Research Collection School Of Computing and Information Systems

School of Computing and Information Systems

7-2017

# Fast adaptation of activity sensing policies in mobile devices

Mohammad Abu ALSHEIKH
*Nanyang Technological University*

Dusit NIYATO
*Nanyang Technological University*

Shaowei LIN
*Singapore University of Technology and Design*

Hwee-Pink TAN
*Singapore Management University*, hptan@smu.edu.sg

Dong In KIM
*Sung Kyun Kwan University*

# Fast Adaptation of Activity Sensing Policies in Mobile Devices

Mohammad Abu Alsheikh, Dusit Niyato, Shaowei Lin, Hwee-Pink Tan, and Dong In Kim

*Abstract*—With the proliferation of sensors, such as accelerometers, in mobile devices, activity and motion tracking has become a viable technology to understand and create an engaging user experience. This paper proposes a fast adaptation and learning scheme of activity tracking policies when user statistics are unknown a priori, varying with time, and inconsistent for different users. In our stochastic optimization, user activities are required to be synchronized with a backend under a cellular data limit to avoid overcharges from cellular operators. The mobile device is charged intermittently using wireless or wired charging for receiving the required energy for transmission and sensing operations. Firstly, we propose an activity tracking policy by formulating a stochastic optimization as a constrained Markov decision process (CMDP). Secondly, we prove that the optimal policy of the CMDP has a threshold structure using a Lagrangian relaxation approach and the submodularity concept. We accordingly present a fast Q-learning algorithm by considering the policy structure to improve the convergence speed over that of conventional Q-learning. Finally, simulation examples are presented to support the theoretical findings of this paper.

*Index Terms*—Activity tracking, fast adaptation, Internet of Things, Markov decision processes, wireless charging.

## I. INTRODUCTION

Activity tracking promises to revolutionize mobile user experience and helps in understanding the big data of today's world [1], [2]. Specifically, activity and motion data is required in many applications such as home security and automation, healthcare systems, contextual advertising, and smart vehicle technologies. Using mobile devices, such as mobile phones and Internet of Things (IoT) gadgets, for activity tracking has many benefits over conventional wearable sensor and body networks in terms of reachability, flexibility, and financial cost. Firstly, the mobile phone market has been rapidly scaling with more than $63\%$ international penetration rate in 2015 and 4.7 billion unique mobile subscribers [3]. Secondly, modern mobile devices are equipped with high-quality built-in sensors that can measure various physical quantities such as orientation, motion, ambient light, and location. Thirdly, mobile devices support efficient data transmission using cellular networks which facilitates backend integration and data synchronization. Fourthly, application stores, such as Google Play, enable the reach of a huge customer base for mobile crowdsensing in activity-aware systems and support software and patch upgrades.

Continuous activity and motion tracking is being deterred by the energy and monetary cost of mobile sensing. Firstly, mobile devices are battery-powered, and they deplete their energy within a few hours when in-device sensors operate continuously [4]–[8]. Wireless charging is gaining an increasing attention from hardware manufacturing companies as a seamless recharging method of mobile devices. The authors in [9] showed that a mobile device can be remotely charged using magnetic resonance coupling while being in user's pocket. Nonetheless, wireless charging is intermittent due to mobility and is not available at all locations. Secondly, cellular data plans are generally expensive, and continuous activity tracking can cause significant bill overcharges by cellular operators for data transmission and synchronization with a backend. Data synchronization is typically required for an up-to-date tracking of user activities and motion over time, and hence provide customized mobile services accordingly.

To address these issues, this paper proposes an adaptive activity tracking policy for mobile devices, and considers the intermittent (wired and wireless) charging and cellular data usage. The mobile sensing optimization is designed to minimize the detection error of user activities subject to a data usage limit. The main contributions and results of this paper are summarized as follows:

- In Section III, the activity tracking problem is formulated as a stochastic optimization using constrained Markov decision processes (CMDPs). A CMDP model [10] is a variant of Markov decision processes (MDPs) for stochastic optimization subject to a constraint on problem variables and feasible solutions. The temporal correlation of user activities is modeled as a discrete-time Markov chain (DTMC), and the CMDP tracking policy minimizes the detection error of user activities subject to a predefined data usage constraint.

- Using a Lagrangian relaxation approach [11], [12], we relax the CMDP formulation to an unconstrained MDP as discussed in Section IV. Then, the optimal tracking policy is found using conventional solution methods such as the value iteration algorithm. A Lagrange multiplier in the unconstrained optimality equation is found recursively to capture the data usage constraint.

M. Abu Alsheikh is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798, and also with the Sense and Sense-abilities Programme, Institute for Infocomm Research, Singapore 138632 (e-mail: stumyhaa@i2r.a-star.edu.sg).

D. Niyato is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (e-mail: dniyato@ntu.edu.sg).

S. Lin is with the School of Engineering Systems and Design Pillar, Singapore University of Technology and Design, Singapore 487372 (e-mail: shaowei_lin@sutd.edu.sg).

H.-P. Tan is with the School of Information Systems, Singapore Management University, Singapore 188065 (e-mail: hptan@smu.edu.sg).

D. I. Kim is with School of Information and Communication Engineering, Sungkyunkwan University, Korea 440-746 (e-mail: dikim@skku.ac.kr).

- Based on the unconstrained MDP, the threshold structure of the policy is proved in Section V. The MDP activity tracking policy is shown to be monotonically non-decreasing in the battery level, and hence the CMDP policy can be represented as a mixture of two threshold MDP policies. Accordingly, fast adaptation of Q-learning can be achieved based on the proved threshold structure of the activity tracking policy.

The remainder of this paper is organized as follows. Section II reviews related works in the literature. Section III presents a CMDP activity tracking policy for mobile devices. Then, the CMDP activity tracking policy is transformed to an unconstrained MDP using a Lagrange-based method as presented in Section IV. Based on the unconstrained MDP problem, the threshold structure of the activity tracking policy is proved in Section V, and a threshold Q-learning method that leverages the structure of the activity tracking policy is also discussed. The performance evaluation is presented in Section VI. Finally, the paper is concluded in Section VII.

## II. RELATED WORK

In this section, we first review related works on the applications of activity tracking systems. Then, we discuss the wireless charging technologies available for mobile devices. Finally, we review related works on mobile sensing optimization.

### A. Activity Tracking Using Mobile Devices

Mobile devices can be programmed to sense and adapt to the physical environment. The authors in [13] presented an algorithm for detecting human contexts, e.g., activities, disposition, and habits, which can be integrated with social networking services. A security method that uses human gestures for continuous authentication was proposed in [14]. In [15], a mobile sensing application in healthcare systems was discussed. The application monitors human physical activities, e.g., heart activity, to generate continuous feedback on health and behavior conditions. The authors in [16] proposed a method that infers user activities for automatic image tagging. Specifically, the rich tags include information about user activities and location, and surrounding ambient light and sound.

### B. Wireless Chargers for Mobile Devices

Wireless charging of mobile devices has seen great advancements in the last few years. This enables the remote charging of devices at a distance of a few meters, i.e., the mobile device is not required to be on the wireless charging pad as in the old technology. The authors in [9] proposed a wireless charging system, called "MagMIMO", for mobile devices based on the technology of magnetic resonance coupling. Similar to beamforming in multiple-input-multiple-output (MIMO) antennas, the proposed system embeds multiple coils in the power charger, and hence forms the magnetic field as a beam focused towards the mobile device. MagMIMO enables effective charging of one mobile device at a distance of $0.4$m from the

charger. The authors in [17] introduced "MultiSpot", a wireless charging system based on magnetic resonance which can charge up to 6 mobile devices simultaneously. The effective charging distance is $0.5$m. MultiSpot uses multiple coils in the wireless charger to beam the charging signal towards the mobile devices. "Wattup" [18] is a wireless charger of mobile devices that uses radio frequency (RF) radiation with an effective charging distance of 15 feet ($4.57$m). The mobile devices are first located using low-energy Bluetooth signals. After the successful localization, an RF signal, similar to the WiFi signal, is sent in the direction of the mobile devices. Wattup comes with a controlling software to select the devices to be charged. Cota [19] is another product that is based on the RF radiation technology. The effective charging distance is 10m.

These recent advancements have encouraged many companies to support wireless chargers in their products and stores. For example, IKEA, the international furniture retailer, has established a new production line that embeds wireless chargers in the furniture and home accessories [20]. Starbucks has started to install wireless chargers in some of its coffee shops worldwide [21].

### C. Mobile Sensing Optimization

Optimal mobile sensing of user activities is typically designed to maximize the detection accuracy under a resource constraint. For example, the authors in [4] used the knowledge about user's motion, location, and surrounding environment to manage sensor activation for detecting various activities. Specifically, the sensor activation is semi-automated and is based on manual settings and an apriori distribution. A related MDP-based method was also presented in [22] to continuously model the user mobility. The continuous sensing is avoided by exploring the location information. The authors in [5] presented an algorithm that uses accelerometer, microphone and GPS sensors to detect human activities and balances the detection performance and energy consumption. In [6], the mobile sensing problem was formulated as a CMDP. The design objective is to maximize the detection accuracy under a given energy constraint. Similarly, the authors in [7] proposed mobile sensing algorithms for accelerometer-based systems using CMDP and partially observable MDP models. The user behavior is assumed to be time-varying which is captured by statistical methods, e.g., the entropy-production rate. In [8], the sensor activation of a mobile tracking system was formulated using a hidden Markov model. The mobility pattern, residual energy, and cellular connection are evoked in predicting a schedule for sensor activation, e.g., a GPS sampling schedule.

This paper substantially differs from existing works in terms of the problem formulation, optimization objectives and constraints, and results. Existing works on activity sensing and tracking in the literature do not consider the user adaptation of tracking policies. Therefore, learning a policy for a particular user with conventional methods requires a large number of iterations which is expensive in mobile devices. This paper has clear novelty in providing fast user adaptation of activity tracking policies. In particular, the theoretical analysis employs

TABLE I: List of frequently used symbols throughout the paper.

| Symbol | Definition |
|---|---|
| $\mathcal{U}$ | User activity state space |
| $\mathcal{E}$ | Energy charging state space |
| $\mathcal{B}$ | Battery level state space |
| $\psi^n = (u^n, e^n, b^n)$ | System state at time $n$ containing the user activity $u^n$, number of acquired energy units $e^n$, and battery level $b^n$ |
| $\Delta = \{\delta_0, \delta_1\}$ | Action space defining the sleep $\delta_0$ and active $\delta_1$ modes |
| $c(\cdot)$ | Detection error function of the user activities |
| $g(\cdot)$ | Probability of connectivity to an access network |
| $\mathbb{P}\left(\psi^{n+1}|\psi^n, \delta^n\right)$ | Transition probability from state $\psi^n$ to state $\psi^{n+1}$ after taking action $\delta^n \in \Delta$ at time $n$ |
| $d(\cdot)$ | Data usage function for data synchronization with a backend |
| $\pi$ | Activity sensing policy defining the sensing action at each state |
| $\mathcal{J}(\cdot)$ | Average detection error under policy $\pi$ |
| $\mathcal{D}(\cdot)$ | Average data usage under policy $\pi$ |
| $\lambda$ | Lagrange multiplier |
| $\beta$ | Discount factor in the unconstrained MDP formulation |
| $\bar{b}$ | Average battery level |
| $\rho$ | Probability of successful data synchronization |
| $\tau$ | Probability of battery overflow |

a Lagrange relaxation approach along with the concept of submodularity to prove that the CMDP policy is a randomized mixture of two threshold MDP policies that are monotonically non-decreasing in the battery level. Our threshold analysis (a) enables fast online policy learning, e.g., using Q-learning, by substantially reducing the search space of the optimal activity tracking policy, and (b) curtails the storage space of the policy.

## III. Problem Formulation and Optimization

### A. Overview

Activity and motion tracking is becoming an integral functionality in modern mobile platforms such as Google Fit for Android[1], Apple Health for iOS[2], and Motion Data 2.0 for Windows Phone[3]. The detected user activities can be shared with all applications installed in a mobile device to provide interactive user experience. A modern example of activity-aware schemes is targeted advertising for dynamically delivering an advertisement based on user's activities and disposition which increases the revenue of both publishers and advertisers [23]. Figure 1 shows the system model as considered in this paper. The mobile sensing is systematically managed based on the user activity, battery level, and battery charging state of a mobile device. These parameters are used for taking decisions on optimal working modes. There are two working modes: (i) an active mode during which battery charging, data synchronization, and activity sensing can be

[1]https://fit.google.com/

[2]http://www.apple.com/sg/ios/health/

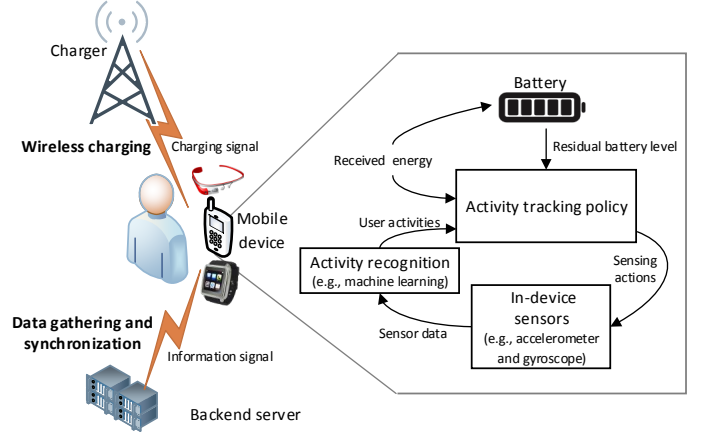[3]http://windows.microsoft.com/en-gb/windows-10/motion-data-privacy-faq



Fig. 1: System model including data synchronization and intermittent charging, e.g., wireless charging.

performed, and (ii) a sleep mode during which only battery charging is performed. The activity detector maps time series data into the most probable user activity using supervised machine learning techniques. This mapping process is beyond the scope of this paper. Nonetheless, we refer interested readers to [24], [25] for some pertinent results. The list of symbols used in this paper are summarized in Table I.

In this section, the mobile sensing problem is formulated as a finite state Markov model. To summarize, this section describes the following steps.

- Similar to previous works in [6], [22], [26], [27], user activities are assumed to evolve as a discrete-time Markov chain (DTMC). Then, the activity and motion tracking is formulated as an infinite horizon CMDP model with a single constraint. The CMDP consists of: decision epochs, system states, actions, transition probabilities, detection error and data usage functions, and a cellular data limit. The objective is minimizing the detection error of user activities subject to a cellular data limit.
- It follows from the results in [10], [28] that the CMDP sensing problem presented as a linear program can be solved in polynomial time to obtain a randomized, stationary, and optimal activity tracking policy.

### B. System Model

In what follows, an activity-aware mobile device is assumed to operate under a discrete time fashion with decision epochs denoted by $\mathcal{N} = \{1, 2, \ldots, N\}$, where $N$ is the sequence termination time. At each decision epoch $n \in \mathcal{N}$, the system selects a mobile sensing action based on its current system state and then transits to a new state. The system state space $\Psi$ of the mobile sensing problem is defined as follows:

$$\Psi = \bigg\{ (\mathcal{U}, \mathcal{E}, \mathcal{B}),$$
$$\mathcal{U} \in \{0, \ldots, U\}; \mathcal{E} \in \{0, 1\}; \mathcal{B} \in \{0, \ldots, B\} \bigg\}, \quad (1)$$

where $\mathcal{U}, \mathcal{E}$, and $\mathcal{B}$ represent the user activity state, energy charging state, and battery level of the mobile device, respectively. $U$ is the maximum number of supported user activities,

and $B$ is the maximum capacity of the battery of energy units. Consequently, state $\psi^n \in \Psi$ at time $n \in \mathcal{N}$ is defined using a 3-tuble as $\psi^n = (u^n, e^n, b^n)$ which includes the current user activity $u^n$, the number of newly acquired energy units $e^n$ ($e^n \in \{0, 1\}$), and the battery level $b^n$ at that decision epoch. Similar to previous works in [6], [22], [26], [27], the user activity states are assumed to evolve as a Markov chain with transitions that are stochastically involved. Under discrete time model, this assumption is typical as user activities have short memory.

The action space $\Delta = \{\delta_0, \delta_1\}$ includes two actions as follows:

$$\begin{cases} \delta_0 = 0, & \text{switch to the sleep mode,} \\ \delta_1 = 1, & \text{switch to the active mode.} \end{cases} \quad (2)$$

During the active mode, the mobile device can measure samples using its in-device sensors, e.g., an accelerometer, gyroscope, digital compass, microphone, and GPS. Additionally, the data synchronization can only be performed during the active mode. The mobile device is assumed to consume one unit of energy during the active mode while no energy is depleted in the sleep mode. It is important to note that these modes of operations are restricted to the activity-aware system and are separated from other applications running on the device. The detection error function of user activities $c(\cdot)$ is defined as

$$c(\cdot) : \Psi \times \Delta \to \mathbb{R}_+. \quad (3)$$

$c(\cdot)$ is a decreasing function on the user activities which is defined by considering the detection error of each activity by machine learning algorithms, i.e., $u_0 \in \mathcal{U}$ has the highest detection error. For example, the authors in [24] used decision trees and supervised neural networks to classify daily human activities with varying detection errors of $3.0 - 42.0\%$ and $4.0 - 78.0\%$, respectively. In [25], a deep learning model is designed to detect human activities from crowdsensing data which scores $3.0 - 47.0\%$ of varying detection errors. Clearly, user activities cannot be detected during the sleep mode $\delta_0$ such that $c(\psi, \delta_0) = 1.0$.

Another practical advantage of the proposed model is its consideration of the mobile device connectivity to a backend. This is important as the data connectivity depends on the available access networks in the area [29]. The probability of having wireless connection to an access network at each system state $g(\cdot)$ is defined as

$$g(\cdot) : \Psi \times \Delta \to [0, 1], \quad (4)$$

where $\times$ is the Cartesian product. $g(\cdot)$ can be defined based on the user activity, battery level, and the actions of the activity-aware mobile device. The connectivity probability to the activity-aware backend is zero when the mobile device is switched to the sleep mode, i.e., $g(\psi, \delta_0) = 0$, as no data transmission is allowed.

For simplicity, the battery charging probability is assumed to be sampled from a Bernoulli distribution with a success probability of $\mathbb{P}(e = 1)$. This charging probability can be construed as the probability that a mobile device is able to receive energy from a wireless charger. Nonetheless, other

more complex distributions can be adopted without affecting the problem formulation. The battery capacity is finite, takes integer values only, and follows Lindley equation [30] which is given as follows:

$$b^{n+1} = \min \left( [b^n - \delta^n]^+ + e^n, B \right), \quad (5)$$

where $[\cdot]^+$ is defined as $[z]^+ = z$ when $z > 0$, and it returns $0$ otherwise. During one time epoch of the battery charging, one energy unit is added to the battery of the mobile device unless the battery is full, i.e., the maximum capacity of a battery is finite and is replenished by wired or wireless charging. The mobile device consumes one unit of energy during a time epoch of active mode. This energy is used for both activity sensing, processing, and synchronization. We remark that our optimization model can be extended straightforwardly for arbitrary number of units of energy consumption and depletion, e.g., the received energy can change based on the distance between the wireless charger and the mobile device [9].

With the above setup, the transition probability $\mathbb{P}\left(\psi^{n+1}|\psi^n, \delta^n\right)$ from state $\psi^n = (u^n, e^n, b^n) \in \Psi$ at time $n \in \mathcal{N}$ to state $\psi^{n+1} = (u^{n+1}, e^{n+1}, b^{n+1}) \in \Psi$ at time $n + 1 \in \mathcal{N}$ after taking action $\delta^n \in \Delta$ at time $n \in \mathcal{N}$ is found as follows:

$$\begin{aligned} \mathbb{P}\left(\psi^{n+1}|\psi^n, \delta^n\right) = \ & \mathbb{P}(u^{n+1}|u^n)\mathbb{P}(e^{n+1}) \\ & \times \bigg[ \mathbb{1}(b^{n+1} = b^n + e^n - \delta^n)g(\psi^n, \delta^n) \\ & + \mathbb{1}(b^{n+1} = b^n + e^n)\left(1 - g(\psi^n, \delta^n)\right) \bigg], \quad (6) \end{aligned}$$

where $\mathbb{1}(\cdot)$ is an indicator function which is used to maintain consistent battery levels over time due to energy consumption and depletion. $\mathbb{P}(u^{n+1}|u^n)$ is the probability of transiting between user activities. Furthermore, $\mathbb{P}(e^{n+1})$ is the probability of the mobile device to be in the charging mode.

Recall that the system is also assumed to be pertaining under a data usage constraint $D$. This constraint is important to avoid overcharges by cellular operators for data synchronization to a backend. Therefore, we define $d(\cdot)$ as the data usage function which returns a non-negative value based on the taken actions $d(\cdot) : \Delta \to \mathbb{R}_+$. Mathematically, $d(\cdot)$ is defined as follows:

$$d(\psi, \delta) = \begin{cases} d(\psi, \delta_1), & b > 0 \text{ and } \delta = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

If the battery is not empty, i.e., $b > 0$, and the active action $\delta_1$ is selected, activity data packets are generated and transmitted to the backend. Here, there is an important connection between the cellular data limit $D \in \mathbb{R}_+$ and the data generated during the active mode $d(\psi, \delta_1)$. Specifically, a mobile device transmits activity data to a backend with probability $\xi$ of the total epochs such that

$$\xi = \frac{D}{d(\psi = [u, e, b], \delta_1)}. \quad (8)$$

For example, the mobile device transmits activity data during one fourth of its total decision epochs when $d(\psi = [u, e, b], \delta_1) = 1$ and $D = 0.25$.

Before proceeding further with the problem solution, we define a *decision rule* $\pi_n$ at time epoch $n$ as a mapping between the current system state and the optimal action $\pi_n : \Psi \to \Delta$. A *policy* $\pi \doteq (\pi_1, \pi_2, \ldots, \pi_N)$ is a sequence composition of optimal decision rules through all decision epochs. A policy is called as a *stationary policy* if its decision rules are not changing over time. A major objective of the activity tracking policy is to minimize the overall error of monitoring the user activities by selecting optimal actions $\delta \in \Delta$ over time. Therefore, we denote the optimal, stationary tracking policy as $\pi^*(\psi, \delta)$ which maps state $\psi \in \Psi$ and action $\delta \in \Delta$.

### C. Optimal Activity Tracking Policy

As our system design imposes a constraint on the data usage, the mobile sensing problem is formulated as a CMDP which is expressed as follows:

$$\min_{\pi} \quad \mathcal{J}(\pi) = \lim_{N \to \infty} \sup \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}(c(\psi^n, \delta^n)), \quad (9)$$

$$\text{s.t.} \quad \mathcal{D}(\pi) = \lim_{N \to \infty} \sup \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}(d(\psi^n, \delta^n)) \leq D, \quad (10)$$

where $\psi^n \in \Psi$ and $\delta^n \in \Delta$ are the state and action at time $n$, respectively. $\mathcal{J}(\cdot)$ and $\mathcal{D}(\cdot)$ are the average detection error and data usage under policy $\pi$, respectively. $\mathbb{E}(\cdot)$ is the expectation function, and $\pi^*(\psi, \delta)$ is an optimal activity tracking policy that defines the probability of taking action $\delta$ at state $\psi$. $D \in \mathbb{R}_+$ is the data usage limit for the transmission of user activities to the backend. The objective function in (9) minimizes the detection error subject to a data usage limit given by (10).

It has been shown in [10], [28] that a CMDP model can be solved using linear programming (LP) in polynomial time. Let $\phi(\psi, \delta)$ denote the stationary probability of state $\psi$ and action $\delta$. The mobile sensing problem can be formulated as follows:

$$\min_{\phi(\psi,\delta)} \quad \sum_{\psi \in \Psi} \sum_{\delta \in \Delta} \phi(\psi, \delta) c(\psi, \delta), \quad (11)$$

$$\text{s.t.} \quad \sum_{\psi \in \Psi} \sum_{\delta \in \Delta} \phi(\psi, \delta) d(\psi, \delta) \leq D, \quad (12)$$

$$\sum_{\delta \in \Delta} \phi(\psi', \delta) = \sum_{\psi \in \Psi} \sum_{\delta \in \Delta} \phi(\psi, \delta) \mathbb{P}(\psi'|\psi, \delta), (13)$$

$$\sum_{\psi \in \Psi} \sum_{\delta \in \Delta} \phi(\psi, \delta) = 1, \phi(\psi, \delta) \geq 0, \quad (14)$$

where $\psi' \in \Psi$. The solution of this problem is the optimal stationary probability $\phi^*(\psi, \delta)$. The objective function in (11) minimizes the activity detection error. The constraint in (12) maintains the cellular data usage below a target level $D$. Then, the constraint in (13) ensures the ergodic transition between system states. The constraints in (14) assert on stationary probability requirements. Solving (11)-(14) using an LP solver gives the optimal stationary probability $\phi^*(\psi, \delta)$. The optimal policy is then found for each state and action pair as $\pi^*_{\text{CMDP}}(\psi, \delta) = \frac{\phi^*(\psi,\delta)}{\sum_{\delta' \in \Delta} \phi^*(\psi,\delta')}$. $\pi^*_{\text{CMDP}}$ is a randomized, stationary policy [10], i.e., $\pi^*_{\text{CMDP}}$ is randomized over available actions and does not change over time.
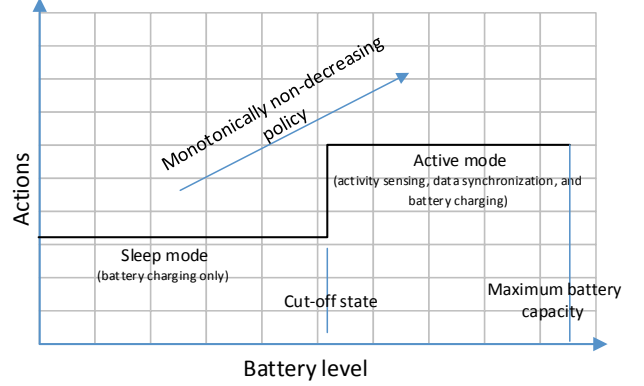


Fig. 2: Threshold activity tracking policy with two actions corresponding to the sleep and active modes. The sleep mode is preferred at low battery levels while data sensing and synchronization are performed at high levels.

### D. Threshold Activity Tracking Policy

Our optimization problem is intentionally designed to derive a threshold activity tracking policy. A *threshold activity tracking policy*, which is an MDP solution that follows a monotone pattern with system states, facilitates problem solution and implementation [31]. In this paper, the optimal policy of the MDP is shown to be a threshold in the battery level. Figure 2 shows the desired structure of the activity tracking policy. A *cut-off state* of a threshold policy is the battery level beyond which the selected action is increased to a new value. This is important due to the following benefits:

- *Fast adaptation of online policy learning*: Solution methods for computing an optimal activity tracking policy can be customized to explore the threshold structure of the intended policy. This is important as activity statistics are unknown a priori, varying with time, and inconsistent for different users. Specifically, this customization significantly improves the convergence of conventional solution methods [32], [33].
- *Low memory and communication overheads*: Saving the threshold policy into the mobile device's memory in a compact form is significantly efficient, as threshold cut-off states are sufficient for policy execution. This requires low memory footprint compared with unstructured policies which are saved using look-up tables with state-action pairs. Similarly, the overhead in transferring the learned policy between the system components is also minimized by only sending the threshold cut-off states.
- *Simple implementation*: The threshold policy helps in developing simple and lightweight algorithms. Clearly, the selection of optimal online actions can be done by comparing the system state with the cut-off value, e.g., using a simple if-then-else statement, and no look-up search is needed.

## IV. UNCONSTRAINED ACTIVITY TRACKING POLICY: A LAGRANGE RELAXATION APPROACH

In this section, the CMDP formulation is transformed into its unconstrained MDP form using a Lagrange multiplier.

Firstly, the unconstrained MDP is required to prove the threshold structure of the CMDP policy. Particularly, utilizing the threshold structure of the CMDP for the mobile sensing in (11)-(14) is complex due to the data usage term. Secondly, the complexity of the value iteration algorithm is lower than that of the algorithm to solve the LP problem [34]. Therefore, the CMDP formulation must be first transformed into an unconstrained MDP. We adopt the transformation approach that relies on using the Lagrange multiplier algorithm [11]. The Lagrangian relaxation approach introduces a Lagrange multiplier $\lambda$ and the resulting *Lagrangian error function* is defined as follows:

$$c(\psi, \delta; \lambda) = c(\psi, \delta) + \lambda d(\psi, \delta), \tag{15}$$

where $\lambda > 0$. Accordingly, the *Lagrangian average error* $\mathcal{J}(\pi; \lambda)$ is given by

$$\mathcal{J}(\pi; \lambda) = \lim_{N \to \infty} \sup \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}(c(\psi^n, \delta^n) + \lambda d(\psi^n, \delta^n)). \tag{16}$$

This Lagrangian average error enables the solution of the CDMP problem using any of the conventional MDP solution methods such as the value iteration or policy iteration algorithms. Thus, the MDP activity tracking policy is found by minimizing (16) such that

$$\pi^*_{\text{MDP}}(\psi) = \arg\inf \mathcal{J}(\pi; \lambda). \tag{17}$$

In the following, we discuss solution methods of the unconstrained MDP in (17). In particular, the unconstrained problem is formulated as an LP using which the optimal Lagrange multiplier $\lambda^*$ is selected. Then, given $\lambda^*$, the problem is solved using the value iteration algorithm.

### A. Solution Using Linear Programming

It is important to note that $\pi^*_{\text{MDP}}$ can still be solved using LP solvers as in (11)-(14) after dropping (12) and replacing $c(\psi, \delta)$ in (11) with $c(\psi, \delta; \lambda)$. The resulting LP problem is as follows:

$$\min_{\phi(\psi, \delta)} \quad \sum_{\psi \in \Psi} \sum_{\delta \in \Delta} \phi(\psi, \delta) c(\psi, \delta; \lambda), \tag{18}$$

$$\text{s.t.} \quad \sum_{\delta \in \Delta} \phi(\psi', \delta) = \sum_{\psi \in \Psi} \sum_{\delta \in \Delta} \phi(\psi, \delta) \mathbb{P}(\psi'|\psi, \delta), \tag{19}$$

$$\sum_{\psi \in \Psi} \sum_{\delta \in \Delta} \phi(\psi, \delta) = 1, \phi(\psi, \delta) \geq 0, \tag{20}$$

where $\psi' \in \Psi$. A special attention should be given to the selection of the Lagrange multiplier $\lambda$ to ensure that the resulting unconstrained activity tracking policy is an accurate transformation of the optimal CMDP activity tracking policy. Therefore, finding an optimal Lagrange value $\lambda^*$ is discussed next.

### B. Finding $\lambda^*$

Clearly, there is a strong connection between the chosen value of the Lagrange multiplier $\lambda$ and the cellular data limit $D$. Explicitly, $\lambda^*$ is found as follows [11], [12]:

$$\lambda^* = \inf \{\lambda : \mathcal{D}(\pi^*; \lambda) \leq D\}. \tag{21}$$

---

**Algorithm 1** Lagrange multiplier estimation of the unconstrained activity tracking policy.

---

**Input:** $\Psi, \Delta, \mathbb{P}, c(\cdot), d(\cdot)$
**Output:** $\lambda^*$
1: Initialize an iteration counter $i = 0$
2: Initialize $\lambda_0$ to a random number greater than $0$
3: Solve (18)-(20) using LP to find $\phi^*_{\lambda_i}(\psi, \delta)$, $\psi \in \Psi$ and $\delta \in \Delta$
4: Find current data usage $\mathcal{D}(\pi^*; \lambda_i) = \sum_{\psi \in \Psi} \sum_{\delta \in \Delta} \phi^*_{\lambda_i}(\psi, \delta) d(\psi, \delta)$
5: Update $\lambda_{i+1} = \lambda_i + \frac{1}{\sqrt{i+1}} (\mathcal{D}(\pi^*; \lambda_i) - D)$
6: if $|\lambda_{i+1} - \lambda_i| < \varepsilon$, terminate and go to Step 8
7: Increment counter $i = i + 1$ and go to Step 3
8: **return** $\lambda^* = \min_{0 \leq j \leq i+1} \{\lambda_j : \mathcal{D}(\pi^*; \lambda_j) \leq D\}$

---

An optimal value of $\lambda^*$ can be obtained by iterative methods as shown in Algorithm 1 such that the policy always meets the cellular data limit $D$. At each iteration $i$ of Algorithm 1, Step 3 solves the LP problem given in (18)-(20) based on the Lagrange multiplier estimation $\lambda_i$. This solution is used to find the corresponding data usage $\mathcal{D}(\pi^*; \lambda)$ as in Step 4 which updates the Lagrange multiplier in Step 5. This update rule in Step 5 is based on the Robbins–Monro algorithm for stochastic approximation which has been used in previous studies for the Lagrange multiplier estimation [32], [35]. The algorithm terminates when the difference between two estimations of the Lagrange multiplier is below a small error value $\epsilon$.

### C. Discounted Cost Solutions

$\pi^*_{\text{MDP}}$ can also be found using discounted cost, infinite horizon MDP solution methods [31]. Based on the previous unconstrained formulation, the *expected total discounted error* for a policy $\pi$ and a discount factor $\beta$, where $0 \leq \beta < 1$, can be expressed as follows:

$$J(\pi; \lambda, \beta) = \lim_{N \to \infty} \sup \left( \mathbb{E} \left[ \sum_{n=1}^{N} \beta^{n-1} c(\psi, \delta; \lambda) \right] \right), \tag{22}$$

which can be solved using the value iteration algorithm. Accordingly, the optimal value function $v(\psi; \lambda, \beta)$ for each state $\psi \in \Psi$ is

$$v(\psi; \lambda, \beta) = \min_{\delta \in \Delta} \left\{ c(\psi, \delta; \lambda) + \beta \sum_{\psi' \in \Psi} \mathbb{P}(\psi'|\psi, \delta) v(\psi'; \lambda, \beta) \right\}, \tag{23}$$

which is the solution of the Bellman equation with stationary policy $\pi^*_{\text{MDP}}$ defined as follows:

$$\pi^*_{\text{MDP}} = \arg\min_{\delta \in \Delta} \left\{ c(\psi, \delta; \lambda) + \beta \sum_{\psi' \in \Psi} \mathbb{P}(\psi'|\psi, \delta) v(\psi'; \lambda, \beta) \right\}. \tag{24}$$

The Bellman equation can be recursively solved using the value iteration algorithm. In particular, the *value function* $v(\psi; \lambda, \beta)$ and the *state-action cost function* (or called the Q-function) $Q(\psi, \delta; \lambda, \beta)$ are arbitrarily initialized and then updated in each iteration of the value iteration algorithm as follows:

$$v^{i+1}(\psi; \lambda, \beta) =$$
$$\min_{\delta \in \triangle} \left\{ c(\psi, \delta; \lambda) + \beta \sum_{\psi' \in \Psi} \mathbb{P}(\psi'|\psi, \delta) v^i(\psi'; \lambda, \beta) \right\}, \quad (25)$$

$$Q^{i+1}(\psi, \delta; \lambda, \beta) =$$
$$c(\psi, \delta; \lambda) + \beta \sum_{\psi' \in \Psi} \mathbb{P}(\psi'|\psi, \delta) v^i(\psi'; \lambda, \beta), \quad (26)$$

where $i \in \{0, 1, 2, \ldots\}$ is an iteration counter of the algorithm, $v^i(\psi; \lambda, \beta)$ is the minimum achievable cost value for state $\psi$ at iteration $i$, and $Q^i(\psi, \delta; \lambda, \beta)$ is the minimum cost value for taking action $\delta$ during state $\psi$ at time $i$. Here, it is important to note that the value function is obtained for each state, while the state-action cost function is found for each state-action pair. After convergence, the policy $\pi_{\text{MDP}}^*$ is deterministic and stationary [31], [36]. Conversely, $\pi_{\text{CMDP}}^*$ is a randomized, stationary policy [10]. Thereby, a key note is that $\pi_{\text{CMDP}}^*$ with the average cost is not directly estimated by $\pi_{\text{MDP}}^*$ that uses the discounted cost. Instead, $\pi_{\text{CMDP}}^*$ will be shown in the next section to be a randomized mixture of two perturbed policies of $\pi_{\text{MDP}}^*$.

To this end, the constrained MDP activity tracking policy was transformed as a discounted cost and unconstrained MDP by introducing a Lagrange error function given in (15). The unconstrained transformation was solved using conventional method, such as the value iteration algorithm, which meets the Bellman equation (23). In the following section, the threshold structure of the activity tracking policy is analytically proved by using the discounted cost MDP formulation and the concept of submodularity.

## V. RANDOMIZED THRESHOLD ACTIVITY TRACKING POLICY: MONOTONICITY ANALYSIS

This section proves the monotone, threshold structure of the proposed activity tracking policy. In summary, this section includes the following contributions:

- The concept of submodularity [37] is used to prove the monotone structure of the unconstrained Lagrange formulation of the activity tracking policy. This analysis is based on the discounted cost MDPs in (23). This proof requires two major steps: (a) the value function $v(\psi; \lambda, \beta)$ must be monotone, and (b) the state-action cost function $Q(\psi, \delta; \lambda, \beta)$ must be submodular.
- After proving the threshold structure of the unconstrained MDPs, the CMDP policy is shown to be a mixture of two threshold MDP policies. Accordingly, this proves the threshold structure for the optimal activity tracking policy.

- A threshold Q-learning method that exploits the threshold structure of the activity tracking policy is presented. This method enables a fast convergence compared with the conventional Q-learning method.

### A. Threshold Structure of the Unconstrained Policy

To prove the threshold structure, this section follows the same approach as in [35], [38]. Firstly, the concept of submodularity will be defined to prove the monotone structure of a policy as submodularity is a sufficient condition for proving threshold policy structure [31]. In summary, our main objective is to prove that the state-action cost function $Q(\psi = [u, e, b], \delta; \lambda, \beta)$ is submodular in $(b \in \mathcal{B}, \delta \in \triangle)$, and hence the optimal activity tracking policy is monotonically non-decreasing in $b$.

**Definition 1.** For any two sets $\mathcal{X} \subseteq \mathbb{R}$ and $\mathcal{Y} \subseteq \mathbb{R}$, a function $f(\cdot)$ that is defined as $f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ is called *submodular* in $(x \in \mathcal{X}, y \in \mathcal{Y})$ if the inequality condition $f(x_1, y_1) + f(x_2, y_2) \leq f(x_1, y_2) + f(x_2, y_1)$ holds for all $x_1 \geq x_2, y_1 \geq y_2, x_1, x_2 \in \mathcal{X}$, and $y_1, y_2 \in \mathcal{Y}$.

This definition is important as the submodularity of $f(\cdot)$ is a sufficient condition for the non-decreasing monotonicity of $y = \arg\min f(x, y)$ [37].

**Lemma 2.** *For a given optimal Lagrange multiplier $\lambda^* > 0$ and a discount factor $\beta \in [0, 1)$, the optimal value function $v(\psi; \lambda, \beta)$ of the mobile sensing is monotonically non-decreasing in the battery level $b$.*

*Proof:* See the Appendix. ∎

**Theorem 3.** *For a given optimal Lagrange multiplier $\lambda^* > 0$ and a discount factor $\beta \in [0, 1)$, the optimal discounted cost MDP policy $\pi_{MDP}^*$ of the mobile sensing is monotone and does not decrease as the battery level $b$ increases.*

*Proof:* See the Appendix. ∎

Based on Theorem 3 and for a given threshold value $b_{\text{cut}} \in \mathcal{B}$, the optimal discount policy $\pi_{\text{MDP}}^*$ can be written in the *compact form* as follows:

$$\pi_{\text{MDP}}^*(\psi = [u, e, b]) = \begin{cases} 0, & 0 < b \leq b_{\text{cut}}, \\ 1, & b_{\text{cut}} \leq b \leq B. \end{cases} \quad (27)$$

This compact form of the threshold policy facilitates the development of the activity tracking policy as discussed in Section III. $\pi_{\text{MDP}}^*(\psi)$ is called a *binary threshold policy* as it only selects between two possible actions $\delta_0$ and $\delta_1$. The intuition of the non-decreasing structure is that when the battery level is higher, the mobile device will take the activation action due to the lower cost.

### B. Threshold Structure of the Constrained Policy

The threshold structure of the discounted cost MDP problem is proved as in Theorem 3. Correspondingly, the infinite horizon average cost CMDP problem with a single constraint has an optimal policy $\pi_{\text{CMDP}}^*$ that is a mixture of two threshold MDP policies in the form [11], [12] presented as follows:

$$\pi_{\text{CMDP}}^* = \gamma \pi_{\text{MDP}}^+ + (1 - \gamma) \pi_{\text{MDP}}^-, \quad (28)$$

where $\pi_{\text{MDP}}^{+}$ and $\pi_{\text{MDP}}^{-}$ are the stationary discounted cost MDP policies for perturbed values of $\lambda^{+} = \lambda^{*} + \Delta\lambda$ and $\lambda^{-} = \lambda^{*} - \Delta\lambda$, where $\Delta\lambda$ is the perturbation value of $\lambda$. $\gamma \in [0,1]$ is the probability of selecting $\pi_{\text{MDP}}^{+}$ and $1 - \gamma$ is for selecting $\pi_{\text{MDP}}^{-}$. Even though the MDP policy is deterministic and stationary, (28) is important as it enables the randomized selection of optimal actions in a randomized manner, i.e., (28) stochastically selects actions as a CMDP policy. Here, the probability $\gamma$ can be calculated using the rule $\gamma = \frac{\mathcal{D}(\pi_{\text{MDP}}^{-}) - D}{\mathcal{D}(\pi_{\text{MDP}}^{-}) - \mathcal{D}(\pi_{\text{MDP}}^{+})}$.

### C. Fast Adaptation of Q-Learning

In online activity tracking, the transition probabilities between user activities could be (i) unknown at design time, (ii) changing over time for the same user, or (iii) distinct for different users. Therefore, an online algorithm that can adapt with the changing parameters is required.

The Q-learning algorithm [39] is widely considered as one of the most powerful *model-free* methods of reinforcement learning. The gradual learning feature of the Q-learning allows the policy formulation of an activity-aware mobile device when the transition matrix $\mathbb{P}(\psi'|\psi, \delta)$ is unknown. Additionally, this enables the customization of the activity tracking policy to match each user's temporal behavior. Q-learning estimates (26) using stochastic approximation by starting with a random (or zero) approximations of each state-action cost $\{Q^0(\psi, \delta; \lambda, \beta) : \forall \psi \in \Psi$ and $\delta \in \Delta\}$. Then, the conventional Q-learning update rule can be expressed as follows:

$$Q^{i+1}(\psi, \delta; \lambda, \beta) = Q^i(\psi, \delta; \lambda, \beta) + \frac{1}{\sqrt{i+1}}$$
$$\times \left[ c(\psi, \delta; \lambda) + \beta \min_{\delta' \in \Delta} Q^i(\psi', \delta'; \lambda, \beta) - Q^i(\psi, \delta; \lambda, \beta) \right],$$
(29)

where $\beta \in [0,1)$ is a discount factor, and $i \in \{1, 2, \dots, L\}$ is an iteration counter that is limited to an upper bound $L$, e.g., $L = 10^5$ iterations. (29) is a greedy update rule that the lowest-cost action in the next state is used to update the state-action cost factor $Q^{i+1}(\psi, \delta; \lambda, \beta)$ of the current state.

A known limitation of the conventional Q-learning algorithm is the long sequence of iterations required in practical applications [32], [33]. Recall that the monotonically non-decreasing structure of the activity tracking policy was proven in Theorem 3. This structure enables a faster convergence of the Q-learning algorithm by (a) initializing the state-action cost factor of all states such that $Q^0(\psi = [u, e, b], \delta; \lambda, \beta) < Q^0(\psi = [u, e, b+1], \delta; \lambda, \beta)$ for all states $\psi \in \Psi$ and action $\delta \in \Delta$ (see [33]), and (b) projecting the final policy such that it preserves the threshold structure. These steps limit the search of the optimal policy to a subset of all possible solutions, and hence avoid the brute-force search in conventional Q-learning.

## VI. NUMERICAL RESULTS AND DISCUSSION

This section presents numerical analysis of the optimal activity tracking policy. Firstly, parameter settings using a real-
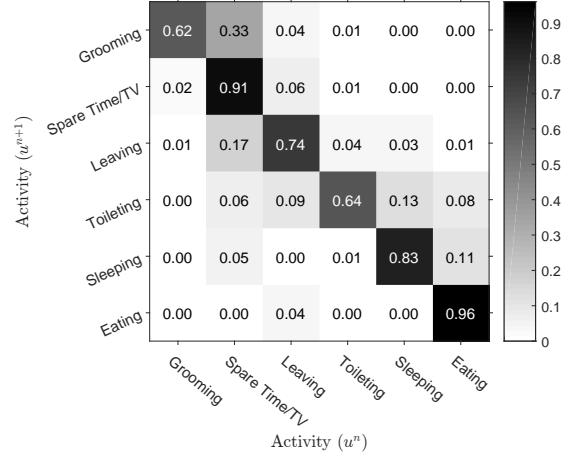


Fig. 3: Transition probabilities $\mathbb{P}_{\text{user}}$ between user activities.

world dataset are presented. Then, the threshold structure validation results are summarized. Finally, performance measures and performance evaluation of the CMDP policy is given.

### A. Parameter Setting

We run experiments while using a real-world dataset [40] to extract the system parameters. Specifically, we consider an activity tracking scenario of six activities in daily home routines such that $\mathcal{U}$ is defined as follows:

$$\mathcal{U} = \big\{ 0 : \text{grooming}, 1 : \text{spare time/watching TV},$$
$$2 : \text{leaving}, 3 : \text{sleeping}, 4 : \text{toileting/showering}, 5 : \text{eating} \big\}.$$
(30)

Unless otherwise stated, the battery charging can be in one of two modes: (i) a charging mode when the mobile device is connected to a wired or wireless charger with a probability of $\mathbb{P}(e = 1) = 0.15$, and (ii) no-charging mode with a probability of $\mathbb{P}(e = 0) = (1 - \mathbb{P}(e = 1)) = 0.85$. For convenience, we denote the user transition probability matrix as $\mathbb{P}_{\text{user}} = [P(u'|u), \forall u \in \mathcal{U}$ and $u' \in \mathcal{U}]$ which is given in Figure 3. The immediate detection error in (3) is defined such that $c(\psi = [u = 0, e, b], \delta_1) = 0.28$, $c(\psi = [u = 1, e, b], \delta_1) = 0.25$, $c(\psi = [u = 2, e, b], \delta_1) = 0.18$, $c(\psi = [u = 3, e, b], \delta_1) = 0.12$, $c(\psi = [u = 4, e, b], \delta_1) = 0.1$, and $c(\psi = [u = 5, e, b], \delta_1) = 0.08$. The user activities are not tracked during the sleep mode such that $c(\psi, \delta_0) = 1$. The data usage function is defined as follows:

$$d(\psi = [u, e, b], \delta) = \begin{cases} 1, & \delta = 1 \text{ and } b > 0, \\ 0, & \text{otherwise.} \end{cases}$$

This indicates that one data packet is generated during the active mode. The connectivity probabilities in (4) are $g(\psi = [u = 0, e, b], \delta) = 0.5\delta$, $g(\psi = [u = 1, e, b], \delta) = 0.55\delta$, $g(\psi = [u = 2, e, b], \delta) = 0.6\delta$, $g(\psi = [u = 3, e, b], \delta) = 0.65\delta$, $g(\psi = [u = 4, e, b], \delta) = 0.68\delta$, and $g(\psi = [u = 5, e, b], \delta) = 0.7\delta$. Clearly, data transmission cannot be performed during the sleep mode $g(\psi = [u, e, b], \delta_0) = 0$. Finally,
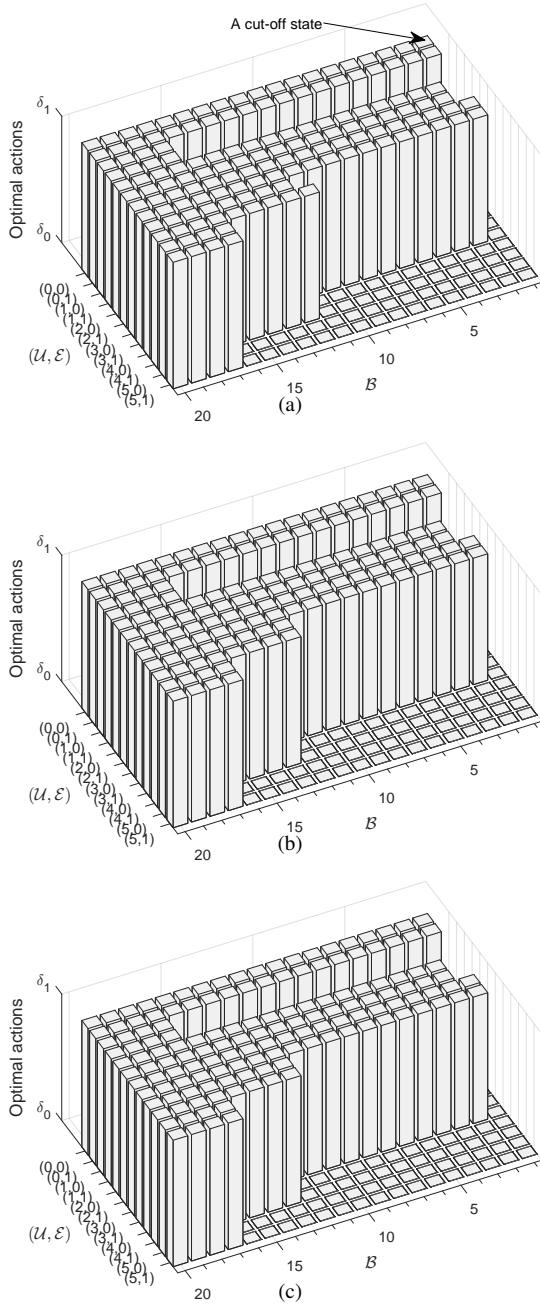
Fig. 4: Threshold policies for an activity-aware mobile device. Recall that $\mathcal{U}, \mathcal{E}$, and $\mathcal{B}$ are the user activity, energy charging, and battery level state spaces of the mobile device, respectively. (a) The constrained MDP solution to the mobile sensing problem using LP, (b) the unconstrained MDP solution using LP with $\lambda = 0.128$, and (c) the unconstrained MDP solution using value iteration with $\lambda = 0.128$.

the discount factor in the discounted cost MDP formulation is $\beta = 0.99$ to ensure high precision solutions.

### B. Threshold Structure

We first analyze the threshold structure of the activity tracking policy. The feasible battery levels are $\mathcal{B} = \{0, \ldots, 20\}$, and the cellular data limit is fixed as $D = 0.25$. This indicates that

the system senses the user activities in one fourth of the total decision epochs $\mathcal{N}$. We use the following steps to obtain the optimal CMDP policy and its unconstrained MDP estimation: (i) the problem is solved using the CMDP formulation in (11)-(14) and LP, (ii) the CMDP problem is then transformed into the unconstrained MDP form given in (18)-(20), (iii) the optimal Lagrange multiplier value $\lambda^* = 0.128$ is found using Algorithm 1, and (iv) the unconstrained problem is solved using the value and policy iteration algorithms. Figure 4 shows the resulting policies of the constrained and unconstrained MDP formulations. Two important results from Figure 4 can be highlighted as follows:

1) The optimal activity tracking policy has a threshold structure. In particular, the policy is a threshold policy and is monotonically non-decreasing in the battery level $b$. This observation represents the outcome from Theorem 3. In this case, the optimal actions taken by the mobile device change from $\delta_0$ to $\delta_1$ as the battery level increases. The cut-off states are the only data required by the mobile device for selecting the optimal actions.

2) The discounted cost MDP solution using (28) provides an accurate transformation of the CMDP problem. This simulation result is consistent with the theoretical analysis in Sections III and V, where the CMDP policy is a randomized mixture of the discounted cost MDP policies.

### C. The Impact of Setting the Lagrange Multiplier

Figure 5 shows the Lagrange multiplier $\lambda$ updates over iterations of Algorithm 1 with $\epsilon = 10^{-4}$ and $D = 0.25$. Based on this experiment, the optimal value is found as $\lambda^* = 0.128$ which satisfies the optimality condition in (21). The data usage is found at each iteration as $\mathcal{D}(\pi^*; \lambda_i) = \sum_{\psi \in \Psi} \sum_{\delta \in \Delta} \phi^*_{\lambda_i}(\psi, \delta) d(\psi, \delta)$.

The Lagrange optimal value selection is critical for the accuracy of the unconstrained discounted cost MDP solutions. Figure 6 shows the policy when the Lagrangian multiplier is incorrectly set as $\lambda = 0.25$. Intuitively, an incorrect value of the Lagrange multiplier results in a poor estimation of the CMDP optimal policy. This because the Lagrangian error function defined in (15) does not accurately capture the data usage constraint of the CMDP activity tracking policy given in (11)-(14).

### D. Fast Adaptation of Q-learning

The key objective of proving the monotone threshold structure of the activity tracking policy in a mobile device is for adopting low-complexity estimation methods as discussed in Section V-C. Figure 7 shows the online policy learning by applying the Q-learning algorithm. In particular, Figure 7a shows the policy which is generated after $10^6$ update iterations of the conventional Q-learning method. Clearly, the conventional Q-learning does not consider the monotone structure of the policy and it initializes the state-action function to random or zero values. Then, it searches through the whole solution space, i.e., a brute-force search. This causes the poor performance of the
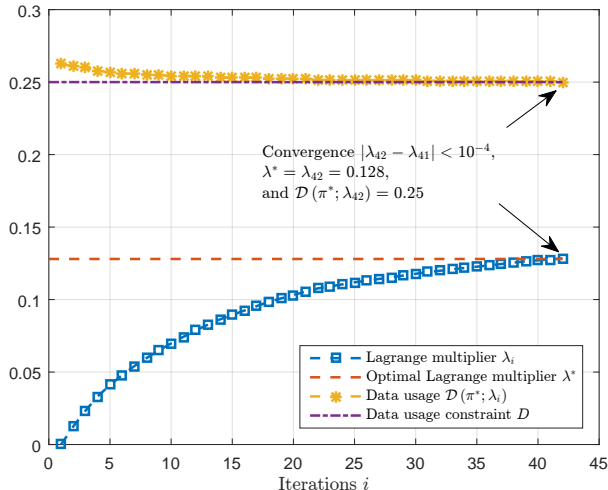
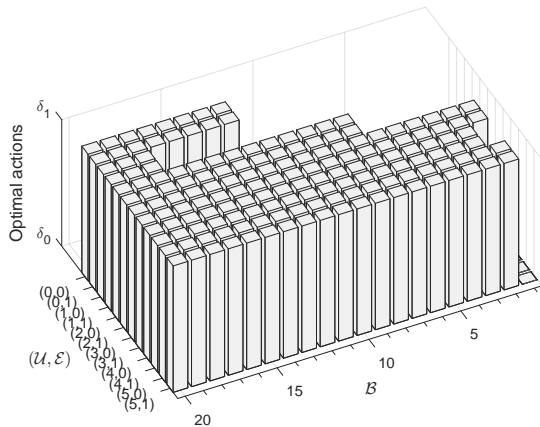Fig. 5: Setting the optimal Lagrange multiplier $\lambda^*$ using Algorithm 1.



Fig. 6: The unconstrained MDP solution using value iteration with an incorrect Lagrange multiplier of $\lambda = 0.25$ resulting in a poor estimation solution.

conventional method. By contrast, the structured Q-learning algorithm starts with an initialization that considers the actual monotonicity of the state-action function and projects the final policy into a threshold form. Consequently, this significantly improves the performance as shown in Figure 7b.

### E. Performance Metrics

We consider three performance metrics of online activity tracking systems under a data usage limit.

*1) Average Battery Level:* This measure is important to ensure that the mobile device has enough energy when performing the activity tracking task. The average battery level $\bar{b}$ during the active mode $\delta_1$ can be obtained as follows:

$$\bar{b} = \sum_{u \in \mathcal{U}} \sum_{e \in \mathcal{E}} \sum_{b \in \mathcal{B}} b \phi^*(\psi = [u, e, b], \delta_1). \tag{31}$$

*2) Probability of Successful Data Synchronization:* Successful data synchronization to a backend server requires

(i) the selection of the activation action, i.e., $\delta_1$, and (ii) the availability of access network connection determined by $g(\psi, \delta)$. Therefore, we define the probability of successful data synchronization $\rho$ as follows:

$$\rho = \sum_{u \in \mathcal{U}} \sum_{e \in \mathcal{E}} \sum_{b \in \mathcal{B}} g(\psi, \delta_1) \phi^*(\psi = [u, e, b], \delta_1). \tag{32}$$

*3) Probability of Battery Overflow:* Assuming that the mobile device is only used for activity tracking, this metric measures the probability of overcharging the battery of the mobile device by adding an energy unit to a fully charged battery, i.e., the probability of wasting energy. This can happen in two cases: (i) charging a full battery during the sleep mode, or (ii) charging a full battery when no access network coverage is available during the active mode. Then, the probability of battery overflow $\tau$ is given by

$$\tau = \sum_{u \in \mathcal{U}} \phi^*(\psi = [u, 1, B-1], \delta_1) [1 - g(\psi, \delta_1)]$$
$$+ \phi^*(\psi = [u, 1, B-1], \delta_0). \tag{33}$$

### F. Performance Evaluations

In this section, the optimal CMDP policy is compared with a baseline activity tracking policy. We consider the baseline policy that also guarantees monetary access cost through a cellular data limit $D$, and hence it achieves the data transmission probability $\xi$ of the total epochs as in (8). We propose a *constrained uniform policy (CUP)* as a baseline method, such that the mobile device is activated based on a fixed stationary probability and $\phi^*_{\text{CUP}}(\psi, \delta)$ is uniform for all states $\psi \in \Psi$ and action $\delta \in \Delta$. Recall that the policy should sample the user activity with a probability $\xi$ of its total decision epochs. Thereby, the uniform probability is simply found as follows:

$$\phi^*_{\text{CUP}}(\psi, \delta) = \begin{cases} \frac{1-\xi}{U \times B \times 2}, & \delta = \delta_0, \\ \frac{\xi}{U \times B \times 2}, & \delta = \delta_1. \end{cases} \tag{34}$$

where the constraints $\sum_{\psi \in \Psi_{\text{CUP}}} \sum_{\delta \in \Delta} \phi_{\text{CUP}}(\psi, \delta) = 1$ and $\phi(\psi, \delta) \geq 0$ are satisfied. In the following, we vary the system parameters and observe their impact on the performance of the data usage-constrained policies.

*1) Detection Error:* Figure 8 shows the performance of the optimal CMDP policy and the constrained uniform policy when the data usage limit $D$, capacity of storage battery $B$, and charging probability $\mathbb{P}(e = 1)$ are varied. Several important results can be observed. Firstly, as $D$ becomes more relaxed, the optimal policy senses the user activity more frequently which decreases the detecting error $\mathcal{J}(\pi)$. This indicates that if a user decides to set a low data usage setup, the system tracking of that particular user will be poor. Secondly, the detection error will be slightly decreased when the capacity of storage battery $B$ is increased. This is intuitive as the extra battery storage helps in decreasing the battery overflow during the charging process. Thirdly, when the charging probability is high, the detection error is low due to the increased energy budget of the mobile device. It is important to note that the charging probability depends on the number of charger
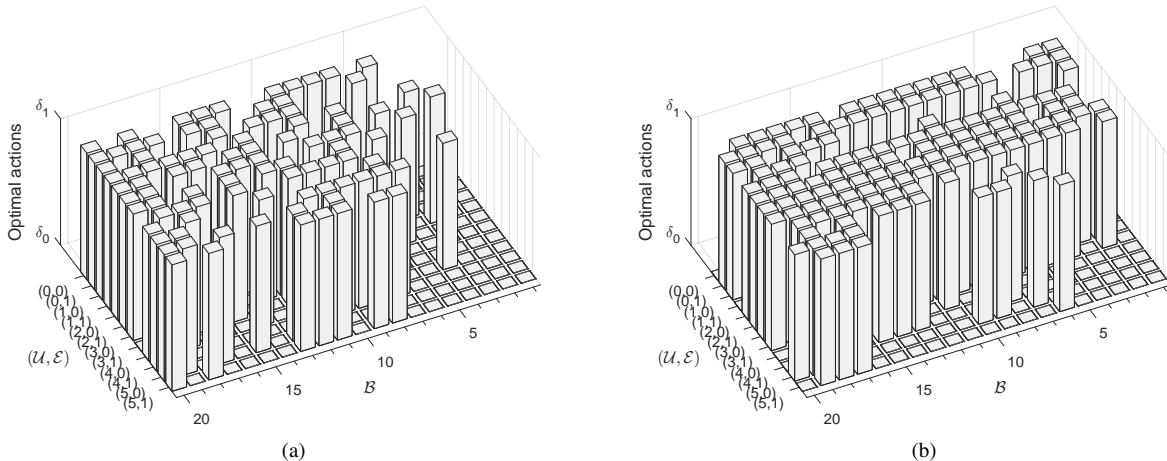
(a)



(b)

Fig. 7: Online learning of the optimal activity tracking policy using $10^6$ gradual iterations. (a) The poor estimation using conventional Q-learning, and (b) a better estimation using the structured Q-learning algorithm.

deployed in the movement locations. It can be observed that the optimal policy outperforms the constrained uniform policy in all scenarios.

Figure 9 shows the average detection error for each activity under varied data usage limit $D$. It can be noted that the detection error of each activity decreases as the data usage limit $D$ is increased. This is due to the increased rate of synchronization probability as defined in (8). However, it can be noted that the average detection error does not uniformly decrease for all activities. Instead, the optimal policy defines the optimal activation based on the detection error and transition probabilities of different activities with the objective of minimizing the total detection error of the tracking system.

*2) Maximum Capacity of Battery:* Figure 10 shows the performance of optimal CMDP policy and the constrained uniform policy when the maximum capacity of storage battery $B$ is varied. When $B$ is high, the average battery level $\bar{b}$ increases as the battery can store more energy units. Likewise, the probability of successful data synchronization $\rho$ slightly increases. The probability of battery overflow $\tau$ slightly decreases. The optimal policy outperforms the constrained uniform policy in all performance metrics.

## VII. Conclusions and Future Work

In this paper, we have presented an activity tracking policy with its threshold structure for mobile devices with intermittent energy charging. We have first modeled the activity and motion tracking as a stochastic optimization using constrained Markov decision processes (CMDPs). The objective is to minimize the detection error of human activities subject to a data usage limit. The CMDP-based problem has been then transformed into an unconstrained, discounted cost MDP with infinite time horizon by using a Lagrange relaxation method. Specifically, a Lagrange multiplier is used to capture the cellular data limit, and therefore ensures the monetary access requirement in the unconstrained activity tracking policy. The CMDP policy has been proved to be a randomized mixture of two threshold MDP policies. Equally important, the CMDP policy has been shown
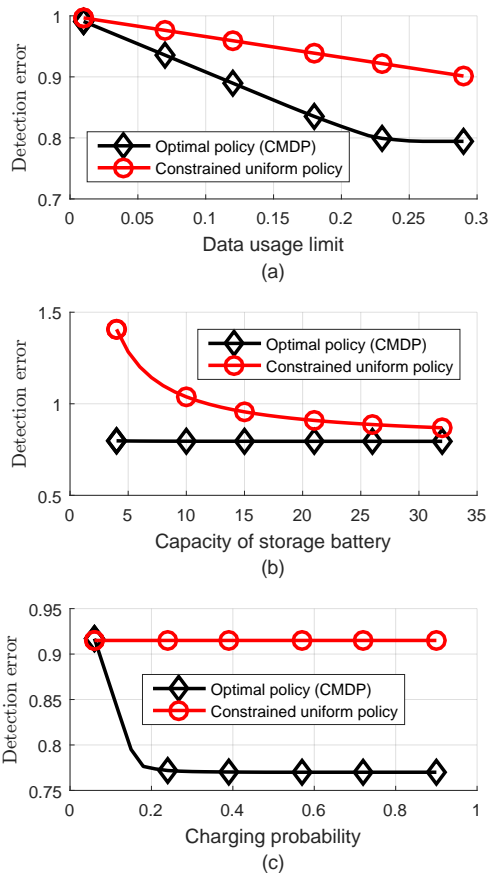


(a)



(b)



(c)

Fig. 8: Detection error $\mathcal{J}(\pi)$ under varied data usage limit $D$, capacity of storage battery $B$, and charging probability $\mathbb{P}(e=1)$.
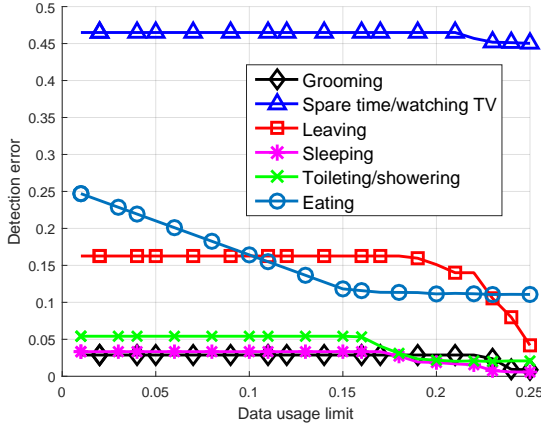
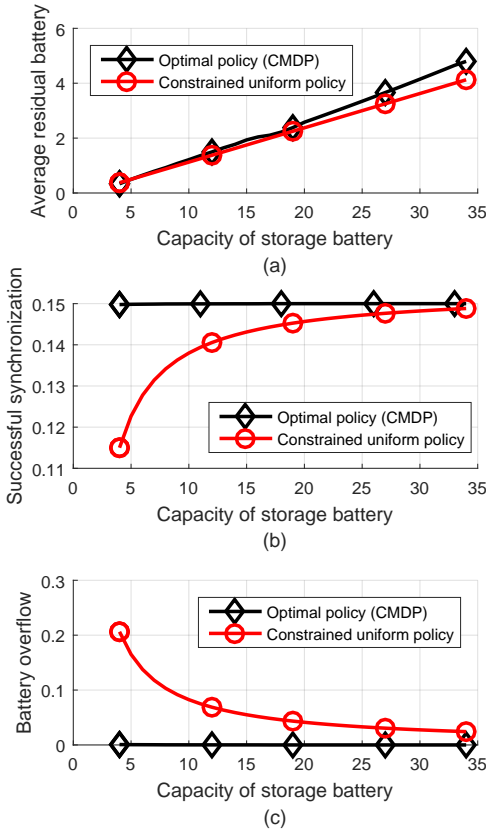Fig. 9: Detection error per activity of the optimal tracking policy under varied data usage limit $D$.



Fig. 10: Average battery level $\bar{b}$ (in energy units), probability of successful data synchronization $\rho$, and probability of battery overflow $\tau$ under varied maximum capacities of battery storage $B$ (in energy units). The data usage limit $D$ is fixed at 0.15.

to be monotonically non-decreasing in the battery level. This monotone threshold structure enables a fast and online learning of the optimal activity tracking policy when the user statistics are unknown a priori, time-varying, and user-defined.

For the future work, dynamic energy pricing can be introduced for wireless charging and the activity tracking policy has to optimize this cost.

## APPENDIX

### A. Proof of Lemma 2

The monotonic structure of $v(\psi; \lambda, \beta)$ can be proved by showing that the two terms of (23) are monotone in the system states. In other words, by proving that (a) the Lagrangian error function is monotone, and (b) the transition probability summation is also monotone. Firstly, $c(\psi, \delta; \lambda)$ is monotonically decreasing in the user activity $u$ and monotonically non-decreasing in battery level $b$. This is clear from the definitions of the cost and data usage functions in (3) and (7), respectively. Secondly, the transition probability summation $\sum_{\psi' \in \Psi} \mathbb{P}(\psi'|\psi, \delta)$ is also monotonically non-decreasing in the battery level $b$ as the transition probabilities are assumed to satisfy the first-order stochastic dominance rule.

### B. Proof of Theorem 3

The optimal discounted cost MDP policy $\pi^*_{\text{MDP}}$ can be shown to be monotone by inductively proving that the state-action cost function $Q(\psi, \delta; \lambda, \beta)$, calculated using (26) for $i \in \{0, 1, 2, \ldots\}$, is a submodular function in $(b, \delta)$. Mathematically, this can by proved by showing that

$$
\begin{aligned}
Q^{i+1}(\psi &= [u, e, b], \delta_1; \lambda, \beta) \\
&- Q^{i+1}(\psi = [u, e, b], \delta_0; \lambda, \beta) \geq \\
Q^{i+1}(\psi &= [u, e, b+1], \delta_1; \lambda, \beta) \\
&- Q^{i+1}(\psi = [u, e, b+1], \delta_0; \lambda, \beta), \quad (35)
\end{aligned}
$$

which indicates that $Q^{i+1}(\psi, \delta; \lambda, \beta)$ is submodular in $(b, \delta)$ for all $i \in \{0, 1, 2, \ldots\}$. Using (26) and (6), the left hand side (LHS) of (35) can be rewritten as follows:

$$
\begin{aligned}
Q^{i+1} &(\psi = [u, e, b], \delta_1; \lambda, \beta) \\
&- Q^{i+1}(\psi = [u, e, b], \delta_0; \lambda, \beta) \\
= c &(\psi = [u, e, b], \delta_1; \lambda) - c(\psi = [u, e, b], \delta_0; \lambda) \\
&+ \beta \sum_{\psi' \in \Psi} \mathbb{P}(u'|u)\mathbb{P}(e')g(\psi, \delta_1) \\
\times \left[ v^i \left( \psi' = [u', e', b'-1]; \lambda, \beta \right) - v^i \left( \psi' = [u', e', b']; \lambda, \beta \right) \right],
\end{aligned}
$$
$$(36)$$

where $c(\psi = [u, e, b], \delta_0; \lambda)$ is equal to zero by the problem setup. Then, the inequality condition in (35) can be proved by showing that $v^i(\psi; \lambda, \beta)$ is non-decreasing in $b$ for all $i \in \{0, 1, 2, \ldots\}$ such that

$$
\begin{aligned}
v^{i+1}([u, e, b+1]; \lambda, \beta) - v^{i+1}([u, e, b]; \lambda, \beta) \geq \\
v^{i+1}([u, e, b]; \lambda, \beta) - v^{i+1}([u, e, b-1]; \lambda, \beta),
\end{aligned}
$$
$$(37)$$

or equivalently

$$v^{i+1}([u,e,b+1];\lambda,\beta) - v^{i+1}([u,e,b];\lambda,\beta)$$
$$- [v^{i+1}([u,e,b];\lambda,\beta) - v^{i+1}([u,e,b-1];\lambda,\beta)] \geq 0. \tag{38}$$

Recall that $v(\psi;\lambda,\beta)$ is defined for states, and $Q(\psi,\delta;\lambda,\beta)$ is defined for state-action pairs. For actions $\delta^0,\delta^1,\delta^2 \in \Delta$ which are selected such that

$$v^{i+1}([u,e,b-1];\lambda,\beta) = Q^{i+1}([u,e,b-1],\delta^0;\lambda,\beta), \tag{39}$$

$$v^{i+1}([u,e,b];\lambda,\beta) = Q^{i+1}([u,e,b],\delta^1;\lambda,\beta), \text{ and} \tag{40}$$

$$v^{i+1}([u,e,b+1];\lambda,\beta) = Q^{i+1}([u,e,b+1],\delta^2;\lambda,\beta), \tag{41}$$

which means that $\delta^0,\delta^1$, and $\delta^2$ are optimal actions at states $[u,e,b-1], [u,e,b]$, and $[u,e,b+1]$, respectively. Then, the right hand side (RHS) of (38) can be expressed as follows:

$$Q^{i+1}([u,e,b+1],\delta^2;\lambda,\beta) - Q^{i+1}([u,e,b],\delta^1;\lambda,\beta)$$
$$-Q^{i+1}([u,e,b],\delta^1;\lambda,\beta) + Q^{i+1}([u,e,b-1],\delta^0;\lambda,\beta). \tag{42}$$

By adding $Q^{i+1}([u,e,b],\delta^2;\lambda,\beta)$ and $Q^{i+1}([u,e,b],\delta^0;\lambda,\beta)$ with their negative values, the expression in (42) becomes

$$\underbrace{Q^{i+1}([u,e,b+1],\delta^2;\lambda,\beta) - Q^{i+1}([u,e,b],\delta^2;\lambda,\beta)}_{\text{Term 1}}$$
$$+\underbrace{Q^{i+1}([u,e,b],\delta^2;\lambda,\beta) - Q^{i+1}([u,e,b],\delta^1;\lambda,\beta)}_{\text{Term 2}}$$
$$+\underbrace{Q^{i+1}([u,e,b],\delta^0;\lambda,\beta) - Q^{i+1}([u,e,b],\delta^1;\lambda,\beta)}_{\text{Term 3}}$$
$$-\underbrace{[Q^{i+1}([u,e,b],\delta^0;\lambda,\beta) - Q^{i+1}([u,e,b-1],\delta^0;\lambda,\beta)]}_{\text{Term 4}}. \tag{43}$$

Terms 2 and 3 are positive in magnitude by the assumptions given in (39)-(41) of optimal actions. In particular, $\delta^1$ is defined as an optimal action at state $[u,e,b]$, and hence it has the minimum state-action value. Moreover, Term 4 is less than Term 1, which can be shown by expanding these terms as in (36) and knowing that $v(\psi;\lambda,\beta)$ is monotonically nondecreasing in the battery level $b$. This proves that the condition in (38) is satisfied, and hence $Q^{i+1}(\psi,\delta;\lambda,\beta)$ is submodular in $(b,\delta)$ for all $i \in \{0,1,2,\ldots\}$. This proves that the optimal discounted cost MDP policy $\pi^*_{\text{MDP}}$ is also monotonically nondecreasing in $(b,\delta)$.

### Acknowledgment

### References

[1] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1192–1209, March 2013.

[2] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context aware computing for the Internet of things: A survey," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 414–454, January 2014.

[3] GSMA Corporation, "The mobile economy 2016," 2016, http://www.gsmamobileeconomy.com.

[4] Y. Wang, J. Lin, M. Annavaram, Q. A. Jacobson, J. Hong, B. Krishnamachari, and N. Sadeh, "A framework of energy efficient mobile sensing for automatic user state recognition," in *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services*. ACM, 2009, pp. 179–192.

[5] H. Lu, J. Yang, Z. Liu, N. D. Lane, T. Choudhury, and A. T. Campbell, "The Jigsaw continuous sensing engine for mobile phone applications," in *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*. ACM, 2010, pp. 71–84.

[6] Y. Wang, B. Krishnamachari, Q. Zhao, and M. Annavaram, "Markov-optimal sensing policy for user state estimation in mobile devices," in *Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks*. ACM, 2010, pp. 268–278.

[7] O. Yurur, C. Liu, C. Perera, M. Chen, X. Liu, and W. Moreno, "Energy-efficient and context-aware smartphone sensor employment," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 9, pp. 4230–4244, September 2014.

[8] Y. Chon, Y. Kim, H. Shin, and H. Cha, "Adaptive duty cycling for place-centric mobility monitoring using zero-cost information in smartphone," *IEEE Transactions on Mobile Computing*, vol. 13, no. 8, pp. 1694–1706, August 2014.

[9] J. Jadidian and D. Katabi, "Magnetic MIMO: How to charge your phone in your pocket," in *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*. ACM, 2014, pp. 495–506.

[10] E. Altman, *Constrained Markov decision processes*. Chapman & Hall/CRC, 1999, vol. 7.

[11] F. J. Beutler and K. W. Ross, "Optimal policies for controlled Markov chains with a constraint," *Journal of Mathematical Analysis and Applications*, vol. 112, no. 1, pp. 236–252, July 1985.

[12] L. I. Sennott, "Constrained average cost Markov decision chains," *Probability in the Engineering and Informational Sciences*, vol. 7, no. 01, pp. 69–83, January 1993.

[13] E. Miluzzo, N. D. Lane, S. B. Eisenman, and A. T. Campbell, "Cenceme–injecting sensing presence into social networking applications," in *Smart Sensing and Context*. Springer, 2007, pp. 1–28.

[14] T. Feng, J. Yang, Z. Yan, E. M. Tapia, and W. Shi, "TIPS: Context-aware implicit user identification using touch screen in uncontrolled environments," in *Proceedings of the 15th Workshop on Mobile Computing Systems and Applications*. ACM, 2014, p. 9.

[15] M. Milošević, M. T. Shrove, and E. Jovanov, "Applications of smartphones for ubiquitous health monitoring and wellbeing management," *Journal of Information Technology and Applications*, vol. 1, no. 1, 2011.

[16] C. Qin, X. Bao, R. R. Choudhury, and S. Nelakuditi, "Tagsense: Leveraging smartphones for automatic image tagging," *IEEE Transactions on Mobile Computing*, vol. 13, no. 1, pp. 61–74, January 2014.

[17] L. Shi, Z. Kabelac, D. Katabi, and D. Perreault, "Wireless power hotspot that charges all of your devices," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 2015, pp. 2–13.

[18] O. Inc, "Wattup," http://www.energous.com/, 2016, online; accessed September 2016.

[19] E. Corp, "Cota wireless power," http://www.ossia.com/cota/, 2016, online; accessed September 2016.

[20] IKEA, "Wireless chargers for wherever you are," http://www.ikea.com/gb/en/products/wireless-charging/, 2016, online; accessed September 2016.

[21] S. Corporation, "Powermat wireless charging: Discover wireless charging for your mobile." http://www.ikea.com/gb/en/products/wireless-charging/, 2016, online; accessed September 2016.

[22] Y. Chon, E. Talipov, H. Shin, and H. Cha, "SmartDC: Mobility prediction-based adaptive duty cycling for everyday location monitoring," *IEEE Transactions on Mobile Computing*, vol. 13, no. 3, pp. 512–525, March 2014.

[23] M. C. Sala, K. Partridge, L. Jacobson *et al.*, "An exploration into activity-informed physical advertising using pest," in *Pervasive Computing*. Springer, 2007, pp. 73–90.

[24] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen, "Activity classification using realistic data from wearable sensors," *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 1, pp. 119–128, 2006.

[25] M. Abu Alsheikh, D. Niyato, S. Lin, H.-P. Tan, and Z. Han, "Mobile big data analytics using deep learning and Apache Spark," *IEEE Network*, vol. 30, no. 3, pp. 22–29, May 2016.

[26] B. D. Ziebart, A. L. Maas, A. K. Dey, and J. A. Bagnell, "Navigate like a cabbie: Probabilistic reasoning from observed context-aware behavior," in *Proceedings of the 10th International Conference on Ubiquitous Computing*. ACM, 2008, pp. 322–331.

[27] O. Yurur, M. Labrador, and W. Moreno, "Adaptive and energy efficient context representation framework in mobile sensing," *IEEE Transactions on Mobile Computing*, vol. 13, no. 8, pp. 1681–1693, August 2014.

[28] D. P. Heyman and M. J. Sobel, *Stochastic models in operations research: Stochastic optimization*. Courier Corporation, 2003, vol. 2.

[29] A. Doufexi, E. Tameh, A. Nix, S. Armour, and A. Molina, "Hotspot wireless LANs to enhance the performance of 3G and beyond cellular networks," *IEEE Communications Magazine*, vol. 41, no. 7, pp. 58–65, 2003.

[30] D. Gross, *Fundamentals of queueing theory*. John Wiley & Sons, 2008.

[31] M. Puterman, *Markov decision processes: Discrete stochastic dynamic programming*. John Wiley & Sons New York, NY, 2005.

[32] D. V. Djonin and V. Krishnamurthy, "Q-learning algorithms for constrained Markov decision processes with randomized monotone policies: Application to MIMO transmission control," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 2170–2181, May 2007.

[33] S. Kunnumkal and H. Topaloglu, "Exploiting the structural properties of the underlying Markov decision problem in the Q-learning algorithm," *INFORMS Journal on Computing*, vol. 20, no. 2, pp. 288–301, 2008.

[34] M. L. Littman, T. L. Dean, and L. P. Kaelbling, "On the complexity of solving markov decision problems," in *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 1995, pp. 394–402.

[35] M. H. Ngo and V. Krishnamurthy, "Monotonicity of constrained optimal transmission policies in correlated fading channels with ARQ," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 438–451, January 2010.

[36] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific Belmont, MA, 2012, vol. 2.

[37] D. M. Topkis, *Supermodularity and complementarity*. Princeton university press, 1998.

[38] D. V. Djonin and V. Krishnamurthy, "MIMO transmission control in fading channels–a constrained Markov decision process formulation with monotone randomized policies," *IEEE Transactions on Signal Processing*, vol. 55, no. 10, pp. 5069–5083, October 2007.

[39] C. J. Watkins and P. Dayan, "Q-learning," *Machine learning*, vol. 8, no. 3-4, pp. 279–292, May 1992.

[40] F. J. Ordóñez, P. de Toledo, and A. Sanchis, "Activity recognition using hybrid generative/discriminative models on home environments using binary sensors," *Sensors*, vol. 13, no. 5, pp. 5460–5477, April 2013.

**Mohammad Abu Alsheikh** (S'14) received his B.Eng. in computer systems engineering from Birzeit University, Palestine, in 2011. Between 2010 and 2012, he was a Software Engineer working on developing robust web services, Ajax-based web components, and smartphone applications. He is currently a Ph.D. candidate in the School of Computer Engineering, Nanyang Technological University, Singapore. His research interests include machine learning in big data analytics, mobile sensing technologies, and sensor-based activity recognition.

**Dusit Niyato** (M'09–SM'15) is currently an Associate Professor in the School of Computer Science and Engineering, at Nanyang Technological University, Singapore. He received B.Eng. from King Mongkuts Institute of Technology Ladkrabang (KMITL), Thailand in 1999 and Ph.D. in Electrical and Computer Engineering from the University of Manitoba, Canada in 2008. His research interests are in the area of energy harvesting for wireless communication, Internet of Things (IoT) and sensor networks.

**Shaowei Lin** received his Ph.D. in Mathematics under Bernd Sturmfels in 2011 from the University of California, Berkeley, where he analyzed singularities in statistical models over large data sets through the lens of modern algebraic geometry. This work was continued at Stanford University in a one-year DARPA postdoctoral collaboration with Andrew Ng's lab to explore mathematical challenges in deep learning. In 2012, he returned to Singapore to join the Institute for Infocomm Research (A*STAR) where he started the Sense-making Group in the Sense and Sense-abilities (S&S) programme. The group focused on exploiting machine learning techniques in sensor networks to create resource-efficient algorithms that exhibit higher-order intelligence. Before joining Singapore University of Technology and Design (SUTD), he oversaw deep science activities in S&S as the Deputy Head for Research.

**Hwee-Pink TAN** (S'00–M'04–SM'14) is currently an Associate Professor of Information Systems (Practice) as well as the Academic Director of the TCS-SMU iCity Lab at the Singapore Management University. Prior to joining SMU in March 2015, he was the SERC Programme Manager of the A*STAR Sense and Sense-abilities Program, Institute for Infocomm Research ($I^2R$), Singapore. Before returning to Singapore to join $I^2R$ in March 2008, he was a Research Fellow at Center for Telecommunications Value-chain Research (CTVR) in the Emerging Networks (EN) strand, led by Dr. Linda Doyle. Between December 2004 and June 2006, he was a post-doctoral researcher at EURANDOM in the research group, Queueing and Performance Analysis (QPA), under the guidance of Prof. Onno Boxma and Dr. Ivo Adan. He also visited the Telecommunications Group at the University of Ferrara from September-October 2004 and was hosted by Prof. Michele Zorzi, who is a Professor of Telecommunications in the Department of Information Engineering, University of Padua, Italy.

**Dong In Kim** (S'89-M'91-SM'02) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 1990. He was a Tenured Professor with the School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada. Since 2007, he has been with Sungkyunkwan University (SKKU), Suwon, Korea, where he is currently a Professor with the College of Information and Communication Engineering. Dr. Kim is a first recipient of the NRF of Korea Engineering Research Center in Wireless Communications for Energy Harvesting Wireless Communications (2014-2021). From 2002 to 2011, he served as an Editor and a Founding Area Editor of Cross-Layer Design and Optimization for the IEEE Transactions on Wireless Communications. From 2008 to 2011, he served as the Co-Editor-in-Chief for the IEEE/KICS Journal of Communications and Networks. He served as the Founding Editor-in-Chief for the IEEE Wireless Communications Letters from 2012 to 2015. From 2001 to 2014, he served as an Editor of Spread Spectrum Transmission and Access for the IEEE Transactions on Communications, and then serving as an Editor-at-Large of Wireless Communication I for the IEEE Transactions on Communications.