Collaborative Mobile E-Health Employing Data Mining and Social Networking

Yidan Liu, Xiping Hu, Chunsheng Zhu, Boon-Chong Seet, Victor C. M. Leung, Terry H. S. Chu and Henry C. B. Chan

Abstract—In this paper, we present a novel design and implementation of a mobile e-health system that employs mobile data mining to reduce the cost of processing sensing data, and social networking to facilitate collaborations between patients and care-givers. Our design integrates a Bluetooth v4.0 wireless radio, mobile data mining, and social networks techniques to provide a power-efficient solution for mobile e-health that optimizes the processing of multiple and diverse sensing data in mobile devices, and supports group-level activities for collaborative mobile e-health. Experimental results show the effectiveness of our proposed system.

Index Terms—Mobile e-health; Bluetooth; Mobile data mining; Social network

I. INTRODUCTION

RAPID developments in information and communication technologies have improved people’s lives in many ways, among which the progress made in mobile e-health may be the most significant. Wearable wireless sensors and mobile devices are used to help provide instant and secure access to the health status of users regardless of time and location. These devices are able to collect a patient’s physiological data as well as data regarding the patient’s activities [1]. Therefore, mobile e-health systems are increasingly being utilized to enable ambulatory health care to be delivered in a cost-effective manner. However, current mobile e-health systems are hampered by several challenges that we will address in this paper.

The first is the power efficiency problem. Due to the limited battery capacity of mobile devices, it is likely that e-health applications can only be run for a couple of hours in the mobile devices even if their batteries are fully charged, hence constraining the ubiquitous usefulness of the mobile applications for e-health. Therefore, a solution is needed to reduce the power consumption of mobile devices when they are running e-health applications.

Also, diverse types of data from various sensors are used in e-health [2, 3]. Such data are collected from various wireless sensors as distinct data streams rather than being mixed indiscriminately. Healthcare applications need to exploit the synergy and related conditions between different sensor data streams to provide a better understanding of the subject’s health condition; e.g., a patient may appear normal if only his heart rate or blood pressure is checked separately, while he is actually in a state of disorder if these two parameters are considered together. Considering the constraints of the computational capability and battery capacity of these sensors, extensive and synergistic processing and analyzing of sensing data at the sensors to directly extract real-time and appropriate information relating to healthcare is not feasible. In addition, even though sensing data can be transferred to remote servers through mobile devices for processing, the potentially long delay and high cost of bandwidth consumption may not be acceptable. Thus, it is of interest to investigate an efficient and effective data mining mechanism to process and analyze the sensing data for e-health in the mobile devices, so as to provide critical healthcare-related information to users in a timely manner through their mobile devices to enable early detection of medical disorders.

Furthermore, patients may not know how to interpret their e-health related information without help from their care-givers. However, substantial costs in human and communication resources may be incurred by direct bilateral connections between health care providers and patients. Therefore, the collaborative sharing of health information and group-level communications need to be investigated as a more cost-effective approach to enable mobile e-health.

In this paper, we address the challenges mentioned above. Our main contributions are threefold. First, we adopt Bluetooth v4.0 [4] for communications between mobile devices and wireless sensors, and show that the low-energy feature of a Bluetooth v4.0 can provide a more power-efficient solution for mobile e-health applications than existing solutions that use Bluetooth 2.1or ZigBee. Second, to facilitate the processing of diverse e-health related data in mobile devices, we present an optimized Naïve Bayes algorithm to efficiently classify the different types of data. Finally, we design a system that allows users to collaborate over social networks and perform group-level activities for mobile e-health with a low investment in manpower and communication resources. To the best of our knowledge, our system is the first work to integrate Bluetooth v4.0 wireless technology, mobile data mining techniques, and a social networks application to improve the performance of mobile e-health.

In the rest of this paper, we first review the background and
related techniques of Bluetooth Low-Energy, Bayes Classification, and social networks in Section II. Then, we present detailed descriptions of the design of our mobile e-health system in Section III. An experimental evaluation of our system design is presented in Section IV. Related work is reviewed in Section V. Section VI concludes the paper.

II. BACKGROUND AND RELATED TECHNIQUES

A. Bluetooth Low Energy

Some of the key features of Bluetooth v4.0 include an ultra-low peak, average and idle-mode power consumption, and an enhanced communications range. Bluetooth Low Energy (BLE) is a subset of Bluetooth v4.0 with an entirely new protocol stack for the rapid establishment of simple data links for wireless sensor applications. Since wireless sensor devices employing BLE are capable of operating for months or even years on tiny, coin-cell batteries, they are ideally suited for e-health applications [5]. The architecture of the BLE protocol stack is shown in Fig. 1.

![BLE Protocol Stack](image)

In Fig. 1, the Generic Access Profile (GAP) layer is responsible for handling the access modes and procedures of the device, including device discovery, link establishment, link termination, initiation of security features, and device configuration. The Generic Attribute Profile (GATT) layer is used by the application to establish data communications between two connected devices. A software development kit used by the application to establish data communications is available to support the development of BLE applications. Between two connected devices, a software development kit is used by the application to establish data communications.

![General Packet Format of BLE](image)

III. THE MOBILE E-HEALTH SYSTEM

In this section, we present our mobile e-health system design incorporating mobile data mining and social network applications, and introduce the design and implementation of the system components. Fig. 3 shows the architecture of the system design.

![Architecture of the Proposed Mobile e-Health System](image)

In our system, wireless sensor devices that are able to sense the context and sample physiological parameters are attached to the patient’s body, and the sensing data are conveyed to a mobile device (e.g., an Android smart phone) via BLE connections. In the system, we integrate the health sensors (i.e., wearable and peripheral sensors) with the BLE SoC. The signals sampled by the health sensors are sent to the mobile device (e.g., an Android smart phone) via BLE. The signals are then sent to the mobile device via BLE, and the mobile device processes the data.

B. Bayes Classification for Mobile Data Mining

A Naïve Bayes classifier is a simple probabilistic classifier based on applying Bayes’ theorem with strong independence assumptions [7]. It features computational simplicity, and is understandable, fast, and requires only a single pass through the data if the attributes are discrete. Naïve Bayes also has the limitation that the assumption of independent attributes may be unrealistic, which may lead to incorrect estimations of probability if there are some strong dependencies among some attributes. Considering the computational capacity of mobile devices, the Naïve Bayes classifier could potentially provide an efficient and affordable approach for the processing of multiple and diverse health sensing data on mobile devices.

C. Social Networks

Many social networking services are offered over online platforms to facilitate social relationships and networking between people who have similar interests or needs. In some social networks, relationships between individuals are privileged [8]. Research on social networks has penetrated into health care analytics, in both epidemiological studies and models of patient collaboration. By leveraging the advantages of social networks to enhance health care services, mutual support and information exchanges among patients and between patients and care-givers can be efficiently provided so that patients may engage in self-management and address psychosocial concerns. Therefore, social networking could have a significant impact on the efficiency and cost-effectiveness of e-health systems [9].
automatically processes the sensing data stored in the local device using a data mining technique, and the user can directly obtain the health information from the Android device. The processed result can be further uploaded to the Internet and shared with other social network users. Moreover, all of the data can be uploaded to a remote healthcare center that can support the storage and management of all of the healthcare data both from the Android devices of users and the social networks.

A. Sensor Deployment and BLE Communications

All embedded software for the BLE SoC is developed using Embedded Workbench for 8051 version 8.10.4 from IAR Software.

In our system, blood pressure sensors, thermometer sensors, and heart rate sensors are applied as wearable devices. For the benefit of some elderly people, we also apply moisture sensors, which can be fixed on the thigh to detect urinary incontinence. In order to collect data with six degrees of freedom motion information, we fix a three-axes, digital accelerometer (Analog Devices ADXL345) and a three-axes, digital gyroscope (Invernsense ITF-3200) on each sensor expansion board. Thus, a data analysis subsystem contains information acquired by both types of sensors so as to implement an inertial position and limb motion pattern estimation scheme to discover abnormal symptoms [1, 2, 3]. To establish a connection between BLE SoC and an Android device with BLE, we first need to initialize the BLE module. The initialization of the module occurs in two phases. First, parameters in the peripheral profile, the GAP, and the GAP bond manager are configured by calling the related functions. Second, we set the BLE SoC to an advertising state, which enables it to be found by other devices. If the Android device with BLE sends a connection request to BLE SoC, it will accept the request and go into the connected state as a slave. If no connection request is received, it will only remain discoverable for 30.72 seconds, before going into the standby state. Once connected to the Android device, the service discovery procedure is performed for each service of interest.

Once the connection is established and measurements have been enabled, the BLE SoC will be able to send data packets of physical parameters when configured to do so by the Android device through BLE communications. We can cycle through different optional measurement formats by through the BLE SoC.

To make our system smarter so that it is aware of the environment, we incorporate various types of sensors into our network and deploy them in the environment in which our human subjects go about their everyday activities. For instance, we use Berkeley MicaZ motes to emulate the temperature sensors and smoke sensors in the home environment to provide comprehensive information regarding the parameters of the environment [10, 11]. In addition, by deploying RFID readers at key locations such as doors and corridors, and subjects carrying RFID tags can be identified together with their locations [12]. By taking people’s locations and environmental conditions into consideration, this system can enable better e-health, particularly for the elderly and disabled.

B. Mobile Data Mining Algorithm

Since our application is run on a mobile device with limited computational power, its algorithm should not be too complex in terms of both time and space complexity. Therefore we consider the Naïve Bayes classifier, which is famous for its computational simplicity and accuracy [13]. To increase the accuracy of our classifier, we make some modifications to it. The following figure gives a general description of our mobile data mining process:

In our system, the format of health data monitored by wireless sensors is stored as numerical features. In order to simplify the work of classification that occurs afterwards, we apply Fayyad and Iran’s MDL method [14] to discretize the data by converting them from numerical to nominal, in order to decrease the computational complexity of the classification work that takes place afterwards.

One noticeable limitation of Naïve Bayes is that it assumes that every attribute is independent of others, which is rarely true in the real world. To weaken this assumption of attribute independence, we consider the Lazy Learning of Bayes Rules (LBR) [15]. For each \( x = (x_1, \ldots, x_n) \) to be classified, a set \( W \) of the attribute values is selected using a heuristic wrapper approach. Independence is assumed of the remaining attributes. Given the information that the number of total classes is \( y \), \( x \) can be classified by selecting:

\[
\arg\max_y \left( \hat{p}(y | W) \prod_{i=1}^{n} \hat{p}(x_i | y, W) \right)
\]

Thus, the entire work of classification depends both on classes and attributes in \( W \). However, according to LBR, the selection of \( W \) is an operation of time complexity \( O(tkn^3) \), in which \( t \) is the number of training examples, \( k \) is the number of classes, and \( n \) is the number of elements to be classified. In practice, when large numbers of samples are to be classified, the computational burden becomes prohibitive [15]. Thus, in the end we make a trade-off between accuracy and complexity, by selecting the subset of attributes randomly instead of applying the heuristic wrapper approach. By doing so, we can weaken the assumption of attribute independence while keeping the computational complexity within an acceptable range for mobile devices.

C. Mobile Social Network Module

To perform group-level activities and provide an economical solution for mobile e-health, we apply social networks to our system. Here we simulate the situations of group-level activities without and with social networks:

Discretize
Select attributes randomly
Optimize Naïve Bayes

Fig. 4. Mobile Data Mining Process

Using Fayyad and Iran’s MDL method
To weaken attribute independence assumption of Naïve Bayes

argmax

\[
\arg\max_y \left( \hat{p}(y | W) \prod_{i=1}^{n} \hat{p}(x_i | y, W) \right)
\]

Thus, the entire work of classification depends both on classes and attributes in \( W \). However, according to LBR, the selection of \( W \) is an operation of time complexity \( O(tkn^3) \), in which \( t \) is the number of training examples, \( k \) is the number of classes, and \( n \) is the number of elements to be classified. In practice, when large numbers of samples are to be classified, the computational burden becomes prohibitive [15]. Thus, in the end we make a trade-off between accuracy and complexity, by selecting the subset of attributes randomly instead of applying the heuristic wrapper approach. By doing so, we can weaken the assumption of attribute independence while keeping the computational complexity within an acceptable range for mobile devices.

C. Mobile Social Network Module

To perform group-level activities and provide an economical solution for mobile e-health, we apply social networks to our system. Here we simulate the situations of group-level activities without and with social networks:
As seen in Fig. 5, we find that applying social networks for mobile e-health can lead to a reduction in cost in terms of human and financial resources, as collaborative healthcare can be supported and the cost for individuals can be reduced. Moreover, by sharing information, a wider range of information such as environmental elements may be taken into consideration. We make use of the existing popular social network, Twitter [16], to connect mobile e-health users. We mainly use APIs from the Twitter4j library to achieve this function. Twitter4J is an unofficial Java library for the Twitter API. It can run on the Android platform and is Google App Engine ready.

Every time a user tries to upload his/her health information to Twitter, a web task of OAuth Authentication [17] is initialized, in which Users grant access to their Protected Resources without sharing their credentials with the Consumer. OAuth Authentication is done in three steps. First, the Consumer obtains an unauthorized Request Token. Second, the User authorizes the Request Token. Finally, the Consumer exchanges the Request Token for an Access Token. The Request Token is used by the Consumer to ask the User to authorize access to the Protected Resources. The Access Token is used by the Consumer to access the Protected Resources on behalf of the User. After successfully receiving the Access Token and Token Secret, the Consumer is able to access the Protected Resources on behalf of the User.

By providing the username and password of the user’s Twitter account as inputs, our application will be authorized to gain access to Twitter. After the information is confirmed, the user will be able to connect to Twitter, update the profile, and communicate with other users about their health information. Thus, a many-to-many relationship is formed between users without having to input a lot of manpower and material resources.

IV. EVALUATIONS AND APPLICATION INSTANCES

As our system design is mainly concerned with BLE, mobile data mining, and social networks, we evaluate and demonstrate our system from the above three aspects.

A. BLE Power Efficiency Experiment

First, we conduct experiments to prove the power efficiency of BLE devices. We use LG Optimus 4X HD in this experiment as it is equipped with the BLE module. Thus, we use it to test the power efficiency of BLE, which is adopted in our system. Since the format of the data in the payload of the sensor packets is undocumented and we have not been able to determine any obvious format to the data, we simply use the individual bytes from these packets directly as numerical features. Due to experimental and technical constraints, such as a lack of sufficient multiple health sensors on hand, we use binary data packets, which contain health sensing data on heart rate and blood pressure to replace the individual bytes in the sensor packets [18, 19]. In a series of experiments, we transfer streams of data packets with total sizes of 30MB, 50MB, and 100MB, from sm Asus TF101 to LG Optimus 4X HD and from LG VS910 to LG Optimus 4X HD. The time cost and power consumption are recorded directly from the “Settings -> Battery” statistics provided by LG Optimus 4X HD, as illustrated in Fig. 6(a) and Fig. 6(b) for the battery level before and after the file transfer, respectively.

Then, we calculate the percentage of file transfer power consumption via BLE on LG Optimus 4X HD, which can be calculated by subtracting the power level after the BLE file transfer from the power level before the file transfer. In Fig. 6(a), the battery power consumed before the file transfer is 24%, of which the Bluetooth file transfer has taken up 5%. In Fig. 6(b), the battery power consumed after the file transfer is 25%, with the Bluetooth file transfer having taken up 4%. Therefore, the percentage of BLE power consumption is 25%×5% - 24%×4% = 0.29%.

<table>
<thead>
<tr>
<th>Size of Data (MB)</th>
<th>Power Consumption of BLE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>0.17%</td>
</tr>
<tr>
<td>100</td>
<td>0.81%</td>
</tr>
<tr>
<td>150</td>
<td>1.01%</td>
</tr>
</tbody>
</table>

TABLE I

**Power consumption during file transfer from ASUS TF101 to LG Optimus 4X HD**

<table>
<thead>
<tr>
<th>Size of Data (MB)</th>
<th>Time Used (min)</th>
<th>Power Consumption of BLE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>7</td>
<td>0.2%</td>
</tr>
<tr>
<td>100</td>
<td>20</td>
<td>0.73%</td>
</tr>
<tr>
<td>150</td>
<td>33</td>
<td>1.08%</td>
</tr>
</tbody>
</table>

TABLE II

**Power consumption during file transfer from LG VS910 to LG Optimus 4X HD**
The above tables indicate the power consumption with respect to different sizes of data stream. Take, for instance, the sending of 100MB of data. We can estimate the battery lifetime under a prolonged BLE file transfer: \(27.5 \text{ min} / 0.77\% \approx 3571.4 \text{ min} \approx 59.5 \text{ h}\). Although this estimate is obtained under the unrealistic assumption that all of the power in LG Optimus 4X HD is provided to BLE, it does indicate the power efficiency of BLE on LG Optimus 4X HD.

### B. Mobile Data Mining Evaluation

#### TABLE IV

**Confusion Matrix of Naïve Bayes**

<table>
<thead>
<tr>
<th>Diabetes</th>
<th>Ionosphere</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 407</td>
<td>93 b</td>
</tr>
<tr>
<td>1 96</td>
<td>172 g</td>
</tr>
</tbody>
</table>

#### TABLE V

**Confusion Matrix of Our Algorithm**

<table>
<thead>
<tr>
<th>Diabetes</th>
<th>Ionosphere</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 423</td>
<td>77 b G</td>
</tr>
<tr>
<td>1 105</td>
<td>163 g</td>
</tr>
</tbody>
</table>

#### TABLE VI

**Evaluation Result of Naïve Bayes**

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart Rate</td>
<td>0.814</td>
<td>0.642</td>
<td>0.754</td>
</tr>
<tr>
<td>Blood Pressure</td>
<td>0.865</td>
<td>0.804</td>
<td>0.826</td>
</tr>
</tbody>
</table>

#### TABLE VII

**Evaluation Result of Our Classifier**

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart Rate</td>
<td>0.846</td>
<td>0.608</td>
<td>0.763</td>
</tr>
<tr>
<td>Blood Pressure</td>
<td>0.865</td>
<td>0.822</td>
<td>0.838</td>
</tr>
</tbody>
</table>

Second, we evaluate our algorithm and compare the result with the original Naïve Bayes Classifier. We mainly carry out this experiment on those data packets that have been used in BLE power consumption experiments. Confusion matrices for the test results of the Naïve Bayes Classifier and our classifier are listed in Table IV and Table V. Based on the confusion matrices, we can further calculate the sensitivity, specificity, and accuracy of both classifiers, which are shown in Table VI and Table VII.

From the tables listed above, we can see that our algorithm has an increased performance compared to Naïve Bayes in terms of its sensitivity, specificity, and accuracy.

### C. Social Network Application Instances

The following figures show the Graphical Use Interface (GUI) of our application developed for Android devices with the BLE module. Fig. 7(a) shows the main layout of the application. To perform group-level activities with social networks, the user presses the “Update my Profile” button and logs into her Twitter account as shown in Fig. 7(b). After the authorization process, the user is free to gain secure access to Twitter and to share her information. Fig. 7(c) and 7(d) show the application of mobile data mining. To begin a classification task, the user should select between “Cross Validation” and “Training”, as shown in Fig. 7(c). Fig. 7(d) represents some operations that will take place afterwards if “Cross Validation” is chosen. Here we use the same data packets as in the BLE power consumption experiments. The Twitter GUI for sharing health information is shown in Fig. 7(e) and Fig. 7(f).
V. RELATED WORK

Several wireless sensor systems [7, 14, 20, 21] have been built to study human activities such as walking, lying down, or climbing stairs; e.g., in [7], a platform named iLearn is presented, which can classify human activities using the accelerometers in Apple iPhones. This shows that human activities can be sensed using a commodity device instead of custom hardware. The eWatch system [14] is a custom-made system with a wristwatch form factor, which aims to detect similar activities such as walking, lying down, or climbing stairs. It also provides an online nearest location classification to identify and recognize a set of frequently visited locations. Yet most of these system designs have not considered the problem of power consumption, nor provided the support of group-level activities.

There are also other e-health related systems that use laptops to perform computation work. The system in [22] is designed to detect the bed postures of the elderly and bedridden. The system applies Bayesian classification and Gaussian distribution to process sampled data. Kurtosis and skewness are also estimated to represent the pressure contour. The system in [23] uses five wireless accelerometers, a heart rate monitor, and a laptop to track similar activities in real time. Most of these systems employ wired sensors, or need a laptop to collect and process sensing data, which make them awkward for supporting mobile e-health. Compared with these systems, our system performs its main functions using wearable wireless sensors and portable mobile devices. Thus, our system can be deployed for ubiquitous applications.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we have proposed a novel design of a collaborative mobile e-health system that incorporates mobile data mining and social networking. The design utilizes cutting-edge BLE modules to enable superior power efficiency and a long system run time. To optimize the processing of diverse sensing data, we have adopted a mobile data mining technique employing a Naïve Bayes algorithm, and increased its accuracy by weakening the assumption of independent attributes. Furthermore, we have provided a cost-effective platform based on social networking for users to share healthcare information, which enables collaborative e-health between patients and care-givers. Our design provides an efficient, affordable, and seamless solution for future mobile e-health. As our work on this is currently at the initial stage, an overall system implementation and an evaluation of the complete system is still lacking. Thus, we plan to extend our current work in several aspects in the future. For instance, we are working on the extension of our classifier to fit more human activities and to find an optimization method for continuous recognition. We also intend to generalize the system design to incorporate a larger variety of BLE devices and support a wider range of applications.

REFERENCES


[18] Systolic blood pressure dataset http://www-users.york.ac.uk/~mb55/datasets/bp.dct


