KNOWME: An Energy-Efficient, Multimodal Body Area Network for Physical Activity Monitoring

GAUTAM THATTE, MING LI, SANGWON LEE, ADAR EMKEN, SHRIKANTH NARAYANAN, URBASHI MITRA, DONNA SPRUIJT-METZ, MURALI ANNAVARAM
University of Southern California, Los Angeles

Submitted to WHS'09 Special Issue

The use of biometric sensors for monitoring an individual's health and related behaviors, continuously and in real time, promises to revolutionize health care in the near future. In an effort to better understand the complex interplay between one's medical condition and social, environmental and metabolic parameters, this paper presents the *KNOWME* platform, which is a complete, end-to-end, body area sensing system that integrates off-the-shelf biometric sensors with a Nokia N95 mobile phone to continuously monitor the metabolic signals of a subject. With a current focus on pediatric obesity, KNOWME employs metabolic signals to monitor and evaluate physical activity. KNOWME development and in-lab deployment studies have revealed three major challenges: (1) the need for robustness to highly varying operating environments due to subject-induced variability such as mobility or sensor placement, (2) balancing the tension between achieving high fidelity data collection and minimizing network energy consumption, and (3) accurate physical activity detection using a modest number of sensors. The KNOWME platform described herein directly addresses these three challenges. Design robustness is achieved by creating a three-tiered sensor data collection architecture. The system architecture is designed to provide robust continuous multichannel data collection, and scales without compromising normal mobile device operation. Novel physical activity detection methods which exploit new representations of sensor signals provide accurate and efficient physical activity detection. The physical activity detection method employs personalized training phases and accounts for inter-session variability. Finally, exploiting the features of the hardware implementation, a low-complexity sensor sampling algorithm is developed, resulting in significant energy savings without loss of performance.

Categories and Subject Descriptors: C.3 [**Special-Purpose and Application-Based Systems**]: Real-time and embedded systems

General Terms: Algorithms, Design, Human Factors, Measurement, Performance

Additional Key Words and Phrases: Mobile Phones, People-centric Sensing

1. INTRODUCTION

Accurate detection of a person's state is critical in many healthcare domains, as it can be used to measure a person's physical activity (PA) level. PA has many health benefits, including reduced obesity, reduced risk for cardiovascular disease,

osteoporosis and type 2 diabetes, reduction of breast cancer risk by as much as 40%, reduction of risk for many other cancers including colon and endometrial cancer, and improved mental health and well-being [Spruijt-Metz et al. 2008]. Obesity, the health risk focused on in this paper, is a growing health care concern for youth and adults [Ogden et al. 2008; Caballero 2007] and outranks both smoking and drinking in its deleterious effects on health and health costs [Sturm 2002]. Promoting PA to reduce obesity is a key component to reverse this epidemic [PAGAC 2008]. Thus, the ability to record and interpret PA continuously with minimal intervention from the subject is the first step in understanding the complex interplay of environmental, genetic and socio-economic reasons for weight gain [Levine et al. 2000; Sirard and Pate 2001]. In spite of its significant potential, detecting and recording PA continuously remains notoriously difficult [Spruijt-Metz et al. 2008].

One type of sensor system which has been extensively used for assessing PA in preventive health applications is accelerometry. Unfortunately, such systems currently have many drawbacks in implementation [Spruijt-Metz et al. 2008]: participants must wear accelerometers for several days and then return them to researchers for data download, processing and interpretation. Also, current accelerometer systems for PA assessment are custom-made with proprietary software and only provide processed measurements, not raw accelerometer data; as a result, data cannot be compared across systems [Chen and Bassett 2005] and have limited utility for further signal processing. Additionally, calibration of accelerometers for PA typically employs high-overhead methodologies, such as whole-room calorimeters [Chen and Sun 1997], energy consumption measurements using oxygen consumption [Crouter et al. 2006] or doubly-labeled water [Sun et al. 2008; Pfeiffer et al. 2006].

Further problems in current approaches to PA, and ones which are independent of the type of system implemented, include the lack of a single accepted methodology for identifying PA, though pattern recognition is being used as one approach for acquiring easily interpretable and robust measures [Bonomi et al. 2009]. Furthermore, age, gender, and the amount of body fat versus lean tissue present are all subject-specific complicating factors in PA detection. The KNOWME platform remedies the drawbacks discussed above by developing real-time measurement of PA and providing immediate feedback, which can be used in adaptive interventions that are tailored to the individual needs of the subject [Collins et al. 2004].

Given this context, we believe that a combination of health sensors and mobile phones will play a significant role in overcoming the difficulties of practical, accurate PA state detection, which in turn will lead to improved assessment of PA levels. It is with this in mind that the KNOWME system is under development as an end-to-end mobile health platform [Annavaram et al. 2008] that interfaces wireless sensors with a Nokia N95 mobile phone via Bluetooth to precisely monitor heart rate using ECG sensors, blood oxygen levels using a pulse oximeter, and motion using the mobile phone's built-in accelerometer. Our current implementation collects also information from other phone-internal sensors including Global Positioning Systems (GPS) measurements and audio and video tags, which we plan to incorporate into our state detection algorithms in the future. We exploit the mobile phone's communication capabilities to transmit the combined health record information to a data server in real time, and its computing capabilities to process

the sensor data in order to detect and classify current physical activity.

Based on the development of our prototype system and in-lab deployment studies, this paper addresses three major research challenges:

(1) *Robust System Design.* The mobile phone interfaces with the sensors to collect and store sensor data in addition to acting as the primary computation node for processing the stored sensor data. Using a mobile phone for computation results in several unique challenges: mobile phones are highly sensitive to small increases in the memory footprint, which causes application instability. Mobile devices also rely on software implementation of complex floating point operations and hence using these operations, particularly during a signal processing stage, can lead to dramatic slowdowns in the application. In addition to computational issues, subject mobility and incorrect sensor placement can lead to poor signal quality. Hence, the first critical challenge that this paper addresses is design robustness.Thus, the system architecture is three-tiered, and designed to provide robust, continuous, multi-channel data collection, and to scale without compromising normal mobile device operation.

(2) *Accurate PA Detection and Classification.* Once data from multiple sensors is collected, highly accurate PA state detection is the next challenge to address. This goal is achieved via novel sensor signal representations and high performance classification algorithms. Our methods employ personalized training phases and account for inter-session variability, including inconsistent sensor placement by the subject. In particular, this paper describes new algorithms developed using multi-modal sensing, wherein data from multiple sensors are combined to achieve 85.5% accuracy in user-state detection.

(3) *Energy Efficiency.* The most precious resource on a mobile phone is the battery, which can be drained dramatically by simultaneous and continuous communication with multiple biometric sensors and sensor data analysis. Therefore, balancing the conflicting demands of high fidelity data collection and the need to conserve battery life is the final challenge addressed in this paper in the form of an algorithm for optimal sampling that minimizes the probability of misclassification between multiple PA states. KNOWME then uses a low-complexity approximation of the sampling procedure which can be easily implemented on the mobile phone, and which yields near-optimal performance.

The remainder of this paper is organized as follows. The three primary components of KNOWME and their relation to the current state of the art are outlined in Section 2. The first component, namely KNOWME's current hardware and software implementation, is presented in Section 3. The multimodal state detection algorithm is presented in Section 4. Section 5 details the final component, an optimal sensor-sampling algorithm and its low-complexity approximation which develops KNOWME as an energy-aware system. The data collection protocol and system performance using real data, are presented in Section 6. Section 7 outlines lessons learned and research in progress for the KNOWME system as a whole.

2. PREVIOUS WORK RELEVANT TO THE KNOWME COMPONENTS

The KNOWME Network is a complete end-to-end platform for health care monitoring, which in its current implementation focuses on a holistic approach for physical

activity monitoring with three tightly coupled components: (1) hardware and software implementations, (2) physical activity detection and (3) optimal sampling for energy-efficiency. Each of these components directly addresses important challenges which must be overcome in order to deploy our system for longer-term continuous health-monitoring. In this section, we overview the state of the art in each of these areas, and highlight the contributions of the KNOWME Network.

Mobile Phones and Health Care: The use of mobile phones in health care is emerging as a major research area [Patrick et al. 2008]. Recent studies have demonstrated the efficacy of mobile phones for monitoring alcohol abuse [Collins et al. 2003] and controlling tobacco [Rodgers et al. 2005] and drug addiction [Freedman et al. 2006]. These first-generation studies have employed mobile phones for the provision of advice or support via SMS or voice messages; mobile phones are also used for user-provided statistics via simple mobile device applications. In contrast, the KNOWME platform integrates external sensors with the phone's internal sensors to provide a comprehensive, continuous, real-time view of the user's state.

Multi-Sensor Physical Activity and User State Detection: Promising recent results have revealed the efficacy of wearable body accelerometers (single or multiple) for PA detection employing naive Bayes classifiers and C4.5 decision trees [Maurer et al. 2006; Bao and Intille 2004; Ravi et al. 2005], support vector machines (SVMs) [Ravi et al. 2005; Huynh et al. 2007], and hidden Markov models (HMMs) [Krause et al. 2005]. The detection performances of these classification techniques for activity-detection are found to be comparable [Jatoba et al. 2008]; we employ SVM classifers as described in Section 4. Five accelerometers offered a 35% PA recognition rate over one in [Bao and Intille 2004]. However, while the use of multiple accelerometers may be acceptable in some usage scenarios, they are cumbersome when used for continuous monitoring. Therefore, KNOWME uses a smaller, heterogeneous set of sensors which provide complementary information for 85.5% accuracy of detection.

Prior approaches based on multiple integrated sensor modalities have also been proposed [Lester et al. 2006; Parkka et al. 2006; Ermes et al. 2008; Miluzzo et al. 2008], but these have normally focused on system design, and have not explored the novel signal processing techniques for PA detection or energy efficiency employed herein. In addition to a tri-axial accelerometer, we use biometric signals from an ECG sensor[1] for PA recognition. Using different types of sensors will enable us to detect more sophisticated types of activities and motions; for example, there typically exists information that is extracted from the ECG signals that is not available in the accelerometer signals. Another previous method [Maguire and Frisby 2009] has also employed accelerometry and ECG signal for activity detection via kNN classifers, but employs a strict subset of the features considered herein. A principal component analysis (PCA) for ECG signals to study body movement was considered in [Pawar et al. 2007], but inter-session variability was ignored. This is an important complicating factor, as continuous deployment of WBAN systems invariably involves daily disparities in sensor placement and physical activity levels associated with a specific type of activity, *e.g.* walking at different rates. Ignoring

---

[1]Though KNOWME does track blood oxygen levels via a pulse oximeter (OXI), this sensor was found to have limited utility for PA classification and so is currently not involved in PA detection.

this inter-session variability thus adversely affects the detection accuracy. Herein, we employ Hermite polynomial modeling [Linh et al. 2003] for the purposes of PA recognition; this type of modeling has previously been employed in arrhythmia detection, and its use here allows for variations in the shape of the ECG waveform due to different activities and address inter-session variability, resulting in higher detection accuracy.

Energy Efficient Sensing: Another important challenge in WBANs is improving system energy-efficiency, since long-term deployment is limited by battery life. This problem is further exacerbated for WBANS employing Bluetooth or GPS. In contrast to most traditional sensor networks, KNOWME achieves energy efficiency by minimizing the use of Bluetooth communication between the mobile phone and the sensors: having the mobile phone *listen less* to sensors (while maintaining performance) directly translates to prolonged battery life. Energy saving strategies for mobile and/or sensor networks in general are a well-studied problem [Shih et al. 2002; Kang et al. 2008], with recent extensions to WBANs. A joint sampling and duty cycle schedule is examined in [Benbasat and Paradiso 2007]; similarly a task-scheduling scheme is considered in [Liu et al. 2007]. An energy optimization scheme which manipulates the connection latency of Bluetooth and Zigbee for WBANs is studied in [Yan et al. 2007]. While energy efficient context monitoring is considered in [Kang et al. 2008; Shin et al. 2009], these works focus on the provision of application programming interfaces for sensor management while identifying changes in PA state. In contrast, the KNOWME platform conducts people-centric, user-state recognition, rather than PA state change detection. Furthermore, KNOWME achieves energy-efficiency via optimized detection performance (probability of misclassification) for PA recognition.

Integrating User State in Mobile Applications: User state recognition has utility beyond WBANs for health monitoring, and is being employed in a new class of mobile cooperative services: real time traffic monitoring [Hoh et al. 2008] and social networking (such as Facebook and MySpace). User state can be defined by biometric signals, motion, location and background. Access to such information can facilitate application adaptation, as in the case of more context-aware applications which endeavor to exploit user state as defined by sensors embedded on mobile phones. For example, CenceMe [Miluzzo et al. 2008] uses the mobile phone on-board accelerometer to sense four PA states (walking, running, sitting, standing), which are then securely shared with "buddies" within a defined social network. KNOWME, on the other hand, provides higher-precision PA state classification for more states by employing multi-modal signal processing and heterogeneous sensors (accelerometers and ECG). While mobile devices are used to monitor mobility patterns in [Ryder et al. 2009] and change mobile device settings (ring tone, alert type) in [Siewiorek et al. 2003] via a context-aware mobile application, decision making in both these works is external to the mobile phone; Sensay [Siewiorek et al. 2003] does its sensing external to the mobile phone as well. In contrast, KNOWME mobile device fusion center is responsible for both data collection and decision making.

Thus, a review of the prior art suggests that that a holistic mobile energy-efficient health-monitoring system that provides accurate state detection and real-time user feedback with moderate complexity, and real-time implementation does not exist.
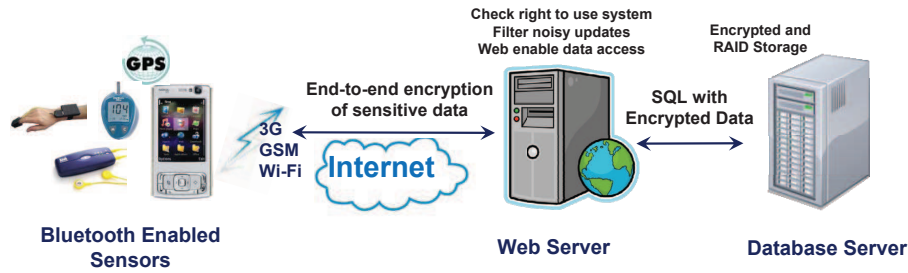
Fig. 1. Three-tier KNOWME architecture.

However, the successful development of the hardware and software platforms, as well as algorithms for physical activity detection and energy-efficient sampling, suggest that the prototype KNOWME platform is a step in this direction.

3. SYSTEM ARCHITECTURE

The KNOWME Network is built using a three-tier architecture as shown in Figure 1. The first tier is the WBAN layer that connects multiple sensors with a mobile phone. The second tier is a web server that receives data from the WBAN and performs data consistency checks and encryption. The web server then transmits the data to the back-end database server that stores the data. In our current implementation, the web server and database server are primarily used for data storage over long time horizons (multiple weeks) and do not participate directly in PA detection. In the future, these layers will enable remote visualization and long term historical data analysis.

3.1 Sensor Tier and Mobile Application

The WBAN layer is comprised of the on-body sensors and the mobile device, and thus experiences operating conditions and usage models which vary from user to user and over time. A WBAN node in KNOWME consists of a Nokia N95, and Bluetooth-enabled oximeter (OXI) and electrocardiograph (ECG) sensors from Alive Technologies [AliveTechnologies 2008]. The ECG and OXI sensors can sample at 300 Hz or 75Hz; sensor data is transmitted via Bluetooth to the N95 or locally stored on flash memory. The core of the WBAN is a mobile application which allows the N95 to pair with the external sensors via Bluetooth for sensor data collection. Additionally, sensor data is collected from in-built N95 sensors: an accelerometer (ACC) and GPS. End-user annotation, which is sensor specific, via audio and video tags is also enabled by the mobile application using user-specified semantics. KNOWME does not currently actively use GPS or audio/video tags; however ongoing research is examining the correlations between geo-spatial tags and metabolic data. The mobile application must gather data from multiple sensors with minimal user intervention and no interruption to regular mobile device functionality. Continuous, long-term data collection (*e.g.* 12 hours/day for multiple weeks) is envisioned and thus robustness is also necessary for the mobile application.

The mobile application is divided into two components: a server component and a client interface application for configuring the sensors and data visualization.
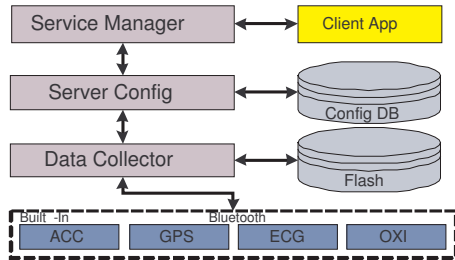
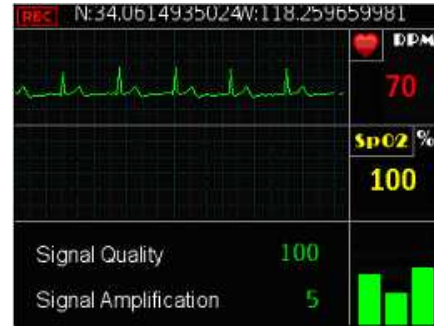Fig. 2. Server-client interactions in the mobile application.

Fig. 3. Mobile (Nokia N95) visualization client screenshot.

After considering the tradeoffs between productivity and efficiency, we chose the S60 Python for the server side and the J2ME for the client interface. Since these languages are not restricted by platform, our application is not tied to the Nokia N95. The server component of the mobile application comprises of four tasks: server configuration manager, data collector, device manager and service manager. Figure 2 shows the various components in the server application interacting with the client. During the initialization phase of KNOWME, the configuration manager reads the sensor configuration database to identify all the potential external sensors that can form the WBAN. The configuration database consists of information such as sensor names and Bluetooth MAC addresses necessary for pairing with the mobile phone. Note that not all the sensors in the database are active.

The configuration manager then creates a collector thread. The collector thread exploits the configuration database to initiate a Bluetooth scan to determine the list of active external sensors at that time. For each sensor, the collector thread allocates a memory buffer for sensor data and creates a device manager thread. The device manager thread is responsible for pairing with the corresponding sensor and to receive sensor data continuously. The device manager thread also deals with the sensor vagaries, such as lost connections, noisy data *etc.* When a connection is lost, as identified by no data reception for an extended period of time, the device manager thread initiates a new pairing request for the sensor. Finally, the server component also creates a single service thread to interface between the sensor data and any external application that may request the sensor data. For instance, the J2ME client visualization application requests the service thread to provide sensor data which it then plots.

The client component is designed to provide a user interface. Furthermore, new sensors are configured through the client component: the client application scans for a new Bluetooth device and finds the associated MAC address. These values are then communicated to the service thread within the server manager which will then store these values in the device database. The values are retrieved during the server initialization and configuration, as described earlier. Figure 3 shows the screen shot of the client visualization software.
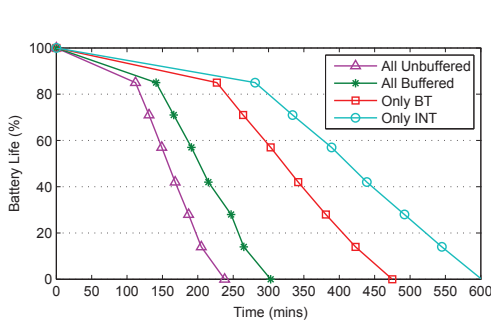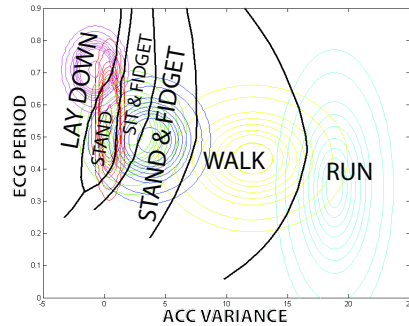
Fig. 4. Battery Usage of KNOWME.　　Fig. 5. Gaussian decision regions.

Sensors	# of packets per hour	Packet loss (%)
ECG & ECG	14999/14998	0/0
Oximeter & ECG	36008/14931	0/0.0045
Oximeter & ECG & ECG	36007/7351/7312	0/50.9/51.2

Table I. Data loss on Bluetooth channel

3.2 Sensor Energy Management and Data Integrity

Herein we quantify the severity of the energy consumption problem using the KNOWME platform which motivates our energy efficient sensor sampling algorithm derived in Section 5. In KNOWME, the sensors simply transmit data to the mobile phone fusion center; the Nokia N95 performs all the coordination, processing and computation tasks, as described in the previous section.

Figure 4 shows the battery level of the mobile phone as we turn on various sensors used in KNOWME. The curve labeled ALL_UnBuffer shows the battery drain over time on the mobile phone as we collect data from all the sensors (ECG, OXI, ACC, GPS) and write them to the local mobile phone flash drive without any buffering. The battery last just about 4 hours. For comparison, N95 has 10 hours of rated talk time and over 200 hours of standby time. The ALL_Buffer curve considers the case where we buffer the writes to the flash drive and send large packets to write to the flash. Thus, we find that buffering improves the battery life from 240 minutes to 299 minutes, a 25% improvement.

The two remaining curves contrast the effect of using the internal versus the external sensors on the Nokia N95. The curve labeled Only_BT shows battery life when using only the Bluetooth enabled sensors (ECG, OXI) giving a measure of the energy cost of Bluetooth alone: 475 minutes. If WBANs are to become ubiquitous, we expect that judicious sampling of external sensors to conserve mobile phone battery (and the external sensor battery as well) will be critical. Finally, the graph labeled Only_INT shows the battery drain when using the N95 internal sensors only (ACC, GPS), which results in a battery life of 600 minutes; a significant improvement over using the Bluetooth sensors. However, most of the energy consumption in the internal sensors is due to GPS: a single GPS reading typically consumes 6.616 J whereas an ACC reading costs 0.359 J.

A source of data loss is the inability of the mobile phone to keep up with the

data reception rate from the external sensors. Due to limitations in buffering while receiving data from multiple Bluetooth channels data packets are lost. Table I shows the data packets lost per hour due to this limitation.

Each instance of "ECG" in Table I denotes a distinct Alive Technologies ECG sensor communicating with the mobile device via Bluetooth at the 300 Hz sampling rate. Although, Bluetooth allows a maximum of six connections, a relatively high data loss is seen with the addition of a third sensor. The packet loss is negligible for the OXI sensor in the three-sensor scenario due to the significantly shorter packets employed relative to the ECG sensor. When the external sensors are paired with a traditional computing platform such as a desktop no data losses are observed. We conclude that the data loss is primarily due to the limited buffer size for Bluetooth communication in the mobile phone. As such, sensor data transmission rate plays a role in data integrity. The buffer limitation is also a reflection of the fact that current generation mobile phones are not designed for use in high data rate WBAN designs. There is also a fundamental tradeoff between the number of allowed concurrent connections and power consumption of the radio. Thus, both battery life and data integrity issues motivate developing energy-aware sensing and sampling protocols, which are addressed in Section 5.

3.3 Lessons from System Design

Based on our extensive system design effort over the past two years we make the following observations regarding the use of mobile phone in KNOWME. An inherent challenge is developing a mobile application to reside on a mobile device not originally designed for use in a WBAN. Limited memory and its scheduling presents a major challenge. Available memory is small (128 MB on the Nokia N95) and typically fixed. A recent enhancement is the provision of virtual memory (VM); without VM, the programmer must manage application memory usage. However, this task is difficult as the scheduling of tasks is not known *a priori.* Thus, the mobile phone may schedule a higher priority task, such as an incoming call, competing for memory with KNOWME and resulting in a non-repeatable memory allocation bug for KNOWME. This problem is exacerbated when the user launches multiple applications, resulting in KNOWME application instability. We separated the client visualization application from the primary sensing task using a separate server component. Most of the memory allocation in the server component is to buffer the sensor data and perform a bulk write to the phone's flash memory. As shown in Figure 4 buffering improves energy efficiency by 25%.

Typical signal processing methods employ complex floating point computations which are not currently supported by the N95 hardware. Such operations are executed as a software routine, consuming both significant power as well as time. Thus, naive implementation of signal processing algorithms on the mobile phone can cause dramatic application slowdowns. We used either approximations or precomputed values to reduce the impact.

The choice of available programming paradigms on mobile phones are also limited. For instance, the N95 supports Python, J2ME or Native Symbian programming to develop applications. Each of these options provide a tradeoff between programmer productivity and execution overhead. It is relatively simple to develop J2ME based mobile applications, but the Java runtime layer adds significant compu-

tational (and thus energy) overhead. We wrote the server component using Python S60 and the client visualization application using J2ME. We are currently exploring developing the server component as a native Symbian application to further reduce the Python overhead. The debugging capability is also severely constrained on the Nokia N95, as it does not support on-device application debugging. Thus, one must employ an emulator which may not faithfully capture mobile phone behavior. Hence, most of the system design effort was focused on the WBAN design with the primary goal of providing robustness under unpredictable operating conditions.

4. MULTIMODAL PHYSICAL ACTIVITY RECOGNITION

The KNOWME Network recognizes physical activities by fusing multimodal biometric signals from the accelerometer and ECG sensors. High-dimensional support vector machine (SVM) classifiers [Cristianini and Shawe-Taylor 2000] are employed, and use novel features extracted from the ECG signal, in addition to more traditional features, resulting in more accurate state detection in a multihypothesis testing framework. We use personalized models, which result in a higher detection accuracy compared to using a common set of biometric parameters for all subjects. Furthermore, our methods account for the inherent variability in a single individual's movements as a function of context, denoting this variation as *inter-session variability*, which further improves the detection accuracy.

4.1 Multihypothesis Testing for Activity-Detection

A feature is a characteristic measurement, transform or structural mapping of the input data which captures important patterns of desired phenomena (PA in our case) with reduced dimension. For example, the standard deviation of an accelerometer reading or the mean of the instantaneous heart rate via the ECG. In the sequel, we will explicitly outline three sophisticated methods for feature extraction from the ECG sensor. Using sets of features with complementary characteristics normally offers improvement in recognition accuracy [Ross et al. 2006].

For both the state-detection problem and the optimal sampling problem addressed in Section 5, we consider an M-ary hypothesis testing problem wherein each distinct activity maps to a hypothesis. We consider the following nine PA states: lying down, sitting, standing, sitting and fidgeting, standing and fidgeting, playing Wii tennis, slow walking, brisk walking and running. Our methods are easily extensible to more states. The feature set (see Section 4.2) is collected into a vector $\mathbf{y}$ and our PA detector outputs an integer which corresponds to one of the nine states, *i.e.* $g(\mathbf{y}) = i$, where $i \in \{1, 2, \cdots 9\}$. The goal of any M-ary detector is to take the feature space and divide it into M distinct regions, whenever a multiple-dimension feature vector falls into one of the regions, the activity state associated with that region is declared. Figure 5 provides an example of two-feature decision regions for a seven state scenario. This figure was generated using the data of a single individual and a simplified detector, outlined in Section 5.1; however it facilitates highlighting the fact that there are complex interactions between features which yield non-linear decision regions. We further note that increasing the number of good features improves detection accuracy, as is evidenced in the case of the ECG sensor results in Table II (see Section 6.2). And finally, the SVM classifiers that we employ, determine near-optimal decision regions for very large dimensional
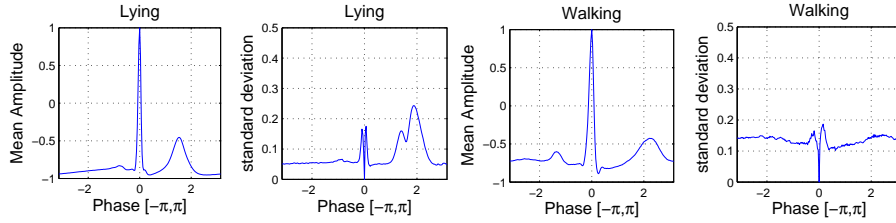
Fig. 6.   The mean and standard deviation of normalized ECG signals.

feature sets (we use 128 features).

4.2   Feature Extraction

We consider four types of features. The first set of features, which we denote as *conventional*, were selected based on their efficacy as demonstrated in the literature regarding wireless body area sensor networks; for the accelerometer, the conventional ACC features are shown in Box 1 (see Section 4.2.5), and for the ECG sensor, the mean and variance of the instantaneous heart-rate constitute the conventional features. The other three features sets are comprised of features that were chosen because their inclusion improves detection accuracy. These features result from more complex processing of the ECG signal: (i) the Hermite polynomial expansion (HPE) coefficients and (ii) the principal component analysis (PCA) error vector, which have been previously studied in [Linh et al. 2003] for a different context and [Pawar et al. 2007], respectively, and (iii) the standard deviation of multiple normalized beats which is novel to our work. These techniques model the underlying signals and the resultant model parameters are the features. First, we describe the required pre-processing of the collected biometric ECG signal, followed by feature extraction, and finally outline the temporal accelerometer features.

4.2.1   *Pre-processing of the ECG Signal.* Each type of body movement induces a particular type of motion artifact in the ECG signal. For the jth heartbeat observation under the ith hypothesis, or activity, the continuous-time recorded ECG signal, $r_{ij}(t)$, is modeled as [Pawar et al. 2007]

$$r_{ij}(t) = \theta_i(t) + \chi_{ij}(t) + \eta_{ij}(t), \tag{1}$$

where $\theta_i(t)$ is the cardiac activity mean (CAM) which is the normal heart signal, $\chi_{ij}(t)$ is an additive motion artifact noise (MAN) due to i^{th} class of activities, and $\eta_{ij}(t)$ is the sensor noise present in the ECG signal. Since the length of each heartbeat is different due to inherent heart rate variability, the first step of pre-processing normalizes each heartbeat waveform to the same time duration (in the phase domain) and amplitude range [Pawar et al. 2007; Clifford et al. 2006].

The D-dimensional vector representation of $r_{ij}(t)$ over one heartbeat is denoted $\mathbf{r}_{ij}$ and the D-dimensional vector representations of the corresponding CAM, MAN, and sensor noise components are $\boldsymbol{\theta}_i$, $\boldsymbol{\chi}_{ij}$, and $\boldsymbol{\eta}_{ij}$, respectively. Figure 6 shows the mean and standard deviation of the normalized ECG signal for different activities. One of our innovations over [Pawar et al. 2007] is the recognition that **both** CAM and MAN carry discriminative information between different PAs.

4.2.2 *Principal Component Analysis.* Principal component analysis (PCA) is used for MAN component χ_{ij} feature extraction. For the ith activity class, we use ν_i heartbeats to estimate the CAM $\boldsymbol{\theta}_i$ (as in [Pawar et al. 2007]),

$$\tilde{\boldsymbol{\theta}}_i = \frac{1}{\nu_i} \sum_{j=1}^{\nu_i} \boldsymbol{r}_{ij}. \tag{2}$$

We note that the number of heartbeats available for training, ν_i, is different for each of the activities. Subtracting the CAM from the signal $\boldsymbol{r}_{ij}$ yields residual activity vectors $\acute{\boldsymbol{r}}_{ij}$:

$$\acute{\boldsymbol{r}}_{ij} = \boldsymbol{r}_{ij} - \tilde{\boldsymbol{\theta}}_i = \chi_{ij} + \acute{\boldsymbol{\eta}}_{ij}, \tag{3}$$

where $\acute{\boldsymbol{\eta}}_{ij}$ includes both the sensor noise and the CAM estimation noise induced by the session variability. As noted in [Pawar et al. 2007], although the signal component due to MAN has smaller amplitude than CAM, it has much greater amplitude than the sensor noise, *i.e.* $|\boldsymbol{\eta}| \ll |\chi_i| < |\boldsymbol{\theta}_i|$, $\forall i$ (where $|\cdot|$ is the 2-norm). Thus, the MAN has a dominant influence on the shape of the residual activity vector $\acute{\boldsymbol{r}}_i$. For each activity class i, we now compute eigenvectors and eigenvalues using the eigen-decomposition of the covariance matrix $\boldsymbol{\Sigma}_i$ of $\acute{\boldsymbol{r}}_{ij}$. Let $\mathbf{E}_i = [\mathbf{e}_{i0}, \mathbf{e}_{i1}, \cdots, \mathbf{e}_{i\kappa_i}]$ be a set of eigenvectors corresponding to the $\kappa_i < D$ largest eigenvalues, and let $\mathbf{p}_{uj}$ be a vector representation of the jth normalized testing ECG heartbeat after pre-processing. We subtract the class mean $\tilde{\boldsymbol{\theta}}_i$, see (2), from $\mathbf{p}_{uj}$ to yield $\tilde{\mathbf{p}}_{ij}$. Thus, a measure of the reconstruction error in i^{th} activity's residual vector eigenspace, for the jth ECG heartbeat observation, is defined as:

$$\mathrm{RE}_j^{\mathrm{PCA}}(i) = \left| \tilde{\mathbf{p}}_{ij} - (\mathbf{E}_i \mathbf{E}_i^T) \tilde{\mathbf{p}}_{ij} \right|^2, \tag{4}$$

which is summed over ν_F heartbeat observations. The PCA residual error vector $\mathbf{RE}^{\mathrm{PCA}} = [\mathrm{RE}^{\mathrm{PCA}}(1) \ \mathrm{RE}^{\mathrm{PCA}}(2) \ \ldots \ \mathrm{RE}^{\mathrm{PCA}}(M)]$ is a feature for activity-detection, and accounts for inter-session variability.

4.2.3 *Hermite polynomial expansion (HPE) Coefficients.* A Hermite polynomial expansion (HPE) is used to model the CAM component $\boldsymbol{\theta}_i$ of the sampled ECG signal, and the resulting coefficients are another feature set for classification. Hermite polynomials are classical orthogonal polynomial sequence representations [Arfken et al. 1985] and have been successfully used to describe ECG signals for arrhythmia detection [Linh et al. 2003], but do not appear to have been previously used for PA detection. As shown in Figure 6, the shape of the CAM component for each activity is different and thus these signals can be used to distinguish between different PA states. Denote the pre-processed D-dimensional ECG curve vector and polynomial order by $x[n]$ and L, respectively. The HPE of the heartbeat can be expressed as [Linh et al. 2003]

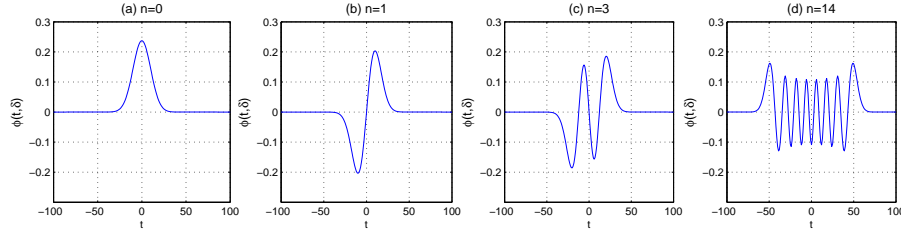$$x[n] = \sum_{l=0}^{L-1} c_l \psi_l(n, \delta), \qquad n \in \left[-\frac{(D-1)}{2}, \frac{(D-1)}{2} \right], \tag{5}$$

Fig. 7. Hermite basis functions with $\delta = 10$ and $D = 201$: (a) n=0, (b) n=1, (c) n=3, (d) n=14.

where l is an integer, $\{c_l\}$, $l = 0, 1, \cdots, L-1$, are the HPE coefficients, and $\psi_l(n, \delta)$ are the Hermite basis functions defined as:

$$\psi_l(n, \delta) = \frac{1}{\sqrt{\delta 2^l l! \sqrt{\pi}}} e^{-n^2/2\delta^2} H_l(n/\delta). \tag{6}$$

The functions $H_l(n/\delta)$ are the Hermite polynomials [Arfken et al. 1985] which are defined recursively by:

$$H_0(t) = 1, \quad H_1(t) = 2t, \tag{7}$$

$$H_l(t) = 2t H_{l-1}(t) - 2(l-1) H_{l-2}(t). \tag{8}$$

Figure 7 presents several Hermite basis functions for different orders; the higher the order, the greater the frequency of the signal. The HPE basis functions can be denoted by a $D \times L$ matrix $\mathbf{B} = [\boldsymbol{\psi}_0 \ \boldsymbol{\psi}_1 \ \cdots \ \boldsymbol{\psi}_{L-1}]$; the expansion coefficients $\mathbf{c} = [c_0 \ c_1 \ \cdots \ c_{L-1}]$ are obtained by minimizing the sum squared error E:

$$E = \left\| x[n] - \sum_{l=0}^{L-1} c_l \psi_l(n, \delta) \right\|_2^2 = \left\| \mathbf{x} - \mathbf{Bc} \right\|_2^2 \ \rightarrow \ \mathbf{c} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{x}. \tag{9}$$

Now let ν_F denote the fixed number of normalized heartbeats in each running window. The CAM component of this window is estimated using (2), but with $\nu_F = 10$ heartbeats instead of ν_i heartbeats. Figure 8 shows an example of HPE based reconstruction which is nearly identical to the original waveform.

4.2.4 *Standard deviation of multiple normalized beats.* We had previously observed that the variance of the ACC measurements offered discrimination capability [Annavaram et al. 2008]; this feature for the ECG signal also has utility. From Figure 6 it is clear that higher intensity states (walking) have a larger standard deviation than lower (lying). If the user is lying down or sitting, then the normalized heartbeat shapes are more consistent or similar within the whole processing window, but if the user is walking or running, then the normalized ECG shape can vary dramatically and become noisy. Thus, the sum of standard deviations for all the normalized bins (D bins) in the window is also employed as a classification feature. To our knowledge, this feature has not been previously used for PA classification.

4.2.5 *Temporal Accelerometer Features.* For the tri-axial ACC, a set of conventional temporal features (see Box 1) are extracted from the recorded signals of each axis, in a specified window (see Section 6.2). These features have been previously
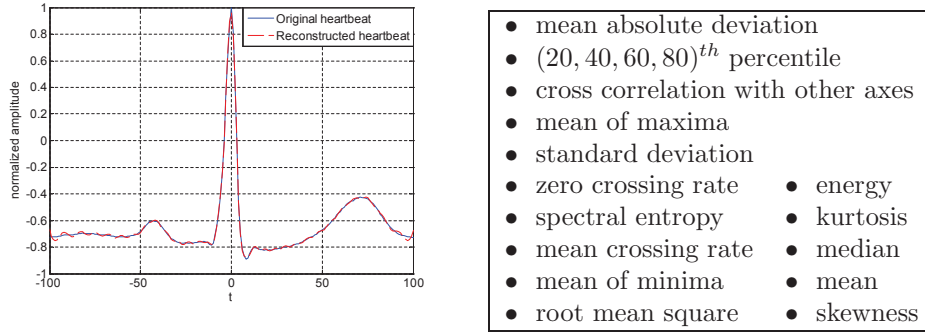
Fig. 8. The original and the reconstructed ECG heartbeat from HPE.

- mean absolute deviation
- $(20, 40, 60, 80)^{th}$ percentile
- cross correlation with other axes
- mean of maxima
- standard deviation
- zero crossing rate   • energy
- spectral entropy   • kurtosis
- mean crossing rate   • median
- mean of minima   • mean
- root mean square   • skewness

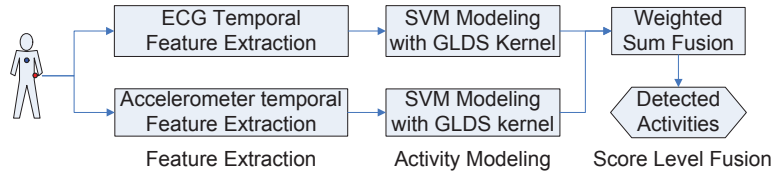Box 1: *Conventional* temporal accelerometer features.

Fig. 9. The proposed physical activity recognition system overview

studied in [Maurer et al. 2006; Bao and Intille 2004; Huynh 2008; Godfrey et al. 2008], which employ various subsets of the features listed, as well as frequency-domain features that are not considered in this work.

In the sequel, we explain the SVM classifier used. The feature vector that we employ, $\mathbf{y}$, consists of two components: $\mathbf{y}_E$ and $\mathbf{y}_A$, which denote the features from the ECG signal and accelerometer signal, respectively. In total 128 features are employed; 72 features from the ECG and 56 features from the tri-axial accelerometer.

4.3 Activity-Detection via SVM Classifiers

A support vector machine (SVM) using the generalized linear discriminative sequence (GLDS) kernel [Campbell et al. 2006] is used for classification. The accelerometer features $\mathbf{y}_A$ are computed for each of the N_1 processing windows of the signal, $\mathbf{s}_i^n$ or $s_i^n(t)$, for the nth window and i^{th} activity-class. Similarly, the ECG features $\mathbf{y}_E$ are computed using ν_F heartbeats for each of the N_2 windows, $\{\mathbf{r}_{ij}\}_{j=\nu_F n+1}^{(n+1)\nu_F}$. The ACC and ECG SVM classifier outputs are fused at the score level; Figure 9 overviews the activity-modeling and classification procedure. The SVM is a binary classifier constructed from sums of a kernel function $K(\cdot, \cdot)$, and specified as

$$f(\mathbf{y}) = \sum_{l=1}^{N_{\text{SVM}}} \alpha_l t_l K(\mathbf{y}, \mathbf{y}_l) + d, \qquad (10)$$

where $\mathbf{y}_l$, $l = 1, \ldots, N_{\text{SVM}}$ denotes the lth support vector, $\mathbf{y}$ is the feature vector from the testing set, t_l is the ideal output, and d is a bias term that is determined via the training data. The GLDS kernel, using first-order polynomials, allows

computation of the score function via a simple inner product, which is very computationally efficient (see [Campbell et al. 2006] for details); the LIBSVM tool [Chang and Lin 2001] is used for the SVM model training. Fusion at the score level uses a weighted sum, wherein the weights are determined by a logistic regression based on the training data [Ross et al. 2006]. This method is particularly useful when the individual sub-systems, *i.e.* the ACC and ECG feature sets, have significantly different performances on the desired task [Ross et al. 2006]. This discrepancy in performance between the two sensors is leveraged in the development of the optimal sampling protocol, which is derived in the next section, to yield an energy-efficient KNOWME health-monitoring application.

5. ENERGY-EFFICIENCY VIA OPTIMAL SAMPLING

The problem considered is allocating fixed number of samples N between K sensors to minimize probability of misclassification. The energy-efficiency method for the KNOWME Network differs from the energy-aware systems described in Section 2 in two key ways: (1) energy-efficiency is achieved via performance optimization and (2) energy minimization is focused on the mobile device, and not the sensors or the network as a whole. The mobile device is the bottleneck for long-term deployment. The mobile device can turn off its Bluetooth for sensors whose measurements are not currently being utilized, resulting in energy-efficient application development. We first present a simplified signal model for the features $\mathbf{y}$ received at the mobile device, and then outline our optimization problem which is motivated by the battery life and data integrity issues discussed in Section 3.2.

The optimal sampling protocol dictates how many measurements to collect from each of the sensors. Deriving an optimal sampling protocol based on the SVM classifier in Section 4.3 is intractable for two reasons. First, there is no closed form expression for the performance of the SVM to optimize. Second, the SVM uses far too many features and thus optimal feature selection would be a combinatorial problem which would incur prohibitive complexity. Thus, we use a single exemplary feature per sensor and Gaussian models to simplify the computations. While using only a subset of the features is not ideal, we observe that the application of our optimized sampling protocol using a subset of the features does provide substantial energy savings when applied to our novel PA detection methods outlined in Section 4. Furthermore we underscore that the full set of sensors are employed for both detection and optimized sampling.

Obtaining the optimal allocation of measurements requires an exhaustive search over all possible K-partitions of the N total measurements, since all possible integer allocations of measurements amongst the heterogeneous sensors must be considered to find the optimal allocation. As the number of total available measurements and number of sensors increases, a exhaustive search becomes computationally expensive, requiring $\mathcal{O}(N^{K-1})$ function evaluations with a typical N beings several tens of samples. To this end, we develop an analogous optimization problem which yields an approximately optimal solution, but which can be solved using lower complexity continuous-valued vector optimization techniques which requires $\mathcal{O}(K)$ function evaluations. This optimal allocation is then applied to the SVM classifier and shown to yield significant energy savings with no performance loss in Section
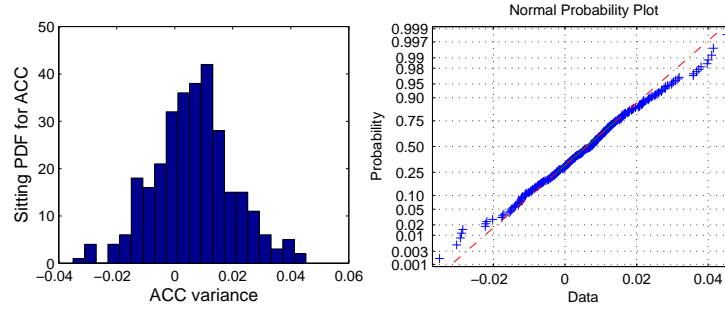
Fig. 10. Histogram and normality plot for the Sitting hypothesis using the ACC variance feature.
This motivates the Gaussian models adopted for the development of the optimal sampling protocol.

6. Furthermore, the low complexity optimization is more easily implemented on
the mobile device.

5.1 Re-visiting the Signal Model

In Section 4.3, the feature vectors $\{\mathbf{y}_A, \mathbf{y}_E\}$ are directly used by the non-parametric
high-dimensional SVM for classification. In contrast, for developing an energy-
efficient algorithm, we consider a simpler parametric model for a subset of features.
We make the following key assumptions: (i) based on our prior work [Annavaram
et al. 2008], we approximate the statistics of the key features as Gaussian, and (ii)
to capture the temporal correlation of the features, we further impose an auto-
regressive (AR), order 1 model. The data histogram and the fit of the Gaussian
model via a normality plot is shown in Figure 10, for the Sitting hypothesis us-
ing the ACC sensor. We notice that the normality plot validates the choice of
Gaussian models for the features $\mathbf{y}$. Also, the AR model has been previously em-
ployed to estimate electroencephalogram (EEG) signals [Mohammadi et al. 2006]
and physiological hand-tremors [Zhang and Chu 2005]. In addition to the temporal
correlation, we allow for sensing and communication noise.

For our selected features, the correlation between features (ACC and ECG) was
found to be weak [Annavaram et al. 2008], thus different sensor signals are modeled
as uncorrelated[2]; this assumption is consistent with our observation that certain
sensors better discriminate between some subsets of activities than others and is
borne out by the numerical results. Furthermore, this assumption allows us to
develop the low-complexity implementation (see Section 5.3), which is critical to
on-phone implementation. Additionally, our preliminary simulations indicated that
assuming independence between features does not impact performance. We now
propose the following signal model for the decoded and processed samples, or fea-
tures, received by the fusion center:

$$y_l - \mu_{jk} = \phi(y_{l-1} - \mu_{jk}) + n_l \ , \quad l = 1, \ldots, N_k, \tag{11}$$

where ϕ is the AR(1) parameter, μ_{jk} is the mean of the feature for that the par-
ticular activity, and n_l is the zero-mean noise with variance σ_{jk}^2. Since the features

---

[2]Due to our jointly Gaussian feature model, lack of correlation implies statistical independence.

are modeled as Gaussian, and given that the AR(1) model is *linear*, the M-ary hypothesis test introduced in Section 4.1 is equivalent to the generalized Gaussian problem, using the model in (11).

We denote $\mathbf{m}_i$ and $\boldsymbol{\Sigma}_i, i = 1, 2, \ldots, M$ to be the mean vectors and covariance matrices of the observations under the each of the M hypotheses, respectively. For completeness, we recall the density of the multivariate Gaussian random variable, $\mathbf{x} = [x_1, \cdots, x_N]$ is given by:

$$f_X(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right) \doteq \mathcal{N}(\mathbf{x}; \mathbf{m}, \boldsymbol{\Sigma}), \quad (12)$$

where $\boldsymbol{\mu}$ is the mean vector and $\boldsymbol{\Sigma}$ is the covariance matrix. For each of the K sets of features $A_1, A_2, \ldots, A_K$ from the sensors, the mean vector and covariance matrix for the observations for hypothesis H_i for $i = 1, \cdots, M$ are of the form

$$\mathbf{m}_i = \begin{bmatrix} \mu_{iA_1} \\ \mu_{iA_2} \\ \vdots \\ \mu_{iA_K} \end{bmatrix} \text{ and } \boldsymbol{\Sigma}_i = \begin{bmatrix} \boldsymbol{\Sigma}_i(A_1) & 0 & \cdots & 0 \\ 0 & \boldsymbol{\Sigma}_i(A_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{\Sigma}_i(A_K) \end{bmatrix}, \quad (13)$$

respectively, where μ_{iA_k} and $\boldsymbol{\Sigma}_i(A_k)$ are the single-feature statistics.

Given the signal model in (11), and incorporating zero-mean channel and measurement noise with variance σ_z^2, the covariance matrix for a particular feature A_k can be expressed as

$$\boldsymbol{\Sigma}(A_k) = \frac{\sigma_{A_k}^2}{1 - \varphi^2}\mathbf{T} + \sigma_z^2\mathbf{I}, \quad (14)$$

where $\mathbf{T}$ is a Toeplitz matrix whose first row/column is $\begin{bmatrix} 1 & \phi & \phi^2 & \ldots & \phi^{N_k - 1} \end{bmatrix}$, and $\mathbf{I}$ is the $N_k \times N_k$ identity matrix. This results in the covariance matrices $\boldsymbol{\Sigma}_j, j = 1, \ldots, M$ being block-Toeplitz matrices. To derive a vector optimization that circumvents an exhaustive search, we may approximate the Toeplitz covariance matrices with their associated circulant covariance matrices[3] given by

$$\boldsymbol{\Sigma}(A_k) = \frac{\sigma_{A_k}^2}{1 - \varphi^2}\mathbf{C} + \sigma_z^2\mathbf{I}, \quad (15)$$

where the matrix $\mathbf{C}$ is a circulant matrix whose first row is identical to that of $\mathbf{T}$.

5.2 Misclassification Metric Derivation

We derive a closed-form approximation for the probability of error in the multi-hypothesis case via a union bound incorporating the Bhattacharyya coefficients between pairs of hypotheses. A result by Lianiotis [Lianiotis 1969] provides an upper bound on the probability of error, given as

$$P(\epsilon|\mathcal{S}) \leq \sum_{i<j} \left(P_{i|\mathcal{S}} P_{j|\mathcal{S}}\right)^{1/2} e^{-\rho_{ij}|\mathcal{S}} = P_{\text{ub}}(\epsilon), \quad (16)$$

---

[3]We note that the inverse of the Toeplitz covariance matrix in (14) converges to the inverse of the circulant covariance matrix in (15) in the weak sense. Sun et al [Sun et al. 2003] have derived that a sufficient condition for weak convergence is that the strong norm of the inverse matrices be uniformly bounded, which is the case for the matrix forms in (14) and (15) for $0 < \phi < 1$.
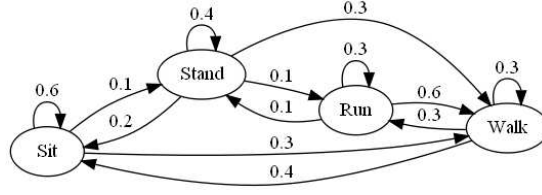
Fig. 11. Finite-state machine with nominal prior probabilities used for optimal sampling algorithm.

where $\mathcal{S}$ is the current state, $P_{i|\mathcal{S}}$ and $P_{j|\mathcal{S}}$ are the *a priori* probabilities for hypotheses H_i and H_j from the current state, and $\rho_{ij|\mathcal{S}}$ is the state-dependent Bhattacharyya coefficient. Thus, the optimization problem considered can be stated as

$$\min_{\mathbf{N}} \quad P_{\text{ub}}(\epsilon) \quad \text{subject to} \quad \sum_{k=1}^{K} N_k = N, \quad N_k \geq 0 \quad \forall k, \tag{17}$$

where $\mathbf{N} = (N_1, N_2, \ldots, N_K)$ is the allocation of samples amongst the K sensors. In the case of the multivariate Gaussian, if $f_i(\mathbf{x}) = \mathcal{N}(\mathbf{x}; \mathbf{m}_i, \boldsymbol{\Sigma}_i)$, the Bhattacharyya coefficient is given by:

$$\rho_{ij|\mathcal{S}} = \frac{1}{8} (\mathbf{m}_i - \mathbf{m}_j)^T \boldsymbol{\Sigma}_h^{-1} (\mathbf{m}_i - \mathbf{m}_j) + \frac{1}{2} \log \frac{\det \boldsymbol{\Sigma}_h}{\sqrt{\det \boldsymbol{\Sigma}_i \cdot \det \boldsymbol{\Sigma}_j}}, \tag{18}$$

where $|\boldsymbol{\Sigma}| = \det \boldsymbol{\Sigma}$, and $2\boldsymbol{\Sigma}_h = \boldsymbol{\Sigma}_i + \boldsymbol{\Sigma}_j$. We first note that $\rho_{ij|\mathcal{S}}$ is a function of the means and covariances associated with the hypotheses H_i and H_j, and is a measure of the confusability of the two hypotheses. The upper bound $P_{\text{ub}}(\epsilon)$, defined in (16), incorporates classification errors that can be made between all **possible** pairs of hypotheses. Note that not all pairs are considered in the case of evolving activities undertaken by a subject. Figure 11 shows an exemplary finite state machine (FSM) that specifies the probabilities of transitioning from one activity-state to another for a subset of activities, and is used to define the evolution of physical activities for the user. For example, as seen in Figure 11, if the subject is "Standing," then all three possible pairs of hypotheses are considered. In comparison, if the subject is "Sitting," the Sit→Run transition probability is 0, and so only two pairs of hypotheses are considered. Therefore, the optimal allocation depends on both the current state and the possible next state transition probabilities.

5.3 Low-Complexity Implementation

Given the block-diagonal structure of the covariance matrix in (13), we first decompose quadratic and determinant terms for each feature, and refer the reader to [Thatte et al. 2009] for details. This decomposition implies that the computation of each of the terms for an individual feature A_k is sufficient to evaluate the upper bound on the probability of error specified in (16). The structure of the covariance matrix in (13) is block-Toeplitz, or approximated as block-circulant, where the k-th block is of size $N_k \times N_k$. For every unique allocation of samples amongst sensors, the structure of the covariance matrix is distinct, and thus a combinatorial search over all possible partitions of the total number of samples is required to find the

optimal allocation of samples to minimize the probability of error.

To evaluate the determinant term in (18), we use the Toeplitz structure in (14), and rewrite the covariance matrix as follows [Ipsen and Lee 2006]:

$$\boldsymbol{\Sigma}(A_k) = \boldsymbol{\Sigma}_{\mathrm{D}}(A_k) + \boldsymbol{\Sigma}_{\mathrm{off}}(A_k). \tag{19}$$

The determinant is subsequently computed using the identity:

$$\det \boldsymbol{\Sigma} = \det \boldsymbol{\Sigma}_{\mathrm{D}} \cdot \det \left(\mathbf{I} + \boldsymbol{\Sigma}_{\mathrm{D}}^{-1} \boldsymbol{\Sigma}_{\mathrm{off}} \right), \tag{20}$$

the approximation of $\log(\mathbf{I} + \mathbf{A})$ for a nilpotent matrix $\mathbf{A}$, and geometric sum identities. To simplify quadratic term in (18), we use the circulant approximation (15) of the inverse of the covariance matrix, and a simple result from [Wilansky 1951] which states that if the sum of elements in each row of a square matrix is c, then the sum of elements in each row of the inverse is $1/c$. Completing these computations, for a single block of the covariance matrix $\boldsymbol{\Sigma}(A_k)$, simplifies the quadratic term as

$$(m_{jA_k} - m_{iA_k})^2 N_k \left[\frac{1}{2} \frac{\sigma_{iA_k}^2 + \sigma_{jA_k}^2}{1 - \phi^2} \frac{1 - \phi^{N_k}}{1 - \phi} + \sigma_z^2 \right]^{-1}. \tag{21}$$

We refer the reader to [Thatte et al. 2009] for the details, and conclude that the combinatorial search over K integer partitions of N samples is converted to a continuous-valued vector optimization, which is minimized with significantly lower complexity.

6. PERFORMANCE ANALYSIS

In this section, we first overview the data collection effort, protocols and test subject details. We then present numerical results for the physical activity recognition and optimal sampling algorithms.

6.1 Data Collection

Traditional health-monitoring experiments (see *e.g.* [Spruijt-Metz et al. 2008]) in preventive health scenarios require test subjects to first wear an accelerometer (or some other sensor) for an extended period of time. At the end of the pre-determined phase, the data is collected and processed before the experiment is repeated, if required. There is no explicit training period, and standard cut-points (decision regions for single features), which may not be applicable to all demographics, are employed to determine the level of user physical activity. KNOWME adopts a different methodology since we use personalized training periods, and offer real-time feedback and user visualization.

The main motivation for using a personalized training phase is activity detection performance: there are variations amongst the biometric signals of different individuals. Thus, using models developed on one set of individuals and then applied to another set of individuals would result in a decrease in detection performance and potentially reduced energy efficiency. Most typical medical interventions begin with intake interviews and baseline data collection. Thus, a brief, personalized training is easily realizable during this initial intervention phase.
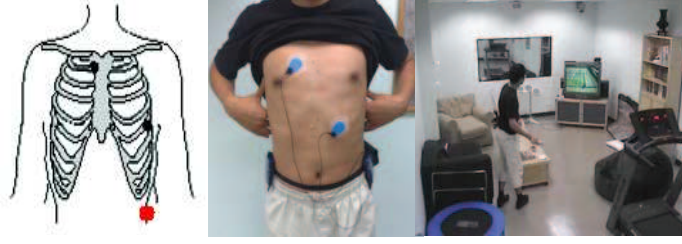
Fig. 12. Placement of the electrodes (black circles) and accelerometer (red square) and data collection environment.

Table II. Performance of SVM system based on temporal ECG features (% correct)

ECG	1 PCA	2 HPE	3 HR+NM	2+3	1+2+3
10 beats	46.9	51.7	43.3	56.7	60.8
20 seconds	49.4	54.4	44.0	60.0	64.2

Data collection was conducted using two Alive heart rate monitors [AliveTechnologies 2008] and the Nokia N95 mobile device. A single lead ECG signal is collected with electrodes on the chest; one heart rate monitor with a built-in accelerometer is placed on the left hip to record the ECG and accelerometer signal, and the second heart rate monitor with built-in accelerometer is placed on the right hip to record only the accelerometer signal. The placement of electrodes and the accelerometer sensor is shown in Figure 12. For each of three sessions, test subjects were required to wear the sensors and perform 9 specific categories of PA, following a predetermined protocol [Spruijt-Metz et al. 2009]. The sequence of the 9 PAs to be performed was the same for each session for every test subject, unless the test subjects were physically unable to perform the required sequence. Data from 5 subjects (2 male, 3 female; ages range from 13-30; body mass index range from 23.0-34.3) is reported, wherein each subject performed 4 sessions on different days and times. The data reflects the variability of electrodes position and a variety of environmental and physiological factors, and we recall that the state-detection algorithms developed in Section 4 account for this inter-session variability.

6.2 Performance of Activity-Detection Classification

For each subject, training was on 3 sessions' data and testing was on the remaining session. Training and testing data are from different days/times, and training data and testing data are rotated 4 times (for cross validation). The performance reported is based on the average of all the subjects and all the rotation tests. Decisions from the SVM classifier are made every 20 seconds; the number of HPE coefficients, the PCA eigenvector dimension and the normalized heartbeat sample length D are empirically chosen to be 60, 40 and 201, respectively.

Table II shows the results of the ECG temporal features based SVM system. Compared to the conventional PCA method [Pawar et al. 2007], the proposed HPE coefficients together with heart rate (HR) mean/variance and noise measurement (NM) features achieved nearly 10% improvement in accuracy. Furthermore, the

Table III.   Performance of SVM system based on temporal Accelerometer features

Window length (seconds)	1.68	3.36	6.72	13.44
Accuracy (% correct)	83.56	84.05	84.85	84.36

PCA approach requires projection and distance calculations for each heartbeat which is computationally expensive. Also, the pre-trained activity mean might be different from the testing condition due to the session variability which can decrease the system performance. The proposed method, based on the linear kernel of SVM and the support vector compression method, is more efficient in terms of both computational cost and model size. Finally, fusing PCA, HPE, HR and NM features together achieves a further 4% PA recognition accuracy improvement. This is because PCA and HPE model the MAN and CAM part of the normalized ECG waveform, respectively, while HR and NM measure the heart rate and inter-beats noise level, and this information is complementary. From Table II, we can also see that the single lead ECG signal has more activity discrimination information than provided by just the heart rate, but the performance is still relatively low compared with accelerometer based methods. Therefore, fusing the information from both modalities is necessary. We find that fusing the accelerometer and ECG modalities yields a detection accuracy of 85.5%, which is a slight improvement ($< 1\%$) from using using only the temporal accelerometer features.

Table III shows the performance of the proposed SVM system based on the GLDS kernel and the temporal accelerometer features listed in Box 1. The accuracy is not sensitive to the processing window size and a 6.72-second window size (corresponding to 504 samples) gives the best performance, which matches the experiment results in [Bao and Intille 2004; Krause et al. 2005; Ravi et al. 2005].

6.3 Performance of Optimal Sampling

As described in Section 5.1, simple Gaussian model based detectors are adopted to develop a low-complexity implementation of the optimal sampling protocol. This algorithm is based entirely on a parametric model (Gaussian), in contrast to data-driven/non-parametric approach used for state-detection using SVM classifiers in Section 4.3. For clarity of exposition, we focus on using two features and discriminating between four hypotheses (Sitting, Standing, Walking and Running). We assume that the mobile device allocates N_1 ACC variance and N_2 ECG period samples, respectively. We recall that the ACC variance feature is the average (over the three axes) variance of the accelerometer signal in a fixed 6.72-second window, and the ECG period feature is the average heart-rate over 10 heartbeats. Each of these features is time-varying due to a number of factors, including the current user state, and the AR(1) model, described in Section 5.1, attempts to capture this temporal correlation. We further assume that the finite-state machine, shown in Figure 11, determines transitions between activities. These are the state-dependent *a priori* probabilities specified in the probability of error upper bound in (16), and are exemplary values, and not based on real data. This is an avenue for future research.

The underlying Gaussian distributions associated with each of the hypotheses for these two features for a particular test subject are shown in Figure 13. Notice that,
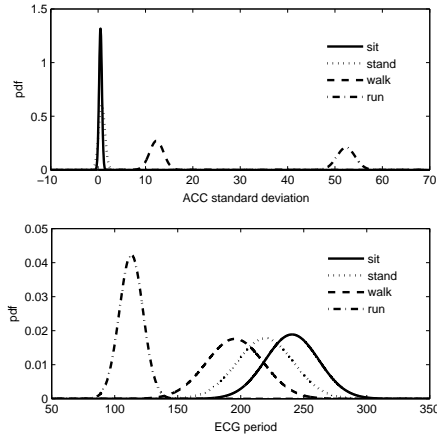
Fig. 13. Gaussian distributions associated with the four activities for the ACC variance and ECG period features.
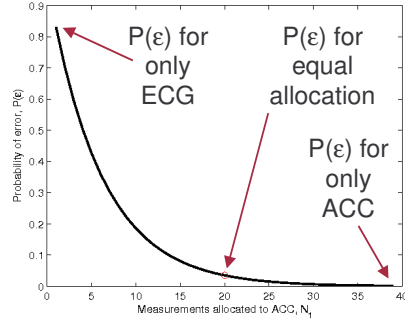

Fig. 14. Optimal allocation of samples between ACC and ECG sensors; allocate all samples to ACC to achieve minimum $P(\epsilon)$.

in Figure 13, that ACC variance is not a good discriminator between the Sit and Stand activities, but the ECG period is. The upper bound on the probability of error is plotted as a function of all possible allocations between these two features in Figure 14. Assuming we are initially in the "Run" hypothesis, we find that the optimal allocation corresponds to allocating all available samples to the ACC, which achieves the minimum probability of misclassification. For this scenario, we find that the low-complexity implementation yields a solution which corresponds to optimal solution obtained via the exhaustive search.

Since we have used the high-dimensional SVM for accurate physical activity recognition, and simpler Gaussian model based detectors to derive the optimal allocation, we now compare the energy-savings achieved due to each of these methods. Different methods are used to compute the resulting energy-savings for the high-dimensional SVM classifier and a Gaussian detector that uses two features. For the Gaussian detector, we denote the bounds on the error probability corresponding to the equal and optimal allocations of N samples as $P_{eq}(N)$ and $P_{opt}(N)$, respectively. In order to compute the energy-savings, we now increase the number of available samples to N', and find the minimum $N' > N$ such that $P_{eq}(N') = P_{opt}(N)$. For our collected data, $N = 40$ and $N' = 63$, this translates to an energy savings of approximately 38%. Note that these numbers are averaged across all subjects.

The energy-savings when eight activities are considered (the 9 PAs enumerated in Section 4.1 with the exception of playing Wii tennis) reduces to approximately 20% (from 38% when only four activities are considered), averaged across all subjects. The algorithm optimally allocates, on average 85% of the samples to the ACC and the remaining 15% to the ECG. Furthermore, we find that the optimal allocation is not significantly affected by the initial state of the subject, given the exemplary a priori probabilities.

In order to compute the energy-savings for the SVM classifier, we first compute

the optimal allocation of samples for a particular state using the simpler Gaussian models. Then the testing data that corresponds to this optimal allocation is replayed through the SVM classifier, and we find that 50% energy-savings are achieved. For the subset of activities considered, the detection accuracy is maintained while using only the appropriate one of the two sensors, *i.e.* when the "Sit" state is assumed, only the ECG features are employed, whereas if the current state is "Walk," the ACC features in Box 1 are used.

Although the optimal sample allocation is determined for the simpler, Gaussian case, we see that the resultant allocation yields significant energy savings for the SVM classifier without a loss in performance.

7. CONCLUSIONS AND FUTURE WORK

We have developed a prototype implementation of the KNOWME Network, which is an energy-efficient health-monitoring system for physical activity detection. Our goal is to extend this system to enable true energy-expenditure estimation, with a focus on pediatric obesity intervention. The current implementation of our system uses the Nokia N95 mobile device along with off-the-shelf sensors that communicate via the Bluetooth protocol. Accurate state detection is achieved using high-dimensional SVM classifiers with a GLDS kernel and a novel feature set based on accurate modeling of the ECG signal via Hermite polynomials and a principal component analysis. Bluetooth communication consumes a relatively large amount of power, so an optimal sampling protocol is developed to create a more energy-aware sensing system and to counter the relatively high packet-loss that results when several Bluetooth sensors are connected to the mobile device at one time.

Our hardware choices enable "plug-and-play" capabilities for multimodal sensors, and the software allows for efficient memory management and uninterrupted phone operation. In addition, the specific implementation chosen makes it easy to port the system onto a different mobile platform in the future. The accurate physical activity recognition algorithm accounts for inter-session variability, and employs novel features extracted from the ECG biometric signal and high-dimensional SVM classifiers. Energy efficiency at the mobile device fusion center is achieved via performance optimization, in that the optimal sampling algorithm minimizes the probability of misclassification. We find that significant energy-savings are achieved by the SVM classifier, even when the optimal allocation is computed using simpler Gaussian models and a minimal subset of the features.

We are currently working on enhancing the KNOWME Network on three fronts: first, we are developing algorithms to exploit both frequency-domain and temporal features for physical activity recognition, which should result in a more accurate classifier. We are also developing alternative models that will allow the seamless integration of the optimal sampling protocol and high-dimensional classifiers. Second, we are implementing our algorithms on the mobile device, and investigating the minimization of power consumption due to computation as well as Bluetooth communication. We are also implementing real-time user feedback and introducing location- and context-awareness via GPS and other custom-built sensors. Finally, we are working towards long-term continuous deployment of the KNOWME Network and an expanded data collection effort with targeted youth test subjects in

order to validate the practicality of the system as a modality for intervention in pediatric obesity.

We have recently deployed KNOWME Networks in the field in 12 overweight Hispanic youth aged 14.8 ± 1.9 years (12-18) with an initialized personalized phase and 100% compliance, validating the feasibility of personalized training for a small group of test subjects. The integrated optimal sampling protocols and high-dimensional classifers are currently being implemented on the mobile device for an upcoming field trial. This field trail will use KNOWME's ability to detect cumulative time spend in sedentary behavior in real-time to trigger tailored text messages that we have developed in focus group and individual interviews with overweight Hispanic youth that encourage physical activity. KNOWME will sense and transmit all responses to messaging, evaluate best timing for messaging and the types of messages that elicit the longest and most intense bouts of physical activity.

REFERENCES

ALIVETECHNOLOGIES. 2008. Wireless health monitors from alive technology, www.alivetec.com.

ANNAVARAM, M., MEDVIDOVIC, N., MITRA, U., NARAYANAN, S., SUKHATME, G., MENG, Z., QIU, S., KUMAR, R., THATTE, G., AND SPRUIJT-METZ, D. 2008. Multimodal sensing for pediatric obesity applications. In *Proceedings of UrbanSense08*. Raleigh, NC, 21–25.

ARFKEN, G., WEBER, H., AND WEBER, H. 1985. *Mathematical methods for physicists*. Academic Press, New York.

BAO, L. AND INTILLE, S. 2004. Activity recognition from user-annotated acceleration data. *Lecture Notes in Computer Science 3001*, 1–17.

BENBASAT, A. AND PARADISO, J. 2007. A framework for the automated generation of power-efficient classifiers for embedded sensor nodes. In *Proceedings of SenSys*. Sydney, Australia, 219–232.

BONOMI, A., GORIS, A., YIN, B., AND WESTERTERP, K. 2009. Detection of type, duration, and intensity of physical activity using an accelerometer. *Medicine and Science in Sports and Exercise 41(9)*, 1770–1777.

CABALLERO, B. 2007. The global epidemic of obesity: an overview. *Epidemiologic Reviews 29(1)*, 1–5.

CAMPBELL, W., CAMPBELL, J., REYNOLDS, D., SINGER, E., AND TORRES-CARRASQUILLO, P. 2006. Support vector machines for speaker and language recognition. *Computer Speech & Language 20,* 2-3, 210–229.

CHANG, C.-C. AND LIN, C.-J. 2001. *LIBSVM: a library for support vector machines*. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

CHEN, K. AND BASSETT, D. 2005. The technology of accelerometry-based activity monitors: current and future. *Medicine and Science in Sports and Exercise 37(11 Suppl)*, S490–500.

CHEN, K. AND SUN, M. 1997. Improving energy expenditure estimation by using a triaxial accelerometer. *Journal of Applied Physiology 83(6)*, 2112–2122.

CLIFFORD, G., AZUAJE, F., AND MCSHARRY, P. 2006. *Advanced methods and tools for ECG data analysis*. Artech House.

COLLINS, L., MURPHY, S., AND BIERMAN, K. 2004. A conceptual framework for adaptive preventive interventions. *Prevention Science 5(3)*, 185–196.

COLLINS, R., KASHDAN, T., AND GOLLNISCH, G. 2003. The feasibility of using cellular phones to collect ecological momentary assessment data: application to alcohol consumption. *Experimental and Clinical Psychopharmacology 11*, 73–78.

CRISTIANINI, N. AND SHAWE-TAYLOR, J. 2000. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, Cambridge, UK.

CROUTER, S., CLOWERS, K., AND JR., D. B. 2006. A novel method for using accelerometer data to predict energy expenditure. *Journal of Applied Physiology 100(4)*, 1324–1331.

Ermes, M., Parkka, J., Mantyjarvi, J., and Korhonen, I. 2008. Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. *IEEE Transactions on Information Technology in Biomedicine 12,* 1, 20–26.

Freedman, M., Lester, K., Milby, C. M. J., and Schumacher, J. 2006. Cell phones for ecological momentary assessment with cocaine-addicted homeless patients in treatment. *Journal of Substance Abuse Treatment 30*, 105–111.

Godfrey, A., Conway, R., Meagher, D., and ÓLaighin, G. 2008. Direct measurement of human movement by accelerometry. *Medical Engineering and Physics 30,* 10, 1364–1386.

Hoh, B., Gruteserand, M., Herring, R., Ban, J., Work, D., Herrera, J., Bayen, A., Annavaram, M., and Jacobson, Q. 2008. Virtual trip lines for distributed privacy-preserving traffic monitoring. In *Proceedings of MobiSys*. Breckenridge, CO.

Huynh, D. 2008. Human Activity Recognition with Wearable Sensors. Ph.D. thesis, Darmstadt University of Technology.

Huynh, T., Blanke, U., and Schiele, B. 2007. Scalable recognition of daily activities with wearable sensors. *Lecture Notes in Computer Science 4718*, 50.

Ipsen, I. and Lee, D. 2006. Determinant approximations. *Numerical Linear Algebra with Applications (Under review)*.

Jatoba, L., Grossmann, U., et al. 2008. Context-aware mobile health monitoring: Evaluation of different pattern recognition methods for classification of physical activity. In *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*. Vancouver, Canada.

Kang, S., Lee, J., Jang, H., Lee, H., Lee, Y., Park, S., Park, T., and Song, J. 2008. Seemon: scalable and energy-efficient context monitoring framework from sensor-rich mobile environments. In *Proceedings of MobiSys*. Breckenridge, CO.

Krause, A., Siewiorek, D., Smailagic, A., and Farringdon, J. 2005. Unsupervised, dynamic identification of physiological and activity context in wearable computing. In *IEEE International Symposium on Wearable Computers*. White Plains, NY, 88–97.

Lester, J., Choudhury, T., and Borriello, G. 2006. A practical approach to recognizing physical activities. *Lecture Notes in Computer Science 3968*, 1–16.

Levine, J. A., Schleusner, S. J., and Jensen, M. D. 2000. Energy expenditure of nonexercise activity. *American Journal of Clinical Nutrition 72(6)*, 1451–1454.

Lianiotis, D. 1969. A class of upper bounds on probability of error for multihypotheses pattern recognition. *IEEE Transactions on Information Theory 15(6)*, 730–731.

Linh, T., Osowski, S., and Stodolski, M. 2003. On-line heart beat recognition using Hermite polynomials and neuro-fuzzy network. *IEEE Transactions on Instrumentation and Measurement 52,* 4, 1224–1231.

Liu, Y., Veeravalli, B., and Viswanathan, S. 2007. Critical-path based low-energy scheduling algorithms for body area network systems. In *Proceedings of the 13th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications*. Daegu, Korea.

Maguire, D. and Frisby, R. 2009. Comparison of feature classification algorithms for activity recognition based on accelerometer and heart rate data. In *Proceedings of the 9th Information Technology and Telecommunication Conference*. Dublin, Ireland.

Maurer, U., Rowe, A., Smailagic, A., and Siewiorek, D. 2006. Location and Activity Recognition Using eWatch: A Wearable Sensor Platform. *Lecture Notes in Comp Sci 3864*, 86.

Miluzzo, E., Lane, N. D., Fodor, K., Peterson, R., Lu, H., Musolesi, M., Eisenman, S. B., Zheng, X., and Campbell, A. T. 2008. Sensing meets mobile social networks: The design, implementation and evaluation of the cenceme application. In *Proceedings of SenSys'08*. Raleigh, NC, 337–350.

Mohammadi, G., Shoushtari, P., et al. 2006. Person identification by using ar model for eeg signals. *Proc. of World Academy of Sci., Engr. and Tech. 11*, 5.

Ogden, C., Carroll, M., and Flegal, K. 2008. High body mass index for age among us children and adolescents, 2003-2006. *Journal Of the American Medical Association 299(20)*, 2401–2405.

Pagac, P. A. G. A. C. 2008. Physical activity guidelines advisory committee report. Tech. rep., Department of Health and Human Services, Washington DC.

PARKKA, J., ERMES, M., KORPIPAA, P., MANTYJARVI, J., PELTOLA, J., KORHONEN, I., TECHNOL, V., AND TAMPERE, F. 2006. Activity classification using realistic data from wearable sensors. *IEEE Transactions on Information Technology in Biomedicine 10,* 1, 119–128.

PATRICK, K., GRISWOLD, W. G., RAAB, F., AND INTILLE, S. S. 2008. Health and the mobile phone. *American Journal of Preventive Medicine 35(2)*, 177–181.

PAWAR, T., CHAUDHURI, S., AND DUTTAGUPTA, S. 2007. Body movement activity recognition for ambulatory cardiac monitoring. *IEEE Transactions on Biomedical Engineering 54,* 5, 874–882.

PFEIFFER, K., MCIVER, K., DOWDA, M., ALMEIDA, M., AND PATE, R. 2006. Validation and calibration of the actical accelerometer in preschool children. *Medicine and Science in Sports and Exercise 38(1)*, 152–157.

RAVI, N., DANDEKAR, N., MYSORE, P., AND LITTMAN, M. 2005. Activity recognition from accelerometer data. In *Proceedings of the National Conference on Artificial Intelligence.* Vol. 20(3). Pittsburgh, PA, 1541–1547.

RODGERS, A., CORBETT, T., BRAMLEY, D., RIDDELL, T., WILLS, M., LIN, R.-B., AND JONES, M. 2005. Do u smoke after txt? results of a randomised trial of smoking cessation using mobile phone text messaging. *Tobacco Control 14,* 255–261.

ROSS, A., NANDAKUMAR, K., AND JAIN, A. 2006. *Handbook of multibiometrics.* Springer.

RYDER, J., LONGSTAFF, B., REDDY, S., AND ESTRIN, D. 2009. Ambulation: A tool for monitoring mobility patterns over time using mobile phones. In *Proceedings of the IEEE International Conference on Social Computing: Workshop on Social Computing with Mobile Phones and Sensors: Modeling, Sensing and Sharing.* Vancouver, Canada.

SHIH, E., BAHL, P., AND SINCLAIR, M. 2002. Wake on wireless: an event driven energy saving strategy for battery operated devices. In *Proceedings of MobiCom.* Atlanta, GA.

SHIN, M., TSANG, P. P., KOTZ, D., AND CORNELIUS, C. 2009. Deamon: Energy-efficient sensor monitoring. In *Proceedings of the 6th IEEE Communications Society Conference on Sensor, Mesh, and Ad Hoc Communications and Networks (SECON).* Roma, Italy.

SIEWIOREK, D., SMAILAGIC, A., FURUKAWA, J., MORAVEJI, N., REIGER, K., AND SHAFFER, J. 2003. Sensay: a context-aware mobile phone. In *Proceedings of ISWC.* White Plains, NY.

SIRARD, J. AND PATE, R. 2001. Physical activity assessment in children and adolescents. *Sports Medicine 31(6)*, 439–454.

SPRUIJT-METZ, D., BERRIGAN, D., KELLY, L., MCCONNELL, R., DUEKER, D., LINDSEY, G., ATIENZA, A., MICHEL, S., IRWIN, M., WOLCH, J., JERRETT, M., TATALOVICH, Z., AND REDLINE, S. 2008. *Measures of Physical Activity and Exercise.* Vol. Handbook of Assessment Methods for Obesity and Eating Behaviors, Related Problems and Weight: Measures, Theory and Research. Sage Publications Inc., Thousand Oaks, CA, Chapter 6.

SPRUIJT-METZ, D., LI, M., THATTE, G., SUKHATME, G., ANNAVARAM, M., GHOSH, S., ROZGIC, V., MITRA, U., MEDVIDOVIC, N., BELCHER, B., AND NARAYANAN, S. 2009. Differentiating physical activity modalities in youth using heartbeat waveform shape and differences between adjacent waveforms. In *International Conference on Diet and Activity Methods.* Washington DC.

STURM, R. 2002. The effects of obesity, smoking, and drinking on medical problems and costs. *Health Affairs 21(2)*, 245–253.

SUN, D., SCHMIDT, G., AND TEO-KOH, S. 2008. Validation of the rt3 accelerometer for measuring physical activity of children in simulated free-living conditions. *Pediatric Exercise Science 20(2)*, 181–197.

SUN, F.-W., JIANG, Y., AND BARAS, J. 2003. On the convergence of the inverses of toeplitz matrices and its applications. *IEEE Trans IT 49(1)*, 180–190.

THATTE, G., ROZGIC, V., LI, M., GHOSH, S., MITRA, U., NARAYANAN, S., ANNAVARAM, M., AND SPRUIJT-METZ, D. 2009. Optimal allocation of time-resources for multihypothesis activity-level detection. In *Proceedings of DCOSS.* Marina Del Rey, CA.

WILANSKY, A. 1951. The row-sums of the inverse matrix. *The American Mathematical Monthly 58(9)*, 614–615.

YAN, L., ZHONG, L., AND JHA, N. 2007. Energy comparison and optimization of wireless body-area network technologies. In *Proceedings of the ICST 2nd International Conference on Body Area Networks.* Florence, Italy.

ZHANG, J. AND CHU, F. 2005. Real-time modeling and prediction of physiological hand tremor. In *Proceedings of ICASSP*. Philadelphia, PA.