Optimal Time-Resource Allocation for Activity-Detection via Multimodal Sensing

Gautam Thatte[†]
thatte@usc.edu

Viktor Rozgic[†], Ming Li[†]
{rozgic,mingli}@usc.edu

Sabyasachi Ghosh[†]
sabyasag@usc.edu

Urbashi Mitra[†]
ubli@usc.edu

Shri Narayanan[†]
shri@sipi.usc.edu

Murali Annavaram[†]
annavara@usc.edu

Donna Spruijt-Metz[‡]
dmetz@usc.edu

[†]Ming Hsieh Department of Electrical Engineering, University of Southern California, USA
[‡]Keck School of Medicine, University of Southern California, USA

ABSTRACT

The optimal allocation of measurements for activity-level detection in a wireless body area network (WBAN) for health-monitoring applications is considered. The WBAN with heterogeneous sensors is deployed in a simple star topology with the fusion center receiving a fixed number of measurements from the sensors; the number of measurements allocated to each sensor is optimized to minimize the probability of detection error at the fusion center. An analysis of the two-sensor case with binary hypotheses is presented. Since the number of measurements is an integer, an exhaustive search (grid search) is traditionally employed to determine the optimal allocation of measurements. 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 lower complexity. Numerical case studies based on experimental data for different key activity-states are presented. It is observed that the Kullback-Leibler (KL) distances between the distributions associated with the hypotheses dominate the optimal allocation of measurements.

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). The objectives of such WBANs are manifold from diet monitoring [18], activity detection [4, 3], and health crisis support [9]. We focus on the KNOWME network [2], which supports the development of assessments and interventions for pediatric obesity applications. Pediatric obesity has emerged as a major national and international health crisis: nationally collected data from 2003-2006 show 11.3% of adolescents aged 12–19 years could be designated by some measures as obese, while a further 16% would be classified as overweight and 32% considered at risk for being overweight [16]. In order to truly understand and reverse childhood obesity, we need a multimodal system that will track stress levels, physical activity levels, blood glucose levels and other vital signs simultaneously, as well as anchor these levels to context such as time of day and geographical location. The KNOWME network is a first step towards such a system.

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

A key aspect of the KNOWME network is the unified design and evaluation of multimodal sensing and interpretation for automatically recognizing, predicting and reasoning about human physical activity and socio-cognitive behavior states. On the one hand, the KNOWME network meets the needs of traditional observational research practices in the obesity and metabolic health domain (based on, and validated through, careful expert human coding of data) while on the other, it enables new analysis capabilities such as providing information on user emotional state in conjunction with physical activity and energy expenditure.

The KNOWME network deploys heterogeneous sensors which communicate their measurements via Bluetooth to a Nokia N95 cellphone, shown in Figure 1. The adoption of the Bluetooth standard for communication results in a "serve as available" protocol, *i.e.* every measurement is collected

BodyNets '09 Los Angeles, California USA

from each of the sensors. While this is beneficial for signal processing and activity-level detection at the fusion center, it results in particularly high energy consumption. For instance, the Nokia N95 cellphone can support more than ten hours of telephone conversations on a fully charged battery, but empirical results show that the battery is drained in under six hours if the GPS receiver is turned on [23]. Similarly, the battery drains very quickly if Bluetooth functionality is constantly left on.

Our pilot study [2] was configured to have each sensor transmit an equal number of measurements to the fusion center in each time period via the Bluetooth protocol. The study suggested that some sensors were more effective in distinguishing between certain activities than other sensors. For example, low-level activities (lying down, sitting and standing) were better distinguished between using the electrocardiograph (ECG) sensor, while the accelerometer was more effective when distinguishing between high-level activities (walking and running). In this work, we exploit the dependency of the detection performance on the specific activity-levels being considered, and propose an alternate allocation of the measurements to each of the sensors in order to develop a more energy-efficient detection mechanism.

The goal of this paper is to develop intuition for what the optimal allocation of measurements between heterogeneous sensors should be to minimize the probability of detection error at the fusion center. Achieving a better performance via optimal time-resource allocation then suggests that sensors that need to communicate fewer measurements, or no measurements at all, can turn their Bluetooth off, thus resulting in a more energy-efficient health-monitoring application. In the current work, we develop preliminary results for the binary hypothesis testing problem with two sensors.

The contribution of this work is to describe the optimal allocation of measurements between two sensors for the binary hypothesis testing problem in order to minimize the probability of detection error. Specifically, an approximately optimal continuous-valued solution, which does not require a high-complexity integer grid search, is derived. Not requiring a grid search enables the real-time deployment of the allocation scheme, since the optimal allocation of measurements can be computed faster at the fusion center. This is particularly important as the number of sensors and hypotheses considered increases. The extension to multiple hypotheses is outlined, and we are currently developing an energy-efficient algorithm using this optimal allocation of measurements.

The remainder of this paper is organized as follows. Prior relevant work in activity-level detection and energy-efficient algorithms in WBANs, and its relationship to our work, is presented in Section 2. In Section 3, we describe the specific WBAN employed for activity-level detection, and the corresponding signal model used to develop our optimal time-resource allocation. In Section 4, we outline the framework for minimizing the probability of detection error, and derive an approximately optimal solution. Numerical results and case studies based on experimental data are presented in Section 5. Finally, we draw conclusions and discuss our future work direction in Section 6.

2. RELATED WORK

In recent years, there have been several projects that have investigated activity-level detection in a variety of frameworks. Much of the work appears to center on accelerometer data alone (*e.g.* [4, 11, 13]), with some systems using many accelerometer packages. These studies employ both accelerometer data and the relative positions of the sensors to discriminate between specific high-level and low-level activities. The work in [4] develops a Hidden Markov Model based identification system that uses accelerometer measurements to detect high-level activities, and the relative proximity between accelerometers to distinguish between low-level activities. Multi-sensor systems have also been implemented and deployed for activity-level detection, context-aware sensing and specific health-monitoring applications: the work of Gao et al [9] is tailored for emergency response and triage situations, while Dabiri et al [7] have developed a lightweight embedded system that is primarily used for patient monitoring. The system developed by Jovanov et al [12] is used to assist physical rehabilitation, and Consolvo et al's UbiFit system [5] is designed to promote physical activity and an active lifestyle. In these works, the emphasis is on the 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, while also creating an implementation scheme. This scheme was tested in our pilot study [2], and is currently being further developed.

A variety of 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 is 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.

Energy-saving strategies, well-studied and implemented in the context of traditional sensor and mobile networks [20, 14], have also been incorporated into WBANs for activity-level detection. For example, the goal of Benbasat et al [3] is to determine a sampling scheme (with respect to frequency of sampling and sleeping/waking cycles) for multiple sensors to minimize power consumption. Our approach is different in that the energy-efficiency of the system is a result of optimized performance. In the next section, we describe our signal model and develop the optimal time-resource allocation problem.

3. PROBLEM FORMULATION

The current study employs an Alive Technologies [1] electrocardiograph (ECG). The ECG is a single channel device with 8 bit resolution and a peak sampling rate of 300 samples/second. This sensor also incorporates an accelerometer that provides three-dimensional acceleration measurements over the single channel. The pulse-oximeter, also from Alive, provides non-invasive monitoring of oxygen saturation (SpO_2) and pulse rate. The sensors are deployed in a simple star topology, as shown in Figure 1: each sensor sends its measurements directly to the cellphone fusion center.

Measurements of certain biometrics, *e.g.* heart-rate and pulse, are transmitted with some fixed transmission power to the fusion center via Bluetooth. Both the sensing and the communication of the measurements are assumed to be noisy given the measurement systems and wireless channels, respectively. The features extracted from the biometric signals are modeled as correlated Gaussian random variables,

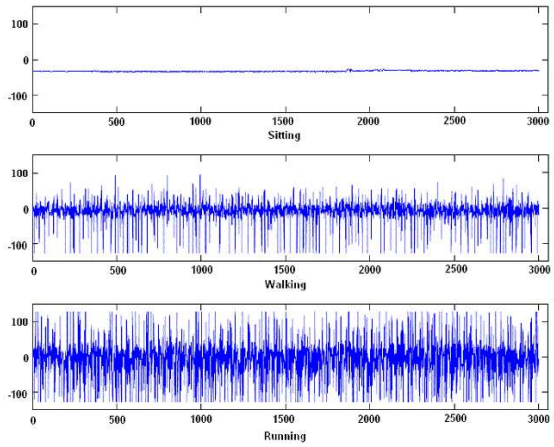

Figure 2: Accelerometer time-series for three hypotheses: Sitting, Walking and Running.

and thus we propose the following signal model for the measurements received by the fusion center:

$$y_i = \sqrt{P_i}(\theta + z_i) + n_i, \ i = 1, \ldots, N_k \tag{1}$$

for the k-th sensor, where P_i represents the transmission power of the i-th measurement, and z_i and n_i represent the independent and identically distributed (iid) zero-mean Gaussian measurement and channel noises, respectively. For a general feature A, θ is a normally distributed random variable specified as

$$\theta_j = \mu_{jA} + w_i \tag{2}$$

for hypothesis H_j. The noise term w_i is modeled using an autoregressive (AR) model of the first order, denoted AR(1), *i.e.*,

$$w_i = \varphi w_{i-1} + \varepsilon \ , \quad i = 2, \ldots, N_k, \tag{3}$$

for the k-th sensor, which has been allocated N_k measurements, and ε is zero-mean Gaussian with variance σ^2_{jA}. The AR(1) process models the temporal dependency of the biometric signals. To simplify notation, we omit the hypothesis subscript $_j$ when expressions and definitions are applied to both hypotheses. We denote the number of measurements sent by the two sensors as N_1 and N_2, respectively, and impose a constraint of N total measurements, *i.e.* $N_1 + N_2 = N$, for a specific time-period.

For example, consider a single set of sample waveforms, as in Figure 2, which show the accelerometer signal received at the fusion center for the Sitting, Walking and Running hypotheses for a single subject. The variance of the time-series over 1-second non-overlapping segments is the feature extracted from the accelerometer data. This feature is modeled, as described in our signal model, as a Gaussian random variable; in this particular case we find

$$\begin{aligned} \text{Sitting} &: \ \mu_S = 0.18, \sigma^2_S = 0.02, \\ \text{Walking} &: \ \mu_W = 27.2, \sigma^2_W = 1.94, \\ \text{Running} &: \ \mu_R = 33.8, \sigma^2_R = 0.76, \end{aligned}$$

and the AR(1) parameter is $\phi = 0.24$. The signals from each of the sensors are processed in a similar manner for each of the required features. The means and variances obtained by modeling the features as Gaussian are subsequently used to determine the optimal allocation of measurements as described in Section 5.

Given our assumptions, the binary hypothesis test using the model in (1) is simply the generalized Gaussian problem, which is specified as

$$\begin{aligned} H_0 : \quad & \mathbf{Y} \sim N(\mathbf{m}_0, \mathbf{\Sigma}_0) \\ H_1 : \quad & \mathbf{Y} \sim N(\mathbf{m}_1, \mathbf{\Sigma}_1) \end{aligned} \tag{4}$$

where $\mathbf{m}_i, \mathbf{\Sigma}_i, i = 0, 1$ are the mean vectors and covariance matrices of the observations under the two hypotheses.

We initially assume that measurements from different sensors are independent of each other for two activity-levels. Thus, for two features A and B from these sensors, the mean vector and covariance matrix of the observations for hypothesis H_j are of the form

$$\mathbf{m}_j = \begin{bmatrix} \mu_{jA} \\ \mu_{jB} \end{bmatrix} \text{ and } \mathbf{\Sigma}_j = \begin{bmatrix} \mathbf{\Sigma}_j(A) & 0 \\ 0 & \mathbf{\Sigma}_j(B) \end{bmatrix}, \tag{5}$$

respectively. Note that μ_{jA} and μ_{jB} are $N_1 \times 1$ and $N_2 \times 1$ vectors, and $\mathbf{\Sigma}_j(A)$ and $\mathbf{\Sigma}_j(B)$ are $N_1 \times N_1$ and $N_2 \times N_2$ matrices, respectively. Given the signal models in (1) and (3) for a particular sensor with feature A, the i-th element of mean vector is $\sqrt{P_i}\mu_A$, and the elements of covariance matrix are specified as

$$\mathbf{\Sigma}_{ii}(A) = P_i \frac{\sigma^2_A}{1-\varphi^2} + P_i\sigma^2_z + \sigma^2_n \tag{6}$$

$$\mathbf{\Sigma}_{ij}(A) = \sqrt{P_i P_j}\frac{\sigma^2_A}{1-\varphi^2}\varphi^{|i-j|}, \tag{7}$$

where ϕ is the AR(1) model parameter which incorporates the temporal dependency of the biometric signal. The covariance block $\mathbf{\Sigma}(A)$, described in (6) and (7), is a Toeplitz matrix which, for equal transmission powers, can be written as

$$\mathbf{\Sigma}(A) = \frac{\sigma^2_A}{1-\varphi^2}\mathbf{T} + (\sigma^2_z + \sigma^2_n)\mathbf{I} \tag{8}$$

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

$$\mathbf{T} = \begin{bmatrix} 1 & \varphi & \cdots & \varphi^{N_i-2} & \varphi^{N_i-1} \\ \varphi & 1 & \varphi & & \varphi^{N_i-2} \\ \vdots & \varphi & 1 & \ddots & \vdots \\ \varphi^{N_i-2} & & \ddots & \ddots & \varphi \\ \varphi^{N_i-1} & \varphi^{N_i-2} & \cdots & \varphi & 1 \end{bmatrix}. \tag{9}$$

This implies the covariance matrices $\mathbf{\Sigma}_0$ and $\mathbf{\Sigma}_1$ are block-Toeplitz. In the next section, we derive a bound for the probability of error which we optimize over the measurements from each of the sensors. Furthermore, a closed-form approximately optimal solution for a simplified case is developed.

4. MINIMIZING ERROR PROBABILITY

The probability of detection error, also known as classification error, *i.e.* deciding H_0 when H_1 is actually true or vice versa, cannot be meaningfully simplified for the general Gaussian hypothesis testing problem in (4) [22]. Since a purely numerical analysis has limited value, we attempt to the Chernoff upper bound for the probability of error, which will then allow us to better investigate the optimal allocation

of measurements between the sensors. We assume that the hypotheses are equally likely and the error costs are equal, and thus we can bound the probability of error for the above hypothesis testing problem as [8]

$$\mathrm{P}(\epsilon) \;\leq\; e^{-k(\beta_m)}, \tag{10}$$

where

$$k(\beta) = -\ln \int_{-\infty}^{\infty} \left[p(\mathbf{Y}|H_1)\right]^{\beta} \left[p(\mathbf{Y}|H_0)\right]^{1-\beta} d\mathbf{Y}, \tag{11}$$

and $0 \leq \beta_m \leq 1$ satisfies $\dot{k}(\beta_m) = 0$. For the hypotheses in (4), we compute

$$\begin{aligned} k(\beta) &= \frac{\beta(1-\beta)}{2}(\mathbf{m}_0 - \mathbf{m}_1)^T \mathbf{\Lambda}^{-1}(\mathbf{m}_0 - \mathbf{m}_1) \\ &\quad + \frac{1}{2}\ln \frac{|\mathbf{\Lambda}|}{|\mathbf{\Sigma}_1|^{\beta}|\mathbf{\Sigma}_0|^{1-\beta}} \;, \end{aligned} \tag{12}$$

where

$$\mathbf{\Lambda} = \beta\mathbf{\Sigma}_1 + (1-\beta)\mathbf{\Sigma}_0 \;. \tag{13}$$

Thus, the optimization problem can be expressed as

$$\min_{N_1, N_2} \; \mathrm{P}(\epsilon) \;\; \text{such that } N_1 + N_2 = N \;. \tag{14}$$

The optimal allocation of measurements between the two sensors that minimizes the probability of error can be obtained via a grid search over the total number of observations since the total number of measurements is an integer, and the structure of the covariance matrix changes for different allocations of measurements between the two sensors. However, this exhaustive search is computationally expensive, and increasingly so as the number of sensors and hypotheses increases. Therefore, we derive a closed-form analysis of a simplified case which yields an approximately optimal allocation of measurements between the individual sensors, while remaining less computationally expensive.

To simplify the analysis, we assume equal transmission powers for each of the measurements, and the sensor covariance matrix in this case is given by (8). Note that for the block-diagonal structure of $\mathbf{\Sigma}_j$ in (5), we have

$$\det \mathbf{\Sigma}_j = \det \mathbf{\Sigma}_j(A) \cdot \det \mathbf{\Sigma}_j(B) \tag{15}$$

and

$$\mathbf{\Sigma}_j^{-1} = \begin{bmatrix} \mathbf{\Sigma}_j^{-1}(A) & 0 \\ 0 & \mathbf{\Sigma}_j^{-1}(B) \end{bmatrix}. \tag{16}$$

In order to derive an analytic expression for the probability of error, we further simplify the Chernoff bound by considering the case in (12) where $k = 1/2$, which yields the Bhattacharyya bound that is rewritten as

$$\begin{aligned} k\left(\frac{1}{2}\right) &= \frac{1}{8}(\mathbf{m}_0 - \mathbf{m}_1)^T \left[\frac{\mathbf{\Sigma}_0 + \mathbf{\Sigma}_1}{2}\right]^{-1} (\mathbf{m}_0 - \mathbf{m}_1) \\ &\quad + \frac{1}{2}\ln \frac{1}{\sqrt{|\mathbf{\Sigma}_1||\mathbf{\Sigma}_0|}} \left|\frac{\mathbf{\Sigma}_0 + \mathbf{\Sigma}_1}{2}\right| \end{aligned} \tag{17}$$

where $(\mathbf{\Sigma}_0 + \mathbf{\Sigma}_1)/2$ is block-diagonal as in (5). Although the Chernoff bound is tighter than the Bhattacharyya bound, it is only marginally so since the optimal value β_m of the Chernoff bound does not normally lie at the ends of the interval $[0, 1]$. On the other hand, this simplification significantly reduces the complexity of the derivation.

We first derive results for a single block of the covariance matrix $\mathbf{\Sigma}(A)$, and then consider the two-sensor case to simplify the expression in (17). To evaluate the determinant terms in (17), we use the Toeplitz structure from (8), and rewrite the covariance matrix as follows [10]:

$$\begin{aligned} \mathbf{\Sigma}(A) &= \mathbf{\Sigma}_{\mathrm{D}}(A) + \mathbf{\Sigma}_{\mathrm{off}}(A) &(18)\\ &= \alpha\mathbf{I} + \frac{\sigma_A^2}{1-\phi^2}(\mathbf{T} - \mathbf{I}), &(19) \end{aligned}$$

where $\alpha = \sigma_A^2/(1-\phi^2) + \sigma_z^2 + \sigma_n^2$, and $\mathbf{T}$ is as defined in (9). Given this expansion, the determinant of the covariance matrix can be computed using

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

We denote $\mathbf{A} = \mathbf{\Sigma}_{\mathrm{D}}^{-1}\mathbf{\Sigma}_{\mathrm{off}}$, and we now evaluate

$$\begin{aligned} \det(\mathbf{I} + \mathbf{A}) &= \exp\left(\mathrm{tr}\left(\log\left(\mathbf{I} + \mathbf{A}\right)\right)\right) &(21)\\ &= \exp\left(\mathrm{tr}\left(\mathbf{A} - \frac{\mathbf{A}^2}{2} + \frac{\mathbf{A}^3}{3} - \cdots\right)\right). &(22) \end{aligned}$$

From the form in (20), and using the geometric progression

$$\sum_{k=0}^{n} k r^k \;=\; r\left[\frac{1 - r^{n+1}}{(1-r)^2} - \frac{(n+1)r^n}{1-r}\right] \quad \text{for } r \neq 1 \;, \tag{23}$$

we evaluate the single feature term $|\mathbf{\Sigma}(A)|$ as

$$\det \mathbf{\Sigma}(A) \;=\; \alpha^{N_k} e^{-C\left[-1+\phi^{2N_k} - N_k(1-\phi^{-2})\right]}, \tag{24}$$

where

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

In order to evaluate the inverse term in (17), we replace the Toeplitz matrices with their associated circulant matrices, based on the well-known fact that Toeplitz matrices asymptotically converge to their associated circulant matrices in the weak sense [21, 19] as the number of total measurements gets large.

Note that a block of the covariance matrix with N_i measurements can be rewritten as

$$\mathbf{\Sigma}(A) = \frac{\sigma_A^2}{1-\varphi^2}\mathbf{C} + \sigma^2\mathbf{I} \tag{25}$$

where $\sigma^2 = \sigma_z^2 + \sigma_n^2$ and $\mathbf{C}$ is a $N_i \times N_i$ circulant matrix of the form

$$\mathbf{C} = \begin{bmatrix} 1 & \varphi & \cdots & \varphi^{N_i-2} & \varphi^{N_i-1} \\ \varphi^{N_i-1} & 1 & \varphi & & \varphi^{N_i-2} \\ \vdots & \varphi^{N_i-1} & 1 & \ddots & \vdots \\ \varphi^2 & & \ddots & \ddots & \varphi \\ \varphi & \varphi^2 & \cdots & \varphi^{N_i-1} & 1 \end{bmatrix} \tag{26}$$

where the first row of the approximated covariance matrix in (25) is identical to that of the original Toeplitz matrix, described in (6) and (7).

We denote $\mathbf{d} = \mathbf{m}_0 - \mathbf{m}_1$ and $\mathbf{\Sigma}_s = (\mathbf{\Sigma}_0 + \mathbf{\Sigma}_1)/2$, and rewrite (17) as

$$k\left(\frac{1}{2}\right) = \frac{1}{8}\mathbf{d}^T\mathbf{\Sigma}_s^{-1}\mathbf{d} + \frac{1}{2}\ln\frac{|\mathbf{\Sigma}_s|}{\sqrt{|\mathbf{\Sigma}_1||\mathbf{\Sigma}_0|}} \tag{27}$$

where

$$\mathbf{\Sigma}_s(A) = \frac{\sigma_{0A}^2 + \sigma_{1A}^2}{2(1-\varphi^2)}\mathbf{C} + \sigma^2\mathbf{I}, \tag{28}$$

and the inverse matrix $\mathbf{\Sigma}_s^{-1}$ is of the form described in (16), and $\mathbf{d}$ has the form

$$\mathbf{d} = \begin{bmatrix} \mu_{0A} - \mu_{1A} \\ \mu_{0B} - \mu_{1B} \end{bmatrix} \quad (29)$$

which results in (17) being further simplified to

$$\begin{aligned} k\left(\frac{1}{2}\right) &= \frac{1}{2}\ln \frac{|\mathbf{\Sigma}_s(A)||\mathbf{\Sigma}_s(B)|}{\sqrt{|\mathbf{\Sigma}_1(A)||\mathbf{\Sigma}_1(B)||\mathbf{\Sigma}_0(A)||\mathbf{\Sigma}_0(B)|}} \\ &+ \frac{1}{8}(\mu_{0A} - \mu_{1A})^2 \sum_{i=1}^{N_1}\sum_{j=1}^{N_1} \left[\mathbf{\Sigma}_s^{-1}(A)\right]_{ij} \\ &+ \frac{1}{8}(\mu_{0B} - \mu_{1B})^2 \sum_{i=1}^{N_2}\sum_{j=1}^{N_2} \left[\mathbf{\Sigma}_s^{-1}(B)\right]_{ij} \ . \quad (30) \end{aligned}$$

Notice that instead of explicitly computing $\mathbf{\Sigma}_s^{-1}$, we simply need the sum of the elements of the inverse matrix. To this end, we employ a simple result by Wilansky [24] 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 $\mathbf{\Sigma}(A)$ can be simplified as

$$\sum_{k=1}^{M} \mathbf{\Sigma}_{nk}(A) = \frac{\sigma_A^2(1-\varphi^M)}{(1-\varphi^2)(1-\varphi)} + \sigma^2 \quad (31)$$

using the identity

$$\sum_{n=0}^{m} x^n = \frac{1-x^{m+1}}{1-x} \quad \text{for } x \neq 1 \ . \quad (32)$$

Thus we can compute

$$\sum_{i=1}^{N_1}\sum_{j=1}^{N_1} \left[\mathbf{\Sigma}^{-1}(A)\right]_{ij} = N_1 \left[\frac{\sigma_A^2(1-\varphi^{N_1})}{(1-\varphi^2)(1-\varphi)} + \sigma^2\right]^{-1} \quad (33)$$

for a single block of the covariance matrix $\mathbf{\Sigma}(A)$. We finally obtain the required continuous-valued objective function by substituting (24) and (33) into the expression in (30). The resulting closed-form expression allows us to evaluate an approximately optimal allocation of measurements without performing an exhaustive search. As this computation executes significantly faster than an exhaustive search, it makes possible the real-time allocation of measurements amongst the heterogeneous sensors. We are currently working on deploying this allocation scheme on the Nokia N95.

4.1 Scalability

Although this paper only considers the binary hypothesis testing problem with two sensors, our results are easily extended both to incorporate multiple sensors and to discriminate amongst multiple hypotheses. Multiple sensors are included by expanding the block diagonal structure of the covariance matrix in (5), wherein the j-th block is of size $N_j \times N_j$. To incorporate multiple hypotheses, we employ a union bound and thus consider pairwise Bhattacharyya bounds, given in (10) for $k = 1/2$, over all the hypotheses considered. Thus, the optimization problem for M hypotheses and K sensors can be expressed as

$$\min_{N_1,\ldots,N_K} \ P(\epsilon) = \sum_{i<j} e^{-k_{ij}(1/2)} \ \text{ subj to } \sum_{k=1}^{K} N_k = N, \quad (34)$$

where $k_{ij}(1/2)$ is the Bhattacharyya bound specified in (17) for hypotheses H_i and H_j. We note that the complexity of the continuous-valued vector optimization problem derived above is significantly lower than an optimization via grid search, especially when the multiple sensor, multiple hypotheses problem is considered. The optimization problem for this more general framework is under preparation; meanwhile, we develop the optimal allocation of measurements using our simplified model for specific case studies from our pilot study in Section 5.

5. PERFORMANCE ANALYSIS

In this section, we illustrate the utility of optimal measurement allocation using Kullback-Leibler (KL) distances between the distributions associated with each of the two hypotheses as a metric. Case studies based on experimental data collected in our pilot study [2] are also presented.

Recall that the KL divergence between two distributions P and Q is a measure of the difference between the two probability distributions. Since the KL divergence is non-commutative, we use the symmetric Kullback-Leibler (SKL) distance, which in the case of two Gaussian distributions $\mathcal{N}(\mu_P, \sigma_P^2)$ and $\mathcal{N}(\mu_Q, \sigma_Q^2)$ is defined as [6]

$$D_{\text{SKL}}(P||Q) = \frac{\sigma_P^2}{\sigma_Q^2} + \frac{\sigma_Q^2}{\sigma_P^2} + (\mu_P - \mu_Q)^2\left(\frac{1}{\sigma_Q^2} + \frac{1}{\sigma_P^2}\right), \quad (35)$$

to quantify the difference between the distributions associated with the hypotheses for each of the features/sensors. Since the magnitude of the symmetric KL distance is proportional to the ability to distinguish between two distributions, we expect the sensor with the larger SKL distance to be more significant in discriminating between the binary hypotheses. In other words, a larger SKL distance corresponds to a lower probability of error, as the distributions associated with the hypotheses are easier to differentiate. Given means and variances for certain activity-levels, the optimal allocation of measurements between the accelerometer and heart monitor is the solution to the continuous-valued vector optimization derived in the previous section.

5.1 Numerical Simulations

To exemplify minimizing the probability of error via the optimal allocation of measurements, we consider the following simple scenario: given two hypotheses H_1 and H_0 and two sensors A and B, we denote $\mathcal{N}(\mu_{1A}, \sigma_{1A}^2)$ as the probability distribution associated with hypothesis H_1 for Sensor A. The distributions associated with Sensor A remain fixed with SKL= 2.5 for both hypotheses, as do distributions associated with hypothesis H_0 for Sensor B. We increase the value of μ_{1B}, thus separating the densities associated with the two hypotheses for Sensor B, in order to investigate the effect of the SKL on the optimal measurement allocation scheme.

Consider a case wherein the symmetric KL distance is incrementally increased from 2.5 to 5.5 for Sensor B. For different scenarios of distributions associated with the hypotheses for the two sensors A and B, we compare the probability of error for equal allocation of measurements and the optimized allocation of measurements. Figure 3 plots the Bhattacharyya bound of the probability of error for the equal allocation of measurements (dashed lines) versus the optimal allocation of measurements (solid lines) for both

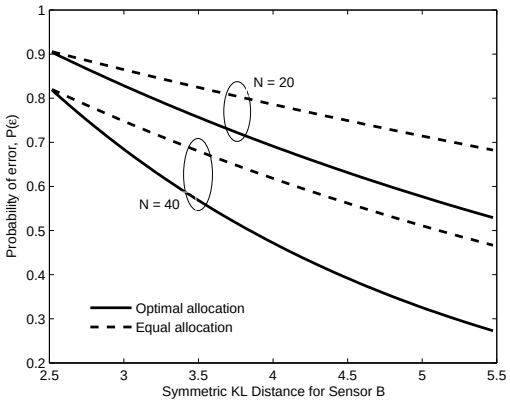

Figure 3: Comparing the optimal allocation to equal allocation as SKL distance between distributions for Sensor B increases.

$N = 20$ and $N = 40$ total measurements. When the SKL distances for the two sensors are equal, equal allocation of the measurements between sensors A and B is optimal. As the SKL distance for sensor B increases while the SKL distance for sensor A stays constant, the optimal allocation of measurements is uniformly better than equally allocating measurements amongst sensors, and the gain due to the optimal allocation increases as the SKL distance for sensor B gets larger. We also find that the probability of error decreases as a greater number of measurements are available.

5.2 Case Studies

The case studies presented here are based on experimental data collected during our pilot study [2]. The initial deployment during this study was with three graduate and two undergraduate student test subjects. Subsequently, data from two child subjects was collected in the Exercise Physiology Lab at the USC Keck School of Medicine. Subjects were fitted with the Alive Technologies heart-monitor and accelerometer at the hip, with electrode cables connecting the heart-monitor unit to standard electrode pads adhered to the chest. The features extracted from the accelerometer and heart-monitor time-series data were the variance of the accelerometer time-series (as shown in Figure 2) and the ECG time-period, respectively. All seven subjects were asked to perform each of the following activities: Lying, Sitting, Sitting&Fidgeting, Standing, Standing&Fidgeting, Walking, and Running. The protocol followed was a modification of Puyau [17] and McKenzie [15], and required subjects to perform each activity for a 10-minute period, with 6 minutes of rest between activities.

We note that the parameters for the Gaussian distributions associated with the different sensors for each of the hypotheses are unique to each of the test subjects, so that as described earlier, a dedicated training period is required to estimate these parameters for a subject before the optimal allocations can be determined. Thus, the case study presented in this section is specific to a particular subject.

In this section, our experimental data is used to obtain means and variances of the distributions corresponding to each of the hypotheses for both sensors. For example, the distributions associated with the Run/Walk and Sit&Fidget

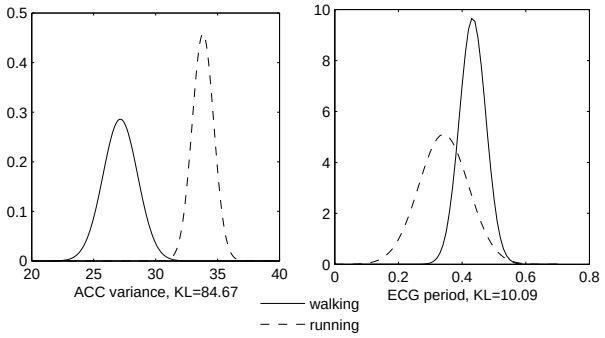

Figure 4: Densities associated with the Walking and Running hypotheses for ACC and ECG sensors.

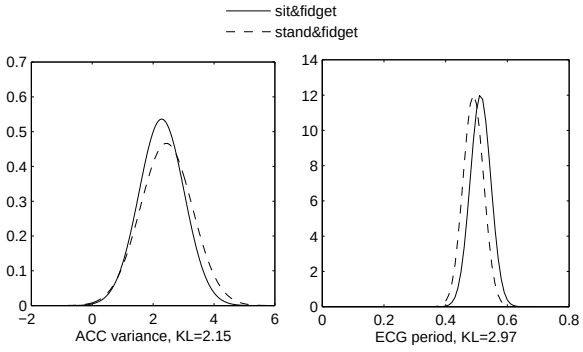

Figure 5: Densities associated with the Sit&Fidget and Stand&Fidget hypotheses for ACC and ECG sensors.

/ Stand&Fidget hypothesis tests for both the ACC and ECG sensors are plotted in Figures 4 and 5, respectively.

The activities listed above are grouped into the following activity-levels:

$High$:	Walking, Running
$Medium$:	Sitting&Fidgeting, Standing&Fidgeting
Low :	Lying, Sitting, Standing

We consider binary hypothesis tests both within the same activity-level and spanning across different activity-levels. The number of measurements allocated to the ACC and ECG sensors are denoted N_1 and N_2, respectively. The continuous-valued objective function, derived in (30), is minimized by varying N_1 and N_2 such that $N_1 + N_2 = N$. The probability of error is computed for all values of $N_1 \in [1, N-1]$, and the optimal allocation of measurements corresponds to the minimum probability of error.

The bivariate Gaussian classifier, described in [2], is used to discriminate between any two hypotheses. As the use of this detector results in a >95% probability of accuracte detection for both the intra-level and inter-level binary hypothesis tests, we investigate the necessity of each of the sensors to accurate detection. In other words, if sensor A is significantly better at discriminating between two hypotheses than sensor B, we expect the probability of error when all measurements are allocated to the sensor A to be minimal.

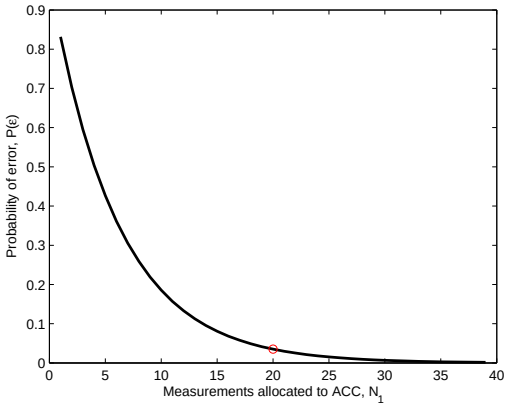

Figure 6: $P(\epsilon)$ for increasing number of measurements allocated to the ACC sensor for Walk vs Run hypothesis test ($\circ$ denotes equal allocation).

Table 1: Optimal allocation of measurements between ACC and ECG for binary hypothesis tests.

Hypothesis Test	Allocation	% Improvement
Run/Walk	100% ACC	>95%
Sit&Fidget/ Stand&Fidget	Any allocation	<1%
Between low activities	100% ECG	5%-10%
High/Medium	100% ACC	>95%
Medium/Low	100% ACC	60%-80%
High/Low	100% ACC	>95%

This means that collecting N measurements from sensor A results in a probability of error that is equal to that achieved by collecting N measurements from *each* of the two sensors. This implies that fewer total measurements may be collected to achieve the same performance, thus yielding an energy-aware application.

Using the continuous-valued vector optimization derived in the previous section, we compute the probability of error for all possible measurement allocations for the Run/Walk hypothesis test. The $P(\epsilon)$ is plotted in Figure 6 as a function of N_1, and we find that the minimum probability of error is achieved when all measurements are allocated to the accelerometer. The distributions associated with the Run and Walk hypotheses for each of the sensors, shown in Figure 4, indicate that the SKL for the ACC sensor is significantly greater than the SKL between the distributions for the ECG sensor. Thus, given our data, we expect it is optimal to allocate all measurements to the accelerometer.

The optimal allocation of measurements is tabulated in Table 1 for binary hypothesis tests both within activity-levels (first three rows) and spanning different activity-levels (last three rows) using the ACC and ECG sensors. For example, the "Medium/Low" entry in Table 1 includes the Sit / Sit&Fidget and Stand / Stand&Fidget hypothesis tests. The third column indicates the percentage by which the probability of error is reduced, in comparison to the equal allocation scenario, by optimal allocation of measurements. A >95% improvement, seen when discriminating between High/Medium and High/Low activities, corresponds to the probability of error approaching zero when all measurements are allocated to the accelerometer. Thus, we find that significant gains result from employing only the accelerometer when discriminating between two hypotheses *if* one or more of the hypotheses is a high-level activity; *i.e.* collecting $1.1N_1$ accelerometer measurements is approximately equivalent to collecting N_1 ACC *and* $N_2 = N_1$ ECG measurements. On the other hand, we find that the ECG sensor is better at distinguishing between low-level activities, but to a lesser degree. The probability of error when optimally allocating all measurements to the ECG sensor is reduced by 5%-10% in comparison to equally allocating measurements between the accelerometer and the heart monitor. Thus, collecting $1.9N$ measurements from the ECG sensor is approximately equivalent to collecting N measurements from each sensor.

In contrast to the above situations, there is no clear optimal allocation of measurements for the Sit&Fidget / Stand& Fidget hypothesis test. As seen in Figure 5, the SKL distances for the two sensors are comparable, so that any allocation of measurements only changes the probability of error by <1%. Thus any allocation of measurements, including the equal allocation between the two sensors, is approximately optimal.

5.3 System Implementation

The optimal allocation of measurements as described in this paper is only a theoretical model at this time. We are currently in the process of testing its implementation on the system described in Section 3 above, which includes the Nokia N95 as the fusion center that collects measurements from the other sensors via the Bluetooth protocol.

Our proposed implementation would involve modifying the software program on the Nokia N95, written in Python for s60 devices, which is multi-threaded, wherein each thread is waiting for data from a particular sensor to become available. Thus, the optimal allocation of measurements is implementable on the Nokia N95 by detecting the activity on-the-fly on the phone and subsequently disconnecting the sensor whose measurements are not useful for discriminating between the current set of activities, and re-connecting to it when the activity-level changes. The optimal allocation scheme is implementable in real-time on the Nokia N95 because the continuous-valued vector optimization can execute faster than an exhaustive search, which would be too processor and time-intensive if multiple hypothesis and multiple sensors with high sampling rates were considered. On the other hand, activity-level detection could also be done on an external server in real-time, which could alert the phone as to which sensors to disconnect. Using this approach would enable us to implement more sophisticated algorithms for real-time activity detection.

6. CONCLUSIONS AND FUTURE WORK

Current approaches to activity-level detection encompass both advanced classification capabilities for use in a variety of health-monitoring applications and energy-saving strategies, since battery life is often a limiting resource in long-term deployment of wireless body-area networks. In this paper, we have shown that an optimal allocation of measurements can achieve better performance when compared to allocating equally between sensors. Furthermore, the fact that measurements from a particular sensor are not required

in some scenarios results in an energy-efficient system, since the Bluetooth communication between that sensor and the fusion center can be turned off.

We have found that the optimal allocation of measurements reduces the probability of error by varying degrees for specific binary hypothesis tests, and that an approximately optimal solution can be obtained using a low-complexity continuous-valued vector optimization which does not require an exhaustive search. We are currently extending our analysis to incorporate multiple sensors and multiple hypotheses via a union bound, and are developing a real-time implementation of our optimal allocation scheme on the KNOWME network infrastructure. Furthermore, more data is being collected which we hope to present at the conference.

7. ACKNOWLEDGMENTS

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

8. REFERENCES

[1] Wireless Health Monitors from Alive Technology, www.alivetec.com, Retrieved on July 24, 2008.

[2] M. Annavaram, N. Medvidovic, et al. Multimodal sensing for pediatric obesity applications. In *Proceedings of International Workshop on Urban, Community, and Social Applications of Networked Sensing Systems (UrbanSense08)*, Raleigh, NC, November 2008.

[3] A. Benbasat and J. Paradiso. A framework for the automated generation of power-efficient classifiers for embedded sensor nodes. In *Proceedings of SenSys*, Sydney, Australia, November 2007.

[4] S. Biswas and M. Quwaider. Body posture identification using hidden markov model with wearable sensor networks. In *Proceedings of BodyNets Workshop*, Tempe, AZ, March 2008.

[5] 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.

[6] T. Cover and J. Thomas. *Elements of Information Theory*. John Wiley and Sons, Inc., New York, NY, 1991.

[7] F. Dabiri, H. Noshadi, H. Hagopian, et al. Light-weight medical bodynets. In *Proceedings of 2nd International Conference on Body Area Networks (BodyNets 2007)*, Florence, Italy, June 2007.

[8] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley, New York, 2001.

[9] 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.

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

[11] S. Jiang, Y. Cao, S. Iyengar, et al. Carenet: An integrated wireless sensor networking environment for remote healthcare. In *Proceedings of BodyNets Workshop*, Tempe, AZ, March 2008.

[12] 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.

[13] 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.

[14] S. Kang, J. Lee, H. Jang, et al. Seemon: scalable and energy-efficient context monitoring framework from sensor-rich mobile environments. In *Proceeding of 6th International Conference on Mobile Systems, Applications, and Services (MobiSys)*, Breckenridge, CO, June 2008.

[15] T. McKenzie, J. Sallis, and P. Nader. Sofit: System for observing fitness instruction time. *Journal of Teaching in Physical Education*, 11:195–205, 1991.

[16] C. Ogden, M. Carroll, and K. Flegal. High body mass index for age among us children and adolescents, 2003-2006. *Journal of the American Medical Association*, 299(20), May 2008.

[17] M. Puyau, A. Adolph, F. Vohra, and N. Butte. Validation and calibration of physical activity monitors in children. *Obesity Research*, 10:150–157, 2002.

[18] 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.

[19] P. J. Sherman. Circulant approximations of the inverses of toeplitz matrices and related quantities with applications to stationary random processes. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(6):1630–1632, December 1985.

[20] E. Shih, P. Bahl, and M. Sinclair. Wake on wireless: an event driven energy saving strategy for battery operated devices. In *Proceedings of 8th Annual Int'l Conference on Mobile Computing and Networking (MobiCom)*, Atlanta, GA, September 2002.

[21] F.-W. Sun, Y. Jiang, and J. S. Baras. On the convergence of the inverses of toeplitz matrices and its applications. *IEEE Transactions on Information Theory*, 49(1):180–190, January 2003.

[22] H. V. Trees. *Detection, Estimation, and Modulation Theory, Part I*. John Wiley, New York, 1968.

[23] Y. Wang, J. Lin, M. Annavaram, et al. A framework of energy efficient mobile sensing for automatic user state recognition. In *7th Annual International Conference on Mobile Systems, Applications and Services (Submitted)*, June 2009.

[24] A. Wilansky. The row-sums of the inverse matrix. *The American Mathematical Monthly*, 58(9):614–615, November 1951.