due to several interpretation issues such as the number of test samples per subject and the data collection methodology.

| Reference | Feature | Method | Subjects | Activity | Acc
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Biel [4]</td>
<td>Fiducial</td>
<td>PCA</td>
<td>20</td>
<td>No</td>
<td>100%</td>
</tr>
<tr>
<td>Shen [30]</td>
<td>Fiducial</td>
<td>Template Matching+DBNN</td>
<td>20</td>
<td>No</td>
<td>100%</td>
</tr>
<tr>
<td>Israel [16]</td>
<td>Fiducial</td>
<td>LDA</td>
<td>29</td>
<td>No</td>
<td>98%</td>
</tr>
<tr>
<td>Wang [34]</td>
<td>Both</td>
<td>KNN+LDA</td>
<td>13</td>
<td>No</td>
<td>96%</td>
</tr>
<tr>
<td>Chiu [9]</td>
<td>Non-Fiducial</td>
<td>Wavelet Distance+LDA</td>
<td>35</td>
<td>No</td>
<td>100%</td>
</tr>
<tr>
<td>Chan [7]</td>
<td>Non-Fiducial</td>
<td>wavelet DM</td>
<td>50</td>
<td>No</td>
<td>89%</td>
</tr>
<tr>
<td>Ours</td>
<td>Both</td>
<td>KNN+Bayesian</td>
<td>17</td>
<td>Yes</td>
<td>88%</td>
</tr>
</tbody>
</table>

Table 1: Comparison of related work with our scheme. The accuracy values represent the percentage of subjects who are correctly identified across a majority of their test samples.

3. METHODOLOGY

In this section, we describe the ECG plus accelerometer-based biometric authentication system we have developed and outline the preprocessing, feature extraction, and classification steps.

3.1 Sensors

As the base sensor board, we use the SHIMMER platform developed by Intel Digital Health Advanced Technology Group. The SHIMMER1 is a compact sensing platform with an integrated 3-axis accelerometer. SHIMMER runs the TinyOS operating system and integrates (via custom cabling) to a commercially available Polar WearLink Plus ECG chest strap2. We ran our initial data collection experiments using the BioMOBIUS software. Sensor data from the SHIMMER’s triaxial accelerometer and the ECG add-on board was sampled at 100Hz and transmitted via Bluetooth to BioMOBIUS which saves the data to a file.

3.2 Preprocessing and Feature Extraction

We segment the ECG and accelerometer traces into 400-sample windows (approximately 4 seconds of data) to obtain the feature windows $w_1, w_2, \ldots, w_n$. We use $w_i(j)$ to denote the $j^{th}$ sample of feature window $w_i$. We chose the size of the feature window so that multiple heart beats are present within a given window (at least 4 in our case). We attempt to reduce the impact of misclassified R-peaks by averaging beat features within the window rather than relying on features extracted from a single beat. Although, this method introduces a time lag of 4s, we can reasonably assume that change of activity or person takes longer than 4s. We use the notation $w_i^a$ and $w_i^b$ to denote the accelerometer data and ECG data respectively within the $i^{th}$ feature window. Any raw ECG trace collected using non-invasive surface electrodes usually has several artifacts, notably a low frequency baseline drift due to respiratory effects, electrode contact noise, and motion artifacts. Typically this noise is removed by high-pass or moving-average filtering techniques [28, 24, 10]. Since we collect our datasets during exercise, including durations of high-intensity activity, the ECG trace was also corrupted with the more troublesome motion artifact noise whose spectrum overlaps the ECG band (see Figure 2). The corresponding signal distortions cannot be easily eliminated by filtering.

We perform baseline correction before non-fiducial feature extraction. We employ an adaptive, beat-based linear interpolation approach to estimate the baseline from the line joining the q-minima. The estimated baseline is then subtracted to align all beats within a window. Baseline correction introduces some sharp discontinuities within the window, so before non-fiducial feature extraction we employ a high pass filter with coefficients adapted from [26].

We extract a combination of the two types of ECG features, fiducial and non-fiducial, from windows of the pre-processed signal. The final feature vector is shown in Table 2. Beat normalization ensures that the effects of beat-to-beat amplitude variations stay minimized. The normalized beats within a window have different periods since the adaptive segmentation in step 4 of Algorithm 1

![Figure 2: The ECG of a healthy subject at the beginning of exercise activity (above) and signal distortions due to motion artifact introduced as the subject proceeds to perform intense exercise activity (below).](image-url)

Table 2: Feature Vector representing the Biometric Profile of an individual

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1-F20</td>
<td>First 20 Normalized AC Coefficients</td>
</tr>
<tr>
<td>F21-F41</td>
<td>Spline interpolate</td>
</tr>
<tr>
<td>F42</td>
<td>Normalized Slope</td>
</tr>
<tr>
<td>F43</td>
<td>Normalized Slope</td>
</tr>
<tr>
<td>F44</td>
<td>R-R interval</td>
</tr>
<tr>
<td>F45-F50</td>
<td>Accelerometer $X,Y,Z$ means and variances</td>
</tr>
</tbody>
</table>

1http://shimmer-research.com/
2http://polarusa.com/
is based on the distance between the R-peaks of the beats on either side of the current one. The underlying beat morphology is obtained from the mean spline (cubic) interpolant (F21–F41) of the beats within the feature window.

**Fiducial analysis:** Our procedure for beat segmentation and fiducial feature extraction is described in Algorithm 1. In addition, we

**Algorithm 1** Detects the sets of beats $B$ and the sets of QRS markers, $q,r,s$, using an existing QRS detector [10]

1: The detected R-peaks are denoted by a set $r$ of sample index-amplitude pairs as $(r_j, u_x^r(r_j))$, for the $j^{th}$ beat of the $t^{th}$ window of the ECG trace.

2: For all $r_j$ in $r$, do

3: Search downhill from each R-peak to locate the Q and S minima as $(q_j, u_x^q(q_j))$ and $(s_j, u_x^s(s_j))$. (The normal width of the QRS peak is known to be 100 ± 20ms [10]. We incorporate this fact by searching between ±6 samples of the detected R-peaks.)

4: Align beat along the detected R-peaks by extracting a sequence of samples of size min$(|r_{j+1} − r_j|, |r_j − r_{j−1}|)$ so that the R-peaks are centered within the extracted beat segments. Discard if current beat $B_j$ is incomplete and cannot be centered.

5: Normalize the beat by clamping the R-peaks to 1 and the Q-minima to 0.

6: end for

7: Discard beats corresponding to poorly detected fiducials that do not contain Q and S within the assumed search interval.

8: return $\{q_j, (r_j, u_x^r(r_j)), (s_j, u_x^s(s_j))\}$

extract the mean R-R peak distances (F44) and the slopes QR(F42) and RS(F43). The RR-interval is normalized using the equation $s_i = \frac{e_i}{w_i}$, where $e_i$ is the unimodal feature vector (F1–F50), and an activity-unaware classifier which uses the unimodal feature vector (F1–F44). We combine estimated pairwise correlation distances for features F1–F20 and F21–F42 and Euclidean distances for all other features to obtain a modified KNN classifier. KNN typically uses Euclidean distance as its distance metric. The additional correlation distance metric represents the similarity in the shapes of two curves. We evaluated the performance of both the Euclidean-distance-based KNN and the modified-KNN (xKNN) classifiers.

**3.3 Classification**

The goal of classification is to identify a subject or to verify an identity claim from the sensor observations. We investigate the performance of two types of classifiers: K-Nearest Neighbor (KNN) and Bayesian network (BN).

**3.3.1 K-Nearest Neighbor**

Initially, we measure the benefit of incorporating activity information using simple KNN classifiers. We trained two KNN classifiers: an activity-aware classifier, which uses the multimodal feature vector (F1–F50), and an activity-unaware classifier which uses the unimodal feature vector (F1–F44). We estimated pairwise correlation distances for features F1–F20 and F21–F42 and Euclidean distances for all other features to obtain a modified KNN classifier. KNN typically uses Euclidean distance as its distance metric. The additional correlation distance metric represents the similarity in the shapes of two curves. We evaluated the performance of both the Euclidean-distance-based KNN and the modified-KNN (xKNN) classifiers.

**3.3.2 Bayesian Network**

Our hypothesis is that explicit modeling of activity states will lead to better recognition performance. We developed two Bayesian network classifiers to allow us to evaluate that hypothesis (shown in Figure 3).

Suppose we have $N$ persons whose identity $P$ is given by the labels $p = p_1, p_2, \ldots, p_N$. The biometric profile for a person $p_i$ is a set of $m$ ECG feature vectors $e = (e_1, e_2, \ldots, e_m)$ and the corresponding activity feature vectors $a = (a_1, a_2, \ldots, a_m)$. We assume that the ECG features are normally distributed, i.e., $P(E | P, H) \sim N(\mu, \Sigma)$, and depends on the person and the activity being performed. The problem of classification is now reduced to that of estimating the parameters of the conditional distributions of Equation (3). We discretize the accelerometer features, $A$, into distinct activity levels $H$ and obtain the Bayesian Network (BN) shown in Figure 3(a). The joint distribution of the BN is defined as


During training, we assume that the BN is fully observed.

The activity levels $H$ were obtained in two ways: from manual annotations and via unsupervised clustering. Manual annotations include three activity levels: still, low-intensity, and high-intensity. We also tested three types of unsupervised clustering techniques: K-means clustering with the Euclidean distance metric, K-means using city block distance and Gaussian Mixture Model (GMM) with diagonal covariance matrices. Ultimately, we chose the GMM which allows soft cluster boundaries and resulted in better performance. Since the manual annotations contained significant human error, there was no reliable ground truth for comparing the activity clustering. So we visually compared the inferred cluster IDs with our manual annotations and found that the inferences were close except for occasional discontinuities (see Figure 4).

We used the Bayes Net Toolbox [3] for training and inference. During testing both nodes $H$ and $P$ were hidden. We are interested in estimating the probability distribution of the hidden variable $P$,
A set of imposters for each individual. Selection of imposters has a large number of individuals, especially by mobile inference systems. A compact probabilistic model and maybe preferable when there are many legitimate, imposter pairs and the pooling of samples from multiple imposters. We describe our verification algorithm in Algorithm 2, in which we denote the training data for persons \( p = \{p_1, p_2, \ldots, p_N\} \) to be \( D(p_1), D(p_2), \ldots, D(p_N) \).

Algorithm 2 Training the pooled imposter model for Verification

1. for all claimants \( p_1 \in p = \{p_1, p_2, \ldots, p_N\} \) do
2. select a set of \( k \) most confused subjects as imposters \( \hat{p} = \{\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_k\} \) such that \( \hat{p}_j \in p \) and \( \hat{p}_j \neq p_i \) for any \( j \)
3. obtain the legitimate claimant dataset \( T \) as \( D(p_i) \)
4. for \( j = 1 \) to \( k \) do
5. add \( D(\hat{p}_j) \) the pooled imposter dataset \( \mathcal{F} \)
6. end for
7. Randomly sample from the pooled imposter dataset such that \( T \) and \( \mathcal{F} \) are of equal sizes.
8. Estimate model parameters for \( p_i \), training data \( \hat{D} = [T, \mathcal{F}] \)
9. end for

Finally, the system uses the verification score, given as the ratio of true and imposter model likelihoods, to make a decision about a claimed identity \( p_1 \):

\[
\frac{P(p_1 \text{ is a legitimate claimant})}{P(p_1 \text{ is an imposter})}
\]

(6)

Since both probabilities are obtained from the claimant’s model the proportional score simply serves to exaggerate differences in their probabilities for comparison against a threshold.

The potential to misclassify subjects who are not represented in the imposter pool is one of the drawbacks of the proposed verification model. But the compact representation and lower computational cost makes it a potentially appealing option. When we tested a dataset of 20 samples on the identification model running on an Intel Core2 Duo machine, it took approximately 0.6 seconds to predict all class labels. A verification decision on the same dataset took around 0.2s.

Combining Multiple Predictions: We refer to the individual prediction for each feature window as window identification/verification. A sequence of window identification/verification decisions are used to make person identification/verification decisions by majority voting. During identification, an entire test sequence is classified as belonging to person \( p_i \) if the majority of the predictions are for \( p_i \). During verification, an entire test sequence is verified for a certain claim \( p_i \) if the majority of the predictions are verified as legitimate. For ongoing verification, instead of majority voting over the entire dataset, we evaluate the performance of the classifier as follows,

1. Divide test dataset in chunks of 10 feature windows each.
2. For each chunk if the majority of windows is classified as legitimate, the claimant is considered legitimate. If not the claimant is an imposter.

4. EXPERIMENTAL RESULTS

To test the feasibility of our approach, we collected data from 17 volunteers under different activity conditions and across different days. To make it easier to compare with related approaches, we present the identification performance of the classifiers in addition to the verification results. We also describe the wearable prototype system that we built in Section 4.3.
4.1 Data Collection

We collected sensor data from normal healthy subjects for the duration of a workout session. Subjects were asked to exercise on the treadmill for 12–15 minutes (training dataset DT) or 5–7 minutes (test dataset DX). We collected test and training data on different days. The subjects were told to pace themselves and slowly work up to a jog; after at least 2 minutes of brisk jogging they were asked to slow down gradually to a halt. The data is manually labeled with annotations about the subjects pace and duration of exercise. We collected sitting (DS) and recovery (DR) data immediately before and after the training work-out. To construct the training dataset (DT, DS, DR) we selected approximately 10000 samples from each of the following activity annotations – sit, walk, run, endure run and recover. These are grouped into either 3 activity level annotations: sit+recover (labeled: still), walk (labeled: low intensity), run+endure run (labeled: high intensity) or 2 activity levels annotations: sit+recover (labeled: still) and walk+run+endure run (labeled: high+low intensity) for the supervised models. Some subjects had fewer than 10000 samples per annotation due to noise. In aggregate the training dataset size for each subject ranged between 40000-50000 samples. We tested our classifiers using samples from DX.

4.2 Identification and Verification Results

In order to compare with related approaches, we first evaluate the performance of the classifiers on data taken from subjects at rest (DS). The results are presented in Table 3 and are comparable to other existing approaches.

<table>
<thead>
<tr>
<th>No. of Activity Levels</th>
<th>KNN</th>
<th>xKNN</th>
<th>BN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>0.96</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>Recall</td>
<td>0.95</td>
<td>0.97</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Table 3: Identification performance for the sitting session (DS). We randomly select 20 windows as the test dataset and the remaining data as the training dataset.

Next we tested the performances of the different activity-aware and activity-unaware classifiers on the test dataset DX. Generally, the activity-aware classifiers outperformed the activity-unaware classifiers (as shown in Table 4) by being able to explain the intra-subject variability seen in the ECG signal. Among the activity-aware classifiers, the KNN classifiers do not explicitly model the effects of different activity levels. Nonetheless, the simple inclusion of the additional accelerometer modality is clearly useful, as seen by the improvement in performance. We use k=1 (one neighbor) based on our results from cross-validation. The BN classifier makes better use of the accelerometer data by explicitly modeling the activities, which leads to slightly better performance numbers even with the manual activity annotations.

Manual annotations are inconvenient, unreliable, subjective and unsuitable for fine-grained activity clustering (beyond a small number of levels). We used activity labels derived from unsupervised activity clustering using Gaussian mixture models. Table 5 shows the performance of the Bayesian classifier based on unsupervised activity clustering. The case where |H|=1 corresponds to the activity-unaware classifier.

Table 6 shows the performance of verification with varying number of imposters. Verification decisions are made according to Equation(5). We can view the false positive rate (FPR) as an indicator of the security of the system. As would be expected, FPR is reduced as more imposters were added to the pool. It should be noted that we evaluate the entire test dataset from all subjects against every combination of claimant and imposter to obtain the results shown in Table 6. When the number of imposters is large, the random sampling did not include sufficient samples from each activity and imposter combination. This resulted in an increase in false acceptances possibly due to an ill-constructed pooled dataset. With too few imposters the pooled dataset does not have enough information to begin with.

Figure 5 shows the ROC curve for the verification model. Depending on the nature of application, we can choose to optimize for fewer false positives or more true positives by selecting different operating points on the curve. To help reduce the number of overall false rejection, we aggregate the window verification decision. As described in Section 3.3.2, we combine 10 window predictions in an ongoing manner in our experiments. We use predictions from the 8 person imposter model. The results in Table 7 show, for each person, how often his identity claims are correctly accepted (legitimate claimant) and how often the claims made by the most confused imposter are accepted. The most confused imposter test provides the worst case verification numbers. Fewer imposter claims are incorrectly accepted for the other imposters in the dataset.
Figure 5: ROC curve for the verification model (8 imposters). The thresholds used are shown in the graph. We can see that with higher thresholds the system rejects too many legitimate users and with lower thresholds too many imposters are accepted. The dotted line shows \( y=x \).

<table>
<thead>
<tr>
<th></th>
<th>Legitimate Claimant</th>
<th>Imposter</th>
</tr>
</thead>
<tbody>
<tr>
<td>p1</td>
<td>100.00</td>
<td>0.00</td>
</tr>
<tr>
<td>p2</td>
<td>88.24</td>
<td>75.00</td>
</tr>
<tr>
<td>p3</td>
<td>90.91</td>
<td>0.00</td>
</tr>
<tr>
<td>p4</td>
<td>100.00</td>
<td>0.00</td>
</tr>
<tr>
<td>p5</td>
<td>100.00</td>
<td>72.27</td>
</tr>
<tr>
<td>p7</td>
<td>81.83</td>
<td>0.00</td>
</tr>
<tr>
<td>p8</td>
<td>100.00</td>
<td>70.00</td>
</tr>
<tr>
<td>p9</td>
<td>85.71</td>
<td>12.50</td>
</tr>
<tr>
<td>p10</td>
<td>80.00</td>
<td>0.00</td>
</tr>
<tr>
<td>p11</td>
<td>81.83</td>
<td>0.13</td>
</tr>
<tr>
<td>p12</td>
<td>93.75</td>
<td>83.3</td>
</tr>
<tr>
<td>p13</td>
<td>91.67</td>
<td>0.00</td>
</tr>
<tr>
<td>p14</td>
<td>100.00</td>
<td>0.00</td>
</tr>
<tr>
<td>p15</td>
<td>77.78</td>
<td>0.00</td>
</tr>
<tr>
<td>p16</td>
<td>100.00</td>
<td>0.00</td>
</tr>
<tr>
<td>p17</td>
<td>60.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 7: Acceptance rates for person verification (8 imposters) by aggregating window verification decisions. A person verification decision is made based on the most number of verified samples within a chunk. The test dataset consists of data from the legitimate claimant and imposter (most confused).

4.3 Prototype Application

To test the feasibility of our biometric verification approach, we implemented a prototype application as shown in Figure 6. Our primary design goal for the prototype was to set up a simple and reliable architecture to relay data from the sensor to an authentication server that runs the pattern recognition algorithms. We implemented a NesC application running on the SHIMMER to send ASCII packets of data from the onboard accelerometer and the add-on ECG sensor board over Bluetooth [6]. A thin client application running on a mobile phone (Nokia N95) wirelessly receives sensor data over Bluetooth and forwards it to the remote authentication server over an IEEE802.11g (Wi-Fi) link. Our protocol forwards chunks of 4000 samples so that data is sent to the authentication server every 40s. We implemented the analysis algorithms in MATLAB. The authentication server periodically analyzed the sensor data and logged the results. Although this approach works, significant development and evaluation still needs to be done to make this application practical and usable.

5. SUMMARY

In this paper we introduced a novel multimodal biometric authentication system based on wearable human electrocardiogram (ECG) and accelerometer sensors. We demonstrated, on data collected from 17 subjects, that activity-aware authentication systems can effectively deal with the ECG variability induced by physical activities performed in the real world. We believe the approached outline in the paper could facilitate ongoing authentication without requiring frequent and active participation from the user. In our future work, we plan include a broader range of activities and potentially more sensors. For instance, a galvanic skin sensor may be used to classify stressful activities. We also plan to develop the mobile verification platform further and perform verification locally on the mobile device.

6. REFERENCES


