

Optimal Allocation of Time-Resources for Multihypothesis Activity-Level Detection

Gautam Thatte^{1*}, Viktor Rozgic¹, Ming Li¹, Sabyasachi Ghosh¹, Urbashi Mitra¹, Shri Narayanan¹, Murali Annavaram¹, and Donna Spruijt-Metz²

¹Ming Hseih Department of Electrical Engineering

²Keck School of Medicine

University of Southern California, Los Angeles, CA

{thatte, rozgic, mingli, sabyasag, ubli, annavara, dmetz}@usc.edu

shri@sipi.usc.edu

Abstract. The optimal allocation of samples for activity-level detection in a wireless body area network for health-monitoring applications is considered. A wireless body area network with heterogeneous sensors is deployed in a simple star topology with the fusion center receiving biometric samples from each of the sensors. The number of samples collected from each of the sensors is optimized to minimize the probability of misclassification between multiple hypotheses at the fusion center. Using experimental data from our pilot study, we find equally allocating samples amongst sensors is normally suboptimal. A lower probability of error can be achieved by allocating a greater fraction of the samples to sensors which can better discriminate between certain activity-levels. As the number of samples is an integer, prior work employed an exhaustive search to determine the optimal allocation of integer samples. However, such a search is computationally expensive. To this end, an alternate continuous-valued vector optimization is derived which yields approximately optimal allocations which can be found with significantly lower complexity.

1 Introduction

Wearable health monitoring systems coupled with wireless communications are the bedrock of an emerging class of sensor networks: wireless body area networks (WBANs). Such networks have myriad applications, including diet monitoring [16], detection of activity [3, 2], and health crisis support [6]. This paper focuses on the KNOWME network, which is targeted to applications in pediatric obesity, a developing health crisis both within the US and internationally. To understand, treat, and prevent childhood obesity, it is necessary to develop a multimodal system to track an individual's level of stress, physical activity, and blood glucose, as well as other vital signs, simultaneously. Such data must also be anchorable to context, such as time of day and geographical location. The

* This research is supported by the National Center on Minority Health and Health Disparities (NCMHD) (supplement to P60 MD002254) and Qualcomm.



Fig. 1. The Nokia N95 cellphone fusion center (A), and the Alive Technologies oximeter sensor (B) and ECG sensor (C).

KNOWME network is a first step in the development of a system that could achieve these targets.

A crucial component of the KNOWME network is the unified design and evaluation of multimodal sensing and interpretation, which allows for automatic recognition, prediction, and reasoning regarding physical activity and socio-cognitive behavior states. This accomplishes the current goals of observational research in obesity and metabolic health regarding physical activity and energy expenditure (traditionally carried out through careful expert human data coding), as well as enabling new methods of analysis previously unavailable, such as incorporating data on the user’s emotional state.

The KNOWME network utilizes heterogeneous sensors, which send their measurements to a Nokia N95 cellphone via Bluetooth, as shown in Figure 1. The Bluetooth standard for data communication uses a “serve as available” protocol, in which all samples taken by each sensor are collected by the fusion center. Though this is beneficial from the standpoint of signal processing and activity-level detection, it requires undesirably high energy consumption: with a fully charged battery, the Nokia N95 cellphone can perform over ten hours of telephone conversation, but if the GPS receiver is activated, the battery is drained in under six hours [20]. A similar decrease in battery life occurs if Bluetooth is left on continuously. One of the aims of this paper is to devise a scheme that reduces the Bluetooth communication, thus resulting in energy savings.

Our pilot study [1] suggested that data provided by certain types of sensors were more informative in distinguishing between certain activities than other types. For example, the electrocardiograph sensor was a better discriminator when the subject was lying down, while data from the accelerometer was more pertinent to distinguishing between higher-level activities. In the present work, we exploit the advantages gained by focusing on particular sensors when selecting from a specific set of hypothesized activity states, thus providing a more energy-efficient detection mechanism.

The goal of the present work is to characterize the optimal allocation of samples for heterogeneous sensors in order to minimize the probability of misclassification at the fusion center. Making optimal time-resource allocation a priority leads us to consider that sensors whose measurements are not currently

being utilized can turn off their Bluetooth, resulting in a health-monitoring application that is more energy-efficient than previous models. To achieve this goal, we consider the centralized approach adopted in [1], wherein detection is performed at the Nokia N95 fusion center, and present numerical results for the M -ary hypothesis testing problem with multiple sensors.

Thus, the contribution of our work is describing the optimal allocation of samples amongst heterogeneous sensors in a WBAN for activity-level detection. Specifically, we derive a lower complexity (see Section 5.2) continuous-valued vector optimization to minimize the probability of misclassification in the multi-hypothesis case. We are currently developing an energy-efficient algorithm using this optimal allocation of samples.

The remainder of the paper is organized as follows: prior relevant work on activity-level detection and energy-efficient algorithms in WBANs, and their relationship to our work is presented in Section 2. An overview of our activity-level detection system is described in Section 3. In this work, we focus on developing the time-resource allocation algorithm. Specifically, in Section 4, we describe the signal model used to develop our optimal allocation, outline the framework for minimizing the probability of misclassification, and derive a lower complexity continuous-valued optimization problem. Numerical results based on experimental data are presented in Section 5. Finally, we draw conclusions and discuss our future work direction and extensions to the optimal time-resource allocation problem in Section 6.

2 Related Work

In recent years, there have been several projects investigating activity-level detection in a variety of frameworks. Much of the work appears to center on accelerometer data alone (*e.g.* [3, 8, 10]) with some systems employing several accelerometer packages and/or gyroscopes. On the other hand, multi-sensor systems have also been implemented and deployed for activity-level detection, context-aware sensing and specific health-monitoring applications. For example, the work of Gao et al [6] is tailored for emergency response and triage situations, while Dabiri et al [5] have developed a lightweight embedded system that is primarily used for patient monitoring. The system developed by Jovanov et al [9] is used in the assistance of physical rehabilitation, and Consolvo et al's UbiFit system [4] is designed to promote physical activity and an active lifestyle. In these works, emphasis is on higher layer communication network processing and hardware design. In contrast, our work explicitly focuses on developing the statistical signal processing techniques required for activity-level detection.

Several context-aware sensing systems and activity-level detection schemes have been designed using multiple accelerometers and heterogeneous sensors. However, the long-term deployment of some systems are constrained by the battery life of the individual sensors or the fusion center. The problem becomes more severe when Bluetooth, GPS measurements, and similar high-energy requirement features and applications are part of the sensor network.

The notion of designing energy-saving strategies, well-studied and implemented in the context of traditional sensor and mobile networks [17, 11], has also been incorporated into WBANs for activity-level detection. For example, the goal of the work by Benbasat et al [2] is to determine a sampling scheme (with respect to frequency of sampling and sleeping/waking cycles) for multiple sensors to minimize power consumption. A similar scheme which minimizes energy consumption based on redistributing un-used time over tasks as evenly as possible is described in the work by Liu et al [13]. Yan et al [22] have investigated the Bluetooth and ZigBee protocols in WBANs, and developed an energy optimization scheme based on the tunable parameters of these protocols, *e.g.* connection latency. Our approach is different in the fact that the energy-efficiency of the system is a result of optimized detection performance. In the next section, we describe our experimental setup, present the signal model used to develop our optimal time-resource allocation, and outline the optimization problem which uses the probability of error metric.

3 KNOWME Activity-Detection System Overview

The target functionality of the KNOWME network for activity-level detection, and the current system implementation, is outlined in this section. We note that our current work derives the optimal time-resource allocation algorithm for the “static” case wherein we do not explicitly account for the current state of the subject, *i.e.* the optimal allocation of samples does not evolve as a function of time. Figure 2 shows an illustrative example of a state-transition diagram that is used to determine the a priori probabilities for a particular activity-level. For example, if we know that the subject is standing, the transition probabilities from the Standing state to the Lying, Sitting, Walking and Standing states are given as 0.05, 0.4, 0.25 and 0.3, respectively. As is seen in Figure 2, the a priori probabilities are distinct for each state. These are incorporated into the probability of misclassification metric that we use to derive the optimal time-resource allocation algorithm in Section 4.2.

3.1 System Implementation

Our current implementation of the KNOWME software connects a Nokia N95 8GB phone to multiple bluetooth-enabled body sensors which continuously collect data and relay it to the phone. The phone, which serves as the fusion center, has a Dual ARM 11 332 MHz CPU, 128MB RAM, and is a Bluetooth 2.0 EDR compliant Class 2 device, running Symbian OS 9.2. The software running on the phone is written for Python for S60 devices. The software is configured to connect to multiple Bluetooth devices; on starting the data collection, the softwares opens a new thread for each device specified. The data received from the sensors can be parsed and analyzed on-the-fly in the write thread on the phone to decide whether to disconnect or disable a sensor or not. The thread handling that sensor can be put in a wait state after disconnecting, and can be

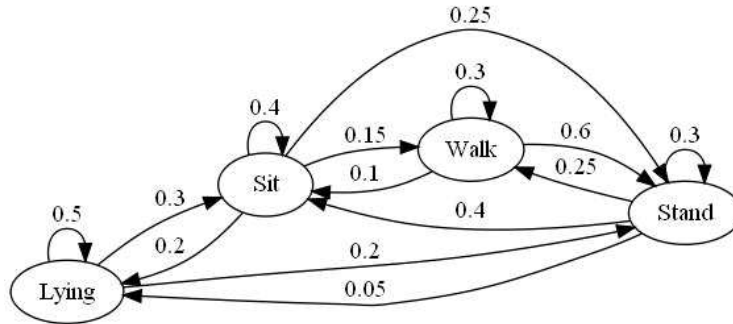


Fig. 2. Example of a state-transition diagram that may be used to determine the transition probabilities in the KNOWME activity-detection system.

signalled later when it is decided that we need data from that sensor again. We note that we are still currently working on implementing the optimal allocation of samples; in its current form, all sensors transmit an equal number of samples to the fusion center.

4 Problem Formulation

In this section, we first present the signal model for our wireless body area network which is deployed in a simple star topology, and then outline our optimization problem: We minimize the probability of misclassification at the fusion center, given samples from all the sensors. We note that obtaining the optimal allocation of samples requires an exhaustive search over the total number of samples, N , since all possible partitions of the samples between the heterogeneous sensors must be considered to find the optimal allocation. This was considered for the binary hypothesis case with two sensors in our previous work [19]. As the total number of available samples and number of sensors increases, the combinatorial search becomes undesirably computationally expensive.

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. The derivation of the alternate problem is outlined, and the algorithmic complexities of the two optimizations are compared; numerical simulations are presented in the following section.

4.1 Signal Model

Heterogeneous sensors are deployed in a simple star topology as shown in Figure 1. Each sensor sends its biometric samples directly to the cellphone via the Bluetooth protocol. There has been extensive research in the modeling of

physiological time-series data, and the autoregressive (AR) model has been employed in a variety of contexts. Physiological hand-terms are represented using an AR(3) model in [23], while the works in [14, 15] use an AR model to estimate ElectroEncephaloGram (EEG) signals. The AR model, while not always an exact match to the data, is one of the less complicated models used since it allows for a more thorough analysis and development of solutions. In our work, we use the AR(1) model to represent a biometric time-series. Furthermore, both the sensing and the communication of the measurements are assumed to be noisy given the measurement systems and wireless channels, respectively.

We now propose the following signal model for the decoded and processed samples received by the fusion center:

$$y_i = \theta + z_i, \quad i = 1, \dots, N_k \quad (1)$$

for the k -th sensor, where z_i represents the independent and identically distributed (iid) zero-mean Gaussian measurement and channel noise. For a general feature/sensor A_k , θ is a normally distributed random variable, specified as

$$\theta_j = \mu_{jA_k} + w_i \quad (2)$$

for hypothesis H_j , where w_i represents the *biometric noise* and is modeled using the AR(1) model, *i.e.*

$$w_i = \varphi w_{i-1} + \varepsilon, \quad i = 2, \dots, N_k, \quad (3)$$

for k -th sensor which has been allocated N_k samples, and ε is zero-mean Gaussian with variance $\sigma_{jA_k}^2$. The choice of the Gaussian model is motivated by several factors. An analysis of the data collected during our pilot study [1] suggested that the Gaussian model is a relatively good fit for the biometric samples collected by the fusion center. A more complicated model with a better fit to the data may be chosen, but the Gaussian model lends itself to analysis, and the development of tractable solutions. To simplify notation, we omit the hypothesis subscript j when expressions and definitions are applied to a generic hypothesis. We denote the number of samples sent by the K sensors as N_1, N_2, \dots, N_K , respectively, and impose a constraint of N total samples, *i.e.* $N_1 + N_2 + \dots + N_K = N$, for a specific time-period.

Since the features are modeled as Gaussian, and given that the AR(1) model is *linear*, the M -ary hypothesis test using the model in (1) is simply the generalized Gaussian problem which is specified as

$$H_i : \mathbf{Y} \sim N(\mathbf{m}_i, \mathbf{\Sigma}_i), \quad i = 1, \dots, M, \quad (4)$$

where $\mathbf{m}_i, \mathbf{\Sigma}_i, i = 1, 2, \dots, M$ are the mean vectors and covariance matrices of the observations under the each of the M hypotheses. For completeness, we recall the density of the multivariate Gaussian given by:

$$f_X(x_1, \dots, x_N) = \frac{1}{(2\pi)^{N/2} |\mathbf{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(x - \mu)^T \mathbf{\Sigma}^{-1}(x - \mu)\right), \quad (5)$$

where μ is the mean vector and Σ is the covariance matrix. Empirical data from our pilot study [1] suggests that the conditional correlation between features is relatively small; and thus, for the K features A_1, A_2, \dots, A_K from the sensors, the mean vector and covariance matrix for the observations for hypothesis H_j for $j = 1, \dots, M$ are of the form

$$\mathbf{m}_j = \begin{bmatrix} \mu_{jA_1} \\ \mu_{jA_2} \\ \vdots \\ \mu_{jA_K} \end{bmatrix} \text{ and } \Sigma_j = \begin{bmatrix} \Sigma_j(A_1) & 0 & 0 & \cdots & 0 \\ 0 & \Sigma_j(A_2) & 0 & \cdots & 0 \\ 0 & 0 & \Sigma_j(A_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \Sigma_j(A_K) \end{bmatrix}, \quad (6)$$

respectively. Note that μ_{jA_i} is a $N_i \times 1$ vector and $\Sigma_j(A_i)$ is a $N_i \times N_i$ matrix. We have assumed that the samples from different sensors are independent; this is further supported by the fact that certain sensors yield a better performance when discriminating between some subsets of activities.

Given the signal models in (1) and (3), for a particular feature A_k , the covariance matrix can be expressed as

$$\Sigma(A_k) = \frac{\sigma_{A_k}^2}{1 - \varphi^2} \mathbf{T} + \sigma_z^2 \mathbf{I}, \quad (7)$$

where \mathbf{T} is a Toeplitz matrix of the form

$$\mathbf{T} = \begin{bmatrix} 1 & \phi & \phi^2 & \cdots & \phi^{N_k-1} \\ \phi & 1 & \phi & \cdots & \phi^{N_k-2} \\ \phi^2 & \phi & 1 & \cdots & \phi^{N_k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi^{N_k-1} & \phi^{N_k-2} & \phi^{N_k-3} & \cdots & 1 \end{bmatrix}, \quad (8)$$

and \mathbf{I} is the $N_k \times N_k$ identity matrix. This results in the covariance matrices $\Sigma_j, j = 1, \dots, M$ being block-Toeplitz matrices. To derive a vector optimization that circumvents an exhaustive search, we may replace the Toeplitz covariance matrices with their associated circulant covariance matrices¹ given by

$$\Sigma(A_k) = \frac{\sigma_{A_k}^2}{1 - \varphi^2} \mathbf{C} + \sigma_z^2 \mathbf{I}, \quad (9)$$

¹ We note that the inverse of the Toeplitz covariance matrix in (7) converges to the inverse of the circulant covariance matrix in (9) in the weak sense. Sun et al [18] have derived that a sufficient condition for weak convergence is that the strong norm of the inverse matrices be uniformly bounded. We find that this is the case for the matrix forms in (7) and (9) for $0 < \phi < 1$.

where the matrix \mathbf{C} is of the form

$$\mathbf{C} = \begin{bmatrix} 1 & \phi & \phi^2 & \dots & \phi^{N_k-1} \\ \phi^{N_k-1} & 1 & \phi & \dots & \phi^{N_k-2} \\ \phi^{N_k-2} & \phi^{N_k-1} & 1 & \dots & \phi^{N_k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi & \phi^2 & \phi^3 & \dots & 1 \end{bmatrix}. \quad (10)$$

4.2 Probability of Error Derivation

We derive a closed-form approximation for the probability of error in the multihypothesis case via a union bound incorporating the Bhattacharyya coefficients between pairs of hypotheses. A result by Lianiotis [12] provides an upper bound on the probability of error, independent of the prior probabilities, given as

$$P(\epsilon) \leq \sum_{i < j} (P_i P_j)^{1/2} \rho_{ij}, \quad (11)$$

where ρ_{ij} is the Bhattacharyya coefficient defined as

$$\rho_{ij} = \int [f_i(x) f_j(x)]^{1/2} dx. \quad (12)$$

Thus, the optimization problem considered can be stated as

$$\min_{\mathbf{N}} P(\epsilon) \quad \text{subject to} \quad \sum_{k=1}^K N_k = N, \quad N_k \geq 0 \quad \forall k, \quad (13)$$

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

$$\rho_{ij} = \exp \left(- \left[\frac{1}{8} (m_i - m_j)^T \Sigma_h^{-1} (m_i - m_j) + \frac{1}{2} \log \frac{|\Sigma_h|}{\sqrt{|\Sigma_i| |\Sigma_j|}} \right] \right), \quad (14)$$

where $|\Sigma| = \det \Sigma$, and $2\Sigma_h = \Sigma_i + \Sigma_j$.

Given the block-diagonal structure of the covariance matrix in (6), we first decompose the terms of the Bhattacharyya coefficient ρ_{ij} for generic hypotheses H_i and H_j as follows:

$$\det \Sigma_i = \prod_{k=1}^K \det \Sigma_i(A_k), \quad (15)$$

and

$$(\mu_j - \mu_i)^T \Sigma_h^{-1} (\mu_j - \mu_i) = \sum_{k=1}^K (\mu_{jA_k} - \mu_{iA_k})^T \Sigma_h^{-1}(A_k) (\mu_{jA_k} - \mu_{iA_k}). \quad (16)$$

Thus, computing each of the terms for an individual feature A_k is sufficient to evaluate the probability of error specified in (11). The structure of the covariance matrix in (6) 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 an exhaustive 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 term in (15), we use the Toeplitz structure in (7), and rewrite the covariance matrix as follows [7]:

$$\begin{aligned} \mathbf{\Sigma}(A_k) &= \mathbf{\Sigma}_D(A_k) + \mathbf{\Sigma}_{\text{off}}(A_k) \\ &= \alpha \mathbf{I} + \frac{\sigma_{A_k}^2}{1 - \phi^2} \begin{bmatrix} 0 & \phi & \phi^2 & \dots & \phi^{N_k-1} \\ \phi & 0 & \phi & \dots & \phi^{N_k-2} \\ \phi^2 & \phi & 0 & \dots & \phi^{N_k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi^{N_k-1} & \phi^{N_k-2} & \phi^{N_k-3} & \dots & 0 \end{bmatrix}, \end{aligned} \quad (17)$$

where $\alpha = \sigma_{A_k}^2 / (1 - \phi^2) + \sigma_z^2$. Given this expansion, the determinant of the covariance matrix can be computed using

$$\det \mathbf{\Sigma} = \det \mathbf{\Sigma}_D \cdot \det (\mathbf{I} + \mathbf{\Sigma}_D^{-1} \mathbf{\Sigma}_{\text{off}}), \quad (19)$$

wherein, using $\mathbf{A} = \mathbf{\Sigma}_D^{-1} \mathbf{\Sigma}_{\text{off}}$, we now evaluate

$$\det (\mathbf{I} + \mathbf{A}) = \exp (\text{tr} (\log (\mathbf{I} + \mathbf{A}))) \quad (20)$$

$$= \exp \left(\text{tr} \left(\mathbf{A} - \frac{\mathbf{A}^2}{2} + \frac{\mathbf{A}^3}{3} - \frac{\mathbf{A}^4}{4} + \dots \right) \right). \quad (21)$$

From the form in (19), and using the geometric progression

$$\sum_{k=0}^n kr^k = r \left[\frac{1 - r^{n+1}}{(1 - r)^2} - \frac{(n + 1)r^n}{1 - r} \right], \quad \text{for } r \neq 1, \quad (22)$$

we evaluate the single feature term in (15) as

$$\det \mathbf{\Sigma}(A_k) = \alpha^{N_k} e^{-C[-1 + \phi^{2N_k} - N_k(1 - \phi^{-2})]}, \quad (23)$$

where

$$C = \frac{1}{\alpha^2} \left[\frac{\sigma_{A_k}^2}{1 - \phi^2} \right]^2 \frac{\phi^{-2}}{(1 - \phi^{-2})^2}.$$

To evaluate the term in (16), the circulant approximation in (9) is employed and we can simplify (16) as

$$(\mu_{yA_k} - \mu_{xA_k})^2 \sum_{i=1}^{N_k} \sum_{j=1}^{N_k} [\mathbf{\Sigma}_y^{-1}(A_k)]_{ij}, \quad (24)$$

which shows that we do not need to compute Σ_y^{-1} , but only require the sum of all the elements of the inverse matrix. To this end, we employ a simple result by Wilansky [21] 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 $1/c$. Note that the sum of the elements of the n -th row of $\Sigma(A_k)$ can be simplified as

$$\sum_{j=1}^{N_k} [\Sigma(A_k)]_{nj} = \frac{\sigma_{A_k}^2}{1-\phi^2} \cdot \frac{1-\phi^{N_k}}{1-\phi} + \sigma_z^2 + \sigma_n^2 \quad (25)$$

using the geometric progression,

$$\sum_{k=0}^n r^k = \frac{1-r^{n+1}}{1-r} \quad \text{for } r \neq 1. \quad (26)$$

Thus, for a single block of the covariance matrix $\Sigma(A_k)$, we can evaluate the term in (16) as

$$(\mu_{yA_k} - \mu_{xA_k})^2 N_k \left[\frac{\sigma_{yA_k}^2 (1-\phi^{N_k})}{(1-\phi^2)(1-\phi)} + \sigma_z^2 + \sigma_n^2 \right]^{-1}. \quad (27)$$

Note that the simplified expressions in (23) and (27) are used to compute the bound on the probability of error in (11), but are independent of the discrete block-diagonal structure of the covariance matrix in (6). Thus, an exhaustive search over K integer partitions of N total samples is converted to a continuous-valued vector optimization over $[0, 1]^K$ which is minimized with lower complexity. The optimization problem may be rescaled with no loss of generality.

5 Performance Analysis

In this section, we present a numerical analysis of experimental data collected during our pilot study [1], and compare the complexity of the optimal exhaustive search and the approximately optimal vector optimization. Data was collected using graduate student test subjects; training and testing periods of 10 and 4 minutes, respectively, are collected for seven activities (Sitting, Standing, Sitting and Fidgeting, Standing and Fidgeting, Lying down, Walking and Running) from two accelerometers and an ECG monitor.

5.1 Optimal Allocation using Sensor Data

We consider the case with four hypotheses (Sit, Stand, Sit&Fidget, and Stand&Fidget), and three available features. The features considered, extracted from the three sensors, are accelerometer mean, accelerometer variance, and the ECG time-period, which are allocated N_1 , N_2 and N_3 samples, respectively. The underlying distributions for these hypotheses, for a single participant, is shown in Figure 3 for each of the sensors/features.

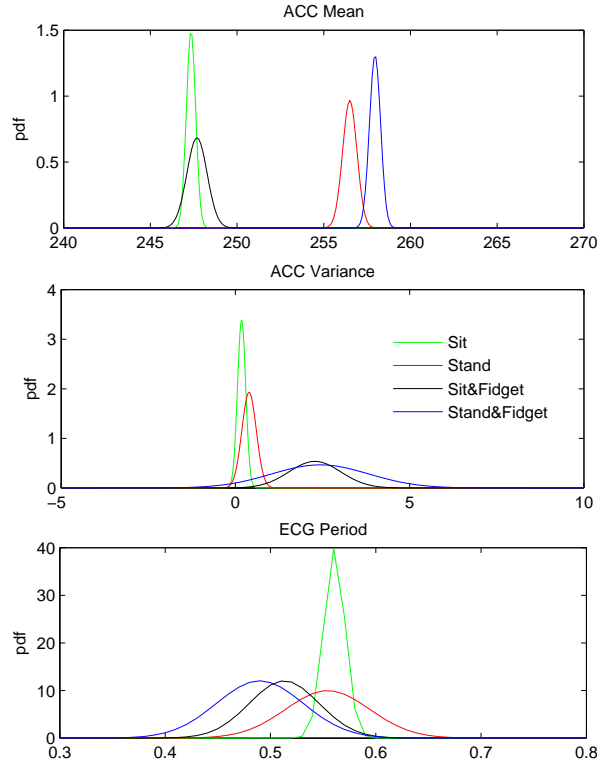


Fig. 3. Distributions for four hypotheses and three features; ACC mean, ACC variance and ECG period that are allocated N_1, N_2 and N_3 samples, respectively.

We assume that “Sit” is the current state, and that the transition probabilities to the states (Sit, Stand, Sit&Fidget, and Stand&Fidget) are $[0.3, 0.15, 0.5, 0.05]$. We first compute the optimal solution via an exhaustive search over all possible sample allocations, and then compute the approximately optimal solution via the lower complexity vector optimization. Figure 4 shows the results of an exhaustive search for the numerical case specified above with $N = 30$ samples and $\phi = 0.25$. We have assumed that each of the sensors is allocated at least one sample, and we see that the lowest probability of error is achieved for $\{N_1 = 2, N_2 = 27, N_3 = 1\}$. Thus, the optimal allocation of samples in this case is to allocate most of the samples to the accelerometer variance feature, and no samples to the ECG period.

The probability of error is also minimized via the vector optimization based on the closed-form approximations derived in the previous section. The minimization problem is rescaled, without loss of optimality, so that $N = 1$ and we use an initial guess of $N_0 = (1/3, 1/3, 1/3)$. We employ Matlab’s `fmincon` optimization function, and find $N^* = (0.08, 0.92, 0)$, which corresponds to an al-

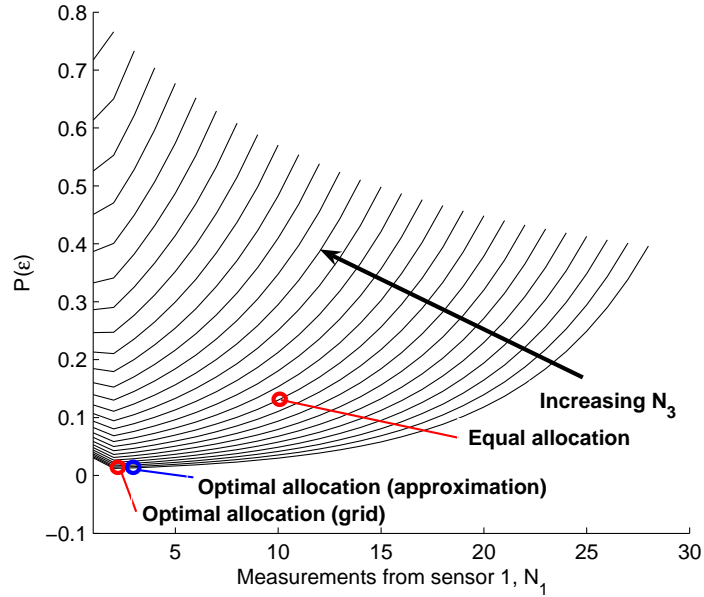


Fig. 4. Minimizing the probability of error $P(\epsilon)$ via an exhaustive search and the real-valued continuous optimization yields approximately equal solutions.

location of $(4, 26, 0)$ amongst the three sensors. The optimal and approximately optimal solutions are indicated on Figure 4, along with the probability of error corresponding to an equal allocation of samples. We find that the approximately optimal solution ($P(\epsilon) = 0.017$) obtained via the vector optimization provides an improvement compared to the equal allocation case ($P(\epsilon) = 0.138$). This vector optimization requires 29 function evaluations, and the exhaustive search requires 406 function evaluations for $N = 30$ samples.

We recall that the optimal allocation of samples will be different when the current state changes, and for different values of transition probabilities from the current state to the next possible state.

5.2 Algorithm Complexity and Scalability

Computing the optimal allocation of samples via the vector optimization is much lower complexity than the exhaustive search. Given N samples and K sensors, the number of function evaluations required are

$$\binom{N + K - 1}{K - 1} \sim \mathcal{O}(N^{K-1}). \quad (28)$$

A functional evaluation is defined as a single computation of the Bhattacharyya coefficient in (14). The notation $\mathcal{O}(\cdot)$ denotes the complexity of an algorithm

in the asymptotic sense, *i.e.* a complexity of $\mathcal{O}(N^2)$ means that as the number of samples, N , gets very large, the exhaustive search will require N^2 function evaluations.

On the other hand, the vector optimization via Matlab’s `fmincon` function, for the $N = 30$ case, requires between 30-60 function evaluations, and does not depend on the number of total samples available since the optimization problem has been rescaled. As N gets large, computing the optimal allocation of samples via the continuous optimization yields a significant reduction in complexity; in the $K = 3$ case considered herein, the number of function evaluations is reduced from $\mathcal{O}(N^2)$ to $\mathcal{O}(1)$; the latter notation means that the number of function evaluations required will remain approximately constant, regardless of the total number of available samples, N .

The model and derivation presented in this work incorporates 3 sensors and multiple hypotheses. As more sensors are incorporated, the decreased complexity due to the vector optimization over an exhaustive search is much more pronounced as is suggested by (28). Furthermore, an increase in the total number of samples will not affect the optimization.

6 Conclusion and Future Work

The problem of minimizing the probability of misclassification when discriminating between multiple hypotheses in the context of activity-level detection in wireless body area networks has been considered. We find that equally allocating samples amongst heterogeneous sensors is not optimal, and the probability of error can be minimized by allocating more samples to sensors that can better discriminate between a subset of hypotheses. Given a fixed number of samples, the optimal allocation of samples may be obtained via an exhaustive search which is computationally expensive. An alternative vector optimization derived yields an approximately optimal solution with significantly lower complexity.

Since not all sensors are required to transmit all their measurements in a specific time-period, the Bluetooth communication between a particular sensor and the fusion center can be turned off. We are working on quantifying energy savings that occur as a result of the optimized performance. Furthermore, we are developing the KNOWME activity-detection system, described in Section 3, that incorporates time-varying optimal allocations. During its development, these algorithms will be trained and tested in overweight minority youth participants, with the aim of providing accurate, user-friendly, interpretable real-time information on hard to measure outcomes such as physical activity. In the future, the KNOWME body-area network will provide excellent tools to accurately measure real-time obesity-related outcomes as well as a modality for intervention.

References

1. M. Annavaram, N. Medvidovic, et al. Multimodal sensing for pediatric obesity applications. In *Proceedings of UrbanSense08*, Raleigh, NC, November 2008.

2. A. Benbasat and J. Paradiso. A framework for the automated generation of power-efficient classifiers for embedded sensor nodes. In *Proc. of SenSys*, November 2007.
3. S. Biswas and M. Quwaider. Body posture identification using hidden markov model with wearable sensor networks. In *Proc. of BodyNets*, March 2008.
4. S. Consolvo, D. W. McDonald, T. Toscos, et al. Activity sensing in the wild: A field trial of ubifit garden. In *Proceedings of the Conference on Human Factors in Computing Systems*, Florence, Italy, April 2008.
5. F. Dabiri, H. Noshadi, H. Hagopian, et al. Light-weight medical bodynets. In *Proc. of BodyNets*, Florence, Italy, June 2007.
6. T. Gao, C. Pesto, L. Selavo, et al. Wireless medical sensor networks in emergency response: Implementation and pilot results. In *Proceedings of Int'l Conference on Technologies for Homeland Security*, May 2008.
7. I. Ipsen and D. Lee. Determinant approximations. *Numerical Linear Algebra with Applications*, (Under review), 2006.
8. S. Jiang, Y. Cao, et al. Carenet: An integrated wireless sensor networking environment for remote healthcare. In *Proc. of BodyNets*, Tempe, AZ, March 2008.
9. E. Jovanov, A. Milenkovic, C. Otto, and P. C. de Groen. A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, 2:6, March 2005.
10. A. Kalpaxis. Wireless temporal-spatial human mobility analysis using real-time three dimensional acceleration data. In *Proc. of International Multi-Conference on Computing in the Global Information Technology*, March 2007.
11. S. Kang, J. Lee, et al. Seemon: scalable and energy-efficient context monitoring framework from sensor-rich mobile environments. In *Proc. of MobiSys*, June 2008.
12. D. Lianiotis. A class of upper bounds on probability of error for multihypotheses pattern recognition. *IEEE Transactions on Information Theory*, November 1969.
13. Y. Liu, B. Veeravalli, and S. Viswanathan. Critical-path based low-energy scheduling algorithms for body area network systems. In *Proc. of RTCSA*, August 2007.
14. G. Mohammadi, P. Shoushtari, et al. Person identification by using ar model for eeg signals. *Proc. of World Academy of Sci., Engr. and Tech.*, 11, February 2006.
15. J. Pardey, S. Roberts, and L. Tarasenko. A review of parametric modeling techniques for eeg analysis. *Medical Engineering and Physics*, 18(1):2–11, 1996.
16. S. Reddy, A. Parker, J. Hyman, et al. Image browsing, processing, and clustering for participatory sensing: Lessons from a dietsense prototype. In *Proceedings of Workshop on Embedded Networked Sensors*, Cork, Ireland, June 2007.
17. E. Shih, P. Bahl, and M. Sinclair. Wake on wireless: an event driven energy saving strategy for battery operated devices. In *Proc. of MobiCom*, September 2002.
18. F.-W. Sun, Y. Jiang, and J. S. Baras. On the convergence of the inverses of toeplitz matrices and its applications. *IEEE Trans IT*, 49(1):180–190, January 2003.
19. G. Thatte, V. Rozgic, et al. Optimal time-resource allocation for activity-detection via multimodal sensing. In *Proc. of BodyNets*, Los Angeles, CA, April 2009.
20. Y. Wang, J. Lin, M. Annavaram, et al. A framework of energy efficient mobile sensing for automatic user state recognition. In *Submitted to 7th Annual International Conference on Mobile Systems, Applications and Services*, June 2009.
21. A. Wilansky. The row-sums of the inverse matrix. *The American Mathematical Monthly*, 58(9):614–615, November 1951.
22. L. Yan, L. Zhong, and N. K. Jha. Energy comparison and optimization of wireless body-area network technologies. In *Proceedings of the ICST 2nd International Conference on Body Area Networks*, Florence, Italy, June 2007.
23. J. Zhang and F. Chu. Real-time modeling and prediction of physiological hand tremor. In *Proceedings of ICASSP*, Philadelphia, PA, March 2005.