Markov Decision Model for Data Offloading in Mobile Cloud Computing

Dongqing Liu, Lyes Khoukhi, and Abdelhakim Hafid

Abstract

Cellular network is facing severe traffic overload problem caused by phenomenal growth of mobile data. Offloading part of the mobile data traffic from cellular network to alternative networks is a promising solution. In this paper, we study mobile data offloading problem under the architecture of mobile cloud computing (MCC), where mobile data can be delivered by WiFi network and device-to-device (D2D) communication. In order to minimize the overall cost for data delivery task, it is crucial to reduce cellular network usage while satisfying delay requirements. In our proposed model, we formulate the data offloading task as a finite horizon Markov Decision Process (FHMDP). We first propose a hybrid offloading algorithm for mobile data with different delay requirements. Moreover, we establish the sufficient conditions for the existence of threshold policy. Then, we propose a monotone offloading algorithm based on threshold policy in order to reduce the computational complexity. The simulation results show that the proposed offloading approach can achieve minimal communication cost compared with other three offloading schemes.

Index Terms

Mobile data offloading, device-to-device communication, mobile cloud computing, Markov decision process.

I. Introduction

With the increase of the number of smart mobile devices and data heavy mobile applications, such as video streaming and cloud backup, global mobile data traffic has been growing dramatically in recent years. The global mobile traffic grew 74% in 2015, while mobile network

Dongqing Liu is with the Department of Computer Science and Operational Research, University of Montreal, QC, Canada and Environment and Autonomous Networks Lab (ERA), University of Technology of Troyes, France, e-mail: dongqing.liu@utt.fr

Lyes Khoukhi is with the Environment and Autonomous Networks Lab (ERA), University of Technology of Troyes, France, e-mail: lyes.khoukhi@utt.fr

Abdelhakim Hafid is with the Department of Computer Science and Operational Research, University of Montreal, QC, Canada, ahafid@iro.umontreal.ca

(cellular) connection speeds only grew 20% [1]. The growing speed of mobile traffic will push the current cellular network to the limit. The Quality of Experience (QoE) of mobile services will not be guaranteed without the high-speed and stable network connections. However, it is impractical to keep extending the current cellular network infrastructure to improve QoE, given the corresponding expensive investment. In order to cope with this problem, mobile data offloading technology can be an alternative solution. Mobile data offloading can opportunistically use alternative networks (e.g., WiFi network and D2D communnication) to reduce the network congestion.

Compared with data offloading, applications involving computation offloading usually are more delay-sensitive. This is because computation offloading includes two data delivery processes, i.e., uploading computation data and downloading computation results. In many cases, data offloading can improve these two processes by using alternative networks with higher data rates than cellular network, e.g., WiFi network [2]. Thus, data offloading can be used in MCC to improve the performance of computation offloading.

WiFi offloading is considered as a promising solution to reduce mobile data traffic in cellular network. WiFi Access Points (APs) can efficiently reduce cellular traffic [3, 4]. It is shown that about 65% of cellular traffic can be offloaded through WiFi APs[5]. Although WiFi APs can provide better data rate than cellular network, their coverage area is much smaller than cellular network [6, 7].

Another mobile offloading method, called opportunistic offloading, is based on D2D communication [8]. Opportunistic offloading uses the store-carry-forward strategy, where some mobile users can store data in the buffer (called mobile helpers, MHs), carry the data when they are moving, and forward the data to other mobile users (called mobile subscribers, MSs) [9, 10]. When mobile network operator (MNO) wants to deliver data to MSs, it can first send the data to MHs. Then, MHs will transmit data to MSs using opportunistic connections. With more than half a billion mobile devices and connections added in 2015 [1], D2D communication is becoming an important data delivery scheme. However, the data rate of D2D communication is low and the mobility patterns of MHs or MSs are difficult to predict.

In this paper, we propose two mobile data offloading schemes based on Finite Horizon Markov Decision Process (FHMDP). Our objective is to minimize the communication cost for delivering mobile data with different delay sensitivities through multiple wireless networks, i.e., cellular network, WiFi network and D2D communication.

A preliminary version of this work has been published in [11]. We extend this work by analyzing the threshold structures of optimal policy and proposing new monotone offloading algorithm. The main contributions of our paper can be summarized as follows:

- We propose a hybrid offloading model, where multiple wireless networks are used to transfer mobile data. MNO can minimize the total communication cost by selecting different networks.
- We formulate the data offloading problem in hybrid wireless networks as an FHMDP model, and propose an offloading algorithm that can support different delay requirements (i.e., loose and tight delay tolerant).
- We prove that there exit threshold structures in the optimal policy and propose a monotone offloading algorithm for generating monotone policy with lower computational complexity.
- The simulation results demonstrate that our proposed schemes achieve the lowest communication cost as compared with three offloading schemes.

The rest of this paper is organized as follows. Section II reviews the related work. Section III describes the system model. Section IV formulates the mobile data offloading problem as an FHMDP model. Section V proposes a hybrid offloading algorithm. Section VI establishes the sufficient conditions for the existence of threshold policy and proposes a monotone offloading algorithm based on threshold policy. Section VII evaluates the performance of the proposed offloading algorithms. Finally, Section VIII concludes the paper.

II. Related work

MCC can support mobile devices access cellular, WiFi and D2D networks [12]. Most existing contributions [13, 14] focus on mobile computation offloading problem, which means migrating resource-intensive computations from MS to cloud. In this paper, we consider how to implement mobile data delivery in MCC by employing offloading technology. In the following, we present a survey of prior work aiming to offload cellular traffic to other mobile networks, including WiFi network and D2D communication.

Several contributions have shown the benefits of offloading mobile data from cellular network to WiFi network. Song *et al.* [15] investigated offloading schemes for cellular and WLAN integrated networks. They considered the WLAN-first resource allocation scheme where WLAN connection is used whenever possible, in order to benefit from low cost and large bandwidth of WLAN. Siris *et al.* [16] investigated the methods for enhancing mobile data offloading from mobile networks to WiFi APs by using mobility prediction and prefetching techniques. They evaluated these methods in terms of offloading ratio, data transmission time and cache size when using prefetching. Cheng *et al.* [17] presented an analytical framework for offloading cellular traffic to WiFi network using queuing theory. They evaluated the offloading performance in terms of average service delay. Mehmeti *et al.* [18] evaluated the performance of on-the-spot mobile data offloading. They analyzed the performance improvement by WiFi-based offloading using queuing theory. Jung *et al.* [19] proposed a network-assisted user-centric WiFi-offloading model in a heterogeneous network; the objective was to maximize throughput for each MS by utilizing network information.

Other contributions have shown the possibility of offloading mobile data from cellular network to D2D network. The main idea is to transmit mobile data using opportunistic communication among MSs; this has been shown to provide significant wireless capacity gains. Vinicius *et al.* [20] proposed a multi-criteria decision-making framework for data offloading from 3G network to D2D network. The framework avoids changes in the infrastructure by employing only user knowledge to select MHs. It shows that delay tolerant applications can offload six-fold mobile data compared to delay sensitive applications. Sciancalepore *et al.* [21] considered data offloading in D2D network with heterogeneous node mobility patterns. They used an optimization method to minimize cellular network traffic while satisfying the applications' constraints. Filippo *et al.* [22] proposed a method, called DROiD, to control popular data distribution in D2D network; the aim was to minimize the usage of infrastructure resources. They did show that the proposed method can offload a significant amount of data from cellular network to D2D network under tight delivery delay constraints. Andreev *et al.* [23] investigated the offloading method from cellular network to D2D network. They demonstrated that assisted offloading of cellular user sessions into D2D links improves the degree of spatial reuse and reduces the impact of interference.

Since the coverage area of D2D network is flexible with the movement of MHs, it can help offload data when WiFi connections are not available, especially for transmitting small size data, due to the short connection time and low data rate. However, since the data rate of WiFi network is higher than that of D2D network and WiFi network is more stable than D2D network, WiFi based offloading generally outperforms D2D based offloading from MS's perspective [7]. Notice that D2D based offloading can offload significant mobile data from MNO's perspective, since the number of MHs can be quite large. In the simple case, WiFi APs can be considered as a special kind of MHs. Compared with MHs, WiFi APs are installed at some fixed locations and

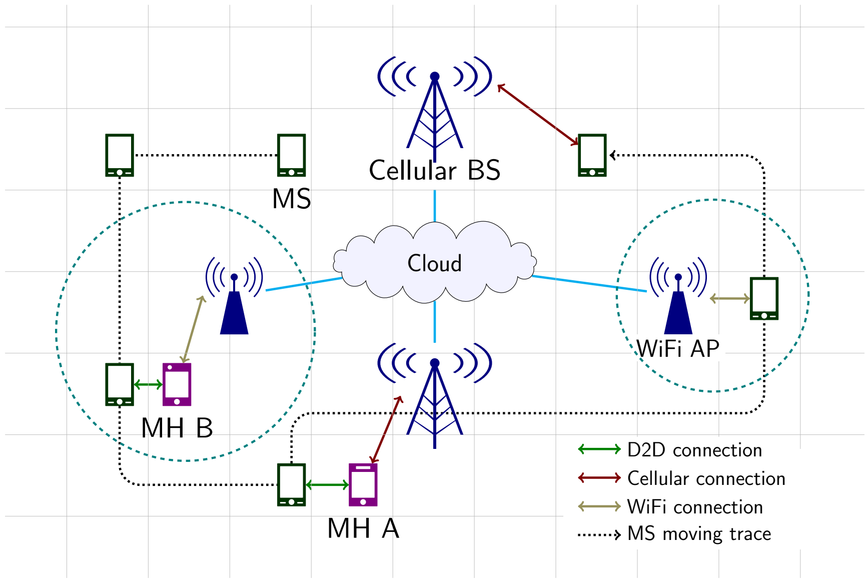


Fig. 1. The system model of mobile data offloading.

have more bandwidth.

We conclude that most existing contributions are based on WiFi offloading or opportunistic networks, without considering the combination of different mobile networks. In this paper, we consider a hybrid offloading model, where mobile data can be offloaded through WiFi offloading and D2D communication. Our objective is to minimize the overall cost for data delivery while satisfying delay requirements of different user types.

III. System Model

In this section, we present our system model to enhance data offloading in mobile cloud. In our model, we consider that mobile devices can access cloud services through multiple wireless networks, as shown in Fig. 1: (1) WiFi network. WiFi APs provide opportunistic WiFi communication (e.g., WLAN) for MS within its working coverage, and connect to distant cloud infrastructure through wired network; (2) Cellular network. Cellular base stations (BSs) provide seamless cellular communication (e.g., 4G) for MS, and connect to cloud through wired network; (3) D2D network. MHs (e.g., MH A and MH B in Fig. 1) provide opportunistic D2D communication for MS, and connect to nearby WiFi APs or Cellular BSs through WiFi or cellular communication.

In our data offloading system, mobile helpers (MHs) are chosen to work as data providers for mobile subscribers (MSs). Incentives for MHs to participate in data offloading can be provided by using some micro-payment scheme, or MNO can offer participants a reduced cost for the service. In this paper, we choose the micro-payment scheme [7, 9], where MHs can get rewards by participating in data offloading. The price for transmitting a data unit (i.e., χ_4) is set by MNO. MNO first announces the price to mobile users and then chooses MHs from those users who accept the price and are willing to participate in data offloading. It is worth noting that a full analysis of such process is not the focus of this paper.

The mobile data being received from cloud to MS is divided into a sequence of data units. The data units are predetermined by MNO. Delivering data means transmitting data of size K to MS before deadline D. K is the number of total data units and D is the maximum available time for data transmission. The data delivery is completed when non-transmitted data size k (i.e., k is the size of data that has not been received by MS) is zero before D. Conventionally, without WiFi APs and MHs, MS receives all mobile data through cellular communication. However, in our model, MS has an option to receive parts of the data through nearby WiFi or D2D communication, which may offer higher data rates and lower communication cost.

Upon arrival of a data delivery request from MS, MNO decides whether to transmit data by cellular network or offload it to WiFi and D2D networks according to data characteristics and network performance. The possibility of offloading depends on the delay characteristic (i.e. delay tolerant or not) of mobile data. If data is delay tolerant, MNO can defer data transmission to increase the possibility of offloading. Otherwise (i.e., data is delay sensitive), MNO will have less opportunities to offload mobile data from cellular network. Moreover, if data rates of WiFi and D2D networks are higher than that of cellular network, MNO can shorten the delivery time by cellular data offloading.

The main idea of data offloading is to use delay tolerance of mobile data and mobility of mobile users to seek opportunities to use WiFi and D2D networks.

We assume that a time slot T is long enough for MS to receive at least one data unit from cellular, WiFi or D2D network. An offloading decision (i.e., selecting a network) is made at the beginning of each time slot. The time when offloading decision is determined is denoted by d; it is called decision epoch. Thus, a network is selected at each decision epoch and will be the working network during the time slot.

At each decision epoch, MNO observes the current system states, i.e., the location of MS,

TABLE I
NOTATIONS

Symbol	Description
K	Total size of mobile data to be transmitted.
D	Total length of time for data transmission.
k	Size of mobile data that is not transmitted.
d	Decision epoch: time for making offloading decision.
$\mathcal{A}$	Set of transmission actions.
$\mathcal{U}$	Set of mobile user types.
a	Transmission action.
u	Mobile user type.
L	Total number of grids.
T	Length of one time slot.
μ	Stable factor: probability of MS staying at the same location in two sequential decision epochs.
ν_a^l	Data rate for action a at grid l.
χ_a	Unit price for transmitting data by action a.

the non-transmitted data size and the locations of available MHs. Based on the observed system state, MNO computes the communication cost for available networks. Then, MNO makes an offloading decision of either transmitting data using cellular network or offloading data to other network (i.e., WiFi or D2D network).

In this paper, we propose a Finite Horizon Markov Decision Process (FHMDP) to formulate this problem, with the aim to minimize communication costs and satisfy delay constraints by offloading mobile data as much as possible with WiFi network and D2D communication. Markov decision process is a useful model for sequential decision making, where MNO needs to take a sequence of actions (wireless network selection). FHMDP is a Markov decision process with a finite number of decision epochs [24]. Since every data delivery task should be finished before a given deadline, FHMDP will plan data offloading decisions at each decision epoch. FHMDP planning phase can be implemented in remote cloud and ease the heavy burden of complex data offloading management by MS.

It is worth noting that the locations of WiFi APs and the base station are stationary, while MHs are moving around in the coverage area of base station. MHs can be considered as supplementary to WiFi APs because of their mobility.

IV. PROBLEM FORMULATION

In this section, we formulate the mobile data offloading problem as an FHMDP problem. Table I shows the notations used in the rest of this paper. In our model, mobile data is initially delivered to one or more MSs through cellular and WiFi networks. Additionally, any MH who carries a copy of the data can opportunistically transmit it to MSs using D2D communication. For each MS, data of size K needs to be transmitted before deadline D. MNO will select a wireless network for MS, at each decision epoch $d \in \mathcal{D} = \{1, \cdots, D\}$, based on the system state at that time.

A. System State and Action Space

The system state for multiple MSs and multiple MHs is defined as $\boldsymbol{s} = (\mathcal{M}, \mathcal{H})$, where $\mathcal{M}$ and $\mathcal{H}$ are the sets of states for MSs and MHs, respectively. More specifically, $\mathcal{M} = \{\boldsymbol{m}_i, i \in \{1, \cdots, M\}\}$ includes all the states of MSs, where $\boldsymbol{m}_i = (l_i, u_i, k_i)$ denotes the possible state of MS i; l_i denotes the location of MS i, u_i denotes the user type, and k_i is the size of data to be transmitted. $\mathcal{H} = \{l_j, j \in \{1, \cdots, N\}\}$ is a set that includes all the locations of MHs, where l_j is the location of MH j.

In the following, we omit the subscripts i and j of state parameters l, u and k for simplification. The state parameter $l \in \mathcal{L} = \{1, \cdots, L\}$ denotes the index of grid (or location), where L is the number of possible grids that MSs may reach before D. We assume that cellular network can provide seamless coverage to all grids. All grids are classified into four disjoint catagories at decision epoch d based on available WiFi or D2D connections. $\mathcal{L}_d^1$ denotes the grids covered by only cellular network, $\mathcal{L}_d^2$ contains the grids covered by both cellular and WiFi networks, $\mathcal{L}_d^3$ represents the grids covered by both celluar network and D2D communication, and $\mathcal{L}_d^4$ denotes the grids covered by celluar, WiFi networks and D2D communication. Due to the mobility of MHs, $\mathcal{L}_d^2$, $\mathcal{L}_d^3$ and $\mathcal{L}_d^4$ change over decision epoch d.

The state parameter $u \in \mathcal{U} = \{1, 2, \cdots, U\}$ represents the mobile user type (e.g. loose delay or tight delay), where U represents the number of different user types. We consider that different user types have different delay requirements resulting in different deadlines. To simplify the model, we consider two sets of user types, each of which has different QoS requirements. More specifically, user types that are delay-sensitive are in set $\mathcal{U}^1$; the other types (e.g. software update) are in set $\mathcal{U}^0$. Thus, $\mathcal{U} = \mathcal{U}^0 \cup \mathcal{U}^1$.

We divide the data, to be transmitted, into K equal portions; the state parameter $k \in \mathcal{K} = \{0, 1, \cdots, K\}$ represents the number of data portions still to be transmitted. If $k = 0$ when $d \leq D$, the data delivery process is completed.

After defining the system state of FHMDP, we next introduce the action space of MNO in mobile data offloading system. At each decision epoch, MNO selects one of the offloading actions for data transmission. There are four actions in the action space corresponding to four offloading decisions. Formally, action $a \in \mathcal{A} = \{1, 2, 3, 4\}$: (1) $a = 1$ (waiting action): MS will wait for a chance to receive data from WiFi or D2D network; (2) $a = 2$ (cellular action); (3) $a = 3$ (WiFi action); and (4) $a = 4$ (D2D action): MS can receive data from cellular network,WiFi network and D2D connection, respectively.

We observe that WiFi action is available when MS is in WiFi coverage and D2D action is available when MS can access a nearby MH. Thus, the available actions depend on the state parameter l. We also notice that the mobile user type u impacts the avaliable actions, i.e., D2D action is not available for delay sensitive data due to the low data rate. $\mathcal{A}(l, u) \subseteq \mathcal{A}$ representing the set of available actions at grid l for data with type u, is defined as follows:

$$\mathcal{A}(l,u) = \begin{cases} \{1,2\}, & l \in \mathcal{L}_d^1,\ u \in \mathcal{U}, \\ \{1,2,3\}, & l \in \mathcal{L}_d^2,\ u \in \mathcal{U}, \\ \{1,2,4\}, & l \in \mathcal{L}_d^3,\ u \in \mathcal{U}^0, \\ \{1,2,3,4\}, & l \in \mathcal{L}_d^4,\ u \in \mathcal{U}^0. \end{cases} \tag{1}$$

B. Transition Cost and Transition Probabilities

The transition cost for MS i, $c_d(\boldsymbol{m}_i, a)$, is independent from current state $\boldsymbol{m}_i$ and decision epoch d. It is equal to the action cost function $cost(a)$, which is defined as follows:

$$c_d(\boldsymbol{m}_i, a_i) = cost(a_i) = \nu_{a_i}^{l_i} \cdot T \cdot \chi_{a_i}, \tag{2}$$

where ν_a^l is the network data rate at grid l with action a, T is the period of time between two consecutive decision epochs, and χ_a is the cost to transmit a data unit by action a, i.e., χ_2, χ_3 and χ_4 are incurred by the usage of cellular, WiFi and D2D actions, respectively. The benefit of mobile data offloading is based on the fact that $\chi_3 < \chi_2$ and $\chi_4 < \chi_2$. This means that the cost to send data using cellular network is higher than that of using WiFi network and D2D communication. The total cost of transmitting data of size K is the sum of costs incurred at each period during the total transmission process.

There may be some data transmission tasks that cannot be completed before the deadline. For failed data transmissions (i.e. $k > 0$ when $d > D$), the penalty cost function is defined as follows:

$$c_{D+1}(\boldsymbol{m}_i) = penalty(u, k) = k^{(u+1)}, \tag{3}$$

where u and k finish the state parameters of $\boldsymbol{m}_i$. $k^{(u+1)}$ is an increasing function of k and u and reflects the fact that a larger remaining data size k and a tighter delay sensitivity u lead to a larger penalty.

In the following, we derive the transition probability between system states, which is the probability that current state $\boldsymbol{s}$ changes into $\boldsymbol{s}'$ in the next decision epoch by taking action $\boldsymbol{a}$. $\boldsymbol{a} = \{a_1, \cdots, a_M\}$ is the set of actions for all MSs. Since each MS or MH changes its state independently, we obtain the following state transition function.

$$\mathbb{P}(\boldsymbol{s}'|\boldsymbol{s}, \boldsymbol{a}) = \prod_{i \in \mathcal{M}} \mathbb{P}(\boldsymbol{m}_i'|\boldsymbol{m}_i, a_i) \cdot \prod_{j \in \mathcal{N}} \mathbb{P}(l_j'|l_j), \tag{4}$$

where

$$\mathbb{P}(\boldsymbol{m}_i'|\boldsymbol{m}_i, a_i) = \mathbb{P}(l_i', u_i', k_i'|l_i, u_i, k_i, a_i) = \mathbb{P}(l_i'|l_i) \cdot \mathbb{P}(k_i'|l_i, u_i, k_i, a_i). \tag{5}$$

For MS i, the next grid l_i' depends only on the current grid l_i and the user type u_i does not change during the offloading process, i.e., $u_i' = u_i$. The remaining data size k_i' depends on current location l_i, user type u_i, data size k_i, and selected action a_i.

$\mathbb{P}(l_i'|l_i)$ is the probability that MS will move from grid l_i to grid l_i'. We consider a two dimensions memoryless mobility pattern of MS as follows:

$$\mathbb{P}(l_i'|l_i) = \begin{cases} \mu, & \textit{if } l_i' = l_i, \\ \rho_j, & \textit{otherwise}, \end{cases} \tag{6}$$

where μ is the stable factor that denotes the probability that MS i stays at the same grid in two sequential decision epochs. Alternatively, MS can move randomly to an adjacent location with probability $\rho_j, j \in \{1, 2, 3, 4\}$, where j represents one of four possible moving directions (i.e., north, south, east and west). μ and ρ_j satisfy the relation $\mu + \sum_{j \in \{1,2,3,4\}} \rho_j = 1$.

$\mathbb{P}(k_i'|l_i, u_i, k_i, a_i)$ is the probability describing the change of remaining data size k_i and is defined as follows:

$$\mathbb{P}(k_i'|l_i, u_i, k_i, a_i) = \begin{cases} 1, & \textit{if } k_i' = \max\{k_i - \nu_{a_i}^{l_i}, 0\} \textit{ and } a_i \in \mathcal{A}(l_i, u_i), \\ 0, & \textit{otherwise}, \end{cases} \tag{7}$$

where the availability of a_i depends on grid l_i and user type u_i. Fig. 2 illustrates MS state transition graph considering parameters l and k; the terminal states are those with $k = 0$.

V. Hybrid Offloading Algorithm

In this section, we propose an algorithm, called hybrid offloading algorithm, to compute the optimal offloading policy.

A policy in FHMDP is denoted by $\boldsymbol{\pi} = \{\pi_i\}_{i\in\mathcal{M}}$, where $\pi_i : \boldsymbol{M}_i \times \mathcal{D} \rightarrow \mathcal{A}$ is the policy for MS i, which can decide an action based on $\boldsymbol{m}_i$ and the decision epoch d. The feasible domain for $\boldsymbol{\pi}$ is denoted by $\boldsymbol{\Pi}$. The objective function is defined as follows:

$$\min_{\boldsymbol{\pi}\in\boldsymbol{\Pi}} \sum_{i\in\mathcal{M}} E^{\pi_i}_{\boldsymbol{m}_{i,1}} \left[\sum_{d=1}^{D} c_d\big(\boldsymbol{m}^{\pi_i}_{i,d}, \pi_i(\boldsymbol{m}^{\pi_i}_{i,d}, d)\big) + c_{D+1}(\boldsymbol{m}^{\pi_i}_{i,D+1}) \right]. \tag{8}$$

The objective function aims to minimize the expected cost to deliver data of size K for all MSs. Before presenting our offloading algorithm, we define the value function as follows.

$$V^*_d(\boldsymbol{m}_i) = \min_{a_i\in\mathcal{A}(l_i,u_i)} Q_d(\boldsymbol{m}_i, a_i), \tag{9}$$

where

$$\begin{aligned} Q_d(\boldsymbol{m}_i, a_i) &= \sum_{\boldsymbol{m}'_i\in\boldsymbol{M}_i} \mathbb{P}(\boldsymbol{m}'_i|\boldsymbol{m}_i, a_i)\cdot\big(c_d(\boldsymbol{m}_i, a_i) + V^*_{d+1}(\boldsymbol{m}'_i)\big) \\ &= c_d(\boldsymbol{m}_i, a_i) + \sum_{\boldsymbol{m}'_i\in\boldsymbol{M}} \mathbb{P}(\boldsymbol{m}'_i|\boldsymbol{m}_i, a_i)\cdot V^*_{d+1}(\boldsymbol{m}'_i) \\ &= \nu^{l_i}_{a_i}\cdot T\cdot\chi_{a_i} + \sum_{l'_i\in\mathcal{L}}\sum_{k'_i\in\mathcal{K}} \mathbb{P}(l'_i|l_i)\cdot\mathbb{P}(k'_i|l_i,u_i,k_i,a_i)\cdot V^*_{d+1}(l'_i,u_i,k'_i) \\ &= \nu^{l_i}_{a_i}\cdot T\cdot\chi_{a_i} + \sum_{l'_i\in\mathcal{L}} \mathbb{P}(l'_i|l_i)\cdot V^*_{d+1}(l'_i,u_i,(k_i-\nu^{l_i}_{a_i}T)). \end{aligned} \tag{10}$$

The value function $V^*_d(\boldsymbol{m}_i)$ denotes the minimal expected cost for MS i in state $\boldsymbol{m}_i$ in decision epoch d to finish the data delivery process. $Q_d(\boldsymbol{m}_i, a_i)$ is a one step forward function that calculates the minimal expected cost if MS i selects action a_i; it is the sum of current cost $c_d(\boldsymbol{m}_i, a_i)$ and expected future cost. Eq (10) is derived from Eqs. (2), (5) and (7). The optimal policy is defined as

$$\pi^*_d(\boldsymbol{m}_i) = \operatorname*{argmin}_{a_i\in\mathcal{A}(l_i,u_i)} Q_d(\boldsymbol{m}_i, a_i). \tag{11}$$

Due to the mobility of MHs, we are interested in the expected number of MHs in grid l at

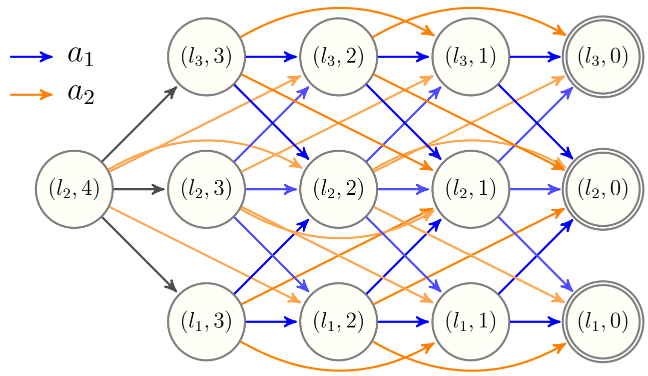


Fig. 2. A sample reduced state transition for MS, where the first component is the location of MS l, and the second component is data size k. Here, $\mathcal{L} = \{l_1, l_2, l_3\}$ and $K = 4$. Action a_1 can transfer 1 data unit each time, while a_2 can transfer 2 data units each time. The state with double circle is the terminal state, i.e., state $(l_i, 0)$, $i \in \mathcal{L}$.

decision epoch d, denoted by $\mathbb{N}(d, l)$.

$$\mathbb{N}(d,l) = \begin{cases} \sum\limits_{j \in \mathcal{N}} \delta(l_j^1, l), & \textit{if } d = 1, \\ \sum\limits_{l' \in \mathcal{L}} \mathbb{P}(l|l') \cdot \mathbb{N}(d-1, l'), & \textit{if } d = 2, \cdots, D, \end{cases} \tag{12}$$

where l_j^1 is the initial location of MH j. The function $\delta(l_j^1, l)$ returns 1, if $l_j^1 = l$; otherwise, it returns 0.

Our hybrid offloading algorithm, illustrated in Algorithm 1, consists of three phases: initialization phase (steps 1-3), planning phase (steps 4-12) and offloading phase (steps 13-22). In the initialization phase, we calculate the expected number of MHs in different locations and decision epochs using Eq. (12). $\mathbb{N}(d, l)$ is used to indicate the availability of D2D action in planning phase. We consider that MSs with the same user type have the same offloading policy. Thus, we generate the optimal policy based on the user type. Since the state transition graph, illustrated in Fig. 2, is an acyclic graph, we use the backward induction method to obtain the optimal offloading policy. In the planning phase, we first calculate the value function in $D+1$ (steps 4-6). Then, we calculate the value function $V_d^*(l, u, k)$ and optimal policy $\pi_u^*(l, u, k, d)$ from decision epoch D to 1 (steps 7-12).

In the offloading phase, MNO determines the offloading action at each decision epoch. Step 16 gets the location of MS i at decision epoch d, denoted by l_i^d. Step 17 obtains the idle MHs (MHs that are not serving other MSs) in proximity to MS i. If $\mathbb{N}(d, l_i^d) \geq 1$, then D2D action

Algorithm 1 Hybrid Offloading Algorithm

1: **for** $d \leftarrow \{1, \cdots, D\}$ and $l \in \mathcal{L}$ **do**
2:     Compute $\mathbb{N}(d, l)$ using Eq. (12)
3: **end for**
4: **for** $l \in \mathcal{L}$, $u \in \mathcal{U}$ and $k \in \mathcal{K}$ **do**
5:     $V^*_{D+1}(l, u, k) \leftarrow c_{D+1}(l, u, k)$
6: **end for**
7: **for** $u \in \mathcal{U}$ **do**
8:     **for** $d \leftarrow \{D, \cdots, 1\}$, $k \leftarrow \{0, \cdots, K\}$ and $l \in \mathcal{L}$ **do**
9:         Compute $V^*_d(l, u, k)$ using Eq. (9)
10:         Compute $\pi^*_u(l, u, k, d)$ using Eq. (11)
11:     **end for**
12: **end for**
13: **for** $i \in \mathcal{M}$ **do**
14:     Set $d \leftarrow 1$ and $k \leftarrow K$
15:     **while** $d < D + 1$ and $k > 0$ **do**
16:         Get current location l^d_i of MS
17:         Get current idle $\mathbb{N}(d, l^d_i)$
18:         Set $a \leftarrow \pi^*_{u_i}(l^d_i, u_i, k, d)$
19:         **if** $k - (D - d) \cdot \bar{\nu}_2 \cdot \kappa(u) > 0$ **then**
20:             Set $a \leftarrow \mathrm{argmax}_{a \in \mathcal{A}(l^d_i, u_i)} \nu^l_a$
21:         **end if**
22:         Set $k \leftarrow k - \nu^l_a \cdot T$
23:         Set $d \leftarrow d + 1$
24:     **end while**
25: **end for**

is available. Step 18 sets a to the action provided by optimal policy. In order to counteract the prediction error caused by the mobility of MSs and MHs, steps 19-21 first check whether data of size k can be transmitted using cellular network before deadline. If the response is yes, MNO will take action according to the optimal policy. Otherwise, the network with highest data rate will be selected (step 20). $\bar{\nu}_2$ is the average data rate of cellular network. Notice that the function $\kappa(u)$ (step 19) is used to control the delay sensitivity of different user types. Generally, user type accepting D2D communication results in larger prediction error, leading to a higher value of $\kappa(u)$. We set $\kappa(u)$ to 1.2 and 1 for $u \in \mathcal{U}^0$ and $u \in \mathcal{U}^1$, respectively [25]. However, these two values can be adjusted according to different situations. The main computational complexity of Algorithm 1 is associated with the planning phase, that is $\mathcal{O}(UDLK)$.

VI. Monotone Policy and Offloading Algorithm

Since the state space becomes extremely large with the increase of deadline D, data size K and the number of grids L, Algorithm 1 will take more time and resources to solve the problem. In order to reduce the computational complexity for generating the optimal policy, we provide sufficient conditions under which the offloading policy is monotone (non-decreasing or non-increasing) in terms of data size k and decision epoch d, called monotone policy. The monotone policy enables efficient computation due to the existence of threshold structure in optimal policy. Threshold structure has several boundaries between different offloading decisions according to k and d, as discussed in Section VII-B. Thus, instead of generating the optimal actions for all the system state, we only need to determine the threshold states, which can greatly reduce the computational complexity.

The rest of this section is organized as follows. Subsection VI-A presents our assumptions. Subsection VI-B discusses the properties of optimal policy. Subsection VI-C shows the special case where the monotone policy degrades into a single action a^* that does not change with k and d. Subsection VI-D shows the general case where the monotone policy has threshold structures. Subsection VI-E presents an algorithm for generating and executing monotone policy, called monotone offloading algorithm.

A. Assumptions

We make the following assumptions for deriving the monotone policy.

Assumption 1. *Eqs.* (2) *and* (3) *satisfy the following relation:*

$$c_d(\boldsymbol{m}, a) < c_{D+1}(\boldsymbol{m}),\ \forall d \in \mathcal{D}, i \in \mathcal{M}, a \in \mathcal{A}. \tag{13}$$

Eq. (13) indicates that MNO should take an action $a \in \mathcal{A}$ at each decision epoch d, since the penalty cost is larger than the action cost.

Assumption 2. *The unit costs of cellular, WiFi and D2D networks (χ_2, χ_3 and χ_4) satisfy the relation $\chi_3 < \chi_4 < \chi_2$.*

Assumption 3. *The data rates of cellular, WiFi and D2D networks (ν_2, ν_3 and ν_4) are location independent.*

We underline that our model can be extended to location dependent rates by considering the same action with different data rates as different actions. For example, the data rate of action a in grid l, denoted by ν_a^l, is location dependent. We replace action a with actions $(a_1, \cdots, a_W)$, each of which represents the WiFi action with a different data rate. Thus, the data rate of WiFi action becomes location independent, denoted by $\nu_{a_i}, i \in \{1, \cdots, W\}$, where W denotes the number of different data rates that can be used in the case of WiFi action.

B. Properties of the optimal policy

We discuss some properties of the optimal policy under above assumptions.

Lemma 1. *The penalty function $c_{D+1}(l, u, k)$ satisfies the following relation:*

$$c_{D+1}(l, u, k) - c_{D+1}(l, u, (k - \nu_a T)) \geq c_d(l, u, k, a), \tag{14}$$

for all $d \in \mathcal{D}$, $l \in \mathcal{L}$, $u \in \mathcal{U}$, $k \geq \nu_a T$ and $a \in \mathcal{A}$.

Proof. Note that the penalty cost (defined in Eq. (3)) is greater than the action cost (defined in Eq. (2)) for the same data size $\nu_a T$, as shown in *Assumption 1*.

Moreover, given u, the penalty cost is a power function with respect to data size k (see Eq. (3)). Thus, we obtain that, $\forall k \geq \nu_a T$,

$$c_{D+1}(l, u, k) - c_{D+1}(l, u, (k - \nu_a T)) \geq c_{D+1}(l, u, \nu_a T) - c_{D+1}(l, u, 0) = c_{D+1}(l, u, \nu_a T) \tag{15}$$

From Eqs. (13), (15) and the definition of action cost, we get Eq. (14). □

Lemma 2. *The value function $V_d^*(l, u, k)$ is non-decreasing in the remaining data size k, $\forall l \in \mathcal{L}, u \in \mathcal{U}, d \in \mathcal{D}$.*

Proof. We prove the result using backward induction. Since the penalty function $c_{D+1}(l, u, k)$ is non-decreasing in k, the value function $V_{D+1}^*(l, u, k) = c_{D+1}(l, u, k)$ is non-decreasing in k. Assume that $V_{d'}^*(l, u, k)$ is non-decreasing in k for $d' \in \{d+1, \cdots, D\}$. Based on Eq (10), we get

$$V_d^*(l, u, k) = Q_d(l, u, k, a^*) = \nu_{a^*} T \chi_{a^*} + \sum_{l' \in \mathcal{L}} \mathbb{P}(l'|l) \cdot V_{d+1}^*(l', u, (k - \nu_{a^*} T)). \tag{16}$$

By induction hypothesis, $V^*_{d+1}(l', u, (k - \nu_{a^*}T)$ is non-decreasing in k. Thus, $V^*_d(l, u, k)$ is non-decreasing in k. $\square$

Lemma 3. *The value function $V^*_d(l, u, k)$ is non-decreasing in decision epoch d, $\forall l \in \mathcal{L}, u \in \mathcal{U}, k \in \mathcal{K}$.*

Proof. We first show that $V^*_{D+1}(l, u, k) \geq V^*_D(l, u, k)$. From Eq (9), we obtain

$$\begin{aligned} V^*_D(l, u, k) &= c_D(l, u, k, a^*) + c_{D+1}(l', u, (k - \nu_{a^*}T)) \\ &\leq c_{D+1}(l, u, k) = V^*_{D+1}(l, u, k), \end{aligned} \tag{17}$$

where the inequality is based on *Assumption 1*. Next, we assume that $V_{d'}(l, u, k)$ is non-decreasing for $d' \in \{d + 1, \cdots, D\}$. By induction hypothesis, $V^*_{d+1}(l', u, (k - \nu_{a^*}T)$ is non-decreasing in Eq. (16). Thus, $V^*_d(l, u, k)$ is non-decreasing in d. $\square$

Lemma 2 reflects the fact that the expected cost is higher when the non-transmitted data size k is larger. *Lemma 3* shows that a larger decision epoch d (i.e. the deadline is closer) results in higher expected cost. The proofs of *Lemma 1*, *Lemma 2* and *Lemma 3* are available online [26].

C. Single action monotone policy

In this subsection, we show the special case where the monotone policy degrades into a dominant action policy, i.e., $\pi^*_d(l, u, k) = a^*_{(l,u)}$.

Definition 1. *Given $l \in \mathcal{L}$ and $u \in \mathcal{U}$, $a^*_{(l,u)} \in \mathcal{A}(l, u)$ is a dominant action if*

$$\pi^*_d(l, u, k) = \underset{a_{(l,u)} \in \mathcal{A}(l,u)}{\operatorname{argmin}} Q_d(l, u, k, a) = a^*_{(l,u)}, \tag{18}$$

for all $d \in \mathcal{D}$ and $k \in \mathcal{K}^+$,

$\mathcal{K}^+ = \mathcal{K} \setminus \{0\}$; $k = 0$ indicates that the data transmission process is finished and thus no action is chosen. Notice that $a^*_{(l,u)}$ is different from a^*. $a^*_{(l,u)}$ is the optimal action for all $d \in \mathcal{D}$ and $k \in \mathcal{K}^+$, while a^* is the optimal action for some k and d. Next, we establish conditions under which a dominant action exists.

Theorem 1. *Given $u \in \mathcal{U}$ and $l \in \mathcal{L}^2_d \cup \mathcal{L}^4_d$ where WiFi action is available, if $\nu_2 < \nu_3$ and $\nu_4 < \nu_3$, then $a^*_{(l,u)} = 3\,(WiFi)$ is a dominant action, for all $d \in \mathcal{D}$ and $k \in \mathcal{K}^+$. The optimal* *policy is*

$$\pi_d^*(l,u,k) = a_{(l,u)}^* = 3\,(WiFi). \tag{19}$$

Proof. We first prove that $Q_d(l,u,k,3) \leq Q_d(l,u,k,2)$. From Eq. (16), we get

$$Q_d(l,u,k,3) = \nu_3 T\chi_3 + \sum_{l'\in\mathcal{L}} P(l'|l)\cdot V_{d+1}^*(l',u,(k-\nu_3 T)), \tag{20}$$

$$Q_d(l,u,k,2) = \nu_2 T\chi_2 + \sum_{l'\in\mathcal{L}} P(l'|l)\cdot V_{d+1}^*(l',u,(k-\nu_2 T)). \tag{21}$$

Subtracting Eq. (20) from Eq. (21), we get Eq. (22a). Since $\nu_2 < \nu_3$, let $\nu_3 = \nu_2 + \delta$ and $\delta > 0$. Replacing ν_3 with $\nu_2 + \delta$ in Eq. (22a), we get Eq. (22b).

$$\begin{aligned}
&Q_d(l,u,k,2) - Q_d(l,u,k,3) \\
&= \nu_2 T\chi_2 - \nu_3 T\chi_3 + \sum_{l'\in\mathcal{L}} P(l'|l)\cdot \Big(V_{d+1}^*\big(l',u,(k-\nu_2 T)\big) - V_{d+1}^*\big(l',u,(k-\nu_3 T)\big)\Big) && \text{(22a)}\\
&= \nu_2 T(\chi_2-\chi_3) - \delta T\chi_3 + \sum_{l'\in\mathcal{L}} P(l'|l)\cdot \Big(V_{d+1}^*\big(l',u,(k-\nu_2 T)\big) - V_{d+1}^*\big(l',u,(k-\nu_2 T-\delta T)\big)\Big) && \text{(22b)}\\
&> \sum_{l'\in\mathcal{L}} P(l'|l)\cdot \Big(V_{d+1}^*\big(l',u,(k-\nu_2 T)\big) - V_{d+1}^*\big(l',u,(k-\nu_2 T-\delta T)\big) - \delta T\chi_3\Big) && \text{(22c)}\\
&= \sum_{l'\in\mathcal{L}} P(l'|l)\cdot \Big(\sum_{a\in\mathcal{A}} p_a\delta T\chi_a - \delta T\chi_3\Big) && \text{(22d)}\\
&\geq \sum_{a\in\mathcal{A}} p_a\delta T\chi_3 - \delta T\chi_3 && \text{(22e)}\\
&= 0. && \text{(22f)}
\end{aligned}$$

Note that: (1) Based on *Assumption 2*, where $\chi_2 > \chi_3$, we get $\nu_2 T(\chi_2-\chi_3) > 0$ in Eq. (22b). By eliminating $\nu_2 T(\chi_2-\chi_3) > 0$, we get Eq. (22c); (2) Eq. (22d) is obtained based on *Lemma 2*. Since $(k-\nu_2 T) \geq (k-\nu_2 T-\delta T)$, we get Eq. (23).

$$V_{d+1}^*\big(l',u,(k-\nu_2 T)\big) - V_{d+1}^*\big(l',u,(k-\nu_2 T-\delta T)\big) = \Delta \geq 0, \tag{23}$$

where Δ is the cost for transmitting data of size δT; it is defined in Eq. (24).

$$\Delta = \sum_{a\in\mathcal{A}} p_a\delta T\chi_a, \tag{24}$$

where p_a is the percentage of data size δ transmitted by choosing action a. Except for the waiting action $a = 1$, where $p_1 = 0$, p_a is unknown for other actions. From *Assumption 2*, χ_3 is the minimum cost, we get Eq. (25).

$$\Delta = \delta T \sum_{a\in\mathcal{A}\backslash\{1\}} p_a \chi_a \geq \delta T \sum_{a\in\mathcal{A}\backslash\{1\}} p_a \chi_3 = \delta T \chi_3. \tag{25}$$

By Eq. (22), we have proved that $Q_d(l,u,k,3) \leq Q_d(l,u,k,2)$. Similarly, if $\nu_4 < \nu_3$, we can prove that $Q_d(l,u,k,3) \leq Q_d(l,u,k,4)$. Moreover, since $\nu_1 = 0 < \nu_3$, we obtain $Q_d(l,u,k,3) \leq Q_d(l,u,k,1)$. According to *Definition 1*, $a^*_{(l,u)} = 3\,(WiFi)$ is the dominant action.

□

The proof of *Theorem 1* is available online [26]. Based on *Theorem 1*, where the waiting action, cellular action and D2D action are dominated by WiFi action, we obtain the following corollary.

Corollary 1. *(1) Given $l \in \mathcal{L}$ and $u \in \mathcal{U}$, for all $a^1_{(l,u)} \in \mathcal{A}(l,u)$ and $a^2_{(l,u)} \in \mathcal{A}(l,u)$, if $\chi_{a^1_{(l,u)}} > \chi_{a^2_{(l,u)}}$ and $\nu_{a^1_{(l,u)}} < \nu_{a^2_{(l,u)}}$, then $a^1_{(l,u)}$ is dominated by $a^2_{(l,u)}$. (2) Given $l^1, l^2 \in \mathcal{L}$ and $u \in \mathcal{U}$, for all $a^1_{(l^1,u)} \in \mathcal{A}(l^1,u)$ and $a^2_{(l^2,u)} \in \mathcal{A}(l^2,u)$, if $\chi_{a^1_{(l^1,u)}} > \chi_{a^2_{(l^2,u)}}$, $\nu_{a^1_{(l^1,u)}} < \nu_{a^2_{(l^2,u)}}$ and $a^2_{(l^2,u)} \in \mathcal{A}(l^2,u) \setminus \mathcal{A}(l^1,u)$, then $a^1_{(l^1,u)}$ is potentially dominated by $a^2_{(l^2,u)}$.*

This corollary includes two parts. The first part implies the dominant relationship between two actions in same grid, while the second part reveals the potentially dominant relationship between actions in different grids.

D. General monotone policy

In this subsection, we show the general case where the monotone policy exists in dimensions k and d. We first introduce the basic definitions and properties of superadditive function and illustrate that $Q_d(l,u,k,a)$ is a superadditive function in $\mathcal{K}\times\mathcal{A}$ and $\mathcal{D}\times\mathcal{A}$. Then we derive the optimal monotone policy $\pi^*_d(l,u,k)$ in dimensions k and d.

Definition 2. *A real valued function $f(m,a)$ is superadditive in $\mathcal{M}\times\mathcal{A}$, if*

$$f(m^+,a^+) - f(m^+,a^-) \geq f(m^-,a^+) - f(m^-,a^-), \tag{26}$$

for $\forall m^+, m^- \in \mathcal{M}$ and $\forall a^+, a^- \in \mathcal{A}$, where $m^+ \geq m^-$ and $a^+ \geq a^-$.

Given the definition of superadditive function, we next illustrate its properties summarized in *Lemmas 4* and *5* [27].

Lemma 4. *If $f_1(m,a)$ and $f_2(m,a)$ are superadditive functions in $\mathcal{M} \times \mathcal{A}$, then the function $h(m,a) = f_1(m,a) + f_2(m,a)$ is superadditive in $\mathcal{M} \times \mathcal{A}$.*

Lemma 5. *If $f(m,a)$ is a superadditive function in $\mathcal{M} \times \mathcal{A}$, then the function $g(a)$ defined below is monotone increasing in m.*

$$g(a) = \operatorname*{argmin}_{a \in \mathcal{A}} f(m,a). \tag{27}$$

Lemma 4 shows that the sum of two superadditive functions satisfies superadditive property. *Lemma 5* states that a superadditive function including two variables can be considered as a monotone increasing function having one variable, which implies the theoretical basic of monotone policy. In our model, we consider $Q_d(l,u,k,a)$ as a superadditive function in $\mathcal{K} \times \mathcal{A}$ and $\mathcal{D} \times \mathcal{A}$. The monotone policy in dimensions k and d is summarized as follows.

Theorem 2. *The optimal monotone policy $\Pi^* = \left\{\pi_d^*(l,u,k) = a^*, \forall l \in \mathcal{L}, u \in \mathcal{U}, k \in \mathcal{K}, d \in \mathcal{D}\right\}$ has threshold structure in both k and d as follows:*

(a) *For location $l \in \mathcal{L}^1$ with only cellular network and $u \in \mathcal{U}$, we get $\mathcal{A}(l,u) = \{1,2\}$ by Eq.(1). There is one threshold for both k and d. That is $\forall d \in \mathcal{D}$,*

$$\pi_d^*(l,u,k) = \begin{cases} 2\,(cellular), & if\ k \geq k^*(l,u,d), \\ 1\,(waiting), & otherwise, \end{cases} \tag{28}$$

and $\forall k \in \mathcal{K}$,

$$\pi_d^*(l,u,k) = \begin{cases} 2\,(cellular), & if\ d \geq d^*(l,u,k), \\ 1\,(waiting), & otherwise. \end{cases} \tag{29}$$

(b) *For location $l \in \mathcal{L}^2$ with cellular and WiFi networks, and $u \in \mathcal{U}$, we get $\mathcal{A}(l,u) = \{1,2,3\}$ by Eq.(1). If $\nu_2 > \nu_3 > \nu_4$, there is one threshold for both k and d. That is $\forall d \in \mathcal{D}$,*

$$\pi_d^*(l,u,k) = \begin{cases} 2\,(cellular), & if\ k \geq k^*(l,u,d), \\ 3\,(WiFi), & otherwise, \end{cases} \tag{30}$$

and $\forall k \in \mathcal{K}$,

$$\pi_d^*(l,u,k) = \begin{cases} 2\,(cellular), & if\ d \geq d^*(l,u,k), \\ 3\,(WiFi), & otherwise. \end{cases} \tag{31}$$

(c) *For location $l \in \mathcal{L}^3$ with cellular network and D2D communication, and $u \in \mathcal{U}^0$, we get $\mathcal{A}(l,u) = \{1,2,4\}$ by Eq. (1). If $\nu_2 > \nu_4$, there are two thresholds for both k and d. That is $\forall d \in \mathcal{D}$,*

$$\pi_d^*(l,u,k) = \begin{cases} 1\,(waiting), & if\ k \leq k_1^*(l,u,d), \\ 2\,(cellular), & if\ k \geq k_2^*(l,u,d), \\ 4\,(D2D), & otherwise, \end{cases} \tag{32}$$

and $\forall k \in \mathcal{K}$*,*

$$\pi_d^*(l,u,k) = \begin{cases} 1\,(waiting), & if\ d \leq d_1^*(l,u,k), \\ 2\,(cellular), & if\ d \geq d_2^*(l,u,k), \\ 4\,(D2D), & otherwise. \end{cases} \tag{33}$$

(d) *For location* $l \in \mathcal{L}^4$ *with cellular network, WiFi network and D2D communication, and* $u \in \mathcal{U}^0$*, we get* $\mathcal{A}(l,u) = \{1,2,3,4\}$ *by Eq.(1). If* $\nu_2 > \nu_4 > \nu_3$*, there are two thresholds for both* k *and* d*. That is* $\forall d \in \mathcal{D}$*,*

$$\pi_d^*(l,u,k) = \begin{cases} 3\,(WiFi), & if\ k \leq k_1^*(l,u,d), \\ 2\,(cellular), & if\ k \geq k_2^*(l,u,d), \\ 4\,(D2D), & otherwise, \end{cases} \tag{34}$$

and $\forall k \in \mathcal{K}$*,*

$$\pi_d^*(l,u,k) = \begin{cases} 3\,(WiFi), & if\ d \leq d_1^*(l,u,k), \\ 2\,(cellular), & if\ d \geq d_2^*(l,u,k), \\ 4\,(D2D), & otherwise. \end{cases} \tag{35}$$

Proof. We prove the result of *Theorem 2.a*. First, we show that the transition cost $c_d(s,a)$ defined in Eq. (14) is superadditive in $\mathcal{M} \times \mathcal{A}$. From Eq. (14), we get

$$c_d(m^+,a^+) - c_d(m^+,a^-) = \nu_{a^+} T \chi_{a^+} - \nu_{a^-} T \chi_{a^-}, \tag{36}$$

$$c_d(m^-,a^+) - c_d(m^-,a^-) = \nu_{a^+} T \chi_{a^+} - \nu_{a^-} T \chi_{a^-}. \tag{37}$$

From Eqs. (36) and (37), we get

$$c_d(m^+,a^+) - c_d(m^+,a^-) = c_d(m^-,a^+) - c_d(m^-,a^-). \tag{38}$$

Since Eq. (38) satisfies the definition of superadditive function (*Definition 2*), $c_d(s,a)$ is superadditive in $\mathcal{M} \times \mathcal{A}$.

Then, we show that $V_{d+1}^*(l',u,(k-\nu_a))$ is superadditive in $\mathcal{M} \times \{a^-,a^+\}$, where $a^- = 1$ and $a^+ = 2$.

According to *Lemma 2*, we get Eq. (39). Thus, $V_{d+1}^*(l',u,(k-\nu_a T))$ is superadditive in $\mathcal{M} \times \{a^-,a^+\}$.

$$
\begin{aligned}
&V^*_{d+1}(l', u, (k^+ - \nu_{a^+}T)) - V^*_{d+1}(l', u, (k^+ - \nu_{a^-}T)) \\
&\geq V^*_{d+1}(l', u, (k^- - \nu_{a^+}T)) - V^*_{d+1}(l', u, (k^- - \nu_{a^-}T)).
\end{aligned} \tag{39}
$$

Based on *Lemma 4*, we obtain that $Q_d(l, u, k, a)$ defined in Eq. (40) is superadditive in $\mathcal{M} \times \{a^-, a^+\}$.

$$
Q_d(l, u, k, a) = c_d(m, a) + \sum_{l' \in \mathcal{L}} P(l'|l) \cdot V^*_{d+1}(l', u, (k - \nu_a T)) \tag{40}
$$

Based on *Lemma 5*, we obtain that the optimal policy $\pi^*_d(l, u, k)$ defined in Eq. (41) is monotone increasing with s.

$$
\pi^*_d(l, u, k) = \operatorname*{argmin}_{a \in \{a^-, a^+\}} Q_d(l, u, k, a) \tag{41}
$$

Thus, $\pi^*_d(l, u, k)$ is a step function of the form Eq. (28). $k^*(l, u, d)$ is a state at which the optimal policy switches from $a^- = 1$ to $a^+ = 2$, called threshold state. Similarly, we can prove Eq. (29) by showing that $Q_d(l, u, k, a)$ is superadditive in $\mathcal{D} \times \mathcal{A}$ by *Lemma 3*. □

The proof of *Theorem 2.a* is available online [26]. We can derive *Theorem 2.(b)-(d)* by *Corollary 1* and then prove them the same way as *Theorem 2.a*. For example, considering *Theorem 2.b*, where $\chi_3 < \chi_4 < \chi_2$ (by *Assumption 2*) and $\nu_2 > \nu_3 > \nu_4$, action 4 is potentially dominated by action 3. This is because that $\chi_3 < \chi_4$ and $\nu_3 > \nu_4$ when $3 \in \mathcal{A}(l, u)$ and $4 \notin \mathcal{A}(l, u)$. Notice that the waiting action 1 is used to delay data transmission by seeking better offloading action. However, we don't need to delay now, since action 4 (the only action not in $\mathcal{A}(l, u)$) is potentially dominated by action 3. Moreover, action 2 and action 3 is not dominated by each other. Thus, $\mathcal{A}(l, u) = \{2, 3\}$ and there is one threshold in $l \in \mathcal{L}^2$. Since $\chi_3 < \chi_2$, we first choose action $3\,(WiFi)$ which has lower cost. When exceeding the threshold $k^*(l, u, d)$ or $d^*(l, u, k)$, action $2\,(cellular)$ which has higher data rate is used, as shown in Eqs. (30) and (31).

The monotone offloading algorithm will search the threshold states from the state space. In order to reduce the searching complexity, we make use of the following corollary.

Corollary 2. *In the monotone policy,* $\forall l \in \mathcal{L}, u \in \mathcal{U}$*,* $i \in \{1, 2\}$ *is the index of thresholds,*

(1) $k^*_i(l, u, d) \geq k^*_i(l, u, d+1)$ *and* $k^*_2(l, u, d) \geq k^*_1(l, u, d)$*,* $\forall d \in \mathcal{D}$*;*

(2) $d_i^*(l,u,k) \geq d_i^*(l,u,k+1)$ *and* $d_2^*(l,u,k) \geq d_1^*(l,u,k)$*,* $\forall k \in \mathcal{K}$*.*

E. Monotone offloading algorithm

We propose monotone offloading algorithm to calculate the general monotone policy with a lower computational complexity, compared with hybrid offloading algorithm. By taking advantage of the threshold structure, monotone offloading algorithm only searches for the threshold states, instead of computing the optimal action for every system state. The threshold states are calculated in dimension k, denoted by $\Pi = \left\{k_i^*(l,u,d) = k, \forall\, l \in \mathcal{L}, u \in \mathcal{U}, d \in \mathcal{D}, i \in \{1,2\}\right\}$.

Algorithm 2 consists of two phases: (1) planning phase (steps 3-5): *Algorithm 3* is used to calculate the threshold states; and (2) running phase (steps 7-14): The offloading action is decided by MS's grid l. For grid with WiFi coverage (step 10), the optimal action is WiFi action; For grid with only cellular coverage (step 11), the optimal action is determined by *Theorem 2.a*; For grid with D2D coverage (step 12), the optimal action is determined by *Theorem 2.c*.

Algorithm 2 Monotone Offloading Algorithm

1: *Planning Phase*
2: Initialize $V_{d+1}^*(s)$ with Eq. (3) and $\Pi \leftarrow \varnothing$
3: **for** $l \in \mathcal{L}, \quad u \in \mathcal{U}, \quad d \leftarrow \{D, ..., 1\}$ **do**
4:     Call *Calculate Threshold States Algorithm*
5: **end for**
6: *Running Phase*
7: Set $d \leftarrow 1$ and $k \leftarrow K$
8: **while** $d < D+1$ and $k > 0$ **do**
9:     Get the location of MS as l
10:     **if** $l \in \mathcal{L}_d^2 \cup \mathcal{L}_d^4$, set $a \leftarrow 3$ by *Theorem 1*, **end if**
11:     **if** $l \in \mathcal{L}_d^1$, choose action by *Theorem 2.a*, **end if**
12:     **if** $l \in \mathcal{L}_d^3$, choose action by *Theorem 2.c*, **end if**
13:     $k \leftarrow k - \nu_a, \quad d \leftarrow d+1$
14: **end while**

Algorithm 3 calculates the threshold states for $l \in \mathcal{L}_d^1 \cup \mathcal{L}_d^3$. Notice that there is one threshold in $l \in \mathcal{L}_d^1$ and two thresholds in $l \in \mathcal{L}_d^3$, as illustrated in Section VII-B. *Algorithm 3* first determines the number of thresholds based on l (step 1). Then it calculates the threshold states $k_i^*(l,u,d)$ in dimension k. Notice that we only consider k that starts from $k_i^*(l,u,d+1)$ (step 3) instead of $\forall\, k \in \mathcal{K}$, due to the fact that the threshold states cannot be found in $k < k_i^*(l,u,d+1)$ based on *Corollary 2*. *Algorithm 3* can find the threshold states within a constant number of loops. Indeed, the maximum number of outer loops (steps 2-10) is 2. The number of inner loops (steps

Algorithm 3 Calculate Threshold States

1: Calculate threshold states in dimension k
  if $l \in \mathcal{L}_d^1$, set $numThreshold \leftarrow 1$ and $a_1 \leftarrow 2$ **endif**
  if $l \in \mathcal{L}_d^3$, set $numThreshold \leftarrow 2$, $a_1 \leftarrow 4$ and $a_2 \leftarrow 2$ **endif**
2: **for** $i \leftarrow \{1, ..., numThreshold\}$ **do**
3:   Set $k \leftarrow k_i^*(l, u, d+1)$ and $flag \leftarrow 0$
4:   **while** $k \leq K$ and $flag == 0$ **do**
5:     Calculate $Q_d(s, a)$, $\forall a \in \mathcal{A}(l, u)$ using Eq. (10)
6:     Set $\pi_d^*(l, u, k) \leftarrow \text{argmin}_{a \in \mathcal{A}(l,u)} Q_d^*(s, a)$
7:     Set $V^*(s, d) \leftarrow Q_d^*(s, \pi_d^*(l, u, k))$
    if $\pi_d^*(l, u, k) == a_i$, set $\Pi \leftarrow \Pi \cup \left\{k_i^*(l, u, d) \leftarrow k\right\}$ and $flag \leftarrow 1$ **endif**
8:     Set $k \leftarrow k + 1$
9:   **end while**
10: **end for**

4-9) is determined by $\max_{d \in \mathcal{D}}\{k_i^*(l, u, d) - k_i^*(l, u, d+1)\}$, which is a constant number. Thus, the complexity of *Algorithm 3* is $\mathcal{O}(1)$. The complexity of *Algorithm 2* is mainly due to the planning phase, that is $\mathcal{O}(LUD)$.

VII. Performance Evaluation

This section evaluates the performance of our proposed approaches to implement efficient mobile data offloading. More specifically, we aim to evaluate the impact of mobile data size and delay tolerance on the performance of our offloading methods. First, we describe the parameters' settings used in our numerical analysis. Next, the threshold structures in monotone policy are illustrated. At last, we evaluate the performance metrics considered in our analysis.

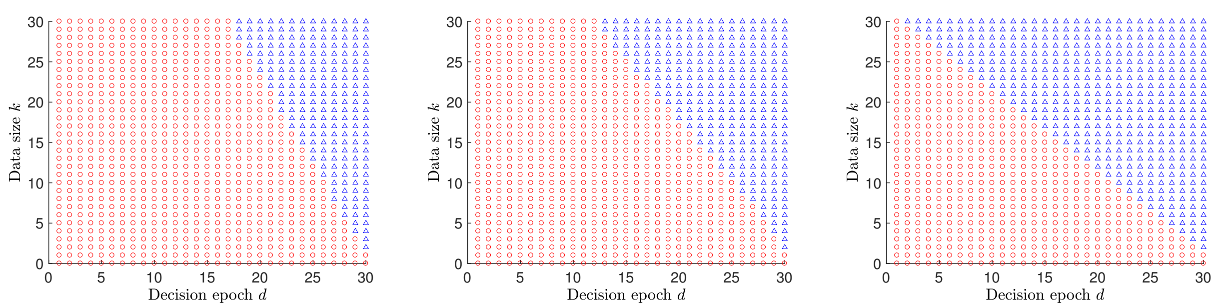
A. Experimental Setup

We have solved the optimal offloading policy and implemented the proposed offloading methods using MATLAB. For each choice of parameters' settings, we run the simulations 1000 times and show the average values. The number of grids L equals to 10×10. MSs and MHs move in the grids according to the memoryless mobility pattern, where the stable factor $\mu = 0.6$ and $\rho_i = 0.1, \ \forall i \in \{1, 2, 3, 4\}$ [28]. The number of WiFi APs and MHs are denoted as n_w and n_h, respectively. The locations of WiFi APs and MHs are generated randomly. The data rates of WiFi, cellular and D2D actions follow normal distribution with means $\nu_2 = 16$ Mbps, $\nu_3 = 24$ Mbps and $\nu_4 = 8$ Mbps, respectively, and standard deviations equal to 5 Mbps, which is a rational setting based on [4, 29].

Similar to [28], the cellular unit cost is set to $\chi_2 = 1$ serving as a baseline. We set the unit cost for WiFi network and D2D communication in terms of the reserve price (i.e., the price that MNO is willing to pay at most for offloading one data unit), where $\chi_3 = [0.05, 0.08]$ and $\chi_4 = 0.2$ [7]. The reserve price χ_4 is set by MNO. If MNO sets a high χ_4 for D2D communication, MHs will get high rewards and thus will be willing to help in D2D offloading. In our evaluation, we set $\chi_3 = 0.08$ and $\chi_4 = 0.2$. The length of a time slot T is 10 *seconds* unless stated otherwise.

B. Threshold Structures in Monotone Policy

In our settings, WiFi action has the highest data rate and lowest cost. Thus, according to *Theorem 1*, WiFi action is the dominant action in $l \in \mathcal{L}_d^2 \cup \mathcal{L}_d^4$. The optimal policies having threshold structures in $\mathcal{L}_d^1$ and $\mathcal{L}_d^3$, based on *Theorem 2* (a) *and* (c), are shown in Figs. 3 and 4, respectively.

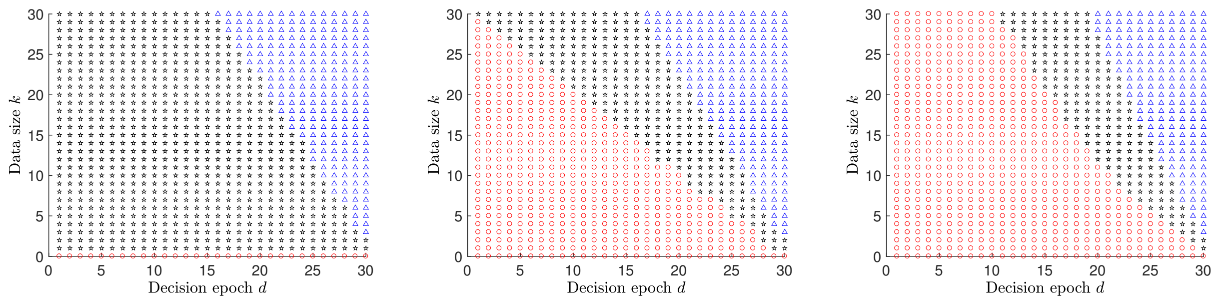


(a) $n_w = 33, n_h = 66$ (b) $n_w = 10, n_h = 66$ (c) $n_w = 0, n_h = 66$

Fig. 3. Monotone policy in $l \in \mathcal{L}^1$: $K = 30$ *Mbytes* and $D = 30$ *seconds*. The dots (○) and triangles (△) represent the waiting and cellular action, respectively.

Fig. 3 illustrates the optimal policy for MSs in $l \in \mathcal{L}_d^1$, where $\mathcal{A}(l, u) = \{1, 2\}$. Thus, MNO has two actions, i.e., waiting action and cellular action. As shown in Fig. 3, a single threshold exits in dimension d. It shows that cellular action is selected when d is large enough. Otherwise, MNO chooses the waiting action. For example, in Fig. 3(a), given $k = 10$, the optimal action changes from waiting action to cellular action when $d > 26$. The single threshold that exits in dimension k has similar observation. Fig. 3 shows that our monotone policy will delay the usage of cellular network until the threshold state ($k^*(l, u, d)$ or $d^*(l, u, k)$). This is rational since MNO seeks to use WiFi network or D2D communication before deadline.

We observe that the threshold changes with the number of WiFi APs and MHs. With the belief that the number of WiFi APs in Fig. 3(a) is higher than that in Fig. 3(b), MS in Fig. 3(a) has larger WiFi connection probability, which implies higher offloading potential (the size of data can be transmitted using WiFi network or D2D communication before deadline). Thus, the waiting action area in Fig. 3(a) is larger than that in Fig. 3(b). Since Fig. 3(c) has the smallest offloading potential, the waiting action area is smaller than that in Figs. 3(a) and 3(b).



(a) $n_w = 0, n_h = 66$ (b) $n_w = 33, n_h = 66$ (c) $n_w = 66, n_h = 66$

Fig. 4. Monotone policy in $l \in \mathcal{L}^3$: $K = 30\,Mbytes$ and $D = 30\,seconds$. The dots ($\circ$), triangles ($\triangle$), and stars ($\star$) represent the waiting, cellular and D2D action, respectively.

Fig. 4(a) shows that MNO chooses D2D action when d or k is small, instead of waiting action (see Fig. 3(c)), with the same knowledge of network setting (i.e., $n_w = 0$ and $n_h = 66$). Since $\chi_4 < \chi_2$, MNO chooses D2D action to minimize the transmission cost. However, when d or k is large, cellular action is selected to ensure that data transmission will be completed before deadline, due to the fact that $\nu_2 > \nu_4$. Figs. 4(b) and 4(c) show that two thresholds separate three actions (waiting action, cellular action and D2D action.) in dimensions k and d, since $l \in \mathcal{L}_d^3$ and $\mathcal{A}(l,u) = \{1, 2, 4\}$. Compared with two actions (D2D and cellular actions) shown in Fig. 4(a), additional waiting action occurs in Figs. 4(b) and 4(c) with the knowledge of potential WiFi offloading. Both D2D and cellular action areas in Fig. 4(c) are smaller than those in Fig. 4(b), while the waiting action area in Fig. 4(c) is larger than that in Fig. 4(b); this can be explained by the fact that there are more WiFi APs in the case of Fig. 4(c).

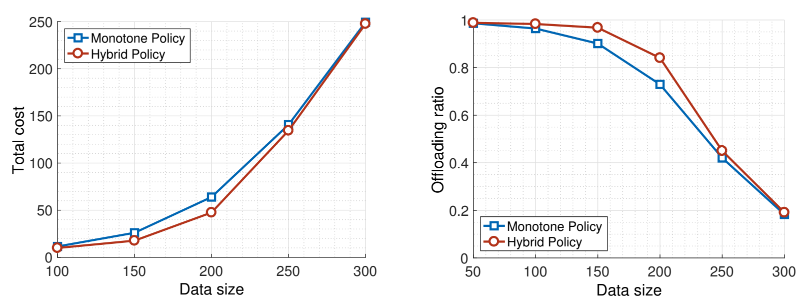
C. Performance Comparisons among Different Schemes

In order to evaluate the performance of our proposed offloading methods, we consider following performance metrics: (1) *Total cost.* The total network cost for data transmission; (2) *Completion time.* The total time used for data transmission; (3) *Offloading ratio.* The percentage of cellular traffic that MNO transmits through WiFi or D2D networks; and (4) *Time usage percentage (TUP).* The ratio of completion time to the deadline. We compare our proposed

TABLE II
DIFFERENT OFFLOADING SCHEMES

Abbreviation	Schemes
D4	Delayed optimal offloading with 4 actions
D3	Delayed optimal offloading with 3 actions
ND4	Non-Delayed offloading with 4 actions
ND3	Non-Delayed offloading with 3 actions
NO	No Offloading

scheme (we name $D4$) with four benchmark schemes (see Table II). The abbreviation of each scheme includes a digit that indicates the available actions for a system state (i.e., 4 indicates that $\mathcal{A} = \{1, 2, 3, 4\}$ and 3 indicates that $\mathcal{A} = \{1, 2, 3\}$). The benchmark schemes include: (1) optimal delayed WiFi offloading scheme ($D3$) [28]; (2) on-the-spot WiFi offloading scheme ($ND3$) [30]: data transmission is switched between WiFi and cellular networks. WiFi network is used whenever available; and (3) on-the-spot WiFi and D2D offloading scheme ($ND4$) [31]: WiFi network is used wherever available; D2D communication is used when MH is available and MS's state satisfies $k < \bar{\nu}_2 * (D - d)$. (4) no offloading scheme (NO): MS only uses cellular network. The results are averaged for a single MS.

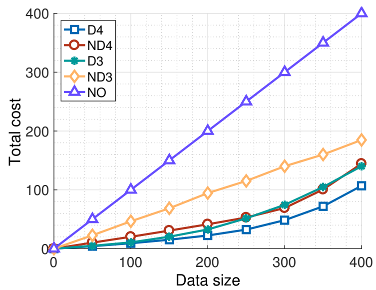


(a) Total cost versus data size K $(Mbytes)$, the deadline D is 200 seconds.

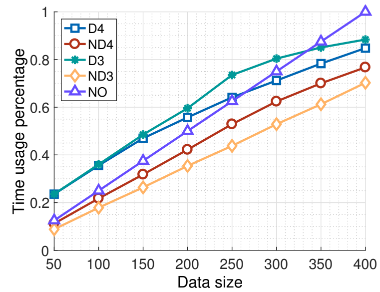
(b) Offloading ratio versus data size K $(Mbytes)$, the deadline D is 200 seconds.

Fig. 5. Performance comparison of Hybrid and Monotone Policies.

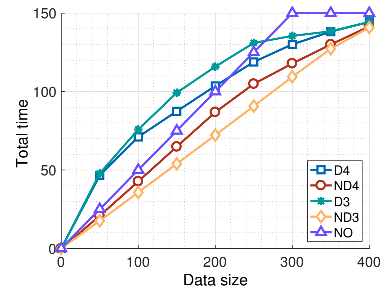
1) Performance Comparison of Hybrid and Monotone Policies: Fig. 5 shows the performance comparison of our proposed hybrid offloading policy and monotone policy. We observe that hybrid policy has better performance in total cost (see Fig. 5(a)) and offloading ratio (see Fig. 5(b)). This is because that Algorithm 1 generates the hybrid policy based on location dependent data rates of different networks, while Algorithm 2 calculates the monotone policy based on the average data rates of different networks. Thus, hybrid policy achieves the better performance at



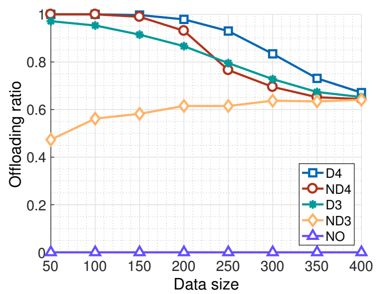
(a) Total cost versus data size K, the deadline D is 200 seconds.



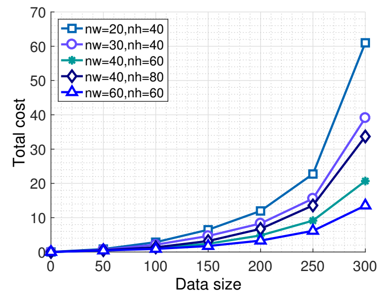
(b) Time usage percentage versus data size K, the deadline D is 140 seconds.



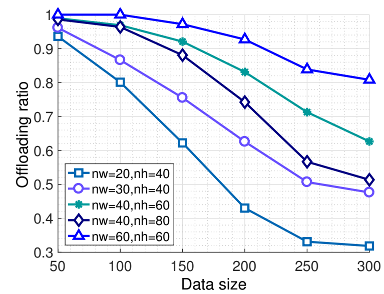
(c) Completion time versus data size K, the deadline D is 150 seconds.



(d) Offloading ratio versus data size K, the deadline D is 140 seconds.



(e) Total cost versus data size K, the deadline D is 100 seconds.



(f) Offloading ratio versus data size K, the deadline D is 100 seconds.

Fig. 6. Performance comparison with fixed deadline.

the cost of higher computational complexity.

2) Performance Comparison with Fixed Deadline: We compare the performance metrics of different schemes with a given deadline D. Fig. 6(a) shows that the total cost increases with data size. We observe that $D4$ outperforms the other schemes by achieving the lowest total cost for any data size. Note that when $K < 250$ Mbytes, $D3$ outperforms $ND4$; it is not the case when $K > 250$ Mbytes. This can be explained by the fact that $ND4$ can use D2D action to offload more data than $D3$ when $K > 250$ Mbytes, while $D3$ can use delayed WiFi offloading to offload more data than $ND4$ when $K < 250$ Mbytes.

Fig. 6(b) shows that TUP increases with data size, since it takes longer to transmit data of larger size. We observe that TUP of non-delayed schemes (i.e., $ND4$ and $ND3$) are smaller than that of delayed schemes (i.e., $D4$ and $D3$). This is because delayed schemes use additional time to wait for offloading opportunities. TUP of $D4$ is smaller than that of $D3$, since $D4$ can use D2D action to offload mobile data when $D3$ is waiting for another WiFi connection.

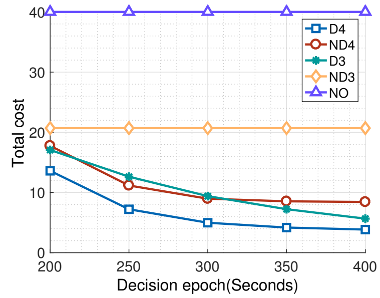
Fig. 6(c) shows that completion time increases with data size. We observe that non-delayed schemes (i.e., $ND3$ and $ND4$) achieve short transmission time compared to delayed schemes (i.e., $D3$ and $D4$). This is because the objective of delayed schemes is to use cellular action, only when it is necessary; indeed, they wait for opportunities to use WiFi and D2D actions (in opposition to non-delayed schemes). We also observe that $D4$ outperforms $D3$ because it makes use of D2D action. However, $ND3$ outperforms $ND4$ for any data size due to the fact that the data rate of D2D action is low. Using D2D action can increase the total time in non-delayed schemes (e.g., $ND3$ outperforms $ND4$) while decreasing the total time in delayed schemes (e.g., $D4$ outperforms $D3$).

Fig. 6(d) shows that the offloading ratio decreases when data size increases except for $ND3$. This is because $ND3$ transmits data based on the available networks without considering current data size k and decision epoch d. We observe that the offloading ratio of $D3$ drops rapidly with the increase of data size, while that of $D4$ and $ND4$ drop slowly. This is because $D4$ and $ND4$ use alternative D2D action to offload data. Notice that $D4$ has the highest offloading ratio. We also observe that the offloading ratios for delayed and non-delayed schemes are the same when $K = 400$ Mbytes. This can be explained by the fact that 400 Mbytes is the transmission limit under the setting used in our simulations. Since all offloading schemes try to complete data delivery before deadline, they use WiFi network wherever possible and cellular network when WiFi network is not available, which is the offloading policy used by $ND3$. It means that all other offloading polices (i.e., $D4$, $ND4$, and $D3$) degenerate to the policy $ND3$. Notice that, $ND4$ can offload more data than $D3$ when $K < 235$ Mbytes, while $D3$ can offload more data than $ND4$ when $K > 235$ Mbytes. This is because $D3$ uses delayed policy, while $ND4$ uses alternative D2D action.

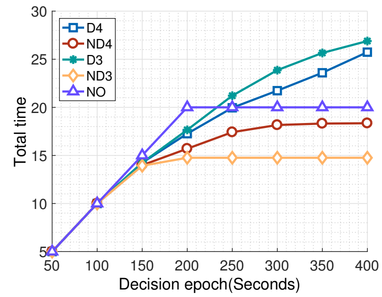
In Figs. 6(e) and 6(f), we investigate the impact of n_w and n_h on total cost and offloading ratio for $D4$. We observe that the total cost decreases and the offloading ratio increases with the increase of the number of WiFi APs and MHs.

3) Performance comparison with fixed data size: We compare the performance metrics of different schemes with a given data size K.

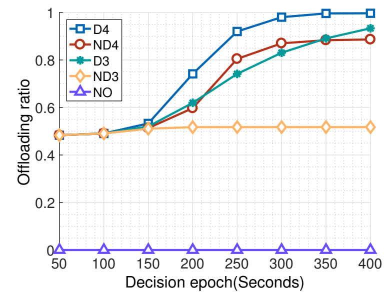
Fig. 7(a) shows that, for delayed schemes (i.e., $D4$ and $D3$) and $ND4$, the total cost decreases when the deadline increases; indeed, larger deadlines give more opportunities (i.e. more time) to look for WiFi and D2D actions to transmit data. We observe that the total cost of $D3$ is larger than that of $ND4$ when $D < 300$ Mbytes. However, the situation changes when $D > 300$ Mbytes. This is because $D3$ uses waiting based strategy for seeking WiFi offloading opportunities; larger deadlines imply more WiFi offloading opportunities. Note that $D4$ incurs the minimum total



(a) Total cost versus deadline D, data size K is 400 Mbytes.



(b) Completion time versus deadline D, data size K is 400 Mbytes.



(c) Offloading ratio versus deadline D, data size K is 400 Mbytes.

Fig. 7. Performance comparison with fixed data size.

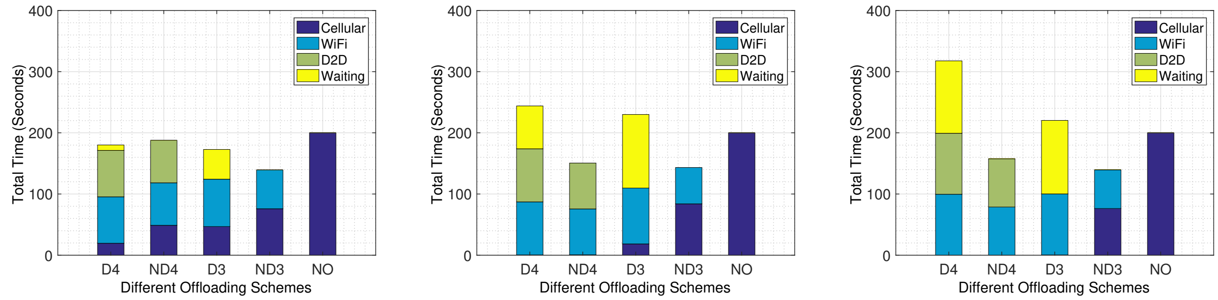
cost compared to other schemes.

Fig. 7(b) shows that the total transmission time increases with the deadline. We observe that, for $ND3$, the total time does not increase when $D > 200$ seconds. This is because $ND3$ uses on-the-spot strategy and cannot make use of the delay tolerance. The total time for delayed schemes (e.g., $D3$ and $D4$) increases almost linearly with the deadline; indeed $D3$ and $D4$ use delayed time to seek for offloading opportunities. Moreover, $D4$ uses slightly less total time than $D3$ when the deadline increases. However, $D4$ can offload more data than $D3$, as shown in Fig. 7(c). This is because D2D action can be used to transmit data when WiFi action is not available.

In Fig. 7(c), we observe that offloading ratio increases with the deadline for delayed schemes (i.e., $D3$ and $D4$). This is because delayed schemes can take advantage of the delay tolerance to seek offloading opportunities through WiFi and D2D actions. We also observe that offloading ratio for $ND3$ does not increase with the deadline, while offloading ratio for $ND4$ increases with the deadline. This is because, for $ND4$, MS has more opportunities to use D2D action as the deadline increases. We conclude that $D4$ achieves the maximum offloading ratio while satisfying data transmission deadline.

4) Simulation Results for Impact of Delay Tolerance: We evaluate the impact of delay tolerance on the transmission time and the amount of data transmitted by different networks, as shown in Figs. 8 and 9, respectively. Three scenarios with low, middle, and high delay tolerance are considered by setting D to 200, 300 and 400 seconds, respectively. The data size K is set to 400 Mbytes.

In Fig. 8, we observe that higher delay tolerance results in better data offloading performance

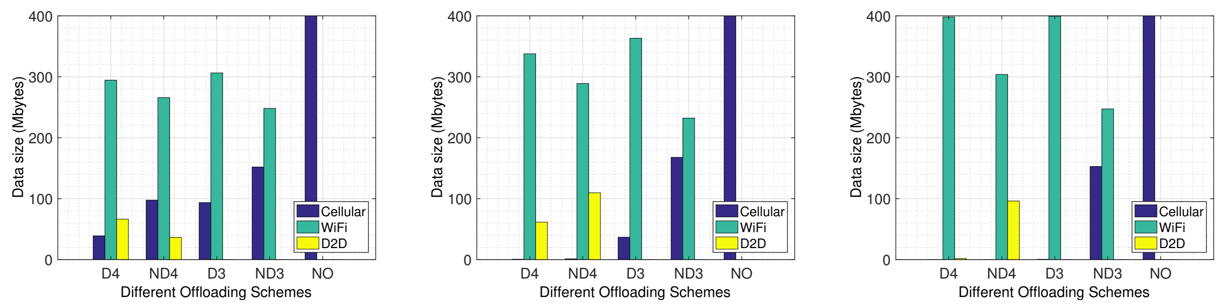


(a) $K = 400$ Mbytes, and $D = 200$ seconds.

(b) $K = 400$ Mbytes, and $D = 300$ seconds.

(c) $K = 400$ Mbytes, and $D = 400$ seconds.

Fig. 8. Total completion time comparison for different schemes, with same data size $K = 400$ Mbytes and different deadline D.



(a) $K = 400$ Mbytes, and $D = 200$ seconds.

(b) $K = 400$ Mbytes, and $D = 300$ seconds.

(c) $K = 400$ Mbytes, and $D = 400$ seconds.

Fig. 9. Data transmitted comparison for different schemes, with same data size $K = 400$ Mbytes and different deadline D.

for delayed schemes (i.e., $D4$ and $D3$) at the cost of longer completion time. To offload the same size of data, less cellular time is used in high delay tolerance scenario. We also observe that the completion time for delayed schemes is larger than non-delayed schemes in high delay tolerance scenario. This is because delayed schemes look for more offloading opportunities by extending the waiting time. Fig. 8 shows that the waiting time for $D4$ and $D3$ grows significantly when D increases. Note that the completion time of $ND3$ is always smaller than that of NO, due to the fact that $\nu_3 > \nu_2$. This implies that using WiFi network can reduce the completion time. However, the completion time of $ND4$ is not always smaller than that of NO, as shown in Fig. 8(a). This can be explained by the fact that although WiFi network can reduce the completion time, the usage of D2D action may increase the completion time, due to the fact that $\nu_4 < \nu_2$. Fig. 8 shows that $ND4$ uses less cellular time than $D3$ when $D = 200$ and $D = 300$ seconds. However, $D3$ outperforms $ND4$ when $D = 400$ seconds. This is because $D3$ can offload more cellular traffic with higher delay tolerance. Furthermore, Fig. 8 shows that $D4$ uses the minimum cellular time in all situations.

In Fig. 9, we evaluate the impact of delay tolerance on the amount of data transmitted by different networks. For each scheme, the total amount of data transmitted by different networks is equal to 400 Mbytes. We observe that the cellular data size of delayed schemes (e.g., $D4$ and $D3$) decreases as the deadline increases, due to the fact that a higher delay tolerance can provide more opportunities of delivering data using WiFi or D2D action. Figs. 9(a), 9(b) and 9(c) shows that for $D4$, with the increase of WiFi data size, cellular data size decreases accordingly. However, D2D data size first increases and then decreases with the increase of deadline D. The reason behind this phenomenon is that when higher delay tolerance is allocated, it is more likely to have more D2D connections or larger D2D data size. However, higher delay tolerance also increases the WiFi data size with lower cost $\chi_3 < \chi_4$. Thus, part of D2D offloading is replaced by WiFi offloading when more WiFi connections are possible. Moreover, compared to other schemes, the cellular data size of $D4$ decreases faster with higher delay tolerance, as shown in Figs. 9(a), 9(b) and 9(c). This demonstrates that the proposed $D4$ can offload more cellular data in practice.

VIII. CONCLUSION

This paper proposed a hybrid data offloading model, where MNO can use WiFi network and D2D communication to offload mobile data of MSs. We formulated the mobile data offloading problem as an FHMDP and proposed a hybrid offloading algorithm for delay sensitive and delay tolerant applications. Moreover, we established sufficient conditions for the existence of thresholds in monotone policy and proposed a monotone offloading algorithm which can reduce the computational complexity caused by large data size and long deadline. The simulation results demonstrate that, compared to existing offloading schemes, our proposed schemes can achieve minimal data offloading cost and maximum offloading ratio.

REFERENCES

[1] Cisco, “Cisco visual networking index: Global mobile data traffic forecast update, 2015-2020,” Cisco, White Paper, 2016.

[2] X. Chen, L. Jiao, W. Li, and X. Fu, “Efficient multi-user computation offloading for mobile-edge cloud computing,” *IEEE/ACM Transactions on Networking*, vol. 24, no. 5, pp. 2795–2808, October 2016.

[3] K. Lee, J. Lee, Y. Yi, I. Rhee, and S. Chong, “Mobile data offloading: How much can WiFi deliver?” *IEEE/ACM Transactions on Networking (TON)*, vol. 21, no. 2, pp. 536–550, 2013.

[4] J. Lee, Y. Yi, S. Chong, and Y. Jin, "Economics of WiFi offloading: Trading delay for cellular capacity," *Wireless Communications, IEEE Transactions on*, vol. 13, no. 3, pp. 1540–1554, 2014.

[5] O. B. Yetim, M. Martonosi, O. Bilgir Yetim, and M. Martonosi, "Adaptive usage of cellular and WiFi bandwidth: An optimal scheduling formulation," in *Proceedings of the seventh ACM international workshop on Challenged networks*. ACM, 2012, pp. 69–72.

[6] G. Iosifidis, L. Gao, J. Huang, and L. Tassiulas, "A double-auction mechanism for mobile data-offloading markets," *IEEE/ACM Transactions on Networking (TON)*, vol. 23, no. 5, pp. 1634–1647, 2015.

[7] X. Zhuo, W. Gao, G. Cao, and S. Hua, "An incentive framework for cellular traffic offloading," *Mobile Computing, IEEE Transactions on*, vol. 13, no. 3, pp. 541–555, 2014.

[8] K. Fall, "A delay-tolerant network architecture for challenged internets," in *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*. ACM, 2003, pp. 27–34. [Online]. Available: http://dl.acm.org/citation.cfm?id=863955.863960

[9] Y. Li, M. Qian, D. Jin, P. Hui, Z. Wang, and S. Chen, "Multiple mobile data offloading through disruption tolerant networks," *IEEE Transactions on Mobile Computing*, vol. 13, no. 7, pp. 1579–1596, July 2014.

[10] R.-G. Cheng, N.-S. Chen, Y.-F. Chou, and Z. Becvar, "Offloading multiple mobile data contents through opportunistic device-to-device communications," *Wireless Personal Communications*, vol. 84, no. 3, pp. 1963–1979, 2015.

[11] D. Liu, L. Khoukhi, and A. Hafid, "Data offloading in mobile cloud computing: A markov decision process approach," in *Communications (ICC), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1–6.

[12] M. Chen, Y. Hao, Y. Li, C.-F. Lai, and D. Wu, "On the computation offloading at ad hoc cloudlet: architecture and service modes," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 18–24, 2015.

[13] M. V. Barbera, S. Kosta, A. Mei, and J. Stefa, "To offload or not to offload? the bandwidth and energy costs of mobile cloud computing," *Proceedings - IEEE INFOCOM*, pp. 1285–1293, 2013.

[14] S. Deng, L. Huang, J. Taheri, and A. Y. Zomaya, "Computation Offloading for Service Workflow in Mobile Cloud Computing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 12, pp. 3317–3329, 2015.

[15] W. Song and W. Zhuang, *Interworking of wireless LANs and cellular networks*. Springer Science & Business Media, 2012.

[16] V. A. Siris and D. Kalyvas, "Enhancing mobile data offloading with mobility prediction and prefetching," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 17, no. 1, pp. 22–29, 2013.

[17] N. Cheng, N. Lu, N. Zhang, X. S. Shen, and J. W. Mark, "Opportunistic WiFi offloading in vehicular environment: A queueing analysis," in *Global Communications Conference (GLOBECOM), 2014 IEEE*. IEEE, 2014, pp. 211–216.

[18] F. Mehmeti and T. Spyropoulos, "Performance Analysis of Mobile Data Offloading in Heterogeneous Networks," *GLOBECOM - IEEE Global Telecommunications Conference*, pp. 1577–1583, 2013.

[19] B. H. Jung, N.-O. O. Song, and D. K. Sung, "A network-assisted user-centric WiFi-offloading model for

maximizing per-user throughput in a heterogeneous network," *Vehicular Technology, IEEE Transactions on*, vol. 63, no. 4, pp. 1940–1945, 2014.

[20] V. F. S. V. F. S. Mota, D. F. Macedo, Y. Ghamri-Doudanez, and J. J. M. S. Nogueira, "Managing the decision-making process for opportunistic mobile data offloading," in *Network Operations and Management Symposium (NOMS), 2014 IEEE*. IEEE, 2014, pp. 1–8.

[21] V. Sciancalepore, D. Giustiniano, A. Banchs, A. Hossmann-Picu, A. Picu, and A. Hossmann-Picu, "Offloading cellular traffic through opportunistic communications: Analysis and optimization," *Selected Areas in Communications, IEEE Journal on*, vol. 34, no. 1, pp. 122–137, 2016. [Online]. Available: http://arxiv.org/abs/1405.3548

[22] F. Rebecchi, M. Dias de Amorim, and V. Conan, "Droid: Adapting to individual mobility pays off in mobile data offloading," in *Networking Conference, 2014 IFIP*. IEEE, 2014, pp. 1–9.

[23] S. Andreev, O. Galinina, A. Pyattaev, K. Johnsson, and Y. Koucheryavy, "Analyzing assisted offloading of cellular user sessions onto D2D links in unlicensed bands," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 1, pp. 67–80, 2015.

[24] A. Kolobov, "Planning with Markov decision processes: An AI perspective," *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 6, no. 1, pp. 1–210, 2012.

[25] F. Calabrese, G. D. Lorenzo, and C. Ratti, "Human mobility prediction based on individual and collective geographical preferences," in *13th International IEEE Conference on Intelligent Transportation Systems*, Sept 2010, pp. 312–317.

[26] D. Liu, L. Khoukhi, and A. Hakim, "Data offloading in mobile cloud computing based on markov decision process," 2017. [Online]. Available: https://hollyliu2008.github.io/liudongqing.github.io

[27] V. Krishnamurthy, *Partially Observed Markov Decision Processes*. Cambridge University Press, 2016.

[28] M. H. Cheung and J. Huang, "Dawn: Delay-aware wi-fi offloading and network selection," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 6, pp. 1214–1223, June 2015.

[29] C.-H. H. Yu, K. Doppler, C. B. C. B. Ribeiro, and O. Tirkkonen, "Resource sharing optimization for device-to-device communication underlaying cellular networks," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 8, pp. 2752–2763, 2011.

[30] F. Mehmeti and T. Spyropoulos, "Performance analysis of 'on-the-spot' mobile data offloading," *GLOBECOM - IEEE Global Telecommunications Conference*, no. February 2016, pp. 1577–1583, 2013.

[31] S. Andreev, M. Gerasimenko, O. Galinina, Y. Koucheryavy, N. Himayat, S. P. Yeh, and S. Talwar, "Intelligent access network selection in converged multi-radio heterogeneous networks," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 86–96, December 2014.