Low-Cost Mobile Personal Clouds

Milena Radenkovic  
The University of Nottingham  
milena.radenkovic@nottingham.ac.uk

Vu San Ha Huynh  
The University of Nottingham  
eeyvsh@nottingham.ac.uk

Abstract: We propose a mobile peer to peer personal cloud architecture which allows users to capture, store, analyse, interact with and share different types of personal and context data with no privacy leakage. Our mobile personal cloud can host multiple different services which are intelligent, distributed, dynamic and operate in real time. In this paper we describe one service that we designed and deployed on our mobile personal cloud called Mobile Wellbeing Companion Cloud (MWCC). Using low-cost, off-the-shelf hardware components and open-source software, our MWCC combines several sensor network technologies to allow users to monitor and interact with their personal data and environment in real time without privacy leakage. MWCC augments heterogeneous sensors data with state of the art machine learning algorithms for signal filtering, fast classification and analysis and provides interactive data visualisation for transparent user interaction. We show that our MWCC is easy to use and highly accurate while managing to keep resource costs low.

Keywords — Opportunistic Disruption Tolerant Networks, Mobile computing and applications.

I. INTRODUCTION

With continuous increase of different cloud services available at seemingly no cost to the users, large volumes of user data, some of which personal, is being stored, processed, controlled and owned by third party companies. Even though the majority of users are not aware of the technical details of what happens to their data once they use cloud services, they are increasingly aware that many service providers analyze their personal data to build profiles which they "trade" with other agencies such as advertising or credit companies [1]. In this paper, we address this problem by proposing a mobile low cost, intelligent and adaptive human-centered, personal cloud architecture which allows users to host their services, collect, store and analyse as well as interact with their data and services. We then propose a novel Mobile Wellbeing Personal cloud service hosted in our mobile personal cloud which combines 1) heterogeneous sensing technologies that allow users to understand their context and interact with their data and services in privacy preserving manner via a real time dashboard, 2) smart machine learning algorithms to analyse and predict context and 3) reliable intelligent underlying communication network protocols. Even though similar products (e.g. Fitbit Surge, Basis Peak) exist, they are not always affordable (typically costing over £100 and offering only 2 or 3 features), our MWCC reduces the cost by over half and provides more features. The paper begins by providing overview of the related work in section II, section III introduces low cost mobile personal cloud architecture, section IV describes Mobile Wellbeing Companion Cloud (MWCC) testbed design, deployment and testing; section V gives conclusions. Our testbed shows that we manage to achieve high accuracy of our real time machine learning algorithms, keep low resource costs and allow rich real time user interaction.

II. RELATED WORK

In this section, we describe related work that has inspired ours. Databox [1] proposes a trusted platform which is designed to allow users to manage and use their own data, grant controlled access to others and offer different support incentive policies for external users. RaspPiCloud [2] builds on this to propose a novel low cost personal cloud approach hosted ion devices such as Raspberry Pi which allows users to collect, store, query and share their own data. [2] argues that there is a need for deploying personal clouds in both remote and urban areas where the network connectivity can be intermittent but where the users may still require and have the right to their privacy. Other related research efforts in human-centered systems typically address challenges on how to improve their performance and usability. For example, Octopus[3], BodyCloud[4], CocaMAAL[5], BDCaM[6] and [7,8,9,10,11,12,14,15] describe various features recognition and detection techniques that may allow users to gather and process data from sensing devices with high accuracy. However, they do not address the cost and limited resources challenges (such CPU, storage, battery, network cost) in order to achieve that high performance. Research on context aware, intelligent agents such as[16,17] allows users to view and share their sensitive personal data using assumed-trusted service providers but disregard questions about how and where users’ data is stored, processed and any other cost related issues. Emerging work on anticipatory mobile computing [13] argues that context prediction and anticipatory decision-making will be integral part of future mobile systems but identifies that processing costs can be high (suitable usually for the top of the range mobile devices) and that the challenge of real-world implementations of such systems which are almost non-existent (as most of the work is done via simulations).

Delay tolerant and opportunistic networking approaches address the lack of continuous network connectivity [18, 19,20,21,22,32,33,34] and were shown to have high efficiency and reliability as core data transfer protocols both mobile social and mobile ad hoc vehicular networks. Research in [32,33,34] has shown that both communication performance results and achieved levels of user privacy and anonymity of
newly proposed adaptive opportunistic DTN protocols are high even in the presence of high numbers of malicious nodes. [35,36,38] combine mobility of sensor sources, data carrier and clouds to enhance routing/forwarding protocol in terms of energy efficiency, security, connectivity and throughput. We adopt such communication protocols for our inter-mobile personal cloud communication.

III. LOW COST MOBILE PERSONAL CLOUD ARCHITECTURE

We envisage a low cost mobile personal cloud architecture that enables rich interactivity with the user, is context aware and uses intelligent network communication for increased reliability as described in [37]. At the core of the mobile personal cloud is the idea of open source evolving platform where the user owns and hosts all the services and data they use and can add any new ones when they need. Specifically, the user stores their data and services locally and may analyze and share them in a fully controlled privacy conscious manner according to their own requirements. Examples of the services that the mobile personal cloud can host include: personal social networks, personal well-being, or personal financial services. Each of these stores and manages different types of data requiring different levels of privacy. As such each is hosted in a separate virtual container and has its real time user interface what allows rich interactions with the environment, data and services to the user. An early demonstration of the feasibility of this work deployed on off the shelf Raspberry PIs was shown in [2]. This paper extends our previous work to focus on integrating suitable machine learning algorithms into the cloud, allowing the user to have rich real time interaction, describing how the cloud can sense and interact with the environment and how the data can be shared across multi-hop communication with the remote destinations. In order to improve reliability of data communication, we use P2P DTN for to support communication between multiple personal clouds as proposed in [32,33,34]. This enables multiple personal clouds to cooperate and communicate together without the need of infrastructure and thus preserving higher levels of privacy. We consider privacy as our core motivation for the architecture and aim to support different kinds of data, in different situations, will have different privacy level requirements. For example, users would normally not publish their health information to public. However, in emergency cases, for example, a heart attack occurrence, the MWCC may be required to generate a notification/alert to its neighborhood in addition to the hospital. In addition to this, a smart routing protocol should be proposed to seek the best route for sending these alerts directly to third-party organizations like hospital, police, fire-station.

IV. MOBILE WELLBEING COMPANION CLOUD TEST-BED

We describe the design and deployment of Wellbeing Companion Personal Cloud which is one service our Personal Cloud offers to the users. Mobile Wellbeing Companion Personal Cloud (MWCC) integrates various heterogeneous wireless and wired sensor networks to capture and analyse real time information about the user (i.e. motion, heart rate, etc) and their environment (such as temperature, humidity or pressure) in order to predict. Our test-bed comprises a range of components such as body-worn sensors (heart rate monitoring), smart-phone that provides user interface and inertial sensors, separated (MEMS sensor board, or RF temperature sensors), Raspberry Pi device as a head node which gathers, stores and processes incoming signals in real time.

The head node consists of 4 main hierarchical layers as described in Figure 1. User interacts with sensor's data via MWCC on the Application Layer which allows them to access, monitor (either query requests or subscribe periodically) and share their personal data. In order to increase reliability and scalability, the data may be synchronized with more powerful devices such as personal laptop or home computer. One of the challenges we aim to tackle is how to integrate smart knowledge-based decision making algorithms with different contextual sensed data into meaningful information when the resources are limited and sensed data is noisy. Consider a scenario where Wellbeing Companion Personal Cloud aims to provide the following user centric context driven interrelated services: real time human activity recognition correlated with heart rate monitoring and external sensors such as room temperature, allow the user to interact with all of these (e.g. viewing, querying, sharing) and receive notifications. Our main challenge is how to achieve optimal tradeoff between accuracy, resource utilization and user experience. Namely, more sophisticated algorithms may result in better accuracy of activity monitoring but lower resources and cause delays to the user. Poor algorithms may result in wrong detections, thus causing inaccurate feedback to the user and false notifications. We first investigate most commonly used human activity recognition algorithms in terms of accuracy achieved and time taken as measured in a real test-bed scenario. We then explore the accuracy and CPU cost of real time user heart rate visualization and analysis (in the terms of normal or anomalous behavior) for different user activities.

A. Human Activity Recognition

We build on an openly available dataset [14] which includes raw accelerometer and gyroscope data collected from 30 subjects who performed 6 basic activities (walking, walking upstairs, walking downstairs, standing, sitting and lying). We identify transitioning activities that occurred (stand-to-sit, lie-to-stand, etc.). We separate 30-subject data randomly as a ratio
70:30 into a training set and a new-user test set so that we can perform cross-validation after building classification models. We include a real world user who uses the Wellbeing Companion Cloud in the new user set. We apply low CPU filter to 3-D accelerometer and gyroscope sensor signals to remove noise. More specifically, we separate filtered accelerometer $a(t)$ into body linear accelerometer $la(t)$ and gravity force $g(t)$ using low-pass filter [25,26] with cut-off constant $\alpha = 0.3$ (in line with [14]) as below.

$$
g(t) = \alpha * a(t) + (1 - \alpha) * g(t - 1)$$

$$
a(t) = la(t) + g(t)
$$

As signal frequency = 50Hz, we argue that if $\alpha$ is too large (e.g. 0.5), almost everything will get passed and nothing will be filtered out. If $\alpha$ is too small (e.g. 0.1), signal details will be lost. Thus, we choose $\alpha = 0.3$ to balance the trade-off. To reduce chance of miss classification between walking, walking upstairs and downstairs that occur frequently [14,23,24], we propose to merge these 3 states of motion into one, walking which improves accuracy while minimizing overheads induced by miss classification data. A vector of 12 features (in which $x, y$-axis accelerometer, $x$-axis linear accelerometer and $y$-axis gravity are proved that significantly affect performance) is generated in time domain and labeled in 5 classes.

We then carry out experiments in our test-bed using benchmarking classification algorithms such as traditional Naive Bayes classifier, Bayes Network, Random Tree, Random Forest, Support Vector Machine (SVM) and Lazy Learning (IBk). The comparison between these algorithm based on activity recognition accuracy and model building time is shown in Figure 2. We show that the shorter a model building an algorithm takes, the lower CPU cost it has and the faster its processing completes.

We observe that Random Tree and IBK have highest accuracy (above 97%) and require lowest time to build their models (less than 10s). Random Forest has high accuracy but its model building process consumes significantly longer time (roughly 74.8s). The model building time $T$: $T(\text{IBk}) < T(\text{Random Tree}) < T(\text{Random Forest})$ can be explained by the complexity of IBk, Random Tree and Random Forest algorithm which are $O(mn)$, $O(nm \log(n))$, $O(n \log(n))$ respectively for $n$ attributes and $n$ instances. However, the model building time is not significantly important, thus long model building time of Random Forest is acceptable, due to the fact that the model would be built only once in the test-bed and users may only suffer this for the first time the Human Activity Recognition module being installed.

To observer our training models’ performance with unknown users, we perform the new user cross-validation on our top-3 accurate models above and compare the result based on accuracy, load average and time required for each classification.

<table>
<thead>
<tr>
<th>Activity</th>
<th>W</th>
<th>S</th>
<th>L</th>
<th>T</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking (W)</td>
<td>49100</td>
<td>54</td>
<td>0</td>
<td>107</td>
<td>99.7%</td>
</tr>
<tr>
<td>Sitting (SI)</td>
<td>28</td>
<td>364</td>
<td>0</td>
<td>199</td>
<td>97.3%</td>
</tr>
<tr>
<td>Standing (S)</td>
<td>353</td>
<td>19013</td>
<td>0</td>
<td>108</td>
<td>96.9%</td>
</tr>
<tr>
<td>Lying (L)</td>
<td>0</td>
<td>0</td>
<td>19912</td>
<td>210</td>
<td>99.0%</td>
</tr>
<tr>
<td>Transition (T)</td>
<td>1097</td>
<td>204</td>
<td>273</td>
<td>366</td>
<td>72.72</td>
</tr>
<tr>
<td>Precision</td>
<td>97.1%</td>
<td>99%</td>
<td>97%</td>
<td>98.2%</td>
<td>92.1%</td>
</tr>
</tbody>
</table>

Table 1. Confusion matrix using Random Forest on test set. Rows are actual classes and Columns are predicted classes.

The classification recall for Walking, Sitting, Standing And Lying is high (97% - 99%) with minor miss-classification
between walking and standing. This shows that our Human Activity Recognition module would achieve extremely high performance if users were calibrated.

If users are completely unknown or not calibrated, our Human Activity Recognition module still gets reasonably high performance (87.1% accuracy) as shown in Table 2.

<table>
<thead>
<tr>
<th>Activity</th>
<th>W</th>
<th>SI</th>
<th>S</th>
<th>L</th>
<th>T</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking (W)</td>
<td>97232</td>
<td>46</td>
<td>195</td>
<td>0</td>
<td>3287</td>
<td>96.5%</td>
</tr>
<tr>
<td>Sitting (SI)</td>
<td>252</td>
<td>23975</td>
<td>10288</td>
<td>0</td>
<td>1701</td>
<td>66.2%</td>
</tr>
<tr>
<td>Standing (S)</td>
<td>1134</td>
<td>6820</td>
<td>30450</td>
<td>0</td>
<td>393</td>
<td>78.5%</td>
</tr>
<tr>
<td>Lying (L)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>36297</td>
<td>501</td>
<td>98.6%</td>
</tr>
<tr>
<td>Transition (T)</td>
<td>3540</td>
<td>510</td>
<td>817</td>
<td>1082</td>
<td>14224</td>
<td>70.5%</td>
</tr>
<tr>
<td>Precision</td>
<td>95.2%</td>
<td>76.5%</td>
<td>72.9%</td>
<td>97.1%</td>
<td>70.7%</td>
<td>87.1%</td>
</tr>
</tbody>
</table>

Table 2. Confusion Matrix using Random Forest on new user cross-validation.

The clusters of different activity classification are visualized in Figure 5. We show that the clusters of 2 static motions, sitting and standing, are partly overlapped. This is because of the similarity of postures in Sitting and Standing activity which can make it challenging to always accurately categorize them. On the other hand, the other static motion, lying, which has different posture, gives high activity recognition (98.6%). The cluster of lying only overlaps trivially with transitioning. We observe that there is also a minor mis-classification of 2 dynamic activities, walking and transitioning which, however, does not affect the recognition accuracy of Walking (96.5%).

In summary, by merging walking, walking upstairs, walking downstairs into walking as their nature, adding transitioning as one of user's state and applying Random Forest Model, our human activity recognition module of Wellbeing Companion Cloud balances the trade-off between accuracy and CPU cost. This allows Wellbeing Companion Cloud to operate in real-time while not requiring a lot of pre-processing data e.g. 1 raw reading generates 12 features to achieve 1 activity recognition.

B. Real time Fusing Heart Rate Monitoring and Human Activity Recognition in the MWCC

We propose that there is a need for our Wellbeing Companion Cloud to provide automatic calibration of users resting heart rate (RHR) and max heart rate (MHR). This is important in order to calculate their threshold heart rate for each activity users take part in, as argued in [29]. For example, for people who work-out or play sports, Wellbeing Companion Cloud will help to keep track of their heart rate so that they could adjust their training intensity. In emergency situations, such as heart attack or unusually high/low high heart rates for current activity, Wellbeing Companion Cloud will issue notifications or alerts.

Some existing applications allow users to manually set their RHR, MHR and Age in order to calculate HRZ However, our Wellbeing Companion Cloud assumes that the majority of users have no knowledge about their heart rate parameters and supports automatic calibration of users heart rate based on merging activity recognition with prediction of heart rate parameters in the following way.

RHR is defined as how fast a heart beats when user is at complete state of rest [29,30]. Our Wellbeing Companion Cloud calibrates RHR by getting average user's current heart rate whenever it recognizes user's current motion as lying. Previous research proposed that the calibration should be performed over 3 days to get the highest accurate measurements [30] and this is the accuracy that we target in our Wellbeing Companion Cloud.

Pseudocode to calibrate Resting Heart Rate (RHR):

```java
function calibrateRestingHeartRate ():
for t = 0 to N:
    activity = getHumanActivityRecognition();
    if (activity == Lying AND isUserActivityStable()):
        heartRate = getHeartRate();
        sum += heartRate;
    restingHeartRate = sum/count
end if;
```

The MHR is another core heart rate parameter defined as the fastest users' heart beat rate per minute. As it would be unsafe to empirically measure, we have developed algorithm for formulas proposed in [27,28] to estimate MHR based on age and weight which are easy for users to set up . In this paper, we apply common age predicted formula, to estimate MHR and allow the users to review the results manually with doctor's advice if necessary.

Formula:

\[ MHR = c - \text{Age} \quad c = \begin{cases} 
220 & \text{if men} \\
226 & \text{if women} 
\end{cases} \]

Pseudocode to calibrate Max Heart Rate (MHR):

```java
function calibrateMaxHeartRate ():
age = getAgefromUI();
gender = getGenderFromUI();
if (gender == Male) maxHeartRate = 220 - getHeartRate();
else maxHeartRate = 226 - getHeartRate();
```

Another core heart rate parameter we consider is heart rate zone (HRZ) which is defined as the heart rate limit which indicate the estimating lower bound and upper bound of users' heart rate depending on their current activity. Our Wellbeing Companion Cloud is able to separate HRZ into 5 Zones by implementing Karvonen formula proposed in [29]

\[ HRZ = \alpha \cdot MHR + (1 - \alpha) \cdot RHR \]

As shown in Table 3, \( \alpha \) varies from 50% to 100% for Zones 1 to Zone 5. While Zone 1 refers to human activities like sitting, standing, walking slowly (lower limit) or walking quickly (upper limit), Zone 5 refers high impact user activities such as intensive training.

![Figure 5. Activity Clustering Visualization](image-url)
<table>
<thead>
<tr>
<th>Zone</th>
<th>Lower Bound $\alpha$</th>
<th>Upper Bound $\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50%</td>
<td>60%</td>
</tr>
<tr>
<td>2</td>
<td>60%</td>
<td>70%</td>
</tr>
<tr>
<td>3</td>
<td>70%</td>
<td>80%</td>
</tr>
<tr>
<td>4</td>
<td>80%</td>
<td>90%</td>
</tr>
<tr>
<td>5</td>
<td>90%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 3. Constant $\alpha$ values on different activity zones.

Based on the heart rate zones, Wellbeing Companion Cloud issues emergency notifications and alerts both locally (to the user) and remotely (to the hospital or friend). The warning level depends on a gap identified between the currently measured heart rate, estimated normal heart rate for the current user activity and the estimated user’s HRZ (shown in Figure 8).

Pseudocode to detect abnormal users’ health behaviours:

```python
function analyzeUserHealth()
    while (true)
        activity = getHumanActivityRecognition()
        heartRate = getHeartRate()
        switch (activity):
            case Lying:     lowerAlpha = 0.5; upperAlpha = 0.55
            case Sitting, Standing:  lowerAlpha = 0.55; upperAlpha = 0.6
            case Walking:      lowerAlpha = 0.6; upperAlpha = 0.7
            default:          lowerAlpha = 0.5; upperAlpha = 0.7
        lowerHeartRateBound = lowerAlpha*MHR + (1- lowerAlpha)*RHR
        upperHeartRateBound = upperAlpha*MHR + (1- upperAlpha)*RHR
        if (heartRate < lowerHeartRateBound AND isSustained()):
            notifyMessage("User's current heart rate is low");
            measureCriticalLevelBasedOnGap(heartRate, lowerHeartRateBound);
        if (heartRate > upperHeartRateBound AND isSustained()):
            notifyMessage("User's current heart rate is high");
            measureCriticalLevelBasedOnGap(heartRate, upperHeartRateBound);
        visualizeInGraph(activity, heartRate);
```

C. Context sensing and Interaction with the Environments

In addition to user activity and bio sensor data sensing and analysis, Mobile Wellbeing Companion Cloud interacts with the environments. It senses a range of environmental sensors and fuses them with the rest of the user data. In our test-bed, we use wireless temperature sensors that provide room temperature data either on demand or regularly (e.g. where the sleep times are some fixed period e.g. 30 seconds). Using MWCC knowledge base that identified optimal temperature for different user activities and heart rate conditions, the MWCC combines the level of external temperature with current heart rate to give notification about potential significant change in temperature which may harm users' health [31].

D. MWCC Communication Layer and Realtime Interactive Dashboard

In our testbed, we have 7 MWCCs connected via adaptive smart opportunistic and DTN communication protocol that is resource aware, social and privacy aware - CafREP[32]. Each MWCC communicates via intermediary MWCC or the infrastructure to gather sensor measurements as well as to share different data driven by both data requirements and network in different ways. The Figure 6 and 7 below show real-time interactive dashboard with varying heart rate when user performs different activity. The data from source node are shared by source node and visualized on the destination node with 100% successful rate using CafREP protocol for increases reliability.

For emergency situation, the notifications are forwarded via CafeREP to other nodes as shown in Figure 8.

Figure 6. Dashboard Heart Rate = 72 bmp when user is lying

Figure 7. Dashboard Heart Rate = 74 bmp when user is walking slowly

Figure 8: MWCC dashboard notification to the user

In order to estimate the cost for sensing, processing, visualizing and communicating data in real time in our test-bed, MWCC tracks and visualizes load average (as combined value for CPU, memory, I/O, etc.) as given in Figure 9. We observe that the load average remains below 5 for different types personal cloud services for long period of time.

Figure 9. Load average when running the Well-being Companion Cloud

V. CONCLUSION

We proposed a novel Mobile P2P Personal Cloud that can host different intelligent distributed services and described a novel “feel and communicate” mobile wellbeing companion service hosted on it. We explain a range of functionalities and multi-layer smart algorithms that we designed, built and deployed to
support users to sense the context, analyze, store, inter- with and share their sensitive data with others. We argue that our approach will help improve users trust and compliance with the new technologies which maintain the reasonable performance while significantly improving reliability and privacy. In future work, an integration of smart knowledge-based system with different contextual sensed data to generate meaningful information when the resources are limited and sensed data is noisy or conflicted, will be investigated further. Moreover, a promising future direction is to integrate our low-cost mobile personal cloud with cognitive radio sensor networks to address resource limitations as argued in [39].

VI. REFERENCES
[23] Girija Chettya, Matthew Whiteb, Farnaz Akther. Smart Phone Based Data Mining For Human Activity Recognition, International Conference on Information and Communication Technologies, 46(1), 1181-1187. 2015.