Data Harmony

Junwon Jang and Matthew Zhang

I. ABSTRACT

In recent clinical studies, it has been shown that regular moderate-to-vigorous physical activity can establish significant health benefits for individuals. However, tracking this information can be fairly difficult - though there exist gold standards for tracking activity levels through devices like the Actigraph, they can be costly and difficult to use. We propose that mHealth devices, such as wearables and mobile devices, can be a much cheaper and accessible alternative, while also retaining the same gold standard classifications of the Actigraph.

In this study, we hypothesized that accelerometer data that is collected through an LG watch, an Android wearable, can be mapped onto METs given by an Actigraph, the currently accepted standard of physical activity used by researchers, to correctly determine levels of moderate-vigorous physical activity. We also hypothesized that this validation process can also be replicated among other mHealth devices, which will provide easier access for gathering physical activity data among clinicians and other health studies.

In our study, we had 6 adult males and adult females wear each of the devices in a supervised array of ten 1 minute activities. These activities will include sedentary behavior, light activity, moderate activity, and vigorous activity. We then synthesized and classified moderate-vigorous physical activity (defined as 3 METs) using the accelerometer data using regression models in our mHealth application. Afterwards, we compared the classification of our mHealth system to the actual classification of participant activities, which had an F-score of 0.937, which demonstrated that our mHealth system could validate classification data against the Actigraph fairly robustly.

II. INTRODUCTION

In modern medicine, it can be very difficult to monitor whether or not people are fulfilling the recommended amount of physical activity. It is recommended that adults have at least 150 minutes of moderate-vigorous physical activity each week. Precision medicine is a new field in health that is attempting to tackle this problem by improving our ability to analyze each persons unique lifestyles and catering to their individual needs. However, obtaining accurate and reliable data for each individual is difficult because of subject interpretations of activity between subjects and lack of standardized definitions of physical activity. Mobile health sensors such as smartphones and smartwatches are increasingly becoming a promising new approach for addressing this issue as these devices become more available and widespread. However, these sensors currently do not provide data that adheres to any standard of physical activity. Hence, the goal of this project will be to develop an mHealth system that can collect, synthesize, and standardize data from multiple sensors to provide a precise and accurate measurement of people's physical activity. If a universal measure of fitness can be developed through a harmonized mHealth application, we also hypothesize that physicians can more easily prescribe individual achievable fitness goals, as well as more easily and accurately monitor progress towards these fitness goals. These factors should improve a patients self efficacy towards meeting physical activity requirements, which is a key component of the Social Cognitive Theory construct from the Transtheoretical Model.

III. RELATED WORK

There has been a general medical consensus that moderate-to-vigorous physical activity has established health benefits for individuals [1] - however, with over 76 studies across 36 countries [2], the methods of classifying physical activity and harmonizing data are extremely varied.

Several categories of work have attempted to harmonize data across established medical data-collecting devices. For instance, one study measured raw data from medical accelerometers such as the ActiGraph and Hookie, and using Mean Amplitude Deviation (MAD) and the Bland-Altman method, found that the physical activity classifications derived from the raw data were comparable with activity classifications from heart rate monitors such as the M61. [3] However, while MAD and Bland-Altman methods are useful methods of classifying and harmonizing data, this particular category of study incorporates medical devices that require the use of chest or belt straps, which are impractical for the general population to wear on a daily basis.

The closest related work to our study attempts to harmonize data between an Android smartphone and an ActiGraph accelerometer [4]. This study utilized both laboratory studies and free-living studies, where participants would wear both the Android and ActiGraph devices while performing activities that could be classified as sedentary or moderate-to-vigorous physical activity. The researchers then mapped the raw accelerometer data using a low-pass filter and area-under-curve measurements to units that could be validated against the ActiGraphs counts, which is how the ActiGraph measures and classifies physical activity. Using cut-point derived classifications and Bland-Altman methods, the study found that the Android smartphone could be correlated with the ActiGraphs measurements ($p < .001$).

However, there are a few factors in which our study differs. While we intend to use the same process for mapping accelerometer data to units that can be validated against the ActiGraph, we are using an Android smartwatch rather than a smartphone. According to previous studies comparing activity classifications between hip and wrist accelerometer devices [5], wrist devices register a much lower ROC and accuracy score when compared to hip devices. As wrist worn fitness devices such as the FitBit and LG Watch gain popularity, it is important to ensure that the measurements from these devices can be validated as well, to ensure that data can be harmonized across the majority of mHealth devices that the general population uses. Additionally, the study focused on an older user population (mean = 55 years old), whereas we intend to focus on a younger user population, who are more likely to be using fitness-tracking smart devices on a regular basis.

IV. SYSTEM

The system is comprised of three levels: on-body sensors, smartphone and server database, and a UI display. First, sensor data is collected through on-body sensors from the LG watch and actigraph. The accelerometer data from the LG watch is then sent to the smartphone, from which the data then gets sent to a backend server along with the corresponding timestamp. The server will then run the above algorithm to classify the times at which a person is moderately to vigorously physically active and the times when a person is sedentary. A monitoring system will also be added so that doctors will be able to track their patients physical activity throughout the week.

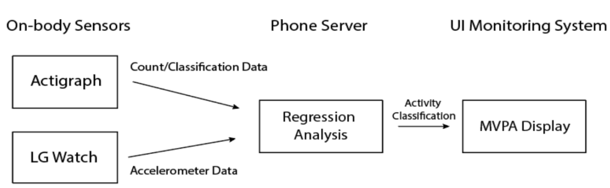

Fig. 1. DataHarmony system diagram

V. EXPERIMENTAL SETUP

Our method assumes that a person is wearing a watch with a built-in accelerometer from which data can be collected. For this study, we used the data from the accelerometer embedded into an LG watch and the sensors in the Actigraph. An LG watch was chosen because it is easily programmable, equipped with a built in accelerometer with a ready-to-use API. Furthermore, Accelerometers have been used successfully in the past for measuring physical activity, and so what this study attempts to do is use this accelerometer data to determine how physically active individuals are based on the MVPA metric provided by the Actigraph. The possible classifications for MVPA are as follows: sedentary, light, moderate, vigorous.

A. Data Collection

An LG watch was used to collect data to develop and evaluate our algorithm. The subject wore both the LG watch and Actigraph throughout the experiment. A custom application was written for the LG watch that collects the raw accelerometer data as shown in Figure 2. All the data collected from this was then recorded along with the timestamp.

Fig. 2. Accelerometer data collection and physical activity display using the LG Watch

In our experiment, a sliding window was used to divide up the sensor streams into 5-second intervals. In each interval, the Actigraph labels the MET value of each exercise. In that same interval, the LG Watch collects the average acceleration across its three axes during the window:

We performed this experiment and collected data in two separate batches.

In the first batch, we collected data from 4 individuals (2 male, 2 female) in order to develop an initial model for our algorithm to map accelerometer data onto the MVPA measure given by the Actigraph. Each subject ran the following protocol twice:

- Stand for 1 minute
- Walk 3.5 mph for 1 minute
- Run 5 mph for 1 minute
- Run 7 mph for 1 minute

- Run 9 mph for 1 minute

The accelerometer data was classified using the MET labellings from the Actigraph based on the corresponding timestamps between the devices. This was then used to develop the mapping algorithm from average accelerometer values retrieved from the LG Watch within the same timestamp window into MET values, and then later the MVPA classifications.

Then, in the second batch, we ran the same protocol on 2 individuals (1 male, 1 female) who are wearing both the LG watch and Actigraph to test the validity of our regression and classification algorithms.

B. Evaluation

Our evaluation metric measures the validity of our algorithm by looking at the total amount of time that is correctly classified. The true labellings are given by the Actigraph. Then the true positives (TP) are equal to the number of intervals that were correctly classified as moderate or vigorous physical activity; the false positives (FP) are the number of light or sedentary intervals that were classified as moderate or vigorous; the true negatives (TN) are the number of light or sedentary intervals that were classified correctly; and the false negatives (FN) are the number of light or sedentary intervals that were classified incorrectly.

VI. METHODS/ALGORITHM

The data collected from the LG watch consisted of the average acceleration values in the x, y, z directions within a 5-second sliding window, which was calculated as follows:

$$ACCEL_{avg} = \sqrt{x^2 + y^2 + z^2}$$

.

The goal of our algorithm is to use these accelerations values to classify the times where an individual is engaged in either light/sedentary or moderate/vigorous physical activity. To do this, a regression model was used to map the magnitude of the average acceleration values onto an MET value. Then, because MET values that are lower than 3.0 are defined to be sedentary/light activity, we used 3.0 as a threshold value for the regression. And so for all time intervals that mapped to an MET less than 3.0, the times were classified as light/sedentary and the intervals that mapped to an MET greater than 3.0 were classified as moderate/vigorous.

VII. RESULTS

A. Regression

Using our acceleration and MET values gathered from the experiment, we developed several regression models to map acceleration to MET.

As our regression model with a degree of 2 had the highest R^2 value, we decided to use this one to develop our classification system on the LG Watch. The regression equation for this model was as follows:

$$MET = -0.0433*(AccelMAG)^2 + 2.4417*(AccelMAG) - 17.3015$$

TABLE I
REGRESSION MODELS VS. R^2 VALUES

Degree	R^2
a.) 1	0.7625
b.) 2	0.7975
c.) 3	0.7801

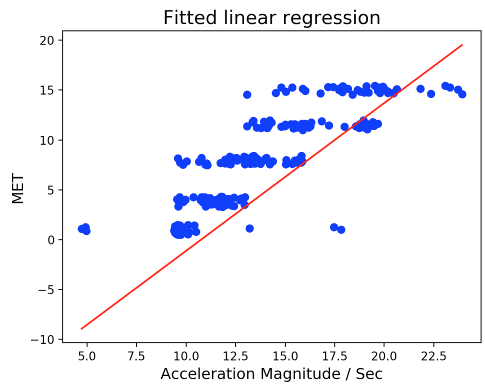

Fig. 3. a.) Degree = 1

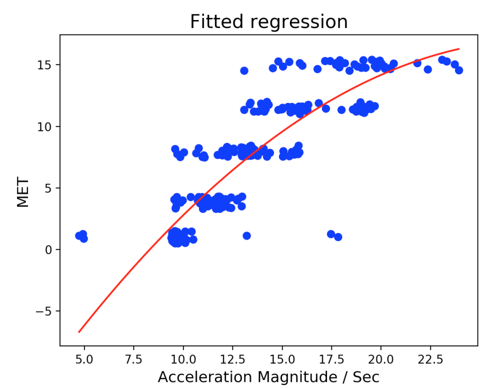

Fig. 4. b.) Degree = 2

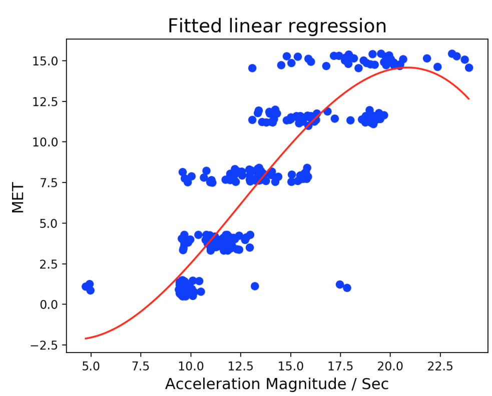

Fig. 5. c.) Degree = 3

B. Classification

Using the regression model developed in the previous section, we built the regression equation into the LG Watch to test the validity of our model. We had 2 participants, 1 male and 1 female, run through the same protocol of activities used to develop the initial training model. Within a 5-second time window, if the activity was light/sedentary, and the LG Watch classified the activity as below 3 METs, we classified it as a true negative; and if the activity was moderate/vigorous, and the LG Watch classified it as above 3 METs, we classified it as a true positive.

TABLE II

CONFUSION MATRIX OF ACTIVITY CLASSIFICATIONS

	Light/Sedentary	Moderate/Vigorous
Light/Sedentary	79	13
Moderate/Vigorous	10	185

The F-score from the confusion matrix above was

$$F1 = 0.9372$$

which demonstrated that our classification system was fairly robust in classifying these two categories.

VIII. LIMITATIONS/FUTURE WORK

As our study was only tested across 6 subjects in a 18-21 age range, we would like to be able to expand our study to a larger subject range in order to create more accurate models, as well as ensure that our system remains robust across different subject variations. Additionally, as our study is geared towards clinicians and researchers, we would like to be able to refine and conduct user studies with the monitoring UI of our Data Harmony application to see how useful our application is towards their purposes.

Finally, although our study was designed to be easily replicated across a variety of mHealth devices, we only validated an LG Watch against the Actigraph in this paper. Given more time, we would like to conduct further validation studies with other mHeatlh devices, such as the iPhone, Android phones, or Fitbits.

IX. CONCLUSION

In this study, we successfully developed a regression model that was able to map accelerometer values from the LG Watch, an mHealth device, onto MET values given by an Actigraph, the currently accepted standard of physical activity used by researchers. Through our Data Harmony application, we were also able to correctly classify if activities were light-sedentary or moderate-vigorous with a high F-score of 0.937. All together, we've demonstrated with high evidence that our hypothesis that an mHealth system could be validated against the Actigraph "gold standard" classification is correct, and that further validation studies with other mHealth devices should be undertaken.

X. ACKNOWLEDGMENTS

Thanks to Professor Nabil Alshurafa and Professor Amanda Paluch for their guidance throughout the project, as well as to the HABITS lab and the Northwestern EECS department for the funding.

REFERENCES

[1] Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet. 2012; 380(9838): 219 29. [PMC free article] [PubMed]

[2] Troiano, Richard P, et al. Evolution of Accelerometer Methods for Physical Activity Research. 29 Apr. 2014.

[3] Aittasalo, Minna, et al. Mean Amplitude Deviation Calculated from Raw Acceleration Data: a Novel Method for Classifying the Intensity of Adolescents' Physical Activity Irrespective of Accelerometer Brand. BMC Sports Science, Medicine and Rehabilitation, BioMed Central, 7 Aug. 2015, bmcsportsscimedrehabil.biomedcentral.com/articles/10.1186/s13102-015-0010-0.

[4] Hekler, Eric B, et al. Validation of Physical Activity Tracking via Android Smartphones Compared to ActiGraph Accelerometer: Laboratory-Based and Free-Living Validation Studies. JMIR MHealth and UHealth, JMIR Publications Inc., 15 Apr. 2015, www.ncbi.nlm.nih.gov/pmc/articles/PMC4414958/.

[5] Rosenberger, ME, et al. Estimating activity and sedentary behavior from an accelerometer on the hip or wrist. Medicine and Science in Sports and Exercise, http://europepmc.org/abstract/MED/23247702