Enabling Mobile Traffic Offloading via Energy Spectrum Trading

---

Citation:
Tao Han and Nirwan Ansari, "Enabling Mobile Traffic Offloading via Energy Spectrum Trading," *IEEE Transactions on Wireless Communications*, vol.13, no.6, pp.3317-3328, June 2014.

URL:
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6805337

Enabling Mobile Traffic Offloading via Energy Spectrum Trading

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

***Abstract*—Green communications has received much attention recently. For mobile networks, the base stations (BSs) account for more than 50% of the energy consumption of the networks. Therefore, reducing the power consumption of BSs is crucial to greening mobile networks. In this paper, we propose a novel energy spectrum trading (EST) scheme which enables the macro BSs to offload their mobile traffic to Internet service providers' (ISPs') wireless access points by leveraging cognitive radio techniques. Since the ISP's wireless access points are usually closer to the mobile users, the energy and spectral efficiency of mobile networks are enhanced. However, in the EST scheme, achieving optimal mobile traffic offloading in terms of minimizing the energy consumption of the macro BSs is NP-hard. We thus propose a heuristic algorithm to approximate the optimal solution with low computation complexity. We have proved that the energy savings achieved by the proposed heuristic algorithm is at least 50% of that achieved by the brute-force search. Simulation results demonstrate the performance and viability of the proposed EST scheme and the heuristic algorithm.**

***Index Terms*—Mobile traffic offloading, energy efficient wireless networks, green communications, cellular networks, WiFi hotspots.**

I. Introduction

OWING to the direct impact of greenhouse gases on the earth environment and the climate change, the energy consumption of Information and Communications Technology (ICT) is becoming an environmental and thus social and economic issue. Mobile networks are among the major energy hoggers of communication networks. 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 of data traffic in mobile networks. The mobile data traffic surges result in a dramatic increase of energy consumption of mobile networks for provisioning higher network capacity [1]. Therefore, greening cellular networks is crucial to reducing the carbon footprints of ICT [2].

Mobile traffic offloading, which is referred to as utilizing complementary communication networks to deliver mobile traffic, is a promising technique to improve the energy and spectral efficiency of mobile networks [3]. In order to offload mobile traffic, mobile network operators usually deploy small cell base stations (BSs), e.g., pico-BSs, femto-BSs and WiFi hot spots, in the area where the mobile traffic intensity is high. Such mobile network deployments, referred as to heterogeneous mobile networks, can efficiently offload mobile traffic from macro BSs, thus reducing the energy consumption of mobile networks [4]. However, deploying small cell BSs requires backhaul networks which connect the small cell BSs and the mobile core networks. The energy consumption of the backhaul networks may neutralize the increased energy efficiency. Thus, 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.

With strong revenue growth in wireless data markets, internet service providers (ISPs) such as Comcast and Optimum are densely deploying WiFi hot spots to provide WiFi connectivity to their customers in urban and suburban areas [5]. Therefore, it is desirable to utilize the hotspots deployed by ISPs to offload mobile data traffic. However, since carrying mobile traffic introduces additional operation cost to ISPs' networks, without proper incentives, the ISPs are not willing to open their networks to mobile network subscribers.

In this paper, we propose a novel mobile traffic offloading scheme by leveraging cognitive radio techniques referred to as energy spectrum trading (EST). The EST scheme exploits the merits of both mobile networks and ISPs' networks. One of the advantages of the mobile networks is that the networks are operating on licensed spectrum, which are not accessed by unlicensed users. Therefore, by proper spectrum management, mobile networks are able to provide their subscribers a variety of services with different QoS levels. However, as compared with the hotspots deployed by ISPs, the BSs of mobile networks are usually sparsely deployed. Such deployments are not efficient in terms of the energy and spectral utilization. One of the merits of ISPs' hotspots is that they are densely deployed, and are able to provide high speed data rates to their subscribers. However, operating on unlicensed spectrum, the (QoS) of data services may not be guaranteed. The EST scheme enables mobile networks to offload data traffic to ISPs' networks to improve energy and spectral efficiency, and allows ISPs' hotspots access to the licensed spectrum to provide ISPs' data services with different QoS levels.

The proposed scheme is illustrated in Fig. 1, where the primary BS (PBS) is defined as the macro BS owned by the mobile network operator while the secondary BSs (SBSs) are referred to as the hotspots owned by ISPs. We assume both the PBS and SBSs are able to dynamically access the spectrum by

Manuscript received May 25, 2013; revised September 17, 2013, November 24, 2013, and January 28, 2014. The associate editor coordinating the review of this paper and approving it for publication was S. Bahk.

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

This work was supported in part by the NSF under grant no. CNS-1218181 and no. CNS-1320468.

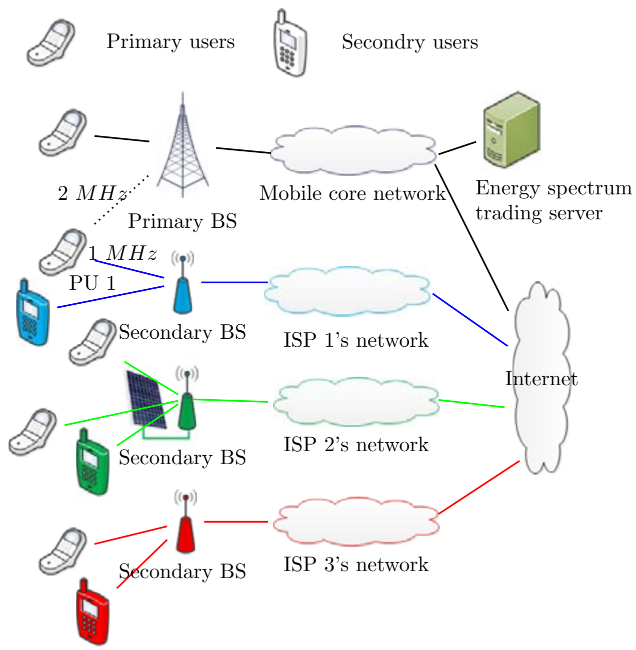


Fig. 1. Illustration of the energy spectrum trading scheme.

leveraging cognitive radio techniques. There are two types of users: primary users (PUs) and secondary users (SUs). PUs are subscribers of the mobile networks while SUs are subscribers of ISPs. Different SUs may subscribe to different ISPs. The energy spectrum trading server manages the spectrum sharing and mobile data offloading between the mobile networks and ISPs' networks.

The PBS has the exclusive access to the licensed band. However, owing to the wireless channel fading between the PBS and PUs, providing high data rates to the PUs, especially to those located at the cell edge, is both bandwidth and power consuming. As compared with the PBS, the SBSs which are closer to the PUs may experience less wireless channel fading and have higher spectral and energy efficiency in providing data services to the PUs. In the EST scheme, the PBS shares a certain amount of licensed bandwidth with SBSs while SBSs provide data services to PUs within their coverage area using the allocated bandwidth. Since SBSs are close to PUs, the SBSs can satisfy PUs' QoS requirements by utilizing only a portion of the allocated bandwidth. The residual bandwidth can be utilized to fulfill SUs' data rate requirements. For example, in Fig. 1, if PU 1 is associated with the PBS, the PBS should allocate 2 MHz bandwidth to the PU to satisfy its minimum data rate requirement. If associated with the SBS, PU 1 may only require 1 MHz to ensure its minimum data rate. If the PBS offloads PU 1 to the SBS and grants the SBS 2 MHz bandwidth, then the SBS spends 1 MHz bandwidth to serve PU 1, and the other $1MHz$ bandwidth can be utilized to enhance QoS of its SUs. Therefore, the EST scheme enables the PBS to reduce its power consumption by offloading some of the PUs to SBSs, and allows the SBSs to enhance their QoSs to SUs by utilizing the licensed bandwidth. Since SBSs usually have a low transmit power, the power consumption and the spectrum usages of mobile networks in providing data services to PUs is reduced. Thus, the EST scheme enhances both the energy efficiency and the spectral efficiency of mobile networks.

The EST between the PBS and SBSs can be either event driven or traffic driven. For the event driven EST, the PBS triggers an EST process when a cell edge user initiates data service requests. For traffic driven EST, the PBS monitors its traffic intensity from cell edge users. When the traffic intensity is beyond a threshold, an EST process is triggered. In this paper, assuming the PBS experiences heavy traffic load from the cell edge users, we design algorithms to optimize the EST between the PBS and SBSs to minimize the PBS's energy consumption when an EST process is triggered. However, minimizing the energy consumption of the PBS in the EST scheme is not trivial. On the one hand, in order to minimize the power consumption, the PBS has to maximize the number of users offloading to SBSs. Meanwhile, since the total amount of licensed spectrum is limited, the PBS aims to minimize the amount of bandwidth allocated to SBSs because the less bandwidth allocated to SBSs, the more bandwidth is reserved for the PUs associated with the PBS, and therefore the PBS consumes less power. On the other hand, the PBS has to give the SBS sufficient incentives in term of the amount of licensed spectrum to incentivize SBSs to provide data services to the PUs. Therefore, solving the power consumption minimization (PCM) problem is to find user-BS associations and bandwidth allocations to minimize the power consumption of the PBSs while satisfying PUs' minimum data rates and SBS's bandwidth requirements. In fact, the PCM problem is an NP-hard problem. Therefore, we propose a heuristic algorithm to approximate the optimal solution achieved by the brute-force search. The heuristic algorithm first finds the PUs whose user-BS associations are not determined, and then iteratively associates the PU, whose power-bandwidth ratio is the largest, with SBSs.

If the power consumption of the PBS is reduced, the PU is associated with SBSs; otherwise, the PU is associated with the PBS. As compared with the brute-force search, the heuristic algorithm achieves at least 50% power consumption savings when the PBS experiences heavy traffic load from cell edge users.

The rest of the paper is organized as follows. In Section II, we provide an overview on related research efforts. In Section III, we define the system model. Section IV formulates and analyzes the power consumption minimization problem. Section V presents a heuristic algorithm to approximate the optimal solution and analyzes the performance of the algorithm. Section VI shows the simulation results, and concluding remarks are presented in Section VII.

II. Related Works

In this section, we briefly overview the related research on mobile traffic offloading and the solutions for user-BS associations in heterogeneous mobile networks.

A. Mobile Traffic Offloading

Based on the network access mode, the mobile traffic offloading schemes can be classified into infrastructure based traffic offloading and ad-hoc based traffic offloading. The infrastructure based traffic offloading is most related to our work. Therefore, we provide a brief overview of the infrastructure based traffic offloading. In the infrastructure based mobile traffic offloading, mobile traffic can be offloaded to small cell base stations (BSs), e.g., pico-BSs, femto-BSs and WiFi hot spots [6]. Small cell BSs usually consume much less power than the macro BSs. Therefore, offloading mobile traffic to small cell BSs can significantly enhance the energy efficiency of mobile networks [4]. On the other hand, in order to reduce CO_2 footprints, mobile traffic can be offloaded to the BSs powered by green energy such as sustainable biofuels, solar and wind energy [7]–[10]. In this way, the green energy utilization is maximized, and thus the consumption of the on-grid energy is minimized.

B. User-BS Associations in Heterogeneous Mobile Networks

Heterogeneous network is a promising network architecture which may significantly enhance the spectral and energy efficiency of mobile networks. One of the most important issues in heterogeneous cellular networks is to properly associate mobile users with the serving base stations (BSs), referred to as the user-BS association problem. In heterogeneous cellular networks, the transmit power of small cell BSs are significantly lower than that of macro BSs. Thus, mobile users are more likely associated with the macro BS based on the strength of their received pilot signal. As a result, small cell BSs may be lightly loaded, and do not contribute much on traffic offloading. To address this issue, many user-BS association algorithms have been proposed [11]–[13]. Kim *et al.* [11] proposed a framework for the user-BS association in cellular networks to achieve flow level load balancing under spatially heterogeneous traffic distribution. Jo *et al.* [12] proposed cell biasing algorithms to balance traffic loads among macro BSs and small cell BSs. The cell biasing algorithms perform user-BS association according to the biased measured pilot signal strength, and enable the traffic to be offloaded from macro BSs to small cell BSs. Corroy *et al.* [13] proposed a dynamic user-BS association algorithm to maximize the sum rate of the network and adopted cell biasing to balance the traffic load among BSs. Fooladivanda *et al.* [14] studied the joint resource allocation and user-BS association in heterogeneous mobile networks. They investigated the problem under different channel allocation strategies, and the proposed solution achieved global proportional fairness among the users. Madan *et al.* [15] studied the user-BS association and interference coordination in heterogeneous mobile networks, and proposed heuristic algorithms to maximize the sum utility of average rates.

The existing mobile traffic offloading scheme does not consider the traffic offloading among different service providers. In addition, the available user-BS association algorithms in heterogeneous networks usually assume that the macro BS and small cell BSs belong to the same service provider. Therefore, the existing traffic offloading scheme and user-BS association algorithms are not enabling the traffic offloading between the mobile network operators and ISPs.

III. System Model

Consider an area consisting of one PBS and several SBSs from various ISPs as shown in Fig. 1. The PUs are randomly distributed in the area. Denote $\mathcal{U}$ and $\mathcal{S}$ as the set of PUs and SBSs, respectively. The PBS provides data service to the PUs within its coverage area via licensed spectrum. SBSs are randomly deployed in the area. We assume that SBSs are able to dynamically access the licensed spectrum by utilizing cognitive radio techniques.

A. Communications Model

In the EST scheme, the PBS aims to offload data traffic to SBSs to reduce its energy consumption, and is willing to grant a portion of the licensed spectrum to incentivize SBSs to allow PUs to access their networks. Meanwhile, SBSs aim to dynamically utilize the licensed spectrum to enhance QoS of data services to their subscribers. Thus, SBSs are willing to allow PUs to access their networks in exchange for the access of the licensed spectrum.

We assume the total amount of licensed spectrum is W which can be split into orthogonal channels, e.g., OFDMA, with variable amount of bandwidth to avoid interference. Each channel is allocated to an individual PU as needed. For simplicity, we assume both PUs and SUs experience frequency flat fading. Therefore, we focus on the amount of bandwidth allocated to PUs and SBSs instead of specifying which part of the spectrum to be allocated. Users' locations are assumed to be static during an EST procedure. We assume the channel fading changes slowly and can be considered as a constant within the duration. Therefore, the wireless channel is modeled as a slow-fading channel which reflects the large-scale fading between BSs and users.

At the beginning of an EST procedure, the kth SBS calculates its bandwidth requirements, denoted as $\phi_{k,i}$, for serving the ith PU. The calculation of $\phi_{k,i}$ consists of two steps. First, the kth SBS calculates the required bandwidth, $\phi^P_{k,i}$, to satisfy the ith PU's minimum data rate, r_i^{min}. Assuming the kth SBS's transmit power-spectral density is p^s, and the channel fading between the kth SBS and the ith PU is $h^s_{k,i}$, $\phi^P_{k,i}$ can be derived by solving

$$r_i^{min} = \phi^P_{k,i} \log{(1 + \frac{p^s |h^s_{k,i}|^2}{\mathcal{N}_0})}. \tag{1}$$

Second, the kth SBS calculates the required bandwidth, $\phi^S_{k,i}$, to compensate for its cost in serving the ith PU. The kth SBS's cost includes the SBS's energy consumption and backhaul usages for serving the ith PU. The cost may be different for different ISPs. For example, in Fig. 1, the second ISP utilizes green energy powered access point, which may reduce the energy cost. Thus, as compared with other ISPs, the second ISP may incur a smaller cost in serving one PU. However, how to calculate $\phi^S_{k,i}$ is beyond the scope of the paper. We assume $\phi^S_{k,i}$ is a constant. Then,

$$\phi_{k,i} = \phi^P_{k,i} + \phi^S_{k,i}. \tag{2}$$

The energy spectrum trading server collects $\phi_{k,i}, \forall k \in \mathcal{S}$, $\forall i \in \mathcal{U}$, and optimizes the user-BS associations and bandwidth allocations to minimize the energy consumption of the PBS.

B. Energy Consumption Model

The PBS's power consumption consists of two parts: the static power consumption and the dynamic power consumption [16]. The static power consumption is the power consumption of a BS without any traffic load. The dynamic power consumption refers to the additional power consumption caused by traffic load on the BS. We consider the PBS's static power consumption, p^{fix}, as a constant, and focus on reducing the dynamic power consumption of a PBS by offloading its traffic to SBSs. The dynamic power consumption of a macro BS depends on the traffic load on the BS and can be expressed as a linear function of the BS's transmit power [17]. Therefore, we model the PBS's power consumption as

$$C = \sum_{i \in \mathcal{U}} \alpha \mu_i p_i w_i + p^{fix} \tag{3}$$

Here, α is a coefficient which reflects the relationship between the PBS's dynamic energy consumption and the summation of the PBS's transmit power toward its associated PUs. The value of α depends on the characteristic of the BS [17]. μ_i is an indicator function. If PU is associated with the PBS, $\mu_i = 1$; otherwise, $\mu_i = 0$. w_i is the amount of bandwidth allocated to the ith PU, and p_i is the transmit power-spectral density in w_i.

IV. Problem Formulation and Analysis

In the EST scheme, the PBS aims to minimize its power consumption by offloading data traffic to SBSs. Therefore, the power consumption minimization (PCM) problem can be formulated as follows:

$$\begin{aligned} \min_{(\mu_i, \beta_{k,i}, w_i, p_i)} \quad & \sum_{i \in \mathcal{U}} \alpha \mu_i p_i w_i + p^{fix} & (4) \\ subject\ to: \quad & \sum_{i \in \mathcal{U}} (\mu_i w_i + \sum_{k \in \mathcal{S}} \beta_{k,i} \phi_{k,i}) = W, \\ & r_i \geq r_i^{min}, \ \forall i \in \mathcal{U}, \\ & \mu_i p_i \leq p^{max}, \ \forall i \in \mathcal{U} \\ & \mu_i + \sum_{k \in \mathcal{S}} \beta_{k,i} = 1, \ \forall i \in \mathcal{U}. & (5) \end{aligned}$$

Here, $\beta_{k,i}$ is an indicator function. If the ith PU is associated with the kth SBS, $\beta_{k,i} = 1$; otherwise, $\beta_{k,i} = 0$. p^{max} is the PBS's maximum transmit power-spectral density. If a PU is offloaded to a SBS, the SBS should satisfy the PU's minimum data rates. Thus, $r_i = r_i^{min}$ when $\mu_i = 0$. Therefore,

$$r_i = \begin{cases} w_i \log(1 + \frac{p_i |h_i^P|^2}{\mathcal{N}_0}), & \mu_i = 1; \\ r_i^{min}, & \mu_i = 0. \end{cases} \tag{6}$$

Here, h_i^P is the channel fading between the PBS and the ith PU. The PCM problem consists of four constraints. The first constraint is that the sum of the allocated bandwidth should not be larger than the total amount of bandwidth. The second constraint is that the PU's minimum data rate should be satisfied. The third constraint is that the PBS's transmit power should not be larger than its maximum transmit power. The fourth constraint is that a PU can only access either the PBS or one of the SBSs.

When $\mu_i = 1$,

$$p_i = \frac{\mathcal{N}_0 (2^{\frac{r_i}{w_i}} - 1)}{|h_i^P|^2}. \tag{7}$$

Given the amount of bandwidth, w_i, the derivative of p_i with respect to r_i can be expressed as

$$\frac{\partial p_i}{\partial r_i} = \frac{\mathcal{N}_0 2^{r_i/w_i} \ln 2}{w_i |h_i^P|^2} > 0. \tag{8}$$

Since $\frac{\partial p_i}{\partial r_i} > 0$, given the amount of bandwidth, the power consumption increases as the data rate increases. Therefore, to minimize the PBS's energy consumption, PUs are served at the minimum data rate. Thus, $r_i = r_i^{min}$ in the PCM problem. If the ith PU is associated with the PBS, the minimum required bandwidth, w_i^{min}, is derived by solving

$$r_i^{min} = w_i^{min} \log(1 + \frac{p^{max} |h_i^P|^2}{\mathcal{N}_0}). \tag{9}$$

Therefore, we replace the second and third constraints with the minimum bandwidth constraint. In addition, since a PU can be associated with at most one SBS, we select the SBS with the smallest $\phi_{k,i}, \ \forall i \in \mathcal{U}$. Here, $k = \arg\min_{j \in \mathcal{S}} \phi_{j,i}$. Thus, the PCM problem can be rewritten as

$$\begin{aligned} \min_{(\mu_i, \beta_{k,i}, w_i)} \quad & \sum_{i \in \mathcal{U}} \alpha \mu_i p_i w_i + p^{fix} & (10) \\ subject\ to: \quad & \sum_{i \in \mathcal{U}} (\mu_i w_i + \beta_{k,i} \phi_{k,i}) = W, \\ & \mu_i w_i \geq \mu_i w_i^{min}, \ \forall i \in \mathcal{U} \\ & \mu_i + \beta_{k,i} = 1, \ \forall i \in \mathcal{U}. & (11) \end{aligned}$$

Lemma 1. *The optimal solution to the problem in Eq. 10 is the optimal solution to the PCM problem in Eq. 4.*

Proof: The problem in Eq. 10 has the same objective function as that in Eq. 4. For the constraints, the data rate constraint and the transmit power constraint in Eq. 4 can be translated to the minimum bandwidth constraint in Eq. 10. Thus, proving Lemma 1 is equivalent to prove that $k = \arg\min_{j \in \mathcal{S}} \phi_{j,i}$ is a necessary condition of the optimal solution to the problem in Eq. 4. This can be proved by contradiction. Assume the optimal solution to the problem in Eq. 4 offloads the ith PU to the k^*th SBS and $k^* \neq \arg\min_{j \in \mathcal{S}} \phi_{j,i}$. Let $k = \arg\min_{j \in \mathcal{S}} \phi_{j,i}$. In this case, since the kth SBS requires less bandwidth from the PBS, offloading the ith PU to the kth SBS increases the available bandwidth in the PBS. According to the Shannon−Hartley theorem, to achieve a given data rate, increasing the bandwidth reduces the requirement of the transmitting power. Thus, offloading the ith PU to the kth SBS can further reduce the power consumption of the PBS. Therefore, offloading the ith PU to the k^*th SBS is not the optimal solution to the problem in Eq. 4. ■

Theorem 1. *The PCM problem is an NP-hard problem.*

Proof: We prove the theorem by transforming a simplified PCM (SPCM) problem into a knapsack problem which is an NP-hard problem [18]. We simplify the PCM problem by setting $p_i = p^{max}$, $\forall i \in \mathcal{U}$. Then, the SPCM problem can be expressed as

$$\min_{(\mu_i, \beta_{k,i})} \quad \sum_{i \in \mathcal{U}} \alpha \mu_i p^{max} w_i^{min} + p^{fix} \tag{12}$$

$$\begin{aligned} subject\ to: \quad & \sum_{i \in \mathcal{U}} (\mu_i w_i^{min} + \beta_{k,i} \phi_{k,i}) = W, \\ & \mu_i + \beta_{k,i} = 1,\ \forall i \in \mathcal{U}. \end{aligned} \tag{13}$$

Denote $\Delta W = W - \sum_{i \in \mathcal{U}} w_i^{min}$ as the maximum amount of bandwidth that can be utilized by the PBS as the incentives to SBSs for traffic offloading. Define $\Delta \phi_{k,i} = \max\{\phi_{k,i} - w_i^{min}, 0\}$ as the required incentives for the kth SBS to offloading the ith PU. The PBS's power savings by offloading the ith PU equals to $p^{max} w_i^{min}$. The SPCM can be transformed to

$$\max_{\beta_{k,i}} \quad \sum_{i \in \mathcal{U}} \beta_{k,i} p^{max} w_i^{min} \tag{14}$$

$$subject\ to: \quad \sum_{i \in \mathcal{U}} \beta_{k,i} \Delta \phi_{k,i} \leq \Delta W. \tag{15}$$

The above formulation is actually a knapsack problem. Therefore, the SPCM problem can be transformed into a knapsack problem which is an NP-hard problem. Thus, the PCM problem is an NP-hard problem. ∎

V. A Heuristic Power Consumption Minimization Algorithm

In this section, we propose a heuristic power consumption minimization (HPCM) algorithm to approximate the optimal solution of the PCM problem with low computational complexity, and prove that the maximum power savings achieved by the HPCM algorithm is at least 50% of that achieved by the brute force search.

A. The HPCM Algorithm

For the PCM problem, if user-BS associations are determined, then μ_i and $\beta_{k,i}$ are known. The amount of available bandwidth in the PBS can be derived as

$$W^P = W - \sum_{i \in \mathcal{U}} \sum_{k \in \mathcal{S}} \beta_{k,i} \phi_{k,i}. \tag{16}$$

Define $\mathcal{U}^P = \{i | \mu_i = 1, \forall i \in \mathcal{U}\}$ as the set of PUs associated with the PBS. Then, the PCM problem becomes a bandwidth allocation (BA) problem as follows:

$$\min_{w_i} \quad \sum_{i \in \mathcal{U}^P} \alpha p_i w_i + p^{fix} \tag{17}$$

$$\begin{aligned} subject\ to: \quad & \sum_{i \in \mathcal{U}^P} w_i = W^P \\ & w_i \geq w_i^{min}, \forall i \in \mathcal{U}^P. \end{aligned} \tag{18}$$

Let $f(\boldsymbol{w}) = \sum_{i \in \mathcal{U}^P} \alpha p_i w_i + p^{fix}$ and $\boldsymbol{w} = (w_1, w_2, \cdots, w_{|\mathcal{U}^P|})$. When $\boldsymbol{w} > 0$,

$$\frac{d^2 f(\boldsymbol{w})}{d w_i^2} = \frac{\alpha \mathcal{N}_0 \mu_i (r_i^{min})^2 (\ln 2)^2 2^{r_i^{min}/w_i}}{|h_i^P|^2 w_i^3} > 0. \tag{19}$$

Thus, $f(\boldsymbol{w})$ is a convex function of $\boldsymbol{w}$. Therefore, the objective function of the BA problem is convex. The constraints of the BA problem satisfy the Slater's conditions, and therefore the Karush-Kuhn-Tucher (KKT) conditions provide necessary and sufficient conditions for the optimality of the BA problem [19]. Hence, we can derive optimal bandwidth allocations by solving the KKT conditions of the BA problem.

The PCM problem, thus, can be solved in two steps. In the first step, the user-BS associations are determined. Then, the PCM problem is reduced to the BA problem. In the second step, the BA algorithm is solved by solving its KKT conditions. Since the BA problem can be easily solved, the major difficulty of solving the PCM problem is to optimize the user-BS associations.

When $\sum_{i \in \mathcal{U}} \phi_{k,i} \leq W$ and $k = \arg\min_{j \in \mathcal{S}} \phi_{j,i}$, then all the PUs are offloaded to the SBSs, where the ith PU is offloaded to the kth SBS. In this case, the PBS does not provide data service to any PU, and its dynamic power consumption is zero.

When $\sum_{i \in \mathcal{U}} \phi_{k,i} > W$ and $k = \arg\min_{j \in \mathcal{S}} \phi_{j,i}$, not all PUs can be offloaded to SBSs, and the user-BS associations are to be optimized to minimize the PBS's power consumption. In this case, the PUs can be classified into three categories based on their minimum data rates, their channel conditions and the amount of compensating bandwidth required by the SBSs.

The first category of PUs pertains to the PUs which can only be associated with the PBS. For example, if the ith PU is out of the coverage area of all SBSs or $\phi_{k,i} > W$, $k = \arg\min_{j \in \mathcal{S}} \phi_{j,i}$, then the ith PU can only be associated with the PBS. The second category of PUs involves the PUs which have to be associated with SBSs in order to achieve the optimal solution. For example, if $\phi_{k,i} < w_i^{min}$ and $k = \arg\min_{j \in \mathcal{S}} \phi_{j,i}$, the ith PU is associated with SBSs because by such association, the ith PU consumes less amount of bandwidth and zero dynamic power from the PBS. The third category of PUs refers to the PUs whose user-BS associations are to be determined. These PUs, if associated with the SBSs, consume zero dynamic power from the PBS. However, the PSB has to allocate more bandwidth to SBSs in order to incentivize them to provide data services to these PUs. This results in a reduction of the amount of bandwidth that can be allocated to the PUs which are associated with the PBS, and thus the overall power consumption on the PBS may increase. Therefore, determining user-BS associations for the third category of PUs is the essential task of the HPCM algorithm. Thus, we first present the user filtering algorithm which classifies PUs into three user sets, $\mathcal{U}^P$, $\mathcal{U}^S$, and $\mathcal{U}^T$, which are denoted as the first, the second, and the third category of PUs, respectively. The pseudo code is shown in Algorithm 1.

The user-BS associations of the PUs belonging to $\mathcal{U}^P$ and $\mathcal{U}^S$ are associated with the PBS and SBSs, respectively. The HPCM algorithm is to determine the user-BS associations for the PUs in $\mathcal{U}^T$. When $\sum_{i \in \mathcal{U}} \phi_{k,i} > W$ and $k = \arg\min_{j \in \mathcal{S}} \phi_{j,i}$, the PBS is unable to offload all the PUs to SBSs. Thus, the HPCM algorithm is to iteratively offload the PU, which consumes the largest amount of dynamic power from the PBS, to SBSs.

Assuming the PUs belonging to $\mathcal{U}^T$ are associated with the PBS, the amount of bandwidth, w_i^t, allocated to the ith PU,

Algorithm 1: The User Filtering Algorithm

for *i=1 to* $|\mathcal{U}|$ **do**
  if $\phi_{k,i} \le w_i^{min},\ k = \arg\min_{j\in\mathcal{S}} \phi_{j,i}$ **then**
    Assign the ith PU in the user set $\mathcal{U}^S$;
  else if $\phi_{k,i} > W,\ k = \arg\min_{j\in\mathcal{S}} \phi_{j,i}$ **then**
    Assign the ith PU in the user set $\mathcal{U}^P$;
  else
    Assign the ith PU in the user set $\mathcal{U}^T$;
Return $\mathcal{U}^P$, $\mathcal{U}^S$, and $\mathcal{U}^T$.

$i \in \mathcal{U}^T$, can be derived by solving the BA problem. If $w_i^t \ge \phi_{k,i}$ and $k = \arg\min_{j\in\mathcal{S}} \phi_{j,i}$, it indicates that associating the ith PU with SBSs does not require more bandwidth than associating the PU with the PBS. Meanwhile, associating the ith PU with SBSs reduces the dynamic power consumption of the PBS. Therefore, the ith PU is associated with SBSs. If $w_i^t < \phi_{k,i}$ and $k = \arg\min_{j\in\mathcal{S}} \phi_{j,i}$, by associating the ith PU with SBSs, the PBS reduces its dynamic power consumption. However, the PBS has to allocate an additional amount of bandwidth to SBSs to incentivize them to provide data services to the ith PU. This reduces the amount of available bandwidth for the PUs which are associated with the PBS, and may result in an increment of the PBS's dynamic power consumption in serving these PUs.

When $w_i^t < \phi_{k,i}$ and $k = \arg\min_{j\in\mathcal{S}} \phi_{j,i}$, the PBS's dynamic power savings on offloading the ith PU to the kth SBS depend on two factors. The first one is the PBS's dynamic power consumption in serving the ith PU. The second one is the difference between w_i^t and $\phi_{k,i},\ k = \arg\min_{j\in\mathcal{S}} \phi_{j,i}$. We denote $\Delta w_i^t = \phi_{k,i} - w_i^t$ as the difference. A smaller Δw_i^t indicates that offloading traffic of the ith PU reduces a less amount of bandwidth from the PBS's total bandwidth, and thus results in a less increment on the PBS's dynamic power consumption in serving the rest of PUs. The ratio between the two factors is utilized by the HPCM algorithm to reflect the amount of potential power savings that can be achieved by the PBS in offloading the traffic of a PU to a SBS. The larger the ratio, the more power savings may be achieved by the PBS. We refer to this ratio as the power-bandwidth ratio (PBR). Denote p_i^t and w_i^t as the PBS's transmit power-spectrum density and the corresponding bandwidth allocation toward the ith PU, respectively. The PBR of the ith PU can be expressed as

$$\rho_i = \frac{\alpha p_i^t w_i^t}{\triangle w_i^t} \tag{20}$$

The idea of the HPCM algorithm is to iteratively find a PU with the largest PBR, and offload its traffic to SBSs if power savings can be achieved by the PBS. The HPCM algorithm terminates when the user set $\mathcal{U}^T$ is empty.

At the beginning of each iteration, the HPCM algorithm assumes all the PUs in $\mathcal{U}^T$ are associated with the PBS, and calculates the PBS's total power consumption, C, its transmit power-spectrum density toward the ith PU, p_i^t, and the corresponding bandwidth allocation, w_i^t, $\forall i \in \mathcal{U}^T$. The HPCM algorithm finds the largest $\rho_i,\ i \in \mathcal{U}^T$. Assuming $m = \arg\max \rho_i, i \in \mathcal{U}^T$, the HPCM algorithm associates the mth PU with the kth SBS. Here, $k = \arg\min_{j\in\mathcal{S}} \phi_{j,m}$. Then, the HPCM algorithm calculates the total power consumption of the PBS, which is denoted as C^m. If $C^m < C$, the HPCM algorithm offloads the traffic of the mth PU to the kth SBS and assigns $C = C^m$; otherwise, the mth PU is associated with the PBS.

To ensure its performance, the HPCM algorithm, before the iteration begins, associates the mth PU, $m = \arg\max_{i\in\mathcal{U}^T} \alpha p_i^t w_i^t$, with the kth SBS, $k = \arg\min_{j\in\mathcal{S}} \phi_{j,m}$. Then, the HPCM algorithm calculates the total power consumption of the PBS, C^{max}. At the end, the algorithm compares C^{max} with C, and returns the user-BS associations that achieve the minimum power consumption of the PBS.

The pseudo code of the HPCM algorithm, as described above, is shown in Algorithm 2.

Algorithm 2: The HPCM Algorithm

Assign all PUs in $\mathcal{U}$;
Calculate $\phi_{k,i}$ and w_i^{min}, $k = \arg\min_{j\in\mathcal{S}} \phi_{j,m}$, $\forall i \in \mathcal{U}$;
if $\sum_{i\in\mathcal{U}} \phi_{k,i} \le W$ **then**
  $C = p^{fix}$, and all PUs are associated with SBSs;
else
  $(\mathcal{U}^P, \mathcal{U}^S, \mathcal{U}^T)$ = User Filter Alg. $(\phi_{k,i},\ w_i^{min})$;
  Calculate W^P and derive C, w_i^t, and p_i^t by solving the BA problem with $\mathcal{U}^P = \mathcal{U}^P \cup \mathcal{U}^T$;
  if $\phi_{k,i} \le w_i^t,\ \forall i \in \mathcal{U}^P$ **then**
    Assign the ith PU in $\mathcal{U}^S$;
  Find $m = \arg\max_{i\in\mathcal{U}^T} \alpha p_i^t w_i^t$;
  Calculate C^{max} by solving the BA problem with $\mathcal{U}^P = \mathcal{U}^P \cup \mathcal{U}^T \setminus \{m\}$;
  Assign $\mathcal{U}^P_{max} = \mathcal{U}^P \cup \mathcal{U}^T \setminus \{m\}$;
  while $\mathcal{U}^T$ *is not empty* **do**
    Calculate C, w_i^t, and p_i^t by solving the BA problem with $\mathcal{U}^P = \mathcal{U}^P \cup \mathcal{U}^T$;
    Calculate $\rho_i,\ \forall i \in \mathcal{U}^T$;
    Find $m = \arg\max_{i\in\mathcal{U}^T} \rho_i$;
    if $\sum_{i\in\mathcal{U}^P\cup\mathcal{U}^T\setminus\{m\}} w_i^t + \sum_{i\in\mathcal{U}^S\cup\{m\}} \phi_{k,i} \le W$ **then**
      Calculate C^m by solving the BA problem with $\mathcal{U}^P = \mathcal{U}^P \cup \mathcal{U}^T \setminus \{m\}$;
      if $C^m < C$ **then**
        Offload the mth PU to $\mathcal{U}^S$;
        Assign $C = C^m$;
      else
        Assign the mth PU into $\mathcal{U}^P$;
    else
      Set primary user m in $\mathcal{U}^P$;
    Update $\mathcal{U}^T = \mathcal{U}^T \setminus \{m\}$;
  if $C^{max} < C$ **then**
    Assign $\mathcal{U}^P = \mathcal{U}^P_{max}$, and $\mathcal{U}^S = \mathcal{U} \setminus \mathcal{U}^P$
Derive w_i by solving the BA algorithm;
Return $\mathcal{U}^P$, $\mathcal{U}^S$ and $w_i, \forall i \in \mathcal{U}^P$.

B. Performance Analysis

Since the user-BS associations of PUs in both $\mathcal{U}^P$ and $\mathcal{U}^S$ are determined, the HPCM algorithm optimizes the user-BS associations for the PUs in $\mathcal{U}^T$. If the brute-force search is applied, the total number of possible combinations of the user-BS associations for the PUs in $\mathcal{U}^T$ is $2^{|\mathcal{U}^T|}$. The computation complexity of solving the KKT conditions of the BA problem is $O(|\mathcal{U}|)$. Therefore, the computation complexity of the brute-force search, in the worst case, is $O(|\mathcal{U}|2^{|\mathcal{U}^T|})$. When $|\mathcal{U}^T|$ is large, the brute-force search is very inefficient, and is even impossible to solve the PCM problem within a reasonable time. As compared with the brute-force search, the HPCM algorithm incurs significantly less computational complexity. The $while$ loop at most requires $|\mathcal{U}^T|$ iterations. In the $while$ loop, the finding the PU with the largest PBR requires at most $|\mathcal{U}^T|\log|\mathcal{U}^T|$ iterations. The complexity of solving the BA problem is $O(|\mathcal{U}|)$. Therefore, the worst case computation complexity of the HPCM algorithm is $O(|\mathcal{U}^T|(|\mathcal{U}|+|\mathcal{U}^T|\log|\mathcal{U}^T|))$.

Although with significantly less computation complexity, the HPCM algorithm's performance in terms of minimizing the PBS's power consumption is not compromised very much. In fact, the PBS's power savings achieved by the HPCM algorithm is at least 50% of that achieved by the brute-force search when the PBS experiences heavy traffic load from cell edge users.

Lemma 2. *If $\rho_m > \rho_j,\ m, j \in \mathcal{U}^T$, the mth PU's user-BS association does not depend on the jth PU's user-BS association.*

Proof: The proof is presented in Appendix A. ■

Lemma (2) is important to guarantee the correctness of the HPCM algorithm. According to the HPCM algorithm, when $\rho_m > \rho_j$, the mth PU's user-BS association is determined prior to the jth PU's. In this case, if the mth PU's user-BS association depends on the jth PU's user-BS association, then the HPCM algorithm cannot determine the mth PU's user-BS association before determining the jth PU's, which contradicts the procedure of the HPCM algorithm. However, Lemma (2) proves that the mth PU's user-BS association does not depend on the jth PU's, which ensures the correctness of the HPCM algorithm.

Let $m = \arg\max_{j\in\mathcal{U}^T}\rho_j$. Denote ρ_i and ρ_i^m as the ith PU's PBR before and after determining the mth PU's serving BS, repetitively.

Lemma 3. *When the PBS experiences heavy traffic from both cell edge users and inner cell users, if $\rho_i \geq \rho_k, \forall i, k \in \mathcal{U}^T \setminus \{m\}$, $\rho_i^m \geq \rho_k^m$.*

Proof: The proof is presented in Appendix B. ■

Theorem 2. *When $\sum_{i\in\mathcal{U}}\phi_{k,i} \leq W$ and $k = \arg\min_{j\in\mathcal{S}}\phi_{j,i}$, both the HPCM algorithm and the brute-force search achieve the same solution.*

Proof: If $\sum_{i\in\mathcal{U}}\phi_{m,i} \leq W$ and $m = \arg\min_{j\in\mathcal{S}}\phi_{j,i}$, then all the PUs can be offloaded to SBSs, and the PBS's dynamic power consumption is zero. For this scenario, both the HPCM algorithm and the brute-force search achieve the same solution. ■

Define a relaxed PCM problem where the PUs can be partially associated with SBSs as the RPCM problem. In this case, if the mth PU is partially associated with the kth SBS, and γ_m percent of the mth PU's data service is provided by the kth SBS, and the other portion of data service is provided by the PBS, then the kth SBS's bandwidth requirement to partially serve the mth PU is defined as $\phi_{m,k}^p = \gamma_m\phi_{m,k}$. We assume $(m-1)$ PUs are fully offloaded to SBSs. The mth PU is defined as the PU with the largest PBR after offloading the $(m-1)$ PUs to the SBS. Denote E^P as the PBS's power savings. We assume that the mth PU is partially offloaded to the SBS. Define the PBS's power savings achieved by solving the PCM problem with the brute-force search as E^B. We assume the PBS experiences heavy traffic load from cell edge users and the SBS's compensating bandwidth is properly selected. In this case, there always exists a PU which can be offloaded to the SBS, thus resulting in the reduction of energy consumption of the PBS.

Lemma 4. *When $\sum_{i\in\mathcal{U}}\phi_{k,i} > W$ and $k = \arg\min_{j\in\mathcal{S}}\phi_{j,i}$, $E^P >= E^B$.*

Proof: Since the RPCM problem is a relaxed version of the PCM problem, every solution for the PCM problem is feasible for the RPCM problem. Since we assume the PBS experiences heavy traffic load from the cell edge users and the SBS's compensating bandwidth is properly selected, there always exists a PU which can be offloaded to the SBS, thus resulting in the reduction of energy consumption of the PBS. In other worlds, we assume that given the maximum transmit power constraint is satisfied, offloading traffic load to the SBS can reduce the PBS's power consumption. Therefore, $E^P >= E^B$. ■

Theorem 3. *When $\sum_{i\in\mathcal{U}}\phi_{k,i} > W$ and $k = \arg\min_{j\in\mathcal{S}}\phi_{j,i}$, the maximum power savings achieved by the HPCM algorithm is at least 50% of that achieved by the brute-force search.*

Proof: When $\sum_{i\in\mathcal{U}}\phi_{k,i} > W$ and $k = \arg\min_{j\in\mathcal{S}}\phi_{j,i}$, not all the PUs can be offloaded to SBSs. In this case, we show that the maximum power savings achieved by the HPCM algorithm is at least 50% of that achieved by the brute-force search.

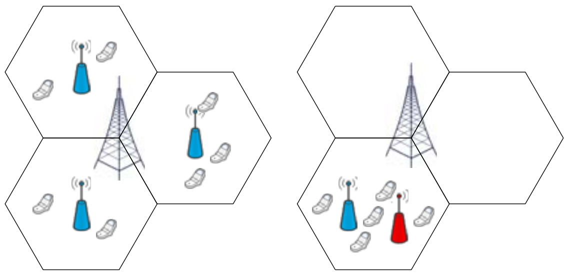
Denote E^H, E^M, and E^{max}as the PBS's power savings by offloading the first $(m-1)$th PUs, the mth PU and the PU with the maximum power consumption, respectively. Then, the power saving achieved by the HPCM algorithm is $\max\{E^H, E^{max}\}$. Since E^P is the power savings by partially offloading the mth PU, $E^H + E^M \geq E^P$.

$$\begin{aligned}\frac{\max\{E^H, E^{max}\}}{E^B} &\geq \frac{\max\{E^H, E^M\}}{E^P}\\ &\geq \frac{\max\{E^H, E^M\}}{E^H + E^M}\\ &\geq 0.5. \end{aligned}\tag{21}$$

■

VI. Simulation Results

Two simulation scenarios are set up to evaluate the performance of the proposed EST scheme and the HPCM algorithm.



(a) Topology 1 (b) Topology 2

Fig. 2. The simulation topology.

In the simulations, we adopt COST 231 Walfisch-Ikegami [20] as the propagation model with 9 dB rayleigh fading and 5 dB shadowing fading for both the PBS and SBSs. The carrier frequency is 2110 MHz, the antenna feeder loss is 3 dB, the transmitter gain is 1 dB, the noise density is 10^{-10} w/Hz, and the receiver sensitivity is -97 dB. The total amount of licensed bandwidth is 20 MHz and the PBS's maximum transmit power is 20 w ($43dBm$)[17]. Thus, the PBS's maximum transmit power-spectral density is 1 $\mu w/Hz$. Based on the measurement results in [17], we set $\alpha = 25$ and $p^{fix} = 700$ w. SBSs are assumed to have the same energy consumption model as that of the PBS. The SBSs' static power consumption is 14 w, and the SBSs' coefficient between the dynamic power consumption and their transmit power is 2. We assume the SBS's transmit power-spectral density is 20 $\mu w/kHz$ [21].

The simulation typologies are shown in Fig. 2. A radio cell, which is covered by the PBS, is divided into three sectors. The radius of the radio cell is 1.5 km, and the PBS is located at the center of the radio cell. The PU's minimum data rate is 500 $kbps$. In the simulations, the energy efficiency (EE) is calculated by dividing PUs' total data rate by the sum of the PBS's power consumption and the SBSs' dynamic power consumption in serving PUs; the spectrum efficiency (SE) is calculated by dividing PUs' total data rate by the sum of the bandwidth allocated to PUs by both the PBS and SBSs.

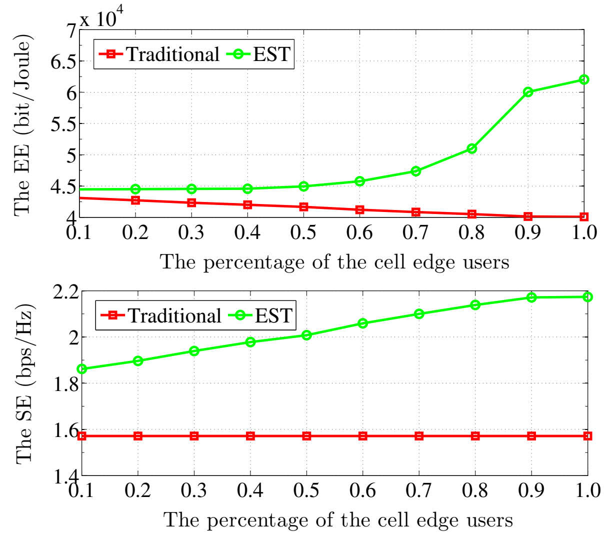


Fig. 3. The performance of the EST v.s. the percentage of the cell edge users (topology 1).

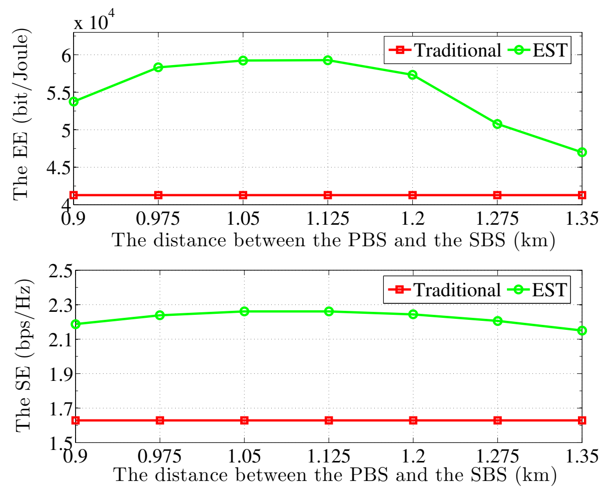


Fig. 4. The performance of the EST v.s. the distance between the PBS the SBSs (topology 1).

A. Simulation Scenario One

In this simulation scenario, we consider a radio cell with one PBS and one SBS in each sector, as shown in Fig. 2(a). The SBSs have the same operation parameters such as the transmit spectral-power density, the per-PU compensating bandwidth, and the distance between the SBS and the PBS. We define the cell edge users as the users whose distances from the PBS are larger than 0.9 km.

Fig. 3 shows the EE and SE of the network versus the percentage of the cell edge users. In this simulation, the total number of mobile users in each sector follows the Poisson distribution with the mean equaling to 20. As the percentage of the cell edge users increases, the EE of the traditional scheme decreases because serving cell edge users usually requires more energy consumption. The EE of the EST scheme increases because more users are offloaded to the SBS. For the same reason, the SE of the EST scheme also increases.

Since the EST scheme aims to offload the cell edge users to enhance the network efficiency, in the following simulations, we assume that PUs are randomly distributed at the edge of each sector. The number of mobile users in each sector follows Poisson distribution with the mean equaling to 20.

Fig. 4 compares the EE and SE of the traditional scheme and the EST scheme versus the distance between the PBS and SBSs. Here, the traditional scheme refers to as the scheme in which all PUs are served by the PBS. In this simulation, we assume the distances between the PBS and three SBSs are the same. For all the SBSs, the per-PU compensating bandwidth is 100 kHz. As shown in Fig. 4, the EST scheme enhances the EE and the SE by 43.60% and 38.79%, respectively. When the distance between the PBS and SBS is 1.05 km or 1.125 km, the EE and SE achieve the maximum values, respectively. As shown in Fig. 5, as the distance between the PBS and SBSs increases, the average distance between the mobile users and the SBS decreases until reaching its minimum, and then

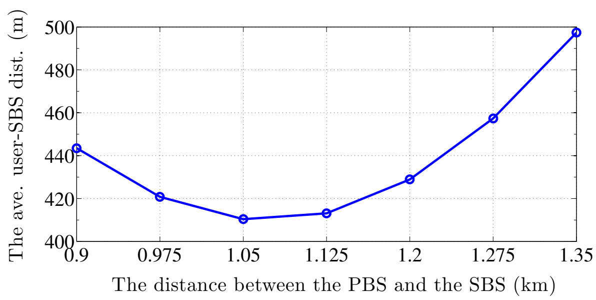


Fig. 5. The average distance between mobile users and the SBSs (topology 1).

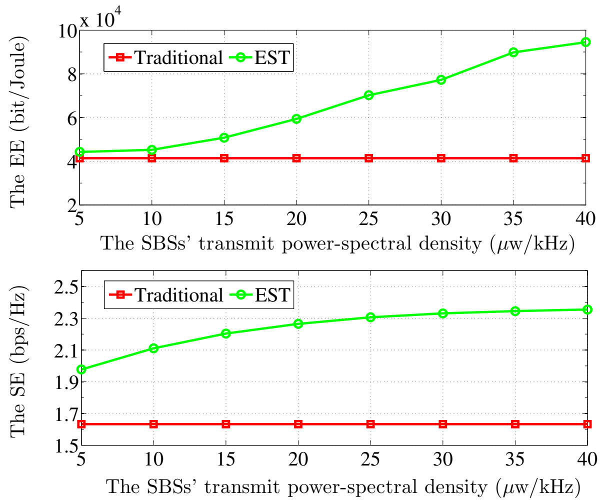


Fig. 6. The performance of the EST v.s. the SBSs' power-spectral density (topology 1).

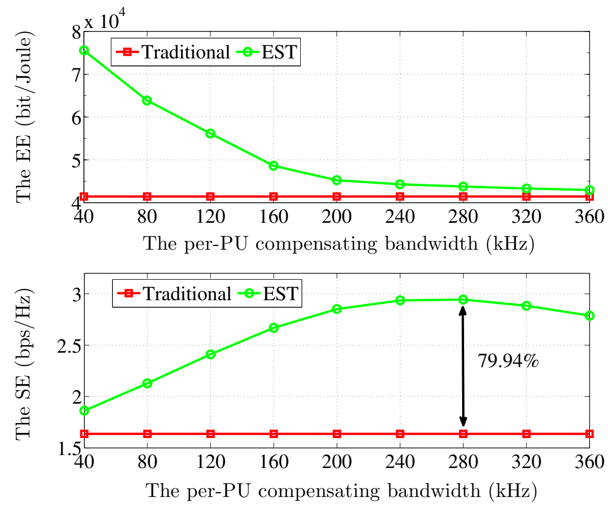


Fig. 7. The performance of the EST v.s. the SBSs' compensating bandwidth (topology 1).

increases. When the average distance between the mobile users and the SBS decreases, the SBSs may offload more PUs from the PBS and may require less bandwidth for offloading the same PU. Thus, both the EE and SE increase. After the EE and SE reach their peaks, they decrease as the average distance between the mobile users and the SBSs increases. The EE decreases faster than the SE. As the average distance between the mobile users and the SBS increases, for offloading the same PU, the SBS requires more bandwidth. As a result, less bandwidth is available in the PBS, which increases the PBS's power consumption.

Remark 1. This observation indicates that the location of the SBS significantly impacts the performance of the mobile network under the EST scheme in terms of the EE and SE. For network planning, the locations of SBSs deployed by ISPs may not be optimized for the purpose of enhancing the EE and SE of mobile networks. However, if the EST scheme is considered, in order to maximize their profits from utilizing the licensed bandwidth, the ISPs are desired to maximize their traffic loading by optimizing the locations of the SBSs. Furthermore, if the EST scheme is considered, the mobile network operators and ISPs can jointly optimize their network planning to maximize their profits.

Fig. 6 shows the EE and the SE of the network versus the SBS's transmit power-spectral density. In this simulation, the distance between the PBS and SBSs is 1.125 km, and the SBS's per-PU compensating bandwidth is 100 kHz. When the SBS's transmit power-spectral density increases, more PUs fall into the coverage area of SBSs and the SBSs' bandwidth requirements in serving the PUs are reduced. As a result, more PUs are offloaded to SBSs. Therefore, both the EE and the SE improves as the SBSs' transmit power-spectral density increases.

Fig. 7 shows the performance of the EST versus the SBSs' compensating bandwidth. In this simulation, the distance between the PBS and SBSs is 1.125 km, and the SBS's transmit power-spectral density is 20 $\mu w/kHz$. As the SBS's per-PU compensating bandwidth increases, the EE of the network decreases because less PUs are offloaded to SBSs. When the SBS's per-PU compensating bandwidth increases, the SE increases, peaks when the per-PU compensating bandwidth equals to 280 kHz, and then decreases. This is because when the SBS's per-PU compensating bandwidth increases, although the number of offloaded PUs decreases, the total amount of the bandwidth obtained by SBSs increases due to the larger per-PU compensating bandwidth. Thus, the SE shows the concavity. This observation indicates that the per-PU compensating bandwidth should be properly selected to optimize the trade-off between the EE and the SE of the network. In addition, when the SBSs' compensating bandwidth equals to 360 kHz, the EE of the network under the EST is only slightly larger than that under the traditional scheme while the SE of the network under the EST is significantly larger than that under the traditional scheme. The improvement of the SE indicates that some PUs are still offloaded to the SBSs even when the compensating bandwidth is as large as 360 kHz. The traditional scheme and the EST show the similar EE because a large compensating bandwidth leads to a reduced available bandwidth in the PBS. As a result, the PBS has to increase its transmit power density to satisfy users' data rate requirement. Therefore, the PBS's power consumption increases and the EE of the network under the EST decreases. As the compensating bandwidth increases, the EE of the network under the EST

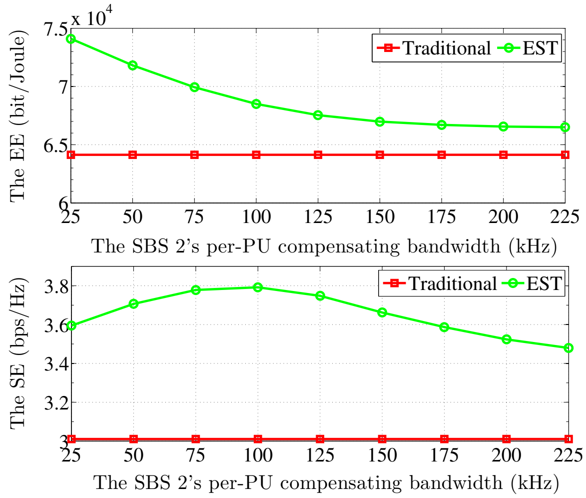


Fig. 8. The performance of the EST v.s. the SBSs' compensating bandwidth (topology 2).

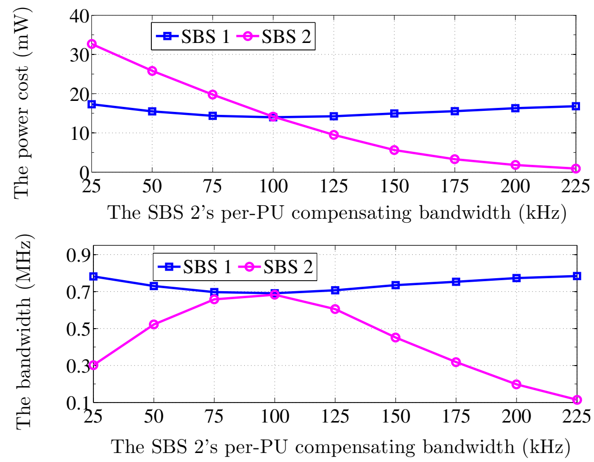


Fig. 9. The SBSs' power cost and obtained bandwidth v.s. the SBSs' compensating bandwidth (topology 2).

keeps decreasing until converging to the EE of the traditional scheme. When the traditional scheme and the EST have the same EE, the performance of the SE reflects whether PUs are offloaded to the SBS.

B. Simulation Scenario Two

As shown in Fig. 2(b), we consider only one sector of the radio cell in this simulation scenario. PUs are randomly distributed at the edge of each sector. The distances between the PUs and the PBS are larger than 0.9 km. The number of mobile users in the sector follows Poisson distribution with the mean equaling to 40. Two SBSs are deployed in the sector. The distance between the PBS and the SBSs is 1.125 km, and the distance between the two SBSs is also 1.125 km. Both the SBSs' transmit power-spectral density is 20 $\mu w/kHz$. The per-PU compensating bandwidth of one of the SBSs, SBS 1, is set to 100 kHz. We vary the per-PU compensating bandwidth of the other SBS to show the interactions among PBS and the SBSs.

Fig. 8 show the network's EE and SE when SBS 2 varies its per-PU compensating bandwidth. When SBS 2 increases its per-PU compensating bandwidth, the EE of the network decreases because less PUs are being offloaded to SBSs. Meanwhile, the SE of the network shows the concavity for the same reason shown in Fig. 7.

Fig. 9 shows the interaction between two SBSs. As SBS 2's per-PU compensating bandwidth increases, its dynamic power cost decreases because less PUs are offloaded to it. At the same time, SBS 1's power cost increases because more PUs are associated with SBS 1. This indicates that as the SBS 2's per-PU compensating bandwidth increases, the PUs, who are originally offloaded to SBS 2, are associated with SBS 1. On the other hand, as the SBS 2's per-PU compensating bandwidth increases, the amount of bandwidth obtained by SBS 1 increases because it serves more PUs. Meanwhile, the amount of bandwidth obtained by SBS 2 shows the concavity for the same reason as explained before. We can observe from the simulation result that given SBS 1's strategies in terms of transmit power-spectral density and the per-PU compensating bandwidth, the profits of SBS 2 in terms of the amount of obtained bandwidth and the power cost can be maximized by selecting an optimal per-PU compensating bandwidth.

VII. Conclusion

In this paper, we have proposed a novel energy spectrum trading (EST) scheme which enables the mobile traffic offloading between the mobile networks and the ISPs' networks by leveraging cognitive radio techniques. We have shown that achieving optimal mobile traffic offloading is NP-hard. We have proposed a heuristic power consumption minimization (HPCM) algorithm to approximate the optimal solution with low computation complexity. The HPCM algorithm enables the mobile traffic offloading, and significantly enhances the energy and spectral efficiency of mobile networks.

Appendix A
Proof of Lemma 2

Based on the HPCM algorithm, when $\rho_m > \rho_j, m,\ j \in \mathcal{U}^T$, the mth PU's user-BS association is determined prior to the jth PU's. On the one hand, if the mth PU is associated with a SBS by the HPCM algorithm, the jth PU's user-BS association does not change the mth PU's user-BS association. On the other hand, if the mth PU is not associated with a SBS by the HPCM algorithm, it is because either 1) $\sum_{i\in\mathcal{U}^P\cup\mathcal{U}^T\setminus\{m\}} w_i^t + \sum_{i\in\mathcal{U}^S\cup\{m\}} \phi_{k,i} > W$, $k = \arg\min_{j\in\mathcal{S}} \phi_{j,i}$, or 2) offloading the mth PU increases the PBS's dynamic power consumption.

For the first case, if the jth PU is associated with the kth SBS, because $\phi_{k,j} - w_j^t > 0$, $\sum_{i\in\mathcal{U}^P\cup\mathcal{U}^T\setminus\{m,\ j\}} w_i^t + \sum_{i\in\mathcal{U}^S\cup\{m\}} \phi_{k,i} + \phi_{k,j} - w_j^t > W$. Therefore, the mth PU still cannot be associated with a SBS.

For the second case, the jth PU's user-BS association does not change the fact that the PBS' power savings in offloading the mth PU to the SBS is negative. Given user-BS associations, in order to minimize the PBS's power consumption, the bandwidth allocations are optimized by solving the BA problem. Therefore, any bandwidth allocation solution which is different from the solution obtained by solving the BA problem will not reduce the PBS's power consumption. Based on this observation, we can prove that even if the jth PU is associated with a SBS, the PBS's power savings in offloading the mth PU to the SBS is still negative.

Assume both the mth PU and the jth PU are associated with the PBS. Let $W^0 = W - \sum_{i \in \mathcal{U}^S} \phi_{k,i}$, $k = \arg\min_{j \in \mathcal{S}} \phi_{j,i}$. Denote w_i^0 and p_i^0 as the bandwidth allocation and the PBS's transmit power density for the ith PU, $i \in \mathcal{U}^P \cup \mathcal{U}^T$, derived by solving the BA problem with $W^P = W^0$ and $\mathcal{U}^P = \mathcal{U}^P \cup \mathcal{U}^T$. Define $C_m^0 = \sum_{i \in \mathcal{U}^P \cup \mathcal{U}^T} \alpha w_i^0 p_i^0 + p^{fix}$;

When the mth PU is associated with the kth SBS and the jth PU is associated with the PBS, the total amount of bandwidth available for the PUs in $\mathcal{U}^P \cup \mathcal{U}^T \setminus \{m\}$ is reduced by $\Delta w_m^0 = \phi_{k,m} - w_m^0$. The bandwidth reduction results in an increase of the PBS's dynamic power consumption in serving its associated PUs. Denote w_i^1 and p_i^1 as the bandwidth allocation and the PBS's transmit power density for the ith PU, $i \in \mathcal{U}^P \cup \mathcal{U}^T \setminus \{m\}$, derived by solving the BA problem with $W^P = W^0 - \Delta w_m^0$ and $\mathcal{U}^P = \mathcal{U}^P \cup \mathcal{U}^T \setminus \{m\}$. Define $C_m^1 = \sum_{i \in \mathcal{U}^P \cup \mathcal{U}^T} \alpha w_i^1 p_i^1 + p^{fix}$; since both w_i^1 and p_i^1 are derived by solving the BA problem, C_m^1 is minimized. Denote $\Delta C = C_m^1 - C_m^0$ as the increased power consumption owing to offloading the mth PU to the SBS. In the process of the minimization, if the bandwidth allocation and the transmit power density toward the jth PU do not change, the power consumption increases do not come from the jth PU. Therefore, whether the jth PU is offloaded to the SBS does not change the mth PU's user-BS association. On the other hand, if the jth PU's power consumption increases owing to the bandwidth reduction, the bandwidth reduction is derived by solving the BA problem to minimize the overall energy consumption. In other words, in this case, if we keep the jth PU's power consumption unchanged during the process of solving the BA algorithm, the PBS's power consumption will be larger than C_m^1. As a result, even if the jth PU is offloaded to the SBS, the mth PU still cannot be offloaded to SBSs.

Appendix B
Proof of Lemma 3

Proving Lemma 3 is to prove the order of PBRs of users in $\mathcal{U}^T$ do not change during the iterations. $\phi_{k,i}$ is determined by the channel condition between the ith PU and the kth SBS, the ith PU's data rate requirement, and the kth SBS's compensating bandwidth. These parameters do not change during the iterations. Thus, ρ_i is determined by p_i^t and w_i^t, which are derived by solving the BA problem. The BA problem is solved by solving its KKT conditions. Then, w_i^t can be derived by solving the following equation array.

$$\begin{cases} \alpha(2^{r_i/w_i^*} - 1 - \dfrac{r_i}{w_i^*} 2^{r_i/w_i^*} ln2) \dfrac{\mathcal{N}_0}{|h_i^P|^2} = \nu^* \\ \displaystyle\sum_{i \in \mathcal{U}^P} w_i^* = W^P. \end{cases} \tag{22}$$

Here, w_i^* and ν^* are the primal and dual optimal points for the BA problem, respectively. Although there is no close form solution for the above equation array, we can derive the structure of the optimal solutions based on which we prove the lemma. Let

$$\psi(w_i) = \alpha(2^{r_i/w_i} - 1 - \frac{r_i}{w_i} 2^{r_i/w_i} ln2) \frac{\mathcal{N}_0}{|h_i^P|^2}. \tag{23}$$

Since $w_i \geq w_i^{min}$ and $\frac{d\psi(w)}{dw} > 0$, $\psi(w_i^{min}) = \min_{w_i} \psi(w_i)$. Based on the first equation in the equation array, w_i^* can be expressed as a function of ν^*.

$$w_i^* = \begin{cases} \varphi(r_i, |h_i^P|^2, \nu^*), & \nu^* > \psi(w_i^{min}) \\ w_i^{min}, & \nu^* \leq \psi(w_i^{min}). \end{cases} \tag{24}$$

Here, $\varphi(r_i, |h_i^P|^2, \nu^*)$ is derived based on the first equation in equation array (22). Since $r_i = w_i^{min} log(1 + \frac{p^{max}|h_i^P|^2}{\mathcal{N}_0})$, $\psi(w_i^{min})$ can be expressed as

$$\psi(w_i^{min}) = \alpha(p^{max} - log(1 + \frac{p^{max}|h_i^P|^2}{\mathcal{N}_0}))(\frac{\mathcal{N}_0}{|h_i^P|^2} + p^{max}) ln2) \tag{25}$$

$\psi(w_i^{min})$ can be considered as a function of $|h_i^P|^2$. When $p^{max}|h_i^P|^2 \gg \mathcal{N}_0$, $\frac{d\psi(w_i^{min})}{d|h_i^P|^2} < 0$. Therefore, a small $|h_i^P|^2$ leads to a large $\psi(w_i^{min})$. When the PBS experiences heavy traffic from both cell edge users and inner cell users, the optimal bandwidth allocations toward cell edge users equal to their minimum required bandwidths. When the proposed scheme is applied to offload cell edge users, it is reasonable to assume that the users in $\mathcal{U}^T$ are the cell edge users. For these users, their optimal bandwidth allocations derived by solving the BA problem are their minimum required bandwidths which do not change during each iteration. Hence, the order of PBR of users in $\mathcal{U}^T$ does not change during the iterations.

References

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

[2] D. Zeller, M. Olsson, O. Blume, A. Fehske, D. Ferling, W. Tomaselli, and I. Gódor, "Sustainable wireless broadband access to the future Internet - The EARTH project," in *The Future Internet*, ser. Lecture Notes in Computer Science, F. Alvarez et al., Eds. Berlin/Heidelberg: Springer, 2013, vol. 7858, pp. 249–271.

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

[4] M. Etoh, T. Ohya, and Y. Nakayama, "Energy consumption issues on mobile network systems," in *Proc. Int. Symp. Appl. Internet (SAINT)*, Aug. 2008.

[5] "Comcast xfinity hotspots map." [Online]. Available: http://hotspots.wifi.comcast.com/

[6] T. Han and N. Ansari, "Opportunistic content pushing via wifi hotspots," in *Proc. 3rd IEEE Int. Conf. Netw. Infrastructure Digital Content (IC-NIDC)*, Sep. 2012.

[7] ——, "Powering mobile networks with green energy," *IEEE Wireless Commun. Mag.*, vol. 21, no. 1, pp. 90–96, Feb. 2014.

[8] ——, "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.

[9] 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, Physical Social Comput.*, Dec. 2010.

[10] 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.

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

[12] 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.

[13] S. Corroy, L. Falconetti, and R. Mathar, "Dynamic cell association for downlink sum rate maximization in multi-cell heterogeneous networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, June 2012.

[14] D. Fooladivanda and C. Rosenberg, "Joint resource allocation and user association for heterogeneous wireless cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 248–257, 2013.

[15] R. Madan, J. Borran, A. Sampath, N. Bhushan, A. Khandekar, and T. Ji, "Cell association and interference coordination in heterogeneous lte-a cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1479–1489, 2010.

[16] 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*, June 2010.

[17] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. Imran, D. Sabella, M. Gonzalez, O. Blume, and A. Fehske, "How much energy is needed to run a wireless network?" *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40–49, 2011.

[18] M. R. Garey and D. S. Johnson, *Computer and Intractability: A Guide to the Thoery of NP-Completeness*. New York: W. H. Freeman and Company, 1979.

[19] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[20] "Evolution of land mobile radio (including personal) ccommunications: Cost 231." [Online]. Available: http://www.awe-communications.com/Propagation/Urban/COST/

[21] "Cisco Aironet 1550 series outdoor access point." [Online]. Available: http://www.cisco.com/en/US/prod/collateral/wireless/ps5679/ps11451/data_sheet_c78-641373.pdf

Tao Han (S'08) received the B.E. in electrical engineering and M.E. in computer engineering from Dalian University of Technology and Beijing University of Posts and Telecommunications, respectively. He is currently a Ph.D candidate in the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, New Jersey. His research interests include wireless communications, mobile and cellular networking, network optimization, green communications, and Internet of things.

Nirwan Ansari (S'78-M'83-SM'94-F'09) received the BSEE (summa cum laude with a perfect GPA) from NJIT, the MSEE from the University of Michigan, Ann Arbor, and the Ph.D. from Purdue University, West Lafayette, IN. 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 has been a Visiting (Chair/Honorary) Professor at several universities. His current research focuses on various aspects of broadband networks and multimedia communications.

Prof. Ansari has served on the editorial/advisory board of nine journals. He was elected to serve on the IEEE Communications Society (ComSoc) Board of Governors as a member-at-large (2013–2015), as well as the IEEE Region 1 Board of Governors as the IEEE North Jersey Section Chair. He has chaired ComSoc technical committees, and has actively organized numerous IEEE International conferences/symposia/workshops, assuming leadership roles as Chair or TPC Chair.

Prof. Ansari is the author of *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 is the editor of *Neural Networks in Telecommunications* (Springer 1994) with B. Yuhas. He has also co-authored 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 is frequently selected to deliver keynote addresses, distinguished lectures, and tutorials. Some of his recent recognitions include a couple of best paper awards, several Excellence in Teaching Awards, the Thomas Alva Edison Patent Award (2010), the NJ Inventors Hall of Fame Inventor of the Year Award (2012), the NCE Excellence in Research Award (2014), and designation as a ComSoc Distinguished Lecturer (2006–2009).