Offloading Mobile Traffic via Green Content Broker

Tao Han, *Student Member, IEEE,* and Nirwan Ansari, *Fellow, IEEE*

Abstract—A continuous surge of mobile data traffic not only congests mobile networks but also results in a dramatic increase in the energy consumption of mobile networks. Device-to-device (D2D) communications is a promising technique for offloading mobile data traffic and enhancing energy efficiency of mobile networks. By enabling D2D communications, the base stations (BSs) may reduce their energy consumption through traffic offloading, and mobile users may increase their quality of services by retrieving contents from their neighboring peers instead of the remote BSs. By leveraging D2D communications, we propose a novel mobile traffic offloading scheme—the content brokerage. In the content brokerage scheme, a new network node called the green content broker (GCB) is introduced to arrange the content delivery between the content requester and the content owner. The GCB is powered by green energy, e.g., solar energy, to reduce the CO_2 footprints of mobile networks. In the scheme, maximizing traffic offloading with the constraints of the amount of green energy and bandwidth is an nondeterministic polynomial time (NP) problem. We propose a heuristic traffic offloading (HTO) algorithm to approximate the optimal solution with low computational complexity. Our simulation results valid the performance of the content brokerage scheme and the HTO algorithm.

Index Terms—Device-to-device (D2D) communications, energy efficiency, mobile networks, mobile traffic offloading.

I. INTRODUCTION

WITH the rapid development of radio access techniques and mobile devices, a variety of bandwidth-hungry applications and services such as web browsing, video streaming, and social networking are gradually shifted to mobile networks, thus leading to an exponential increase in data traffic in mobile networks. The mobile data traffic surges bring about two major problems to current mobile networks. 1) The significant data increase congests mobile networks and leads to a long delay in content delivery [1]. 2) A continuous surge of mobile traffic results in a dramatic increase in energy consumption in mobile networks for provisioning higher network capacity [2].

Mobile traffic offloading, which is referred to as utilizing complementary network communication techniques to deliver mobile traffic, is a promising technique to alleviate congestion and reduce the energy consumption of mobile networks [1]. Based on the network access mode, the mobile traffic offloading schemes can be divided into two categories. The first category is the infrastructure-based mobile traffic offloading,

which refers to as deploying small cell base stations (BSs), e.g., pico-BSs, femto BSs, and WiFi hot spots, to offload mobile traffic from macro BSs (MBS) [1]. Deploying small cell BSs can efficiently offload mobile traffic from MBSs, thus reducing traffic congestion and the energy consumption of mobile networks [3]. However, the lack of cost-effective backhaul connections for small cell BSs often impairs their performance in terms of offloading mobile traffic and enhancing the energy efficiency of mobile networks.

The second category is the *ad-hoc*-based mobile traffic offloading, which refers to applying device-to-device (D2D) communications as an underlay to offload mobile traffic from MBSs. By leveraging Internet of Things (IoT) technologies, smart devices within proximity are able to connect with each other and form a communication network. Data traffic among the devices can be offloaded to the communications networks rather than delivering through MBSs. For example, by enabling D2D communications, some user equipments (UEs) download contents from MBSs while the other UEs may retrieve the contents through D2D connections with their peers. In this way, D2D communications alleviate traffic congestion and reduce the energy consumption of mobile networks. However, D2D communications may suffer from several disadvantages which impair its performance in terms of offloading mobile traffic. First, UEs are battery-powered devices, and therefore the additional energy consumption may prevent UEs from participating in D2D communications. Second, the transmission range for D2D communications among mobile devices may be limited by its low transmission power. For example, if a mobile device experiences a shortage of battery power, the mobile device may restrict its power usage in D2D communications. This leads to a reduced transmission range for the mobile device's D2D communications. Especially, when the mobile devices operate on millimeter wavelength, the transmission range is further limited by the channel propagation characteristics. Third, D2D communications may require complicated radio resource management schemes implemented in mobile devices to avoid the extensive interference introduced to mobile networks. This will further increase mobile devices' power consumption in D2D communications. Fourth, sharing content through D2D communications may reveal users' privacy. To protect their privacy, mobile devices may not participate in D2D communications.

In this paper, we propose a novel mobile traffic offloading scheme referred to as the content brokerage, which synergizes the merits of both the infrastructure-based and the *ad-hoc*-based mobile traffic offloading. In the content brokerage scheme, we introduce a new network node called the green content broker (GCB), which arranges the content delivery between the content owners (the servers), e.g., UEs or BSs, and the content requesters (the clients).

The authors are with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: th36@njit.edu; nirwan.ansari@njit.edu).

The content broker discovers the available contents among the UEs under its coverage, collects the content requests, and forwards the contents from the content owners to the content requester.

The proposed content brokerage scheme exhibits several advantages.

1) As compared with the mobile traffic offloading by deploying small cell BSs, the content brokerage scheme does not require backhaul connections. This not only reduces the operation costs of offloading mobile traffic, but also increases the flexibility of deploying the GCB and increases the efficiency of traffic offloading.

2) As compared with *ad-hoc*-based mobile traffic offloading, the GCB can facilitate the service and content discovery and the D2D communications session management [4]. In addition, the GCB may anonymize mobile traffic to protect user privacy and may maintain content copyrights for the content providers.

3) The content broker is powered by green energy, e.g., solar energy, to avoid the on-grid energy consumption, thus reducing the carbon footprints. Powering by green energy further increases the flexibility of deploying the GCB.

In the content brokerage scheme, the GCB handles multiple content requests from UEs. Given the amount of available radio channels and green energy, the GCB may not be able to serve all the content requests. To maximize the traffic offloading, and thus minimize the energy consumption of the MBS, the GCB should select to serve a portion of content requests. In selecting the serving contents, the GCB should consider the radio channel constraints in both the uplink and downlink transmissions. Here, the uplink transmissions refer to as radio channels used by the GCB to retrieve contents from UEs or the MBS, and the downlink transmissions refer to as the radio channels used by the GCB utilizes to deliver contents to the clients. Since the GCB is powered by green energy, it also experiences the green energy constraints, which limit the amount of mobile traffic that can be offloaded to the GCB. In addition, the requested content may be owned by multiple UEs. Selecting different content owners (UEs) leads to a different uplink radio channel requirement. Considering these constraints, the traffic offloading maximization (TOM) problem is an NP-hard problem. To approximate the optimal solution with low computational complexity, we divide the TOM problem into two subproblems: the serving content selection (SCS) problem and the content owner selection (COS) problem. We design a heuristic traffic offloading (HTO) algorithm which iteratively solves these subproblems, and subsequently solves the TOM problem.

This paper is organized as follows. In Section II, we provide an overview on the related research efforts. In Section III, we define the system model and the assumptions. In Section IV, we formulate the TOM problem and analyze its properties. Section V presents the proposed heuristic algorithm. Section VI shows the simulation results, and concluding remarks are presented in Section VII.

II. RELATED WORK

In this section, we briefly overview the related works on infrastructure-based mobile traffic offloading, *ad-hoc*-based mobile traffic offloading, and the utilization of green energy in cellular networks.

A. Infrastructure-Based Mobile Traffic Offloading

In infrastructure-based mobile traffic offloading, the mobile traffic is offloading to either pico/femto BSs or WiFi hot spots. Deploying pico/femto BSs improves the spectral and energy efficiency per unit area of cellular networks, and thus reduces the network congestion and energy consumption of cellular networks. Traffic offloading between pico/femto BSs and the MBS is achieved by adapting the user-BS associations. Kim *et al.* [5] proposed a user-BS association to achieve flow level load balancing under spatially heterogeneous traffic distribution. Jo *et al.* [6] proposed cell biasing algorithms to balance traffic loads among pico/femto BSs and the MBS. The cell biasing algorithms perform user-BS association according to the biased measured pilot signal strength, and enables the traffic to be offloaded from the MBS to pico/femto BSs.

WiFi hot spots are also effective in terms of offloading mobile traffic. Lee *et al.* [7] pointed out that a user is in WiFi coverage for 70% of the time on average, and if users can tolerate a 2 h delay in data transfer, the network can offload about 70% cellular traffic to WiFi networks. Balasubramanian *et al.* [8] proposed to offload the delay tolerance traffic such as email and file transfer to WiFi networks. When WiFi networks are not available or experiencing blackouts, data traffic is fast switched back to 3G networks to avoid violating the applications' tolerance threshold. Han and Ansari [9] designed a content pushing system which pushes the content to mobile users through opportunistic WiFi connections. The system responds to a user's pending requests or predicted future requests, codes these requested contents by using Fountain codes, predicts the user's routes, and prelocates the coded contents to the WiFi access points along the user's routes. When the user connects to these WiFi access points, the requested contents are delivered to the user via the WiFi connections.

B. Ad-hoc-Based Mobile Traffic Offloading

Ad-hoc-based mobile traffic offloading rely on D2D communications to disseminate contents. Instead of downloading contents directly from BSs, UEs may retrieve contents from their neighboring UEs. Han *et al.* [10] proposed a mechanism to select a subset of UEs based on either UEs' activities or mobilities, and to deliver contents to them through cellular networks, and let these UEs further disseminate the contents through D2D communications to the other users. Mashhadi and Hui [11] proposed a proactive caching mechanism for UEs in order to offload the mobile traffic. When the local storage does not have the requested contents, the proactive caching mechanism will set a target delay for this request, and explores opportunities to retrieve data from the neighboring UEs. The proactive cache mechanism requests data from cellular networks when the target delay is violated. To encourage mobile users participate in the traffic offloading, Zhou *et al.* [12] proposed an incentive framework that motivate users to leverage their delay tolerance for cellular data offloading.
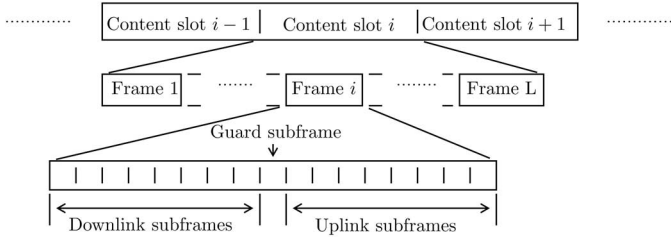
Fig. 1. Content slot structure.

Fig. 2. Communications procedure of the content brokerage scheme.

C. Utilization of Green Energy

Green energy techniques, in which green energy such as sustainable biofuels, solar, and wind energy is able to power BSs, can save on-grid energy consumption of BSs and reduce the CO_2 footprints [13]. Ericson Inc. [14] has developed a wind-powered tower for BSs of cellular networks. To optimize the utilization of renewable energy, Zhou *et al.* [15] proposed the Hand Over (HO) parameter tuning algorithm and the power control algorithm to guide mobile users to access the BSs with renewable energy supply. Han and Ansari [16] proposed an energy aware cell size adaptation algorithm named intelligent cell breathing (ICE), which balances the energy consumption among BSs powered by green energy, and enables more users to be served with green energy. Considering a network with multiple energy supplies, Han and Ansari [17], [18] also proposed to optimize the utilization of green energy, and reduce the on-grid energy consumption of cellular networks by the cell size optimization.

III. SYSTEM MODEL

In this paper, we consider a cellular network consisting of one MBS, N UEs, and one GCB. Denote $\mathcal{U} = (u_1, u_2, \ldots, u_N)$ as the set of UEs. We assume that orthogonal channels are available in the network, e.g., orthogonal frequency-division multiple access (OFDMA), to avoid interference [19] among the MBS, GCB, and UEs. We assume UEs and GCB can transmit and receive simultaneously. The time horizon is divided into multiple content slots with a duration of τ seconds. One content slot contains L transmission frames, which can be further divided into multiple uplink and downlink subframes. The total number of uplink and downlink subframes in one transmission frame is defined as $N_{\max}^u$ and $N_{\max}^d$, respectively. The bandwidth allocated in each subframe is w. The content slot structure is shown in Fig. 1.

At the beginning of the kth content slot, the GCB collects the UEs' content requests and registrations which are submitted in the $(k-1)$th content slot. Here, we assume the UE's content requests are postponed at most one content slot for seeking the traffic offloading opportunities. In other words, the UEs' content requests in the $(k-1)$th content slot either are fulfilled by the GCB in the kth content slot or are served by the MBS. The content request indicates which content a UE is retrieving while the content registration shows which contents are available in the UE. Then, the GCB tries to match the content requests and registrations in offloading traffic from the MBS. If a UE's content
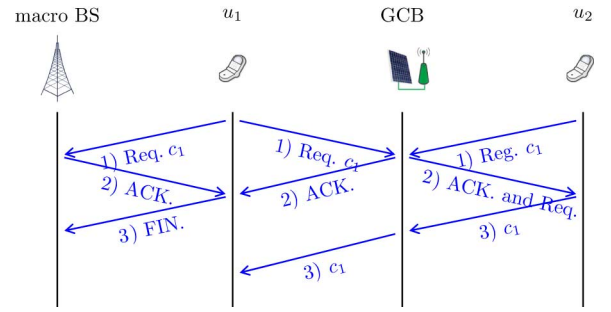
request can be fulfilled locally, the GCB sends an acknowledge message to inform the UE; otherwise, the GCB sends an FIN. message to terminate the connection with the UE. Upon receiving the acknowledgment, the UE sends an FIN. message to the MBS to cancel the content request. Meanwhile, the GCB retrieves the requested content from the UE which registers the content and delivers it to the requesting UE. Fig. 2 shows an example to illustrate the communications procedures in the content brokerage scheme. At the content slot, u_1 requests the content c_1 from the MBS and the GCB. Meanwhile, u_2 registers the content c_1. Then, since c_1 is locally available, the GCB sends an acknowledgment message to u_1 to indicate that its content request can be fulfilled locally. Upon receiving the acknowledgment, u_1 cancels the content request from the MBS by sending an FIN. message. The GCB acknowledges u_2's content registration and retrieves the content c_1 from it. Then, the GCB forwards the content c_1 to u_1.

The content brokerage scheme introduces additional signaling overhead. However, these additional signaling and control messages can be integrated with the control messages required for establishing D2D networks. Forming a D2D network usually involves processes such as the device discovery, device association, and channel allocation. The content request and registration information can be integrated within the device discovery and association messages. For example, when a mobile device is associated with the D2D network, it informs the GCB what content to be cached in the device and what content to be requested by the device. In this way, the control and signaling overhead of the content brokerage scheme will be reduced and will not significantly impact the efficiency of D2D communications.

We assume that there are M contents in the network, and denote $\mathcal{C} = (c_1, c_2, \ldots, c_M)$ and $\mathcal{S} = (s_1, s_2, \ldots, s_M)$ as the sets of the contents and their sizes, respectively. We assume a UE can only request for one content in a content slot, but can register multiple contents in a content slot. Denote $\eta_{i,j}(k)$ and $\varphi_{i,j}(k)$ as a content request indicator and a content registration indicator, respectively. If $\eta_{i,j}(k) = 1$, it indicates the ith UE requests the jth content in the kth content slot. Since a UE can only request one content in a content slot, $\sum_{j \in \mathcal{C}} \eta_{i,j}(k) \leq 1$. If the ith UE registers the jth content in the kth content slot, $\varphi_{i,j}(k) = 1$; otherwise, $\varphi_{i,j}(k) = 0$.

Denote p^c as the transmission power of the GCB, and $h_i^c(k)$ as the channel gain between the ith UE and the GCB in the kth content slot. Here, we assume UEs' channel gains do not change

within a content slot. The downlink data rate of a subframe for the ith UE associated with the GCB is

$$r_i^c(k) = w \log_2 \left(1 + \frac{p^c h_i^c(k)}{w\sigma^2} \right) \tag{1}$$

where σ^2 is the thermal noise density. Supposing the ith UE requests the jth content in the kth content slot, the number of subframes required in the GCB to serve the content request is

$$n_i^c(k) = \frac{s_j}{r_i^c(k)}. \tag{2}$$

Denote p^m as the transmission power of the MBS, and $h_i^m(k)$ as the channel gain between the ith UE and the MBS in the kth content slot. The downlink data rate of a subframe for the ith UE associated with the MBS is

$$r_i^m(k) = w \log_2 \left(1 + \frac{p^m h_i^m(k)}{w\sigma^2} \right). \tag{3}$$

The number of subframes in the MBS required to serve the ith UE's content request is

$$n_i^m(k) = \frac{s_j}{r_i^m(k)}. \tag{4}$$

The GCB's energy consumption consists of two parts: 1) the static power consumption and 2) the dynamic power consumption [20]. The static power consumption is the power consumption of the GCB without any traffic load. The dynamic power consumption refers to the additional power consumption caused by traffic loads in the GCB, which can be well approximated by a linear function of the traffic load [20]. Since a higher traffic load requires a large number of subframes, we model the energy consumption of the GCB as a linear function of the number of transmitting subframes [21]. The more subframes that the GCB is transmitting, the higher the energy consumption of the GCB. Then, the energy consumption of the GCB in the kth content slot can be expressed as

$$C^c(k) = \alpha^c \sum_{j \in \mathcal{C}} (\gamma_j(k) \max_{i \in \mathcal{U}} \{n_i^c(k)\eta_{i,j}(k)\}) + C_s^c \tag{5}$$

where α^c is a linear coefficient which reflects the relationship between the number of transmitting subframes and the GCB's dynamic energy consumption, and C_s^c is the GCB's static energy consumption. $\gamma_j(k)$ is a content serving indicator. If the jth content is served by the GCB in the kth content slot, $\gamma_j(k) = 1$; otherwise, $\gamma_j(k) = 0$.

We adopt the same energy consumption model for the MBS. Denote α^m and C_s^m as the linear coefficient and the static energy consumption of the MBS, respectively. The MBS consumes energy in two cases. For the first case, the MBS consumes energy to deliver contents to the UEs which are not served by the GCB. For the second case, the MBS consumes energy to deliver the requested contents to the GCB. In this case, a UE is served by the

GCB, but the requested content is not registered by the other UEs. Thus, the GCB retrieves the content from the MBS. Considering both cases, the MBS's energy consumption is

$$C^m(k) = \alpha^m \sum_{j \in \mathcal{C}} \left[(1 - \gamma_j(k)) \max_{i \in \mathcal{U}} \{n_i^m(k)\eta_{i,j}(k)\} + \mu_j(k) \frac{s_j}{r_c^m} \right] + C_s^m \tag{6}$$

where $\mu_j(k)$ indicates whether the GCB retrieves the j content from the MBS. If the GCB retrieves the jth content from the MBS, $\mu_j(k) = 1$; otherwise, $\mu_j(k) = 0$. r_c^m is the per subframe data rate for the transmission from the MBS to the GCB, and

$$r_c^m = w \log_2 \left(1 + \frac{p^m h_c^m}{w\sigma^2} \right) \tag{7}$$

where h_c^m is the channel fading between the MBS and the GCB.

In this paper, we assume that the GCB is powered by green energy generated by its solar panel. Denote $e(k)$, $b^{\max}$, and $b^r(k)$ as the energy generation rate, the maximum battery capacity, and the residual energy storage in the kth content slot, respectively. The GCB's available green energy in the kth content slot is

$$b(k) = \min\{e(k)\tau + b^r(k), b^{\max}\} \tag{8}$$

where

$$b^r(k) = \min\{e(k-1)\tau + b^r(k-1) - C^c(k-1), b^{\max}\}. \tag{9}$$

IV. PROBLEM FORMULATION AND ANALYSIS

The content brokerage scheme aims to maximize the traffic offloading, and thus reduces the energy consumption of the MBS. Since the GCB is powered by green energy, the maximum amount of traffic offloading is constrained by available green energy. We assume that the number of downlink subframes in each transmission frame within the same content slot is the same. Define the GCB's capacity as the number of downlink subframes that can be supported in a content slot by green energy. Denote $N_{\max}^c(k)$ as the GCB's capacity in the kth content slot

$$N_{\max}^c(k) = \min \left\{ \frac{C^c(k) - C_s^c}{\alpha^c}, N_{\max}^d L \right\}. \tag{10}$$

Denote p^u as the transmission power of the UE, and $h_i^u(k)$ as the channel gain between the ith UE and the GCB. The data rate of the ith UE in the uplink during the kth content slot is

$$r_i^u(k) = w \log_2 \left(1 + \frac{p^u h_i^u(k)}{w\sigma^2} \right). \tag{11}$$

The number of subframes required for the ith UE to upload the jth content to GCB equals to

$$n_{i,j}^u(k) = \frac{s_j}{r_i^u(k)}. \tag{12}$$

If the ith UE does not have the jth content, we set $n_{i,j}^u(k) = \infty$. According to the system model, maximizing the traffic

offloading equals to minimizing the energy consumption of the MBS. Thus, the TOM problem can be formulated as

$$\min_{\gamma(k),\beta(k)} \quad C^m(k) \tag{13}$$

$$\text{s.t. :} \quad \sum_{j\in\mathcal{C}} \beta_{i,j}(k) \le 1 \quad \forall i \in \mathcal{U}$$

$$\sum_{i\in\mathcal{U}} \beta_{i,j}(k) + \mu_j(k) \le 1 \quad \forall j \in \mathcal{C}$$

$$\sum_{j\in\mathcal{C}} \gamma_j(k) \max_{i\in\mathcal{U}}\{n_i^c(k)\eta_{i,j}(k)\} \le N_{\max}^c(k)$$

$$\sum_{j\in\mathcal{C}} \gamma_j(k) \left(\sum_{i\in\mathcal{U}} \beta_{i,j}(k)n_i^u(k) + \mu_j(k)\frac{s_j}{r_c^m} \right)$$

$$\le N_{\max}^u L \tag{14}$$

where $\gamma(k) = (\gamma_1(k), \ldots, \gamma_j(k), \ldots, \gamma_M(k))$,

$$\beta(k) = \begin{pmatrix} \beta_{1,1}(k), & \cdots & \beta_{1,M}(k) \\ \vdots & \ddots & \vdots \\ \beta_{N,1}(k), & \cdots & \beta_{N,M}(k) \end{pmatrix}. \tag{15}$$

$\beta_{i,j}(k)$ is a content upload indicator. If the ith UE uploads content s_j to the GCB in the kth content slot, $\beta_{i,j}(k) = 1$; otherwise, $\beta_{i,j}(k) = 0$. The TOM problem consists of four constraints. The first constraint is that in one content slot, a UE can only upload one content. The second constraint is that one content can be uploaded by either a UE or the MBS. The third constraint is to guarantee the total number of required subframes is less than the GCB's capacity. The fourth constraint is to make sure that the total number of required uplink subframes is less than the maximum number of uplink subframes.

To solve the TOM problem, two parameters should be determined. The first parameter is $\gamma(k)$ which determines which content is served by the GCB. This problem is referred to as the SCS problem. Given the number of available uplink subframes and its capacity, the GCB may not be able to serve all the requested contents. Different contents may require different resources in terms of uplink and downlink subframes from the GCB. To minimize the MBS's energy consumption, the GCB should select and serve a subset of the requested contents. The second parameter is $\beta(k)$ that determines which UEs are selected to upload the requested contents. This problem is refereed to as the COS problem. Different UEs may have different uplink data rates. Thus, for a given content, selecting different UE to upload the content results in a different requirement on the number of uplink subframes. Therefore, the solution of the SCS problem and the SOC problem is interdependent.

Theorem 1: The TOM problem is an NP-hard problem.

Proof: We prove the theorem by reducing any instance of the double dimensional knapsack problem to a simple case of the TOM problem. The double dimensional knapsack problem is a well-known NP-hard problem.

In the simple case of the TOM problem, we assume that one content has only one content owner in a content slot. In other words, in this simple case, if the serving content is selected, the

corresponding UE which uploads the content to the GCB is determined. Thus, solving the simple case of the TOM problem is equivalent to determine the serving contents. The GCB consumes two resources in serving a content: 1) the uplink subframes and 2) the downlink subframes. Since the total number of uplink and downlink subframes is constrained, the simple case of the TOM problem is actually to offload as many contents as possible to minimize the energy consumption of the MBS with the constraints of the number of uplink and downlink subframes. Therefore, the simple case of the TOM problem can be translated to a double dimensional knapsack problem. In other words, any instance of the double dimensional knapsack problem can be reduced to the simple case of the TOM problem. □

V. HTO ALGORITHM

In this section, we present the HTO algorithm to approximate the optimal solution of the TOM problem. The HTO algorithm iteratively solves the COS problem and the SCS problem, and subsequently solves the TOM problem.

A. Solving the COS Problem

In the content brokerage scheme, a UE may register multiple contents to the GCB, and a content may be registered by multiple UEs. However, in a content slot, a UE can only upload one content to the GCB, and a content can only be retrieved from one content owner by the GCB. From the MBS's point of view, delivering contents to different UEs may consume different amount of energy because 1) the UEs may experience various channel fading, and 2) the size of the content requested by different UEs may be different. Therefore, serving different contents in the GCB results in different amount of energy savings in the MBS. Denote $C_j^s(k)$ as the MBS's energy savings by offloading the jth content to the GCB. Then

$$C_j^s(k) = \alpha^m \left(\max_{i\in\mathcal{U}}\{n_i^m(k)\eta_{i,j}(k)\} - \mu_j(k)\frac{s_j}{r_c^m} \right). \tag{16}$$

UEs may experience different channel conditions when uploading a content to the GCB. Thus, as the content owners, different UEs require different numbers of uplink subframes from the GCB. The number of uplink subframes required by the ith UE to upload the jth content to the GCB is shown in (12). The per uplink subframe energy savings achieved by selecting the ith UE to upload the jth content is

$$\delta_{i,j}(k) = \frac{C_j^s(k)}{n_{i,j}^u(k)}. \tag{17}$$

Since the total number of uplink subframes in the GCB is constrained, to optimize the utilization of the uplink subframes in the term of maximizing the energy savings, the GCB aims to maximize the summation of the per uplink subframe energy savings of all the serving contents. Fig. 3 shows an example of the COS problem. There are three contents in the network, and the MBS's energy savings by offloading these contents to the GCB are 4, 6, and 8, respectively. In Fig. 3, if a UE registers a content, then we connect the UE and the content with an edge. u_1 registers c_1 and c_2, and requires 2 and 4 uplink subframes to upload c_1 and c_2, respectively. u_2 may upload c_1 and c_3 at costs of 4 and 5 uplink
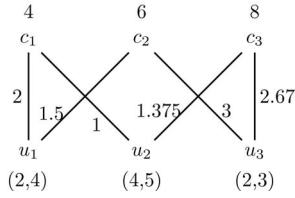
Fig. 3. Illustration of the COS problem.

Fig. 4. Formation of the bipartite graph.

subframes, respectively; and u_3 may upload c_2 and c_3 at costs of 2 and 3 uplink subframes, respectively. The weights on the edges are the per subframe energy savings for the content if it is uploaded by the connected UE. For instance, the weight of the edge between c_1 and u_1 is 2, which indicates the per uplink subframe energy savings is 2 if u_1 upload c_1 to the GCB. The optimal solution of the COS problem is that u_1 uploads c_1, u_2 uploads c_3, and u_3 uploads c_2. In this case, the total number of required uplink subframes, which is 9, is also minimized. Without solving the COS problem, it may requires more uplink subframes to serve these contents, which may violate the constraints of the total number of uplink subframes.

We propose the COS algorithm which translates the COS problem to a maximum weight bipartite matching (MWBM) problem, and solves the COS problem with the MWBM algorithm [22]. Define $C^r = (j | \sum_{i \in \mathcal{U}} \eta_{i,j}(k) > 0 \ \forall j \in \mathcal{C})$ as the set of requested contents in the kth content slot. We form the bipartite graph as shown in Fig. 4. The requested contents, the UEs and the MBS, are the vertices of the graph; if u_i registers c_j, we add an edge between u_i and c_j. Since the MBS is able to deliver any content to the GCB, we add edges between the MBS and every content. We denote u_0 as the MBS during the graph construction. The weights of the edges are calculated according to (17). The pseudo codes of the COS algorithm is shown in Algorithm 1.

Algorithm 1: The COS Algorithm

Input: $\eta_{i,j}(k), \varphi_{i,j}(k), s_j, n_i^m(k), r_c^m \ \forall j \in \mathcal{C} \ \forall i \in \mathcal{U}$;

Output: $\beta(k), \ \mu_j(k) \ \forall j \in \mathcal{C}^r$;

1 Construct a set of $|\mathcal{C}^r|$ vertices corresponding to $\mathcal{C}^r$;

2 Construct a set of $N + 1$ vertices corresponding to $\mathcal{U} \cup \{u_0\}$;

3 Construct a set ε of edges, where $(u_i, c_j) \in \varepsilon$ if $\varphi_{i,j}(k) = 1$ or $i = 0$;

4 for $(u_i, c_j) \in \varepsilon \ \forall u_i \in \mathcal{U} \cup \{u_0\} \ \forall c_j \in \mathcal{C}$ **do**

5   **if** $i == 0$ **then**

6 $\delta_{0,j}(k) = \frac{s_j}{r_c^m}$;

7   **else**

8   Calculate $\delta_{i,j}(k)$;

9 Apply an MWBM algorithm to find the maximum weight matching of the constructed graph;
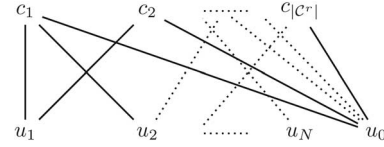
10 Derive $\beta(k)$ and $\mu_j(k) \ \forall j \in \mathcal{C}^r$;

B. HTO Algorithm

Owing to the constraints of green energy and the number of uplink and downlink subframes, the GCB may not be able to serve all the content requests. To maximize the MBS's energy savings, the GCB selects to serve a subset of the content requests. As analyzed in Section IV, when the content owners are selected, the SCS problem can be transformed into a double dimensional knapsack problem. Therefore, we address the SCS problem based on Toyoda's primal effective gradient method in solving the multidimensional knapsack problem [23].

Given $\beta(k)$ and $\mu_j(k)$, the energy savings by offloading the jth content is calculated based on (16), and the number of required uplink and downlink subframes for serving the jth content are derived. We normalize the number of required uplink and downlink subframes of the jth content in the kth content slot with respect to the total number of uplink and downlink subframes, and denote $\psi_j^u(k)$ and $\psi_j^d(k)$ as the normalized number of required uplink and downlink subframes of the jth content, respectively. Then

$$\psi_j^u(k) = \frac{\mu_j(k) \sum_{i \in \mathcal{U}} \beta_{i,j}(k) n_{i,j}^u(k) + (1 - \mu_j(k)) \frac{s_j}{r_c^m}}{N_{\max}^u L} \quad (18)$$

and

$$\psi_j^d(k) = \frac{\max_{i \in \mathcal{U}} \{n_i^m(k) \eta_{i,j}(k)\}}{N_{\max}^d L}. \quad (19)$$

The SCS problem can be expressed as

$$\max_{\gamma(k)} \quad C_j^s(k) \gamma_j(k) \quad (20)$$

$$\text{s.t.:} \quad \sum_{j \in \mathcal{C}} \psi_j^u(k) \gamma_j(k) \leq 1$$

$$\sum_{j \in \mathcal{C}} \psi_j^d(k) \gamma_j(k) \leq 1. \quad (21)$$

Since serving a content consumes both uplink and downlink subframes, it is difficult to evaluate which content consumes less resources in terms of uplink and downlink subframes. In other words, it is difficult to determine offloading which content brings the most profits in the term of energy savings with constrained resources (the uplink and downlink subframes). The evaluation of the resource consumption of a content depends on the current status of the resource consumption. For example, assume $\psi_j^u(k) = 0.4$, $\psi_j^d(k) = 0.1$, $\psi_i^u(k) = 0.1$, and $\psi_i^d(k) = 0.4$. In this case, serving the jth content, the GCB requires more uplink subframes but less downlink subframes than serving the ith content. Without the current status of the usage of the uplink and downlink subframes, it is unable to compare the resource consumption of the jth and that of the ith content. Denote $\mathcal{C}^*$ as the set of contents selected to be served by the GCB. Then,

$\sum_{j \in \mathcal{C}^*} \psi_j^u(k)$ and $\sum_{j \in \mathcal{C}^*} \psi_j^d(k)$ represent the normalized usage of the uplink and downlink subframes, respectively. In order to compare the resource consumption of the different contents, we define $\sum_{j \in \mathcal{C}^*} \psi_j^u(k)$ and $\sum_{j \in \mathcal{C}^*} \psi_j^d(k))$ as the penalty coefficients of the usages of the uplink and downlink subframes, respectively. Then, the aggregate resources requirement on serving the jth content is defined as

$$\chi_j(k) = \frac{\psi_j^u(k) \sum_{i \in \mathcal{C}^*} \psi_i^u(k) + \psi_j^d(k) \sum_{i \in \mathcal{C}^*} \psi_i^d(k)}{\left[\left(\sum_{i \in \mathcal{C}^*} \psi_i^u(k) \right)^2 + \left(\sum_{i \in \mathcal{C}^*} \psi_i^d(k) \right)^2 \right]^{\frac{1}{2}}}. \quad (22)$$

$\chi_j(k)$ reflects the additional burden introduced by serving the jth content. For example, assuming $\sum_{j \in \mathcal{C}^*} \psi_j^u(k) = 0.2$ and $\sum_{j \in \mathcal{C}^*} \psi_j^d(k) = 0.6$, then $\chi_j(k) = 0.22$ and $\chi_i(k) = 0.41$. It indicates that serving the jth content introduces less burden to the GCB than serving the ith content. This result is reasonable because the current usage of the downlink subframes is larger than that of the the uplink subframes. In other words, the GCB has less number of available downlink subframes than that of the uplink subframes. The ith content consumes much more downlink subframes than the jth content does, and therefore, serving the ith content introduces more burden on the GCB than serving the jth content.

To further determine whether the jth content is served by the GCB, we calculate the per unit resource profit of the jth content in terms of reducing the MBS's energy consumption as $C_j^s(k)/\chi_j(k)$. The HTO algorithm iteratively adds the content with the largest per unit resource profit into $\mathcal{C}^*$. Denote $C^f = (j | \psi_j^u(k) \leq 1 - \sum_{i \in \mathcal{C}^*} \psi_i^u(k)$, and $\psi_j^d(k) \leq 1 - \sum_{i \in \mathcal{C}^*} \psi_i^d(k)$ $\forall j \in \mathcal{C})$ as the set of feasible content requests which can be served by the GCB. The pseudo code of the HTO algorithm is shown in Algorithm 2.

In the HTO algorithm, opt_flag indicates the termination of the algorithm. When $\mathcal{C}^f(k) = \emptyset$, opt_flag is set to 1. Then, the HTO algorithm terminates. The loop_flag indicates whether an optimal COS is found for a set of requested contents. When $\mathcal{C}^r(k) = \mathcal{C}^f(k)$, the HTO algorithm sets loop_flag $= 0$, and terminates the loop for computing $\beta(k)$ and $\mu_j(k)$ $\forall j \in \mathcal{C}^r$. If $\mathcal{C}^r(k) \neq \mathcal{C}^f(k)$, it indicates that some of the requested contents cannot be served by the GCB because of the limitation of the uplink and downlink subframes. In this case, the HTO algorithm ignores such content requests by setting $\eta_{i,j}(k) = 0 \, \forall j \in \mathcal{C}^r(k) \backslash \mathcal{C}^f(k) \, \forall i \in \mathcal{U}$. Then, the HTO algorithm calls the COS algorithm to recalculate $\beta(k)$ and $\mu_j(k)$ $\forall j \in \mathcal{C}^r$. This loop of the algorithm guarantees the maximization of the summation of the per uplink subframe energy savings for all the serving contents. For instance, as shown in Fig. 5, the energy savings of offloading c_1 and c_2 are 4 and 6, respectively. u_1 registers both contents at the costs of 2 and 4 uplink subframes, respectively. u_2 only registers c_1 at the cost of 3 uplink subframes. If both requests for c_1 and c_3 are feasible content requests, then the optimal COS in the term of maximizing the summation of the per uplink subframe energy savings is that u_1 uploads c_2 while u_2 uploads c_1. However, if the content request for c_2 is not feasible because of the limitation of the uplink subframes, then the content request for c_1 is the only feasible content request. In this case, u_1 rather than u_2 is the optimal content owner for c_1.

Algorithm 2: The HTO Algorithm

Input: $\eta_{i,j}(k), \varphi_{i,j}(k), s_j, n_i^m(k), r_c^m \, \forall j \in \mathcal{C} \, \forall i \in \mathcal{U}$;

Output: $\mathcal{C}^*(k), \beta(k)$, and $\mu_j(k) \, \forall j \in \mathcal{C}$;

1 Set $\mathcal{C}^* = \emptyset$;

2 while opt_flag *equals to* 0 **do**

3   Set loop_flag $= 1$;

4   **while** loop_flag *equals to* 1 **do**

5   Derive $\beta(k)$ and $\mu_j(k)$ using the COS algorithm;

6   Calculate $\psi_j^u(k)$ and $\psi_j^d(k)$;

7   Calculate $\mathcal{C}^r(k)$ and $\mathcal{C}^f(k)$;

8   **if** $\mathcal{C}^r(k) \neq \mathcal{C}^f(k)$ **then**

9   Set $\eta_{i,j}(k) = 0 \, \forall i \in \mathcal{U} \, \forall j \in \mathcal{C}^r(k) \backslash \mathcal{C}^f(k)$;

10   **else**

11   Set loop_flag $= 0$;

12   **if** $\mathcal{C}^f(k) \neq \emptyset$ **then**

13   **if** $\mathcal{C}^*$ *is empty* **then**

14   Calculate $\chi_j(k) = \frac{\psi_j^u(k) + \psi_j^d(k)}{2^{\frac{1}{2}}} \, \forall j \in \mathcal{C}^f(k)$;

15   Calculate $\frac{C_j^s(k)}{\chi_j(k)} \, \forall j \in \mathcal{C}^f(k)$;

16   **else**

17   Calculate $\chi_j(k)$ based on (22) $\forall j \in \mathcal{C}^f(k)$;

18   Calculate $\frac{C_j^s(k)}{\chi_j(k)} \, \forall j \in \mathcal{C}^f(k)$;

19   Find $j = \arg \max_{j \in \mathcal{C}^f(k)} \frac{C_j^s(k)}{\chi_j(k)}$;

20   Set $\mathcal{C}^*(k) = \mathcal{C}^*(k) \cup \{j\}$;

21   Set $\eta_{i,j}(k) = 0 \, \forall i \in \mathcal{U}$;

22   Find i that satisfies $\beta_{i,j}(k) = 1$;

23   Set $\varphi_{i,j}(k) = 0 \, \forall j \in \mathcal{C}$;

24   **else**

25   Set opt_flag $= 1$;

26 Calculate and Return $\mathcal{C}^*(k), \beta(k)$, and $\mu_j(k) \, \forall j \in \mathcal{C}$;

Theorem 2: Ignoring the infeasible content requests and redoing the COS do not decrease the summation of the per uplink subframe energy savings of the feasible contents.

Proof: Assume $\mathcal{C}^r(k) \backslash \mathcal{C}^f(k) = \{h\}$. Denote $\bar{\eta}_{i,j}(k)$ and $\hat{\eta}_{i,j}(k)$ as the content requests for the contents in $\mathcal{C}^r(k)$ and $\mathcal{C}^f(k)$, respectively,

$$\hat{\eta}_{i,j}(k) = \begin{cases} \eta_{i,j}(k), & j \in \mathcal{C}^f \\ 0, & j = h. \end{cases} \quad (23)$$

Denote $\bar{\beta}(k)$ and $\bar{\mu}_j(k) \, \forall j \in \mathcal{C}$ as the optimization results of the COS algorithm with $\bar{\eta}_{i,j}(k)$ as the input, and denote $\hat{\beta}(k)$ and
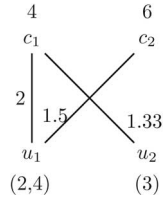
Fig. 5. Illustration of maximization of the summation of the per uplink subframe energy savings.

$\hat{\mu}_j(k) \ \forall j \in \mathcal{C}$ as the optimization results of the COS algorithm with $\hat{\eta}_{i,j}(k)$ as the input. According to (16) and (17), $\bar{\mathcal{C}}_j^s(k)$, $\bar{\delta}_{i,j}(k)$, $\bar{n}_{i,j}^u(k)$, $\hat{\mathcal{C}}_j^s(k)$, $\hat{\delta}_{i,j}(k)$, and $\hat{n}_{i,j}^u(k)$ are derived, respectively. Proving the theorem equals to prove $\sum_{j \in \mathcal{C}^f(k)} \sum_{i \in \mathcal{U}} \hat{\beta}_{i,j}(k)\hat{\delta}_{i,j}(k) \geq \sum_{j \in \mathcal{C}^f(k)} \sum_{i \in \mathcal{U}} \bar{\beta}_{i,j}(k)\bar{\delta}_{i,j}(k)$.

1) If $\bar{\mu}_h(k) = 1$, the GCB retrieves the hth content from the MBS. Ignoring the content request for c_h does not change the COSs of the other contents, and thus $\sum_{j \in \mathcal{C}^f(k)} \sum_{i \in \mathcal{U}} \hat{\beta}_{i,j}(k)\hat{\delta}_{i,j}(k) = \sum_{j \in \mathcal{C}^f(k)} \sum_{i \in \mathcal{U}} \bar{\beta}_{i,j}(k)\bar{\delta}_{i,j}(k)$.

2) If $\bar{\mu}_h(k) = 0$ and $\bar{\beta}_{g,h}(k) = 1$, the GCB receives the content from the gth UE. If $\bar{n}_{g,j}^u(k) < \sum_{i \in \mathcal{U}} \bar{\beta}_{i,j}(k)\bar{n}_{i,j}^u(k)$, $\exists j \in \mathcal{C}^f$, then ignoring the content request for the hth content enables the gth UE to upload the jth content, and thus $\sum_{j \in \mathcal{C}^f(k)} \sum_{i \in \mathcal{U}} \hat{\beta}_{i,j}(k)\hat{\delta}_{i,j}(k) > \sum_{j \in \mathcal{C}^f(k)} \sum_{i \in \mathcal{U}} \bar{\beta}_{i,j}(k)\bar{\delta}_{i,j}(k)$; otherwise, ignoring the content request for c_h does not change the COSs of the other contents, and thus $\sum_{j \in \mathcal{C}^f(k)} \sum_{i \in \mathcal{U}} \hat{\beta}_{i,j}(k)\hat{\delta}_{i,j}(k) = \sum_{j \in \mathcal{C}^f(k)} \sum_{i \in \mathcal{U}} \bar{\beta}_{i,j}(k)\bar{\delta}_{i,j}(k)$.

Therefore, $\sum_{j \in \mathcal{C}^f(k)} \sum_{i \in \mathcal{U}} \hat{\beta}_{i,j}(k)\hat{\delta}_{i,j}(k) \geq \sum_{j \in \mathcal{C}^f(k)} \sum_{i \in \mathcal{U}} \bar{\beta}_{i,j}(k)\bar{\delta}_{i,j}(k)$. The proof is complete. □

C. Computational Complexity

There are two loops in the HTO algorithm. The number of iterations of each loop in the worst case is M. The COS algorithm is called within the inter loop. The major computation complexity of the COS algorithm is from the MWBM algorithm which is called to find the maximum weighted matching between the requested content and the content owner. We adopt the Kuhn–Munkres algorithm [22] to solve the MWBM problem. i.e., finding the MWBM for M contents and N UEs. The computation complexity of the Kuhn–Munkres algorithm for matching M contents and N UEs is $O((N + M)^3)$. Therefore, the overall computation complexity of the HTO algorithm is $O(M^2 (N + M)^3)$.

VI. SIMULATION RESULTS

We set up system level simulations using MATLAB to evaluate the performance of the content brokerage scheme and the HTO algorithm. In the simulation, one MBS and one GCB are deployed, the distance between the MBS and the GCB is 700 m, and 100 UEs are randomly distributed around the GCB. The static power consumption of the MBS and the GCB is 100 W and 20 W, respectively. For simplicity, we set $a^c = 2$ and $a^m = 20$. We adopt the System Advisor Model (SAM) [24] and PVWatts model [25] to estimate the hourly solar energy generation, and set

TABLE I
SIMULATION PARAMETERS

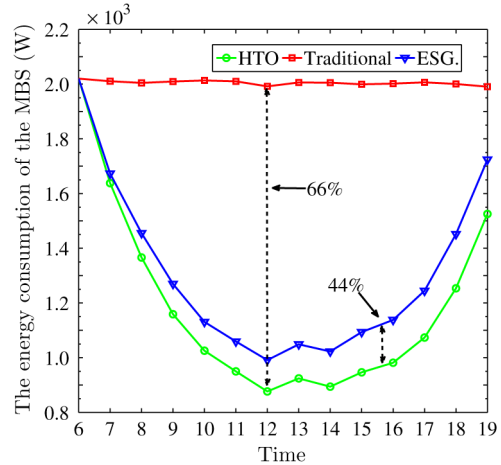| Parameter	Value
MBS's transmit power	20 W
GCB's transmit power	2 W
UEs' transmit power	500 mW
Rayleigh fading	9 dB
Shadowing fading	5 dB
Antenna feeder loss	3 dB
Carrier frequency	2100 MHz
Transmitter gain	1 dB
Noise power level	–104 dBm
Receiver sensitivity	–97 dB

Fig. 6. MBS's energy consumption over time.

$b^{max} = 100$ W. The total number of contents is 100, the content size is randomly distributed between 1 and 10 Mbytes, and the duration of a content slot is 5 s. We adopt COST 231 Walfisch–Ikegami [26] as the propagation model. The communication parameters in the simulation are summarized in Table I.

In the simulations, we compare the MBS's power consumption with and without the content brokerage scheme. We define the content delivery without content brokerage scheme as the traditional scheme. In order to evaluate the content brokerage scheme, we introduce two parameters: 1) the content availability ratio and 2) the content popularity ratio. The content availability ratio is defined as the percentage of the requested contents that are stored in UEs, while the content popularity ratio is defined as the percentage of the UEs who have the requested contents. In addition, to evaluate the HTO algorithm, we compare its performance with an energy saving greedy (ESG) algorithm. The ESG iteratively selects a requested content which leads to the largest energy savings in the MBS, and assigns a content owner to the content. The ESG algorithm terminates when either the uplink or downlink subframes are exhausted.

Figs. 6 and 7 show the power consumption of the MBS and the GCB over time, respectively. In this simulation, we assume the total number of content requests in each content slot is 30. The UEs are randomly selected to submit a content request. If selected, the UE randomly chooses a content, and request it. The total number of uplink and downlink subframes is 50 and 50, respectively. The content availability ratio and popularity ratio are 0.9 and 0.5, respectively. Fig. 6 shows the MBS's power consumption over time. In Fig. 6, each marker represents the MBS's average
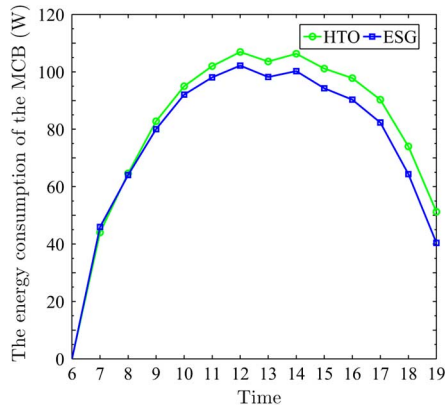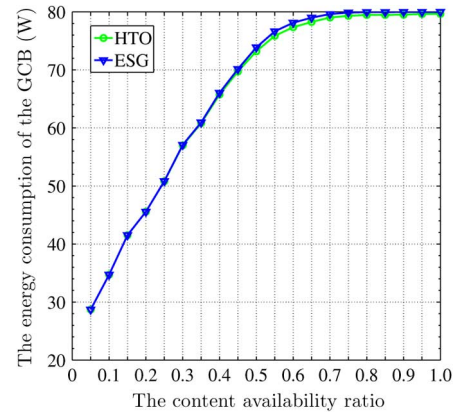
Fig. 7. GCB's energy consumption over time.

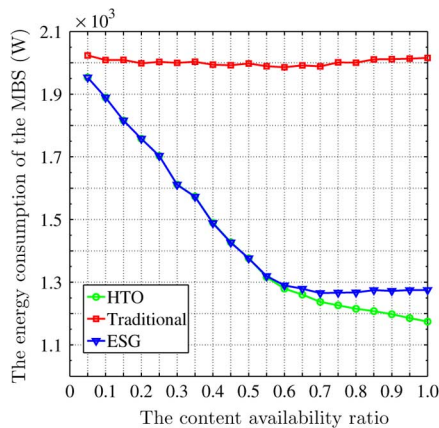Fig. 9. GCB's energy consumption versus different content availability ratio.

Fig. 8. MBS's energy consumption versus different content availability ratio.
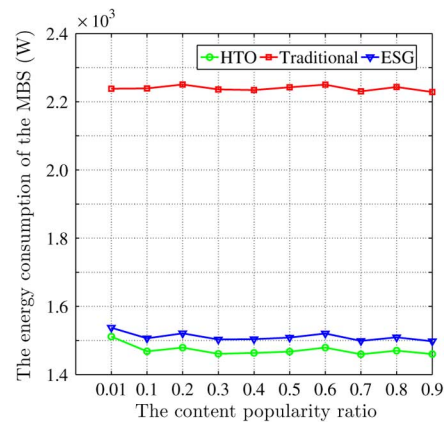

Fig. 10. MBS's energy consumption versus content popularity ratio.

power consumption of all the content slots within an hour. As shown in Fig. 6, the content brokerage scheme achieves upto 66% energy savings in the MBS. The content brokerage scheme achieves the highest energy savings during noon. This is because the GCB is powered by solar energy which peaks during noon. With a larger amount of green energy, the GCB is able to utilize more subframes, and thus can offload more traffic from the MBS. As compared with the ESG algorithm, the HTO algorithm achieves upto 14% energy savings in the MBS. This is because the HTO algorithm optimizes both the COS and the SCS. As a result, with a given amount of resources (the uplink and downlink subframes), the HTO algorithm enables the GCB to offload more contents. Fig. 7 compares the GCB's power consumption under the ESG algorithm and the HTO algorithm. With the HTO algorithm, the GCB consumes more green energy, which indicates that the HTO algorithm enables the GCB to offload more traffic from the MBS.

Figs. 8 and 9 show the performance of the content brokerage scheme at different content availability ratios. In this simulation, we assume the total number of content requests in each content slot is 30. The total numbers of uplink and downlink subframes are 40 and 40, respectively. The content popularity ratio is 0.5. To evaluate the impact of the content availability on the content brokerage scheme, we assume the GCB has sufficient green energy to utilize all the subframes. Thus, the constraints of the GCB on offloading traffic are from the total numbers of subframes. Fig. 8 shows the MBS's energy consumption with different

content availability ratios. As shown in Fig. 8, when the content availability ratio is small, the content brokerage scheme does not achieve much energy savings. However, as the content availability ratio increases, using the content brokerage scheme, the MBS's energy consumption reduces significantly. When the content availability ratio is larger than 0.5, the ESG algorithm is unable to further reduce MBS's energy consumption because of the constraints of the total subframes. However, by optimizing the COS and the SCS, the HTO algorithm is able to maximize the utilization of the uplink and downlink subframes, and thus the HTO algorithm achieves more energy savings than the ESG algorithm does. Fig. 9 shows the GCB's energy consumption with the HTO algorithm and the ESG algorithm. The GCB's energy consumption under both algorithms is almost the same, which indicates the number of subframes utilized by both algorithms is almost the same. Since the HTO algorithm achieves more energy savings, it indicates the HTO algorithm achieves higher resource utilization than the ESG algorithm does.

Fig. 10 shows the MBS's energy consumption with different content popularity ratios. In this simulation, we assume the total number of content requests in each content slot is 30. The total numbers of uplink and downlink subframes are 30 and 30, respectively. The content availability ratio is 0.6. As shown in Fig. 10, the content brokerage scheme significantly reduces the MBS's energy consumption. When the content popularity ratio is 0.01, the performance of the content brokerage scheme is limited

by the number of content owners. Since the total number of UEs is 100, if the content popularity ratio equals to 0.01, it indicates that a requested content is available in at most one UE. The UE may also hold the other requested contents, in this case the content popularity to some extent impairs the content availability because a UE can upload at most one content to the GCB during a content slot. As the content popularity increases, such impairment decreases, and thus the performance of the brokerage scheme is almost stable.

VII. CONCLUSION

In this paper, we have proposed a novel content brokerage scheme which enables traffic offloading from the remote MBS to local UEs, and have designed the HTO algorithm to maximize the traffic offloading with the constraints of resources such as available green energy and the number of uplink and downlink subframes. As demonstrated via extensive simulations, the proposed content brokerage scheme, together with the HTO algorithm, significantly reduces the MBS's energy consumption.

REFERENCES

[1] T. Han, N. Ansari, M. Wu, and H. Yu, "On accelerating content delivery in mobile networks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1314–1333, Jul. 2013.

[2] T. Han and N. Ansari, "On greening cellular networks via multicell cooperation," *IEEE Wireless Commun. Mag.*, vol. 20, no. 1, pp. 82–89, Feb. 2013.

[3] P. Deshpande, X. Hou, and S. R. Das, "Performance comparison of 3G and metro-scale WiFi for vehicular network access," in *Proc. 10th Internet Meas. Conf. (IMC)*, Melbourne, Australia, 2010, pp. 301–307.

[4] L. Lei, Z. Zhong, C. Lin, and X. Shen, "Operator controlled device-to-device communications in LTE-advanced networks," *IEEE Wireless Commun.*, vol. 19, no. 3, pp. 96–104, Jun. 2012.

[5] H. Kim, G. de Veciana, X. Yang, and M. Venkatachalam, "Distributed α-optimal user association and cell load balancing in wireless networks," *IEEE/ACM Trans. Netw.*, vol. 20, no. 1, pp. 177–190, Feb. 2012.

[6] H.-S. Jo, Y. J. Sang, P. Xia, and J. Andrews, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012.

[7] K. Lee, I. Rhee, J. Lee, Y. Yi, and S. Chong, "Mobile data offloading: How much can WiFi deliver?" *SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 4, pp. 425–426, Aug. 2010.

[8] A. Balasubramanian, R. Mahajan, and A. Venkataramani, "Augmenting mobile 3G using WiFi," in *Proc. 8th Int. Conf. Mobile Syst. Appl. Serv. (MobiSys'10)*, San Francisco, CA, USA, Jun. 2010, pp. 209–222.

[9] T. Han and N. Ansari, "Opportunistic content pushing via wifi hotspots," in *Proc. 3rd IEEE Int. Conf. Netw. Infrastruct. Digit. Content (IC-NIDC)*, Beijing, China, Sep. 2012, pp. 680–684.

[10] B. Han, P. Hui, V. A. Kumar, M. V. Marathe, G. Pei, and A. Srinivasan, "Cellular traffic offloading through opportunistic communications: A case study," in *Proc. 5th ACM Workshop Challenged Netw. (CHANTS'10)*, Sep. 2010, pp. 31–38.

[11] A. Mashhadi and P. Hui. (2010). *Proactive Caching for Hybrid Urban Mobile Networks* [Online]. Available: http://www.cs.ucl.ac.uk/research/researchnotes/documents/RN_10_05_000.pdf

[12] X. Zhuo, W. Gao, G. Cao, and Y. Dai, "Win-coupon: An incentive framework for 3G traffic offloading," in *Proc. 19th IEEE Int. Conf. Netw. Protocols (ICNP)*, Vancouver, BC, Canada, Oct. 2011, pp. 206–215.

[13] T. Han and N. Ansari, "Powering mobile networks with green energy," *IEEE Wireless Commun.*, vol. 21, no. 1, pp. 90–96, Feb. 2014.

[14] Ericson Inc., "Sustainable energy use in mobile communications," White Paper, Aug. 2007.

[15] J. Zhou, M. Li, L. Liu, X. She, and L. Chen, "Energy source aware target cell selection and coverage optimization for power saving in cellular networks," in *Proc. IEEE/ACM Int. Conf. Green Comput. Commun. Int. Conf. Cyber Phys. Social Comput.*, Hangzhou, China, Dec. 2010, pp. 1–8.

[16] T. Han and N. Ansari, "ICE: Intelligent cell breathing to optimize the utilization of green energy," *IEEE Commun. Lett.*, vol. 16, no. 6, pp. 866–869, Jun. 2012.

[17] T. Han and N. Ansari, "Optimizing cell size for energy saving in cellular networks with hybrid energy supplies," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2012, pp. 3–7.

[18] T. Han and N. Ansari, "On optimizing green energy utilization for cellular networks with hybrid energy supplies," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 3872–3882, Aug. 2013.

[19] S. Sharma, Y. Shi, Y. Hou, and S. Kompella, "An optimal algorithm for relay node assignment in cooperative ad hoc networks," *IEEE/ACM Trans. Netw.*, vol. 19, no. 3, pp. 879–892, Jun. 2011.

[20] O. Arnold, F. Richter, G. Fettweis, and O. Blume, "Power consumption modeling of different base station types in heterogeneous cellular networks," in *Proc. IEEE Future Netw. Mobile Summit*, Jun. 2010, pp. 1–8.

[21] A. Ambrosy, O. Blume, H. Klessig, and W. Wajda, "Energy saving potential of integrated hardware and resource management solutions for wireless base stations," in *Proc. IEEE 22nd Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, 2011, pp. 2418–2423.

[22] R. Burkard, M. Dell'Amico, and S. Martello, *Assignment Problems*. Philadelphia, PA, USA: Soc. Ind. Appl. Math., 2009.

[23] Y. Toyoda, "A simplified algorithm for obtaining approximate solutions to zero-one programming problems," *Manage. Sci.*, vol. 21, no. 12, pp. 1417–1427, Aug. 1975.

[24] Nat. Renewable Energy Lab. *System advisor model (SAM)* [Online]. Available: https://sam.nrel.gov/

[25] Nat. Renewable Energy Lab. *PVWatts* [Online]. Available: http://www.nrel.gov/rredc/pvwatts/

[26] *EUR 18957—COST Action 231 Digital Mobile Radio Towards Future Generation Systems—Final Report*, Luxembourg, Luxembourg, Office Off. Pub. Eur. Commun., 1999.

Tao Han (S'08) received the B.E. degree in electrical engineering from the Dalian University of Technology, Dalian, China, in 2006, the M.E. degree in computer engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 2009, and is currently working toward the Ph.D. degree in electrical and computer engineering at the New Jersey Institute of Technology, Newark, NJ, USA.

His research interests include mobile and cellular networks, network optimization, energy efficient networking, and Internet of Things.

Nirwan Ansari (S'78–M'83–SM'94–F09) received the B.S.E.E. degree (*summa cum laude*) from the New Jersey Institute of Technology (NJIT), Newark, NJ, USA, in 1982, the M.S.E.E. degree from the University of Michigan, Ann Arbor, MI, USA, in 1983, and the Ph.D. degree from Purdue University, West Lafayette, IN, USA, in 1988.

He joined NJIT in 1988, where he is a Professor of electrical and computer engineering. He has also assumed various administrative positions at NJIT. He was a Visiting (Chair/Honorary) Professor of several universities. He authored *Computational Intelligence for Optimization* (Springer, 1997) with E. S. H. Hou, *Media Access Control and Resource Allocation for Next Generation Passive Optical Networks* (Springer, 2013) with J. Zhang, and edited *Neural Networks in Telecommunications* (Springer, 1994) with B. Yuhas. He has also coauthored over 450 publications, over one-third of which were published in widely cited refereed journals/magazines. He has been granted over 20 U.S. patents. He has also guest-edited a number of special issues, covering various emerging topics in communications and networking. He has been frequently selected to deliver keynote addresses, distinguished lectures, and tutorials. He has served on the Editorial/Advisory Board of nine journals. His research interests include various aspects of broadband networks and multimedia communications.

Dr. Ansari was elected to serve in the IEEE Communications Society (ComSoc) Board of Governors as a Member-at-Large (2013–2014) as well as IEEE Region 1 Board of Governors as the IEEE North Jersey Section Chair. He has chaired ComSoc technical committees, and has been actively organizing numerous IEEE international conferences/symposia/workshops, assuming leadership roles as Chair or Technical Program Committee (TPC) Chair at various conferences, symposia, and workshops. Some of his recent recognitions include a few Best Paper Awards, several Excellence in Teaching Awards, Thomas Alva Edison Patent Award in 2010, NJ Inventors Hall of Fame Inventor of the Year Award in 2012, the Newark College of Engineering (NCE) Excellence in Research Award in 2014, and designation as a ComSoc Distinguished Lecturer from 2006 to 2009.