Improving the Validity of Lifelogging Physical Activity Measures in an Internet of Things Environment

Po Yang, Martin Hanneghan, Jun Qi
School of Computing and Mathematical Sciences
Liverpool John Moores University
Liverpool, UK
p.yang@ljmu.ac.uk

Zhikun Deng, Feng Dong, Dina Fan
Centre of Computer Graphics and Visualisation
Bedfordshire University
Luton, UK
Zhikun.Deng@beds.ac.uk

Abstract—Recently, the popular use of wearable devices and mobile apps makes the effectively capture of lifelogging physical activity data in an Internet of Things (IoT) environment possible. The effective collection of measures of physical activity in the long term is beneficial to interdisciplinary healthcare research and collaboration from clinicians, researchers to patients. However, due to heterogeneity of connected devices and rapid change of diverse life patterns in an IoT environment, lifelogging physical activity information captured by mobile devices usually contains much uncertainty. In this paper, we provide a comprehensive review of existing life-logging physical activity measurement devices, and identify regular and irregular uncertainties of these activity measures in an IoT environment. We then project the distribution of irregular uncertainty by defining a walking speed related score named as Daily Activity in Physical Space (DAPS). Finally, we present an ellipse fitting model based validity improvement method for reducing uncertainties of life-logging physical activity measures in an IoT environment. The experimental results reflect that the proposed method effectively improves the validity of physical activity measures in a healthcare platform.

Keywords—physical activity, life-logging, internet of things

I. INTRODUCTION

Owing to the highly fragmented health system in many countries, gaining access to a consistent personal health record of individual citizens is beneficial to interdisciplinary healthcare research and collaboration [1]. As a key indicator in a number of obesity, diabetes and other chronic diseases, effective measurement and monitoring of physical activity are critical to design programs for preventing/treating metabolic syndrome and obesity [2]. Objectively measuring physical activity and the associated estimates of instantaneous and cumulative energy expenditure (EE) in the long term is valuable for clinical decisions and provides important feedback to people for achieving a healthy lifestyle. Recently, new technologies and concepts such as the ‘Internet of Things’ [3] (IoT) have been emerging as new tools that transform healthcare industry. In the early 21st century, Radio Frequency Identification (RFID) technology [4-6] is regarded as the new IoT solution to solve many healthcare challenges. In the last few 5 years, advance in wearable technology have promoted renewed interests in IoT enabled healthcare fields. A number of physical activity monitors [7-9] have been developed for capturing lifelogging physical activity information and providing continuous, real-time feedback to users. Also, due to the exponential growth of wearable devices [7-11] and mobile apps [12-13], it has become increasingly possible to remotely monitor a patient or citizen’s health by connecting heterogeneous medical devices into an Internet of Things (IoT) healthcare platform [14-15].

However, due to the commercial perspective, nearly all of the popular wearable device and mobile apps in the market focus more on personal fitness and exhibit a lack of compatibility and extensibility. Also, owing to the heterogeneity of connected devices and rapid change of diverse life patterns in an IoT environment, lifelogging physical activity information captured by mobile devices usually contains much uncertainty. Effective validation of these high volume and multi-dimensional data becomes an extremely difficult task. Traditional methods [16-19] use either dedicated wearable sensors [16-17] or advanced machine learning algorithms [18-19] to accurately monitor longitudinal physical activity and access activity patterns and intensity level. Unfortunately most of these methods consider performance optimization of a single sensor or a combination of GPS and accelerometer by analyzing raw sensors’ signals. In IoT based personalized healthcare systems, physical activity data is generated on a daily basis from globally heterogeneous third party devices. Traditional physical activity validation methods hardly deal with these scattered and heterogeneous data sets. In current literature, no methods are reported to successfully validate the heterogeneous physical activity from different resources in an IoT healthcare environment.

In this paper, we firstly give a comprehensive review of existing life-logging physical activity measurement devices, and identify regular and irregular uncertainties of these life-logging physical activity measures in an IoT environment. Then, we project the distribution of irregular uncertainty by defining a walking speed related score named as Daily Activity in Physical Space (DAPS). Finally, we present an ellipse fitting model based validity improvement method for reducing uncertainties of life-logging physical activity measures in an IoT environment. The experimental results reflect that the proposed method effectively improve the validity of physical activity measures in a healthcare platform MHA [20].
II. RELATED WORK

The concept of IoT based personalized healthcare systems [16] uses a set of interconnected devices to create an IoT network devoted to healthcare assessment, patient monitoring and automatically detecting situations. In Fig.1, the general system collects personalized health information from different wearable sensing devices through a middleware that provides interoperability and security needed in the context of IoT for healthcare. These wearable devices are capable of recording multiple types of health data, including physical activity, sleep, heart rate and blood pressure. Among this data, due to the technical and functional maturity of MEMS accelerometer technology and GPS, physical activity is mostly well-observed.

![IoT personalized healthcare systems](adopted from [16])

Recently, many commercial wearable products [7-11] and mobile applications [12-13] have been released that support long term recording and collection of personal health information, particularly on physical activity. Popular mobile apps, such as Moves [12], are based on smartphone 3D accelerometer data and GPS information which allows tracking user movement activities including location, distance and speed. The wearable products, such as Fitbit, Nike+ Fuelband [8], Withings [9], Endomondo [13], are all wristband devices that record steps count, distance, and calories burnt. A brief comparison of above products is listed in Table.1 and explained in detail below:

<table>
<thead>
<tr>
<th>Programs</th>
<th>Data</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobile Apps</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endomondo</td>
<td>route, distance, speed</td>
<td>community sharing, Android and iOS</td>
<td>Short battery longevity, not work indoor</td>
</tr>
<tr>
<td>MyTracks</td>
<td>route, distance, speed</td>
<td>view data live, API support, only Android</td>
<td>Short battery longevity, not work indoor</td>
</tr>
<tr>
<td>Device</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fitbit</td>
<td>steps, calorie, food</td>
<td>low cost, Android and iOS, long battery life</td>
<td>Limited API</td>
</tr>
<tr>
<td>Nike+</td>
<td>steps, calorie, food</td>
<td>reasonable cost, Android and iOS</td>
<td>Variations on accuracy</td>
</tr>
<tr>
<td>Jawbone Up</td>
<td>steps, distance, calorie</td>
<td>reasonable cost, Android and iOS</td>
<td>No API</td>
</tr>
</tbody>
</table>

- **Endomondo** is a popular GPS based mobile application for tracking route, distance, duration, split times and calorie consumption. It offers a full history with previous workouts, statistics and a localized route map for each work out.
- **Google MyTracks** is also based on the use of GPS to record the user’s path, speed, distance and elevation while they walk, run, and cycle (or do any activities) outside.
- **Fitbit** records steps taken, distance travelled, and calories expended. These devices communicate with a host computer using Bluetooth that in turn sends data directly to a user’s account on the Fitbit website.
- **Nike+ Fuelband** is worn on the wrist and records calories, steps, distance, and Nike’s own unit of activity terms “Nike Fuel”. The device connects via USB to a host machine which syncs the data to a user’s account on the Nike+ website.
- **Jawbone Up** calculates steps, distance and calories. Currently the Jawbone up can only be used with a mobile device, drivers for laptop and PCs are not provided.

These wearable devices communicate with a mobile phone via Bluetooth running the relevant mobile application. While the above products have proven their popularity among general users, their usage is limited in the fitness field. This is due to a diversity of life pattern and environmental impacts; personal physical activity data from individual wearable device exhibits remarkable uncertainty. The validity of physical activity data in longitudinal healthcare cases is very challenging. Also, with the exponential growth in the mobile healthcare market, numerous similar wearable products have been developed, which significantly increases the heterogeneity and diversity of devices connected in IoT based personalized healthcare systems.

III. IDENTIFICATION OF UNCERTAINTY

The Internet of Things (IoT) enabled healthcare system is the theoretical cornerstone of validation of physical activity in an IoT environment, as shown in Fig.2. In terms of the concept of IoT, personal health data are accumulated and measured as a cube in three dimensions (3D): Persons, Devices and TimeLine. The increment in any dimension results in an expansion of the health data grid. The products...
like Fitbit Flex [7] or Moves [12] occur on a 2D plane (Persons × TimeLine), which refer to scenarios in which a single device is used by an increasing population over time. Similarly, physical activity recognition with sensor fusion [17-19] [21-22] appears on a 2D plane (Devices × TimeLine) for classifying an individual person’s activities with historical health data.

Fig. 2 Concept of IoT personalized healthcare systems

The uncertainty of physical activity here is categorized into two types:

Irregular uncertainty: Irregular Uncertainty (IU) in physical activity data occurs randomly and accidentally. The causes of these uncertainties may include device malfunctions or faults, breakdown of a third-party server, misuse of mobile apps, or sudden change of personal circumstance for example. The occurrence of irregular uncertainty in physical activity data will appreciably impact the efficiency and accuracy of assessing personal health.

Regular uncertainty: Regular Uncertainty (RU) in physical activity data occurs frequently and persistently. The causes resulting in these uncertainties are mainly from some regular influencing issues, like intrinsic sensors’ errors, differentiation of personal physical fitness and changes of environment. The occurrence of regular uncertainty in physical activity data is inevitable so that it is impossible to completely eliminate these uncertainties.

IV. Ellipse Fitting Model for Removing Irregular Uncertainty

Evaluating the distribution of irregular uncertainties is a primary step in lifelogging physical activity validation. The evaluation of irregular uncertainty distribution is based on the MHA platform, which is an IoT enabled personal healthcare experiment platform connecting Moves, Fitbit and Withings. This platform enables a user to transfer their physical activity data from these third party providers into the MHA server, and then to be able to visualize and analyse this information to gain a better understanding. On the MHA platform, we initially collect daily physical activity (Steps, Distance and Calories) of seven users over six months by three types of wearable device (Withings, One and Moves). All these users (one female and six male) are researchers in university, and their ages are in the range of 30-50 years old.

The features of this raw activity data are:

- All seven people use Moves. Two of them additionally use Withings, and another three people use Flex.
- Missing data occurs frequently in Withings and Flex, because users easily forget they are wearing them.
- Some data in Flex shows lower steps, which is probably because users take of their wearable devices sometime during the day, or devices are out of battery.
- Moves data are more complete than Flex or Withings, but with relatively high errors.

There is a need for a benchmark to represent a person’s physical fitness from completed data sources. Here a walking speed related score is defined to represent a person’s physical fitness, named Daily Activity in Physical Space (DAPS). This score is inspired from earlier work [23] that proposes a Movement and Activity in Physical Space score as a functional outcome measurement for encompassing both physical activity and environmental interaction. Currently, most third party APIs of wearable devices or mobile apps provide functions to assess the intensity of physical activity regarding walking speed. For instance, Fitbit [7] classifies the intensity of daily activities into Very Active, Moderately Active, Lightly Active and Sedentary; Moves [12] records a series of walking segments containing duration, distance and speed. Here, we classify the intensity of daily physical activity into N levels in terms of the ranges of walking speeds (V₁, V₂ … Vₙ). The DAPS formula is created by summing these different level walking speeds:

\[ DAPS = \sum_{i=1}^{N} V_i \]  \hspace{1cm} (1)

Using the data of DAPS and Daily Steps, we calculate \( V_{daw} \), and plot \( S_d \) and \( V_{daw} \) in 2D diagram as in Fig.3.

In order to measure \( T_x \) and \( T_y \) to remove irregular uncertainty physical activity data, we use an ellipse equation (2) to cover 95% of data (C = 0.95).

\[ \frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1 \] \hspace{1cm} (2)

Where:
- \( h \) : Average daily walking speed
- \( k \) : Average daily walking steps
- \( a \) : Error range of average daily walking speed
- \( b \) : Error range of average daily walking steps
A noticeable issue here is that we only consider the lower limits of walking steps and the upper limits of walking speeds as threshold parameters. On some days users might walk distinctly more steps than usually, while the other days might be more sedentary. The threshold parameters are represented in equation (3):

\[ T_y = h + a \]
\[ T_x = k - b \]  

V. EXPERIMENT RESULTS

In this section, we discuss the performance evaluation of our proposed method in a case study on the MHA platform [20]. The criteria for verifying our validation model will concentrate on the efficiency and adaptivity of the method.

We collected an empirical dataset by using the MHA platform. The dataset includes one-year-long daily physical activity information from 14 subjects acquired with three devices: Moves was used by 14 users for nine months; Flex was used by five users for 12 months; Withings was used by three users for three months. These people are healthy in the age range of 30-50 years. The evaluation methodology for verifying the efficiency of proposed model will interview the participants, and collect feedback on reflecting users’ experiences on physical activity uncertainties through different devices. The feedback is used as a standard benchmark to compare the correctness of model.

In order to validate the accuracy of identifying irregular uncertainty, we follow equation (2) and (3) with a confidence interval of 95% to filter data from three different devices. We use the values (130, 1784, 884) of threshold parameter Ts respectively in Moves, One and Withings, for filtering incorrect daily steps data. The results are shown in Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Moves</th>
<th>Flex</th>
<th>Withings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ts Daily Steps</strong></td>
<td>130</td>
<td>1784</td>
<td>1267</td>
</tr>
<tr>
<td>T_x, DAPS Speed (m/s)</td>
<td>0.5</td>
<td>1.50</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Total number of People</strong></td>
<td>14</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Percentage of people with IU</td>
<td>43%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td><strong>Number of IU occurrence</strong></td>
<td>40</td>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td>IU confirmed by User</td>
<td>40</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td><strong>Average number of IU occurrence per person (User Feedback)</strong></td>
<td>6.6</td>
<td>5.4</td>
<td>2.7</td>
</tr>
<tr>
<td><strong>Accuracy of identifying IU (95%)</strong></td>
<td>100%</td>
<td>88.2%</td>
<td>62.5%</td>
</tr>
<tr>
<td><strong>Accuracy of identifying IU (98%)</strong></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Moves has much lower threshold parameters of Daily Steps and DAPS speed than Flex and Withings which are 130 and 0.5 m/s respectively (Table 2). This is because Moves has larger device uncertainties than Withings and Flex as we observed in section IV. Thus the GPS and smartphone internal sensors-based App is not as accurate as an accelerometer only based wrist wearable device. In terms of percentage of people having IU, Moves is much lower than Withings and Flex. It is probably because most of uncertainties from Moves have been classified into regular uncertainties, so its irregular uncertainties became less than for other two devices. However, for average IU occurrence per subject, Moves has higher performance than other two devices (Table 2). The accuracy of identifying IU appears that on the condition with a confidence interval of 95%, the related value of threshold parameter Ts can successfully filter irregular uncertainty in Moves. So Moves has the best IU identification accuracy up to 100%, which means that the incorrect daily steps detected by equation (3) in Moves have been all approved by users. Flex and Withings have accuracy up to 88.2% and 62.5% respectively, which implies that some correct daily steps are eliminated by our method.
For validating the adaptivity of the proposed ellipse fitting model, we consider the whole group of 14 subjects as one group due to the similar professions and backgrounds. We estimate the change of daily steps $T_s$ and DAPS with different periods (from one to 12 months) with a confidence interval of 95%. The results are shown in Fig. 4 and 5.

Fig. 4 Average of daily steps as the function of time period duration

![Fig. 4 Average of daily steps as the function of time period duration](image)

Fig. 4 shows the parameter Daily Steps as the function of time period duration. The value of this parameter is lower for shorter time periods than for longer time periods. The value of this parameter also varies with different devices. For Moves and Withings, the value of this parameter over different periods is slightly growing, but for Fitbit, this parameter dramatically increases after six months. This effect may be influenced by the setting of the confidence interval.

Fig. 5 DAPS as the function of time period duration

![Fig. 5 DAPS as the function of time period duration](image)

Fig. 5 shows little variation of DAPS parameter in the proposed method when the time period duration is changed. There are some minor fluctuations of DAPS on both Moves and Fitbit but in the long term, the value of DAPS is quite stable, which indicates that personal physical fitness does not have significant differences within this group of 14 people.

VI. CONCLUSION

This paper presents an ellipse fitting model based validity improvement method for reducing uncertainties of life-logging physical activity measures in an IoT environment. The experimental result on an IoT-enabled healthcare platform MHA [20] shows that this method can effectively improve the validity of physical activity measures in a group of small number of populations. While efficiency and accuracy of our method require further investigation by more populations and connected devices, our method demonstrates a possible way of improving the validity of life-logging physical activity data in an IoT environment. The future work will focus on extending the proposed method in a large-scale IoT environment, which will include more wearable devices and more people. Also, it will attempt to analysis and process the life-logging data with some machine learning techniques for improving the accuracy of proposed validation method.

REFERENCES


