Increasing Energy Efficiency on Smartphones through Data forecashing

Jiuyun Xu, Dan Yang,
Jiazhen Wang, Chao Guan
School of Comp. & Commun. Eng.
China University of Petroleum
Qingdao, Shandong 266580
Email: jiuyun.xu@ieee.org

Stephan Reiff-Marganiec
Department of Computer Science
University of Leicester
LE1 7RH, Leicester, UK
Email: srm13@le.ac.uk

Huilin Shen
School of Geoscience
China University of Petroleum
Qingdao, Shandong, China 266580
email:hlshen@upc.edu.cn

Abstract—**Smartphones are widely used in daily life to access services and various functions require continuous communication, which leads to increased energy consumption. However, the development of battery and related energy saving technology can not meet the demand for energy consumption. Much of current research work focuses on energy models caring much about energy consumption of every single application. In this paper, we propose a data forecasting-based strategy for increasing energy efficiency on smartphones based on the predictability of data to be accessed. To achieve this, a combination of Collaborative filtering with the k-means algorithm categorize users with similar user groups and speculate use increased for the data users will access. With this model, we also adopt data pre-storing model and dynamic updating model. The simulation results illustrate that our approach is leading to energy saving.**

Index Terms—**energy saving; Data Forecasting; Similarity; k-means Algorithm; Collaborative filtering recommendation Algorithm; data pre-storing model**

I. INTRODUCTION

In the recent years, smartphones are developing rapidly and becoming more and more an indispensable tool in our daily life. Thousands of new applications are deployed everyday. With the diversification of the phone functions, the demand for accessing the Internet using phones is increasing, which in turn leads to a massive increase in energy consumption. Unfortunately, the manufacturing technology for enhancing battery power lags behind, wherever there limitations of battery energy are becoming a bottleneck to smartphone applications to some extent. So reducing the energy consumption in different ways is a research topic requiring solutions to the success of mobile services. Sensing data of the core energy consumption of applications in the smartphones is the of much research work, with power models and sensing data models for single smartphones applications energy.

However, in a smartphone, there is more than one application working. In reality, it is very rational to everyone that smartphone users generally access related data in the same scenarios. So, considering the energy efficiency issue, in this paper we think about reducing the times of connecting to the data source by pre-fetching the data according to similar aspects of a group of similarities user while it might seem counter-intuitive to transfer more data 'on a bunch'

we will show that reducing the number of connections can indeed reduce the energy consumption. To the best of our knowledge, it is the first time to achieve the energy efficiency by considering the times of connecting to the service.

To illustrate our motivation, let us consider the following scenario. In the space of travel, people go to and engage in activities. Customers often have smartphones with many mobile applications. These mobile applications have been used anywhere and in any times. Especially, some travelling apps are very popular for many people. Using these travelling apps people can do a lot of travel related activities, such as search the attract on tickets, book hotels near attractions and others. In this scenario, one user, user A (denoted as U_A) uses a travelling app to search for tickets of the attraction C (denoted as A_C) in the area B (denoted as P_B), while U_B will search for tickets to other attractions in P_B. In most cases U_A will want to choose an accommodation and search the information of hotels in P_B. So the user uses a smartphone to access the data from Internet via wifi and/or sensing data from GPS, which indicate establishing a connection with the server to get the related information. Usually, this process is conducted in two steps. First, user U_A sets up a connection of the Internet to access the resources. And then user U_B also needs to set up the connection to the Internet in the latter time. In this setting it is easy to see there are consume energy due to multiple connection.

In this article, there are two factors taken into consideration. One is that we can identify the groups of similar users according to ones individual information. The other one is that the data of this group is likely to be the data the one will access. Based on the above two factors, we propose a data forecasting-based model for energy efficiency on smartphones. In our model, we use the k-means algorithm to find similar users according to users individual information and use a collaborative filtering recommendation algorithm to forecast the data that users will access according to the data of these similarity group. Then we pre-store the forecast data in the smartphone on first connection in order to acquire data rapidly and accurately when user searches for data late.

In this paper we have made two main contributions as follows: **First**, a data pre-storing model and dynamic updating

model is adopted. Through ensuring the data accuracy, we achieve the goal of energy saving in the way of shortening the communication time and reducing the communication frequency.

Second, With the proposed model, a combination of k-means algorithm and the collaborative filtering recommendation algorithm is proposed to forecast the data that user wants to access in the next step.

Third, evaluation experiments has been conducted using open source tools. With our proposed model and the combination of k-mean and collaborative filtering algorithm, the energy consumption is reduced by more than 15%.

The remainder of the paper is organized as follows: Section II provides an overview of the framework of energy efficiency for smartphones, Section III describes the proposed approach using the combination of k-means algorithm and collaborative filtering method. With the open source tools an experimental analysis study is detailed in section IV. Section V discusses the related work and Section VI concludes the paper.

II. SYSTEM MODEL

A. The framework of Data forecasting-based strategy model for energy efficiency

To make the smartphones more energy efficient, we have proposed a data forecasting-based model. The core function of this model is using the knowledge of characteristic of smartphone users having similar sequential data requirements, which is reducing the times of accessing these data by using fewer requests. To achieve this intention, the data forecasting-based model is using three modules.

In this section, we will give an overview of our proposed model first. Then the detail of our idea of workflow will be presented. Figure.1 depicts the data forecasting-based strategy model for energy efficiency on smartphones which we have proposed. In this model, there are three parts to be considered as main functions for achieving the energy efficiency. In our proposed model, There are some information to be handled. First, we need to calculate the relationship among the users, while the logs of user accessing the resources of Internet are recorded. Then we can pre-fetch the following the Internet resources which the user had been going to visit according the users' trace to store the user's smartphones. There are three parts in our model.

These are the **Data Prediction Module** and the **Data Storage Module** which are located in the Internet server (some of them would be cloud service), and **Data Pre-Storage Module** which is located in normal smartphones. In more detail:

Server Data Storage Module (SerDSM): The part has two main functions. One is storing all the Internet resources, including the known Internet resources and the data in the data prediction center. The other is recoding the user's information when the user uses the smartphone to establish communication, including the user's individual information and behavior information. Considering the privacy protection

issues, we use the user's insensitive information with the Cloaking technology.
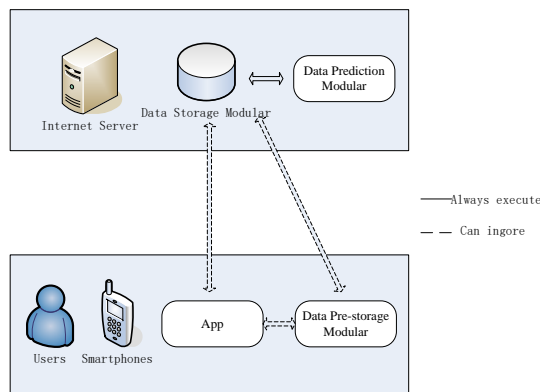

Fig. 1: A Data forecasting-based strategy model for energy efficiency

Server Data Prediction Module (SerDPM): This module is responsible for the classifying known users into groups using k-means algorithm, finding out the similarity of users, forecasting the data a user will access using the collaborative filtering recommendation algorithm according to the similar user group's accessed data and storing the data into the data storage center. The content will be introduced in detail in the next section.

Smartphone data pre-storage Module (SmaDPM): This module is responsible for pre-storing the data from server. When the user searches data for the second time, it plays a role of local cache. Its main contribution is reading the pre-stored data and providing data directly for the user in order to reduce energy consumption due to second establishment of a connection with the server. The premise of this module is ensuring the user experience.

III. THE ALGORITHM OF SERVER DATA PREDICTION MODULE

A. Clustering of similar user group

In the real world, each smartphone user is usually registering with certain kinds of user group, which is consisting of many primitive types of data. Considering these, we identify what kind of user we are dealing with according to the value of age, gender and some other characteristic data. In this paper, we adopt myriad data types-based k-means algorithm to clustering the smartphone users into different user groups, which can predict the data their required in the next step. Figure 2 shows the workflow of establishment of clustering analysis classification model.

To make our proposed model work, Server Data Storage Module (SerDSM) and Server Data Prediction Module (SerDPM) are located in a Server, which is as an Internet node with powerful storage and computation capacity, while the Smartphone data pre-storage Module (SmaDPM) is located in each user's smartphone.

On the Server side in this model, there is a combination of k-means and collaborative filtering algorithm adopted for
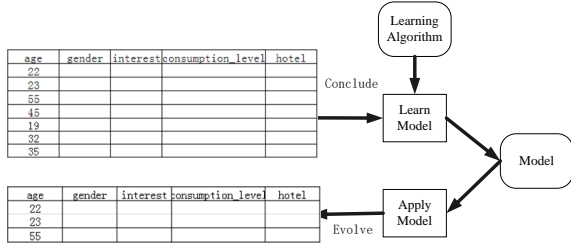
Fig. 2: workflow of establishment of clustering analysis classification model

the recommendation of pre-storage data transferred when a smartphone user requests his(her) data.

In this paper, we use a k-means algorithm to classify user groups according to the similarity [1],[2] of known users. The data is prepared for the next collaborative filtering recommendation. Usually, we consider a person's background (for example, age, hobbies, consumption and favourite things) as influencing the information she searches on the phone to a great extent. In other words, users with similar attributes are likely to search for the same things in the same area. Therefore, it is of great significance to the building of similar user groups.

Firstly, considering the diversity of the user's individual information type, we convert the value of the attributes into the interval $[0, 1]$. The user's individual information (age, gender, hobby, consumption level, favourite hotel type) is denoted as $U = (A, G, I, C, H)$. The search history of the user is denoted as (K, R). All the information can be obtained from the data storage center. Then we use the k-means algorithm to conduct cluster analysis on known users and get the final similar user group. The k-means algorithm used is shown as Algorithm. 1. In our experiment, considering the effectiveness of local data in smartphone and the assumption of users' continuing requiring the dependent data without any deviation, we update the pre-fetched local data every half an hour regularly between smartphone and the server.

After cluster analysis, we will use the collaborative filtering recommendation algorithm to forecast the data a user will access according to the data of similar user group.

B. Design of phone data pre-storage

As the smartphone side: The phone data pre-storage center described is implemented as Android app. The design prototype is shown in Figure. 3.

The purpose of the design of data pre-storage center is to cache the data requested this time and the data that the server predicts the user will access next time when the user uses the app to search for information. When the user uses the app to search for information for a second time, the phone first acquires related data from the data pre-storage center. If the data pre-storage center already has that data, it returns the data and the phone does not connect to the server. The workflow of the data pre-storage center is shown in Algorithm. 3.

Algorithm 1: Establishment of similar user group using K-Means

Input: C_U: The known user set; U: user attributes u, k: cluster number

Output: $\{\mathbf{C_i}, i = 1, k\}$: clusters

1 **for** $random \quad k \quad u_i \in CU$ **do**
2    $clusters\_center \leftarrow random k u_i$;

3 **while** $true$ **do**
4   **for** $u_i \in CU$ **do**
5   **for** $j < k$ **do**
6        $distance[i][j] \leftarrow$ $\sqrt{\sum(u_{im} - cluster\_center_{j_m})^2}$;

7   **for** $u_i \in CU$ **do**
8   **for** $j < k$ **do**
9   **if** $distance[i][j] ==$ *smallest* **then**
10          $clusters\langle j \rangle \leftarrow u_i$;
11   calculate new $clusters\_center$;

12   **if** $new\_clusters\_center == old\_clusters\_center$ **then**
13   **break**;
14   **else**
15        $clusters\_center \leftarrow$ **new** $clusters\_center$;

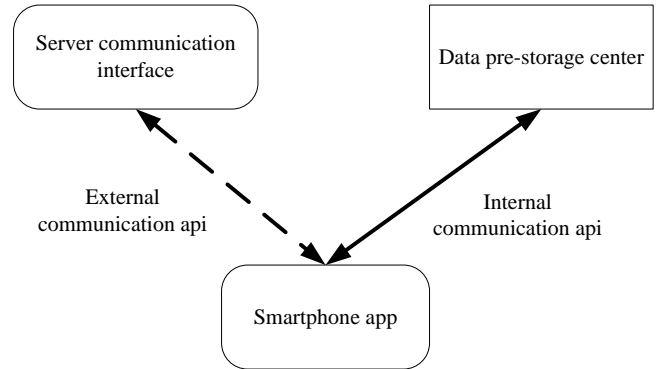16 **return** $\{\mathbf{C_i}, i = 1, k\}$*: clusters*;

Fig. 3: Phone Data pre-storage center

IV. EXPERIMENTS AND ANALYSIS

Due to the lack of a benchmark dataset for this kind of work, we used the social network dataset to extract the users' individual information and combined this with the scenic spot's data information in a travel website. Based on the data, we conducted on the simulation experiment according to our energy saving-based model. In addition, we adopted the real data type in travel app in order to ensure the authenticity of the data.

A. The Establishment of similar user group

We applied IBM Predictive Analytic Software SPSS using k-means algorithm to conduct on cluster analysis based on

Algorithm 2: Selecting pre-storage data with collaborative filtering

---

Input: {**U**: the attributes of user u;
M: The attribute set with m users which user u is belonging to same group according to Algorithm 1;
H:the value set of interesting to keyword k that each user corresponding to the attribute set M;
Output: $\{< Key_i, value_i >, i = 1, p\}$
```
/* the subset C_v which is the subset
   with similar characteristics of
   user u in {C_i, i = 1, k}        */
```
1 forall the $v_i \in M$ **do**
2 $T \leftarrow 0$;
3 $T_\delta \leftarrow 0$;
4   **for** $f \leftarrow 1, p$ **do**
```
        /* p: denotes that user has p
           attributes              */
```
5   **if** $((a_u^f \vee a_v^f == \emptyset) \vee (a_u^f \wedge a_v^f == 0))$ **then**
6 $\delta_{uv}^f \leftarrow 0$;
7   **else** $\delta_{uv}^f \leftarrow 1$;
8   **if** $a_u^f \in \{SetofBi\_value\}$ **then**
9 $d_{uv}^f \leftarrow 0$;
10   **else**
11 $d_{uv}^f \leftarrow 1$;
12   **if** $a_u^f \in \{SetofNumerical\}$ **then**
13 $d_{uv}^{(f)} \leftarrow \frac{|a_u^f - a_v^f|}{(max(a_h^f) - min(a_h^f))}$;
```
           /* h assign each one of all
              possible values of the
              variable f            */
```
14 $T+ \leftarrow \delta_{uv}^f d_{uv}^f$;
15 $T_\delta+ \leftarrow d_{uv}^f$;
16 $duv_i \leftarrow \frac{T}{T_\delta}$;
17 $W_{uv_i} \leftarrow \frac{1}{d_{uv_i}}$;
18 $p_{uk} \leftarrow \sum_{i=1}^m W_{uv_i} H_{v_i k}$;
```
   /* p_uk is an element with the value
      like < key, value>;           */
```
19 $P_{sorted} \leftarrow SortTo(\{P | \forall p_{uk} \in P\})$;
```
   /* Sort the data set P.           */
```
20 $Data(u) \leftarrow First(P_{sorted})$;
```
   /* Select first of P_sorted as a possible
      data which user u need.        */
```
21 return $Data(u)$. /* Return the data $Data(u)$ the user u is possible to access in next step. */

Algorithm 3: Data matching algorithm in data pre-storage center

---

Input: K: Search keyword, $Map\langle key, data \rangle$: Data set in data pre-storage center
Output: Relevant data RD with keyword
1 if *Map contains Key* K **then**
2   search RD via k locally ;
3 $Map.get(K)$;
4 $close(establishment)$;
5 $RD \leftarrow GetRD(K)$;
6   **else**
7 $open(establishment)$;
8   select data RD from the Server side;
9 return RD;

---

the training set. In our experiment, we classify user groups according to hobby consumption level and favourite hotel type mainly. Figure. 4a shows part of input data. Due to the large amount of data, when $k = 4$, part of output result is shown in Figure. 4b. The row QCL_1 shows the grouping of cluster members. The row QCL_2 shows the distance between the object and the center of clustering. Figure. 4b shows that the users with similar hobbies or similar favourite hotel types are likely to be classified into one group. In other words, age and gender have little influence on the classification. The result is consistent with the expected, similar users are likely to acquire the same or similar data. Then we adopt the user's predicted data to replace the data user will access, which can reduce the times user communicates with server and can achieve our goal accordingly.

B. Energy Saving-based Model

We conducted on comparasion experiment. One acquires data from server directly, and one adds data prediction center and data pre-storage model. We used the PowerTutor software [3] for our power measurements.

For our experiments, we used a Google Nexus 5 phone running Android OS 4.4.4 and compare our model with normal Android app for the power consumed in the same environment. Figure 5 shows the results. The result of prediction and pre-storage-based energy saving strategy is consistent with the expected sequence and can reduce energy consumption in the same environment. Figure. 5a shows the result of normal Android app the first time to acquire data. Figure. 5b shows the result of a normal Android app the second time. Figure. 6a shows the result of Android app for which we add prediction and pre-storage-based energy saving strategy for the first time to acquire data. Figure. 6b shows the result of Android app which we add prediction and pre-storage-based energy saving strategy the second time in the same environment.

With the comparison the figure. 5 and figure. 6, we can find that the prototype system with added prediction and pre-storage-based energy saving strategy consumes lower energy

	age	gender	interest	consumption_level	hotel_type
1	32	male	hot spring	1	budget
2	25	female	hot spring	3	three-star
3	19	female	skiing	2	two-star
4	23	female	amusement park	4	three-star
5	26	male	pavilion	4	four-star
6	48	female	cultural relics	2	budget
7	56	female	amusement park	2	three-star
8	37	male	skiing	2	budget
9	29	female	cultural relics	1	budget
10	40	male	cultural relics	4	five-star
11	25	female	pavilion	2	two-star
12	45	male	natural scenery	2	budget

(a) part of input data

	age	gender	interest	consumption_level	hotel_type	QCL_1	QCL_2
1	32	male	hot spring	1	budget	1	.00000
2	25	female	hot spring	3	three-star	2	1.00000
3	19	female	skiing	2	two-star	2	1.73205
4	23	female	amusement park	4	three-star	2	2.00000
5	26	male	pavilion	4	four-star	4	1.41421
6	48	female	cultural relics	2	budget	3	1.00000
7	56	female	amusement park	2	three-star	2	2.82843
8	37	male	skiing	2	budget	1	1.41421
9	29	female	cultural relics	1	budget	3	1.41421
10	40	male	cultural relics	4	five-star	4	.00000
11	25	female	pavilion	2	two-star	2	2.23607
12	45	male	natural scenery	2	budget	3	.00000

(b) part of output result

Fig. 4: Range set and efficient points set of Bi-Objective Optimal.
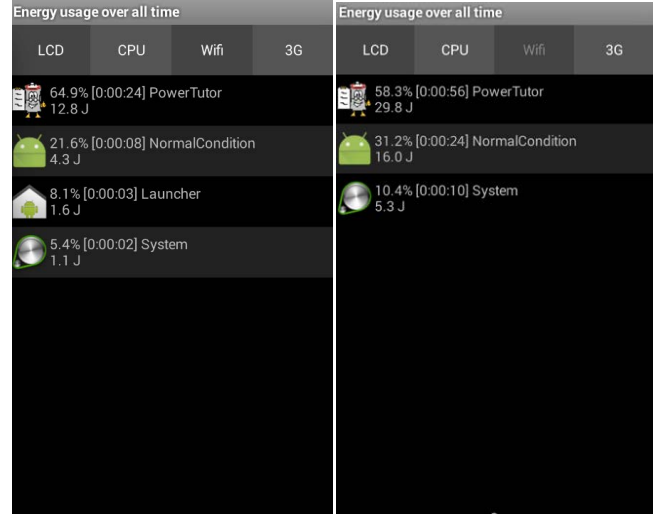
TABLE I: Comparison of the energy consumption

Method	Energy consumption of 1st time to access(Joules)	Energy consumption of second time to access(Joules)	Total
Normal	4.3	16.0	20.3
Predict with pre-store	5.3	12.3	17.6

(a) Smartphone's Energy Consumption in the first time normally
(b) Smartphone's Energy Consumption in the second time normally

Fig. 5: The experimental result of Smartphone's energy consumption normally
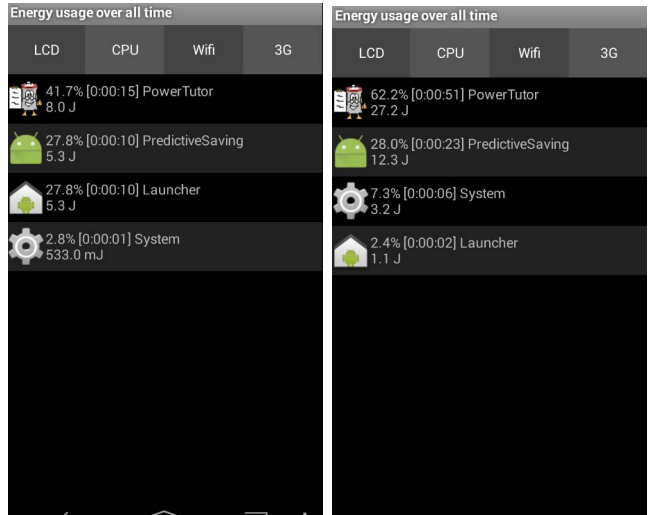
(a) Smartphone's Energy Consumption of our model in the first time
(b) Smartphone's Energy Consumption of our model in the second time

Fig. 6: The experimental result of Smartphone's energy consumption with our method

overall even though it consumes more energy than normal Android app the first time to acquire data. The experiment was run about 10 times and we are presenting averages here. Table. I shows the comparison of energy consumptions in our experiments. We can find that using prediction and pre-storage-based energy saving strategy can reduce energy consumption by 13.3%.

V. RELATED WORK

With the increasing popularity of smartphones, researchers pay more attention to the corresponding energy saving problem. According to the issue, the current solutions are divided into two aspects. One is finding solutions based on hardware, adjusting the power and status according to different demand. The other is reducing the energy consumption of apps through software. Most researchers focus on the software aspect.

In the past few years, much research work has focused on the engergy efficiency of smartphones. Brakmo et al.[4] proposes a Sleep energy saving technology, aiming at using the method of turning the status of CPU into sleepiness rapidly to achieve energy saving. Shih et al. [5] present wake-on-wireless energy saving strategy, controlling the status of phone to reduce the energy consumption of free time. Lin et al. [6] propose an A-loc location system. It chose lower power positioning method by location-based and positioning-based modeling. In [7], Deblauwe at al. adopt the way of switching between higher power and lower power sensors to solve the issue of higher power sensors.Kjrgaad et al. [8] believe that the method of reducing the times and
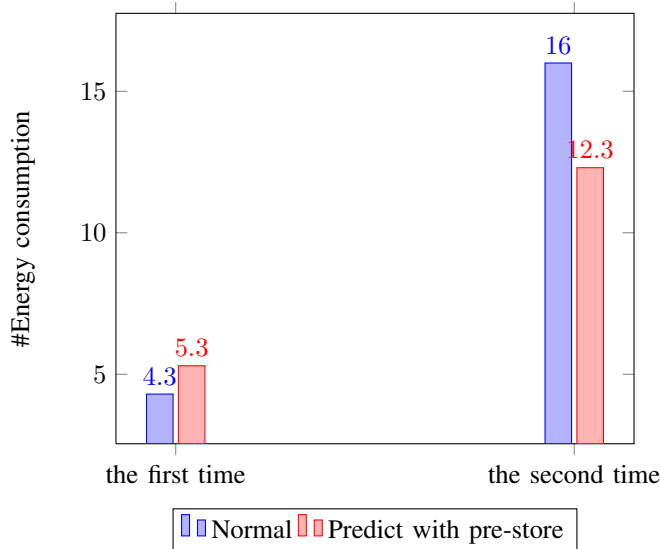
Fig. 7: Comparison of the energy consumption

frequency of positioning can reduce energy consumption.

In addition, Lim [9] introduces the ACQUA framework. The framework replaces the original push model with a pull model to reduce energy consumption due to data transmission. In [10], Nath et al. propose ACE (acquisitional context engine). The system adopts two optimal strategies: the inference caching and speculative sensing. In [11], Burgstahler et al. presents the idea of switching between the push model and pull model in real time. Moamen et al. [12] emphasize that sharing of overlapping feeds between applications can reduce the consumption.

The above solutions do not take account on the influence of the predictability of user's behavior information in the smartphone energy saving and the influence of data pre-storage.

The data forecasting-based strategy model for energy efficiency on smartphones we propose mainly forecasts the data user will access based on the similarity of users and these reduces the times and frequency of communication with the server using data pre-storage to achieve the goal of energy saving.

VI. Conclusion

In this paper, we proposed a model for data pre-fetching for smartphones accessing the expected data in order to reducing the energy consumption. With this model, we present a combination of k-means and collaborative filtering recommendation algorithm to pre-store the data for smartphone user.

To verify the effect of our model, we have conducted a comparative experiment. The experimental result shows that this model is reducing energy consumption with the help of similar characteristics (such as same context etc.). The strategy we propose can achieve the expected goal and ensures the user satisfaction in some well defined environment. In our

experiment, using the prediction strategies of our proposed model, it has led to a significant saving of energy (about 13%).

In the next step, we will consider management approaches to updating usage logs to ensure that this frequent task does not consume undue amounts of resources. Performances of the model on applications will be explored in the future. Finally, we will consider the impact on privacy through the extra data gathering.

References

[1] A. Parashar and J. K. Chhabra, "Clustering dynamic class coupling data to measure class reusability pattern," in *High Performance Architecture and Grid Computing - International Conference, HPAGC 2011, Chandigarh, India, July 19-20, 2011. Proceedings*, 2011, pp. 126–130. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-22577-2_17

[2] A. Verma and Chetna, "Dynamic document clustering using k mean with cosine similarity for effective software test cases reuse," *International Journal of Applied Engineering Research*, vol. 10, no. 15, pp. 36 046–36 054, 2015.

[3] L. Zhang, B. Tiwana, Z. Qian, Z. Wang, R. P. Dick, Z. M. Mao, and L. Yang, "Accurate online power estimation and automatic battery behavior based power model generation for smartphones," in *Proceedings of the eighth IEEE/ACM/IFIP international conference on Hardware/software codesign and system synthesis.* ACM, 2010, pp. 105–114.

[4] L. S. Brakmo, D. A. Wallach, and M. A. Viredaz, "μsleep: a technique for reducing energy consumption in handheld devices," in *Proceedings of the 2nd international conference on Mobile systems, applications, and services.* ACM, 2004, pp. 12–22.

[5] E. Shih, P. Bahl, and M. J. Sinclair, "Wake on wireless: an event driven energy saving strategy for battery operated devices," *Acm Mobicom*, pp. 160–171, 2002.

[6] K. Lin, A. Kansal, D. Lymberopoulos, and F. Zhao, "Energy-accuracy trade-off for continuous mobile device location," in *Proceedings of the 8th international conference on Mobile systems, applications, and services.* ACM, 2010, pp. 285–298.

[7] N. Deblauwe and P. Ruppel, "Combining gps and gsm cell-id positioning for proactive location-based services," in *Fourth International Conference on Mobile & Ubiquitous Systems: Networking&services*, 2007, pp. 1–7.

[8] M. B. Kjrgaard, J. Langdal, T. Godsk, and T. Toftkjr, "Demonstrating entracked a system for energy-efficient position tracking for mobile devices," in *Ubiquitous Computing/Handheld and Ubiquitous Computing*, 2010, pp. 367–368.

[9] L. Lim and A. Misra, "Optimizing sensor data acquisition for energy-efficient smartphone-based continuous event processing," in *Proceedings of the 2011 IEEE 12th International Conference on Mobile Data Management - Volume 01*, 2011, pp. 88–97.

[10] S. Nath, "Ace: exploiting correlation for energy-efficient and continuous context sensing," *IEEE Transactions on Mobile Computing*, vol. 12, no. 8, pp. 1472–1486, 2012.

[11] D. Burgstahler, N. Richerzhagen, F. Englert, R. Hans, and R. Steinmetz, "Switching push and pull: An energy efficient notification approach," in *Mobile Services (MS), 2014 IEEE International Conference on.* IEEE, 2014, pp. 68–75.

[12] A. A. Moamen and N. Jamali, "Sharesens: An approach to optimizing energy consumption of continuous mobile sensing workloads," in *Mobile Services (MS), 2015 IEEE International Conference on*, 2015.