A Reliable and Reconfigurable Signal Processing Framework for Estimation of Metabolic Equivalent of Task in Wearable Sensors

Parastoo Alinia, Student Member, IEEE, Ramyar Saeedi, Student Member, IEEE, Ramin Fallahzadeh, Student Member, IEEE, Ali Rokni, Student Member, IEEE, Hassan Ghasemzadeh, Member, IEEE

Abstract—Wearable motion sensors are widely used to estimate Metabolic Equivalent of Task (MET) values associated with physical activities. However, one major obstacle in widespread adoption of current wearables is that any changes in configuration of the network requires new data collection and re-training of the underlying signal processing algorithms. For any wearable-based MET estimation framework to be considered a viable platform, it needs to be reconfigurable, reliable, and power-efficient. In this paper, aim to address the issues of sensor misplacement, power efficiency, and new sensor addition and propose a reliable and reconfigurable MET estimation framework. We introduce a power-aware sensor localization approach that allows users to wear the sensors on different body locations without need for adhering to a specific installation protocol. Furthermore, propose a novel transductive transfer learning approach, which gives end-users the ability to add new sensors to the network without need for collecting new training data. This is accomplished by transferring the knowledge of already trained sensors to the untrained sensors in real-time. Our experiments demonstrate that our sensor localization algorithm achieves an accuracy of 90.8% in detecting location of the wearable sensors. The integrated model of sensor localization and MET calculation achieves an \( R^2 \) of 0.8 in estimating MET values using a regression-based model. Furthermore, our transfer learning algorithm improves the \( R^2 \) value of MET estimation up to 60%.

Index Terms—Metabolic equivalent of task, Physical activities, Transfer learning, Node localization, Sensor misplacement, Motion sensors.

I. INTRODUCTION

Metabolic equivalent of task (MET) is an approximation of energy expenditure and an indicator of the intensity of physical activities. This measurement is commonly used to assess performance of physical activity interventions associated with many chronic illnesses such as coronary heart disease, type-2 diabetes, and cancer [1]. Healthy lifestyle changes such as diet control and exercise, which maintain a balance between dietary intake and calories burned, are key approaches in reducing complications due to these diseases [2]. This requires real-time tracking of physical activities that individuals at high risk of chronic diseases perform daily [3]. There are several approaches to calculate food intake and level of physical activity, including traditional self-reported questionnaires, indirect calorie meters, doubly labeled water techniques, and electrocardiographs [3], [4]. In recent years, however, accelerometers, gyroscopes, pressure sensors, and heart rate monitors have been used for physical activity detection and energy expenditure calculation [5]–[7] due to their small size, portability, low-power consumption, and low cost [3], [4].

Accelerometers have been widely used to estimate energy expenditure and MET of physical activities [3]–[5], [8]. Although, the current approaches for estimation of MET values using wearable sensors have proven to be accurate [5], [8], they do not take two important issues into consideration for deployment in real-world settings. First, users would naturally tend to add new wearable sensors to the network as new sensors such as smart watches, ankle bracelets and necklaces, become available. Current research requires collection of new labeled training data for the purpose of algorithm development when a new sensor is added to the system. Second, users prefer to carry their mobile devices on various body locations, resulting in a displacement of the sensor [9]. Therefore, claimed accuracy of current MET estimation systems (MES) is dependent on adhering to the deployment protocols; for example, users must wear sensors on predefined body locations. One issue with real time monitoring of the location of the wearable sensors is excessive power usage and frequent need to charge the multiple sensor nodes. To address this issue, power optimization should be considered in different design levels [10]. An important aspect of the low-power system level design and optimization in wearables is to develop efficient signal processing and data reduction algorithms that reduce computation load of the processing units, allowing low-cost processors to be embedded with the wearable device. These requirements are limiting practical use and potentially imposing discomfort for end-users. In order to make wearables of the futures more reliable and reconfigurable, the underlying MET estimation model needs to be updated upon changes in the wearable network [11].

In this paper, we propose a framework to address reliability and reconfiguration challenges of wearable sensor networks. First, an approach is proposed to compensate unreliability due to change of on-body sensor location while taking into account the computation complexity of the devised sensor localization algorithms. Second, a novel transfer learning algorithm is developed to adopt the knowledge of existing nodes in a new configuration of the network. The result is a reliable, power-efficient, and reconfigurable MET calculation system that allows users to change the location of the sensors or
add new sensors to the network in real-time without any need for new data collection or re-training of the underlying signal processing and machine learning algorithms. To the best of our knowledge, the impact of sensor localization and transfer learning on calculation of MET has not been studied previously. Our specific contributions in this paper are as follows: (1) we propose a framework for estimating MET numbers and activity levels to address unreliability and reconfigurability challenges in wearables; (2) we develop a power-aware sensor localization algorithm based on machine learning techniques to automatically detect the location of the wearable sensors; (3) we propose a transductive learning approach to transfer the knowledge of existing trained wearable sensors to a newly added untrained sensor; (4) we develop regression-based algorithms for estimating the MET values of physical activities without need for collecting new training data; (5) we assess the performance of our individual algorithms as well as the entire framework using real-data collected in two experiments involving both daily physical activities and fitness movements.

II. PRELIMINARIES AND RELATED WORK

Metabolic Equivalent of Task is the ratio of the work metabolic rate to the resting metabolic rate. One unit of MET is defined as 1 kcal/kg/hour which is approximately equal to the energy cost of sitting quietly, equivalent to 3.5 ml/kg/min. In other words, MET is an energy expenditure measurement that demonstrates the intensity of physical activities [1]. Our literature review in this section includes a discussion on how gold standard MET values are computed in clinical practice, followed by the state-of-the-art on estimating MET numbers using wearable sensors, and related research on sensor localization and transfer learning.

A. Gold Standard MET Computation

In recent studies, two approaches have been used to compute the gold standard MET values of physical activities. First, users can look up the MET values from the Compendium, which contains MET value of almost 300 daily activities in a table [12], [13]. In spite of the ease of use and wide range of physical activities in the Compendium, there are several limitations to this approach. The major problem is that the users need to perform physical activities in a fully controlled environment with knowledge about the participants' age, weight, and gender. Therefore, the true energy cost for an individual may or may not be close to the stated mean MET level presented in the Compendium. In the second approach, researchers use a metabolic cart to compute MET values. They measure volume of oxygen breathed into the lungs while performing various activities. The following equation has been used in previous studies to compute the actual MET corresponding to each activity [3]:

\[
MET = \frac{VO_2}{f \times m}
\]

where \(VO_2\) (\(ml/min\)) and \(m\) denote the oxygen uptake volume and the mass of the user in kilograms, respectively. The symbol \(f\) represents a factor that depends on the general fitness features of the group participated in the experiment [14]. The main problem users may encounter while using this approach is that the metabolic cart is a heavy and expensive device and an oxygen uptake mask need to be worn while performing the activities. Thus, the utility of the metabolic cart is limited to indoor usage and constrained activities.

B. Sensor-based MET Estimation

To address the difficulties in calculating gold standard MET values, in recent studies, accelerometers have been used to estimate energy expenditure and MET of physical activities [3]–[5], [8]. This approach requires a training phase to develop an MET estimation model. Usually, several accelerometers are placed on different locations on the body and an estimation model (e.g., a regression model) is developed based on the features extracted from the acceleration signals and the gold standard MET numbers. In one application, researchers developed regression models to estimate MET values when playing a soccer game [3]. It demonstrates that the MET value of soccer exergaming movements can reach a value of 7, which is a standard value for actual casual intensity soccer. The authors in [5] compare a wearable multi-sensor with a single-sensor approach for energy expenditure estimation. The results show that a wearable multi-sensor approach outperforms the single-sensor solution using ActiGraph GT3X+ and linear regression. Another study proposes two MET estimation methods, one traditional single and multiple regression models, and one mono-exponential MET estimation method [4]. In all the aforementioned studies, the location of the wearable sensors is fixed and a retraining of the MET estimation model is needed if a new sensor is added to the network.

C. Unreliability Mitigation and Power Awareness

There have been several research efforts on sensor localization for wearable sensors. Those studies can be divided into three categories as follows: (1) sensor placement on different body locations (e.g., back pocket of trousers versus side pocket of jacket) [15]–[20]; (2) sensor displacement within a given coarse location (e.g., shifting from top upper arm to middle upper arm) [21], [22]; and (3) changes in the orientation of the sensors [23], [24]. In particular, there exist several recent studies that develop localization techniques based on machine learning algorithms [11], [15]–[17], [20]–[22], [25], [26]. However, none of these studies investigated sensor localization in the context of MET estimation.

Sensor misplacement can dramatically decrease the accuracy of MET estimation. There has been minimal effort in developing sensor localization algorithms for MET calculation. In one study, authors use a wearable sensor to estimate energy expenditure without relying on the prior knowledge about the location of the sensor [27]. The system, however, uses a single sensor on three predefined locations on the body. Three different models are built for these body-locations. Our goal in this paper is to develop a global model of sensor localization where the developed machine learning algorithms can be used uniformly on all the sensors independent of their on-body location. We construct one model for all the activities users perform while wearing multiple sensors at the same time. The
localization accuracy and MET estimation model in the prior research is based on one activity (i.e., walking).

Efficient power consumption in wearable systems is an important design consideration [10], [28]. There exist several studies on energy efficient design of physical activity monitoring applications using wearable sensors. In one study, authors proposed an energy efficient sensor coverage for physical movement monitoring [29]. The proposed solution is capable of eliminating redundant sensor nodes, while maintaining a lower bound of activity recognition accuracy. Researchers have also proposed a genetic programming-based feature selection algorithm for activity recognition systems [30]. The goal is to find a set of discriminative and variation tolerant features that may reduce the energy requirements of the wearable sensor system and to enhance the robustness of the activity recognition solution. In another study, researchers presented a gesture recognition system that minimizes power consumption while maintaining a run-time application defined performance target through dynamic sensor selection [31]. Using this technique, the network lifetime is extended 4 times while the accuracy remains approximately the same.

D. Transfer Learning

To the best of our knowledge, there has been no prior study on developing transfer learning approaches for MET estimation. However, researchers have conducted much research on transfer learning in the field of machine learning and artificial intelligence. Transfer learning approaches can be divided into three sub-categories of inductive, transductive and unsupervised based on different conditions of the source and target domains and the knowledge extraction tasks [9]. In inductive transfer learning [32], [33], the target and source tasks are different from each other and labeled data are available in the target domain. In transductive transfer learning [34]–[36], the source and target tasks are the same, while their domains are different. In this case, labeled data exist in source domain but not in the target view. In unsupervised transfer learning, similar to inductive transfer learning setting, the target task is different but related to the source task, and there exist no labeled data available in either source or target domain [37]–[39].

In [34], authors derived efficient transductive transfer learning algorithms based on a Support Vector Machine (SVM) paradigm of a large margin hyperplane classifier on a feature space. In another study [35], authors introduce a maximum entropy based technique, called Iterative Feature Transformation (IFT), and show that it achieves comparable performance with the state-of-the-art transductive SVMs in solving the problem of transfer learning for protein name extraction.

III. System Architecture and Methods

We propose a reliable and self-adaptive framework to address the challenges of sensor misplacement, power efficiency, and automatic sensor addition in wearable networks, with a special focus on MET estimation applications. We first develop a reliable MET calculation technique that estimates the MET values of physical activities without requiring the sensors to be placed on predefined locations while minimizing the energy consumption due to feature extractions. We then introduce a novel transfer learning algorithm to allow for addition of new sensors without re-training of the MET estimation model. We define a system configuration \( C(k, l_0, l_1, \ldots, l_k, t) \) as the number of the sensors \( k \), their location on the body \( l_0, l_1, \ldots, l_k \) and whether we know the MET values of physical activities \( t = 1 \) or not \( t = 0 \). Each configuration of the system illustrates a different situation. For example, when a user wears two trained accelerometer sensors, one on the chest, and other on the waist, the configuration of the MES (MET Estimation System) will be \( C(2, \text{chest}, \text{waist}, 1) \). Since in MES, the MET labels are independent of the location of the sensors, we won’t see any configuration with mixed type of untrained and trained sensors.

A. Reliable MET Calculation

In defining reliability in the context of MET calculation, we focus on two properties of the system: (1) changing the location of the sensors should not have a significant effect on the accuracy of MES; and (2) the system need to be power efficient in terms of sensing and feature computation complexity. The goal of our location-independent MET calculation is to demonstrate how detecting the location of the sensors before estimating the MET values would increase the accuracy of the underlying linear regression equation in estimating MET values of physical activities. The power-aware feature selection selects the set of optimal features taking into account their information relevance and computation complexity.

Our reliable MET estimation model contains five main steps: (1) Feature extraction, which computes a set of representative features from the signals captured using wearable accelerometers; (2) Power-aware sensor localization, which is a constructed machine learning algorithm based on the optimal selected features; (3) Label assignment, which assigns predicted location labels to the sensor readings based on the output of the sensor localization algorithm; (4) Linear regression, which uses the extracted features from wearable sensors to estimate the MET values. As shown in Fig. 1, in the training phase, the extracted features from the accelerometers labeled with their locations are fed into a classifier to build a location detection model. A linear regression model is then constructed on the magnitude of the accelerometer signals to estimate MET numbers. (5) Activity level classification, which classify the features extracted from the sensor signals into three groups of light, moderate and vigorous activities based on mapping the exact MET numbers from the Pate model.

1) Power-Aware Sensor Localization: potentially, there are many different features that can be extracted from human activity signals coming from a variety of sensors (e.g. accelerometers, and gyroscopes). Statistical features have shown effectiveness in human activity recognition based on previous studies [40]. For node localization purposes, however, it is rather unknown which features are most effective. Thus, we aim to explore most effective and power-efficient features for node localization. For this purpose, we first extract an exhaustive set of potentially useful features; then minimize the number of features for each sensor to decrease the computation
Our goal is to minimize the power consumption due to computation and sensing in our algorithm. We computed the energy dissipation to sense the data segment as 5.28mJ for each accelerometer according to the work in [11], [41]. The set of features extracted from individual sensor streams and corresponding power consumption for each feature is shown Table I. This list includes ten statistical features and ten morphological features for each data segment. We calculated energy consumption for each feature based on the work in [11], [41].

**TABLE I: Extracted features from sensor signals**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
<th>Energy(nJ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMP</td>
<td>Amplitude of Signal Segment</td>
<td>16386</td>
</tr>
<tr>
<td>MED</td>
<td>Median of the Signal</td>
<td>405159</td>
</tr>
<tr>
<td>MINVALUE</td>
<td>Mean Value of the Signal</td>
<td>8126</td>
</tr>
<tr>
<td>MAX</td>
<td>Maximum Value of Signal</td>
<td>8103</td>
</tr>
<tr>
<td>MIN</td>
<td>Minimum Value of Signal</td>
<td>8108</td>
</tr>
<tr>
<td>P2P</td>
<td>Peak to Peak Amplitude</td>
<td>16291</td>
</tr>
<tr>
<td>STD</td>
<td>Standard Deviation</td>
<td>38846</td>
</tr>
<tr>
<td>VAR</td>
<td>Variance</td>
<td>40431</td>
</tr>
<tr>
<td>RMS</td>
<td>Root Mean Square Power</td>
<td>29705</td>
</tr>
<tr>
<td>S2E</td>
<td>Stand to End Value</td>
<td>83</td>
</tr>
<tr>
<td>MORPH</td>
<td>Morphological Features</td>
<td>45</td>
</tr>
</tbody>
</table>

The input of the classification module is an exhaustive set of features extracted from on-body sensors with their location as labels. We used the greedy algorithm proposed in [11], to select the optimal set of features. In this algorithm, on-body sensor localization is considered as a classification problem which seeks the most prominent features in terms of power consumption from the large set of feature pool, while retaining the localization accuracy at a minimum level. The algorithm ranks the entire set of the features based on the product of computation and the rank of feature, which is computed based on the ranker algorithm implemented in Weka [42]. Afterwards, the number associated with each feature is normalized; and in each step the feature with highest rank rate is added to the target set which is initially empty. The accuracy of the classification is computed based on the selected features in the target set in each step. If the accuracy is improved, the new feature is selected, otherwise, it will be removed from the target set. This process continues as the accuracy reaches a threshold value. The output of is the set of features are those that are most informative and power-efficient among all features in the entire set.

2) *MET Value Estimation:* the feature vector with predicted location label is passed to the linear regression module as input. A linear regression model is then developed to fit the best line on features as input and the MET values as output as given by:

\[
L = \sum_{i=0}^{k-1} A_i f_i
\]

where \(A\) denotes the corresponding coefficients for one feature \(f_i\) (i.e., ‘magnitude’). The symbol \(k\) represents the number of sensors in the network.

3) *Activity Level Classification:* our MET model estimates the activity levels (i.e., the activity intensities: light, moderate, and vigorous). Using the model proposed by Pate et al. [43], the physical activities can be classified into three groups of light activities with < 3 METs, moderate with 3 – 6 METs, and vigorous with > 6 METs. The input to the activity level estimation unit is the set of extracted features and their MET labels. In the first step, the MET numbers are mapped into three groups of light, moderate and vigorous using the Pate model. Then a classification model is built on the feature set and new MET level labels. We set the labels 1, 2, and 3, respectively, for the three activity levels.

**B. The Reconfigurable MET Estimation System**

The objective of the reconfiguration unit is to transfer knowledge from the existing trained sensors in the network to a newly added (untrained) and adapt to the new configuration of the system. Based on the common analogy in machine learning domain, we will call the new sensor, the target sensor, and all other existing sensors in the network, the source sensors. Our approach, shown in Algorithm 1, consists of two main phases of sensor selection and MET transfer. Parameter \(S_{source}\) denotes the set of input feature datasets with MET.
labels extracted from the source sensors. Symbol $s_{\text{target}}$ is the target sensor feature set without MET labels. In the first phase, to avoid the possible negative transfer learning effects, the function MaxCorrelation selects most similar source sensor to the target sensor ($s_{\text{target}}$) using a correlation factor. The value of this metric ranges from 0, for entirely dissimilar sensors, to 1, for sensor with highest correlation. Function $\gamma$ computes the correlation factor between two arbitrarily sensors in the algorithm. Equation (3) calculates the correlation factor between two arbitrarily sensors $s_1$ and $s_2$:

$$\gamma(s_1, s_2) = \sum_{k=0}^{n-1} U(s_1.f_k, s_2.f_k)$$

where $U$ is symmetric uncertainty measure between two feature instance and $n$ denotes the number of the features in the datasets. Symbols $f_k$ represents $k$-the sample in the dataset from each sensor. The symmetric uncertainty between two random variables is given by:

$$U(X,Y) = \frac{2I(X,Y)}{H(X) + H(Y)}$$

where $H(X)$ and $H(Y)$ represents the entropy of random variables $X$ and $Y$, respectively, and $I(X,Y)$ denotes the information gain between them and is computed by:

$$I(X,Y) = H(X) - H(Y)$$

**Algorithm 1: Transfer Label Algorithm**

**Input:** $S_{\text{target}}$ feature set of the new untrained sensor (without MET labels) which has been added to the network of a trained sensors on the body, $S_{\text{source}}$ set of the trained sensors in the body sensor network with MET, $MET_{source}$

```plaintext
1. $minDist \leftarrow +Inf; /* the minimum distance among the pairs of the instances in the feature set of the target and source sensors */
2. $s_{\text{selected}} \leftarrow \text{MaxCorrelation}(s_{\text{target}}, S_{\text{source}})$
3. foreach sample $d_{\text{target}}$ in $s_{\text{target}}$ do
   4.   foreach sample $d_{\text{selected}}$ in $s_{\text{selected}}$ do
      5.     if $\text{Distance}(d_{\text{target}}, d_{\text{selected}}) < minDist$ then
         6.       $minDist \leftarrow \text{Distance}(d_{\text{target}}, d_{\text{selected}})$
         7.       $d_{\text{target.met}} \leftarrow d_{\text{source.met}}$
5. Output: $s_{\text{target}}$ /* the dataset of the new sensor with transferred MET labels */
6. Function MaxCorrelation($s_1, S$)
   7.   $\maxCor \leftarrow -Inf; /* the maximum correlation value between the $s_1$ and the sensors in set of sensors $S$ */$
   8.   foreach sensor $s$ in $S$ do
      9.     if $\text{Correlation}(s_1, s) > \maxCor$ then
         10. $\maxCor \leftarrow \gamma(s_1, s)$
   11. return $s_{\text{selected}}$
```

In the second phase, for each instance in the feature set of target sensor, $s_{\text{target}}$, we perform a Nearest-Neighbor classification to detect the closest instance from the feature set in the selected sensor $s_{\text{selected}}$. Finally, all instances in dataset of the target sensor are assigned with MET labels with the labels from the selected sensor feature set, based on the k-NN classification model which links each instance in $s_{\text{target}}$ to the closest instance in $s_{\text{selected}}$ in a Euclidean feature space given by (6).

$$D(I_1, I_2) = \sqrt{\sum_{k=1}^{n} (I_1.f_k - I_2.f_k)^2}$$

where $I_1$ and $I_2$ are two arbitrarily feature vectors from the sensor datasets. Symbol $f_k$ denotes the $k$th feature in feature vector instances when there are $n$ feature vectors in the dataset. The output of the algorithm is the dataset of the target sensor $s_{\text{target}}$ containing the transferred MET labels.

**IV. DATA COLLECTION AND EXPERIMENTAL PROCEDURES**

We investigate the effect of sensor localization on MET values of the physical activities. We investigate the effectiveness of our transfer learning approach in improving the accuracy of our system in computing MET numbers. We conduct two experiments: (1) exergaming movements [3], and (2) walking on the treadmill. The first dataset is collected from six subjects wearing two accelerometer-based sensors during a clinical experiment. The second dataset includes accelerometer data from three sensors located on the body of 15 participants while walking at three different speeds on a treadmill. Both experiments were approved by Institutional Review Board (IRB).

**A. Exergaming Experiment**

Data was collected from six male subjects, aged between 21 and 30 during a clinical experiment. In the experiment, two data collection modalities were used. Two GCDC ±2g three-axis accelerometers sampling at 50 Hz were attached to the hip and ankle of participants. A metabolic cart that measures the volume of oxygen breathed into the lungs while doing activities was also used to collect ground truth MET values. Since the breathing pattern is not unique, the average oxygen uptake of participants during a 30-second period was reported. We used equation (1) to compute the actual value of MET, corresponding to each activity using the oxygen volume from the metabolic cart. Participants were asked to perform 6 activities (running, sprint, pass, chip, medium shot, and full powered shot). They further were asked to play a simulated game. We allowed the participants to perform each activity for 3 minutes to achieve a steady state in breathing. Subjects were also instructed to have their own desired intensity while performing the activities. Intensity of each action was extracted based on the work in [3], which requires repeating an action for 3 minutes. The data captured by accelerometers are used for sensor localization and MET calculation. We computed the magnitude of each accelerometer sensor data as a feature demonstrating intensity of each action by computing the norm value of each signal sample given by [44]:

$$M = \sqrt{x^2 + y^2 + z^2}$$
Fig. 2: The process of evaluating sensor localization on MET estimation linear model

where $M$ denotes the magnitude of each accelerometer data segment.

In order to synchronize the accelerometer data with $VO_2$ data, we averaged the output signal over a moving window of 3000 samples with one sample overlap. As a result, each axis of the accelerometer signal is averaged over 30 seconds (same as the $VO_2$ data). We calculated the magnitude of each ‘ankle’ and ‘hip’ accelerometer signal segment. The peak of these signal segments, which were collected over time span of 3 minutes, matched to MET value points computed from the metabolic data. Accurate peak detection from noisy signal of the accelerometer, requires an approach similar to Ousto’s proposed in [6], [45] which performs a normalized cross correlation function on the input signal. This method is based on non-linear neighbor filtering in which a sliding neighboring is defined and centered on the signal where the output vector is the local maximum (peak). The peaks of the magnitude of both ankle and hip accelerometers using equation below:

$$\Phi(t - 1) < \Phi(t) > \Phi(t + 1)$$  \hspace{1cm} (8)

where $\Phi(t)$ denotes the feature (magnitude) from accelerometer signal at time $t$. Fig. 3 shows the output of the cross-correlation function on magnitude of ankle accelerometer signal from one participant. We detected seven peaks from each participant corresponding to seven physical activities performed by subjects during the experiments.

B. Treadmill Experiment

In a second experiment, we conducted a two-month experimental study with 15 young healthy adults, five females, and ten males, aged from 21 to 33. As shown in Table II, each participant was asked to walk at three different speeds for a duration of five minutes each.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Walking on Treadmill at 1.5 Mph</td>
</tr>
<tr>
<td>2</td>
<td>Walking on Treadmill at 3 Mph</td>
</tr>
<tr>
<td>3</td>
<td>Walking on Treadmill at 4.5 Mph</td>
</tr>
</tbody>
</table>

TABLE II: Activities performed during the treadmill experiment

Three accelerometer-based motion sensors were used to collect three axis accelerometer signals from the subjects during the experiment. Participants were asked to wear sensor nodes on specific on-body locations (right hand, chest and left jacket pocket) as shown in Fig. 4.

Fig. 4: Three smart phones are placed on three different locations of each participant’s body (chest, right hand, and left jacket pocket) to collect accelerometer signals while walking on the treadmill

Based on prior studies, age and gender hardly affect the energy expenditure in common activities of daily living such as walking [46]. Therefore, the gold standard MET values of walking with different intensities were determined from the Compendium data. The MET values of 7.0, 3.5, and 2.0 were
TABLE III: Energy consumption of various configurations

<table>
<thead>
<tr>
<th>Feature Selection</th>
<th>Accuracy</th>
<th>Energy Computation (nJ)</th>
<th>Sensing (mJ)</th>
<th>Total (mJ)</th>
<th>Power-aware saving</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFS 5-NN</td>
<td>90%</td>
<td>56777</td>
<td>5.28</td>
<td>8.11</td>
<td>34.7%</td>
</tr>
<tr>
<td>Ranker</td>
<td>93.5%</td>
<td>48757</td>
<td>5.28</td>
<td>7.71</td>
<td>31.4%</td>
</tr>
<tr>
<td>Greedy stepwise</td>
<td>91.1%</td>
<td>8637</td>
<td>5.28</td>
<td>9.6</td>
<td>44.8%</td>
</tr>
<tr>
<td>Power-aware</td>
<td>92%</td>
<td>270</td>
<td>5.28</td>
<td>5.29</td>
<td>-</td>
</tr>
</tbody>
</table>

selected for walking at speed of 4.5 mph, 3.0 mph and 1.5 mph, respectively [12], [13].

V. Results

In this section, we discuss the performance of the proposed model in terms of its reliability and reconfigurability.

A. Reliable MET Calculation

In this section, the accuracy of our MET calculation approach is compared with a baseline MET calculation method (that does not consider sensor localization). We discuss the results of the sensor localization algorithm and then describe the results of MET calculation. First, as shown in Fig. 2, the optimal set of features are extracted and selected from the accelerometer signals. Next, we utilize the localization algorithm to detect the location of the sensor from which the feature vector is obtained. Finally, the MET value of each activity is estimated using the linear regression model discussed previously.

1) Power-Aware Sensor Localization: the data collected in the exergaming experiment was used to extract 40 features from ‘ankle’ and ‘hip’ sensors. The features were extracted from a moving window of 100 samples (equivalent of 2s of data) on the dataset. A subset of 6 optimal features out of 40 were selected to perform sensor localization. We used WEKA machine learning tool [42] to develop the sensor localization algorithms and to classify the signals into two classes of ‘ankle’ and ‘hip’.

Moreover, the sensor localization was performed on the dataset of treadmill experiment. In the first step, we extracted time frequency features (as mentioned before from each 100 samples), afterwards we performed a power-aware feature selection. Lastly, ANN and 5-NN classification algorithms were applied to the optimal feature set in order to divide them into three groups of ‘wrist’, ‘pocket’, and ‘waist’.

Table III shows energy consumption of various configurations in the treadmill experiment. We compared the power-aware feature selection algorithm with the forward feature selection, ranker and greedy stepwise algorithms in terms of power consumption. As it can be observed from this table, for accuracy thresholds of 90% to 93% for 5-NN classifier and 94% to 97% for the ANN classifier, we acquire total energy saving of 31.5% to 44.8%, respectively, compared to the other feature selection approaches. The reason for the considerable differences among the computation power is that almost all the selected features by the algorithm are the morphological which require very low computation energy comparing to other statistical features.

Furthermore, in Table IV the accuracy of the classifiers in identifying the location of the sensors in both experiments are shown. We were able to classify the sensor signal segments with average accuracy of 94.6% in exergaming experiment and 86.2% in treadmill experiment using 10-fold cross validation method. Between the classifiers, 5-NN shows better superior performance with 97.3% and 99.6%, in the two experiments.

2) MET Value Estimation: in order to find the most accurate model in computing MET values, we performed linear regression on various possible combination of the sensors on ankle and hip. In [3] sum of magnitude of the accelerometer signals was used as a single feature for linear regression. We, however, use the magnitude of the accelerometer signals as two separate features to perform the regression.

Table V compares the $R^2$ statistics and error of different fusion on the sensors such as using a single sensor ankle or hip, using summation of both ankle and hip sensors and using ankle, and hip sensors separately together. As shown, using ankle and hip signals, improves the average $R^2$ value from 0.71 to 0.80 and reduces the amount of the error from 1.37 to 1.15 on average compared to the approach used in [3].

A LOSOCV was performed on the dataset, due to differences in training pattern of each participant. This validation approach provides a generalized and unbiased estimation of the performance of our model and can be implemented very efficiently in the case of linear regression [47]. With 6 subjects in our experiment, in each iteration, we performed leave-one-subject-out validation, then measured mean squared error (RMSE) and normalized mean squared error (NMSE) as validation measurements using following equation:

$$\text{NMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{MET}_i - MET_i)^2}{n}} \times \frac{1}{\lambda}$$  \hspace{1cm} (9)

where MET and $\hat{MET}$ are the actual and estimated MET values respectively. Symbols $n$ and $\lambda$, respectively, are the size and range of the measured value. Table VI lists the result of LOSOCV. The NMSE values in all cases are smaller than 0.12, which demonstrates the acceptable performance of the MET estimation model for a separate dataset from the training data.

3) Activity Level Classification: in addition to estimating the exact MET values of physical activities, we propose a classification model that predicts the intensity level of physical activity (i.e., light with $< 3$ METs, moderate with $3-6$ METs, and vigorous with $> 6$ METs labeled as 1, 2, and 3 in our data set, respectively).

Table VII lists the accuracy and root mean squared value of estimating activity level by conducting the classifiers 5-NN
TABLE IV: Accuracy of sensor localization

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Exergaming</th>
<th>Treadmill</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10-fold</td>
<td>LOSOCV</td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td>RMSE</td>
</tr>
<tr>
<td>5-NN</td>
<td>97.3%</td>
<td>11.7%</td>
</tr>
<tr>
<td>ANN</td>
<td>92%</td>
<td>20.9%</td>
</tr>
<tr>
<td>Average</td>
<td>94.6%</td>
<td>16.3%</td>
</tr>
</tbody>
</table>

TABLE V: The comparison between the accuracy of the regression on different combination of ankle and hip sensors

<table>
<thead>
<tr>
<th>Sub.</th>
<th>Ankle</th>
<th>Hip</th>
<th>Ankle+Hip</th>
<th>Ankle, Hip</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>RMSE</td>
<td>R</td>
<td>RMSE</td>
</tr>
<tr>
<td>1</td>
<td>0.11</td>
<td>3.25</td>
<td>0.71</td>
<td>1.85</td>
</tr>
<tr>
<td>2</td>
<td>0.22</td>
<td>2.5</td>
<td>0.84</td>
<td>1.25</td>
</tr>
<tr>
<td>3</td>
<td>0.63</td>
<td>1.58</td>
<td>0.84</td>
<td>1.02</td>
</tr>
<tr>
<td>4</td>
<td>0.23</td>
<td>2.01</td>
<td>0.60</td>
<td>1.43</td>
</tr>
<tr>
<td>5</td>
<td>0.60</td>
<td>1.02</td>
<td>0.53</td>
<td>1.11</td>
</tr>
<tr>
<td>6</td>
<td>0.76</td>
<td>1.23</td>
<td>0.44</td>
<td>1.17</td>
</tr>
<tr>
<td>All</td>
<td>0.60</td>
<td>1.38</td>
<td>0.32</td>
<td>1.81</td>
</tr>
<tr>
<td>Average</td>
<td>0.45</td>
<td>1.85</td>
<td>0.61</td>
<td>1.37</td>
</tr>
</tbody>
</table>

TABLE VI: Leave-one-subject-out cross validation test

<table>
<thead>
<tr>
<th>Subject</th>
<th>MSE</th>
<th>NMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.41</td>
<td>0.08</td>
</tr>
<tr>
<td>2</td>
<td>0.29</td>
<td>0.06</td>
</tr>
<tr>
<td>3</td>
<td>0.33</td>
<td>0.05</td>
</tr>
<tr>
<td>4</td>
<td>0.59</td>
<td>0.12</td>
</tr>
<tr>
<td>5</td>
<td>0.39</td>
<td>0.11</td>
</tr>
<tr>
<td>6</td>
<td>0.58</td>
<td>0.09</td>
</tr>
<tr>
<td>Average</td>
<td>0.43</td>
<td>0.08</td>
</tr>
</tbody>
</table>

TABLE VII: Accuracy of the estimating the activity level using 5-NN and ANN classification algorithms with LOSOCV

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Accuracy</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-NN</td>
<td>79.8%</td>
<td>0.77</td>
</tr>
<tr>
<td>ANN</td>
<td>81.5%</td>
<td>0.8</td>
</tr>
<tr>
<td>Average</td>
<td>80.65%</td>
<td>0.785</td>
</tr>
</tbody>
</table>

and ANN on the features extracted from the node localization section and mapped MET levels as labels. We were able to estimate the level of the physical activities with average accuracy of 80.65%. Based on the accuracy and RMSE value of each classifier, the ANN shows a slightly better performance than 5-NN in terms of activity level classification.

4) The Impact of Sensor Localization: we extracted a linear regression model form the magnitude of the accelerometer signals and actual MET values computed from the VO2 data. In real life scenarios, there is no guarantee that users wear sensors on predefined locations. Therefore, we integrate sensor localization with MET calculation to develop a robust MET calculation algorithm. In order to compare the result of MET value estimation, first we assume the location of each signal segment is known using localization algorithm from the previous section. Second, we consider one of the most probable mistakes in which user might wear the hip accelerometer sensor on ankle and ankle accelerometer sensor on hip. We apply this replacement by swapping the detected labels for each corresponding signal in ankle and hip accelerometers. As a result, we have two data sets, one is the baseline with estimated MET values for the ankle and hip, and the other one with the same features but swapped MET values for the ankle and hip accelerometers.

Table VIII shows the result of linear regression on each participant, separately. We used the linear regression model to develop a linear formula from the accelerometer signals of each subject in the experiment assuming the location of the sensors is known to the model using our localization algorithm. The R^2 values are reported in Table VIII. There is variability among the models per subject which is originated from different possible intensities of the same activity performed by individuals or distinct patterns of body movements during the same activity per subject. For instance, some people may move their hip during walking while others do not. Based on the results, our linear regression model represents all the subjects with R^2 value of 0.76.

TABLE VIII: R^2 values from linear regression on MET vs Ankle and Hip Accelerometers

<table>
<thead>
<tr>
<th>Subject</th>
<th>R^2</th>
<th>Linear Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.83</td>
<td>11.7μa + 16.9μh − 25.9</td>
</tr>
<tr>
<td>2</td>
<td>0.84</td>
<td>−2.3μa + 30.3μh − 24.8</td>
</tr>
<tr>
<td>3</td>
<td>0.92</td>
<td>2.7μa + 7.3μh − 7.4</td>
</tr>
<tr>
<td>4</td>
<td>0.71</td>
<td>2.2μa + 17.9μh − 18</td>
</tr>
<tr>
<td>5</td>
<td>0.82</td>
<td>1.8μa + 1.7μh − 2.5</td>
</tr>
<tr>
<td>6</td>
<td>0.79</td>
<td>3.0μa + 2.5μh − 3.2</td>
</tr>
<tr>
<td>All</td>
<td>0.76</td>
<td>4.9μa + 4.0μh − 8.0</td>
</tr>
</tbody>
</table>

Since the R^2 correlation coefficient of a regression model only reports how well the line fits to the data, in order to better report the accuracy, we defined an estimation error (EE) parameter by differentiating the actual MET values computed from the metabolic data (MET) and estimated MET values (MET̂) from the regression model, which is shown by ϵ in
The R-squared statistic of linear regression models with sensor localization ranges from 71% to 92%, while it dramatically decreases down to negative values when the location of the sensors is unknown. Negative value for $R^2$ correlation demonstrates a poorly fitted model on the dataset. Based on the computed average estimation error, locating the position of the sensor nodes on different body parts improves the $R^2$ of the MES with a factor of 2.3 comparing to when there is no prior knowledge on the location of the sensors.

Furthermore, we evaluated the effect of the sensor localization in estimating activity levels using the classification model. Table X depicts the result of classifying the features extracted from the activity signals into three groups of light-intensity, moderate and vigorous physical activities in both cases of known sensor location and when there is no information of the location of the sensors.

### B. Reconfigurable Design

We proposed an algorithm to transfer the knowledge of the source sensors to the target sensor(s) in a wearable sensor network. We evaluate the performance of our transfer learning approach on both exergaming and treadmill datasets. In both cases, we consider two system configurations. In the first configuration the user wears all the sensors in the network except one while performing the physical activities with known MET values. Then, in the second configuration, the user decides to add another sensor in a new location on the body, but this time the MET values to the physical activities performed in the experiment, are unknown. We aim to examine if our transfer learning approach is able to estimate the MET numbers in second configuration with an acceptable accuracy, despite the fact that the MET labels are unknown.

In order to show the performance of our approach, we compare its performance against the upper bound and a baseline approach. The upper bound on the MET estimation assumes that we have the actual MET numbers to the physical activities with the second configuration. This approach estimates the MET labels by fitting the linear regression model on the actual MET labels of the second configuration. The baseline approach merely uses the model trained on the first configuration for MET label estimation.

1) **Treadmill Experiment**: In the treadmill experiment, participants wore three accelerometers in three on-body locations (right hand, chest, and left jacket pocket). In the first configuration the user wears two sensors while walking on the treadmill with known MET values from the Compendium. The sensors on the chest and in the left jacket pocket were selected arbitrarily for the first configuration ($C(2, \text{chest}, \text{pocket}, 1)$). In the second configuration ($C(3, \text{chest}, \text{pocket}, \text{waist}, 0)$), the user adds another sensor in the right hand (not selected sensor) while walking on the treadmill, however we don’t have the actual MET values of the walking on the treadmill (0 in the configuration means we have not the actual labels).

Fig. 5 shows the $R^2$ value of MET estimation in four different configuration of sensors in the wearable network from the following scenario.

In the first step of the transfer learning algorithm, we require to detect the most correlated sensor from the sensors in the first configuration (chest and left jacket pocket sensors) to the target sensor (right hand sensor). Based on the correlation factor values between the right hand sensor (target) and other two (source sensors), it is more correlated to the chest rather than the left pocket sensor. Therefore, we perform the transfer learning algorithm on the chest and right hand sensors to estimate the MET labels of the second configuration.

![Fig. 5: Performance comparison of our transfer learning approach versus the baseline and upper bound regression model estimation in treadmill dataset](image)

As shown in Fig. 5, the upper bound on the average $R^2$ squared value of the linear regression on the treadmill dataset, in 1000 iterations, is 79.9%. Without the actual MET numbers, we can estimate them with the existing model. In this case, the $R^2$ value decreases significantly to 44.9%. Using the transfer learning algorithm, the average $R^2$ value increased to 74% which is only 8% drop comparing to the case when we have the actual MET values for the activities in the experiment.

2) **Exergaming Experiment**: In the exergaming experiment, the participants wore two accelerometer sensors, one on the ankle and other on the hip. We compare the $R^2$ value of estimating MET values of the aforementioned three approaches, as shown in Fig. 6. In the first configuration, user wears only one sensor on the ankle while performing the exergaming movements, with known MET values from the metabolic cart ($C(1, \text{Ankle}, 1)$). Later, in the second configuration, user adds another sensor on the hip during the experiment. However, we don’t have the actual MET values of the activities user is performing ($C(2, \text{Ankle}, \text{Hip}, 0)$).

As you can observe in the Fig. 6, using the exergaming dataset of two sensors and actual MET values from the metabolic cart, we were able to estimate MET numbers with average $R^2$ value of 74% in 1000 iterations in our transfer learning algorithm, which is an upper-bound to the demanding $R^2$ value for the transfer learning approach. We
TABLE IX: Comparing $R^2$ values and error of linear regression on MET vs Ankle and Hip Accelerometers with and without sensor localization

<table>
<thead>
<tr>
<th>Subject</th>
<th>$R^2$ with NL</th>
<th>EE with NL</th>
<th>$R^2$ without NL</th>
<th>EE without NL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.83</td>
<td>0.29</td>
<td>0.68</td>
<td>0.30</td>
</tr>
<tr>
<td>2</td>
<td>0.84</td>
<td>0.12</td>
<td>−19.30</td>
<td>0.34</td>
</tr>
<tr>
<td>3</td>
<td>0.92</td>
<td>0.09</td>
<td>−0.10</td>
<td>0.34</td>
</tr>
<tr>
<td>4</td>
<td>0.71</td>
<td>0.18</td>
<td>−14.41</td>
<td>0.96</td>
</tr>
<tr>
<td>5</td>
<td>0.82</td>
<td>0.13</td>
<td>0.34</td>
<td>0.15</td>
</tr>
<tr>
<td>6</td>
<td>0.79</td>
<td>0.20</td>
<td>0.50</td>
<td>0.37</td>
</tr>
<tr>
<td>Average</td>
<td>0.80</td>
<td>0.18</td>
<td>−4.98</td>
<td>0.41</td>
</tr>
</tbody>
</table>

TABLE X: Accuracy of the estimating the activity level using classification

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Accuracy with NL</th>
<th>RMSE with NL</th>
<th>Accuracy without NL</th>
<th>RMSE without NL</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-NN</td>
<td>79.8%</td>
<td>0.77</td>
<td>0.36%</td>
<td>1.02</td>
</tr>
<tr>
<td>ANN</td>
<td>81.5%</td>
<td>0.8</td>
<td>0.47%</td>
<td>0.95</td>
</tr>
<tr>
<td>Average</td>
<td>80.65%</td>
<td>0.78</td>
<td>41.5%</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Fig. 6: Performance comparison of our transfer learning approach versus the baseline and upper bound regression model estimation in exergaming dataset

could compensate the decreased $R^2$ value using the model based on the first configuration to 20%. However, with our transfer learning approach, the $R^2$ value reached to 62%.

VI. DISCUSSION AND FUTURE WORK

This study illustrates how the presence of unreliability in the system (e.g., sensor misplacement) can affect the result of MET estimation using wearable sensors. Sensor misplacement is not the only concern regarding the use of wearables. There are several other unreliability factors such as displacement of the sensors. We are currently working on developing algorithms to resolve the issue of sensor displacement of wearables in the context of MET estimation.

Adding a new sensor to the network of wearable sensors can decrease the $R^2$ value of MET estimation significantly if the new sensor does not learn MET labels properly. Therefore, we are able to improve the increased $R^2$ value by transferring knowledge about MET numbers of the most relevant sensor among the already trained sensors in the network. As a future work, we intend to explore other transfer learning methodologies to further improve the performance of our MET estimation system.

Furthermore, people tend to remove existing sensors in the wearable sensor network while doing physical activities which might decrease the $R^2$ value of the base model significantly. In order to address this issue, our future work involves studying dynamic reconfiguration of underlying signal processing models to compensate for sensor removal.

VII. CONCLUSION

In this paper, a reconfigurable and location-independent MET estimation system (MES) was proposed. Our system detects the location of the sensors using machine learning algorithms based on features extracted from wearable sensors. This work includes a comparison of the results of estimating MET values of physical activities in two different situations. 1) The location of the wearable sensor nodes is known using localization algorithms; 2) The location of the sensor nodes are unknown. The MES further deployed a regression model to determine the MET values corresponding to soccer exergaming movements. Based on the result, the average error of estimating MET values corresponding to exergaming movements with sensor localization, is 2.3 times less than the case with no sensor localization. We proposed a localization algorithm based on the KNN classifier to detect the location of on-body sensors with an accuracy of 92%. We were able to estimate the MET values of several exergaming movements with $R^2$ value 74% using two accelerometer sensors on hip and ankle, and the MET values computed from the metabolic cart.

One important feature of our system is reconfiguration of the underlying estimation model when the user adds a new untrained sensor to the network, using an algorithm based on the transductive transfer learning approach. Using our transfer learning algorithm, we improved the $R^2$ value when a new sensor was added to the wearable sensor network without any need for knowing the MET values in both exergaming and treadmill datasets. We achieved the average accuracy of 66% and 74% in MET estimation, respectively, in exergaming and treadmill experiments. The results demonstrate that we can compensate for the decreased $R^2$ value by 60% on average in both experiments which proves the effectiveness of our transfer learning approach.
ACKNOWLEDGMENT

This work was supported in part by the Washington State University New Faculty Seed Grant #123191.

REFERENCES


Parastoo Alinia received her B.S. in Computer Engineering from Sharif University of Technology, Tehran, Iran in 2014. She is currently a PhD Student in Computer Science in school of Electrical Engineering and Computer Science at Washington State University. The focus of her research is on algorithm design on embedded and pervasive systems with applications in healthcare and wellness. She is a student member of the IEEE.

Ramyar Saeedi received his B.S. degree in computer engineering from University of Isfahan, Isfahan, Iran, with honors, and MS.c. degree in computer architecture from University of Tehran, Tehran, Iran, in 2008 and 2011, respectively. Currently, he is working toward the PhD degree in electrical and computer engineering at Washington State University. He was with University of Kurdistan, and University of Tehran from 2011-2013, where he served as a lecturer and laboratory instructor. His current research interests include embedded systems, signal processing, low power design, and machine learning. He is a student member of the IEEE.

Ramin Fallahzadeh received his B.S. degree in computer engineering from Sharif University of Technology, Tehran, Iran in 2014. Currently, he is a PhD student in computer science at Washington State University. His current research interests include embedded systems, pervasive computing, machine learning, and wireless sensor networks. The focus of his research is on algorithm design and power optimization of networked wearable sensors with applications in healthcare. He is a student member of the IEEE.

Ali Rokni received his B.S. degree in Computer Engineering from Amirkabir University of Technology, Tehran, Iran in 2005. In 2008, he received his MS.c. in Computer Science from Sharif University of Technology, Tehran, Iran. From 2008 to 2014, he had been involved in industrial development and research in software engineering and machine learning. Currently, he is a PhD student in Computer Science at Washington State University. His current research focus is on multi-view learning algorithm development for embedded and pervasive systems.

He is a student member of IEEE.