Efficient 3G/4G Budget Utilization in Mobile Sensing Applications

Chao Xu, Shaohan Hu, Wei Zheng, Tarek F. Abdelzaher, Pan Hui, Zhiheng Xie, Hengchang Liu, and John A. Stankovic, *Fellow, IEEE*

Abstract—This paper explores efficient 3G/4G budget utilization in mobile sensing applications. Distinct from previous research work that either relied on limited WiFi access points or assumed the availability of unlimited 3G/4G communication capability, we offer a more practical mobile sensing system that leverages potential 3G/4G budgets that participants contribute at will, and uses it efficiently customized for the needs of multiple mobile sensing applications with heterogeneous sensitivity to environmental changes. We address the challenge that the information of data generation and WiFi encounters is not *a priori* knowledge, and propose an online decision making algorithm that takes advantage of participants' historical data. Three typical mobile sensing applications, vehicular application, mobile health and video sharing application are explored. Experimental results demonstrate that our proposed algorithms lead to significantly better system performance compared to alternative solutions for both applications.

Index Terms—Mobile sensing, vehicular application, mobile health, video sharing, system performance optimization

1 Introduction

In this paper, we develop a novel smart phone based mobile sensing system that achieves efficient utilization of limited 3G/4G budgets to improve system performance. This work is motivated by the emergence of multiple types of mobile sensing applications [1], [2], [3], where data are collected from smartphones and wearable devices, stored locally, then offloaded to backend servers via WiFi or 3G/4G. We assume that users will typically not allow mobile sensing applications to use 3G/4G communication without limitation, since unlimited data plans are no longer prevalent [4], [5]. The WiFi-based store-and-forward approach, on the other hand, may result in large latency motivating the work described in this paper.

Vehicular applications and mobile health are two most important mobile sensing applications nowadays. Vehicles become popular mobile sensing platforms mainly due to two reasons. First, their natural mobility increases coverage for many mobile sensing applications [6]. Second, our daily commute itself has become a target of many research efforts, such as those that aim to save fuel consumption [3], find available parking positions [7], avoid traffic jams or routes in bad condition [1], [2], [8], or share general road-side events [9]. With the help of wearable devices and smartphones, mobile health generates tremendous amounts of location-rich, real-time, high-frequency data [10], [11]. Remote monitoring on common chronic diseases such as diabetes [12], asthma [13], [14], drop foot [15], and depression [10] has been extensively studied in recent years. The main difference between these two mobile sensing applications lies in that mobile health is much more safety-critical and thus needs real time response when emergency situation occurs. In addition, video based application is a special type of mobile sensing, as the generated data are typically with large size, results in potential data offload failure due to mobility. Video sharing applications are becoming more popular in the social sensing research field.

Exploiting users' own phones avoids additional investment costs to participants. Compared to placing conventional PC-like devices in cars/buses or people ourselves [1], [9], [16], smart phones are more pervasive and easy to use, while meeting application requirements of sensing, computation, and storage.

The philosophy underlying our work is that we believe many participants are indeed capable of contributing a budget of 3G/4G data. They either still use an unlimited data plan, or have a limited data plan but only use a small portion of it every month. The incentive for these participants to contribute a 3G/4G budget in support of mobile sensing applications is that they want to have their own vehicular or health data delivered and analyzed more reliably and quickly without extra cost, leading to improved feedback services for themselves.

Advances of technologies have made smart phones nowadays powerful enough to run multiple mobile sensing applications simultaneously. These applications typically have heterogeneous sensitivity/tolerance to environmental

- C. Xu and H. Liu are with the Department of Computer Science, University of Science and Technology of China, Hefei 230026, China. E-mail: chaoxu03@mail.ustc.edu.cn, hcliu@ustc.edu.cn.
- S. Hu and T.F. Abdelzaher are with the Department of Computer Science, University of Illinois at Urbana-Champaign, Champaign, IL 61801. E-mail: shu17@illinois.edu, zaher@cs.uiuc.edu.
- W. Zheng is with Sanofi-Aventis US LLC, Bridgewater, NJ 08807. E-mail: 18.weizheng@gmail.com.
- P. Hui is with the Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Hong Kong, Telekom Innovation Laboratories, Berlin 10587, Germany, and Aalto University, Helsinki 02150, Finland. E-mail: panhui@cse.ust.hk.
- Z. Xie and J.A. Stankovic are with the Department of Computer Science, University of Virginia, Charlottesville, VA 22904. E-mail: zx3n@virginia.edu, stankovic@cs.virginia.edu.

Manuscript received 16 Apr. 2015; revised 4 June 2016; accepted 28 July 2016. Date of publication 10 Aug. 2016; date of current version 3 May 2017.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TMC.2016.2599156

changes, for example, applications for traffic monitoring or parking lot availability are no doubt more sensitive than those for finding all Starbucks in a city or raising greener driving habits for aggressive drivers. For a patient with diabetes, applications for monitoring glucose have no doubt higher priority than those for detecting high blood pressure or monitoring pulse. This difference of properties among applications, typically represented by *utility* of received data, brings further opportunity to optimize the quality of information (QoI) during offloading process under the limited 3G/4G budget constraint.

The main contribution in this paper lies in that we develop a novel communication framework in mobile sensing applications, in which a decision making algorithm is designed to assign the limited 3G/4G budget to the sensory data of multiple mobile sensing applications for better overall utility. The biggest challenge behind this problem is that the information of data generation and encounter of WiFi access points is not a *priori* knowledge, thus conventional deterministic resource allocation methods are not applicable here. Instead, our solution predicts these information by taking advantage of participants' historical data, and provides an online decision making algorithm to decide which sensory data should be offloaded via 3G/4G communication while others wait for WiFi access points. Our approach is applied to three typical mobile sensing applications: vehicular application, mobile health and video based application.

For vehicular applications, our solution is evaluated by experimental results from a campus-wide deployment with 30 participants, each driving at least 100 miles. Simulation results of replaying the generated sensory data and WiFi encounters in the deployment demonstrate that our proposed solutions successfully improve the overall utility of received data, and can be tailored for heterogeneous needs of multiple mobile sensing applications. For mobile health, we fully implement our solution on off-the-shelf Google Nexus 5 phones, and our solution is evaluated by experimental results from a six-week-long deployment with 10 participants. Results of replaying the generated healthcare data and WiFi encounters in the deployment demonstrate that our approach successfully achieves better system performance, especially increases timely data delivery significantly for high-risk healthcare data. For video based application, our solution is evaluated by the FCVID Video Dataset [17] with 91,223 different video clips of 239 categories. Experimental results show that our solution successfully improves the total score of uploaded videos by 728 percent when budget is 900 MB.

The remainder of this paper is organized as follows. We compare our work with state of the art in Section 2 and present the system design for vehicular and mobile health applications in Sections 3 and 3.3, respectively. The evaluation for our proposed solution is discussed in Section 4. Finally, we conclude the paper in Section 5.

2 Related Work

Most prior vehicular mobile sensing applications have focused on leveraging smart phones placed in vehicles. For example, the Nericell project [2] presents a system that performs rich sensing using smartphones that users carry with them in normal course, to monitor road and traffic conditions. The GreenGPS system [3] provides a service that computes fuel-efficient routes for vehicles between arbitrary end-points, by exploiting vehicular sensor measurements available through the On Board Diagnostic (OBD-II) interface of the car and GPS sensors on smart phones. SignalGuru [1] is a software service that relies solely on a collection of mobile phones to detect and predict the traffic signal schedule, producing a Green Light Optimal Speed Advisory (GLOSA). These systems rely on WiFi access points, and assume transmitting data through 3G/4G networks is not desirable. However, open public WiFi is becoming less prevalent as more access points are becoming private or secure, resulting in a big delay time for generated sensory data to be delivered. Our paper aims to overcome this drawback by allowing participants who have many remaining 3G/4G data every month to contribute a reasonable 3G/4G budget without extra cost to help improve the system performance.

Most prior work on mobile health applications have focused on healthcare data analytics to assist patients with chronic diseases. For example, the Empath project [10] presents a real-time depression system for the home through monitoring various data including sleep, weight, activities of daily living and speech prosody. DexterNet [11] is an telemonitoring system based on a wireless body sensor network, which focuses on monitoring activity patterns and cumulative exposures to air pollution, transferring of this data to a health information system, and feedback of information to the user on how to manage activity, and reduce the potential for asthma exacerbation. The FIDES project [18] is a telemedicine solution that allows the pharmacy customers to perform self-monitoring operations using a predefined set of medical analysis devices. The collected data are used to build what can be considered a sort of electronic case history of the user that is stored in a centralized database bank. Similar to those work in vehicular applications, they encounter the problem of WiFi scarcity and 3G/4G cost, and we propose in this paper a novel solution that utilized predefined 3G/4G budget to reduce data delivery latency.

Consumption of mobile data by the pervasive usage of smart phones is forcing carriers to find ways to offload the network. Since the modern smart phones have been introduced worldwide, more and more users have become eager to engage with mobile applications and connected services. This eagerness has boosted up sales in the market more than 64 percent up annually worldwide in Q2 2010 [19]. Simultaneously, smartphone owners are using an increasing number of applications requiring the transfer of large amounts of data to/from mobile devices. As a consequence, the traffic generated by such devices has caused many problems to 3G/4G network providers. AT&T's subscribers in USA were getting extremely slow or no service at all because of network straining to meet iPhone users' demand [4]. The company switched from unlimited traffic plans to tiered pricing for 3G/4G data users in summer 2010. Similarly, Dutch T-Mobile infrastructure has not been able to cope with intense 3G/4G traffic, forcing the company to issue refunds for affected users [5]. Meanwhile, carriers are willing to use more pervasive technologies, such as Wi-Fi access points and hot spots. The proliferation of modern

Wi-Fi enabled smartphones, together with the network providers tendency towards already existing technologies has turn Wi-Fi offloading into a reality. However, it is reported that the WiFi coverage is quite limited, usually under 20 percent, even in big cities [20].

Several existing work have investigated making use of different types of communications for data dissemination/collection purposes. Wiffler [20] is a system to augment access to 3G/4G network through WiFi offloading, by leveraging delay tolerance and fast switching of devices. However, it focuses only on Internet access from moving vehicles. Han et al. proposed MoSoNet [21], the first work to exploit opportunistic communication to alleviate 3G/4G traffic, achieved by using a target set and 3G/4G recovery. However, this approach only works for data dissemination and is not applicable for data collection process as in our mobile sensing applications. The VIP-delegation work [22] proposed a data dissemination/collection model based on social groups. Lee et al. measured the performance of 3G/4G mobile data offloading through WiFi networks, and emphasized the incentive of delayed offloading to save traffic and energy [23]. The MultiNets [24] is a system capable of switching between wireless network interfaces (e.g., 3G/4G and WiFi) on mobile devices in real-time, to achieve higher throughput and save energy. Besides the differences in both the nature of the problem and the application scenario compared to our work, these related work consider all data are of the same type and importance, while we take one step forward and address the scenarios of multiple applications running simultaneously, which is the trend as smart phones are more powerful nowadays. Moreover, all of them considering 3G/4G communications assume there is no limitation for 3G/4G usage, which is not reasonable in the near future; instead, we take advantage of those who are capable of contributing a 3G/4G budget to the mobile sensing applications, and address the problem of efficient utilization of these precious 3G/4G data.

3 System Design

In this section, we present the system design for efficient 3G/4G budget utilization for various mobile sensing applications. We first describe the system model, then explain our proposed algorithms in detail for three different mobile sensing applications.

3.1 System Model

Our system is designed to operate in a mobile sensing network of n mobile nodes (vehicles, patients, or smart phones) that can generate sensor data via smart phones of participants. Each sensor node generates N types of data packets. The sampling rates of these packets are denoted by $\lambda_1, \lambda_2, \ldots, \lambda_N$, respectively. Some of these packets are more sensitive/tolerant to environmental changes than others, and the utility functions of these packets are $U_1, U_2, \ldots, U_N$. We assume that, for every data packet, its utility is a monotonically decreasing function of time t. All nodes have buffers of fixed size C_0 that can be used to store packets.

The mobile sensing area is partially covered by WiFi access points. When a sensor node moves into the range of an access point, data packets in its buffer are offloaded to a backend server through WiFi communications. Without loss of generality, we assume that the WiFi capacity is large enough to offload all data packets in the buffer. This assumption is reasonable because participants usually can only access those WiFi networks at their homes, offices, and some public places like shopping malls, and spend quite a long time in these places. In addition, the size of mobile sensing data typically are small as long as no large amount of video clips are generated.

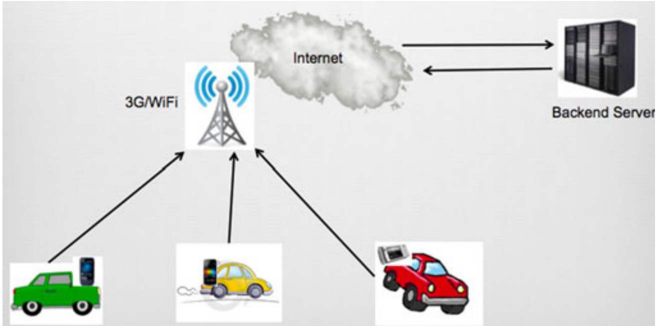


Fig. 1. Mobile sensing system model.

Each sensor node is associated with an amount of 3G/4G budget. This budget is determined by the participant itself, based on his/her 3G/4G data plan, living style, and willingness to contribute to the mobile sensing applications. For simplicity, the mobile sensing process is divided into multiple cycles, and the budget is equally assigned into each cycle. For instance, 600 bytes of budget per month can be transferred to 20 bytes per day. This assignment is additive, which means the remaining budget from previous cycles is accumulated into the new one.

Under this system model (as shown in Fig. 1), we are interested in solving the following problem: At any time point in a cycle, whether packets in the buffer should be transmitted through 3G/4G communications if there are still 3G/4G budget remaining? If so, how much and which of them should be sent? This problem is challenging mainly because the information of when new sensor data are generated and the participants can encounter an access point is NOT *a priori* knowledge, and utility functions are different among multiple mobile sensing applications. In the rest of this section, we first describe the 3G/4G-budget online algorithm to predict these information based on the participants' historical data, then present a heuristic algorithm running in a simple and effective fashion.

3.2 Optimization in Vehicular Applications

3.2.1 *Online 3G/4G-Budget Algorithm*

As detailed above, in practice the global information about data generation and WiFi access point encounters is unknown in advance. In this section, we develop an online algorithm that does not assume the availability of these information and estimate them based on historical data of participants.

Let each data plan period have K time cycles and each time point t_0 in a data plan period can be expressed as time t of cycle c_t. At time t of c_t within a data plan period, we assume there are n data clusters of size $q_1, \ldots, q_n$ in queue. Each of the data clusters is consisted of data packets generated at the same time and from the same type of application.

These n data clusters have utilities $u_1, \ldots, u_n$ at time t_0. The objective of this online algorithm is to decide at this time point, which of the n data clusters should be uploaded and how much of them should be uploaded using 3G/4G communications. Assume that when WiFi is connected for the first time after time t_0, the utilities of them become $v_1, \ldots, v_n$. After time t_0 and within the current data plan period, there are m data clusters generated with size $Q_1, \ldots, Q_m$ and utilities $U_1, \ldots, U_m$. These data clusters have utilities $V_1, \ldots, V_m$ when WiFi is first connected after the data are generated. All of $m, v_1, \ldots, v_n, Q_1, \ldots, Q_m, U_1, \ldots, U_m$ and $V_1, \ldots, V_m$ are unknown.

Assume that x_i out of q_i is uploaded at time t using 3G/4G network and the rest of ith data cluster will be uploaded via wireless when an access point is encountered. The reason that besides x_i, no data is uploaded using 3G/4G network is, given the fixed budget, the uploaded data will cost the same amount of budget and an earlier upload would have larger utility.

Therefore, for the data in queue at time t_0, the total uploaded utility is

$$\begin{aligned} F_t(x_1, \ldots, x_n) &= \sum_{i=1}^{n} x_i u_i + \sum_{i=1}^{n} (q_i - x_i) v_i \\ &= \sum_{i=1}^{n} v_i q_i + \sum_{i=1}^{n} (u_i - v_i) x_i. \end{aligned}$$

Then after time t_0, the total budget left is $B - \sum_{i=1}^{n} x_i$. Therefore, the maximum uploaded utility for the data packets generated after time t_0 is

$$\begin{aligned} & G_t(x_1, \ldots, x_n) = \\ & \max_{y_1, \ldots, y_m} \quad \sum_{i=1}^{m} V_i Q_i + \sum_{i=1}^{m} (U_i - V_i) y_i \\ & s.t. \qquad 0 \leq y_i \leq Q_i, i = 1, \ldots, m \\ & \qquad\qquad \sum_{i=1}^{m} y_i \leq B - \sum_{i=1}^{n} x_i. \end{aligned}$$

The solution for $x_1, \ldots, x_n$ is to maximize

$$E[F_t(x_1, \ldots, x_n) + G_t(x_1, \ldots, x_n)],$$

where the expectation is over $v_1, \ldots, v_n$, m, $Q_1, \ldots, Q_m$, $U_1, \ldots, U_m$ and $V_1, \ldots, V_m$.

In practice, the optimization is computationally expensive, so instead of maximizing the expectation of the total uploaded utility, we estimate the unknown m, $v_1, \ldots, v_n$, $Q_1, \ldots, Q_m$ and $V_1, \ldots, V_m$ using the historical data and then solve for $(x_1, \ldots, x_n)$. Here the estimations from the historical data are based on individual cycles. To estimate m and Q_i's, the mean cluster size generated from each type of application at every time point within a cycle using historical data is computed. Then the estimator of m, $\hat{m}$ is simply the number of data clusters with non-zero mean size. The estimators for Q_i's, $\hat{Q}_i$'s are the mean cluster sizes multiplied by $(K - c_t)$ for the time points prior to or at time t and multiplied by $(K - c_t + 1)$ for the time points after t. To estimate $v_1, \ldots, v_n$ and $V_1, \ldots, V_{\hat{m}}$, the mean time to wireless connection for each time point within a cycle is computed from historical data and the estimators $\hat{v}_1, \ldots, \hat{v}_n, \hat{V}_1, \ldots, \hat{V}_{\hat{m}}$ are obtained using utility functions.

Then $x_1, \ldots, x_n$ can be estimated by solving

$$\max_{x_1, \ldots, x_n, y_1, \ldots, y_{\hat{m}}} \quad \sum_{i=1}^{n} v_i q_i + \sum_{i=1}^{n} (q_i - x_i) \hat{v}_i \tag{1}$$

$$+ \sum_{i=1}^{\hat{m}} \hat{V}_i \hat{Q}_i + \sum_{i=1}^{\hat{m}} (\hat{U}_i - \hat{V}_i) y_i \tag{2}$$

$$s.t. \qquad 0 \leq x_i \leq u_i, i = 1, \ldots, n \tag{3}$$

$$0 \leq y_i \leq \hat{Q}_i, i = 1, \ldots, \hat{m} \tag{4}$$

$$\sum_{i=1}^{n} x_i + \sum_{i=1}^{\hat{m}} y_i \leq B. \tag{5}$$

We note that this is a standard linear programming problem. In addition, with the assumption that all data packets have approximately the same size, the number of data packets that should be uploaded via 3G/4G at time t_0 can be solved using following approach,

1) sort the all data clusters, including data clusters in current queue and future data clusters estimated from historical data, in the decreasing order of utility loss, i.e., $u_i - \hat{v}_i$ and $\hat{U}_i - \hat{V}_i$
2) if the budget B allows, the data packets with larger $u_i - \hat{v}_i$ or $\hat{U}_i - \hat{V}_i$ will have higher priorities of being uploaded via 3G/4G network, according to the order derived above. If a future data cluster has priority and the budget also allows, even though it is not generated at time t_0, budget will be reserved for it. Thus there will be less budget for the data behind it. A data packet in queue at time t_0 will be uploaded at time t_0 via 3G/4G network if all data clusters in front of it have not taken up all available budget.

The online 3G/4G-budget algorithm needs to keep the following historical data to process: assuming that there are n time points in a cycle, and the total number of data types is m. First, a $n \times m$ matrix is used to record average amount of data generated for each data type at any time point in previous cycles. Then another $n \times m$ matrix to keep the number of cycles for the first matrix is required. Similarly, two vectors with length n are needed to record the average time to meet WiFi access point at each time point in previous cycles and the number of cycles for these records. Therefore, totally $2n(m + 1)$ records need to be kept, and the total storage overhead highly depends on n indicating the frequency of running the algorithm.

3.2.2 Heuristic Algorithm

One drawback of the online 3G/4G-budget algorithm is that it requires a large amount of computation to update the matrixes for historical data and run the algorithm at all time points, therefore, it may consume a big amount of storage and energy for resource-constraint mobile phone platforms.

Based on this observation, we propose a heuristic algorithm to provide a simple and effective solution. The idea is to split the overall 3G/4G budget in each cycle into two pieces: reserved budget B_1 and flexible budget B_2. Namely, B_1 is reserved for those applications that are sensitive to environmental changes, denoted by *SENSITIVE*.

applications in *SENSITIVE* can be selected by setting a threshold by the application provides about when they think the data is not interested any more, and a predefined threshold can be used to differentiate *SENSITIVE*. Based on the historical information, the average amount of data generated for *SENSITIVE* in a cycle can be calculated, denoted by n'. To make the reservation more conservative, we also set a balance coefficient, denoted by α, to n'. Thus, B_1 can be obtained by

$$B_1 = \alpha \cdot n'. \tag{6}$$

The flexible budget, B_2, can be used by those applications that are not sensitive to environmental changes, denoted by *Non-SENSITIVE*. Moreover, applications in *SENSITIVE* can use this flexible budget as well if B_1 runs out and there are still remaining B_2.

When a new cycle starts and there are remaining data in the queue from the last cycle, these data will be uploaded using the new budget following the same rule as above. The order of upload in *SENSITIVE* and *Non-SENSITIVE* is based on the application thresholds set by different providers, and those with more sensitive properties have higher priority.

Within a cycle, the algorithm only runs at time points when new data are generated and the budget is not empty. Decisions are made for data in the queue in a greedy fashion: If there is budget in the right category, send that data via 3G/4G; otherwise, wait for WiFi communications.

3.3 Optimization in Mobile Health

Mobile health Data are divided into three categories based on predefined rules: high risk, medium risk, and low risk. We assume that the utility of data with high and medium risk monotonically decreases with time because these data are the fundamental evidence of disease diagnosis and treatment, and meanwhile that the utility of low-risk data is a constant over time; and each mobile node has sufficient space to store packets, which is no doubt a reasonable assumption for off-the-shelf smartphones today. We also assume that when the data is generated from the wearable devices, additional supportive data may also be generated. The supportive data may be provided by the users or generated by the smartphones, which could be used to help the clinicians further diagnose or treatment the disease. For example, the users may provide a brief voice record that describes users' diet or workout patterns at a particular time that could explain the clinical abnormalities detected by the wearable devices.

We first introduce the notations for the analytical model. We assume that at current time, data of size $s_0^{(1)}$ is generated from a medical wearable device and supportive data of $s_0^{(2)}$ is also generated. We assume that there is D days left in the current budget cycle and the total 3G/4G budget reserved for health applications within the current cycle is B. We denote the size of the data generated from different medical devices in the future time within the current cycle by $s_i^{(1)}$, and the time that the data is generated by $t_i^{(1)}$, $i = 1, \dots, N$, where $t_1^{(1)} < t_2^{(1)} < \cdots < t_N^{(1)}$. We also denote the size of the supportive data by $s_i^{(2)}$, $i = 1, \dots, N$.

We assume that when the data of size $s_i^{(1)}$ is generated from a medical wearable device at time $t_i^{(1)}$, a risk score r_i is also generated based upon the nature of the data. The risk score r_i can be categorized into three classes by two thresholds v_1 and v_2: high risk when $r_i > v_2$, medium risk when $v_1 < r_i \le v_2$, and low risk when $r_i \le v_1$. The thresholds v_1 and v_2 should depend on the user's general health conditions and medical experts' recommendations. For example, a cancer patient that is going through chemotherapies may be more likely to have some common side effects, thus the clinicians may recommend different values of v_1 and v_2 for this patient. We assume that the utility for the low-risk data remains a constant l_1 over time. For the high-risk data, the initial utility is l_3 and due to the importance of the high-risk data, the utility decreases to 0 immediately at the next time point. For the medium-risk data, the initial utility is l_2 and it immediately decreases to $l_2/2$ at the next time point. Here, we assume that $2l_1 < l_2 < l_3$.

We also assume that at current time, there are also data generated previously awaiting uploading. Their sizes are denoted by a_i, $i = 1, \dots, m$; their utilities are denoted by b_i's; and their risk scores are denoted by c_i's.

In order to best utilize the 3G/4G budget, it is to identify whether the data just generated and the data in the queue should be uploaded using 3G/4G. Obviously the if the data is to be uploaded via 3G/4G, it should be uploaded as early as possible to ensure better overall utility. Then it is to solve

$$\begin{aligned}
\max &\sum_{i=1}^{m}\left[x_i a_i + (1-x_i)\left(\frac{l_2}{2} I_{(v_1 < c_i \le v_2)} + l_2 I_{(c_i \le v_1)}\right)\right] \\
&+ \sum_{i=0}^{N}\left[y_i s_i + (1-y_i)\left(\frac{l_2}{2} I_{(v_1 < r_i \le v_2)} + l_2 I_{(r_i \le v_1)}\right)\right] \\
s.t.\quad &\sum_{i=1}^{m} x_i a_i + \sum_{i=0}^{N} y_i s_i \le B \\
&x_i \in \{0,1\}, y_i \in \{0,1\},
\end{aligned}$$

where $s_i = s_i^{(1)} + s_i^{(2)}$ for $r_i > v_1$ and $s_i = s^{(1)}$ for $r_i \le v_1$. This is equivalent to upload the data with the biggest utility difference between the one at generation time and the smallest possible utility. Thus the optimization problem above is also to solve the following optimization problems in the order of M_1, M_2 and M_3

$$\begin{aligned}
M_1:\quad &\max_{H,H_0} \sum_{i=0}^{N} I_{(i\in H)} I_{(r_i > v_2)} s_i + \sum_{i=1}^{m} I_{(i \in H_0)} I_{(c_i > v_2)} a_i \\
M_2:\quad &\max_{H,H_0} \sum_{i=0}^{N} I_{(i\in H)} I_{(v_1 < r_i \le v_2)} s_i \\
&\quad + \sum_{i=1}^{m} I_{(i \in H_0)} I_{(v_1 < c_i \le v_2)} a_i \\
M_3:\quad &\max_{H,H_0} \sum_{i=0}^{N} I_{(i\in H)} I_{(r_i \le v_1)} s_i + \sum_{i=1}^{m} I_{(i \in H_0)} I_{(c_i \le v_1)} a_i \\
s.t.\quad &\sum_{i=0}^{N} I_{(i\in H)} s_i + \sum_{i=1}^{m} I_{(i\in H)} a_i \le B,
\end{aligned}$$

where H is a subset of $\{0, 1, \dots, N\}$, H_0 is a subset of $\{1, \dots, m\}$ and they contain the indexes of the data that are uploaded via 3G/4G at current time. If $s_i^{(1)}$'s, $s_i^{(2)}$'s and r_i's are known, the solution to the optimizations is straight forward. The data with high-risk scores should be uploaded immediately and the corresponding personals should be notified. The data with medium-risk scores should be uploaded via 3G/4G too if there is already enough budget for all the high-risk data. For the low-risk data, they would only be uploaded via 3G/4G if enough budget has been reserved for the high-risk and medium-risk data and the main objective of collecting them is to build up the health profile of the user.

However in practice, these variables are unknown as the data are generated in future. We treat them as random variables and re-formulate the problem using the distribution estimated from the historical data. When it is to decide whether certain data should be uploaded via 3G/4G, we reserve certain budget for the high-risk and medium-risk data generated in future and take it into consideration. It is obvious that if the current data is of high risk, we will upload it via 3G/4G immediately. When it is of medium-risk or low-risk, we will need to check if there is budget available besides B_H and B_M. Let A_H be the total size of the high-risk data, including the supportive data; let A_M be the total size of the medium-risk data, including the supportive data; and let A_L be the total size of the low-risk data, i.e., $A_H = \sum_{i=1}^N s_i I_{(r_i > v_2)}$, $A_M = \sum_{i=1}^N s_i I_{(v_1 < r_i \leq v_2)}$ and $A_L = \sum_{i=0}^N s_i^{(1)} I_{(r_i \leq v_1)}$. Let B_H and B_M be the reserved budget for high-risk and medium-risk budget. The parameters B_H and B_M are determined by the probabilities that the future high-risk and medium-risk data will be covered, i.e.,

$$
\begin{aligned}
&P\big[B_H \geq \min(A_H, B - s_0 I_{(r_0 > v_2)})\big] \geq 1 - \alpha_1 \\
&P\big[B_M \geq \min(A_M, \max(B - s_0 I_{(r_0 > v_1)} - B_H, 0))\big] \\
&\geq 1 - \alpha_2.
\end{aligned}
$$

Then when the medium-risk data is generated, it is to check if there is enough budget besides reserved B_H for the future high-risk data. If there is not enough budget, then the data is assigned to the medium-risk data queue. For the low-risk data, we use a slightly different strategy to better utilize the budget. If the current data is of low risk, we put it in the low-risk data queue and it is to wait for time T before it decides whether this data is uploaded via 3G/4G. This strategy helps the scenario when there is a unusual amount of data with high/medium risk in the near future.

To estimate B_H and B_M, we estimate the distribution of A_H and A_M from the historical data. In this paper, we assume that the amount of high-risk data and medium-risk data over time follow a homogeneous Poisson process, i.e., for any given day, the size of high-risk data C_H and medium-risk data C_M have the following distribution:

$$
\begin{aligned}
P(C_H = k) &= \exp(-\lambda_H) \frac{\lambda_H^k}{k!} \\
P(C_M = k) &= \exp(-\lambda_M) \frac{\lambda_M^k}{k!}.
\end{aligned}
$$

Then λ_H and λ_M can be estimated directly from historical data, denoted by $\hat{\lambda}_H$ and $\hat{\lambda}_M$. Then A_H and A_M follow $Poisson(\hat{\lambda}_H T_L)$ and $Poisson(\hat{\lambda}_M T_L)$, where T_L denotes the time left in current cycle in days. The parameters B_H and B_M can be derived using the distribution percentiles of the Poisson distribution. We may also use a second-order Markov Chain model to predict the amount of high-risk and medium-risk data based on the historical data generated shortly before the current time.

3.4 Optimization for Video Sharing

In the video sharing application, data are uploaded when users share the video clips recorded on their smartphones. Since some videos are similar, and video clips usually have large size and partially uploaded clips are not able to be shared to other participants, we have to consider the size of video data as well. Therefore, we need to reduce the similar uploaded videos for better utilization of 3G/4G budget.

We first introduce the notations for the analytical model. We assume that at current time, a video, denoted by v_i, with size of l_i and category of c_i is generated, and there are already m videos in the server. The category of video is a hierarchical name space, which is analogous to a path prefix in a UNIX directory tree. We assume that the left budget is B. $\parallel c_i \parallel$ is the length of c_i.

Score is the measure of the dissimilarity between videos. The current video score can be computed as

$$
\begin{aligned}
s_i &= \frac{d_i}{l_i} \\
d_i &= \frac{\sum_{j=1}^m d_{ij}}{m} \\
d_{ij} &= 1 - \frac{\parallel P(c_i, c_j) \parallel}{\parallel c_j \parallel}.
\end{aligned}
$$

In order to diversify the uploaded videos with limited 3G/4G budget, the total score of uploaded videos is maximized, d_{ij} is the dissimilarity between v_i and v_j, $P(c_i, c_j)$ is the common sub prefix of c_i and c_j. Then it is to solve

$$
\begin{aligned}
&\max \sum_H s_i \\
&s.t. \quad \sum_H l_i \leq B,
\end{aligned}
$$

where H is the uploaded video set in the end.

We compare the score of current video clip and the expected score of the future one video clip, which is computed as

$$
\begin{aligned}
s_n &= E[score] \\
&= \sum_{k=1}^C (p_{n,k} \cdot s_k) \\
s_k &= \frac{d_k}{l_k},
\end{aligned}
$$

where C is number of category that already appeared. $p_{n,k}$ indicates the probability of next video come from c_k, which can simply be estimated by the number of video generated of c_k. l_k is the size of next video from c_k, which can also be estimated by historical size data of c_k with the assumption that the size of the videos follows Exponential distribution. The video size distribution assumption is based on historical video size observation. Considering the amount of

Fig. 2. The hardware set.

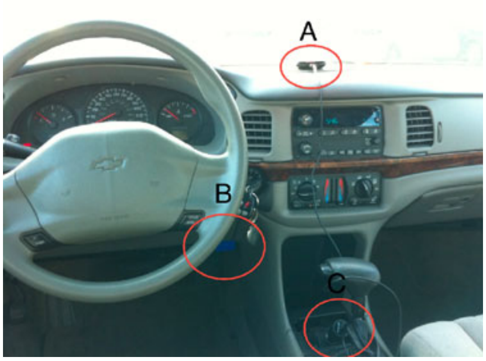


Fig. 3. System in deployment.

videos generated in future could be large, the original optimization problem could be time consuming. Thus we only compare the current and the next immediate video clip to decide whether a video clip should be uploaded using 3G/4G, and this method can only get a suboptimal performance. The algorithm is summarized as follows:

1) When v_i generated. If WiFi is connected, upload via WiFi immediately; Otherwise skip next
2) Get the score of v_i and forecast the score of next one video
3) Compare the s_i and s_n. If $s_i \geq s_n$ and $B > 0$, upload v_i via 3G/4G immediately and update B; Otherwise discard v_i

4 Evaluation

In this section, we evaluate the performance of our approaches for all three mobile sensing applications.

4.1 Experimental Setting

We now give detailed information on the experiments and dataset we use in the evaluation of our solutions. To evaluate the proposed communication scheme, we conducted two human subjects study in 2012 in vehicular application and 2014 in mobile health. In addition, we use the FCVID Video Dataset for evaluation in video sharing applications.

The study in vehicular application involved 30 participants (drivers), and each of them was required to drive at least 100 miles. While we expect a mobile sensing application to run on participants' own phones, in this study we gave our subjects phones pre-loaded with our software (of which 15 were Galaxy Nexus phones [25], shown at the bottom right in Fig. 2, and 15 were Nexus S phones [26], shown at the top right). These phones were placed under the windshield of their vehicles, as shown in circle A in Fig. 3. Participants included students from various departments. To emphasize the performance in vehicular environments, they were asked to keep our phone in their car when driving, but otherwise carry on their daily routines as usual. They were not restricted to any specific routes and were not asked to change their normal driving habits. The phones were kept charged through a lighter-to-USB charger (Circle C in Fig. 3) to support a long-term experiment and reduce the odds of individuals otherwise forgetting to charge a phone they do not use. When the phone was not charging (i.e., the engine is off), sensor data collection stopped but the communication component continued to run. The mobile sensing application in this deployment was to collect various types of sensory data about the vehicles while participants were driving, including GPS locations (GPS), on-board diagnostics (OBD) data [27], acceleration values (ACC), and gyros (GYRO) data. Collection of this data set was a good experimental methodology because these data can be used for multiple purposes, such as enhancing greener driving habits and analyzing traffic conditions.

Engine data were collected by the ELM 327 OBD-to-bluetooth adaptor (bottom center in Fig. 2), connected to the OBD-II port, which is a standard feature in all vehicles sold in the US since 1996, and usually located under the dashboard (as shown in Circle B in Fig. 3). Collected engine information included engine speed, RPM, mass air flow, and throttle position. This information was transmitted to the smart phone via bluetooth. The GPS, acceleration, and gyro data were collected via corresponding sensors on the phone respectively. GPS samples were obtained at 1 sample/second for both types of phones. The sampling rate of the OBD data, accelerometer, and gyroscope sensors were all set to be as fast as the hardware allowed. We observed that the actual data rate varied depending on car engine and phone. These generated data were also displayed on a logging interface screen so as to notify the participants that the system was running. All data were saved in a local database on the phone, and offloaded eventually to our backend server, when an available WiFi access point was present. We used the UIUC-campus-wide IllinoisNet as the only authentication-required WiFi that was available to the phones, which implies that participants can upload their data when they drive to office but not when they are at home. All open public WiFi access points could be used as well.

To record each meaningful event and generate a complete dataset, a log file was stored locally on each smart phone. It included the timestamp and content of GPS, OBD, ACC, and GYRO data, the start and end timestamps of each WiFi access point encounter, as well as which sensory data packets were offloaded to the backend server during each encounter. In addition, the timestamps when the car engine was ON/OFF were also recorded for statistics purposes. Finally, we noticed that 10 participants finished this deployment by taking one long-distance trip instead of driving locally for a long period of time, thus we excluded them out and only focus on the other 20 participants.

The study in mobile health involved 10 participants including professors and students at the University of

Fig. 4. Hygeia user interface on the Nexus 5 phone.

Science and Technology of China to mimic chronic patients, and each of them was required to carry a Google Nexus 5 smartphone. The whole data collection process lasted for one and a half month. We argue that a total of 10 participants is sufficient because the performance comparison of alternative solutions is independent of mobility patterns for different participants. In this study we gave our subjects phone pre-loaded with our software. To emphasize the performance, they were asked to keep carrying our phone in their daily life, but otherwise keep up their daily routines as usual. They were not restricted to any specific routes and not asked to change their normal habits. The phones were kept charged to support a long-term experiment when participants were sleeping at night. The application in this deployment was to collect various types of sensory data about the participants, including heart rate, blood pressure and blood glucose. These data can be used for disease monitoring, such as heart disease, high blood pressure and diabetes. Fig. 4 shows the Hygeia user interface on a Google Nexus 5 phone.

The data were collected by external wearable sensors, and transmitted to smartphone via Bluetooth. The sampling rate of heart rate, blood pressure and blood glucose were set to be as fast as the hardware allowed. All data were saved in a local database on the phone, and offloaded eventually to our backend server, when an available WiFi access point was present. In order to emphasize the mobility characteristics, we intentionally disconnected the WiFi connections at participants' homes, while allowing them to offload data at other WiFi places, such as offices, cafes, and supermarkets. In this way, timely data delivery via 3G/4G communications becomes the dominating manner.

To record each meaningful event and generate a complete dataset, a log file was stored locally on each smartphone. It included the starting and ending timestamps of each WiFi access point encounter. At the server side, the arrival time of data packets and their values were recorded.

4.2 Methodology

We proceed to describe the methodology adopted in our evaluation.

For vehicular application, we first obtain the whole dataset by combining the local log files of 20 participants. Then we divide each log file into multiple cycles for 3G/4G budget assignment. In our evaluation we set the length of a cycle to 24 hours. In each cycle, the detailed information of generated data and WiFi encounters is available. We define the GPS data as the *SENSITIVE* category to simulate real-time traffic monitoring applications, in which the data is not fresh any more after 30 minutes; and the other three types of data as the *NON-SENSITIVE* category to simulate fuel consumption analysis and driving habit improvement applications, in which the data can last as long as a month. For simplicity of calculation, we define the utility function, donated by $f(t)$, for a piece of data in *SENSITIVE* (*NON-SENSITIVE*) as: $f(\cdot)$ starts from 1 when the data is generated, and decreases linearly as time continues, finally achieves a predefined threshold at time of 30 minutes (30 days). We set the threshold to 0.001 in our experiment.

We compare the system performance of three candidate solutions: *Baseline*, *3G/4G-Budget*, and *Heuristic*. *Baseline* represents the default case in which 3G/4G communications are not applicable and data can only be offloaded by WiFi access points. The results for *Baseline* can be calculated directly from the dataset. The other two solutions, on the other hand, require the replay of the dataset by running their own algorithms as described in Sections 3.2.1 and 3.2.2, respectively. Varying amounts of 3G/4G budget are adopted to evaluate its impact on system performance for applications in different categories.

For mobile health, We first obtain the whole dataset by combining the local log files of 10 participants. Then we divide each log file into multiple cycles for 3G/4G budget assignment. In our evaluation we set the length of a cycle to 24 hours. In each cycle, the detailed information of generated data and WiFi encounters is available. We collected three types of data, including heart rate, blood pressure(diastolic blood pressure and systolic pressure) and blood glucose. As explained earlier in Section 3.3, we classify each kind of data into three different ranks, including high risk, middle risk and low risk, which indicates the rank of potential danger to the participant. For each type of data, its rank of risk is determined by predefined thresholds, v_1 and v_2. For heart rate data, $v_1 = 100/\mathrm{minute}$ and $v_2 = 120/\mathrm{minute}$. For blood glucose data, $v_1 = 7.0\ \mathrm{mmol/L}$ and $v_2 = 9.0\mathrm{mmol/L}$. For diastolic blood pressure, $v_1 = 140\ \mathrm{mmHg}$ and $v_2 = 180\ \mathrm{mmHg}$. For systolic pressure, $v_1 = 90\mathrm{mmHg}$ and $v_2 = 110\ \mathrm{mmHg}$. α_1 and α_2 are set to 0.1 and 0.2, respectively.

We compare the system performance of three candidate solutions: BestEffort, Baseline and Hygeia. Baseline represents the default case in which 3G/4G communications are not applicable and data can only be offloaded by WiFi access points, and aims to display the lower bound of system performance. The results for Baseline can be calculated directly from the dataset. The other two solutions, on the other hand, require the replay of the dataset by running their own algorithms. The BestEffort solution is an alternative solution tailored for real time monitoring purposes. It also allows participants to assign a 3G/4G budget in each cycle, however, it forces each generated data, no matter what risk rank it is, to be offloaded via WiFi or 3G/4G within a very short time (10 seconds in our experiment), otherwise this piece of data will be considered as obsolete and discarded. Varying amounts of 3G/4G budget are adopted to evaluate its impact on system performance for applications in different categories.

For video sharing application, the metric we use is total score of video clips uploaded to the backend server. The

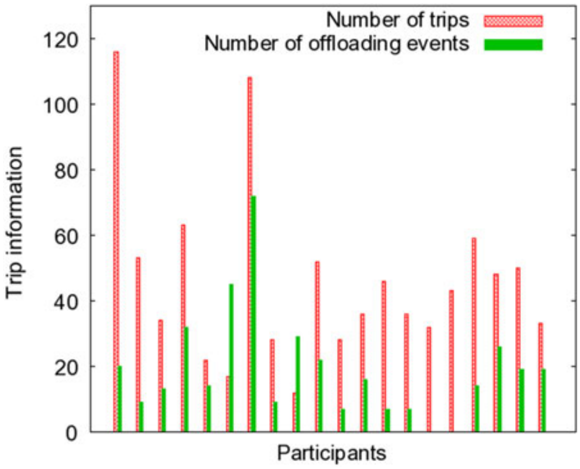


Fig. 5. Trip information for participants.

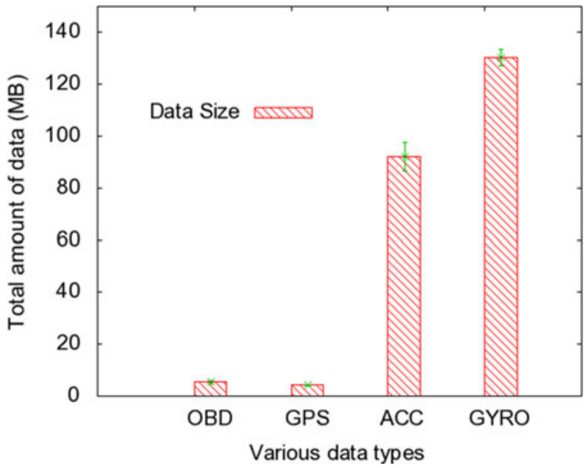


Fig. 6. Comparison of total data amount among various data types.

amount of budget for 3G/4G communication ranges from 0 to 2,400 MB, stepping by 300 MB. Five participants are involved. We compare our algorithm, *forecast*, with two other algorithms, *forward* and *baseline*. The *baseline* algorithm only uses WiFi to upload video clips. The *forward* algorithm uploads a video clip directly when it is generated as long as there are remaining 3G/4G budgets.

4.3 Results

In this section, we present some statistics from our deployment, and analyze how they are related to the performance of our proposed solutions.

4.3.1 *Vehicular Application*

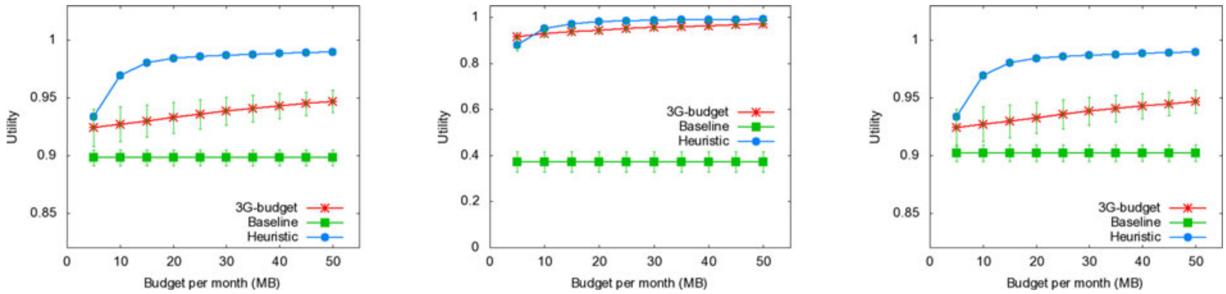
Fig. 5 presents the number of active trips and offloading events (where by offloading, we mean upload data to server) for the 20 participants. Here active trips are defined by the intervals in which sensory data are generated during driving. We set the threshold between two consecutive active trips to 30 minutes. Note that offloading events refer to those really have data uploaded to backend server, not including when the WiFi connection is established but there is not enough time to finish the socket connection and transfer data. We observe that both numbers vary a lot for different users. Some users take more trips than others in the deployment, within the range from 22 to 116 trips. This result indicates that some participants drive more frequently than others. The frequency of offloading events also varies for different users. For instance, we found that two participants were never able to finish an offloading event, because there are no WiFi coverage on their normal driving routines. On the other hand, one participant finished 45 offloading events in 17 active trips, which implies that his/her frequent driving route is well covered by WiFi. Overall, the number of offloading events is small compared to the number of active trips for most participants, this is reasonable because most students go to office/class by walk or school bus, and there are no IllinoisNet coverage when they drive to other off-campus places such as grocery stores or shopping malls. These results indicate that the utility for uploaded sensory data can be very low due to the fact that open wireless access points are not widespread. Our proposed solutions take advantage of the quick delivery of 3G/4G communications and improve the systems performance in terms of utility of uploaded data.

We also recorded the total amount of data in different types for each participant in the deployment. This gives us the comparison of data size in the *SENSITIVE* and *NON-SENSITIVE* categories. The results of total amount of data are shown in Fig. 6. We can see that the average total amount of data for OBD, GPS, ACC, and GYRO are 5.38, 4.26, 92.3, and 130.22 mega bytes (MB), respectively. These results indicate that there are 53.5 times of data generated for *Non-SENSITIVE* than the *SENSITIVE*, which implies that the *Non-SENSITIVE* will be dominating in the overall performance.

We now compare the performance of the candidate solutions, *Baseline*, *3G/4G-Budget*, and *Heuristic* (with balance coefficient set to 1), that we described in Section 4.2. The metrics are the utility of data received at the backend server. The amount of 3G/4G budget every month ranges from 5 to 50 MB, stepping by 5 MB. We choose this range mainly based on two reasons. First, the total amount of GPS data was around 5 MB on average. Therefore, with this smaller budget, we can investigate the scenario in which some *SENSITIVE* data are on hold in queue for the next cycle when new budget is available. Second, participants may not want to reserve too much budget from their limited data plan. Taking the popular AT&T 200 MB/month data plan [28] for example, one quarter of the limit, which is 50 MB, may be the largest possible budget contributed by participants. The budget is reset to 0 after each 30 cycles to reflect that data plan starts for a new month. The dataset for each participant is applied to the three candidate solutions, and statistics such as mean and standard deviation are calculated.

Fig. 7 shows the results of data utility with error bars using different candidate solutions under varying amount of 3G/4G budget. First, we notice that the overall utility for *Baseline* is 0.89, and only 0.37 for *SENSITIVE*. These results indicate that data offloading only by WiFi access points is not sufficient enough, especially for those applications that are quite sensitive to environmental changes.

Second, we observe that both *3G/4G-budget* and *Heuristic* bring a performance gain to the *Baseline* solution, even when the 3G/4G budget is 5 MB. The average utility increases from 0.37 to 0.93 using *3G/4G-budget* and 0.88



(a) Overall. (b) *SENSITIVE*. (c) *Non-SENSITIVE*.

Fig. 7. Utility Comparison: *Baseline*, *3G/4G-budget*, and *Heuristic*.

using *Heuristic* for *SENSITIVE*, which implies 151.4 percent and 137.8 percent improvements, and from 0.90 to 0.96 using *3G/4G-budget* and 0.93 using *Heuristic* for *NON-SENSITIVE*. These results demonstrate that 3G/4G communications are indispensable in mobile mobilesensing applications, and a small portion of data plan (e.g., 2.5 percent of 200 MB in our case) can lead to a boost of system performance especially for those applications that are more sensitive to environment changes.

Third, it is obvious to see that, the performance goes up as the amount of 3G/4G budget increases for *3G/4G-budget* and *Heuristic* in all cases. This is because more packets are offloaded faster through 3G/4G communications. Especially, the utility for *SENSITIVE* increases from 0.934 at 5 MB to 0.961 at 50 MB using *3G/4G-budget* and from 0.879 at 5 MB to 0.991 at 50 MB using *Heuristic*. The variance of data utility decreases too as the amount of budget increases. These results indicate that the more budget participants are willing to contribute to the mobile sensing applications, the better system performance it will lead to.

Fourth, the performance of *3G/4G-budget* is slightly better than *Heuristic* for *SENSITIVE* when the budget is very small (5 MB) compared to the amount of generated data. The utility is 0.93 for *3G/4G-budget* and 0.88 for *Heuristic*. This is mainly because only a portion of budget is reserved to *SENSITIVE* using *Heuristic*, and the other part is quickly consumed by data in *Non-SENSITIVE*, while *3G/4G-budget* holds on the budget for a while to wait for the generation of new *SENSITIVE* data. When the budget increases from 10 to 50 MB, *Heuristic* outperforms *3G/4G-budget* because more budget are assigned to *SENSITIVE* in *Heuristic*, but the data running *3G/4G-budget* sometimes are on hold if the decision making algorithm estimates that a WiFi encounter is approaching. We also notice that *Heuristic* results in better utility than *3G/4G-budget* for *Non-SENSITIVE*. The main reason is that the driving pattern for participants are not regular enough for accurate estimation in *3G/4G-budget*. We noticed that most participants did not drive much on weekdays, probably because they went to office/class by walk or school bus. This makes it a common case that much more data are generated on weekends than weekdays, therefore, the estimation is affected by anticipating that more data will be generated later in the data plan period with larger utility loss and thus decides to hold the budget, but finally that does not happen.

Fig. 8 depicts the average frequency ratio of *Heuristic* to *3G/4G-budget* in cycles with new data generated for each participant. For instance, a value 0.36 for Participant 3 indicates that *Heuristic* runs 36 percent time of *3G/4G-Budget* on average for this user. In general, we observe that *3G/4G-budget* runs a lot more than *Heuristic* for most users. The ratio ranges from 0.03 to 0.88 and the average is 0.39, which implies that *Heuristic* only runs less than 40 percent than *3G/4G-budget* on average.

In addition, both *3G/4G-budget* and *Heuristic* perform well in terms of customized for the needs of different types of mobile sensing applications. The comparison of Figs. 7b with 7c clearly shows that data in *SENSITIVE* obtained more performance gain than those in *Non-SENSITIVE*, based on their different utility functions. These results demonstrate that our proposed solutions are capable of dealing with multiple heterogeneous mobile sensing applications and provide tailored solution to improve the overall utility in the system.

Finally, we compare the utility performance using *Heuristic* under varying balance coefficients, ranging from 1 to 1.2 stepping with 0,05. The 3G/4G budget is 5 MB. The results are shown in Fig. 9. We can see that the utility for *SENSITIVE* increases from 0.879 to 0.893 while the utility for Non-SENSITIVE decreases from 0.933 to 0.932. These results indicate that, when reserving more budget for *SENSITIVE*, more data in *SENSITIVE* were offloaded faster through 3G/4G communication and thus the improved performance; while on the other hand, data in *Non-SENSITIVE* have to wait longer on average, resulting in poorer performance. Due to the relatively smaller 3G/4G budget compared to the total amount of data in *Non-SENSITIVE*, the decreasing ratio is very small.

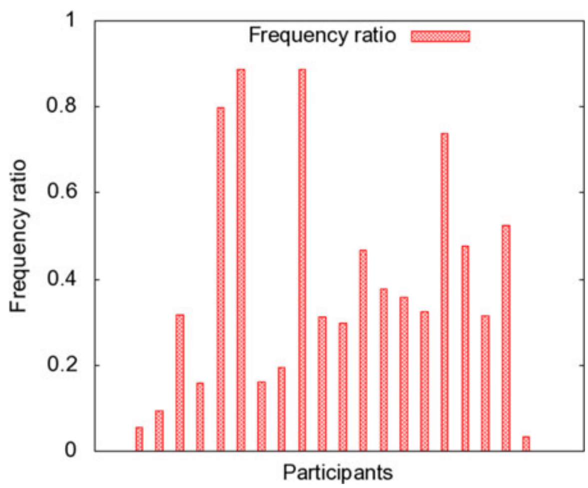


Fig. 8. Frequency ratio of *Heuristic* to *3G/4G-budget* on active cycles.

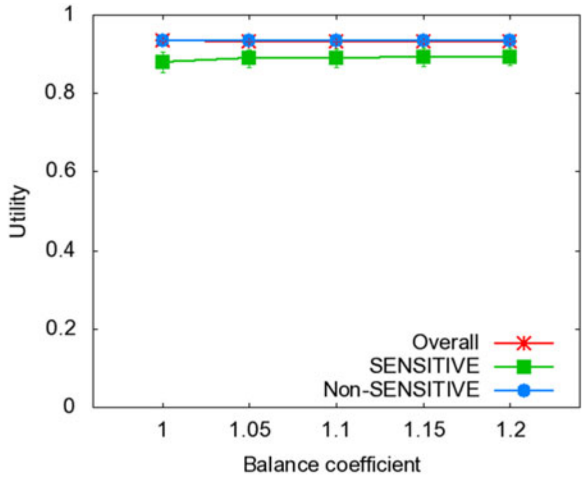


Fig. 9. Utility performance under varying balance coefficient.

4.3.2 Mobile Health

We proceed to present the experimental results in mobile health. The metric we use is the percentage of data uploaded to the backend server. The amount of budget for 3G/4G communication rages from 1 to 15 MB, stepping by 2 MB. We choose this range mainly based on the following reason. We observe that the total amount of high risk data was around 0.24 MB on average. Therefore, with this budget setting we can investigate the scenario in which medium risk and low risk data are on hold in queue even when there is still 3G/4G budget available, to reserve sufficient resources for high risk data. The dataset for each participant is applied to the three candidate solutions, then average value and standard deviations are calculated.

Figs. 10 and 11 presents the size of three types of generated healthcare data: blood glucose, blood pressure, and heart rate; and the average size of offloaded data for the 10 participants, including high risk, medium risk, and low risk data, using different candidate solutions under varying amount of 3G/4G budget. First, we notice that the size of offloaded data increases as the budget increases, for example, from 1.48 to 11.65 MB as the budget changes from 1 to 15 MB. These results indicate that the more budget participants are willing to contribute to the mobile health applications, the better system performance it will lead to.

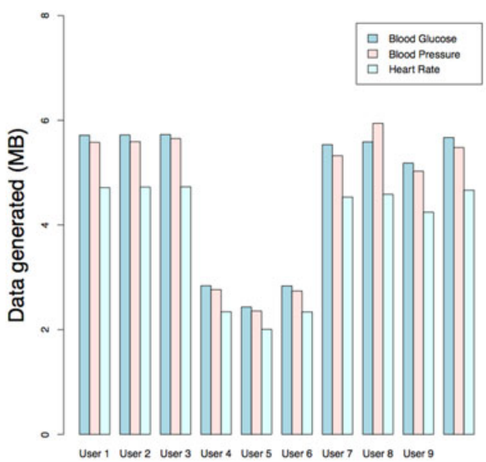


Fig. 10. Three types of generated healthcare data.

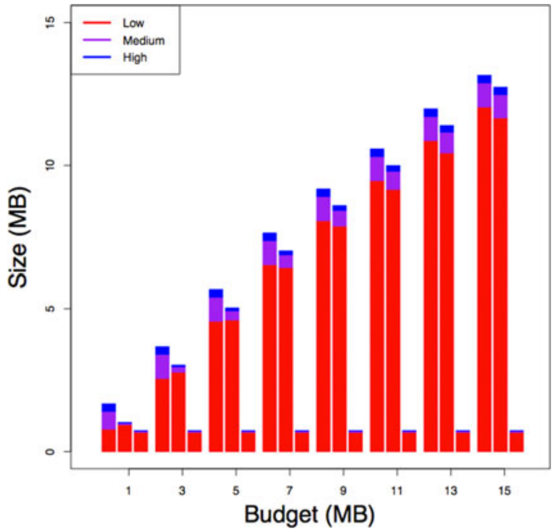


Fig. 11. Total size of offloaded data (candidates from left to right: Hygeia, BestEffort, and Baseline).

Second, we observe that both Hygeia and BestEffort bring a performance gain to the Baseline solution, even when the 3G/4G budget is just 1MB. This is clear because free WiFi access points, except at home, are very few. The average offloaded data increases from 1.48 to 11.65 MB using Hygeia, which implies 2.3 and 23.3 times improvements, and from 1.01 to 11.29 MB using BestEffort, which implies 1.4 and 22.8 improvements. Especially, improvements on high-risk data achieves 25.3 times using Hygeia and 24.4 times using BestEffort, respectively. These results demonstrate that 3G/4G communications are indispensable in mobile health applications, and even a small portion of data plan (e.g., 1 MB in our case) can lead to a boost of system performance.

Third, the performance of Hygeia is slightly better than BestEffort in all cases. This is mainly because Hygeia stores low risk data in local queues and offloads them when WiFi APs are available, while BestEffort operates in an "offload it, or discard it" manner, thus it less utilizes WiFi APs than Hygeia.

Fig. 12 depicts the average percentage of offloaded data to the generated data in total with error bars. For instance,

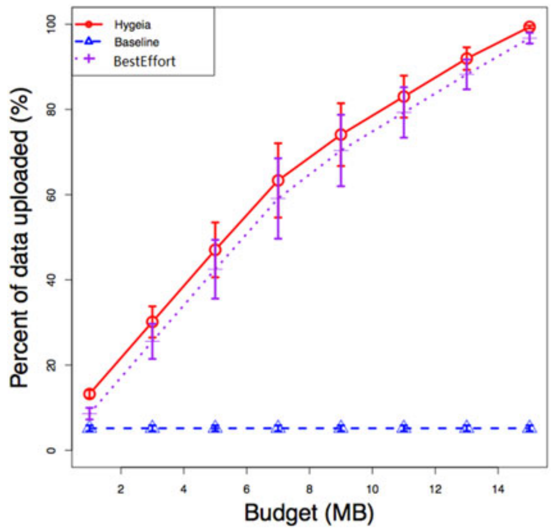
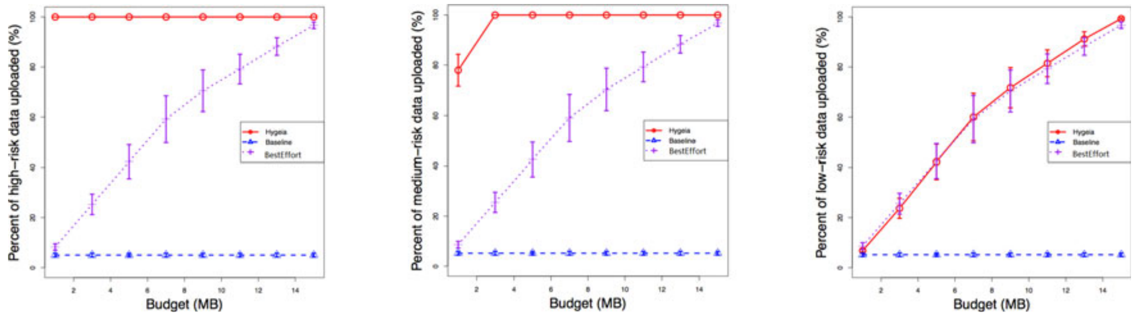


Fig. 12. Percentage of offloaded data in total.



(a) *High-risk*. (b) *Medium-risk*. (c) *Low-risk*.

Fig. 13. Percentage of offloaded data: *High-risk*, *medium-risk*, and *low-risk*.

a value of 13.5 indicates that Hygeia offloads 13.5 percent of generated data on average when the budget is set to 1 MB. In general, we observe that the percentage of offloaded data using Hygeia is slightly higher than using BestEffort, since a little more data are offloaded via WiFi APs. In addition, we observe that as the budget increases, the standard deviation for both Hygeia and BestEffort goes up to some point, then starts decreasing.

Figs. 13a, 13b, and 13c shows the average percentage of offloaded data to all generated data with error bars for high risk, medium risk, and low risk data, respectively. First, we are excited that high risk data always achieves 100 percent delivery ratio using Hygeia, which is the most important factor for data collection platforms in mobile health. The Baseline solution only delivers 3.8 percent high risk data, which is totally undesirable in this type of safety-critical application. The percentage of offloaded high-risk data using BestEffort solution increases from 9.34 percent with 1 MB budget to 96.73 percent with 15 MB budget. This is mainly because the BestEffort approach takes each piece of data as the same and offloads it as long as 3G/4G budget is still available. Therefore, the delivery ratio of high risk data drops significantly as the budget size decreases.

Second, the delivery ratio of medium risk data when using Hygeia stays at 100 percent when the budget size is lager than or equal to 3 MB, then drops to 83.1 percent when the budget size becomes 1 MB. This results indicates that our proposed *budMH* algorithm effectively reserves sufficient resources for possible high risk data based on historical data, as illustrated in Fig. 13a.

Finally, we notice that the percentage of received low risk data using Hygeia is almost the same with using BestEffort. This is reasonable because low risk data occupy the majority of generated data. Meanwhile, we can see that when budget size is very small, say 1 or 3 MB, the percentage using BestEffort is slightly higher than Hygeia, due to the fact that Hygeia utilize the precious 3G/4G resources more cautiously for potential high risk and medium risk data. This result validates our design philosophy behind Hygeia, which is, to guarantee the offloading of high risk data, transfer medium risk data as much as possible if budget allows, and only offload low risk data when 3G/4G resource is quite sufficient to cover all generated healthcare data.

4.3.3 Video Sharing

We proceed to present the experimental results in Video Sharing. Fig. 14 shows the total score of uploaded video clips. We observe that *Forecast* achieves the largest score compared to *Forward* and *Baseline*. For instance, when the budget is 900 MB, the scores of *Forecast*, *Forward* and *Baseline* are 51.6, 7.2, and 3.5, respectively. when the budget is 2,400 MB, the scores of *Forecast*, *Forward* and *Baseline* are 215.7, 15.6, and 3.5, respectively. This result show that our proposed *Forecast* solution successfully optimizes the usage of limited 3G/4G budget and improves the overall system performance. Fig. 15 shows the average score of uploaded video clips. We can see similar trends and *Forecast* achieves the best performance.

4.4 Discussions

Experimental results in all three applications show that our proposed solutions achieve better system performance in

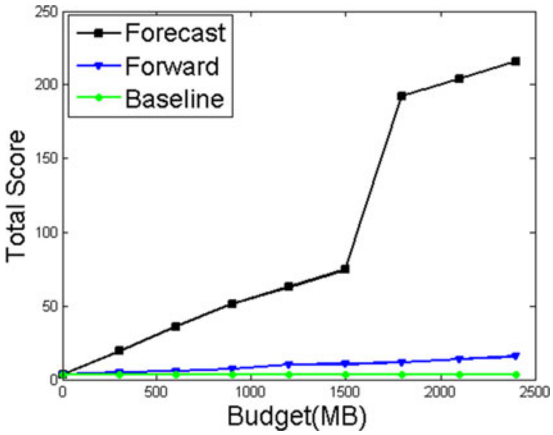


Fig. 14. Total score of uploaded video clips.

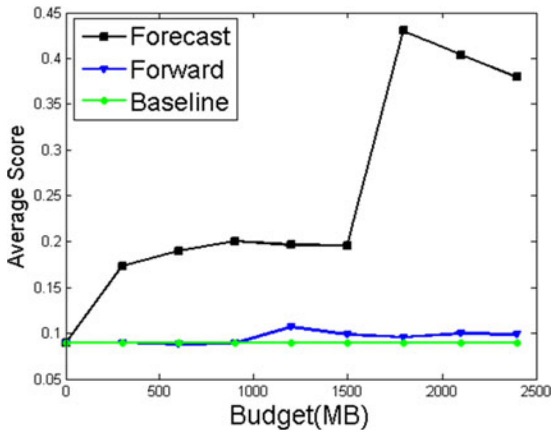


Fig. 15. Average score of uploaded video clips.

terms of data utility. Note that we only considered this utility function for the specific scenario, performance when using other possibilities of the utility function is beyond the scope of this paper. Our communication framework can be further optimized by other potential research directions. For instance, there are potential opportunities to combine our work with the smart-phone based vehicular networking techniques that have been emerging recently [29]. Due to the unknown *priori* information, opportunistic peer-to-peer sharing can help offloading data packets. Finally, although we focus on 3G/4G communication in this paper, our design is general enough for latest 4G/5G communications as well.

5 Conclusion

In this paper, we presented the design, implementation, and evaluation of a novel mobile sensing system which leverages the 3G/4G budget that participants contribute and is customized to the heterogeneous needs of multiple mobile sensing applications. Our proposed algorithms improve the utility of uploaded mobile sensing data and no user involvement is needed. Experimental results from three different mobile sensing applications demonstrate that our proposed algorithms lead to significantly better system performance compared to alternative solutions.

Acknowledgments

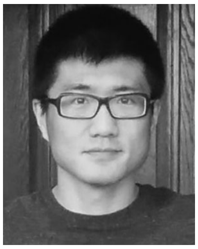
This work was partially supported by NSFCGG2260080042, and NSFC-BJ2260080039. Chao Xu and Hengchang Liu are the corresponding author.

References

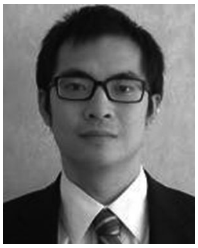
[1] R. Balan, N. Khoa, and J. Lingxiao, "Real-time trip information service for a large taxi fleet," in *Proc. 9th Int. Conf. Mobile Syst. Appl. Services*, 2011, pp. 99–112.
[2] P. Mohan, V. N. Padmanabhan, and R. Ramjee, "Nericell: Using mobile smartphones for rich monitoring of road and traffic conditions," in *Proc. SenSys*, 2008.
[3] R. Ganti, N. Pham, H. Ahmadi, S. Nangia, and T. F. Abdelzaher, "GreenGPS: A participatory sensing fuel-efficient maps application," in *Proc. 8th Int. Conf. Mobile Syst. Appl. Services*, 2010, 151–164.
[4] Customers Angered as iPhones Overload AT&T. NY Times, Sep. 2009. [Online]. Available: http://goo.gl/KtuInV, Accessed on: 2013.
[5] iPhone overload: Dutch T-Mobile issues refund after 3G issues. Ars Technica, Jul.2010. [Online]. Available: http://goo.gl/6DCNYh, Accessed on: 2013.
[6] E. Miluzzo, et al., "Sensing meets mobile social networks: The design, implementation and evaluation of the CenceMe application," in *Proc. 6th ACM Conf. Embedded Netw. Sensor Syst.*, 2008, pp. 337–350.
[7] S. Mathur, et al, "ParkNet: Drive-by sensing of road-side parking statistics," in *Proc. 8th Int. Conf. Mobile Syst. Appl. Services*, 2010, 123–136.
[8] A. Thiagarajan, et al., "VTrack: Accurate, energy-aware road traffic delay estimation using mobile phones," in *Proc. 7th ACM Conf. Embedded Networked Sensor Syst.*, 2009, pp. 85–98.
[9] B. Hull, et al., "CarTel: A distributed mobile sensor computing system," in *Proc. 4th Int. Conf. Embedded Networked Sensor Syst.*, 2006, pp. 125–138.
[10] R. Dickerson, E. Gorlin, and J. Stankovic, "Empath: A continuous remote emotional health monitoring system for depressive illness," in *Proc. WirelessHealth*, 2011.
[11] E. W. Seto, et al., "A wireless body sensor network for the prevention and management of asthma," in *Proc. IEEE Int. Symp. Ind. Embedded Syst.*, 2009, pp. 120–123.
[12] D. Preuveneers and Y. Berbers, "Mobile phones assisting with health self-care: A diabetes case study," in *Proc. 10th Int. Conf. Human Comput. Interaction Mobile Devices Services*, 2008, pp. 177–186.
[13] S. Alkobaisi, W. Bae, and S. Narayanappa, "A novel health monitoring system using patient trajectory analysis: Challenges and opportunities," in *Proc. 4th Int. Conf. Adv. Geographic Inf. Syst. Appl. Services*, 2012, pp. 147–151.
[14] S. Gupta, P. Chang, N. Anyigbo, and A. Sabharwal, "MobileSpiro: Portable open-interface spirometry for android," in *Proc. 2nd Conf. WirelessHealth*, 2011, Art. no. 24.
[15] T. Seel, D. Laidig, M. Valtin, C. Werner, J. Raisch, and T. Schauer, "Feedback control of foot eversion in the adaptive peroneal stimulator," in *Proc. 22nd Mediterranean Conf. Control Autom.*, 2014, pp. 1482–1487.
[16] N. Banerjee, M. D. Corner, and B. N. Levine, "Design and field experimentation of an energy-efficient architecture for DTN throwboxes," *IEEE Trans. Netw.*, vol. 18, no. 2, pp. 554–567, Apr. 2010.
[17] Y.-G. Jiang, Z. Wu, J. Wang, X. Xue, and S.-F. Chang, "Exploiting feature and class relationships in video categorization with regularized deep neural networks," *arXiv preprint arXiv:1502.07209*, 2015.
[18] G. Benelli, G. L. Daino, A. Pozzebon, R. Sesto, and R. Zambon, "Health monitoring and wellness for all, a multichannel approach through innovative interfaces and systems," in *Proc. 7th Int. Conf. Body Area Netw.*, 2012, pp. 96–99.
[19] *Canalys*, 2010. [Online]. Available: http://www.canalys.com/pr/2010/r2010081.html
[20] A. Balasubramanian, R. Mahajan, and A. Venkataramani, "Augmenting mobile 3G using wiFi," in *Proc. 8th Int. Conf. Mobile Syst. Appl. Services*, 2010, pp. 209–222.
[21] B. Han, P. Hui, V. S. A. Kumar, M. V. Marathe, G. Pei, and A. Srinivasan, "Cellular traffic offloading through opportunistic communications: A case study," in *Proc. 5th ACM Workshop Challenged Netw.*, 2010, pp. 31–38.
[22] M. V. Barbera, J. Stefa, A. C. Viana, M. D. de Amorim, and M. Boc, "VIP delegation: Enabling VIPs to offload data in wireless social mobile networks," in *Proc. Int. Conf. Distrib. Comput. Sensor Syst. Workshops*, 2011, pp. 1–8.
[23] K. Lee, I. Rhee, J. Lee, Y. Yi, and S. Chong, "Mobile data offloading: How much can WiFi deliver?" in *Proc. ACM SIGCOMM Conf.*, 2010, pp. 425–426.
[24] S. Nirjon, A. Nicoara, C.-H. Hsu, J. Singh, and J. Stankovic, "MultiNets: Policy oriented real-time switching of wireless interfaces on mobile devices," in *Proc. IEEE 18th Real Time Embedded Technol. Appl. Symp.*, 2012, pp. 251–260.
[25] Galaxy Nexus, 2013. [Online]. Available: http://goo.gl/zdd9f
[26] The Google Nexus S phone, 2013. [Online]. Available: http://www.google.com/phone/detail/nexus-s
[27] R. Birnbaum and J. Truglia, *Getting to Know OBD II*, Ralph Birnbaum, New York, 2001.
[28] The AT&T data plan. [Online]. Available: www.att.com/shop/wireless/plans.html
[29] S. Trifunovic, B. Distl, D. Schatzmann, and F. Legendre, "WiFi-Opp: Ad-hoc-less opportunistic networking," in *Proc. 6th ACM Workshop Challenged Netw.*, 2011, pp. 37–42.

Hengchang Liu received the PhD degree from the University of Virginia, in 2011, under the supervision of professor John A. Stankovic. He is now an assistant professor with USTC. He also worked as a postdoc at UIUC with professor Tarek Abdelzaher. His research interests mainly include cyber physical systems, mobile systems, named data networking, and wireless (sensor) networks.

Chao Xu received the BS degree from Anhui University, in 2014. He is currently working toward the graduate degree in Prof. Hengchang Liu's Group in the Department of Computer Science and Technology, USTC. His research interests mainly include cyber physical system and data mining.



Shaohan Hu received the MS degree from the Department of Computer Science, Dartmouth College. He is working toward the PhD degree in the Department of Computer Science, University of Illinois, Urbana-Champaign, where he is a research assistant. His main research interests lie in designing and building mobile sensing and computing systems



Wei Zheng received the PhD degree from the Department of Statistics, University of Wisconsin-Madison. He is now a statistician at Sanofi-Aventis US LLC. His research interests mainly include statistical modeling in clinical trials, data mining in mobile systems, and various types of networks.

Tarek F. Abdelzaher received the PhD degree from the University of Michigan, Ann Arbor, in 1999, under the supervision of Professor Kang Shin. He was an assistant professor with the University of Virginia, from August 1999 to August 2005. He then joined the University of Illinois at Urbana-Champaign as an associate professor with tenure, where he became a full professor in 2011. His interests lie primarily in systems, including operating systems, networking, sensor networks, distributed systems, and embedded realtime systems. He is especially interested in developing theory, architectural support, and computing abstractions for predictability in software systems, motivated by the increasing software complexity, and the growing sources of non-determinism. Applications range from sensor networks to large-scale server farms, and from transportation systems to medicine.

Pan Hui received the PhD degree from the University of Cambridge, under the supervision of Prof. Jon Crowcroft. He is with the Department of Computer Science and Engineering, Hong Kong University of Science and Engineering, Hong Kong, and Telekom Innovation Laboratories, Berlin, Germany. His research interests mainly include mobile augmented reality, mobile computing, and other interesting topics related to social network.

Zhiheng Xie received the master's degree in software engineering from Tsinghua University, under the supervision of Professor Chi-hung Chi. He received the PhD degree from the Department of Computer Science, University of Virginia, under the supervision of Professor John A. Stankovic. He is now a software developer at Groupon Inc. His research interests mainly include wireless sensor networks, mobile systems, cloud computing, and inertial navigation.

John A. Stankovic received the PhD degree from Brown University. He is the BP America professor in the Computer Science Department, University of Virginia. He served as chair of the department for eight years. He won the IEEE Real-Time Systems Technical Committee's Award for Outstanding Technical Contributions and Leadership. He also won the IEEE Technical Committee on Distributed Processing's Distinguished Achievement Award (inaugural winner). His research interests include real-time systems, distributed computing, wireless sensor networks, and cyber physical systems. He is a fellow of the IEEE and the ACM.