Optimal Wireless Energy Charging for Incentivized Content Transfer in Mobile Publish-Subscribe Networks (Extended version with full appendices)

Yang Zhang, *Member, IEEE*, Dusit Niyato, *Senior Member, IEEE*, Ping Wang, *Senior Member, IEEE*, Dong In Kim, *Senior Member, IEEE*, and Zhu Han, *Fellow, IEEE*

Abstract—In mobile social networks, mobile users may help each other transfer contents when they move and meet each other. In this paper, we consider a mobile publish-subscribe network in which a mobile user (i.e., a content source) incentivizes other users (i.e., messengers) to forward its content by transferring energy to them. The messengers can utilize the received energy for content forwarding and/or their own use. From the perspective of a rational content source, to optimize the content delivery performance, we formulate the problem of energy charging and content transferring as a Markov decision process (MDP), a framework for decision making in stochastic systems. The objective is to maximize the expected utility in terms of the benefit gained by the content source from delivering the content and cost from energy consumption and transfer. The solution to the MDP is proved to be a threshold policy. Performances of the MDP-based scheme are examined in different system scenarios. The numerical results show that the MDP-based scheme outperforms the conventional baseline schemes.

Index Terms—Wireless energy charging, mobile publish-subscribe network, Markov decision process.

I. Introduction

In distributed mobile social networks, mobility of users can be used to transfer contents [1]. A publish-subscribe network [2] is one typical form of the mobile social network, where contents are transferred unidirectionally from sources (i.e., publishers) to destinations (i.e., subscribers) directly or with the help of relay nodes, which can find its applications in systems such as vehicular networks where traffic information messages are disseminated from fixed stations (i.e., content sources) to vehicles and pedestrians. By successfully received the contents, the content sources will be rewarded. A content source can transfer its content to a mobile router or a messenger. The messenger moves and forwards the content to the destination when they meet each other. Typically, the content source will choose the messenger that is socially close to the destination because it is likely that the messenger will meet the destination and be able to deliver the content efficiently.

Y. Zhang is with the School of Computer Science and Technology, Wuhan University of Technology, China (E-mail: yangzhang@whut.edu.sg).

D. Niyato and P. Wang are with the School of Computer Science and Engineering, Nanyang Technological University, Singapore (E-mail: dniyato@ntu.edu.sg, wangping@ntu.edu.sg).

D. I. Kim is with the School of Information and Communication Engineering, Sungkyunkwan University (SKKU), Suwon, Korea (E-mail: dikim@skku.ac.kr).

Z. Han is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004 USA (E-mail: zhan2@uh.edu).

Similarly, the messenger will be willing to spend its resource to help the content source if it has a social relationship. However, without social relationship, an incentive mechanism has to be employed to motivate the messengers to participate in the mobile social networks and to help content sources deliver the content. An incentive can be money that the content source pays to the messengers. However, paying money becomes challenging in practice as it requires sophisticated and secure infrastructure to support. Alternatively, energy is a scarce and important resource in a mobile network environment. With the advancement of energy transfer and charging techniques, e.g., inductive coupling and magnetic resonant coupling, it becomes possible to use energy as an incentive in the mobile publish-subscribe network. However, energy is also limited for the content source, and thus it has to be optimized for performance and cost.

We consider the distributed mobile publish-subscribe networks supported by wireless energy charging.[1] The content source generates contents and stores the contents in a queue. The content source can forward the content to a messenger together with the energy as an incentive when they meet each other. The messenger uses the energy for content delivery, e.g., to store the content, to search, and to forward the content to the destination. The content source optimizes energy replenishment from wired or wireless chargers and energy transfer to messengers. Thus, the content source has to consider the following factors. Firstly, the receiving energy from different chargers may cost the content source differently. Secondly, the capability of messengers to deliver the content is different. For example, some messengers are likely to meet the destination more frequently. After all, the mobility of the content source is random. Therefore, it has to optimize when to receive energy from the charger and which messenger to forward the content and energy to. In this paper, we propose a Markov decision process (MDP) approach to model the content delivery of the content source and messengers in the distributed mobile publish-subscribe networks. An MDP is a framework to assist the decision maker, i.e., the content source, to observe the system states and make decisions accordingly. By solving the MDP, the content source obtains optimal decisions to manage energy utilization and content delivery. In particular, there are different energy chargers with various energy prices in

---

[1]Note that the proposed model in this paper is also directly applicable to wired charging.

the network for the content source to replenish energy. There are also messengers with different probabilities of successful content delivery. The content source makes its decision by observing the current system state of the contact, energy storage, and the contents in the queue so that the utility is maximized.

The contributions of this paper are summarized as follows:

- We propose the concept of an energy incentive in mobile publish-subscribe networks. This is a novel incentive mechanism for mobile nodes to help each other forward contents, improving the network performance. However, because of intermittent energy replenishment from the chargers and energy transfer to messengers, the energy management issues in such mobile publish-subscribe networks needs to be reconsidered.
- We develop an MDP model for the content source to optimally manage energy charging content delivery. In particular, the content source makes decisions both on the energy charging and content forwarding to the messenger to maximize the expected utility.
- We study the structure of the optimal decisions of energy charging and content delivery obtained from the MDP model, and prove that the optimal MDP policy obtained is a threshold policy.
- We develop a partial iteration algorithm for the content source to make decisions, given the existence of a threshold policy. The proposed partial iteration algorithm can yield the near-optimal performance with substantially reduced complexity in practical systems, compared with the classical algorithm.

The rest of this paper is organized as follows. We review related work in Section II. The system model of the distributed mobile publish-subscribe network with energy transfer is described in Section III. Section IV presents the Markov decision process (MDP) formulation for the content source. The solution method of MDP is discussed in Section V. In Section VI, the existence of threshold policy in the optimal policy solved from the MDP model is proved. Based on the threshold policy, a partial iteration algorithm as an approximation decision making scheme is also proposed. Numerical results are presented in Section VII. Section VIII concludes this paper.

II. RELATED WORK

A. Mobile Publish-Subscribe Networks

The authors in [1] review the concept of mobile social networks, as well as survey their architectures and applications. In distributed mobile social networks, content dissemination processes can be done by opportunistic contacts [1]. Relays have an important role in mobile social networks in helping other mobile nodes forward and deliver contents. According to [3], the contact rate to content destinations is an essential quality of a relay. In publish-subscribe social networks, the social relationship and metrics has to be taken into account when performing content delivery [4]. In [5], a virtual "flea market" is proposed for vehicular networks to publish and disseminate query information generated by mobile devices (e.g., vehicles).

With the feature of vehicle mobility, query information can be transferred to destinations (i.e., subscribers) by opportunistic vehicle-to-vehicle contacts. The authors in [6] propose a hybrid content relaying mechanism in vehicular publish-subscribe networks, where the relay can be infostations to keep and forward contents. However, in [5], information dissemination processes by vehicle-to-vehicle communications are treated as free-of-charge processes. Similarly, users in the discussed networks in [6] are assumed to be cooperative. Those assumptions may not be applicable for systems with socially selfish mobile users. Relays are introduced in [4] as self-interested brokers when contacting and communicating with delays, publishers, and subscribers. That is, not all mobile nodes may want to help in content delivery as it costs them resources, e.g., energy and spectrum usage. The social selfishness issue is discussed in [3] that mobile nodes may prefer the contacts with other social mobile nodes with a strong tie [7] over those with weak ties. Based on the social selfishness, the delivery probability of a mobile node is measured, considering its contact rate with content destinations as well as its willingness to forward.

Network nodes in mobile publish-subscribe networks are mobile devices relying on battery power. Therefore, their energy management issues become critical. In [8], the authors use a continuous Markov model to model the opportunistic forwarding of contents. It is shown that, in two-step and epidemic forwarding, a threshold-type policy is the best policy to achieve the highest successful content transfer probability. In [9], peer-to-peer data dissemination in a mobile ad-hoc environment is characterized by three resource constraints, including energy, communication bandwidth, and storage. By imposing the constraints on these resources jointly, an algorithm is proposed to improve the node throughput compared with two existing approaches. In [10], a two-hop system, i.e., a source-relay-destination structure, of a typical mobile publish-subscribe network was studied. The random energy harvesting is considered in the system. The objective is to maximize the transferred contents within a certain period, i.e., throughput.

B. Wireless Energy Charging

Wireless energy charging has been introduced as a method to supply energy for mobile systems [11].

The feasibility of wireless energy charging has been studied. For example, for near-field wireless charging, Qi is introduced as an interface standard [14]. For Qi-compliant devices, 5W energy can be charged with the distance up to 4cm. For far-field wireless charging efficiency, RF-to-DC conversion efficiency can reach 50% [15]. In [13], a prototype for wireless sensor networks that harvest ambient RF energy is proposed. From the experiments, $20\mu W$ power can be harvested from a TV tower located 6.6km away over the UHF band. Compared to ambient light power as an alternative wireless energy source, the RF energy harvesting is 1% energy efficiency [16]. However, it can work without light, e.g., for lowlight indoor environments or at night. The experiments have been done in [17] to measure the RF transferred power. The results show that, with a WiFi AP transferring energy via 2.4GHz ISM

band, the corrected average power received at the distances of 2m and 10m can be 0.052μW and 0.011μW, respectively. A mobile phone working at 0.5W can achieve the power densities of 40mW/m^2, 1.6mW/m^2, and 0.4mW/m^2 at the distances of 1m, 5m, and 10m [18], respectively. A commercial wireless energy charging system has been launched [12], which can support battery-less mobile receivers to have 3.3V system output voltage. Further, battery-less mobile devices are surveyed in [11].

Wireless energy charging and information transfer are both in the form of electromagnetic waves. As a result, energy and information can be transferred simultaneously in the same channel, depending on whether the received signal is processed by the energy harvester or the information decoder of the receiver [19]. Practically, two patterns of simultaneous energy and information transfer are proposed, as in [19], [20] and [21], including a time switching pattern where each time slot is dedicated to energy transfer or infomation transfer only, and a power splitting pattern where the received signal is splitted in to two parts, which are received as energy and decoded as information at the same time. Large-scaled wireless networks are discussed in [20], where relays may exist for a cooperative energy and information transfer scheme. The mobile nodes in [20] are distributed in spatial Poisson process. However, mobility and contact issues of mobile nodes are not adequately considered.

C. Performance Optimization and Markov Decision Process

A Markov decision process (MDP) [22] is widely applied to model the decision making and obtain an optimal policy in mobile networks. In a vehicular delay-tolerant network (VDTN), which is a special form of mobile social networks, an MDP model is introduced for the mobile router to decide whether to forward a content or not [23]. A constrained MDP (CMDP) is employed in [24] to optimize the profit earned by a content provider (i.e., publisher) in terms of maximized number of users to receive fresh contents.

With the emerging wireless energy charging techniques, the optimization can be performed for energy replenishment of mobile nodes. For example, in [25], because wireless energy is randomly harvested, the mobile node has to balance between being idle and sensing the channel, which consumes energy. An MDP is formulated and solved to obtain the mode switching policy. In [26], an MDP is developed to obtain the policy to decide on data transmission given random data generation and energy harvesting processes so that the average delay of data transmission is minimized. The application of an MDP in a body sensor network is discussed in [27]. The MDP takes the energy state of the battery of a sensor node and charging condition to determine an action of different transmission (i.e., reporting) modes. The objective is to maximize the ratio of all the events correctly reported. However, those aforementioned systems have not taken advantages of wireless energy charging in mobile networks.

To the best of our knowledge, in the existing literature, the optimal wireless energy charging and content transferring policy of content providers in publish-subscribe mobile networks are not studied. This paper extends our early work [28]

by considering a threshold policy. Furthermore, the low-complexity algorithm is newly proposed.

III. SYSTEM MODEL

We consider a distributed mobile publish-subscribe network with the energy transfer capability, e.g., supported by wireless charging techniques, as shown in Fig. 1. The network is composed of energy chargers, a content source, mobile content messengers, and a content destination. In the network, the content source visits different locations and makes a contact to a charger or messenger when they meet each other. However, in some cases, the content source may not meet the content destination directly and has to utilize a messenger for the content delivery. For the content source, the set of contact states is denoted by $\mathcal{C} = 0, 1, \ldots, C$. The contact state determines whom the content source is currently visiting or encountering with. We assume that each contact state corresponds to each of a charger or messenger only. Therefore, the transitions of contact states indicate the mobility pattern of the content source. There are N energy chargers in the network. When the content source visits the charger, the content source can decide to charge and receive energy from the charger. The charger requests the content source to pay for the energy with the price $\pi^{CH}(\mathcal{C})$ per unit of energy if the content source decides to charge energy. Here, the price is associated with the contact state $\mathcal{C}$ of the content source.
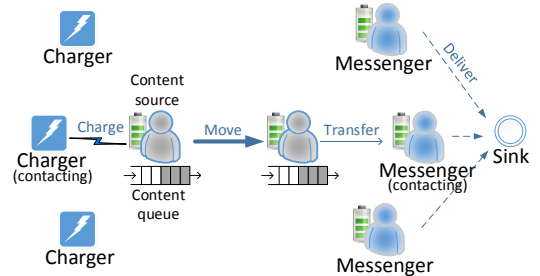

Fig. 1: System description.

The content source can generate new contents. The generated contents will be stored in the queue of the content source. The maximum capacity of the queue is Q contents. The content source aims to deliver the contents to the destinations, called sinks. There are M content messengers in the network that can assist to relay and deliver contents. $p_m(\mathcal{C})$ is defined as the probability that messenger m whom the content source is contacting will successfully deliver the content to the sink. When the content source meets a messenger, the content source can decide to forward contents in the queue to the messenger. Simultaneously, δ units of energy will be transferred to the messenger as a compensation to support the messenger to assist the content delivery[2]. This content and energy transfer can be performed over the same interface, e.g.,

---

[2]Here, the energy transfer loss can be taken into account. In this case, one unit of energy that the content source receives from the charger is defined as $x + \epsilon$, where x is the amount of energy that is actually received by the messenger and ϵ is the amount of energy loss.

TABLE I: Notation descriptions.

Notation	Definition
$\mathcal{S} = (\mathcal{C}, \mathcal{Q}, \mathcal{E})$	Composite state of a content source, including its contact state, queue state, and energy storage state
$\mathcal{A} \in \{0, 1, 2\}$	Possible action of a content source
$P_S(\mathcal{S}, \mathcal{S}'\|\mathcal{A})$	Transition probability from compound state $\mathcal{S}$ to state $\mathcal{S}'$ when action $\mathcal{A}$ is taken
$\mathbf{S}$	Transition matrix of compound state $\mathcal{S}$
$\pi^{CH}(\mathcal{C})$	Price of charging energy at contact state $\mathcal{C}$
$p_m(\mathcal{C})$	Content delivery probablity of messenger m at contact state $\mathcal{C}$
δ	Energy transferred from a content source to a messenger
$F(\mathcal{S}\|\mathcal{A})$	Immediate utility to a content source at state $\mathcal{S}$ when action $\mathcal{A}$ is taken
$H(\mathcal{S}\|\mathcal{A})$	Utility (not optimized) of a content source works from the current system state $\mathcal{S}$
$U(\mathcal{S})$	Optimal utility of a content source works from the current system state $\mathcal{S}$

using the techniques of simultaneous wireless information and power transfer (SWIPT) as in [19]. Alternatively, they can be simply performed over wired connection. The messenger then carries and moves to deliver the content to the corresponding sink.

Due to the movement and energy dissipation, energy transferred to the content messenger m will be consumed with the rate of ν_m. After the transferred energy is exhausted before the sink is met, the content will be discarded by the messenger, and the content delivery fails. Note that we consider the energy used only for the content transfer which is the main focus of this paper. The energy used for running mobile applications and miscellaneous communication by the mobile source is a separate part and will not be considered in the optimization. For the δ units of transferred energy, the content delivery to the sink will be successful in the duration of $\frac{\delta}{\nu_m}$ starting from the moment of content and energy arrival at the messenger. This is defined as the energy depletion time of δ energy units at the messenger. To calculate the probability of successful delivery by the messenger, we assume that the time duration before the energy is completely depleted, i.e., depletion time, to be random and modeled as a phase-type distribution [29]. The phase-type distribution is a relatively general model, which can be employed to model exponential and hyper-exponential distributions. In this case, the probability that the messenger m can successfully meet and deliver the content to the sink before the energy depletion happens is as follows:

$$p_m(\mathcal{C}) = 1 - \psi_m e^{\frac{\delta}{\nu_m}\mathbf{S}_m}\vec{\mathbf{1}}. \quad (1)$$

By definition of phase-type distribution [29], ψ_m is an initial probability row vector, $\mathbf{S}_m$ is a subgenerator matrix, and $\vec{\mathbf{1}}$ is a vector of ones with an appropriate size. The phase-type distribution has been used as a generic model of user mobility [30].

With the aforementioned system model, the content source has to make a decision to keep idle, charge energy, or transfer contents (with energy transferred as a compensation) to the contacted messenger. Some system parameters and notations are summarized as in Table I.

IV. OPTIMIZATION MODEL

We formulate a Markov decision process (MDP) model for the content source. The MDP model includes the following parameters: the system states of the content source, transition matrices of the states, the actions, and the corresponding reward/utility from taking the actions. By optimally solving the MDP, the decision maker, i.e., the content source, can observe or calculate the parameters and make optimal decisions accordingly to maximize the expected utility.

A. State Space and Action Space

The state space of the MDP model is defined as $\mathbb{S} = \{\mathcal{S} = (\mathcal{C}, \mathcal{Q}, \mathcal{E})\}$, where $\mathcal{C} \in \mathbb{C}$. $\mathbb{C} = \{0, 1, \ldots, C\}$ represents the set of contact states. The contact state indicates the current meeting event of the content source in the network, i.e., with what type of components the content source is currently meeting with[3]. The set of contact states is defined as $\mathbb{C} = \mathbb{C}^{\varnothing} \cup \mathbb{C}^{CH} \cup \mathbb{C}^{MG} \cup \mathbb{C}^{SN}$, where $\mathbb{C}^{\varnothing}$, $\mathbb{C}^{CH}$, $\mathbb{C}^{MG}$, and $\mathbb{C}^{SN}$ are the sets of states indicating the content source to be alone, to meet with an energy charger, to meet with a messenger, and to meet with the destination, respectively. We assume that the content source does not meet with an energy charger and messenger at the same time, i.e., $\mathbb{C}^{CH} \cap \mathbb{C}^{MG} = \varnothing$. In this work, we consider that the content source contacts one charger/messenger at a time. This is the scenario that only the cheapest charger or the messenger which is most probably to successful deliver the content will be chosen in the case of multiple accessible chargers/messengers. $\mathcal{Q}$ represents the current queue state of content source, where $\mathcal{Q} \in \mathbb{Q} = \{0, 1, \ldots, Q\}$. $\mathcal{E}$ is the energy level or energy state of the content source, where $\mathcal{E} \in \mathbb{E} = \{0, \ldots, E\}$. The energy level indicates the amount of energy unit in the battery. E is the maximum capacity of the battery.

The content source takes an action in a time-slotted fashion. Each time slot is called a decision period. The action space is defined as $\mathbb{A} = \{0, 1, 2\}$ where $\mathcal{A} = 0$ represents the action of being idle (i.e., not charging energy and not transferring content), $\mathcal{A} = 1$ represents the energy charging action, and $\mathcal{A} = 2$ represents that the content in the queue is transferred to the messenger for delivery. To incentivize the content messenger to relay the transferred content, certain units of energy will be transferred to the messenger as well.

B. Transition Matrices

In the MDP model, the system state transits from the current state $\mathcal{S} = (\mathcal{C}, \mathcal{E}, \mathcal{Q})$ to the next state $\mathcal{S}' = (\mathcal{C}', \mathcal{E}', \mathcal{Q}')$. In the following, we present the transition matrices of the proposed MDP model.

Each row of a transition matrix corresponds to the current state, and each column corresponds to the possible next state. As a result, each element denotes the transition probability from the current state (i.e., row) to the next future state (i.e.,

---

[3]The same type of system components means the components have the same property, e.g., if a content source meets with two messengers with the same probability of content delivery to the corresponding sink, they are treated as the same contact state.

column). In practical system deployments, the transition probabilities in transition matrices can be derived by theoretically modeling the state transitions, e.g., location state transitions derived by mobility modeling as in [31] or queue state transitions derived by assuming that queue content arrivals follow Poisson distribution. On the other hand, the content source can employ existing datasets of system states to estimate the transition probabilities, e.g., by calculating the frequency counts of transition processes [32] or by continually updating the transition matrices from real time training data obtained in an online fashion when the system is in operation [33].

1) Contact State Transition Matrix: Firstly, the transition matrix of the contact state $\mathcal{C}$ is expressed as follows:

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{\emptyset,\emptyset} & \mathbf{C}_{\emptyset,CH} & \mathbf{C}_{\emptyset,MG} & \mathbf{C}_{\emptyset,SN} \\ \mathbf{C}_{CH,\emptyset} & \mathbf{C}_{CH,CH} & \mathbf{C}_{CH,MG} & \mathbf{C}_{CH,SN} \\ \mathbf{C}_{MG,\emptyset} & \mathbf{C}_{MG,CH} & \mathbf{C}_{MG,MG} & \mathbf{C}_{MG,SN} \\ \mathbf{C}_{SN,\emptyset} & \mathbf{C}_{SN,CH} & \mathbf{C}_{SN,MG} & \mathbf{C}_{SN,SN} \end{bmatrix}. \quad (2)$$

The transition matrix $\mathbf{C}$ describes the mobility of the content source to different contact state. Each sub-matrix $\mathbf{C}_{m,m'}$, for $m, m' \in \{\emptyset, CH, MG, SN\}$ is the transition matric corresponding to the contact state in the sets $\mathbb{C}^{\emptyset}$, $\mathbb{C}^{CH}$, $\mathbb{C}^{MG}$, and $\mathbb{C}^{SN}$. The element of matrix $\mathbf{C}$ is denoted by $p_{\mathcal{C},\mathcal{C}'}^{C}$, which is the probability of the content source to change the contact state $\mathcal{C}$ in the current time slot to the state $\mathcal{C}'$ in the next time slot. Note that the transitions of contact states denote the mobility of a mobile node, which is widely modeled as a Markov chain, e.g., as in delay-tolerant networks (DTNs) [34], [35].

2) Queue State Transition Matrix: There are different cases of deriving the transition of the queue state $\mathcal{Q}$. Firstly, when the content source generates contents, the queue state can increase. Let $f^a(k)$, $k = 0, 1, \ldots, +\infty$ denote the probability that k contents are generated in one time slot, where $\sum_{k=0}^{+\infty} f^a(k) = 1$. In this case, the transition matrix is defined as a $(Q+1) \times (Q+1)$ matrix $\mathbf{Q}^+$, as follows:

$$\mathbf{Q}^+ = \begin{bmatrix} f^a(0) & f^a(1) & \cdots & f^a(Q-1) & \sum_{k=Q}^{\infty} f^a(k) \\ & f^a(0) & \cdots & f^a(Q-2) & \sum_{k=Q-1}^{\infty} f^a(k) \\ & & \ddots & & \vdots \\ & & & & 1 \end{bmatrix}. \quad (3)$$

Secondly, when the action to transfer is made by the content source, one content will be transferred. Thus, the number of contents in the queue can decrease. Note that there can still be new contents generated in the same time slot. Here, we assume that the transferring action is taken before the content arrivals at the content source. The transition matrix of the queue for this case is defined as the following $(Q+1) \times (Q+1)$ matrix:

$$\mathbf{Q}^-(\mathcal{A}) = \begin{bmatrix} f^a(0) & f^a(1) & \cdots & f^a(Q-1) & \sum_{k=Q}^{\infty} f^a(k) \\ f^a(0) & f^a(1) & \cdots & f^a(Q-1) & \sum_{k=Q}^{\infty} f^a(k) \\ & f^a(0) & \cdots & f^a(Q-2) & \sum_{k=Q-1}^{\infty} f^a(k) \\ & & \ddots & \vdots & \vdots \\ & & & f^a(0) & \sum_{k=1}^{\infty} f^a(k) \end{bmatrix}, \quad (4)$$

for $\mathcal{A} = 2$, and $\mathbf{Q}^-(\mathcal{A}) = \mathbf{Q}^+$ otherwise. The first row of the matrix indicates that there is currently no content stored in the queue. The other rows of the matrix denote the case that the queue has at least one content that can be transferred. After the content in the queue is transferred, there can be a content arrival. Note that here we assume that the content leaves the queue of the content source if it takes the transferring action. More reliable protocols, e.g., the content leaves the queue only when it is successfully transferred to the sink, can be adopted in the model with slight modification to the transition matrix presented in (4). Due to space limit, we omit this case in our paper.

3) Energy State Transition Matrix: When the content source is with an energy charger, i.e., a contact state is $\mathcal{C} \in \mathbb{C}^{CH}$, and the charging action is taken, i.e., $\mathcal{A} = 1$, the energy state can increase. For each charging action, the content source can receive at most $\Gamma_{\mathcal{C}}$ units of energy from charger at the location associated with the contact state $\mathcal{C}$. The probability of k energy units received by the content source is denoted by $\gamma_{\mathcal{C}}(k)$, $k = 0, \ldots, \Gamma_{\mathcal{C}}$. The parameters $\Gamma_{\mathcal{C}}$ and $\gamma_{\mathcal{C}}(k)$, $k = 0, \ldots, \Gamma_{\mathcal{C}}$, can be obtained from the charger or energy markets in the system model under consideration. The transition matrix of the energy state in this case is denoted by an $(E+1) \times (E+1)$ matrix $\mathbf{E}^+$.

For $\Gamma_{\mathcal{C}} < E$, which is the case that the maximum units of energy arrival is less than the maximum capacity of the battery, we have

$$\mathbf{E}^+ = \left[\begin{array}{cccc|ccc} \gamma_{\mathcal{C}}(0) & \cdots & \cdots & \gamma_{\mathcal{C}}(\Gamma_{\mathcal{C}}) & & & \\ & \ddots & & & \ddots & & \\ & & \gamma_{\mathcal{C}}(0) & \cdots & \cdots & & \gamma_{\mathcal{C}}(\Gamma_{\mathcal{C}}) \\ & & & \gamma_{\mathcal{C}}(0) & \cdots & & \sum_{k=0}^{1} \gamma_{\mathcal{C}}(\Gamma_{\mathcal{C}} - k) \\ & & & & \ddots & & \vdots \\ & & & & & & 1 \end{array} \right]. \quad (5)$$

For $\Gamma_{\mathcal{C}} \geq E$, we have

$$\mathbf{E}^+ = \begin{bmatrix} \gamma_{\mathcal{C}}(0) & \cdots & \cdots & \sum_{k=0}^{\Gamma_{\mathcal{C}}-E} \gamma_{\mathcal{C}}(\Gamma_{\mathcal{C}} - k) \\ & \gamma_{\mathcal{C}}(0) & \cdots & \sum_{k=0}^{\Gamma_{\mathcal{C}}-E+1} \gamma_{\mathcal{C}}(\Gamma_{\mathcal{C}} - k) \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}. \quad (6)$$

When the content source takes the action to transfer a content from its queue, $\delta \in \{1, \ldots, \mathcal{E}\}$ units of energy will be transferred to the content messenger together as a compensation of content delivery to the destination. Considering the energy efficiency, we assume that at least one unit of energy will be transferred. In this case, the energy state $\mathcal{E}$ of the content source will decrease. The transition matrix of the energy state is expressed an $(E+1) \times (E+1)$ matrix $\mathbf{E}^-$ as follows:

$$\mathbf{E}^- = \left[\begin{array}{cc|ccc} \mathbf{I}_{(\delta+1)\times(\delta+1)} & & & & \\ & 1 & & & \\ \hline & & \ddots & & \\ & & & 1 & \mathbf{0}_{1\times\delta} \end{array} \right], \quad (7)$$

where $\mathbf{0}_{1\times\delta}$ is a row sub-matrix of zeros. The sub-matrix $\mathbf{I}_{(\delta+1)\times(\delta+1)}$ is an identity matrix with the dimension of $(\delta+1) \times (\delta+1)$, which indicates that the stored energy in the content source is less than δ, and thus is not enough for

performing further content delivery. The energy state remains the same.

4) Overall System State Transition Matrix: The transition matrix of the current composite state $(\mathcal{E}, \mathcal{Q})$ and the next composite state $(\mathcal{E}', \mathcal{Q}')$ is derived as follows:

$$\mathbf{W}(\mathcal{S}, \mathcal{A}) = \begin{cases} \mathbf{Q}^+ \otimes \mathbf{E}^+, & \mathcal{C} \in \mathbb{C}^{CH}, \mathcal{A} = 1, \\ \mathbf{Q}^- \otimes \mathbf{E}^-, & \mathcal{C} \in \mathbb{C}^{MG}, \mathcal{A} = 2, \delta \leq \mathcal{E}, \\ \mathbf{Q}^+ \otimes \mathbf{I}, & \text{otherwise}, \end{cases} \quad (8)$$

where $\otimes$ is the Kronecker product.

- The first condition is for the case that the charging action is taken while the content source is with a charger. Here, the energy state can increase.
- The second condition is for the case that, given there is enough energy δ to be transferred as an incentive, the content source takes the content (as well as energy) transferring action when it is contacting a messenger.
- The last condition is for the case that there is neither a content nor energy transferred. For example, the content source may choose to be idle and do nothing, or try to charge without contacting an energy charger. Consequently, the queue state only increases, and the energy state remains the same. The matrix $\mathbf{I}$ is an identity matrix with the dimension of $(E+1) \times (E+1)$.

Taking the contact state transition matrix into consideration, the overall transition matrix of the current state $\mathcal{S} = (\mathcal{C}, \mathcal{E}, \mathcal{Q})$ to the next state $\mathcal{S}' = (\mathcal{C}', \mathcal{E}', \mathcal{Q}')$ is obtained as follows:

$$\mathbf{S} = \mathbf{C} \otimes \mathbf{W}(\mathcal{S}, \mathcal{A}). \quad (9)$$

Here, we denote $P_S(\mathcal{S}, \mathcal{S}'|\mathcal{A})$ as the element of row $\mathcal{S}$ and column $\mathcal{S}'$ in the matrix $\mathbf{S}$, when action $\mathcal{A}$ is taken. It is the transition probability from the current state $\mathcal{S}$ to the next state $\mathcal{S}'$.

V. SOLVING THE MDP OPTIMIZATION MODEL

In this section, we first define an immediate utility function of the content source. Then, the method to obtain an optimal policy of the MDP model is presented.

A. Immediate Utility Function

In each decision period, the content source takes the actions of being idle, charging, or transferring a content to the messenger. A reward to the content source is defined as the utility. The immediate utility function $F(\mathcal{S}|\mathcal{A})$, where $\mathcal{S} = (\mathcal{C}, \mathcal{E}, \mathcal{Q})$, is given as follows:

$F(\mathcal{S}|\mathcal{A})$
$$= \begin{cases} -\omega_{ch}\pi^{CH}(\mathcal{C})\Gamma - \omega_{hd}\chi(\mathcal{Q}), & \mathcal{C} \in \mathbb{C}^{CH}, \mathcal{A} = 1, \\ \omega_{tn}\pi^{MG}p_m(\mathcal{C}) - \omega_{hd}\chi(\mathcal{Q}), & \mathcal{C} \in \mathbb{C}^{MG}, \mathcal{A} = 2, \delta \leq \mathcal{E}, \\ -\infty, & \mathcal{C} \notin \mathbb{C}^{CH}, \mathcal{A} = 1, or \\ & \mathcal{C} \in \mathbb{C}^{MG}, \mathcal{Q} = 0, \mathcal{A} = 2, or \\ & \mathcal{C} \in \mathbb{C}^{MG}, \mathcal{E} < \delta, \mathcal{A} = 2, \\ -\omega_{hd} \cdot \chi(\mathcal{Q}), & \text{otherwise.} \end{cases} \quad (10)$$

- The first condition in (10) includes the cost of energy charging from chargers. ω_{ch} and ω_{hd} are the weights, $\omega_{ch}, \omega_{hd} \in [0, 1]$. The cost consists of two components.
  - In the first component, $\pi^{CH}(\mathcal{C})\Gamma$ represents the cost of charging $\Gamma \in \{0, 1, \dots, \Gamma_{\mathcal{C}}\}$ units energy from the charger if the contact state (i.e., with a charger) is $\mathcal{C} \in \mathbb{C}^{CH}$. The cost is a function of the contact state $\mathcal{C}$ since the energy price is different at the different charger.
  - $\chi(\mathcal{Q})$ in the second component is the holding cost of all the contents in the queue at the moment. A typical holding cost can be the delay of the contents waiting in the queue, if the content source is sensitive to content delay. In this case, the holding cost can be defined as $\chi(\mathcal{Q}) = \mathcal{Q}$, which is the increment of delay in the current decision period.
- The second condition in (10) represents the profit gained by transferring the content from the queue, i.e., $\mathcal{A} = 2$, when there are enough energy units at the content source. In this condition, $\omega_{tn} \in [0, 1]$ is the weight coefficient. $\pi^{MG}p_m(\mathcal{C})$ is the revenue of successfully delivering the content with the help of the messenger for the contact state $\mathcal{C}$. As defined in (1), $p_m(\mathcal{C})$ is the probability that the messenger can successfully deliver the content to the sink (i.e., destination) before the transferred energy is depleted.
- Note here that, as shown by the third component, the content source is not allowed to take the charging action if it does not contact with any charger. Likewise, it is not allowed to transfer any content and energy if it does not have a contact with any messenger or it does not have enough energy in its battery. These conditions will yield an infinite negative reward.

B. Solving the MDP Optimization Model with Bellman Equation

With the states and actions, the transition probability matrices, and the immediate utility function defined, the MDP model can be solved by the following Bellman equation [22], [36] to obtain the optimal policy, i.e.,

$$U(\mathcal{S}) = \max_{\phi(\mathcal{S})} H(\mathcal{S}|\mathcal{A}), \quad (11)$$

$$\phi^*(\mathcal{S}) = \arg\max_{\phi(\mathcal{S})} H(\mathcal{S}|\mathcal{A}), \quad (12)$$

$$H(\mathcal{S}|\mathcal{A}) = F(\mathcal{S}|\mathcal{A}) + \beta \sum_{\mathcal{S}' \in \mathbb{S}} P_S(\mathcal{S}, \mathcal{S}'|\mathcal{A})U(\mathcal{S}'). \quad (13)$$

In the Bellman equation given in (11)-(13), the function $U(\mathcal{S})$ is the optimized utility from when the content source works from the current system state $\mathcal{S}$. $\phi(\mathcal{S})$ is defined as a policy function $\phi : \mathcal{S} \mapsto \mathcal{A}$, that is the action to be taken given the current state $\mathcal{S}$. $\phi^*(\mathcal{S})$ denotes the optimal policy. The essential concept of the Bellman equation is to maximize the utility of the decision maker, i.e., the content source, $H(\mathcal{S}|\mathcal{A})$, considering not only the current immediate utility $F(\mathcal{S}|\mathcal{A})$, but also the discounted possible future utilities $\beta \sum_{\mathcal{S}' \in \mathbb{S}} P_S(\mathcal{S}, \mathcal{S}'|\mathcal{A})U(\mathcal{S}')$. $\beta \in [0, 1)$ is the discount factor because of the uncertainty of the future utilities. $P_S(\mathcal{S}, \mathcal{S}'|\mathcal{A})$ is the transition probability from the current state $\mathcal{S}$ to the possible future state $\mathcal{S}'$, which is defined by (9). The value iteration algorithm can be applied to numerically solve the Bellman equation [22] to obtain the optimal utility $U(\mathcal{S})$.

C. Complexity Analyses of the MDP Optimization Model

The complexity of solving the MDP with the value iteration algorithm is $\mathcal{O}(|\mathbb{A}| \cdot |\mathbb{S}|^2)$, where $|\mathbb{A}|$ and $|\mathbb{S}|$ are the sizes of the action space and the state space, respectively. In the proposed system, as aforementioned, the size of state space $|\mathbb{S}|$ is decided by the product of the location, energy, and content state numbers. i.e., $|\mathbb{C}| \cdot |\mathbb{Q}| \cdot |\mathbb{E}|$.

In practical systems, real time solutions of the MDP optimization may subject to the computational capacity of the content source. For example, a content source can be a sensor node, or a laptop on mobile. Two approaches can be applied to solve the MDP optimization accordingly:

I. Given that the system states and state transitions are already available (e.g., by theoretical approaches), MDP solutions can be obtained in advance in an offline fashion. The solution can be downloaded to the memory of a mobile content source, e.g., in the form of a lookup table. The content source only observes the current state $\mathcal{S}$ to make optimal decision $\mathcal{A}$ without keeping past system states.

II. Given that the system states and state transitions are initially unknown to the content source, the content source may traverse over a service area to explore and collect past system data, such as system states and state transition probabilities, e.g., contact states, for training. After then, the collected system states and state transitions are employed for forming and solving the MDP optimization, which can be done in an online manner when the content source has high computational capacity. Otherwise, when the content source is not able to solve the optimization, the collected data can be offloaded to nearby base stations and cloud-like services for further process, where the final MDP optimization solution will be returned to the content source [39].

VI. STRUCTURAL RESULT OF MDP SOLUTION: THRESHOLD POLICY

In this section, we analyze the structural result in terms of a threshold policy of the MDP model for the content source. The threshold policy can help reduce the complexity of solving the MDP model. We firstly show that the action space of the MDP model can be simplified in some scenarios. Then, we prove that, given particular system parameters, the optimal policy solved by the MDP model is a threshold policy.

A. Threshold Policy

The optimal solution (i.e., policy) $\phi^*(\mathcal{S})$ solved by the MDP model is defined to be a threshold policy, if the following holds:

$$\phi^*(\theta, \mathcal{S}_{-\theta}) = \begin{cases} \mathcal{A}_1, & \text{for } \min\theta \preceq \theta \prec \theta_{\text{thr},1}, \\ \mathcal{A}_i, & \text{for } \theta_{\text{thr},i-1} \preceq \theta \prec \theta_{\text{thr},i} \\ & \quad \forall i \in \{2,3,\ldots,|\mathbb{A}|-1\}, \\ \mathcal{A}_{|\mathbb{A}|}, & \text{for } \theta_{|\mathbb{A}|-1} \preceq \theta \preceq \max\theta, \end{cases} \quad (14)$$

where θ denotes a state having a threshold. $\mathcal{S}_{-\theta}$ is a tuple of all the other states except θ. As defined in the Bellman equation given in (11)-(13), $\phi^*(\theta, \mathcal{S}_{-\theta})$ is the optimal decision policy. $\theta_{\text{thr},i}$ is called as the ith threshold state of the

system state θ. From the definition of threshold policy, the physical meaning of the threshold policy is that the action $\mathcal{A}$ monotonically changes as the system state $\mathcal{S}$ changes.

The existence of threshold policy contributes to efficiently solving the MDP model in the following manners. Once the threshold policy is proved to exist, and the particular threshold states are known, the actions corresponding to all the other states will be fixed without further computation, as demonstrated in (14). In particular, given any algorithm that can directly obtain the existing threshold states, the complexity can be reduced significantly, comparing with the complexity of $\mathcal{O}(|\mathbb{A}| \cdot |\mathbb{S}|^2)$ when solving the Bellman equation (11)-(13) by iterating all the system states using the value iteration algorithm [37]. The algorithm for directly deciding the threshold states deserves further research. However, a few studies, such as reinforcement learning [38] and approximation algorithms [39], have been conducted.

To prove that the optimal policy $\phi^*(\mathcal{S})$ in (12) is a threshold policy, the concept of supermodularity/submodularity [40] is applied.

Definition 1: For $x \in \mathbb{X} \subseteq \mathbb{R}$, $y \in \mathbb{Y} \subseteq \mathbb{R}$, a function $f(x,y) \in \mathbb{R}$ is *supermodular* in (x,y) if $f(x_1,y_1) - f(x_1,y_2) \geq f(x_2,y_1) - f(x_2,y_2)$, $\forall x_1, x_2 \in \mathbb{X}, \forall y_1, y_2 \in \mathbb{Y}, x_1 > x_2, y_1 > y_2$. Similarly, $f(x,y)$ is *submodular* in (x,y) if $f(x_1,y_1) - f(x_1,y_2) \leq f(x_2,y_1) - f(x_2,y_2)$, $\forall x_1, x_2 \in \mathbb{X}$, $\forall y_1, y_2 \in \mathbb{Y}, x_1 > x_2, y_1 > y_2$.

The supermodularity/submodularity property of $f(x,y)$ is a sufficient condition of the non-decreasing/non-increasing monotonicity of $y = \arg\max_y f(x,y)$ [36], [40]. Specifically, in the proposed MDP model and Bellman equation given in (11)-(13), for a given state $\theta \in \{\mathcal{C}, \mathcal{Q}, \mathcal{E}\}$, the fact that $H(\mathcal{S}|\mathcal{A})$ is supermodular/submodular in $(\theta, \mathcal{A})$ indicates that $\phi^*(\mathcal{S})$ is non-decreasing/non-increasing in $\theta \in \{\mathcal{C}, \mathcal{Q}, \mathcal{E}\}$.

In the following, we prove that a structure of threshold exists in the proposed MDP solutions.

Lemma 1: The actions of charging $\mathcal{A} = 1$ and content transferring $\mathcal{A} = 2$ will never be taken at a contact state that is without energy chargers (i.e., $\mathcal{C} \notin \mathbb{C}^{CH}$) and without content messengers (i.e., $\mathcal{C} \notin \mathbb{C}^{MG}$), respectively.

The proof of *Lemma* 1 is straightforward by definition and thus omitted for brevity. As stated in *Lemma* 1, the proposed MDP has a binary decision at each system state. The following theorem describes a threshold policy that exists with respect to the queue state $\mathcal{Q}$. The proof of *Theorem* 2 is in Appendix A.

Theorem 2: Given a fixed energy state $\mathcal{E}$ and a contact state $\mathcal{C} \in \mathbb{C}^{MG}$, when the content holding cost (e.g., delay cost) is zero, i.e., $\chi(\cdot) \equiv 0$, the optimal policy of the content source has a threshold structure with respect to the queue state $\mathcal{Q}$. In particular, a threshold $Q_{\text{thr}}(\mathcal{C}, \mathcal{E})$ exists that the content source takes action $\mathcal{A} = 2$ if $\mathcal{Q}_i \geq Q_{\text{thr}}(\mathcal{C}, \mathcal{E})$, and $\mathcal{A} = 0$ otherwise. The intuition of *Theorem* 2 is that if the content source has many contents in its queue, it is likely that the content source will be more willing to transfer the content to a messenger to gain the profit of content delivery. In the similar manner, the following theorems hold:

Theorem 3: Given any fixed queue state $\mathcal{Q}$ and a contact state $\mathcal{C} \in \mathbb{C}^{CH}$, the optimal policy of the content source is a

threshold policy in the energy state $\mathcal{E}$. In particular, a threshold $E_{\mathrm{thr}}(\mathcal{C}, \mathcal{Q})$ exists that the content source takes action $\mathcal{A} = 1$ if $\mathcal{E} < E_{\mathrm{thr}}(\mathcal{C}, \mathcal{Q})$, and $\mathcal{A} = 0$ otherwise.

The intuition of *Theorem* 3 is that if the content source has less energy in its battery, it is likely that the content source will receive energy from a charger. This is because the content source has to obtain sufficient energy for the future content transfer to gain the utility.

Let the contact states of the set $\mathbb{C}^{CH\dagger}$ (i.e., with chargers) be sorted according to the energy price π^{CH}, i.e., $\mathbb{C}^{CH\dagger} = \{\mathcal{C}_1, \ldots, \mathcal{C}_{|\mathbb{C}^{CH\dagger}|}\}$, for $\pi^{CH}(\mathcal{C}_i) \leq \pi^{CH}(\mathcal{C}_{i+1})$ and $i = 1, \ldots, |\mathbb{C}^{CH\dagger}| - 1$. In particular, it is more expensive for the content source to receive energy from a charger at the contact state $\mathcal{C}_{i+1}$ than at the contact state $\mathcal{C}_i$. Then, we have the following theorem.

Theorem 4: Given any fixed queue state $\mathcal{Q}$, energy state $\mathcal{E}$, and the set of contact states $\mathbb{C}^{CH\dagger}$ sorted according to the energy price, the optimal policy of the content source is a threshold policy in the contact state $\mathcal{C} \in \mathbb{C}^{CH\dagger}$. In particular, a threshold $C_{\mathrm{thr}}^{CH\dagger}(\mathcal{E}, \mathcal{Q})$ exists that the content source takes action $\mathcal{A} = 1$ if $\mathcal{C} < C_{\mathrm{thr}}^{CH\dagger}(\mathcal{E}, \mathcal{Q})$, and $\mathcal{A} = 0$ otherwise.

The intuition of *Theorem* 4 is that if the content source is meeting with a charger with a cheap energy price, it is likely that the content source will receive energy from the charger.

Let the contact states of the set $\mathbb{C}^{MG}$ (i.e., with messengers) be sorted according to the contact probability p_m of a messenger to the sink, i.e., $\mathbb{C}^{MG\dagger} = \{\mathcal{C}_1, \ldots, \mathcal{C}_{|\mathbb{C}^{MG\dagger}|}\}$, for $p_m(\mathcal{C}_j) \leq p_m(\mathcal{C}_{j+1})$ and $j = 1, \ldots, |\mathbb{C}^{MG\dagger}| - 1$. In particular, there is a higher chance the content from the content source will be successfully delivered to the sink by the messenger associated with the contact state $\mathcal{C}_{j+1}$ than that by the messenger associated with the contact state $\mathcal{C}_j$. Then, we have the following theorem.

Theorem 5: Given any fixed queue state $\mathcal{Q}$, energy state $\mathcal{E}$, and the sorted set of contact state $\mathbb{C}^{MG\dagger}$ according to the contact probability to the sink, the optimal policy of the content source is a threshold policy in the contact state $\mathcal{C} \in \mathbb{C}^{MG\dagger}$. In particular, a threshold $C_{\mathrm{thr}}^{MG\dagger}(\mathcal{E}, \mathcal{Q})$ exists that the content source takes action $\mathcal{A} = 2$ if $\mathcal{C} \geq C_{\mathrm{thr}}^{MG\dagger}(\mathcal{E}, \mathcal{Q})$, and $\mathcal{A} = 0$ otherwise.

The intuition of *Theorem* 5 is that if the content source is meeting with a messenger that meets with the sink frequently, it is likely that the content source will transfer a content to the messenger.

The proofs of *Theorem* 3, *Theorem* 4, and *Theorem* 5 are similar to that of *Theorem* 2 in Appendix A, and thus we omit them for brevity.

B. Threshold Policy Based Partially Iteration Algorithm

Based on the existence of a threshold policy of the MDP model, approximation algorithms for fast decision making can be developed. For example, our decision making algorithm based on the existence of threshold policy (namely FAST) is proposed in [39]. The basic idea of the FAST algorithm is to equally divide the state space into different parts, e.g., $\mathbb{S}_1, \mathbb{S}_2, \ldots, \mathbb{S}_n$, where $\mathbb{S}_1 \preceq \mathbb{S}_2 \preceq \cdots \preceq \mathbb{S}_n$ for n actions in the action space. Given the existence of a threshold policy,

only one particular action $\mathcal{A}_i$ will be taken for all the states in $\mathbb{S}_i$.

In this paper, we extend the FAST algorithm by introducing an approximation decision making scheme, called partial iteration algorithm (PIA) for the content source to take an action of charging, content transferring, or being idle. The algorithm is a combination of any given myopic scheme with low complexity and the value iteration algorithm with the existence of threshold policy. Algorithm 1 shows the proposed PIA with a binary action space and the threshold policy between actions $\mathcal{A} = 0$ and $\mathcal{A} = 1$. The proposed PIA starts with any given policy, e.g., obtained from FAST. Then, we randomly select a subset of states and update the utilities corresponding to the states in the selected subset. The utility updating is similar to that of the value iteration algorithm [22]. However, unlike the value iteration algorithm that iterates and updates every state with the information of all the future possible states until convergence, the PIA, as an approximation algorithm, only selects a part of the system states and updates their utilities with the knowledge of a subset of future possible states. In particular, the term $\beta \sum_{\mathcal{S}' \in \mathbb{S}_i^F} P_S(\mathcal{S}, \mathcal{S}'|\mathcal{A}) U(\mathcal{S}')$ in Line 7 of Algorithm 1 only includes the adjacent states close to the current states. Based on the utilities updated by limited information, we apply the threshold policy to choose the actions for the rest of system states.

Algorithm 1 Partial iteration algorithm

1: **procedure** GENERATING THE DECISION POLICY $\phi : \mathcal{S} \mapsto \mathcal{A}$
2:     Generate an initial policy $\phi_0(\mathcal{S})$;
3:     Randomly select a subset $\mathbb{S}^U \subseteq \mathbb{S}$;
4:     **for** each state $\mathcal{S}_i \in \mathbb{S}^U$ **do**
5:         Select a subset of future states $\mathbb{S}_i^F \subset \mathbb{S}$ corresponding to the current iterating state $\mathcal{S}_i$;
6:         Update the utility $H(\mathcal{S}|\mathcal{A})$ for the future states in $\mathbb{S}_i^F$:
7: $$H(\mathcal{S}|\mathcal{A}) \leftarrow F(\mathcal{S}|\mathcal{A}) + \beta \sum_{\mathcal{S}' \in \mathbb{S}_i^F} P_S(\mathcal{S}, \mathcal{S}'|\mathcal{A}) U(\mathcal{S}');$$
8:         Select an optimal action $\forall i \; \mathcal{A}_i \leftarrow \arg\max_{\mathcal{A}} H(\mathcal{S}|\mathcal{A})$;
9:         **if** $\mathcal{A}_i = 1$ **then**
10:            Update the policy: set $\phi(\mathcal{S}) \leftarrow 1, \forall \mathcal{S} \succ \mathcal{S}_i$;
11:        **else**
12:            Update the policy: set $\phi(\mathcal{S}) \leftarrow 0, \forall \mathcal{S} \prec \mathcal{S}_i$;
13:        **end if**
14:    **end for**
15:    **return** $\phi : \mathcal{S} \mapsto \mathcal{A}$;
16: **end procedure**

The algorithm starts from improving any given initial policy, as in Line 2. Instead of iterating and updating the utilities of all the system states using the value iteration algorithm, Line 3 selects and updates the utility of only a part of states (e.g., 10%). The selection of the states to be updated can be random or iterative. For example, for $|\mathbb{S}^U| = |\mathbb{S}|$, all the system states will be updated, where $\mathbb{S}^U$ denotes the subset of the selected states. Unlike the Bellman equation given in (11)-(13), which updates the utility function $H(\mathcal{S}|\mathcal{A})$ by adding the immediate utility of the current decision period as well as *all* the future states, i.e., $\mathcal{S}' \in \mathbb{S}$, the proposed PIA only performs the update by considering a subset of the future system states, i.e., $\mathcal{S}' \in \mathbb{S}_i^F$, as in Lines 5 and 7. In practice, the subset can include the states that the content source most probably

visits. For example, for the content source only charging one unit of energy, given the current state $\mathcal{S} = (\mathcal{C}_i, \mathcal{Q}_i, \mathcal{E}_i)$, the subset $\mathbb{S}_i^F$ can include the states $\mathcal{S}_1' = (\mathcal{C}_i, \mathcal{Q}_i, \mathcal{E}_i)$ and $\mathcal{S}_2' = (\mathcal{C}_i, \mathcal{Q}_i, \mathcal{E}_i + 1)$. The threshold policy is applied in Lines 9-12. Since the 0-to-1 threshold policy has been proved to exist (as a precondition to applying the algorithm), given that the action $\mathcal{A}$ to the state $\mathcal{S}_i$ is obtained as $\mathcal{A} = 1$, we can postulate that the actions corresponding to the state $\mathcal{S}_{i'}$ for $\mathcal{S}_{i'} \succ \mathcal{S}_i$ are all $\mathcal{A} = 1$. Similarly, given $\mathcal{A} = 0$ at the state $\mathcal{S}_i$, the state $\mathcal{S}_{i''}$ for $\mathcal{S}_{i''} \prec \mathcal{S}_i$ should have the action of $\mathcal{A} = 0$.

C. Complexity Analyses of PIA

Compared to the complexities $\mathcal{O}(|\mathcal{A}| \cdot |\mathbb{S}|^2)$ and $\mathcal{O}(1)$ of the MDP value iteration algorithm and the FAST algorithm, respectively, the complexity of the PIA is $\mathcal{O}(|\mathcal{A}| \cdot M \cdot K)$, where $M = |\mathbb{S}^U|$ is the size of the subset of the selected states, and $K = \max |\mathbb{S}_i^F|, \forall i, \mathcal{S}_i \in \mathbb{S}^U$, denotes the maximum number of future states to be updated. Since computing the utility in the PIA algorithm can be over a smaller state space, its complexity will be smaller than that of the value iteration algorithm. For example, if only a constant number of states close to the states in $\mathbb{S}^U$ are considered as future states, the complexity for the PIA is $\mathcal{O}(|\mathcal{A}| \cdot |\mathbb{S}|)$, which is one order lower than that of the MDP with value iteration algorithm.

VII. NUMERICAL RESULTS

A. System Settings

1) System Parameters: We compare the performances of different charging and content delivery schemes. Unless otherwise stated, the system parameters are set as follows.

- The battery has the maximum capacity of 10 units of energy, and that of the queue is 10 contents. New content arrival at the content source follows the Poisson distribution. The holding cost of contents is zero.
- There are equal numbers of chargers and messengers in the system that the content source can contact:
  - Contact state $\mathcal{C} = 0$ represents that neither a charger nor a content messenger is encountered;
  - Contact states $\mathcal{C} = 1, \dots, 6$ represent that the content source meets an overall number of 6 different chargers, with the energy prices of 0.0, 1.0, 4.0, 9.0, 16.0 and 25.0, respectively.
  - Contact states $\mathcal{C} = 7, \dots, 12$ represent that the content source may meet with 6 different messengers. By successfully transferring a content, the content source receives the revenue of 15.0. The probabilities of successful content delivery by the messengers associated with the states $\mathcal{C} = 7$ and $\mathcal{C} = 12$ are 0.0 and 1.0, respectively. These are two extreme cases that the content messenger cannot and can always successfully deliver the content. For $\mathcal{C} = 8, \dots, 11$, the phase-type distribution of the time to meet the sink has the following parameters: $\psi_m = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$, and $\mathbf{S}_m$ is an upper bidiagonal matrix, where all the main diagonal elements are $-\lambda$'s, and the other elements are λ. The

physical meaning of ψ_m and $\mathbf{S}_m$ is that the time to meet the sink is an Erlang distribution $E(4, \lambda)$. In our numerical analyses, one unit of energy will be transferred along with one content to any messenger, i.e., $\delta = 1$. The energy depletion rate is $\nu_m = 1$. The parameter $\lambda = 2.30, 3.21, 4.20, 5.52$ is for the messenger $\mathcal{C} = 8, \dots, 11$, which results in the successful content delivery probabilities to be 0.2, 0.4, 0.6, and 0.8, respectively. These probabilities are shown in Fig. 2(a).
  - The content source does not directly contact the sink.
- A new content is generated in each decision period with the probability of 0.4. The probability of successfully receiving energy from a charger is 0.9. The discount factor in the Bellman equation is set to 0.9.

2) Baseline Schemes and Evaluation Criteria: For comparison, we consider three conventional baseline schemes.

1) A greedy (GRD) scheme: The content source takes an action only to maximize the immediate utility of the current decision period. The greedy scheme attempts to achieve the global optimal action decisions by optimize local decisions.
2) A charge-and-transfer (CAT) scheme: The content source always charges when meeting with a charger, and always transfers a content when meeting with a messenger. By adopting this scheme, the content source makes simple decisions only by observing the current contact state.
3) A random (RND) scheme: The content source takes the actions of energy charging and content transferring randomly when they are possible.

To show how the existence of threshold policy can assist the decision making process, we also compare the MDP-based scheme with the FAST algorithm and the proposed PIA, as discussed in Section VI-B.

We obtain and evaluate the following performance measures of the content source:

- *Expected utility*: Suppose the content source start from any initial state $\mathcal{S}_{in} \in \mathbb{S}$ with a uniform distribution [33]. The expected utility is defined as $\mathscr{U} = \sum_{\mathcal{S}_{in} \in \mathbb{S}} \frac{1}{|\mathbb{S}|} \cdot U(\mathcal{S}_{in})$, where $U(\cdot)$ is defined in the Bellman equation given in (11)-(13).
- *Charging and content transferring rates*: At the system steady state, $p_{st}(\mathcal{S})$ denotes the steady state probability that the content source will be at the state $\mathcal{S}$. The energy charging rate of the content source is defined as $\alpha_{ch} = \sum_{\mathcal{S} \in \mathbb{S}} p_{st}(\mathcal{S}) \cdot I(\mathcal{A} = 1)$. The content transfer rate is defined as $\alpha_{mg} = \sum_{\mathcal{S} \in \mathbb{S}} p_{st}(\mathcal{S}) \cdot I(\mathcal{A} = 2)$. Function $I(c)$ returns the value of one if the condition c holds, and zero otherwise. The energy charging and content transfer rates indicate how much the content source prefers any charger and messenger, respectively.

B. Probabilities of Successful Content Delivery by Messengers

In Fig. 2(a), the cumulative probabilities of the phase-type distribution of the time for a messenger to meet the sink
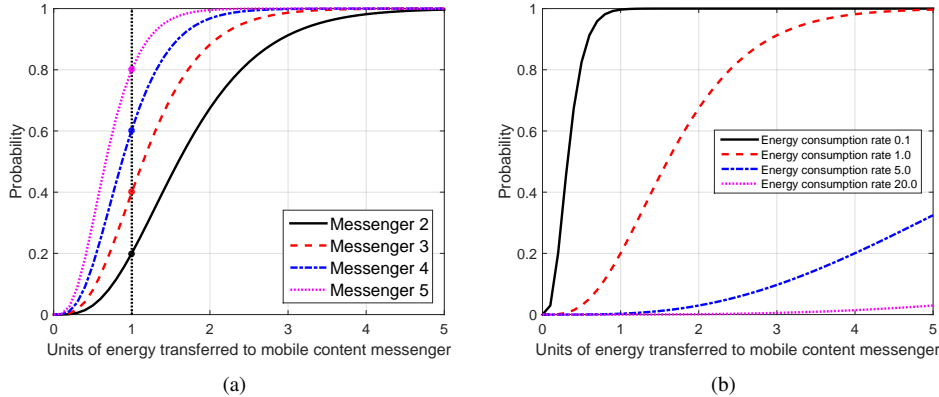
Fig. 2: The probabilities of successful content delivery for (a) Different content messengers, where messengers 2-5 correspond to the contact states $\mathcal{C} = 8, \ldots, 11$ of the content source, and (b) different energy consumption rates of a messenger.

before energy depletion are shown under varied parameters. Generally, more units of energy transferred to the messenger leads to a higher probability of successful content delivery as shown in Fig. 2(a). Here, messengers 2-5 correspond to the contact states $\mathcal{C} = 8, \ldots, 11$ and are set with the parameters $\lambda = 2.30, 3.21, 4.20, 5.52$, respectively. The messenger with a higher value of the parameter λ (e.g., $\lambda = 5.52$ for the messenger 5) will meet the sink more often, and thus the successful content delivery probability is higher.

Fig. 2(b) shows the impacts of energy consumption rate of a messenger to the successful content delivery probability. Clearly, the probability decreases as the energy consumption rate ν_m of the messenger m increases.

C. Threshold Policy

The optimal threshold policy of the content source is shown in Fig. 3. The policy is obtained from solving the MDP model using the value iteration algorithm.

The threshold with respect to the energy state is shown in Figs. 3(a) and (b). When the content source is at a charger (i.e., $\mathcal{C} = 1, \ldots, 6$), as the energy state $\mathcal{E}$ increases, the content source decides to charge and receive the energy from the charger. Basically, the content source monotonically changes its action from being idle ($\mathcal{A} = 0$) to charging ($\mathcal{A} = 1$). For example, as shown in Fig. 3(a), for $\mathcal{C} = 2$, $\mathcal{A} = 1$ when $\mathcal{E} \leq 4$, and $\mathcal{A} = 0$ otherwise. Similarly, Fig. 3(b) shows that, given that the content source has a contact with a charger (e.g., $\mathcal{C} = 1$ in this case), a threshold policy exists for each fixed queue state $\mathcal{Q}$. In the same manner, threshold policies with respect to the contact state $\mathcal{C}$ and the queue state $\mathcal{Q}$ are shown in Fig. 3(a) and Fig. 3(c), respectively.

In the following, we show the performance comparison among different algorithms and schemes including the PIA proposed in Section VI-B. In Fig. 4, we examine the impacts of the number of the selected states for updating, i.e., $|\mathbb{S}^U|$ in the Algorithm 1. Basically, we vary the size of the selected state space from 1 (i.e., one state) to $\frac{2}{24}|\mathbb{S}|, \frac{4}{24}|\mathbb{S}|, \ldots, \frac{20}{24}|\mathbb{S}|$, i.e., the selected state space increases, where $|\mathbb{S}|$ is the number of all the system states. The initial policy is generated by the

FAST algorithm [39]. In the FAST algorithm, when the action space is binary, the system state space is divided into two equal parts, each of which corresponds to one of the actions. As shown in Fig. 4, the FAST scheme achieves around 58% of the utility of the MDP-based scheme. For the PIA, by partially updating the utilities of some system states, the expected utility of the content source can be improved. For example, for the PIA, by updating only $\frac{6}{24}|\mathbb{S}|$ states (i.e., 25% of all the states), the expected utility increases to around 83% comparing with that of the MDP-based scheme. Thus, the proposed PIA offers acceptable expected utility with much lower complexity.

D. Impacts of Battery Capacity

Fig. 5(a) shows the expected utilities obtained from the optimal policy of the MDP model as well as baseline schemes. The optimality of the MDP-based scheme can be observed in terms of the maximum expected utility.

From Fig. 5(a), as the maximum battery capacity of the content source increases from $E = 2$ to $E = 10$, the expected utility increases. This is because the increased battery capacity allows more energy units to be stored to support further content transferring by the content source.

Fig. 5(b) shows the energy charging (dashed curves in the figure) and transferring rates (solid curves) at the system steady state. From the figure, as the battery capacity increases, the charging rate increases, indicating that the content source is more likely to charge at the chargers. Consequently, the content source tends to transfer a content because it has enough energy. However, as the battery capacity E increases above a certain level (e.g., $E = 6$ in Fig. 5(b)), the energy charging and content transferring rates stop increasing. This observation indicates that, due to the cost of energy charging, the amount of energy received by the content source should be limited when adopting the MDP-based scheme. As shown in Fig. 5(c), by employing the MDP-based scheme, the average energy level of the content source reaches around 1.0 after the energy capacity exceeds $E = 6$. That is, relatively large battery capacity of the content source may be unnecessary because the battery can be underutilized. Notably, the energy charging
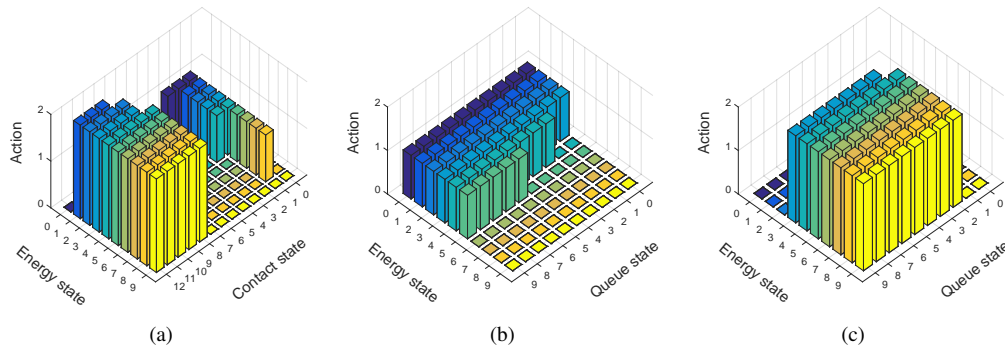
Fig. 3: Threshold policy for different (a) energy state $\mathcal{E}$ and contact state $\mathcal{C}$ (for queue state $\mathcal{Q} = 1$), (b) states $\mathcal{E}$ and $\mathcal{Q}$ (for contact state $\mathcal{C} = 1$, i.e., meeting with a charger), (c) states $\mathcal{E}$ and $\mathcal{Q}$ (for contact state $\mathcal{C} = 8$, i.e., meeting with a messenger).
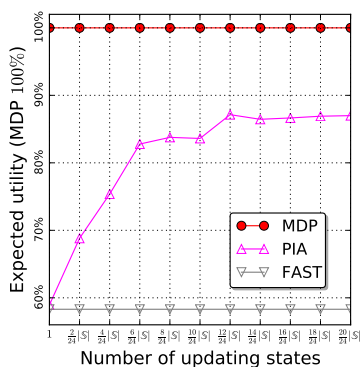

Fig. 4: Comparison of approximation algorithms with the MDP-based scheme.

TABLE II: Experimental cases: Different energy prices.

Energy price level	Energy prices when contacting chargers					
	$\mathcal{C} = 1$	$\mathcal{C} = 2$	$\mathcal{C} = 3$	$\mathcal{C} = 4$	$\mathcal{C} = 5$	$\mathcal{C} = 6$
1	0.0	1.0	2.0	3.0	4.0	5.0
2	1.0	2.0	3.0	4.0	5.0	6.0
. . .			. . .			
9	8.0	9.0	10.0	11.0	12.0	13.0

As the energy price level increases to some extent, e.g., after level 7, the charging cost incurred to the content source becomes too high. In this case, the expected utility shown in Fig. 6(a) stops falling and becomes stable when the MDP-based scheme is adopted. This is because that the content source tends to refuse to charge. The energy charging and content transferring rates decrease and approach zero, as shown in Fig. 6(b). Similarly, as shown in Fig. 6(a), the PIA has acceptable performance metrics compared with the MDP-based scheme, since the algorithm is partially adaptive. For comparison, when the myopic charge-and-transfer (CAT) scheme or the FAST algorithm is applied, the expected utility decreases drastically. This is because the content source will not adjust its policy to charge selectively only at the cheapest chargers.

Fig. 7 shows that the probability of successful messenger-sink content delivery is positively correlated with the expected utility of the content source. We examine four cases, i.e., case 1 to 4. The energy consumption rate ν_m are respectively 10.0, 4.0, 1.0, and 0.2 for the four cases. That is, from case 1 to case 4, the probability increases that the messenger can contact the sink before the transferred energy is drained. As shown in Fig. 7, the energy charging and content transferring rates as well as the expected utility consistently increase.

rates are always larger than the content transferring rates of the same scheme, as shown in Fig. 5(b). This is because the probability of successful energy transfer from a charger is 0.9. By contrast, the action of content transfer consumes one unit of energy. Thus, the charging rate is required to be higher than the transfer rate.

As shown in Figs. 5(a) and (b), although the greedy scheme may achieve a high expected utility (between the results of FAST and PIA), a very low energy charging rate, and consequently a very low content transferring rate are observed. Thus, the greedy scheme is a conservative scheme that constricts the action to transfer a content which may not be beneficial to the content source.

E. Impacts of Energy Prices and Successful Content Delivery Probabilities

We first evaluate the system performance under different energy prices. In particular, we vary the energy prices at the chargers. We consider nine energy price levels as shown in Table II. The energy price of each charger increases ascendingly according to the level. Fig. 6(a) shows that as the energy prices increase, the expected utility decreases due to the cost of energy charging.

F. Discussions on Special System Scenarios

Figs. 8(a) and (b) show the case that the content source can only encounter one charger and one messenger in the system. As shown in Fig. 8(a), the successful content delivery probability of the messenger is varied from 0.0 to 1.0. When the delivery probability is 0.0, the only messenger in the system fails to deliver a content. In this case, the content source halts the charging and content transferring processes.
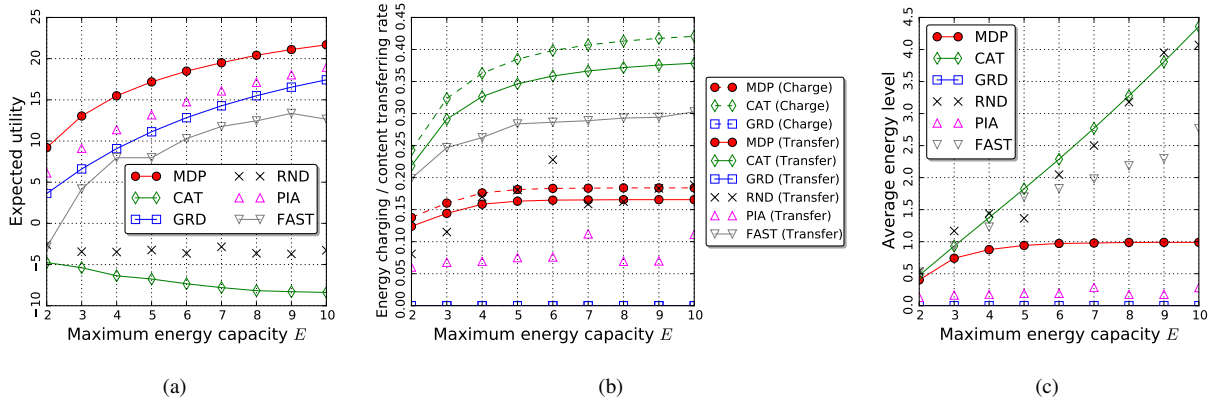
Fig. 5: Impacts of the maximum battery capacity E of the content source to (a) expected utility, (b) energy charging and content transferring rates, and (c) the average energy level in the content source.
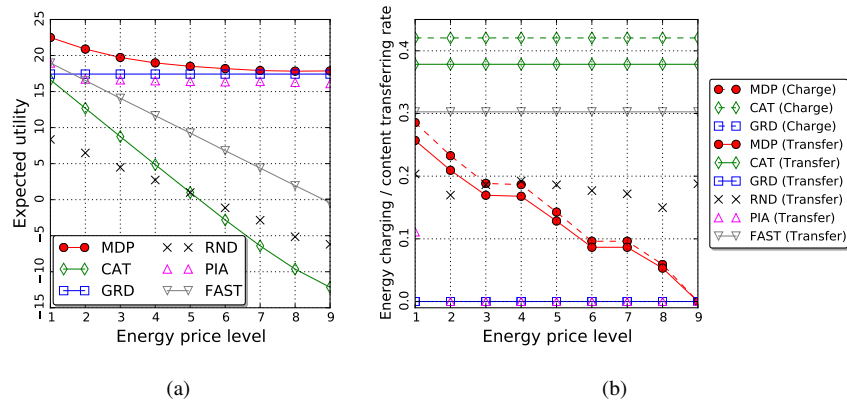

Fig. 6: Impacts of the energy price level of the content source to (a) expected utility, and (b) energy charging and content transferring rates.
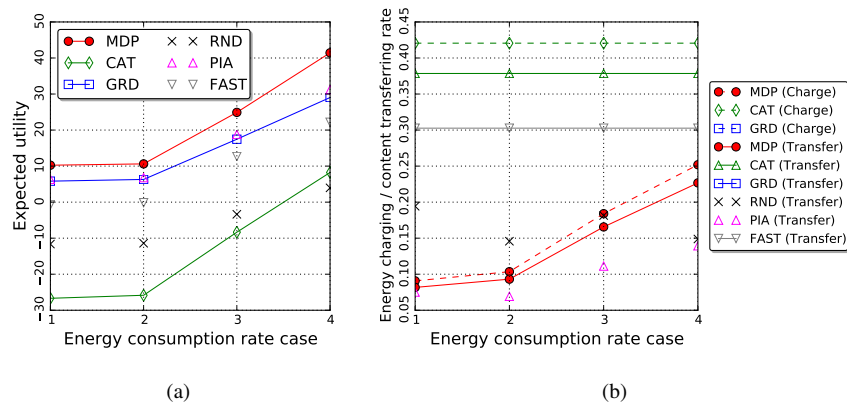

Fig. 7: Impacts of the messenger-sink contact probability level to (a) expected utility, and (b) energy charging and content transferring rates.

Similarly, in Fig. 8(b), the energy charging price asked by the only charger increases from 0 to 10. It is shown in the figure that the charging rate of the content source decreases as the energy price increases. The energy charging and content transferring rates approach 0 when the only charger in the system charges unacceptable prices which the content source cannot afford.

Fig. 8(c) shows the cases where the charging prices at the chargers are all 0, or the profit to the content source by successful delivery is 0. The four experimental cases are the same as in Fig. 7. As shown in Fig. 8(c), the content source always charges given the energy charging prices at the charger is free. However, when the content source cannot gain any profit by sending contents (i.e., the delivery profit always equals 0), the content source will never charge, and thus the content transferring rate is also 0.

VIII. CONCLUSION

In this work, we have proposed a mobile publish-subscribe network where a content source transfers content as well as energy to content messengers for delivering the content to the destination. An MDP has been formulated and solved to help the content source to optimally charge energy and transfer contents. Moreover, we have proved that the optimal energy charging and content transferring policy obtained from the MDP model is a threshold policy. This structural result of the threshold policy can simplify the decision making and solution method of the MDP. Based on this fact, we have introduced a partial iteration algorithm to approximate the optimal policy but with lower complexity. The numerical results have shown that the MDP-based scheme outperforms the baseline schemes.

APPENDIX A
PROOF OF THRESHOLD POLICY IN QUEUE STATE $\mathcal{Q}$

Proof: We employ the supermodularity property of the utility function $H(\mathcal{S}|\mathcal{A})$ to prove the structure of a threshold in an optimal policy with respect to the queue state $\mathcal{Q}$. That is, the following inequality must hold as a sufficient condition for the threshold policy:

$$\underbrace{[H(\mathcal{C},\mathcal{E},\mathcal{Q}+1|\mathcal{A}=2) - H(\mathcal{C},\mathcal{E},\mathcal{Q}|\mathcal{A}=2)]}_{\text{Term L}}$$
$$\geq \underbrace{[H(\mathcal{C},\mathcal{E},\mathcal{Q}+1|\mathcal{A}=0) - H(\mathcal{C},\mathcal{E},\mathcal{Q}|\mathcal{A}=0)]}_{\text{Term R}}. \quad (15)$$

We discuss the supermodularity in different cases of the queue state $\mathcal{Q}$:

- Boundary condition cases: $\mathcal{Q} = 0$, as well as $\mathcal{Q}+1 = Q$ and,
- Non-boundary condition cases: $0 < \mathcal{Q} < Q$ and $0 < \mathcal{Q}+1 < Q$.

When the content source does not have enough energy to transfer a content, $\mathcal{A} = 0$ is always taken. In this case, the threshold exists.

When there is enough energy for content transferring, i.e., $\mathcal{E} \geq \delta$, the following cases happen. Firstly, for the non-boundary case, equation (16) holds:

$$\begin{aligned} &H(\mathcal{C},\mathcal{E},\mathcal{Q}+1|\mathcal{A}=2) - H(\mathcal{C},\mathcal{E},\mathcal{Q}|\mathcal{A}=2)\\ =\ &F(\mathcal{C},\mathcal{E},\mathcal{Q}+1|\mathcal{A}=2) - F(\mathcal{C},\mathcal{E},\mathcal{Q}|\mathcal{A}=2)\\ &+\beta\sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'} \sum_{k=0}^{Q-\mathcal{Q}} f^a_i(k)[U(\mathcal{C},\mathcal{E}-\delta,\mathcal{Q}+k)\\ &\hspace{3cm} -U(\mathcal{C},\mathcal{E}-\delta,\mathcal{Q}-1+k)], \quad (16)\\ =\ &-\omega_{hd}[\chi(\mathcal{Q}+1) - \chi(\mathcal{Q})]\\ &+\beta\sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'} \sum_{k=0}^{Q-\mathcal{Q}} f^a_i(k)[U(\mathcal{C},\mathcal{E}-\delta,\mathcal{Q}+k)\\ &\hspace{3cm} -U(\mathcal{C},\mathcal{E}-\delta,\mathcal{Q}-1+k)], \end{aligned}$$

and in the same manner, (17) holds:

$$\begin{aligned} &H(\mathcal{C},\mathcal{E},\mathcal{Q}+1|\mathcal{A}=0) - H(\mathcal{C},\mathcal{E},\mathcal{Q}|\mathcal{A}=0)\\ =\ &-\omega_{hd}[\chi(\mathcal{Q}+1) - \chi(\mathcal{Q})]\\ &+\beta\sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'} \sum_{k=0}^{Q-\mathcal{Q}-1} f^a_i(k)[U(\mathcal{C},\mathcal{E},\mathcal{Q}+1+k)\\ &\hspace{3cm} -U(\mathcal{C},\mathcal{E},\mathcal{Q}+k)]. \end{aligned} \quad (17)$$

Lemma 6: Given the holding cost of contents $\chi(\cdot) \equiv 0$, the following inequality holds:

$$U(\mathcal{C},\mathcal{E},\mathcal{Q}+1)-U(\mathcal{C},\mathcal{E},\mathcal{Q}) \geq U(\mathcal{C},\mathcal{E}+1,\mathcal{Q}+2)-U(\mathcal{C},\mathcal{E}+1,\mathcal{Q}+1). \quad (18)$$

According to *Lemma 6*, which is proved in Appendix B, the following inequality holds:

$$\begin{aligned} &U(\mathcal{C},\mathcal{E}-\delta,\mathcal{Q}+k) - U(\mathcal{C},\mathcal{E}-\delta,\mathcal{Q}-1+k)\\ \geq\ &U(\mathcal{C},\mathcal{E},\mathcal{Q}+1+k) - U(\mathcal{C},\mathcal{E},\mathcal{Q}+k), \quad k = 0,1,\ldots,Q-\mathcal{Q}-1. \end{aligned} \quad (19)$$

Lemma 7: Given the condition that the holding cost of contents $\chi(\cdot) \equiv 0$, the following inequality holds:

$$\begin{aligned} 0 \leq\ &U(\mathcal{C},\mathcal{E},\mathcal{Q}+1) - U(\mathcal{C},\mathcal{E},\mathcal{Q})\\ <\ &\omega_{tn}\pi^{MG}p_m(\mathcal{C}) - \omega_{hd}[\chi(\mathcal{Q}+1) - \chi(\mathcal{Q})] + (+\infty). \end{aligned} \quad (20)$$

According to *Lemma 7*, which is proved in Appendix C, the following inequality holds:

$$U(\mathcal{C},\mathcal{E}-\delta,Q) - U(\mathcal{C},\mathcal{E}-\delta,Q-1) \geq 0. \quad (21)$$

Substitute (19) and (21) into (16) and (17), the condition in (15) holds, and the supermodularity in (15) is proved for this case.

For the special case of $\mathcal{Q} = 0$, the Term L in (15) is expressed as follows:

$$\begin{aligned} &H(\mathcal{C},\mathcal{E},1|\mathcal{A}=2) - H(\mathcal{C},\mathcal{E},0|\mathcal{A}=2)\\ =\ &F(\mathcal{C},\mathcal{E},1|\mathcal{A}=2) - F(\mathcal{C},\mathcal{E},0|\mathcal{A}=2)\\ &+\beta\sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'} \sum_{k=0}^{Q} f^a_i(k)[U(\mathcal{C},\mathcal{E}-\delta,k) - U(\mathcal{C},\mathcal{E}-\delta,k)],\\ =\ &\omega_{tn}\pi^{MG}p_m(\mathcal{C}) - \omega_{hd}[\chi(1) - \chi(0)] + (+\infty). \end{aligned} \quad (22)$$

For the Term R in (15), according to the upper bound property as given in Lemma 7, i.e., $U(\mathcal{C},\mathcal{E},1+k) - U(\mathcal{C},\mathcal{E},k) \leq P+C+(+\infty)$, for $k = 0,1,\ldots,|Q-1|$. The following chain of inequalities holds:

$$\begin{aligned} &H(\mathcal{C},\mathcal{E},1|\mathcal{A}=0) - H(\mathcal{C},\mathcal{E},0|\mathcal{A}=0)\\ =\ &F(\mathcal{C},\mathcal{E},1|\mathcal{A}=0) - F(\mathcal{C},\mathcal{E},0|\mathcal{A}=0)\\ &+\beta\sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'} \sum_{k=0}^{Q-1} f^a_i(k)[U(\mathcal{C},\mathcal{E},1+k) - U(\mathcal{C},\mathcal{E},k)],\\ \leq\ &-\omega_{hd}[\chi(1) - \chi(0)] + \beta \cdot (+\infty). \end{aligned} \quad (23)$$

That is, from (22) and (23), we have Term L $\geq$ Term R. The supermodularity in (15) exists in the boundary condition
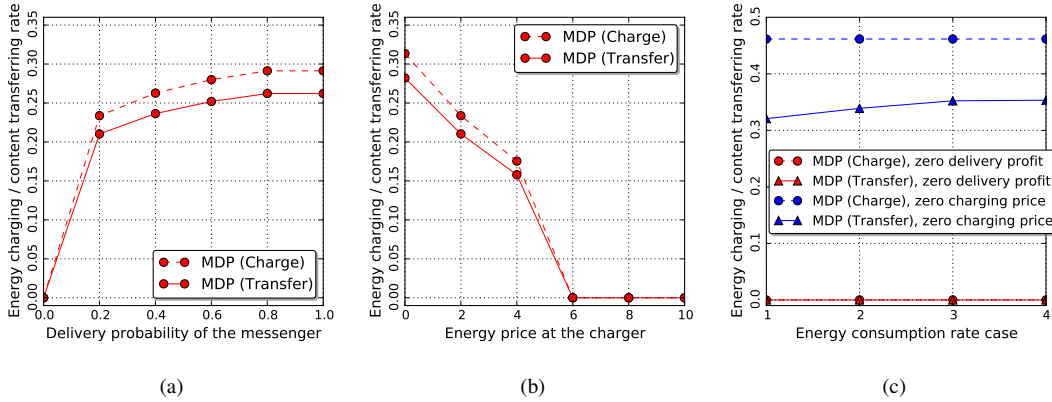
Fig. 8: Actions of the content source in special system scenarios: (a) One charger and one messenger in the system only with (a) different energy prices at the charger, and (b) different content delivery probability of the messenger. (c) Zero delivery profit or zero charging price.

case $\mathcal{Q} = 0$. In the same manner, the boundary condition case $\mathcal{Q} = Q - 1$ can be proved. The supermodularity property of the utility function $H(\mathcal{S}|\mathcal{A})$ is thus proved. ∎

APPENDIX B
PROOF OF *Lemma* 6

Proof: The value iteration algorithm is employed when solving the Bellman equation as in (11)-(13). In each iteration, the value of the optimal utility function $U(\mathcal{S})$ is updated for all the system states $\mathcal{S}$. We denote the nth iteration of $U(\mathcal{S})$ and $H(\mathcal{S}|\mathcal{A})$ to be $U_n(\mathcal{S})$ and $H_n(\mathcal{S}|\mathcal{A})$, respectively.

Let a_1, a_2, a_3, and a_4 be the optimal actions at the system states $(\mathcal{C}, \mathcal{E}, \mathcal{Q}+1)$, $(\mathcal{C}, \mathcal{E}, \mathcal{Q})$, $(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+2)$, and $(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+1)$, respectively, i.e.,

$$U_n(\mathcal{C}, \mathcal{E}, \mathcal{Q}+1) \geq H_n(\mathcal{C}, \mathcal{E}, \mathcal{Q}+1|\mathcal{A}=a_2), \quad (24)$$
$$U_n(\mathcal{C}, \mathcal{E}, \mathcal{Q}) = H_n(\mathcal{C}, \mathcal{E}, \mathcal{Q}|\mathcal{A}=a_2), \quad (25)$$
$$U_n(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+2) = H_n(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+2|\mathcal{A}=a_3), \quad (26)$$
$$U_n(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+1) \geq H_n(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+1|\mathcal{A}=a_3). \quad (27)$$

The two inequalities in (24) and (27) hold because of optimality. From (24)-(27), the following inequality can be derived:

$$
\begin{aligned}
&[U_n(\mathcal{C}, \mathcal{E}, \mathcal{Q}+1) - U_n(\mathcal{C}, \mathcal{E}, \mathcal{Q})] \\
&\quad -[U_n(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+2) - U_n(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+1)] \\
\geq &\underbrace{[H_n(\mathcal{C}, \mathcal{E}, \mathcal{Q}+1|\mathcal{A}=a_2) - H_n(\mathcal{C}, \mathcal{E}, \mathcal{Q}|\mathcal{A}=a_2)]}_{\text{Term A}} \\
&- \underbrace{[H_n(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+2|\mathcal{A}=a_3) - H_n(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+1|\mathcal{A}=a_3)]}_{\text{Term B}}.
\end{aligned}
$$
$$(28)$$

For simplicity of notation, we denote $\Delta^2 U_n = [U_n(\mathcal{C}, \mathcal{E}, \mathcal{Q}+1) - U_n(\mathcal{C}, \mathcal{E}, \mathcal{Q})] - [U_n(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+2) - U_n(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+1)]$.

Induction is employed to prove Lemma 6. We firstly discuss the non-boundary cases, where $\mathcal{Q} + 2 \leq Q$ and $\mathcal{Q} > 0$.

$$
\begin{aligned}
&\text{Term A} \\
=\ & H_{n+1}(\mathcal{C}, \mathcal{E}, \mathcal{Q}+1|\mathcal{A}=a_2) - H_{n+1}(\mathcal{C}, \mathcal{E}, \mathcal{Q}|\mathcal{A}=a_2), \\
=\ & I_{a_2=2} \cdot \Big\{ [F(\mathcal{C}, \mathcal{E}, \mathcal{Q}+1|\mathcal{A}=2) - F(\mathcal{C}, \mathcal{E}, \mathcal{Q}|\mathcal{A}=2)] \\
& + \beta \sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'} \sum_{k=0}^{Q-\mathcal{Q}} f^a_i(k) \big[U_n(\mathcal{C}, \mathcal{E}-\delta, \mathcal{Q}+k) \\
&\hspace{3cm} - U_n(\mathcal{C}, \mathcal{E}-\delta, \mathcal{Q}-1+k)\big] \Big\} \\
& + (1-I_{a_2=2}) \cdot \Big\{ [F(\mathcal{C}, \mathcal{E}, \mathcal{Q}+1|\mathcal{A}=0) - F(\mathcal{C}, \mathcal{E}, \mathcal{Q}|\mathcal{A}=0)] \\
& + \beta \sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'} \sum_{k=0}^{Q-\mathcal{Q}-1} f^a_i(k) \big[U_n(\mathcal{C}, \mathcal{E}, \mathcal{Q}+1+k) \\
&\hspace{3cm} - U_n(\mathcal{C}, \mathcal{E}, \mathcal{Q}+k)\big] \Big\},
\end{aligned}
$$
$$(29)$$

and

$$
\begin{aligned}
&\text{Term B} \\
=\ & H_{n+1}(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+2|\mathcal{A}=a_3) - H_{n+1}(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+1|\mathcal{A}=a_3), \\
=\ & I_{a_3=2} \cdot \Big\{ [F(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+2|\mathcal{A}=2) - F(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+1|\mathcal{A}=2)] \\
& + \beta \sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'} \sum_{k=0}^{Q-\mathcal{Q}-1} f^a_i(k) \big[U_n(\mathcal{C}, \mathcal{E}, \mathcal{Q}+1+k) \\
&\hspace{3cm} - U_n(\mathcal{C}, \mathcal{E}, \mathcal{Q}+k)\big] \Big\} \\
& + (1-I_{a_3=2}) \cdot \Big\{ [F(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+2|\mathcal{A}=0) \\
&\hspace{2cm} - F(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+1|\mathcal{A}=0)] \\
& + \beta \sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'} \sum_{k=0}^{Q-\mathcal{Q}-2} f^a_i(k) \big[U_n(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+2+k) \\
&\hspace{3cm} - U_n(\mathcal{C}, \mathcal{E}+1, \mathcal{Q}+1+k)\big] \Big\}.
\end{aligned}
$$
$$(30)$$

For the initial step $n = 0$, since the value iteration will converge [41] given an arbitrary initial value, we let $U_0(\mathcal{S})$ be 0 for all the states $\mathcal{S}$. From (29) and (30), Term A = Term B = 0, as shown in (28), $\Delta^2 U_n$ = Term A − Term B = 0. Thus, $\Delta^2 U_n \geq 0$ holds.

Induction step: For the nth step, we suppose that $\Delta^2 U_n \geq 0$ already holds. The following inequalities hold for the $(n+1)^{th}$ step, given $\Delta^2 U_n \geq 0$ and Lemma 7, the following chain of

inequalities holds:

$$
\begin{aligned}
&\text{Term A}\\
=\ & H_{n+1}(\mathcal{C},\mathcal{E},\mathcal{Q}+1|\mathcal{A}=a_2) - H_{n+1}(\mathcal{C},\mathcal{E},\mathcal{Q}|\mathcal{A}=a_2),\\
\geq\ & \beta \sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'} \sum_{k=0}^{Q-\mathcal{Q}-1} f_i^a(k)\big[U_n(\mathcal{C},\mathcal{E}-\delta,\mathcal{Q}+1+k)\\
&\hspace{4cm} -U_n(\mathcal{C},\mathcal{E}-\delta,\mathcal{Q}+k)\big],\\
\geq\ & H_{n+1}(\mathcal{C},\mathcal{E}+1,\mathcal{Q}+2|\mathcal{A}=a_3)\\
&\hspace{0.5cm} -H_{n+1}(\mathcal{C},\mathcal{E}+1,\mathcal{Q}+1|\mathcal{A}=a_3),\\
=\ & \text{Term B}.
\end{aligned}
\tag{31}
$$

Therefore, $\Delta^2 U_{n+1} = \text{Term A} - \text{Term B} \geq 0$, the $(n+1)^{th}$ step is proved.

With Lemma 7, the boundary condition cases when $\mathcal{Q}=0$ and $\mathcal{Q}+2=Q$ can be proved in the same manner. ∎

APPENDIX C
PROOF OF *Lemma* 7

Proof: The proof can be conducted by using the induction method similar to that of *Lemma* 6. We start from

$$
\begin{aligned}
& U_{n+1}(\mathcal{C},\mathcal{E},\mathcal{Q}+1) - U_{n+1}(\mathcal{C},\mathcal{E},\mathcal{Q})\\
\geq\ & H_n(\mathcal{C},\mathcal{E},\mathcal{Q}+1|\mathcal{A}=a_2) - U_n(\mathcal{C},\mathcal{E},\mathcal{Q}|\mathcal{A}=a_2),\\
=\ & I_{a_2=2}\cdot\Big\{\big[F(\mathcal{C},\mathcal{E},\mathcal{Q}+1|\mathcal{A}=2) - F(\mathcal{C},\mathcal{E},\mathcal{Q}|\mathcal{A}=2)\big]\\
&+\beta\sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'}\sum_{k=0}^{Q-\mathcal{Q}} f_i^a(k)\big[U_n(\mathcal{C},\mathcal{E}-\delta,\mathcal{Q}+k)\\
&\hspace{3cm}-U_n(\mathcal{C},\mathcal{E}-\delta,\mathcal{Q}-1+k)\big]\Big\}\\
&+(1-I_{a_2=2})\cdot\Big\{\big[F(\mathcal{C},\mathcal{E},\mathcal{Q}+1|\mathcal{A}=0) - F(\mathcal{C},\mathcal{E},\mathcal{Q}|\mathcal{A}=0)\big]\\
&+\beta\sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'}\sum_{k=0}^{Q-\mathcal{Q}-1} f_i^a(k)\big[U_n(\mathcal{C},\mathcal{E},\mathcal{Q}+1+k)\\
&\hspace{3cm}-U_n(\mathcal{C},\mathcal{E},\mathcal{Q}+k)\big]\Big\},
\end{aligned}
\tag{32}
$$

and

$$
\begin{aligned}
& U_{n+1}(\mathcal{C},\mathcal{E},\mathcal{Q}+1) - U_{n+1}(\mathcal{C},\mathcal{E},\mathcal{Q})\\
\leq\ & H_n(\mathcal{C},\mathcal{E},\mathcal{Q}+1|\mathcal{A}=a_1) - U_n(\mathcal{C},\mathcal{E},\mathcal{Q}|\mathcal{A}=a_1),\\
=\ & I_{a_1=2}\cdot\Big\{\big[F(\mathcal{C},\mathcal{E},\mathcal{Q}+1|\mathcal{A}=2) - F(\mathcal{C},\mathcal{E},\mathcal{Q}|\mathcal{A}=2)\big]\\
&+\beta\sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'}\sum_{k=0}^{Q-\mathcal{Q}} f_i^a(k)\big[U_n(\mathcal{C},\mathcal{E}-\delta,\mathcal{Q}+k)\\
&\hspace{3cm}-U_n(\mathcal{C},\mathcal{E}-\delta,\mathcal{Q}-1+k)\big]\Big\}\\
&+(1-I_{a_1=2})\cdot\Big\{\big[F(\mathcal{C},\mathcal{E},\mathcal{Q}+1|\mathcal{A}=0) - F(\mathcal{C},\mathcal{E},\mathcal{Q}|\mathcal{A}=0)\big]\\
&+\beta\sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'}\sum_{k=0}^{Q-\mathcal{Q}-1} f_i^a(k)\big[U_n(\mathcal{C},\mathcal{E},\mathcal{Q}+1+k)\\
&\hspace{3cm}-U_n(\mathcal{C},\mathcal{E},\mathcal{Q}+k)\big]\Big\}.
\end{aligned}
\tag{33}
$$

For the non-boundary cases of $\mathcal{Q}+1 < Q$ and $\mathcal{Q} \geq 0$, *Lemma* 7 can be proved by induction as in Appendix B.

Then we prove the boundary case $\mathcal{Q} = Q - 1$. Firstly, the upper bound of $U(\mathcal{C},\mathcal{E},Q) - U_{n+1}(\mathcal{C},\mathcal{E},Q-1)$ can be proved with the aforementioned induction method. Secondly, the lower bound of $U(\mathcal{C},\mathcal{E},\mathcal{Q}+1) - U(\mathcal{C},\mathcal{E},\mathcal{Q}) \geq 0$ can be proved, given the holding cost of contents $\chi(\cdot) \equiv 0$, as

follows:

$$
\begin{aligned}
& U_{n+1}(\mathcal{C},\mathcal{E},Q) - U_{n+1}(\mathcal{C},\mathcal{E},Q-1)\\
\geq\ & H_n(\mathcal{C},\mathcal{E},Q|\mathcal{A}=a_2) - U_n(\mathcal{C},\mathcal{E},Q-1|\mathcal{A}=a_2),\\
=\ & I_{a_1=2}\cdot\Big\{\big[F(\mathcal{C},\mathcal{E},Q|\mathcal{A}=2) - F(\mathcal{C},\mathcal{E},Q-1|\mathcal{A}=2)\big]\\
&+\beta\sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'}\sum_{k=0}^{1} f_i^a(k)\big[U_n(\mathcal{C},\mathcal{E}-\delta,Q-1+k)\\
&\hspace{3cm}-U_n(\mathcal{C},\mathcal{E}-\delta,Q-2+k)\big]\Big\}\\
&+(1-I_{a_1=2})\cdot\Big\{\big[F(\mathcal{C},\mathcal{E},Q|\mathcal{A}=0) - F(\mathcal{C},\mathcal{E},Q-1|\mathcal{A}=0)\big]\\
&+\beta\sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'} f_i^a(0)\big[U_n(\mathcal{C},\mathcal{E},Q) - U_n(\mathcal{C},\mathcal{E},Q-1)\big]\Big\},\\
\geq\ & -\omega_{hd}[\chi(\mathcal{Q}) - \chi(\mathcal{Q}-1)]\\
&+\beta\sum_{\mathcal{C}'\in\mathbb{C}} p^C_{\mathcal{C},\mathcal{C}'} f_i^a(0)\big[U_n(\mathcal{C},\mathcal{E},Q) - U_n(\mathcal{C},\mathcal{E},Q-1)\big],\\
\geq\ & -\omega_{hd}[\chi(\mathcal{Q}) - \chi(\mathcal{Q}-1)].
\end{aligned}
\tag{34}
$$

When $\chi(\cdot) \equiv 0$, i.e., the holding cost of contents is zero or neglected by the content source, *Lemma* 7 also holds for the boundary condition $\mathcal{Q}+1 = Q$. In summary, *Lemma* 7 holds for all system states $\mathcal{S} \in \mathbb{S}$. ∎

REFERENCES

[1] N. Kayastha, D. Niyato, P. Wang, and E. Hossain, "Applications, architectures, and protocol design issues for mobile social networks: A Survey," *Proceedings of the IEEE*, vol. 99, no. 12, pp. 2130-2158, December 2011.

[2] P. T. Eugster, P. A. Felber, R. Guerraoui, and A.-M. Kermarrec, "The many faces of publish/subscribe", *ACM Computing Surveys (CSUR)*, vol. 35, no. 2, pp. 114-131, June 2003.

[3] Q. Li, S. Zhu, and G. Cao, "Routing in socially selfish delay tolerant networks," in *Proc. of IEEE INFOCOM 2010*, San Diego, CA, USA, 14-19 March 2010.

[4] E. Yoneki, P. Hui, S. Chan, and J. Crowcroft, "A socio-aware overlay for publish/subscribe communication in delay tolerant networks," in *Proc. of the 10th ACM Symposium on Modeling, analysis, and simulation of wireless and mobile systems (MSWiM '07)*, pp. 225-234, Chania, Crete Island, Greece, October 2007.

[5] U. Lee, J. Lee, J.-S. Park, and M. Gerla, "FleaNet: A virtual Mmarket place on vehicular networks," in *IEEE Transactions on Vehicular Technology*, vol. 59, no. 1, pp. 344-355, January 2010.

[6] I. Leontiadis, P. Costa, and C. Mascolo, "A hybrid approach for content-based publish/subscribe in vehicular networks", in *Pervasive and Mobile Computing*, vol. 5, no. 6, pp. 697-713, December 2009.

[7] M. Granovetter, "The strength of weak ties," *The American Journal of Sociology*, vol. 78, no. 6, pp. 1360-1380, May 1973.

[8] Y. Li, Y. Jiang, D. Jin, L. Su, L. Zeng, and D. Wu, "Energy-efficient optimal opportunistic forwarding for delay-tolerant networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 9, pp. 4500-4512, November 2010.

[9] O. Wolfson, B. Xu, and R. M. Tanner, "Mobile peer-to-peer data dissemination with resource constraints," in *Proc. of 2007 International Conference on Mobile Data Management*, pp. 16-23, Mannheim, Germany, May 2007.

[10] O. Orhan and E. Erkip, "Optimal transmission policies for energy harvesting two-hop networks," in *Proc. of 2012 46th Annual Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, USA, 21-23 March 2012.

[11] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," to appear in *IEEE Communications Surveys and Tutorials*.

[12] Powercast [Online], http://powercastco.com .

[13] H. Nishimoto, Y. Kawahara, and T. Asami, "Prototype implementation of ambient RF energy harvesting wireless sensor networks," in *Proc. of 2010 IEEE Sensors*, pp. 1282-1287, Kona, HI, USA, 1-4 November 2010.

[14] D. van Wageningen and T. Staring, "The Qi wireless power standard," in *Proc. of 2010 14th International Power Electronics and Motion Control Conference (EPE/PEMC)*, pp. S15-25-S15-32, Ohrid, Macedonia, 6-8 September 2010.

[15] C. Mikeka and H. Arai, "Design issues in radio frequency energy harvesting system," *Sustainable Energy Harvesting Technologies - Past, Present and Future*, Chapter 10, December 2011.

[16] J. A. Paradiso and T. Starner, "Energy scavenging for mobile and wireless electronics," *IEEE Pervasive Computing*, vol. 4, no. 1, pp. 18-27, January-March, 2005.

[17] K. Gudan, S. Chemishkian, J. J. Hull, M. S. Reynolds, and S. Thomas, "Feasibility of wireless sensors using ambient 2.4 GHz RF energy," in *Proc. of 2012 IEEE Sensors*, Taipei, October 2012.

[18] H. Ostaffe, "Power out of thin air: Ambient RF energy harvesting for wireless Sensors," June 2010, [Online], http://powercastco.com/PDF/Power-Out-of-Thin-Air.pdf .

[19] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 1989-2001, May 2013.

[20] I. Krikidis, "Simultaneous information and energy transfer in large-scale networks with/without relaying," in *IEEE Transactions on Communications*, vol. 62, no. 3, pp. 900-912, March 2014.

[21] X. Chen, C. Yuen, and Z. Zhang, "Wireless energy and information transfer tradeoff for limited-feedback multiantenna systems with energy beamforming," in *IEEE Transactions on Vehicular Technology*, vol. 63, no. 1, pp. 407-412, January 2014.

[22] R. Bellman, "A Markovian decision process," *Journal of Mathematics and Mechanics*, vol. 6, no. 5, pp. 679-684, 1957.

[23] D. Niyato and P. Wang, "Optimization of the mobile router and traffic sources in vehicular delay-tolerant network," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 9, pp. 5095-5104, November 2009.

[24] D. Niyato, P. Wang, E. Hossain, and Y. Li,"Optimal content transmission policy in publish-subscribe mobile social networks," in *Proc. of IEEE GLOBECOM 2010*, 6-10 December, 2010.

[25] A. Sultan, "Sensing and transmit energy optimization for an energy harvesting cognitive radio," *IEEE Wireless Communications Letters*, vol. 1, no. 5, pp. 500-503, October 2012.

[26] V. Sharma, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 1326-1336, April, 2010.

[27] A. Seyedi and B. Sikdar, "Energy efficient transmission strategies for body sensor networks with energy harvesting," *IEEE Transactions on Communications*, vol. 58, no. 7, pp. 2116-2126, July, 2010.

[28] D. Niyato, P. Wang, D. I. Kim, and Z. Han, "Content messenger selection and wireless energy transfer policy in mobile social networks," in *Proceedings of International Conference on Communications (ICC)*, London, UK, June 2015.

[29] M. F. Neuts, *Matrix-geometric solutions in stochastic models: An algorithmic approach*, Dover Publications, January 1995.

[30] A. H. Zahran and B. Liang, "A generic framework for mobility modeling and performance analysis in next-generation heterogeneous wireless networks," *IEEE Communications Magazine*, vol. 45, no. 9, pp. 92-99, September 2007.

[31] D. N. Alparslan, and K. Sohraby, "A generalized random mobility model for wireless Ad hoc networks and its analysis: One-dimensional case," in *IEEE/ACM Transactions on Networking*, vol. 15, no. 3, pp. 602-615, June 2007.

[32] J. A. Whittaker and M. G.Thomason, "A Markov Chain Model for Statistical Software Testing," IEEE Transactions on Software Engineering, vol. 20, no. 10, pp. 812-824, October 1994.

[33] T. Laursen, N. B. Pedersen, J. J. Nielsen, and T. K. Madsen, "Hidden Markov Model based mobility learning fo improving indoor tracking of mobile users," in *Proc. of 2012 9th Workshop on Positioning Navigation and Communication (WPNC)*, pp. 100-104, 15-16 March 2012.

[34] Q. Yuan, I. Cardei, and J. Wu, "An efficient prediction-based routing in disruption-tolerant networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 1, pp. 19-31, January 2012.

[35] A. Picu, T. Spyropoulos, and T. Hossmann, "An analysis of the information spreading delay in heterogeneous mobility DTNs," in *Proceedings of IEEE International Symposium on World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, San Francisco, CA, USA, June 2012.

[36] M. H. Ngo and V. Krishnamurthy, "Optimality of threshold policies for transmission scheduling in correlated fading channels," *IEEE Transactions on Communications*, vol. 57, no. 8, pp. 2474-2483, August 2009.

[37] M. L. Littman, T. L. Dean, and L. P. Kaelbling, "On the complexity of solving Markov decision problems," in *Proc. of Eleventh conference on Uncertainty in Artificial Intelligence (UAI'95)*, Montréal, Québec, Canada, pp. 394-402, 1995.

[38] 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.

[39] Y. Zhang, D. Niyato, and P. Wang, "Offloading in mobile cloudlet systems with intermittent connectivity," in *IEEE Transactions on Mobile Computing*, vol. 14, no. 12, pp. 2516-2529, December 2015.

[40] D. M. Topkis, *Supermodularity and complementarity*, Princeton University Press, 1998.

[41] T. Kamihigashi, "Elementary results on solutions to the Bellman equation of dynamic programming: existence, uniqueness, and convergence," Discussion Paper Series No. DP2012-31, RIEB, Kobe University, 2012.
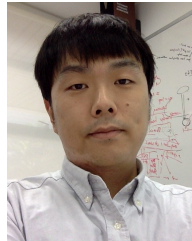
Yang Zhang (M'11) is currently an Assistant Professor in the School of Computer Science and Technology, Wuhan University of Technology. He received the Ph.D. degree from the School of Computer Science and Engineering, Nanyang Technological University (NTU), Singapore, in 2015. He obtained his B.Eng. and M.Eng. degrees from Beihang University (BUAA), Beijing, China, in 2008 and 2011, respectively.


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 Mongkut's Institute of Technology Ladkrabang (KMITL) 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.

Ping Wang (M'08-SM'15) received the PhD degree in electrical engineering from University of Waterloo, Canada, in 2008. Currently she is an Associate Professor in the School of Computer Engineering, Nanyang Technological University, Singapore. Her current research interests include resource allocation in multimedia wireless networks, cloud computing, and smart grid. She was a corecipient of the Best Paper Award from IEEE Wireless Communications and Networking Conference (WCNC) 2012 and IEEE International Conference on Communications (ICC) 2007. She is an Editor of IEEE Transactions on Wireless Communications, EURASIP Journal on Wireless Communications and Networking, and International Journal of Ultra Wideband Communications and Systems.
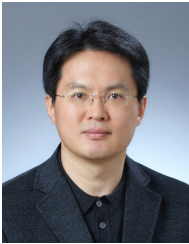
Dong In Kim (S'89-M'91-SM'02) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 1990. He was a tenured Professor with the School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada. Since 2007, he has been with Sungkyunkwan University (SKKU), Suwon, Korea, where he is currently a Professor with the College of Information and Communication Engineering. Dr. Kim has served as an Editor and a Founding Area Editor of Cross-Layer Design and Optimization for the IEEE Transactions on Wireless Communications from 2002 to 2011. From 2008 to 2011, he served as the Co-Editor-in-Chief for the Journal of Communications and Networks. He has 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 in Wireless Communication. He is a first recipient of the NRF of Korea Engineering Research Center (ERC) in Wireless Communications for Energy Harvesting Wireless Communications (2014-2021).

Zhu Han (S'01M'04-SM'09-F'14) received the B.S. degree in electronic engineering from Tsinghua University, in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, in 1999 and 2003, respectively.

From 2000 to 2002, he was an R&D Engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a Research Associate at the University of Maryland. From 2006 to 2008, he was an assistant professor at Boise State University, Idaho. Currently, he is a Professor in the Electrical and Computer Engineering Department as well as in the Computer Science Department at the University of Houston, Texas. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid. Dr. Han received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, IEEE Leonard G. Abraham Prize in the field of Communications Systems (best paper award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. Currently, Dr. Han is an IEEE Communications Society Distinguished Lecturer.