2015

SmartCare: Energy-Efficient Long-Term Physical Activity Tracking Using Smartphones

Hui Liu
School of Software and Institute of Software Engineering, Xidian University, and Science and Technology on Information Transmission and Dissemination in Communication Networks Lab., Xi’an 710126, China.

Rui Li
School of Software and Institute of Software Engineering, Xidian University, and Science and Technology on Information Transmission and Dissemination in Communication Networks Lab., Xi’an 710126, China.

Sicong Liu
School of Software and Institute of Software Engineering, Xidian University, and Science and Technology on Information Transmission and Dissemination in Communication Networks Lab., Xi’an 710126, China.

Shibian Tian
School of Software and Institute of Software Engineering, Xidian University, and Science and Technology on Information Transmission and Dissemination in Communication Networks Lab., Xi’an 710126, China.

Junzhao Du
School of Software and Institute of Software Engineering, Xidian University, and Science and Technology on Information Transmission and Dissemination in Communication Networks Lab., Xi’an 710126, China.

Follow this and additional works at: https://tsinghuauniversitypress.researchcommons.org/tsinghua-science-and-technology

Part of the Computer Sciences Commons, and the Electrical and Computer Engineering Commons

Recommended Citation

This Research Article is brought to you for free and open access by Tsinghua University Press: Journals Publishing. It has been accepted for inclusion in Tsinghua Science and Technology by an authorized editor of Tsinghua University Press: Journals Publishing.
SmartCare: Energy-Efficient Long-Term Physical Activity Tracking Using Smartphones

Hui Liu, Rui Li, Sicong Liu, Shibian Tian, and Junzhao Du*

Abstract: Lack of physical activity is becoming a killer of our healthy life. As a solution for this negative impact, we propose SmartCare to help users to set up a healthy physical activity habit. SmartCare can monitor a user's activities over a long time, and then provide activity quality assessment and suggestion. SmartCare consists of three parts, activity recognition, energy saving, and health feedback. Activity recognition can recognize nine kinds of daily activities. A hybrid classifier that uses less power and memory with satisfactory accuracy was designed and implemented by utilizing the periodicity of target activity. In addition, a learning-based energy saver was introduced to reduce energy consumption by adjusting sampling rates and the set of features adaptively. Based on the type and duration of the activity recorded, health feedback in terms of the calorie burned was given. The system could provide quantitative activity quality assessment and recommend future physical activity plans. Through extensive real-life testing, the system is shown to achieve an average recognition accuracy of 98.0% with a minimized energy expenditure.

Key words: physical activity tracking; hybrid classifier; health feedback

1 Introduction

Lack of physical activity and unhealthy diets can lead to severe chronic diseases such as lumbar disorders, cervical spondylopathy, brain disorders, etc. They have become the fourth most significant cause of death, killing 3.2 million people in 2013, according to World Health Organization statistics. Long-term individual physical activity monitoring is urgently required. The rapid improvement of Cyber-Physical Systems (CPS) is making possible health care systems that integrate computation, communication, and control.

Activity recognition is the key technique in physical activity monitoring systems and has been widely studied[1]. Activity recognition systems using wearable sensing devices[2–4] or small attached sensor chips[5, 6] have been investigated. Recent studies[3, 7] have used Google Glass or wristbands to improve the precision of motion detections. Many wearable devices are commercially available, such as the Apple Watch[8] and Fitbit[9]. These utilize multiple sensors to track users’ movements, and can detect walking, running, and other sporting activities. However, the extra cost and restrictive requirements limit their application in continuous monitoring. Given their range of built-in sensors, activity recognition using smartphones offers a solution to long-term activity tracking, since smartphone ownership is becoming ubiquitous.

A major challenge to smartphones use is the trade-off between measurement accuracy and energy consumption[10]. Many approaches have used multiple sensors such as GPS and barometers to achieve high
levels of accuracy\textsuperscript{14, 11, 12}. However, the continuous use of multiple sensors places demands on the limited battery resources of smartphones, which is unacceptable in long-term activity monitoring. Android and iOS applications such as Moves\textsuperscript{13} and Runkeeper\textsuperscript{14} offer physical activity tracking but also reduce battery life. To reduce energy use, several studies\textsuperscript{15, 16} have proposed sensor transition strategies. Chu et al.\textsuperscript{17} optimized mobile classifiers for saving energy. Yan et al.\textsuperscript{10} used a single accelerometer and decreased the sampling cost by adjusting sampling frequency in real-time. But the results of these studies are still far from being put into practice, and there is a need for future work on sampling, monitoring, and energy consumption.

In this study, we propose SmartCare, an energy-efficient long-term physical activity tracking system using smartphones. SmartCare is an energy-efficient way of precisely monitoring nine basic daily physical activities: walking, jogging, ascending and descending stairs, bicycling, travelling up in an elevator, travelling down in an elevator, using an escalator, and remaining stationary, using the smartphone’s built-in accelerometer and magnetometer sensors. To allow extended use, SmartCare seeks to optimize the trade-off between activity classification accuracy and energy consumption. Using activity-dependent preferred sampling and a classification feature choice strategy, SmartCare adopts a learning-based activity pattern prediction algorithm and a runtime parameter adjustment strategy to decrease energy consumption. By identifying regular features and correlations among activities during long-term monitoring, the subsequent activity pattern can be predicted, allowing SmartCare to manage the sensors and adjusts the parameter tuple of the sampling rate and set of classification features in real time. Furthermore, after collecting sensing data, SmartCare adopts a hybrid classifier methods utilizing decision trees to classify non-periodic and periodic activities based on accelerometer data. Non-periodic activities are further classified into lift use, escalator use, and remaining stationary, based on magnetometer data. Periodic-like activities are classified through a Probabilistic Neural Network (PNN) classifier and periods are calculated using a novel period detection algorithm. Finally, SmartCare follows users’ long-term physical activity habits and gives personalized quantitative health assessment and health regime suggestion. The paper has the following three contributions:

- We describes SmartCare, an energy-efficient long-term physical activity tracking system using smartphones. A practical energy saver and a hybrid classifier are proposed, which enable SmartCare to run over an extended period while achieving satisfactory classification accuracy.
- Users’ daily physical activities and body type are taken into consideration to enable using SmartCare to deliver quantitative health assessment and making suggestions to help establish their healthy physical regimes.
- We undertook extensive experiments, using 8 participants across more than 4 weeks of continuous monitoring. SmartCare was implemented on the Android platform with running on Google Nexus, Samsung Galaxy S2, and Meizu MX3. The energy consumption of SmartCare is about 50% less than that of 100 Hz mode and slightly less than continuous sensing at 16 Hz.

The rest of this paper is organized as follows. Section 2 gives an overview of SmartCare. Section 3 describes the main design features. Section 4 presents the experimental results of the SmartCare system, and Section 5 discusses the limitations of our approach. Section 6 briefly reviews the related literature on activity recognition and health care. Finally, Section 7 concludes the paper and suggests future directions.

2 System Overview

Figure 1 gives an overview of the SmartCare system. The system comprises three modules: activity recognition, health feedback, and learning-based energy saving. The activity recognition module collects the sensing data continuously from the built-in accelerometer and magnetometer to sample physical activities with an adaptive sampling rate. Then the raw data is fed to a noise filter by applying a moving average over the last $n$ readings (e.g., $n = 3$). The results are sent to the hybrid classifier to distinguish different activities. The classifier divides the activities into periodic and non-periodic activities. Three non-periodic activities can be identified: lift use, escalator use, and remaining stationary, each with unique activity templates. Periodic activities are segmented into frames by a Period Detection Algorithm (PDA) using an adaptive sliding window. If the frame is not matched, the activity is regarded as non-periodic. Otherwise,
the activity frames are entered into activity pattern to extract a vector composed of several time-domain and frequency-domain features. The vector is fed into the PNN classifier and the recognition results are output.

These recognition results for the different activities are then input to the health feedback module to assess the users’ activities and give personalized health feedback. In this module, SmartCare adopts the well-defined quantitative measurements of Body Mass Index (BMI), Total Energy Expenditure (TEE), and Total Energy Intake (TEI) calculated from user’s physical information, diet information, and Metabolic Equivalent of Tasks (METS), to give a comprehensive assessment and to offer suggestions.

The learning-based energy saver module runs in a continuous learning cycle mode and works through the activity recognition process. The recognition result is sent as an input to a user’s long-term activity set representing the features of continuous or consecutive activities. Based on the results and the user’s long-term activity features, SmartCare adopts runtime sampling and a feature parameter adaptation strategy to adjust the parameter tuple of the Sampling Rate (SR) and a set of classification features (SF) by switching dynamically between the choice of SR and the set of classification features, which is not equivalent of Tasks (METs), to give a comprehensive physical information, diet information, and Metabolic Equivalent of Tasks (METS), to give a comprehensive assessment and feedback.

3 SmartCare Design
In this section, we describe the design of SmartCare. Firstly, we discuss the sampling and classification in the activity recognition process. Secondly, we describe the learning-based energy saver used by SmartCare. Finally, we discuss quantitative health assessment and feedback.

3.1 Activity recognition
3.1.1 Activity-dependent data collection
When sensing physical activities, the SR of the sensors needs to be precisely set. Using the built-in accelerometer and magnetometer, SmartCare collects raw sensor data in three $X$, $Y$, and $Z$ coordinates. It then extracts a set of classification features and feeds them into an activity classifier that can recognize nine different physical activities. There is a close relationship between the choice of SR and the set of classification features. A higher sampling rate and larger set of features results in more accurate activity classification, but at the cost of a higher energy overhead, especially in long-term continuous monitoring. Most previous studies used the same sampling frequency and set of features for whole activities or states, which is not
optimal\cite{10}. A more energy-efficient approach is to select different combinations of the SR and SF for different activities while preserving accuracy.

Experiments were conducted to determine the optimal parameter selection. By comparing activity recognition accuracy against the energy consumption for every possible combination, we propose activity-dependent preferred parameter tuple selection to optimize the sampling rate and set of classification features for each activity. We first define a criterion for parameter selection.

### 3.1.1.1 Selection criterion

**Definition 1** The preferred parameter tuple \( \langle A, (SR, SF) \rangle \) is a preferred tuple of sampling rate SR and set of classification features SF for activity \( A \). The \( (SR, SF) \) is qualified if it satisfies the condition \( \max \frac{\text{accu}_{SR, SF}}{E} \), where \( \text{accu}_{SR, SF} \geq \max \{N \times \text{accu}, 85\%\} \). \( E \) and \( \text{accu}_{SR, SF} \) represent energy cost and the accuracy of classifying activity \( A \) using specific (SR, SF). \( N \) is the average accuracy for a specific activity among all \( \langle A, (SR, SF) \rangle \). \( N \) is a constraint coefficient (we used \( N=1 \)) and 85% is a base accuracy threshold.

The SR and SF above were carefully chosen before the experiments were run. For sampling rate selection, any value from 0 Hz to 100 Hz is theoretically possible. Since most Android platforms only permit the sampling rate value to be one of the four values, 5 Hz, 16 Hz, 50 Hz, and 100 Hz, we utilized these four values as our SRs and limited the experiment to phones that support all four values. Classification feature selection comprises two categories: time-domain features and frequency-domain features. Our results suggest that both power consumption and classification accuracy are largely determined by whether frequency-domain features are selected or not, while the exact choice of time- or frequency-domain features has little effect. Therefore, instead of trying every possible combination of features, alternatives were simply divided into the following two sets:

- The time-domain features were selected.
- Both the time-domain and frequency-domain features were selected.

We choose widely used time- and frequency-domain features, which are presented in the following section.

### 3.1.1.2 Preferred parameter tuple selection

In order to determine the preferred sampling and classifying parameter tuple \( \langle A, (SR, SF) \rangle \), we conducted experiments to qualify the relationship between energy consumption and classification accuracy. Three participants performed 9 activities, each for a period of 6 minutes, over a period of a day, giving a total of 162 samples. The test activities were walking, jogging, ascending and descending stairs, bicycling, lift travelling up, lift travelling down, escalator use, and remaining stationary. To improve energy consumption monitoring, the network and display were turned off.

In the training phase, the rate was kept fixed at 50 Hz and the data on all individual activities by each participant were imported into the classifier. In the verification phase, the SR was varied and the classification accuracy was estimated by performing 9-fold cross validation of the data at different sampling frequencies and across different feature sets. Figure 2 shows the average classification accuracy and one-hour power consumption for three participants across ten activities at different SRs and two feature choice types. It was shown that: (1) A higher sampling rate led to higher accuracy and energy consumption. The increase in energy consumption was a non-linear, logarithmic function of the rate. (2) Combining time-domain and frequency-domain features yielded higher accuracy than time feature sets alone. The influence on each individual activity was unclear, as shown in Fig. 3.

Figures 3 and 4 show the average accuracy for each activity using different sampling rates and features. Figure 4 uses only time-domain features, while Figure 3 uses both time-domain and frequency-domain features. The choices of sampling rate and domain features affected activity-dependent accuracy. For example, greater than 90% accuracy was achieved for the activities lift traveling up and remaining stationary even at 5 Hz and using only time-domain features. In

![Fig. 2 Average classification accuracy and energy cost of SmartCare.](image-url)
contrast, the activity bicycle required both time- and frequency-domain features and a higher frequency 50 Hz to achieve greater than 90% accuracy. A sampling rate 100 Hz yielded slightly higher accuracy, but at the cost of considerably higher power consumption.

The experimental results show a trade-off between the accuracy and power consumption across the different activities. Using Definition 1, we obtained a preferred selection of sampling frequency along with nine pairs of \( (h_A, SR, SF) \) for each activity, as shown in Table 1.

### 3.1.2 Invalid activity exclusion

Before activity classification can be conducted, invalid activities must be excluded. The experimental results suggest that transient activities such as putting a smartphone into the pocket cause errors in the recognition results. The recognition of nine basic sustained activities required the exclusion of such short period activities. An invalid activity exclusion algorithm was therefore run before recognition and classification was initiated. This is shown in Algorithm 1.

Several test runs were conducted and mean variance was applied to establish the appropriate cut-off duration period based on real data. The threshold for valid activities, except for lift, was established as 20 s. If the system recognizes a lift in any cycle, it stores the result in the local database. If the same activity is recognized in a second cycle, the activity is inserted into the local database.

### 3.1.3 Non-periodic activity recognition

After preprocessed (post noise filter) accelerometer and magnetometer data containing activity information, and invalid activities excluded, the next step is the recognition and classification of activities. We designed a two-layer hybrid classifier combining a first-layer decision tree classifier with a second-layer PNN classifier, by which the activities were divided into two groups in accordance with diverse period characteristic. The work flow of the hybrid classifier is plotted in Fig. 5.

We first recognized three non-periodic activities using a first-layer decision tree classifier. The decision tree classifier first divided the activities into two groups: non-periodic and periodic. Periodic activities displayed sharper acceleration fluctuations than non-periodic ones. Thus, the periodicity of an activity was judged from the standard deviation of its acceleration magnitude over some appropriate time period.

Three non-periodic activities were then identified, based on their unique accelerometer and magnetometer profiles:

- Elevator: A unique accelerometer data pattern was unique recorded when the user is in an elevator. We were further able to determine whether the elevator was moving up or down. A threshold was established to characterize the pattern of pulses uniquely associated with elevator use.

### Table 1 Choice list for each individual activity.

<table>
<thead>
<tr>
<th>No.</th>
<th>Activity</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>upstairs</td>
<td>16 Hz</td>
</tr>
<tr>
<td>2</td>
<td>downstairs</td>
<td>50 Hz</td>
</tr>
<tr>
<td>3</td>
<td>walk</td>
<td>16 Hz</td>
</tr>
<tr>
<td>4</td>
<td>bicycle</td>
<td>50 Hz</td>
</tr>
<tr>
<td>5</td>
<td>jog</td>
<td>16 Hz, 5 Hz</td>
</tr>
<tr>
<td>6</td>
<td>lift</td>
<td>5 Hz</td>
</tr>
<tr>
<td>7</td>
<td>liftdown</td>
<td>5 Hz</td>
</tr>
<tr>
<td>8</td>
<td>stationary</td>
<td>5 Hz</td>
</tr>
<tr>
<td>9</td>
<td>escalator</td>
<td>100 Hz</td>
</tr>
</tbody>
</table>

**Algorithm 1: Invalid activity exclusion algorithm**

**Require:** Current activity recognition result current-result

1. if current_result == lift then
2. last_result ← lift
3. Insert current result into Database
4. else if current_result == last_result then
5. Insert current result into Database
6. else
7. last_result ← current_result
8. end if
Escalator: It proved difficult to distinguish escalator use from a stationary participant, as the magnitudes of acceleration for both activities were close to a certain constant. However, the magnetometer data was proved to be a reliable discriminator. Thus, the standard deviation of magnetometer data is used to distinguish escalator use.

Stationary participant: Once other possibilities are excluded, the remaining activities were classified as stationary.

The periodic-like activities whose variance of accelerations was high were then be entered into the second-layer classifier. If a frame was not found in this phase, the activities containing no frames were classified as non-periodic.

### 3.1.4 Periodic activity recognition

The periodic activities were next processed to establish their periods and features. Five categories were established by the second-layer classifier. Since PNN has achieved a satisfactory level of accuracy without high computational demands, it was selected as the second-layer classifier, in preference to other advanced classifiers. In this section, we present the period detection algorithm and feature extraction part of the PNN classifier.

#### 3.1.4.1 Period detection algorithm

We propose a period detection algorithm to calculate activity periods for periodic-like activities.

We identified a number of frames according to the following two steps. First, all periods $P_{ei}^s$ were checked to find their count $n_e$, starting point $s_i$, end point $e_i$, and ordinal $i$. Second, we calculated the value of $F_{ei}^s$ by Eq. (2). Note that the count of periods in a frame, $n_e$, is already known, and the count of frames, $n_f$, can be determined by Eq. (1).

$$n_f = \left\lfloor \frac{n_e}{n_t} \right\rfloor$$

$$F_{ei}^s = \left( P_{s_i}^{e_i+1}, P_{s_i+n_e-1}^{e_i} \right)$$

where $s_i+1 = e_i$. In this way, SmartCare can quickly decide the location and length of each frame. We rounded down the result if $n_f$ was not an integer to ensure that every frame had $n_t$ periods. The pseudo code for a period detection algorithm is given in Algorithm 2.

---

**Algorithm 2: Period detection algorithm**

**Input:** Raw acceleration data along the three axes ACC\([x, y, z]\).

**Output:** An array of positions of the stride points STRIDE_POS.

1. \text{MAG} = \text{getMagnitude}(\text{ACC}).
2. for \(i \leftarrow 2 \) to \text{length(MAG)} do
3. \quad if $\text{MAG}(i) - \text{MAG}(i-1) > 0$ and $\text{MAG}(i) - \text{MAG}(i + 1) > 0$ then
4. \quad \quad Insert $i$ into the array PEAK_POS
5. \quad Insert $\text{MAG}(i)$ into the array PEAK_MAG
6. \quad end if
7. \quad threshold gets average (PEAK_MAG) \times 1.5
8. \quad $i \leftarrow 1$
9. \quad while $i < \text{length (PEAK_MAG)}$ do
10. \quad \quad if PEAK_MAG($i$) > threshold then
11. \quad \quad \quad Insert PEAK_POS($i$) into the array STRIDE_POS
12. \quad \quad else
13. \quad \quad \quad $i \leftarrow \text{stride_pos} + 25$
14. \quad \quad end if
15. \quad $i \leftarrow i + 1$
16. \quad end while
Three important concepts used in deriving the algorithm are shown in Fig. 6. Stride point is the peak during a period of a periodic activity and two adjacent stride points can determine the length of a period. Note that only one stride point exists in any period and the peak may not be the highest in that the peak depends on the first one. A threshold is used to determine whether a data point is a stride point; only when the data point lies above the current threshold is it considered as a candidate stride point. In the course of the experiment, we often detected several stride points in a single period. This was caused by invalid vibrations. Thus, an illegal window was designed, based on the assumption that people can jog as rapidly as two strides per second (other activities are slower). Thus, the minimum distance between two valid adjacent stride points is defined as 0.5 s and any stride point falling within the illegal window is discarded.

### 3.1.4.2 Feature extraction

Obtaining the period of each possible activity, PNN classifier firstly extracts the classification features in a frame during the training phase.

The classification features can be classified into two categories, i.e., time-domain and frequency-domain features. Time-domain features are calculated on the appropriate frames of sensor streams, while frequency-domain features are from frequency-domain coefficients, which are first obtained by using FFT on each frame. As discussed earlier, we propose an activity-dependent feature selection strategy that each activity has either time-domain features or both time and frequency features. Based on prior works, we summarize and adopt the commonly used features, some of which are shown in Table 2. We explain and list the calculations of three representative features below.

<table>
<thead>
<tr>
<th>Time-domain features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variance and standard deviation of each axis, deviation, mean, binned distribution of each axis correlation Cor(y, z), average period length</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency-domain features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average energy of each axis, spectral entropy, range power</td>
</tr>
</tbody>
</table>

- Average period length $\overline{P}$. From the output of period detection algorithm, we can get the average period length.
- Binned distribution of each axis $\text{Dis} [3]$. We determine the range of values (maximum–minimum) of each frame and divide this range into 5 equal-sized bins. Then we record the count of values falling within each of the bins.
- Average energy of each axis, computed by Eq. (3).

$$\text{AverageEnergy} = \frac{1}{N} \sum_{i=1}^{n} a_i^2 \quad (3)$$

where $N$ is the number of sensor data along each axis in a frame and $a_i$ denotes the accelerometer reading along certain axis at a given time $i$.

Based on the features extracted above, the PNN classifier calculates and obtains different sets of features of different activities and further classifies them into walking, jogging, ascending and descending stairs or bicycling using statistical principle.

### 3.2 Learning-based energy saver

Specific lifestyles, such as student, office worker, and hospital patients, consists a sequence of repeated activities lasting a specified amount of time. For example, a student may get up at 6:30, walk back and forth at home preparing for school for half an hour, and then ride a bicycle for about 40 min to school on weekdays. There are regular features and correlations between activities. Once the activity routine is identified through long-term monitoring, the Learning-based Energy Saver (LES) can be used. The regular activities and correlations such as duration and sequence are identified, and used to predict the subsequent activity. Scheduling the sensors accordingly allows energy consumption to be further decreased.

### 3.2.1 Learning-based activity pattern prediction algorithm

In this section, we describe a learning-based activity pattern prediction algorithm, and give some necessary definitions.
**Definition 2** Timing Data Set $S_t$: a set of data values for regularly occurring activities. $S_t = \{S_{t_1}, S_{t_2}, \ldots, S_{t_n}\}$, $S_{t_i}$ is calculated by Eq. (4).

\[
S_{t_i} = T_n \times 10^2 + C_n \times 10 + T_{\text{dur}}
\]

where $T_n$ represents the value of the mapping set $M_1$, the key is the activity starting time ranging from 6:00 to 22:00 with corresponding values from 1 to 16. $C_n$ represents the value of the mapping set $M_2$ where the key is one of the nine activity types and the value is from 0 to 8 accordingly. $T_{\text{dur}}$ represents the value of the mapping set $M_3$ where the key is activity duration from 0.5 h to 5 h and the associated value is from 0 to 9.

**Definition 3** Sequences Set $S_a$: a set of sequences for regular consecutive activities. $S_a = \{S_{a_1}, S_{a_2}, \ldots, S_{a_s}\}$ where the count of $S_{a_i}$ is greater than 3.

The learning-based activity pattern prediction algorithm is shown in Algorithm 3.

In Step 1, we define three mapping sets $M_1$, $M_2$, and $M_3$. Note that for $M_1$, the starting times of an activity between 6:00 and 7:00 are all mapped to 1, as are all others. For $M_2$, the mapping rules are given in Table 1. For $M_3$, the duration between 0.5 h and 1 h is rounded down to 0.5 h and mapped to 0. Then, the original $S_{t_i}$ is derived by Eq. (4) for every activity. For example, if a participant walked from 9:20 am for 2.5 h, the $S_{t_i} = 4 \times 10^2 + 3 \times 10 + 4 = 434$. For any participant, there are at most sixteen $S_{t_i}$ recording whole day activity information.

**Algorithm 3: Learning-based activity pattern prediction algorithm**

**Input:** Activity classification results across several days including activity category, starting time, and duration.

**Output:** Timing data set $S_t$ and sequences set $S_a$.

1. Define three mapping sets $M_1$, $M_2$, and $M_3$. Merge any continuous identical activity. For each activity (except for liftup and liftdown) that lasts over 10 min, calculate the original Timing Data Set $S_t$.
2. Sort all $S_{t_i}$ into different groups while $S_{t_i} \leq 10$ are the same in any group. Generate a new Timing Data set $S_t$ for which the amount in each group is greater than $\delta_0$.
3. For each $S'_{t_i}$ in set $S_t$, get the value $[T_{\text{dur}}]$ by averaging $T_{\text{dur}}$ and rounding down to calculate the new value $S''_{t_i} = T_n \times 10^2 + C_n \times 10 + [T_{\text{dur}}]$ constituting the final set $S''_{t_i}$.
4. For each day, give a daily activity type sequence in which the activity lasts over half an hour. The mapping set is also $M_2$.
5. Find the common subsequence $S_{a_i}$ among sequences. If $S_{a_i}$ is greater than $\delta_0$, put $S_{a_i}$ into the set, then get a final set $S_{a}$.

In Step 2, the elements in set $S_t$ are divided into different groups according to $S_{t_i}$ values and a new Timing Data Set $S_t$ is generated while in each group the activities occur in the same time segment. Here, we choose 3 as the threshold of $\delta_0$.

In Step 3, we average and round down the $T_{\text{dur}}$ of each element in $S'_{t_i}$, and derive the new set $S''_{t_i}$. If we have 434, 437, 439 in one group, the average duration $[T_{\text{dur}}]$ of this specific activity (i.e., walk) in this time segment is 6, and the new $S''_{t_i}$ is 436. All $S''_{t_i}$ in each group constitute the final set $S_{a}$.

In Step 4, we use $M_2$ to get the daily activity type sequence. For a three-day experiment, we have three-sequences: 48 358 745, 48 338 245, 483 336 735.

In Step 5, we find the public subsequence using dynamic programming and get the public subsequence $S_{a_i} = 483$. Note that the threshold $\delta_0$ is the required minimal length, and here we set it to 3.

Since the system provides long-term feedback, both the timing data set $S_t$ and sequence set $S_a$ use a feedback calibration method to supplement the latest recognition results and thus achieve greater accuracy.

### 3.2.2 Runtime sampling and feature parameters adaptation

As mentioned in Section 3.2.1, we studied the activity patterns and made activity pattern predictions. Then, we applied a series of dynamic schedule strategies to manage the sensors, adopt optimal sampling rate, and adjust computation cycles in an activity dependent way to recognize users’ physical activities with low power and high precision. The details of the framework are as follows. To specify the state transition logic of our framework, we defined the following two parameters.

**Definition 4** The threshold of switching state $\Delta_{\text{conf}}$. When the confidence value of classifier is above a threshold $\Delta_{\text{conf}}$, we regard the current ongoing activity results as a reliable estimation. If not, we turn to the next recognition phase and switch the state accordingly.

As described in Fig. 7, the first step was to seek the $S_{t_i}$ in timing data set $S_t$ during time segment $n$ after collecting activity data from sensors. If $S_{t_i}$ existed, we obtained the possible activity $a_t$ according to the timing data set $S_t$ and then chose SSR as the rate from the $(a_t, (SR, SF))$ tuple. In this case, if the classification confidence of the calculated activity result $a_t$ was greater than $\Delta_{\text{conf}}$, we adjusted the computing cycle to cycle × (here, $\epsilon = 3$). If $S_{t_i}$ did not exist, we converted to the uncertain state and thus set the sampling rate to
the maximum, i.e., $SR_{\text{max}}$. After recognizing activity $a_1$, the sampling rate was adjusted to the $SR_1$ according to $(a_1, (SR, SF))$ tuple. When the classification confidence of the calculated activity result $a_1$ was less than $\Delta_{\text{conf}}$, we moved to the next step.

In Step 2, we started by seeking $S_{i_j}$ in Sequences Set $S_i$ of activity $a_1$. If $S_{i_j}$ existed, the next activity $a_m$ was estimated and then $SR_m$ was chosen as the sampling rate from $(a_m, (SR, SF))$ tuple accordingly. If not, the activity that occurred in the following time segment $n + 1$ in $S_{i_{n+1}}$ was further considered as the most likely activity and we therefore chose the corresponding sampling rate. When the classification confidence of the calculated activity result $a_m$ was less than $\Delta_{\text{conf}}$, we reverted to the uncertain state, adjusted the sampling rate to $SR_{\text{max}}$, and looped the above steps.

Apart from adaptive adjustment of the sampling rate for the sensors, scheduling different sensors was of great significance. In our system, there were two sensors: a magnetometer and an accelerometer. Activities were divided into two categories and the accelerometer sensor was started. When the system recognized the activity as non-periodic, but not matching the lift pattern, the magnetic sensor was started and used to make further classification decisions. The sensor-scheduling part of the framework is described in the activity recognition section.

### 3.3 Health feedback

Physical activity is a key determinant of energy expenditure. Studies such as the health professionals follow-up study\cite{18}, Women’s Health Initiative\cite{19}, Honolulu Heart Program\cite{20}, and others\cite{21} have demonstrated that simple forms of physical activities, such as walking, jogging, and bicycling, substantially reduce the chances of developing heart disease, stroke, and diabetes in different populations. The amount of physical activities that different individuals need depends on their diet, how much muscle and fat they carry on their frame, and how fit they are.

SmartCare gives quantitative health feedback and activity guidance to users by monitoring long-term activity patterns. It calculates quantitative indicators such as TEI and TEE.

#### 3.3.1 Quantitative measures

Physical activity experts measure activities in different ways, using Based Metabolic Rate (BMR) or MET. The BMR is the minimal rate of energy expenditure per unit time at rest. The MET expresses the ratio of energy...
consumption during a specific physical activity to a reference metabolic rate. The MET of each activity is shown in Table 3.

Once an activity is recognized, its TEE can be calculated using Eq. (5).

\[
\text{TEE} = 1.05 \times \text{MET} \times t \times m \quad (5)
\]

where the unit of TEE is kcal. \( t \) represents the duration of the activity, and the unit of \( t \) is hour. \( m \) is the weight of the user in units of kg. Estimates of METs values based on users’ activities are listed in Table 3. \( v \) is a user’s speed when he is walking or jogging in unit of m/min. The value of \( v \) is determined by Eq. (6).

\[
v = f_{\text{main}} \times w \quad (6)
\]

where \( f_{\text{main}} \) is the dominant stride frequency, i.e., the reciprocal of average period length, which can be easily acquired by adaptive sliding window. And \( w \) is the stride length, a constant input by the user, whose unit is m.

The value of the TEI can be calculated according to Eq. (7).

\[
\text{TEI} = \sum_{i=1}^{n} e_i m_i \quad (7)
\]

where \( e_i \) is calories per kilogram of some food of mass \( m_i \), and \( n \) represents the number of types of the food eaten by the user in one day. To help users determine \( e_i \), SmartCare has stored data on some common food types, though users can add extra food types on the basis of their eating habits.

### 3.3.2 Assessment and suggestion model

Using TEE, TEI, and BMI, SmartCare can generate quantitative physical activity measures for each user. SmartCare divides activity levels into excess, moderate, and insufficient. SmartCare daily activity assessments are shown in Table 4.

The detailed activity level discrimination process is as follows.

- First, SmartCare calculates the BMI value of the user to determine his or her constitution in three categories: thinness, fitness, and fatness.
- For a user in the “thin” category, physical activity is excessive if TEI is less than TEE. In this case, SmartCare will not suggest increasing physical activities. For a user categorized as overweight, the activity level can be determined by the relationship between TEE and TEI depending on two constants ±\( \delta \). As shown in the last column of Table 4, \( \delta \) can be obtained using Basal Metabolic Rate (BMR) and its value can be established through Eq. (8).

\[
\text{BMI} = \begin{cases} 
88.362 + 13.397m + 4.799h - 5.677a, & \text{F;} \\
44.593 + 9.247m + 3.098h - 4.330a, & \text{M} 
\end{cases} 
\quad (8)
\]

where \( m \), \( h \), and \( a \) stand for the user’s weight, height, and age, respectively. Similarly, for an overweight user, the activity level is judged based on two constants 0 and 1100, where the constant 1100 is obtained through the Wishnofsky Constant\[^{[22]}\] and the upper limit of safe weight loss (1 kg/week) is derived from Dickey et al.\[^{[23]}\]

- SmartCare gives suggestions based on activity levels. For example, if the user is currently undertaking insufficient physical activity, SmartCare will make appropriate recommendations.

### 4 Experiment and Performance Evaluation

In this section, we evaluate the performance of SmartCare on the basis of experimental data. Our primary results suggest that SmartCare is able to conduct long-term activity monitoring in an energy-efficient manner and can effectively model all health-related activities. SmartCare can further give quantitative physical health data to users.

SmartCare is implemented on the Android platform using a variety of mobile phones including Google Nexus, Samsung Galaxy S2, and Meizu MX3. Eight volunteers participated in the experiments. The group

---

**Table 3** METs of users’ activity

<table>
<thead>
<tr>
<th>User activity</th>
<th>METs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lift, stationary, escalator</td>
<td>1.5</td>
</tr>
<tr>
<td>Walking</td>
<td>0.0272v + 1.2</td>
</tr>
<tr>
<td>Jogging</td>
<td>0.093v - 4.7</td>
</tr>
<tr>
<td>Bicycling</td>
<td>5.5</td>
</tr>
<tr>
<td>Ascending stairs</td>
<td>8.0</td>
</tr>
<tr>
<td>Descending stairs</td>
<td>3.0</td>
</tr>
</tbody>
</table>

**Table 4** METs’ values based on user activity.

<table>
<thead>
<tr>
<th>BMI (kg/m²)</th>
<th>Body size</th>
<th>TEI – TEE (kcal)</th>
<th>Activity level</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,18.5)</td>
<td>Thinness</td>
<td>(−∞, 0)</td>
<td>Excess</td>
</tr>
<tr>
<td>[18.5, 23.9]</td>
<td>Fitness</td>
<td>(−0, +∞)</td>
<td>Moderate</td>
</tr>
<tr>
<td>[23.9, +∞)</td>
<td>Fatness</td>
<td>(−∞, −0)</td>
<td>Insufficiency</td>
</tr>
</tbody>
</table>

Note: \( \Delta = \text{BMR}/100 \)
comprise two teachers, four students, and two elderly persons. Their age composition was representative of the wider population.

4.1 Performance of the hybrid classifier

This section assesses the performance of our hybrid classifier. To achieve this, we developed an additional application for collecting and storing sensor data in a local file. As mentioned in Section 4.3.1, the hybrid classifier comprised a decision tree classifier and a PNN classifier. The task of the decision tree classifier was to distinguish between non-periodic and the periodic-like activities based on accelerometer data. Magnetometer data was then used to distinguish between lift use, escalator use, and remaining stationary. The PNN classifier was used to distinguish each between different periodic activities.

First, we verified the classification accuracy of the decision tree classifier. Because the decision tree classifier is an unsupervised classifier, the eight participants performed activities of interest, each for a duration of about 10 seconds. Ten samples were collected from each participant for each, yielding 80 samples for each activity. The confusion matrix of the decision tree classifier is shown in Table 5, where 1 represents an activity that can be 100% classified. We found that non-periodic activities could be 100% classified.

Next, we conducted the performance evaluation of the PNN classifier. Four participants performed five periodic activities: ascending stairs, descending stairs, walking, jogging, and bicycling. To make the experiments more representative, we chose two female volunteers and two male volunteers; two were under 30 years old while the others were over 30. Participants performed each activity for 5 minutes with smartphones in hand, on which the sensor data collection application was running.

The first 30 seconds and the last 30 seconds of data were discarded to eliminate irrelevant activities; then, the data was divided into samples with a length of 10 seconds. Six folder cross validation was conducted. Each folder contained five different kinds of activities and each activity consisted of 28 samples. Thus each folder contained $28 \times 5 = 140$ samples in total. During the testing phase, we chose 4 folders as the training set and another 2 folders as the test set. We achieved $15 \times 2 \times 28 = 840$ validations (15 for verification combination, 2 for the validation folder, and 28 for samples of each activity).

In the calculating period of the periodic-like activities, different lengths of sliding window give different recognition accuracies. Figure 8 shows the recognition accuracy for different periodic activities with different sliding window lengths. The red dashed line illustrates average variant trend of recognition accuracy. Recognition accuracy reached a maximum when the frame length was 3. Figure 9 shows the average recognition accuracy of five periodic activities with a different number of periods in any frame. An average accuracy of 98.0% was achieved.

4.2 In-situ user studies

To test whether the accuracy of SmartCare activity recognition varies across different users, we invited recruited eight participants in two groups: trained users and novice users. Members of the trained group came from our own institute and had been involved in the development of the project, while the

<table>
<thead>
<tr>
<th>Activity</th>
<th>Escalator</th>
<th>Stationary</th>
<th>Lift</th>
</tr>
</thead>
<tbody>
<tr>
<td>Escalator</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Stationary</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Lift</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5 Confusion matrix of the classification accuracy by decision tree classifier.

Fig. 8 Recognition accuracy with adaptive slide window.

Fig. 9 Measure accuracy with adaptive slide window.
novices were unfamiliar with the project. Before the experiment, the novices were given instructions on how to use the system. Each group member was tested 20 times for each activity, which comprised two phone states, three non-periodic activities, and five periodic activities. In calculating recognition accuracy, the user was asked to confirm the accuracy of the result. If user’s assessment disagreed with that of SmartCare, the user inputs the ground truth and SmartCare then stored the recognition time, recognition result, and ground truth in a local file.

Figure 10 shows the experimental results. The two columns represent the two groups: trained and novice, respectively in each group-bar. The red dashed line describes average accuracy. We found that the accuracy of the novice users was lower than that of the trained users. This was expected. In particular, when the two groups were testing the stationary state, the trained users did nothing, while the novice users would move around occasionally. Recognition accuracy was 100% in both groups: lifting, staying stationary, jogging, and walking. The accuracy of escalator use was 100% in the trained group but significantly lower (about 96.0%) in the novice group. The reason is that the novice users would walk unconsciously when riding an escalator. The accuracies across all activities are very high except for ascending, where the average was about 87.0%. This is understandable as ascending and walking have many similarities, ultimately leading to misclassification. Overall in this experiment, average accuracy was 98.0%.

4.3 Long-term studies for users

Long-term studies for users focus on performance on energy saving and physical health feedback. Eight participants used SmartCare for more than 4 weeks. Volunteers carried their smartphones with SmartCare installed and conducted their daily activities as normal. SmartCare captured all health-related activities of all participants with an accuracy of 98.0% and stored the data for further analysis. We treated continuous activity mixed with other sporadic or short activities as one. Figure 11 shows the total activity instances of each user.

4.3.1 Energy saving

As a long-term physical activity monitoring system, SmartCare explores the trade-off between energy consumption and recognition accuracy. On the basis of good classification accuracy, SmartCare seeks energy savings in two ways. First, SmartCare applies optimal combinations of classification features and sampling rates for each activity. Second, SmartCare uses learning-based activity prediction to set optimal initial sampling rates when a user finishes one activity and switches to another.

We conducted five experiments, four without activity-dependent parameter selection and one with activity-dependent parameter selection. Specifically, for the activity-dependent parameters selection cases, the sampling rate was fixed from one of the values, 5 Hz, 16 Hz, 50 Hz, and 100 Hz. The final case used the optimal activity-dependent parameters combined with an LES. The energy consumption of the LES was approximately 50% less than that of the 100 Hz mode and slightly less than that of the 16 Hz (Table 6).

We further compared battery consumption across three scenarios to assess the energy-saving performance of the LES. The first scenario case was no SmartCare system running on the phone, the second was SmartCare running with LES, and the third was SmartCare running without LES. The non-LES case used a sampling rate of 50 Hz and classification features using both time- and frequency-domain features. We also designed an extra application to record the runtime battery situation. The battery is required to charge overnight so as to ensure identical starting conditions. Figure 12 shows the power drainage time series for Participant A in the three scenarios (1 day per scenario) in the first week of the experiment.
Table 6  The energy consumption of three users in five scenarios.

<table>
<thead>
<tr>
<th>User</th>
<th>Total duration (h)</th>
<th>Energy consumption (kJ)</th>
<th>SmartCare (kJ)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5 Hz</td>
<td>16 Hz</td>
<td>50 Hz</td>
</tr>
<tr>
<td>A</td>
<td>180.5</td>
<td>13.844</td>
<td>22.129</td>
</tr>
<tr>
<td>B</td>
<td>169.8</td>
<td>13.023</td>
<td>20.817</td>
</tr>
<tr>
<td>C</td>
<td>150.4</td>
<td>11.535</td>
<td>18.439</td>
</tr>
</tbody>
</table>

Fig. 12  The SmartCare with LES can save energy efficiently.

As learning and training was continuously accumulated and optimized, the accuracy of activities prediction improved, leading to continuously decreasing energy consumption. We conducted a group of contrasive experiments to show the extent of energy saving changes over time. Figure 13 shows the energy drainage time series for Participant B in the four test conditions. SmartCare-LES2 represents the condition in which SmartCare was running with LES in the second week. Energy consumption was lower than that of the first week. Again, energy saving increases as experience is accumulated.

4.3.2 Health assessment and advice

In this section, we assess the performance of SmartCare in constructing health assessment and making activity recommendations. We discuss the example of 24-year-old participant whose height is 168 cm. SmartCare recorded his physical activity diary for four weeks and provided several visual and quantitative assessments and suggestions for activities.

Figure 14a shows his TEE this month and all of his physical activities today. The detail of today’s physical activities is shown in Fig. 14c: the participant has jogged for 0.92 h and walked for 3.72 h, has been static for 1.2 h, and has ascended stairs for 0.31 h. The TEE for each activity is calculated using Eq. (5). For example, as today’s duration of jogging is 0.92 h and the METS equals \(0.093v - 4.7\), the rate of jog \(v\) is about 100 m/min. SmartCare provides the user with the assessment and makes suggestion about further activity. In this case, based on the values of TEI, TEE, and BMI, the activity level today is already excessive, as shown at the bottom of Fig. 14c.

Figure 14b is a physical activities distribution of this week including nine activities in total. The TEE curve is shown in the lower portion of Fig. 14b and the incentive provided by of SmartCare’s feedback can be seen. The excess physical activities on Monday leads to a reduction of activities from Tuesday to Thursday. SmartCare has helped this participant to choose good physical activities.

5 Discussion

SmartCare is designed to detect nine basic activities sequentially. In the current study, SmartCare can recognize nine activities with an accuracy of 98.0%. Since complex activities were occurring simultaneously, these recognition results may be not reliable. As a long-term physical activity monitoring system, SmartCare focuses on the problem of recognizing nine health-related activities and provides a
quantitative feedback to users. What’s more, SmartCare was also shown to resolve the trade-off between energy saving and recognition accuracy. In practice, there is inevitably a tension between recognition accuracy and energy saving. In future, SmartCare needs to be extended to capture more health-related activities and to pay more attention to energy saving.

6 Related Work

Mobile health care that involves activity monitoring and physical health guidance using activity recognition has achieved sustainable development for several years [1]. A comprehensive survey was presented in Ref. [24]. Activity recognition systems that use on-body wearable sensors have been already investigated [25–27]. The latest work [28] implemented an assistive system based on Google Glass devices to perform real-time scene interpretation combining the first-person image capture and sensing capabilities of Glass with remote processing. Other works include a wearable ring platform to detect finger motion on a surface [2] or other specialized devices including a wristband containing an inertial measurement unit to capture changes in the orientation of a person’s arm and accurately detect smoking gestures [3, 4].

In addition, many commercial activity recognition and health monitoring wearable devices, such as Apple Watch [8] and Fitbit [9], are currently available. Fitbit utilizes multiple sensors to track users’ activities including walking, running, taking the stairs, and certain kinds of sports like golf. It then calculates the movement distance and calories consumed based on heart rate. Apple Watch provides activity tracking with similar functions, sets out exercise goals for users, and calculates the amount of exercises and calories consumed using accelerometer data. Even with portability and good battery life, there are still inevitable limitations on wearable products. Although there may be small screens on the wristband, the details on the screen are hard to read. The accumulated daily exercise information is sent to other devices like phones or pads through Bluetooth or wireless networks and is then analyzed and processed in a visualized, interactive, and comprehensible way. Without these additional devices, the wristband can only have simple functions and be used as a real-time activity collector. It is thus unable to play a role in long-term health monitoring. Also, the uncertainty of wireless network availability and the restricted support for Bluetooth on device systems (above iOS 5 and Android 4.3) mean that many devices are unable to connect with the wristband, forming a bottleneck for continuous and long-term activity monitoring and analysis.

Given their popularity and rich array of built-in sensors, smartphones provide an alternative to long-term activity tracking. Many Android and iOS Apps, such as Moves [13], Runtastic Me [28], Nike+ Running [29], can track daily activities such as walking and bicycling using GPS or accelerometers, and some provide data on calories burned and other results. However, these Apps lack careful algorithm design. Many are unable to guarantee the desired recognition accuracy while saving energy, as the use of GPS and lack of power monitoring make them unable to run for long periods of time.

Recent research efforts have aimed at improving recognition accuracy and decreasing power consumption. Several studies focus on contextualizing activities using multi-sensors with high accuracy. Han et al. [4] used accelerometers, gyroscopes, and GPS to detect indoor or outdoor location-based activities such as walking or waiting for a bus. Similar work [30, 31] recognized fixed activity patterns while comparing different classification algorithms to achieve higher accuracy. A recent study [11] presents an alternative approach for context detection using only the smartphone’s barometer to detect the basic activities like idle, walking, and vehicle. Lane et al. [12] proposed using networked community behavior combining social ties with single individual activities based on relational learning techniques. With increasingly complex activities recognized, the limited battery life of smartphones presents a major barrier [32]. Several studies have proposed different energy-efficient strategies. Early studies [15] comprised a set of sensing pipelines for the accelerometer, microphone, and GPS sensors giving adaptive control and on-demand processing. Later studies [16] simplified the problem by powering only the necessary and most energy efficient sensors and managing the sensors hierarchically to recognize the user state as well as detecting state transitions. Chu et al. [17] concentrated on balancing energy and accuracy by exploring Kobe to optimize mobile classifiers, whereas in the present study, we used only two mobile sensors. Yan et al. [10] used only accelerometer data to infer locomotive activities by adjusting sampling rates in real-time, and the study in Ref. [33] presented a novel method by leveraging
the predictability of human behavior to identify likely future activities and temporarily turn off the non related sensors.

In our work, we focus on balancing the trade-off between accuracy and energy consumption by adjusting the sampling rate in an activity-dependent way and propose an energy-efficient classification strategy based on learning-based activity pattern. In addition, we provide personalized quantitative health feedback to users by calculating caloric intake and expenditure\cite{61272456, 61472312}, the open fund ITD-

7 Conclusions

In this study, we present SmartCare, a health-care system capable of recording a user’s nine physical activities and providing personalized suggestions during long-term running on a smartphone. To balance the trade-off between recognition accuracy and energy consumption, the key components of our system, i.e., activity-dependent parameters selection, hybrid classifier, and an LES, are proposed. Quantitative health feedback is provided based on users’ daily physical activities. We implemented SmartCare on an Android platform and ran it with data collected over four weeks with eight participants. The results show that our system achieves desirable detection accuracy and can run over a long time period as battery consumption is significantly reduced.

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China (Nos. 61190110, 61272456, and 61472312), the open fund ITD-U14004/KX142600011. This work was also supported by the overall innovation project of Shaanxi Province Science and Technology Plan (No. 2012KTZD02-03-03), and the Fundamental Research Funds for the Central Universities (Nos. JB151002, K5051323005, and BDY041409).

References


