A Stochastic Optimization Framework for Energy Efficiency Optimization in 5G Heterogeneous Wireless Access Networks

Chan-Ching Hsu and J. Morris Chang, *Senior Member, IEEE*

Abstract—Fifth generation (5G) wireless networks will provide high-bandwidth connectivity with quality-of-service (QoS) support to mobile users in a seamless manner. In such a scenario, a mobile user will be able to connect to a heterogeneous wireless access networks, comprised of such as a wireless metropolitan area network (WMAN) with different types of bases stations (for example, macrocell and small cell sites), and a wireless local area network (WLAN). We present a stochastic optimization framework for energy efficiency optimization (that is, bandwidth allocation and energy consumption) in such a heterogeneous wireless access environment. First, on a long-term basis, the first stage is to obtain the operation mode for each station and the bandwidth reservation to a service area from the different access networks available in that service area. The first-stage decision variables of this formulation give the optimal reservations which maximizes the energy efficiency of all the resource management in the networks (that is, in all of the service areas). Second, we formulate the second stage to obtain stations to be associated with and the amount of bandwidth allocated to a traffic requirement (in a service area) by the different access networks (on a short-term basis).

Index Terms—

1 Introduction

One of the most important features of the evolving fifth generation (5G) wireless networks will be heterogeneous wireless access in which a mobile will be able to connect to several wireless access networks (for example, wireless local area network (WLAN) and wireless metropolitan area network (WMAN)). The existing forth generation networks even employ small cells to improve the capacity and energy efficiency. This would facilitate high-speed wireless Internet access for services such as multimedia streaming from a fixed server using multiple threads at the application layer. In such a heterogeneous wireless access environment, the radio resource management (RRM) mechanisms for bandwidth allocation, congestion control, and admission control must be designed to satisfy both the mobile users' and the service providers' requirements. After all, a mobile user will be able to connect to the Internet in a seamless manner and also the wireless resources need to be managed efficiently from the service providers' point of view.

A mobile with multiple radio transceivers is able to connect to the different radio access networks. We consider a geographic region that is entirely covered by WWAN base station (i.e., LTE macro cells), partly covered by the small base station, and partly by the WLAN access point, as shown in Fig. 1. Users in the different service areas in this region have an access to different types and different numbers of wireless networks. We present a two-stage stochastic optimization framework for wireless access in a heterogeneous network environment. The objective of the proposed framework is to maximize the system energy efficiency through efficient station active/sleep operation, channel bandwidth reservation to WMAN cells, and resource allocation in a service area.

We formulate the problem of energy efficiency optimization into a two-stage stochastic maximization problem. The limited available bandwidth of a network operator must be reserved to each base station so that the traffic served of the different access networks, which are presumably operated by the same service providers, are maximized. In this competitive environment, therefore, we use the fist-stage decision to obtain the solution of the operation policy and the bandwidth reservation by the different access networks to a station. Again, for stations operating across multiple service areas, a portion of the radio resources in a base station needs to be allocated for traffic from service areas. Since service areas share the available bandwidth in a station, an agreement on bandwidth allocation can be made so that the desired quality-of-service (QoS) performances (for example, user blocking probability) can be achieved. Therefore, we formulate the second-stage decision variables to obtain the capacity (that is, bandwidth) allocation strategies for different scenarios. Both station operation and cpacity reservation can be performed on a long-term basis on the (or steady state) user number and traffic statistics in the different service areas. On the other hand, bandwidth allocation must be performed on a short-term basis and should be adapted upon the user number and traffic requirement in a service area. Again, each network in a service area aims at minimizing its energy consumption while offering bandwidth to covered populations.

Although the problem of bandwidth allocation and admission control in a homogeneous wireless network was extensively studied in the literature (for example, in [1], [2], [3], [4]), it has not been investigated thoroughly in a

- *The authors are with the Department of Electrical and Computer Engineering, Iowa State University, Ames, Iowa, 50011. E-mail: {cchsu, morris}@iastate.edu.*

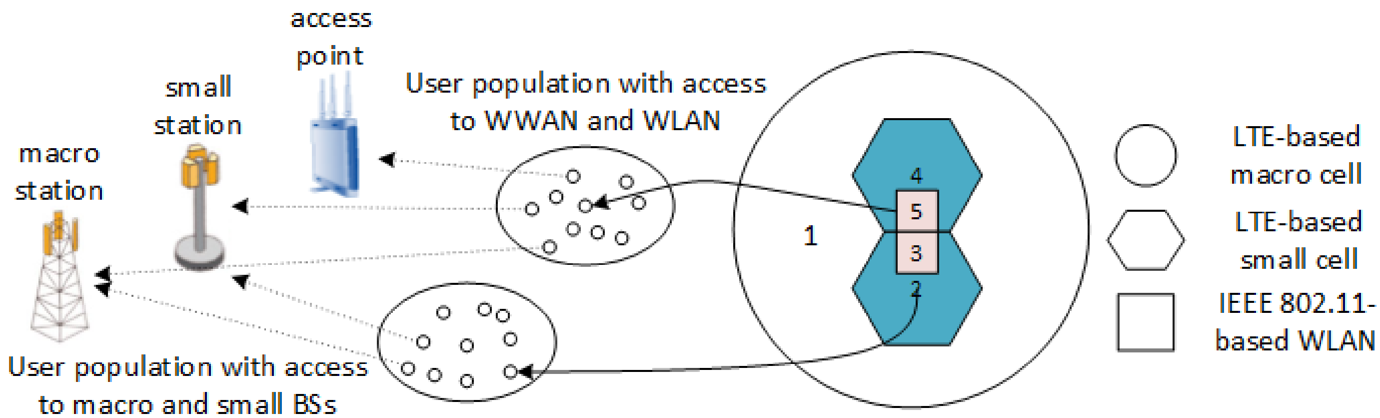


Fig. 1. Service areas under consideration in a heterogeneous wireless access environment

heterogeneous wireless access setting considering both the user-centric and network-centric viewpoints. In a heterogeneous environment, different networks could be operated by the same service provider. The service provider is rational to maximize its revenue. Therefore, the stochastic programming model is useful in modeling and analysis of energy efficiency optimizaiton protocols in this competitive scenario.

The rest of this paper is organized as follows: Related works are reviewed in Section 2. The model for the heterogeneous wireless access network and the components of the proposed energy efficiency optimization framework are described in Section 3. Section 4 presents the operation control mechanisms and the network-level bandwidth allocation and the capacity reservation mechanisms in the proposed energy efficiency maximization framework. Section 5 reviews the general scenario reduction method.Section 6 presents the performance evaluation results. Conclusions are stated in Section 7.

2 Related Work

The issues related to integration of diverse wireless access technologies such as cellular, WLAN, and mobile ad hoc networks (MANET) and provisioning of QoS to the mobile users were studied in [5]. In [6], an adaptive transport layer (ATL) was proposed for heterogeneous wireless networks with the capabilities of adaptive congestion control, multimedia support, and providing fairness of transmission. In [7], a network selection mechanism for 4G wireless networks was proposed. Specifically, gray relation analysis was used to decide which network should be used for each mobile. This decision is based on users' preferences, application requirements, and networks conditions. However, the problem of bandwidth allocation was not considered.

Game theory has been used for resource management in wireless networks. In [8], the admission and rate control problem for CDMA systems was formulated as a noncooperative game. The formulation considered the choice of a user to churn from a current provider to another. The decision on whether a new user can be admitted or not and the allocated transmission are determined from the Nash equilibrium. An admission control game for CDMA networks was formulated in [9] to obtain an efficient and fair resource allocation for multiple classes of traffic. Game theory was also used to solve the power control problem in wireless networks [10], [11], [12]. However, all these works considered the radio resource management problem in a single wireless access network.

On the other hand, small cells have been considered as a promising technique, and have been integrated in current and future radio access networks [13]. The authors of [14] study the problem of downlink power allocation to maximize the capacity in a cellular network where bi-level hierarchy exists. The the Stackelberg game model is used to formulate the optimization problem and Stakelberg equilibrium is solved. The distributed power allocation strategies for a spectrum sharing femtocell network are considered in [15], and the Stackelberg game is formulated to jointly consider the utility maximization of the macrocell and the femtocells. Due to its short transmit-receive distance property, femtocell technique can also reduce energy consumption, prolong handset battery life, and increase network coverage [16], [17]. The problem of energy efficiency for femtocell base station is studied in [18], where a novel energy saving procedure is designed for femtocell base stations. In [19], resource sharing and access control in OFDMA femtocell networks are studied, where user's selfish characteristics are considered and an incentive mechanism is designed for subscribers to share the resource of femtocell vase stations.

In [20], a cooperative game framework was proposed for bandwidth allocation in 4G heterogeneous wireless access networks. The bandwidth allocation problem was formulated as a cooperative game (that is, bankruptcy game) and the solution (that is, the amount of bandwidth offered to a new connection) was obtained from Shapley value. This is different from a stochastic programming approach which maximizes the profit in stages. In a stochastic optimization environment, all stages have their decision variables. More specifically, the difference between cooperative and stochastic programming approaches lies in the fact that the former is group oriented, whereas the latter is scenario oriented. In a cooperative approach, groups of players seek fair resource allocation. On the other hand, in a stochastic optimizing approach, allocation is performed based on the scenario's profit gained from the resource.

In a heterogeneous wireless access environment, a vertical handoff mechanism needs to consider not only the radio link and the physical layer parameters but also the network and the transport layer parameters. In [21], a framework for vertical handoff was presented where the handoff decision metrics include service type, data rate requirement, network condition, and cost of handoff. A dynamic optimization was proposed to provide a QoS guarantee to the mobile users while maximizing the network utilization. In [22], a mobility management solution was proposed for heterogeneous wireless access networks to handle vertical handoff and network roaming. The solution was designed based on a formal policy representation model for a decision-making process for internetwork mobility. However, maximization of energy efficiency (from service providers' point of view) was not considered in these works.

The problem of integrating WLANs into the cellular wireless networks was investigated in the literature. In [23], a hierarchical radio resource management framework was designed to support seamless handoff between a WLAN and a cellular network. A hierarchical and distributed framework for seamless roaming across cellular networks and WLANs was proposed in [24]. The QoS mapping and Internet work message translation mechanisms were designed to support seamless handoff among multiple WLANs and cellular networks. However, a more general heterogeneous network architecture should be considered for radio resource management in 4G wireless networks.

3 Model for Heterogeneous Wireless Access and Energy Efficiency Optimization Framework

The network operator offers two types of wireless access technologies, namely, LTE and WiFi technologies. In particular, the mobile service provider sleep stations and reserve bandwidth in advance to be used by users in locations for a fixed period of time (e.g., a hour). The reserved channels are allocated to accommodate users in covered areas. The network capacity of a station is assumed to be a linear function of the number of bandwidth multiplied by its net utility coefficient.

We consider a service area composed of a set of locations. For example, in the service areas shown in Fig. 1, wireless access to WWAN and WLAN are available (e.g., LTE and IEEE 802.11). While WLAN access is available only in some locations, WWAN access is available in all locations of a service area. Remote and mobile users can connect to either a WLAN or the WWAN. The net utility function for a connection with a macro cell, micro cell and access point are denoted by $u_s, \forall s \in \mathbb{S}$. It is assumed that $u_{s_i} < u_{s_j} < u_{s_k}, \forall i \in \mathbb{S}_{ma}, j \in \mathbb{S}_{mi}, k \in \mathbb{S}_{hs}$

3.1 System Model

We consider a heterogeneous wireless access environment consisting of IEEE 802.11 WLAN, and LTE-Advanced radio interfaces, as shown in Fig. 1. A mobile with multiple radio transceivers (for example, software radio) is able to connect to these radio access networks.

We consider a geographic region that is entirely covered by a WMAN macro base station and partly covered by micro base stations and partly by WLAN access points (APs), as shown in Fig. 1. Users in the different service areas in this region have access to different types of wireless networks. In particular, in area 1, only WMAN service is available. In area 2 and area 4, services from cellular networks and WMAN are available. In area 3 and area 5, a mobile can connect to all three types of networks. Note that the energy efficiency optimization framework in this work can be applied to any other service area setting (different from the one shown in Fig. 1) in a considered geographic region.

Different wireless access networks are operated by the same service provider. We assume that a mobile is able to connect to each of the networks in the corresponding service area and perfect power control is assumed to ensure a uniform available transmission rate across the coverage area. In this heterogeneous wireless access network, we assume that multi-interface mobile terminals are able to connect to three different wireless access networks exclusively. These wireless access networks are IEEE 802.11 WLAN, LTE-Advanced WMAN operated by macro cells and operated by micro cells.

3.1.1 LTE-Based WMAN

We consider a LTE-based WMAN radio interface operating 10-66 GNz band that supports data rate in the range of 10-50 Mbps depending on the bandwidth of operation and the modulation and the coding scheme. The transmission frame is divided into subframes for downlink and uplink transmissions. Although time-division multiplexing (TDM) is used for downlink transmission, time-division multiple access (TDMA) is used for uplink transmission. These subframes are composed of transmission bursts (that is, uplink and downlink bursts), which carry MAC information and users' data. Each transmission burst corresponding to a particular mobile, which contains several MAC layer packet data units (PDUs), is separated from the other by a preamble field.

We consider two types of LTE-based WMAN base stations, macrocell and small cells. In the recent years, heterogeneous network deployment has emerged as a new trend to enhance the capacity/coverage and to save energy consumption for the next generation wireless networks. A heterogeneous network, or HetNet, is a wireless network consisting of nodes with different transmission powers and coverage sizes. High power nodes (HPNs) with large coverage areas are deployed in a planned way for blanket coverage of urban, suburban, or rural areas. Lowe power nodes (LPNs) with small coverage areas aim to compliment the HPNs for coverage extension and throughput enhancement. Furthermore, the infrastructure featuring a high density deployment of LPNs can also greatly improve energy efficiency compared to the one with a low density deployment of fewer HPNs, owing to the higher than linear path loss exponent in a wireless environment.

3.1.2 IEEE 802.11-Based WLAN

We consider an IEEE 802.11 WLAN radio interface with distributed coordination function (DCF)-based medium access control (MAC), which uses a carrier sense multiple access with collision avoidance (CSMA/CA) protocol. However, sine CSMA/CA is a contention-based MAC protocol, we adopt a distributed reservation-based MAC protocol, namely, early backoff announcement (EBA) [25], which is an enhanced version of DCF. EBA is also backeard compatible with IEEE 802.11 DCF. In EBA, by incorporating the backoff information in the MAC header, mobiles can completely avoid collisions.

3.2 Energy Efficiency Optimization Framework

The energy efficiency optimization framework is composed of two components: operation mode and capacity reservation for a station, service offering and bandwidth allocation for the different service areas. In bandwidth allocation, the available bandwidth from different access networks is assigned to the services areas so that service providers and users are satisfied with the allocation. Based on the capacity reservation, bandwidth allocation is used for service differentiation among services areas. In bandwidth allocation, the required amount of bandwidth is allocated to a population in a service area from the different available access networks.

In the system model under consideration, there is a service provider who offers wireless access services to the users. In particular, the service provider aims to maximize the system energy efficiency while allocating bandwidth to the service areas. Since the service provider has his own interest to maximize his payoff, the bandwidth allocation problem can be modeled as the first-stage problem. Therefore, for network-level capacity reservation, the first-stage variables and constraints are formulated and the solution is obtained from the optimum (which will be shown to maximize the total network utility as well).

On the other hand, the problem of bandwidth reservation among service areas can be formulated as a resource-sharing problem. In this case, a solution exists among the areas to achieve efficient sharing. This is practical since all of the users are in the same service area. Therefore, the second-stage formulation is used to obtain the allocation. In this stage, service areas are considered to obtain the allocation thresholds so that the connection-level QoS requirements (that is, connection blocking or dropping probabilities) for the different types of connections are satisfied. In this case, we consider the second-stage decision variables as the solution of the problem. Both network-level bandwidth allocation and capacity reservation are performed based on the number of users and traffic requirements in a service area that can be obtained from the network statistics with uncertainty.

4 Optimization of the Station Operation Problem in Heterogeneous Wireless Networks

We first formulate an optimization problem for the mobile service provider for station operation mode, capacity reservation, service offering and bandwidth allocation as a stochastic programming (SP) problem. The solution of this SP gives the operation mode and capacity at a particular station for a given operation time which maximize the expected energy efficiency of network performance under uncertainty of the number of mobile users and traffic requirements in different locations in a service area. A list of the key mathematical notations is shown in Table 1.

TABLE 1
Frequently Used Notations

$\mathbf{x}$	: station operation mode vector (1st-stage decision), $x_s \in \mathbf{x}, \forall s \in \mathbb{S}$
$\mathbb{S}$	: set of all stations, $\mathbb{S} = \{\mathbb{S}_{ma}, \mathbb{S}_{mi}, \mathbb{S}_{hs}\}$
$\mathbb{S}_{ma}$	: set of macro stations
$\mathbb{S}_{mi}$	: set of micro stations
$\mathbb{S}_{hs}$	: set of hot spot stations
$\mathbf{y}$	: station bandwidth reservation vector (1st-stage decision), $y_s \in \mathbf{y}, \forall s \in \mathbb{S}$
$\mathbf{z}$	: network selection matrix (2nd-stage decision), $z^w_{s,a} = \mathbf{z}_{s,a}(w), \forall w \in \mathbb{W}, s \in \mathbb{S}, a \in \mathbb{A}^{(s)}$
$\mathbb{L}^{(a)}$	: set of service coverage at $a, \forall a \in \mathbb{A}'$
$\mathbb{A}'$	: set of all service areas
$\mathcal{U}(s)$	: network utility function for station s, $u_s = \mathcal{U}(s), s \in \mathbb{S}$
$\hat{n}_{s,a}(w)$	: user number at service subarea a given scenario w, $n^w_{s,a} = \hat{n}_{s,a}(w), \forall w \in \mathbb{W}, s \in \mathbb{S}, a \in \mathbb{A}^{(s)}$
$\mathbb{W}$	: set of scenarios, $w \in \mathbb{W}$
$\mathbb{A}^{(s)}$	: set of subareas in coverage area $s \in \mathbb{S}$
$\mathbf{P}_{act}(s)$	: active mode energy consumption for station s, $P^s_{act} = \mathbf{P}_{act}(s), \forall s \in \mathbb{S}$
$\mathbf{P}_{sle}(s)$	: sleep mode energy consumption for station s, $P^s_{sle} = \mathbf{P}_{sle}(s), \forall s \in \mathbb{S}$
$\mathbb{D}^{(i)}$	: set of adjacent macro base stations of $i, \forall i \in \mathbb{S}_{ma}$
$\mathbb{M}^{(i)}$	: set of micro base stations of $i, \forall i \in \mathbb{S}_{ma}$
$\mathbf{v}$	: allocated bandwidth matrix (2nd-stage decision), $v^w_{s,a} = \mathbf{v}_{s,a}(w), \forall w \in \mathbb{W}, s \in \mathbb{S}, a \in \mathbb{A}^{(s)}$

4.1 Formulation of Stochastic Programming

A two-stage stochastic programming (SP) with recourse [26] is used to obtain the operation modes and operating bandwidth in WLAN and WWAN for running stations. In the first stage, the mobile service provider makes a decision on the operation modes of stations to reduce energy consumption by sleeping stations. Also, the service provider makes the other decision on the capacity reservation among stations to improve the operational energy consumption. The decisions in this stage has to be made without complete information on the random number of users and traffic requirements in each location sharing the allocated bandwidth from the stations. The operation modes to perform at station s for activating and sleeping is denoted by x_s; the number of bandwidth of stations to reserve at station s is denoted by y_s. In the second stage, the operator observes the actual number of users in a particular location (i.e., realization) and makes decisions to allocate bandwidth to users from stations if the stations are active and have channel resources available. The network to be associated with for population at a location for station is denoted by $z_{s,a}$; the number of bandwidth for allocation at a location to access a network is denoted by $v_{s,a}$. The decision in the second stage (i.e., $z_{s,a}$ and $v_{s,a}$) is referred to as a recourse, which is an action used to handle the uncertainty arising due to the mobility of the users.

4.1.1 Decision Variables

The decision variables of the station operation problem are $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ and $\mathbf{v}$. $\mathbf{x} = (x_s), \forall s \in \mathbb{S}$ is a operation mode vector and $\mathbf{y} = (y_s), \forall s \in \mathbb{S}$ is a bandwidth reservation vector; $\mathbf{z} = (z^w_{s,a}), \forall w \in \mathbb{W}, s \in \mathbb{S}, a \in \mathbb{A}^{(s)}$ is a network selection matrix, and $\mathbf{v} = (v^w_{s,a}), \forall w \in \mathbb{W}, s \in \mathbb{S}, a \in \mathbb{A}^{(s)}$ is a bandwidth allocation matrix. The value of a decision variable is decided as follows:

- $x_s = \begin{cases} 1, \text{station } s \text{ in active mode;} \\ 0, \text{otherwise.} \end{cases}$
- $y_s =$ amount of bandwidth to be reserved for s
- $z^w_{s,a} = \begin{cases} 1, \text{population at area } a \text{ is covered} \\ \text{by station } s \text{ in scenario } w; \\ 0, \text{otherwise.} \end{cases}$
- $v^w_{s,a} =$ amount of bandwidth to be allocated for area$a$ from station s in scenario w

4.1.2 Optimization of Energy Efficiency

To identify the economical objective, we first consider the revenue-to-cost ratio, serving as an indicator of the operator's profitability, analogous to the price-earnings ratio on the stock market. Since the way to work out the actual monetary values of the revenue and cost varies from case to case, it is transformed into a generalized objective function that maximizes the ratio of served traffic over energy consumed during the service operation. The goal is to attain a network design which is cost-effective, improving the data rate per resources managed (channel bandwidth and energy). As an efficient network is capable of growing in data volume successfully conveyed, within given cost in terms of energy consumption, the revenue then increases, coming from expanded data services offered. Although a multi-criteria approach explores the relationship between the two objectives, it is not looking to measure and maximize energy efficiency. With the optimal service rate per resources, the most profitable network design is able to be realized.

- **Traffic Service.** We define the network utility as the expected served traffic within areas covered by stations for scenarios, as the resource are utilized to meet traffic requirements respectively. The next expression presents the efficiency achieved as a whole.
$$\sum_{w \in \mathbb{W}, s \in \mathbb{S}, a \in \mathbb{A}^{(s)}} \mathbf{Pr}(w) \hat{r}^w_a z^w_{s,a}$$
- **Energy Consumption.** A station operates in either sleep mode, taking a relatively low energy level, or in active mode, consuming a basic running energy plus the energy for transmission over channel. The energy a sleeping BS consume is P^s_{sle} ; the energy consumed by an active BS comes from the basic active mode energy P^s_{act}, plus the expected energy exploited for using allocated channels to serve populations at different locations, $\sum_{w \in \mathbb{W}, s \in \mathbb{S}, a \in \mathbb{A}^{(s)}} \mathbf{Pr}(w) e_s v^w_{s,a}, \forall s \in \mathbb{S}_{ma} \bigcup \mathbb{S}_{mi}$. The overall energy consumption is expressed as:
$$\mathcal{P} = \sum_{s \in \mathbb{S}} [P^s_{act} x_s + P^s_{sle}(1 - x_s)] + \sum_{w \in \mathbb{W}, s \in \mathbb{S}, a \in \mathbb{A}^{(s)}} \mathbf{Pr}(w) e_s v^w_{s,a}$$

The objective defined in (1) is referred as the recourse function, which has two types of decision variables - the variables in the first stage for the energy consumption (i.e., $\sum_{s \in \mathbb{S}} [P^s_{act} x_s + P^s_{sle}(1 - x_s)]$) and the variables in the second stage due to allocating bandwidths and energy consumed for using channels (i.e., $\mathbf{z}_{s,a}(w)$ and $\mathbf{v}_{s,a}(w)$). The expected recourse value can be obtained from

$$E_{\mathbb{W}} \mathcal{Q}(\mathbf{x}, \mathbf{y}, w) = \frac{\sum_{w \in \mathbb{W}, s \in \mathbb{S}, a \in \mathbb{A}^{(s)}} Pr^w \hat{r}^w_a z^w_{s,a}}{\sum_{s \in \mathbb{S}} [P^s_{act} x_s + P^s_{sle}(1 - x_s)] + \sum_{w \in \mathbb{W}, s \in \mathbb{S}, a \in \mathbb{A}^{(s)}} Pr^w e_s v^w_{s,a}} \tag{1}$$

The SP problem can be formulated as follows:

$$\text{Maximize:} \quad E_{\mathbb{W}} \mathcal{Q}(\mathbf{x}, \mathbf{y}, w) \tag{2}$$

Subject to:

$$\sum_{a \in \mathbb{A}^{(s)}} \mathbf{v}_{s,a}(w) \le y_s, \forall s \in \mathbb{S}, w \in \mathbb{W} \tag{3}$$
$$\sum_{s \in \mathbb{S}, a \in \mathbb{A}^{(s)}} \hat{n}_a(w) \mathbf{z}_{s,a}(w) \ge \gamma \mathbf{N}(w), \forall w \in \mathbb{W} \tag{4}$$
$$\sum_{s \in \mathbb{L}^{(a)}} \mathbf{z}_{s,a}(w) \le 1, \forall w \in \mathbb{W}, a \in \mathbb{A}' \tag{5}$$
$$x_s \le y_s, \forall s \in \mathbb{S} \tag{6}$$
$$y_s \le \pi x_s, \forall s \in \mathbb{S} \tag{7}$$
$$\mathbf{z}_{s,a}(w) \le \mathbf{v}_{s,a}(w), \forall w \in \mathbb{W}, s \in \mathbb{S}, a \in \mathbb{A}^{(s)} \tag{8}$$
$$\pi \mathbf{z}_{s,a}(w) \ge \mathbf{v}_{s,a}(w), \forall w \in \mathbb{W}, s \in \mathbb{S}, a \in \mathbb{A}^{(s)} \tag{9}$$
$$y_i + \sum_{i' \in \mathbb{D}^{(i)}} y_{i'} \le B, \forall i \in \mathbb{S}_{ma} \tag{10}$$
$$y_k \le r_k, \forall k \in \mathbb{S}_{hs} \tag{11}$$
$$y_j + \sum_{j' \in \mathbb{D}^{(j)}} y_{j'} \le B - y_i, \forall i \in \mathbb{S}_{ma}, j \in \mathbb{M}^{(i)} \tag{12}$$
$$\sum_{a \in \mathbb{A}^{(s)}} \mathbf{z}_{s,a}(w) \hat{r}_a(w) \le u(s) y_s, \forall w, s, a \tag{13}$$
$$x_s, y_s \in \{0, 1\}, v^w_{s,a}, z^w_{s,a} \ge 0, \forall s, w, a \tag{14}$$

where $\mathbb{S}$ and $\mathbb{A}^{(s)}$ denote, respectively, the sets of stations and of subareas in coverage area s; $\mathbb{L}^{(a)}$, $\mathbb{D}^{(i)}$ and $\mathbb{M}^{(i)}$ denote the sets of stations users at subarea a is able to select, the set of adjacent macro base stations sharing the frequency band with micro base station s, and the set of bandwidth also available to micro base stations j within macro base station i. And w represents a network scenario. $E_{\mathbb{W}}$ denotes the expectation over all scenarios. In this case, the amount of user requirements $\hat{r}_a(w)$ associated with access networks at location a is a function of w, i.e, $\hat{r}_a(w) = \hat{r}^w_a$, where $\hat{r}^w_a$ is the number of traffic requirements from location a to access a network; e_s denotes the energy a macro or micro cell consumes per channel. The constraint in (3) shows that the bandwidth reservation for station s is fixed but the allocated bandwidth to different location a can be different depending on network scenarios. Constraint (4) is to ensure the ratio of blocked users in each scenario is not higher than the blocking criteria γ (the total user number in scenario w is $\mathbf{N}(w)$). (5) satisfies that the users in a service area can connect with no more than one access network at a time. The next two are used to ensure no bandwidth is reserved by a station in sleep mode and bandwidth reservation must be succussed for an active station (ξ is a large constant number greater than any value of decision variables). Similarly, constraints (8) and (9) are used to ensure that for the population within a service area, it must get services from the station (by channel bandwidth allocation) if it is associated with the station; otherwise both variables ($\mathbf{z}_{s,a}(w)$ and $\mathbf{v}_{s,a}(w)$) have a value of 0. In (10) the constraint is to restrict the channel bands available between a macro base station and its neighboring macro base station and are exclusively used to avoid inter-cell interference. Similarly, in the two next constraints, the available channel resource is limited to WLAN stations; a small cell site is not allowed to reuse a channel reserved to the macro base stations that cover it to prevent intra-cell interference and same bandwidth is not reused in the adjacent small cells (two cells have coverage overlap at any degree). The traffic requirement of an area is ensured to be met by the associated station's reserved channels.

Eq. (2) is the objective function to maximize the energy efficiency in the model. As a result of the energy required for sleep mode, $\mathcal{P}$ has a lower bound even though no associations would be formed by BSs with users. The case, therefore, will not arise that total energy consumption is very small (e.g., goes to zero). However, the problem involves a fractional objective function which cannot be considered as a linear programming model. We reformulate the problem into an ILP problem, so that a software package, such as CPLEX, is able to obtain the optimal solution.

4.1.3 Solving Optimization Problem with a Linear Fractional Objective Function

Eq. (2) has decision variables in both numerator and denominator, which is not the equation in the first degree. To eliminate the nonlinearity of the objective function, the model must be transformed to a model that is pure linear. Besides when the solution is found to this transformed model, the results can be recalculated back to the original model. In our problem the objective function is a ratio of two linear terms, and the constraints are linear. Since the value of the denominator is positive, the discussed method in [27] is applicable to our model. Eq. (2) can be transformed to a linear function, developing an integer linear programming model.

Considering the original problem, the denominator is positive over the entire feasible sets of $x^w_{s,a}$, $y^w_{s,a}$, $z^w_{s,a}$ and $v^w_{s,a}$. In order to realize the model transformation and eliminate the nonlinearity in the original model, variables α_s, β_s, $\theta^w_{s,a}$, $\delta^w_{s,a}$ and ξ are introduced and satisfy: $\alpha_s = \xi \times x_s$, $\beta_s = \xi \times y_s$, $\theta^w_{s,a} = \xi \times z^w_{s,a}$ and $\delta^w_{s,a} = \xi \times v^w_{s,a}$, where $\xi = \frac{1}{\mathcal{P}} > 0$. The objective function is now given by

$$
\text{Maximize:} \quad \sum_{w\in\mathbb{W},s\in\mathbb{S},a\in\mathbb{A}^{(s)}} Pr^w \hat{r}^w_a \theta^w_{s,a}
$$

Subject to:

$$
\begin{aligned}
&\sum_{s\in\mathbb{S}} [P^s_{act}\alpha_s + P^s_{sle}(1-\alpha_s)] \\
&+ \sum_{w\in\mathbb{W},s\in\mathbb{S},a\in\mathbb{A}^{(s)}} Pr^w e_s \delta^w_{s,a} = 1 \\
&\sum_{a\in\mathbb{A}^{(s)}} \delta^w_{s,a} \le \beta_s, \forall s\in\mathbb{S}, w\in\mathbb{W} \\
&\sum_{s\in\mathbb{S},a\in\mathbb{A}^{(s)}} n^w_{s,a}\theta^w_{s,a} \ge P_b \times N \times \xi, \forall w\in\mathbb{W} \\
&\sum_{s\in\mathbb{L}^{(a)}} \theta^w_{s,a} \le \xi, \forall w\in\mathbb{W}, a\in\mathbb{A}' \\
&\alpha_s \le \beta_s, \forall s\in\mathbb{S} \\
&\beta_s \le \pi\times\alpha_s, \forall s\in\mathbb{S} \\
&\theta^w_{s,a} \le \delta^w_{s,a}, \forall w\in\mathbb{W}, s\in\mathbb{S}, a\in\mathbb{A}^{(s)} \\
&\pi\times\theta^w_{s,a} \ge \delta^w_{s,a}, \forall w\in\mathbb{W}, s\in\mathbb{S}, a\in\mathbb{A}^{(s)} \\
&\beta_i + \sum_{i'\in\mathbb{D}^{(i)}} \beta_{i'} \le B\times\xi, \forall i\in\mathbb{S}_{ma} \\
&\beta_j + \sum_{j'\in\mathbb{D}^{(j)}} \beta_{j'} \le B - \beta_i, \forall i\in\mathbb{S}_{ma}, j\in\mathbb{M}^{(i)} \\
&\beta_k \le r_k\times\xi, \forall k\in\mathbb{S}_{hs} \\
&\alpha_s, \beta_s, \theta^w_{s,a}, \delta^w_{s,a}, \xi \ge 0, \forall w\in\mathbb{W}, s\in\mathbb{S}, a\in\mathbb{A}^{(s)}
\end{aligned}
\tag{15}
$$

Provided $\xi > 0$ at the optimal solution, this linear programming model is equivalent to the fractional objective problem stated previously except the binary condition of decision variables. The values of the variables $x^w_{s,a}$, $y^w_{s,a}$, $z^w_{s,a}$ and $v^w_{s,a}$ in the optimal solution to the fractional objective problem are obtained from dividing the optimal α_s, β_s, $\theta^w_{s,a}$ and $\delta^w_{s,a}$ by the optimal ξ.

As some decision variables in the fractional objective model are binary, it is necessary to impose on the transformed model the constraints which reflect the binary essence. We know that $\alpha_s = \xi \times x_s$ and $\theta^w_{s,a} = \xi \times z^w_{s,a}$, in which case if x_s is zero, then α_s must be zero; otherwise, $\alpha_s = \xi$ if $x_s = 1$. For this purpose, with the following constraints, introduced are the binary variables, λ_s and $\nu^w_{s,a}$, and a constant value π holding a value greater than λ_s, $\nu^w_{s,a}$ and ξ.

$$
\begin{aligned}
&\alpha_s - \xi - \pi\lambda_s \ge -\pi, \forall s\in\mathbb{S} \\
&\alpha_s - \xi + \pi\lambda_s \le \pi, \forall s\in\mathbb{S} \\
&\alpha_s \le \pi\lambda_s, \forall s\in\mathbb{S} \\
&\theta^w_{s,a} - \xi - \pi\nu_s \ge -\pi, \forall w\in\mathbb{W}, s\in\mathbb{S}, a\in\mathbb{A}^{(s)} \\
&\theta^w_{s,a} - \xi + \pi\nu_s \le \pi, \forall w\in\mathbb{W}, s\in\mathbb{S}, a\in\mathbb{A}^{(s)} \\
&\theta^w_{s,a} \le \pi\nu_s, \forall w\in\mathbb{W}, s\in\mathbb{S}, a\in\mathbb{A}^{(s)} \\
&\lambda_s, \nu^w_{s,a} \in \{0,1\}, \forall w\in\mathbb{W}, s\in\mathbb{S}, a\in\mathbb{A}^{(s)}
\end{aligned}
\tag{16}
$$

It is recognized in [27] that the problem of Eq. (2)-(14) are equivalent to that of Eq. (15)-(16). In other words, for example, if $\mathbf{A}$ is a solution to the problem of Eq. (15)-(16), then we can obtain a solution $\mathcal{X}$ to the problem of Eq. (2)-(14) from $x_s = \frac{\alpha_s}{\xi}, \forall s\in\mathbb{S}$ that follows from $\xi = \frac{1}{\mathcal{P}}$.

5 Scenario Reduction

Since the computational performance of the stochastic programming models is highly dependent on the size of the scenario set, a compromise between the necessary number

TABLE 2
Experiment Parameters

Parameter	Value
Channel Number	$B = 30$
Blocking ratio	7%
Utility for station	$u_s = 6.6, 9.2, 13.7$
Sleep mode power	$P^s_{sle} = 8, 0.96, 0.028$ Watts
Active mode power	$P^m_{act} = 500, 3.66, 2.79$ Watts
Energy per channel	$e_s = 2.02, 1.52, 0.95$ Watts
User number per location	910 - 2300
Require data rate per location	146.51 - 581.9

of scenarios and the computational burden of the associated stochastic programming model needs to be made, so that the problem can be solved using acceptable computational resources. For this purpose, appropriate scenario reduction techniques are usually applies.

A scenario reduction technique aims at reducing the size of the scenarios as much as possible, while at the same time the stochastic information enclosed in the original set is affected as less as possible. In other words, the optimal solution of the stochastic optimization using the reduced set of scenarios should remain close to the optimal solution obtained using the extended (original) set of scenarios. Various scenario reduction techniques have been reported in the literature so far [28], [29], [30], [31], [32], [33] and [34].

The scenario reduction methodology implemented is based on the concept of the probability distance. In general, the probability distance allows for quantifying how 'close' two different sets of scenarios representing the same stochastic process are. In this context, if a large scenario set is close enough to a reduced one in terms of the probability distance, the optimal solution of the simpler problem (which is formulated and solved using the reduced set of scenarios) is expected to be close to the optimal value of the original problem (which is formulated and solved with the extended set of scenarios).

Because the scenario tree size is usually too large for the optimization problem to be tractable, scenario reduction is applied. The objective of scenario reduction is to obtain a reduced set of scenarios that maintains the statistical properties of the original set to acceptable level. This reduced scenario tree is then used for optimization. The fast forward selection algorithm introduced in [30] is adopted in this work along with discrepancy metric. The algorithm is based on selecting a subset of scenarios from the original set such that the sum of minimum distances between all the members of the two sets is minimized.

6 Numerical Results

This section presents case studies where the wireless local area network and wireless wide area network (WWAN) are assumed such as LTE Advanced using OFDMA interface for the downlink. The case studies are conducted to evaluate the effectiveness of the mathematical model and of adopted scenario reduction in terms of obtained solutions, and the computational efficiencies in multi-station-multi-user scenarios. We firstly compare the scheme performance of maximizing energy efficiency , showing the significance of optimal policies of station operation mode and capacity reservation. The computation costs and solutions obtained by CPLEX for true problems and reduced instances are studied as well. The main system parameters taken into account in the simulations are tabulated in Table 2.

As macro WWAN spans a relatively large area, we separate areas of coverage into three geographic areas, urban (central), rural (peripheral) and suburban (in-between) regions. The area percentages are 65%, 15% and 20%, respectively. Three types of user distribution are considered across the entire area. Type 1 network layout has 80% of the users in the urban area, 12.5% suburban and 7.5% rural separately; Type 2 distribution is the peripheral region with a large population. The last type of user distribution is uniform, for which user numbers are basically the same in the three areas. The traffic requirement of a service area is set up to hold a relationship with the user number by $\kappa \hat{n}^w_a$, where κ is uniform distributed between 0.161 and 0.253.

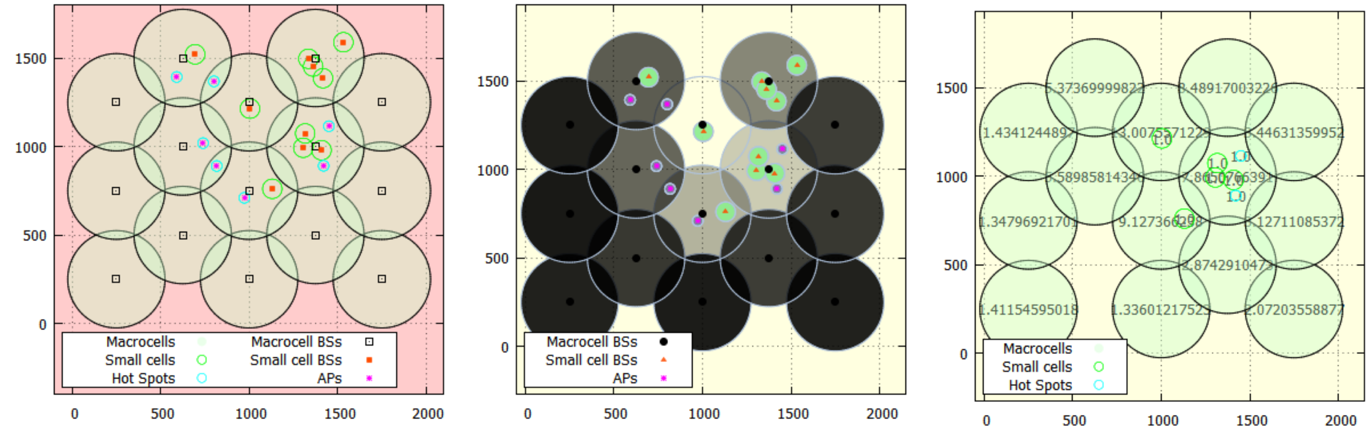
The frequency reuse 1 scheme causes interference for the users at cell boundaries [35]. In cellular networks, however, the groups of frequencies can be reused, provided the same frequencies are not reused in neighboring cells. As the frequency reuse three technique is supported in the WWAN [36], [37], the reuse factor of 3 is considered in our study to manage interference from other macro-cells, that is, three bandwidth parts are reused among cells.

6.1 Solving the Model

Fig. 2 displays the comparisons amongst cases with various user numbers from 910 to 2300 with Type 1 distribution, where there are 15 macro and 10 small base stations, 7 access points, and the blocking ratio is 7%. In Fig. 2(a), the small base stations and access points are deployed mainly in the urban area as in practical. The population within a service area (i.e., access point or a small cell) can be served by any station covering it (i.e., macro or small base stations.

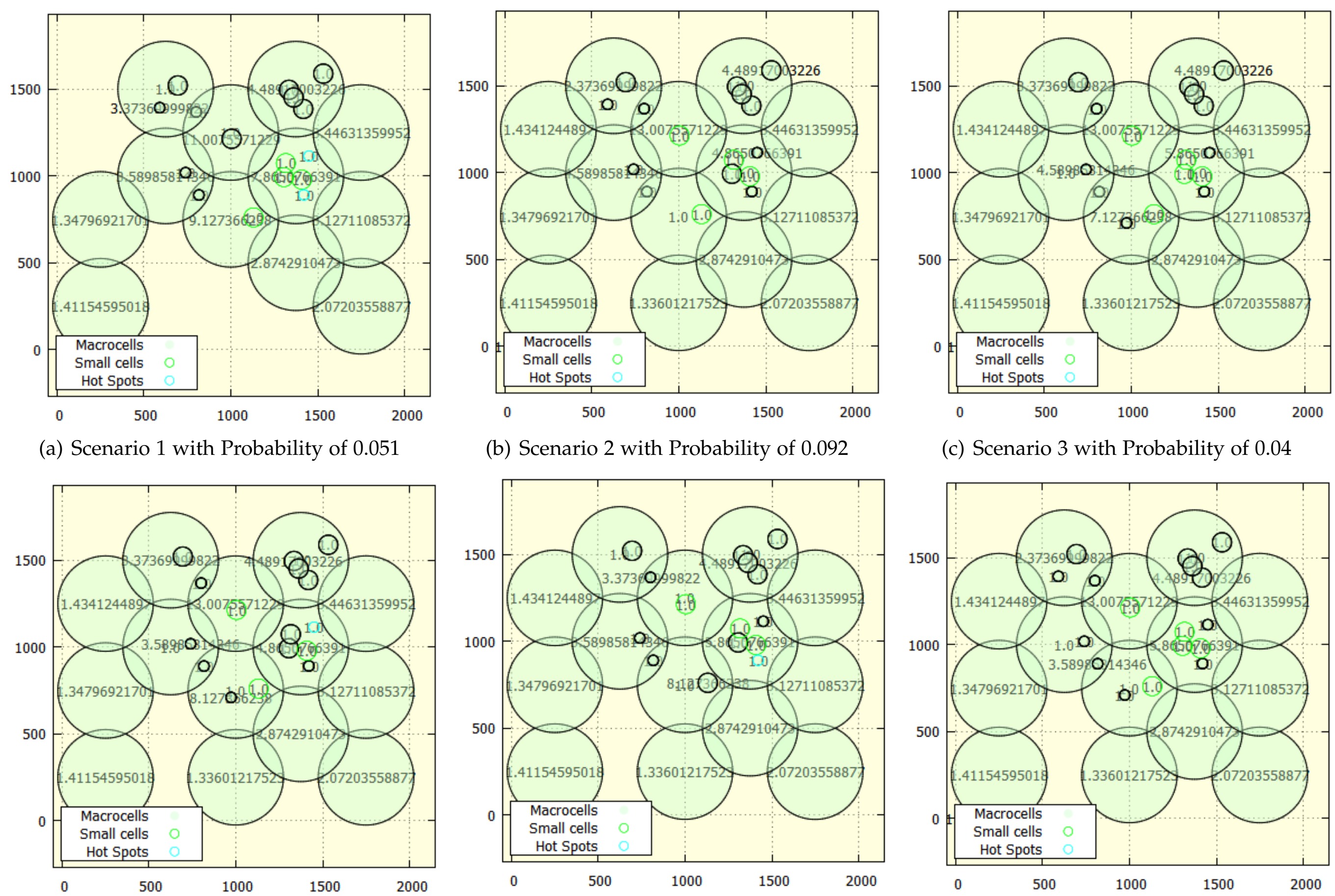
In Fig. 2(b) the user distribution is presented by the transparent level of a circle; the more transparent a circle is, the more user is with the service area relatively. For example, the macro base station has more users in the service area than any other stations. The optimal solution is displayed by 2(c), where several stations are switched off to save energy; 1 macro cell, 5 small cells and 5 access points are in sleep mode. The solution of capacity reservation is shown by the number in the center of each station. It can be verify that any two or three adjacent WWMN cells do not use more than the available licensed channel bandwidth, which is 30 in this work.

Fig. 3 shows bandwidth allocation performance levels in each scenario of the given instance. As shown in the solutions, different user numbers and traffic requirements of service areas result in different solutions; the allocated bandwidth varies according to scenarios and displayed by the numbers in the center of stations. Service areas are labeled by colors to represent the types of stations they associated with; those covered by macro cells are denoted with black circles, small cells with green and access points with blue. For example, for scenario 3 macro and small base stations are employed to cover service areas; on the other hand, the solution to the scenario 4, one access point is on for its service area.



(a) Exemplary Topology Setup (b) User Distribution (c) Optimal Solution

Fig. 2. An Instance Illustration



(a) Scenario 1 with Probability of 0.051 (b) Scenario 2 with Probability of 0.092 (c) Scenario 3 with Probability of 0.04

(d) Scenario 4 with Probability of 0.224 (e) Scenario 5 with probability of 0.189 (f) Scenario 6 with probability of 0.404

Fig. 3. Scenario Solutions for An Instance

6.2 Comparison of True Problem and Reduced Problem

The formulated model is solvable with CPLEX 12.5, which is a software package dedicated to finding the optimal solution to a MILP problem. The optimal results to true problems obtained by CPLEX are taken as the benchmarks to evaluate the solutions to reduced problems; the comparisons are measured on a machine equipped with 8 cores. To show the performance of the scenario reduction algorithm, we compare the sub-optimal solutions with the optimal ones. We find that the solutions from the scenario reduction algorithm are not quite consistent close to the optimum for all the considered cases Although, CPLEX solves the problem costly and solving reduced problems does save the computational cost. As the problem size goes up with larger numbers of scenarios and stations, CPLEX computation time rises significantly. On the other hand, in general the scenario

TABLE 3
Solution Comparison

Nodes N.			Scena. Num.	Gap		Time (Second)	
Ma	Sm	AP		w/o red.	w/ red.	Constructing	Solving
15	12	18	100	-	-	4062.66	3347.64
			40	2%	-	776.79	120.50
			20	48%	49%	194.83	40.28
12	7	12	20	-	-	39.40	3.86
			15	33%	-	24.19	21.20
			10	26%	5%	11.60	8.09
10	8	13	30	-	-	167.36	22.62
			20	8%	-	69.53	22.12
			15	4%	3%	44.74	13.22
9	9	14	40	-	-	172.16	12.48
			30	38%	-	96.32	44.20
			20	27%	1%	44.11	8.27
8	10	15	50	-	-	398.67	36.34
			40	33%	-	206.90	2512.60
			30	34%	1%	118.36	160.18

reduction method solves problems more efficiently; the ratio of the solving time of CPLEX to that of the algorithm may be considerable as the problem size continuously grows.

The performance of the fast forward selection approach is shown in Table. 3, where various instances and the reduced problems are taken into account. In the two columns presenting the gaps, two ratios are calculated and displayed. We first compute the solution differences between a true problems and its reduced problems where we consider the preserved scenarios and their redistributed probabilities. The other calculation is the solution differences between the two reduced problems in order to see whether the solutions are close between them. Regarding times spent on each problem, we list two metrics for comparison, the time for constructing a model by Python programming language and CPLEX APIs, and the time to merely solve the problem. It is indicated that the fast forward selection algorithm can barely provide solutions which are very close to the optimum, but the differences between reduced problems are mostly close. The algorithm does not really demonstrate computation efficiency as expected. The solving times of the fast forward method does not have tendency to be with significant improvement, which we could further verify by more simulation runs and more problems with larger sizes. Difference values between the true problems and reduced ones are not tolerable. The computation time of the adopted algorithm increases with the problem size, sharing the feature of an exponential growth with CPLEX. More scenarios are considered to assess the overall effectiveness of the algorithm. However, due to the general slowness of Python language and the server freezing issue happened in the past half week, as of the submission of the report, more results are being collecting.

7 Conclusion

The number of active macro cells influences largely energy efficiency since whenever one macro BS is activated, its energy consumption for being in active mode will effect reducing the energy efficiency ratio. Although the active mode energy consumption level of a small cell and a access point are relatively low, when the total number of these stations to consider is substantial, sleeping the maximum number of smaller cell sites and access points will save energy significantly, and so increase objective efficiency capacity. In this work, we have formulated the energy efficiency optimization of a HetNet into a two-stage stochastic programming problem, where the operation mode and capacity reservation are decided in the first stage, and given the solution in the previous stage and the realization of user numbers and traffic requirements in a service area, the service offering and bandwidth allocation are determined for each scenario. A fast forward selection algorithm is adopted to reduced the problem into given a problem with smaller specified scenario numbers. Based on the simulation results, we observe that the algorithm needs to further verification as the solutions obtained are not consistently close to the ones to true problems, although reduced problems show some close results between each other. A different probability distance may enhance the solution quality, or a modified heuristic scenario reduction method called forward selection in wait-and-see cluster (FSWC) is able to provide a more satisfactory performance.

7.1 Future Work

Three major perspectives we can work on to improve this work are as follows.

7.1.1 *Probability Distribution of Population in Different Locations*

Beside learning the statistical properties from history traffic logs, a probability distribution of the population and traffic requirement can be modeled mathematically. Considering Fig. 4, we assume that users can move among different locations with different availability of wireless access networks. $\mathbb{A}^{(s)}$ is the set of subareas in coverage area s. For example, in Fig. 1, a set can be defined as follows $A^{(2)} = \{2, 3\}$ whose elements represent the lower small base station, and the lower WLAN base station, respectively. At the station 3 and 5, a user can have access to both WLAN and WWAN, while on station 4, only WWAN connectivity is available. Given a set of locations in a service area, the mobility of users can be modeled by using Markov chain. The state of this Markov chain is the location which is observed at the end of a predefined time period. The transition from location l to l' of a number of users is indicated by probability $T_{l,l'}$, and the probability transition matrix is denoted by $\mathbf{T}$. The steady-state probability that a number of users will be in any location can be obtained by solving $\Delta\mathbf{T} = \Delta$ and $\Delta\mathbf{1} = 1$, where $\mathbf{1}$ is a vector of ones, Δ is a vector defined as follows: $\Delta = [\ \cdots\ \ \delta(l)\ \ \cdots\ ]$, in which $\delta(l)$ is the steady-state probability that the number of users will be at location l. Note that the coverage area of an access network may include multiple coverage areas of another access networks (Fig. 1). For the WWAN coverage area in Fig. 1, this set can be defined as $\mathbb{A}^{(s)} = \{1, 2, 3, 4, 5\}$ since area 1, 2, 3, 4 and 5 are in the coverage area of the WWAN base station.

7.1.2 *Radio Resource Management*

The work can be extended into a radio resource management by considering connection-level issues, for example, ongoing connections, new arrival connections and handoff connections. Considering Fig. 5, what this work covers is the

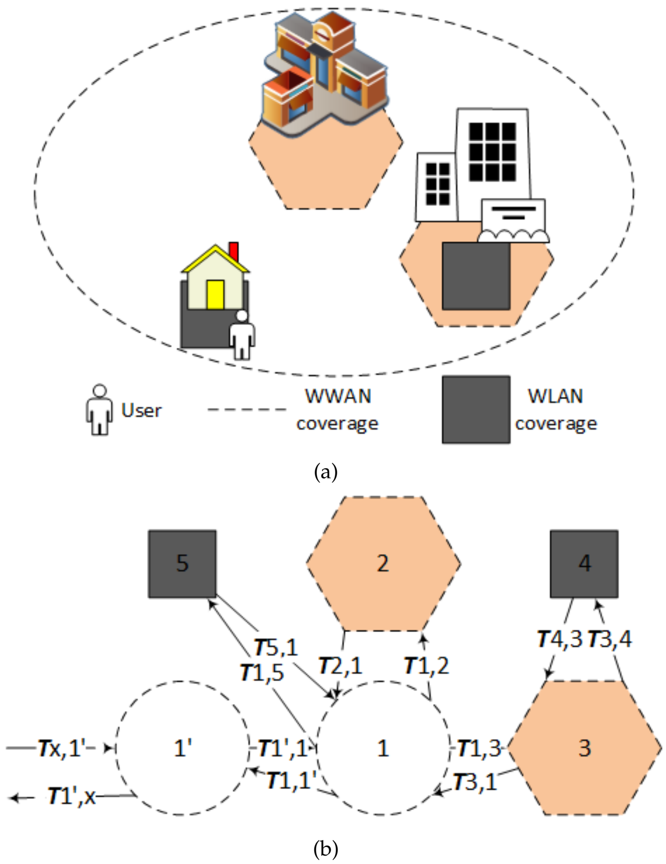


Fig. 4. (a) Sample service area and (b) the corresponding state transition diagram due to user mobility

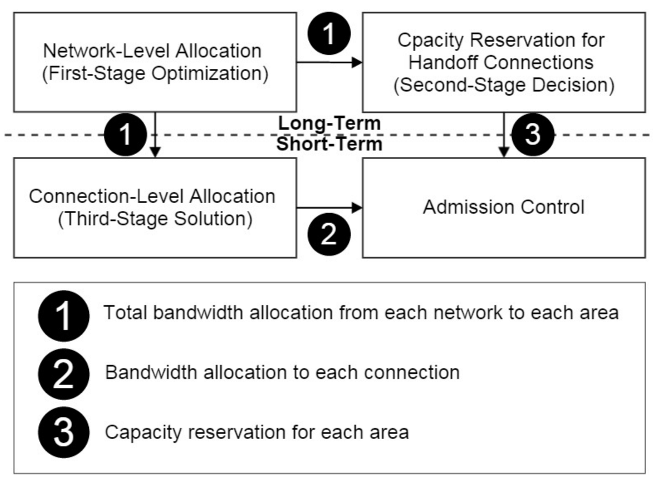


Fig. 5. Components of the proposed RRM framework

long-term capacity for populations in various service areas and bandwidth allocation to each area. The connection-level issues include capacity reservation for handoff connections and connection allocation to individual user. The problem then will involve admission control problem in the third stage of the new multistage stochastic formulation.

7.1.3 Scenario Generation

With the ability to solve nonlinear optimization problem, we can further adopt a scenario generation method into the current work, which will complete the research at more solid degree.

References

[1] Y. Fang and Y. Zhang, "Call admission control schemes and performance analysis in wireless mobile networks," *Vehicular Technology, IEEE Transactions on*, vol. 51, no. 2, pp. 371–382, Mar 2002.

[2] D. A. Levine, I. F. Akyildiz, and M. Naghshineh, "A resource estimation and call admission algorithm for wireless multimedia networks using the shadow cluster concept," *IEEE/ACM Trans. Netw.*, vol. 5, no. 1, pp. 1–12, Feb. 1997. [Online]. Available: http://dx.doi.org/10.1109/90.554717

[3] C.-T. Chou and K. G. Shin, "Analysis of adaptive bandwidth allocation in wireless networks with multilevel degradable quality of service," *IEEE Transactions on Mobile Computing*, vol. 3, no. 1, pp. 5–17, Jan. 2004. [Online]. Available: http://dx.doi.org/10.1109/TMC.2004.1261813

[4] C.-J. Ho, J. Copeland, C.-T. Lea, and G. Stuber, "On call admission control in ds/cdma cellular networks," *Vehicular Technology, IEEE Transactions on*, vol. 50, no. 6, pp. 1328–1343, Nov 2001.

[5] D. Cavalcanti, D. Agrawal, C. Cordeiro, B. Xie, and A. Kumar, "Issues in integrating cellular networks wlans, and manets: a futuristic heterogeneous wireless network," *Wireless Communications, IEEE*, vol. 12, no. 3, pp. 30–41, June 2005.

[6] O. Akan and I. Akyildiz, "Atl: an adaptive transport layer suite for next-generation wireless internet," *Selected Areas in Communications, IEEE Journal on*, vol. 22, no. 5, pp. 802–817, June 2004.

[7] Q. Song and A. Jamalipour, "A network selection mechanism for next generation networks," in *Communications, 2005. ICC 2005. 2005 IEEE International Conference on*, vol. 2, May 2005, pp. 1418–1422 Vol. 2.

[8] H. Lin, M. Chatterjee, S. K. Das, and K. Basu, "Arc: An integrated admission and rate control framework for competitive wireless cdma data networks using noncooperative games," *IEEE Trans. Mob. Comput.*, pp. 243–258, 2005.

[9] J. Virapanicharoen and W. Benjapolakul, "Fair-efficient threshold parameters selection in call admission control for cdma mobile multimedia communications using game theoretic framework," in *Consumer Communications and Networking Conference, 2005. CCNC. 2005 Second IEEE*, Jan 2005, pp. 439–444.

[10] S. Koskie and Z. Gajic, "A nash game algorithm for sir-based power control in 3g wireless cdma networks," *Networking, IEEE/ACM Transactions on*, vol. 13, no. 5, pp. 1017–1026, Oct 2005.

[11] T. Alpcan, T. Basar, and S. Dey, "A power control game based on outage probabilities for multicell wireless data networks," *Wireless Communications, IEEE Transactions on*, vol. 5, no. 4, pp. 890–899, April 2006.

[12] F. Meshkati, M. Chiang, H. Poor, and S. Schwartz, "A game-theoretic approach to energy-efficient power control in multicarrier cdma systems," *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 6, pp. 1115–1129, June 2006.

[13] D. Knisely, T. Yoshizawa, and F. Favichia, "Standardization of femtocells in 3gpp," *Communications Magazine, IEEE*, vol. 47, no. 9, pp. 68–75, September 2009.

[14] S. Guruacharya, D. Niyato, E. Hossain, and D. I. Kim, "Hierarchical competition in femtocell-based cellular networks," in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, Dec 2010, pp. 1–5.

[15] X. Kang, Y.-C. Liang, and H. K. Garg, "Distributed power control for spectrum-sharing femtocell networks using stackelberg game," in *Communications (ICC), 2011 IEEE International Conference on*, June 2011, pp. 1–5.

[16] S. Mclaughlin, P. M. Grant, J. Thompson, H. Haas, D. Laurenson, C. Khirallah, Y. Hou, and R. Wang, "Techniques for improving cellular radio base station energy efficiency," *Wireless Communications, IEEE*, vol. 18, no. 5, pp. 10–17, October 2011.

[17] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell networks: a survey," *Communications Magazine, IEEE*, vol. 46, no. 9, pp. 59–67, September 2008.

[18] I. Ashraf, L. Ho, and H. Claussen, "Improving energy efficiency of femtocell base stations via user activity detection," in *Wireless Communications and Networking Conference (WCNC), 2010 IEEE*, April 2010, pp. 1–5.

[19] C.-H. Ko and H.-Y. Wei, "On-demand resource-sharing mechanism design in two-tier ofdma femtocell networks," *Vehicular Technology, IEEE Transactions on*, vol. 60, no. 3, pp. 1059–1071, March 2011.

[20] D. Niyato and E. Hossain, "A cooperative game framework for bandwidth allocation in 4g heterogeneous wireless networks," in *Communications, 2006. ICC '06. IEEE International Conference on*, vol. 9, June 2006, pp. 4357–4362.

[21] J. McNair and F. Zhu, "Vertical handoffs in fourth-generation multinetwork environments," *Wireless Communications, IEEE*, vol. 11, no. 3, pp. 8–15, June 2004.

[22] P. Vidales, S. Member, J. Baliosian, J. Serrat, and F. Stajano, "Autonomic system for mobility support in 4g networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 2288–2304, 2005.

[23] G. Karetsos, S. Kyriazakos, E. Groustiotis, F. di Giandomenico, and I. Mura, "A hierarchical radio resource management framework for integrating wlans in cellular networking environments," *Wireless Communications, IEEE*, vol. 12, no. 6, pp. 11–17, Dec 2005.

[24] N. Shenoy and R. Montalvo, "A framework for seamless roaming across cellular and wireless local area networks," *Wireless Communications, IEEE*, vol. 12, no. 3, pp. 50–57, June 2005.

[25] J. Choi, J. Yoo, S. Choi, and C. Kim, "Eba: an enhancement of the ieee 802.11 dcf via distributed reservation," *Mobile Computing, IEEE Transactions on*, vol. 4, no. 4, pp. 378–390, July 2005.

[26] P. Kall and S. W. Wallace, *Stochastic Programming*. Chichester, England: John Wiley & Sons, 1994.

[27] O. Yemets, T. Barbolina, and O. Chernenko, "Solving optimization problems with linear-fractional objective functions and additional constraints on arrangements," *Cybernetics and Systems Analysis*, vol. 42, pp. 680–685, 2006.

[28] J. Dupaov, N. Grwe-Kuska, and W. Rmisch, "Scenario reduction in stochastic programming: An approach using probability metrics," 2003.

[29] H. Heitsch and W. Rmisch, "A note on scenario reduction for two-stage stochastic programs," 2007.

[30] ——, "Scenario reduction algorithms in stochastic programming," *Computational Optimization and Applications*, vol. 24, pp. 187–206, 2003.

[31] M. E. Fortet, R., "Convergence de la rpartition empirique vers la rpartition thorique," *Annales scientifiques de l'cole Normale Suprieure*, vol. 70, no. 3, pp. 267–285, 1953. [Online]. Available: http://eudml.org/doc/81682

[32] Y. Feng and S. M. Ryan, "Scenario construction and reduction applied to stochastic power generation expansion planning," *Comput. Oper. Res.*, vol. 40, no. 1, pp. 9–23, Jan. 2013. [Online]. Available: http://dx.doi.org/10.1016/j.cor.2012.05.005

[33] H. Heitsch and W. Rmisch, "Scenario tree modeling for multi-stage stochastic programs," in *Preprint 296, DFG Research Center Matheon, Berlin 2005, and submitted to Mathematical Programming. Contents Heitsch, H., Rmisch, W.: Stability and scenario trees for multistage stochastic programs, Preprint 324, DFG Research Center Matheon, Berlin 2006,*.

[34] J. Morales, S. Pineda, A. Conejo, and M. Carrion, "Scenario reduction for futures market trading in electricity markets," *Power Systems, IEEE Transactions on*, vol. 24, no. 2, pp. 878–888, May 2009.

[35] V. Pauli, J. D. Naranjo, and E. Seidel, "Heterogeneous lte networks and inter-cell interference coordination," *nomorde*, no. Icic, pp. 1–9, 2010.

[36] Y. Wang, S. Kumar, L. Garcia, K. Pedersen, I. Kovacs, S. Frattasi, N. Marchetti, and P. Mogensen, "Fixed frequency reuse for lte-advanced systems in local area scenarios," in *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*, April 2009, pp. 1 –5.

[37] M. Maqbool, M. Coupechoux, and P. Godlewski, "Comparison of various frequency reuse patterns for wimax networks with adaptive beamforming," in *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*, May 2008, pp. 2582 –2586.

Chan-Ching Hsu received the MS degree in computer engineering from Iowa State University. He is a PhD student in the Department of Electrical and Computer Engineering at Iowa State University. His research interests include multicell processing, green radio, optimization, cognitive and heterogeneous cellular systems, and (beyond) 4G systems.

J. Morris Chang received the PhD degree in computer engineering from North Carolina State University. He is an associate professor at Iowa State University. He received the University Excellence in Teaching Award at Illinois Institute of Technology in 1999. His research interests include wireless networks and computer systems. Currently, he is an editor of the Journal of Microprocessors and Microsystems, and of IEEE IT Professional.