Fault-Tolerant Small Cells Locations Planning in 4G/5G Heterogeneous Wireless Networks

Tamer Omar, Zakhia Abichar, Ahmed E. Kamal, J. Morris Chang, and Mohammad Abdullah Alnuem

Abstract—**Fourth/Fifth Generation heterogeneous wireless networks (4G/5G HetNets) use or will use small cells (SCs) to extend network coverage and increase spectrum efficiency. However, the standard and technical specifications do not specify how to plan the locations of the SCs within the network. Several papers introduced strategies for planning the locations of SCs in the 4G HetNet architecture. However, SCs placement strategies to support the self-healing functionality of the 4G/5G self organizing networks framework has not been studied in the literature. The placement of SCs in 4G HetNets such that an SC failure will not interrupt service, hence making the network fault tolerant, is an important design and planning problem that is addressed in this paper. We present an integer linear program formulation for planning operators of managed SC locations with fault tolerance. We allow one SC to fail and by using self-healing, a fault-tolerance service is provided at designated fail-over levels (defined in terms of users throughput). We consider the problem of SC location planning by using offloading in both out-band and in-band modes, and an interference model is presented to consider the in-band mode and to address the effect of interference on SCs placement planning. A novel approach to provide a linear interference model by using an expanded state space to get rid of nonlinearity is introduced. We present numerical results that show how our model can be used to plan the positions of SCs. We also incorporate the existence of obstacles in the planning, such as large structures or natural formations, that might happen in real life. To the best of our knowledge, this is the first work that addresses the planning of SC locations in 4G/5G HetNets in a fault-tolerant manner.**

Index Terms—**Fault tolerance, 4G HetNets, network architecture and design, self-healing, self organizing networks (SON), small cells.**

I. INTRODUCTION

AFTER the success and wide deployment of Wireless Local Area Networks (WLAN) [1], the area of wireless networks has witnessed the standardization process for broadband wireless access networks. The fourth generation (4G) broadband wireless network technologies (LTE and WiMAX) technical specifications provide last-mile connectivity, and they have been touted to fill several needs: last-mile end user access, initial deployment of infrastructure in unwired areas, and providing access to mobile users [2].

The main mobile service providers stations in a 4G HetNets are the base station (BS)/eNodeB (eNB) and the end users stations are called mobile stations (MS)/user equipment (UE) in WiMAX/LTE, respectively. In some areas, voice and data services are provided to end users via wireless networks instead of traditional wire-line infrastructure, which is time-consuming and costly to deploy. Both standards address the utilization of small cells (SCs) in 4G heterogeneous networks (HetNets). The goal of using SCs is to support the connectivity between the BS/eNB on one side and the MSs/UE on the other side. The SC can extend the range of a BS. For example, there could be users that are out of reach of the BS/eNB and cannot connect to the network. With the placement of an SC between the user and the BS/eNB, the user would be able to connect; hence, the range of the BS/eNB is extended. The SC can also be used to enhance the capacity of the BS/eNB. For example, even if all the users are in range within the BS/eNB, placing one or more SCs in the cell allows higher data rates and enhances the cell's capacity as a result. 4G HetNet technologies, however, does not specify how the SCs should be placed in the network.

The model presented in this paper allows for more than two hops communication. One of the advantages of this model is that it accommodates other networks, such as the use of relays in IEEE 802.16 m. There are also advantages in using more than two hops in LTE networks; that is, operator controlled SCs (e.g., Pico Cells) can piggyback on other SCs, hence acting as both SCs, and also relay stations, and therefore achieve some gains in terms of coverage and rate enhancement. It is the goal of this paper to devise a technique for planning the SC locations with fault tolerance to avoid failures in a 4G/5G HetNets.

A. Motivation

In this paper, we consider the operators problem of placing several managed SCs to support a BS/eNB to extend the coverage, improve the rate, and at the same time provide a resilient operation for HetNets. In real life, users expect a reliable service and many businesses rely on the Internet connection to be able to function. If the network is planned with no fault tolerance, a SC failure might result in disconnecting some users. There

---

Manuscript received September 27, 2015; revised February 3, 2016, June 10, 2016, and September 8, 2016; accepted September 24, 2016. Date of publication October 5, 2016; date of current version June 16, 2017. This work was supported in part by the National Plan for Science, Technology and Innovation (MAARIFAH) and in part by King Abdulaziz City for Science and Technology, Kingdom of Saudi Arabia under Award 11-INF1920-02. The review of this paper was coordinated by Prof. C. Assi.

T. Omar is with the Department of Technology Systems, East Carolina University, Greenville, NC 27858 USA (e-mail: omart15@ecu.edu).

Z. Abichar is with the Department of Electrical and Computer Engineering, University of Central Florida, FL 32816 USA (e-mail: zakhia17@ece.ucf.edu).

A. E. Kamal is with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA (e-mail: kamal@iastate.edu).

J. M. Chang is with the Department of Electrical Engineering, University of South Florida, Tampa, FL 33620 USA (e-mail: chang5@usf.edu).

M. A. Alnuem is with the Department of Information Systems, King Saud University, Riyadh 11451, Saudi Arabia (e-mail: malnuem@ksu.edu.sa).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2016.2615325

are different sources for SCs failures, and the different reasons of failures in HetNets can be classified into three categories as follows.

1) The first category is equipment malfunction, which may occur due to hardware or power outage.
2) The second category is link outage due to the failure of the SC back-haul that prevents the SC from relaying traffic to the core network.
3) Outage may occur due to the limited capacity and coverage capabilities of SCs that may become congested by overwhelming traffic from end users or channel impairments.

Due to its importance, the self-healing functionality has been introduced as one of the main functionalities of the self organizing networks (SON) [3]. Using the SON framework, self-healing procedures can be triggered to perform the proper remedy when fault tolerance planning is implemented to restore the service interrupted by any failure from the above three failure categories.

In the model of this paper, we consider the self-healing functionality in the case in which at most one SC might fail at a certain time. It could be any SC among the used SCs. With an adequate level of service, the SC should be repaired before another SC fails. This assumption (allowing only one SC to fail) also allows keeping the cost of the system reasonable. The proposed model is flexible to accommodate the number of assumed failing SCs, albeit with an increased capital expenditure (CAPEX).

To provide fault tolerance, we define for each user a full bit rate and a backup bit rate. When there is no failure among the SCs, users receive service at the full bit rate. However, in the case of a failure, we consider that offering service to affected users at a reduced rate is better than no service at all. Thus, in the case of an SC failure, the users receive service at the backup rate. This definition also allows users who primarily depend on the Internet for business to have a backup rate that can be made equal to the full rate. Thus, these users will function without service degradation even in the case of an SC failure.

The input to our problem is the location of the BS/eNB, the potential locations of the SCs, the location of the users (MSs/UE), and their respective demands represented by the bit rate. To reduce the problem complexity, groups of users are represented by traffic points (TP). For example, if there are several offices located close to each other with demands of 50, 100, and 150 Mb/s, they could be represented by a TP (located in a centric point to the offices) with a demand equal to the total MSs/UE demands of 300 Mb/s.

The planning solution we present in this paper aims at placing the SCs in the network to achieve several goals. The specific goals are as follows.

1) All the service area should be covered with connection to the network. The service area is defined through the TPs; thus, by providing connectivity between all the TPs and the BS/eNB, the service area will be covered.
2) The throughput demand of all the TPs should be satisfied. There should be a connection between the TP and the BS/eNB, with a flow equal to the predefined demand of the corresponding TP. The connections are assumed to be within the licensed carrier's spectrum.
3) The number of SCs placed by our solution should be minimized to reduce the equipment, installation, and operation cost, i.e., both CAPEX and OPEX.
4) In case an SC in the network fails, the network should continue to operate and provide service to the TPs at a predefined level of service, which we call the backup service rate. Thus, our planning method provides fault tolerance and resilience to single SC failures.

B. Contribution

Self-healing is the main functionality of the SON framework that provides fault-tolerant operation. Self-healing mechanism through cooperative clusters is proposed in [4] to deploy and manage the increasing number of small-cell networks. Resource utilization performance in both normal and failure modes of a small-cell network is evaluated and the authors show that their proposed mechanism outperforms other conventional mechanisms.

The study in [5] provided a relay station (RS) planning solution in WiMAX that satisfies only the first three goals listed above. There was no fault tolerance provision in the approach used. Thus, if an RS fails, there is no guarantee that the level of service provided would be adequate to the subscribers. In addition, the authors studied only WiMAX relaying without considering the applicable down-link channels interference as only out-band transmission was concerned. Hence, in this paper, we extend the approach to incorporate the in-band SCs planning with fault tolerance in a different network architecture by modeling the problem of in-band interference in HetNets. The new approach will ensure that users are served adequately in the case of a SC failure in 4G/5G HetNets.

There are several papers in the literature that address the problem of placing relay nodes in 4G networks. These are reviewed in the next section. However, to the best of our knowledge, there is no work on planning the locations of SCs in HetNets with fault tolerance. There have been some approaches on placing relays in a fault-tolerant manner in other types of networks, such as wireless sensor networks (WSN).

We formulate the SCs planning problem using a mixed integer linear program (MILP). We present numerical results by solving our model with CPLEX. We believe that solving the model directly to obtain results is a valid approach since planning is not a real-time operation. The problem is solved and the allocation is made using both out-band (no interference) and in-band (interference due to down-link resources sharing) transmission modes. To address the in-band mode, an interference model is introduced and the maximum link rates are calculated while taking the interference into consideration. Since the interference model results in a nonlinear formulation of the problem, we mapped the formulation to a binary linear formulation by expanding the state space, hence avoiding nonlinearities.

We present numerical results that show how our model finds the number and locations of SCs. Our model also specifies all the links that are used and gives the rate on each link. Also, for every SC that is used in the main topology (used when no SC is in a failure condition), the model gives a corresponding backup topology in case this SC fails.

The rest of this paper is organized as follows. Section II presents the related work and Section III presents the network model. The optimization model with fault tolerance is provided in Section IV for out-band mode and in Section V for in-band operational mode. Numerical results are given in Section VI and the conclusions are given in Section VII.

II. Related Work

This section presents earlier work in the literature that is related to the problem of planning the SC locations with fault tolerance.

a) Planning Locations in LTE

The in-band mode is addressed by authors in [6] as one of the strategies used in 5G HetNets to share the network resources between the eNB and the SCs. This study shows the importance of creating an interference model like the one proposed in this paper to address this issue in HetNets by using in-band strategy. Eguizabal and Hernandez [7] proposed an in-band strategy to multiplex traffic to several relay nodes in LTE. Interference coordination is proposed to increase coverage and improve capacity. However, fault tolerance and self-healing was not discussed by the authors. An in-band mode performance evaluation for different deployments is investigated by the authors in [8]. Different scenarios are presented to show the performance of relaying in LTE networks, and results show that relaying is strongly affected by the back-haul. However, fault tolerance is not addressed in the study to show the effect of failure on network planning.

A clustering algorithm based on uniform cluster concepts is proposed in [9] to select the BS and RS locations from candidate positions, depending on the traffic demands. The authors introduce a scheme that makes adaptive decision for selecting the deployment sites of the BS and RS. Simulation results show that the scheme achieves good performance in terms of network throughput and coverage. Lin and Ho present another RS placement solution in [10], wherein the cooperative transmission paradigm is used in multihop relaying for the purpose of range extension. Also, Lin *et al.* [11], [12] presents an RS placement solution that uses the cooperative transmission technique for the purpose of capacity enhancement.

b) Planning Locations in WiMAX

Our previous work in [5] presents a model for planning the RS locations in a WiMAX network. However, in the previous work, there was no guarantee of service if an RS fails since fault tolerance was not considered. In this paper, we extend our model to provide resilience to relay failures.

A planning model is presented in [13] to find the locations of BSs and RSs in the network. The model is formulated as an optimization problem using integer programming. In this model, there is at most one RS between the SS/MS and the BS, and a maximum of two hops is allowed. Since the standard does not have a limit on the number of hops going through the RSs, this assumption may impose unnecessary restrictions.

Yu *et al.* present an extension of their work in [14]. In this paper, they consider a large coverage area that increases the computation time of the model. To reduce the computation time, they divide the area into clusters and apply the approach above to every cluster. Then, the cases on the boundaries of the clusters are solved to find the overall solution. This paper similarly limits the number of hops to two.

In [15], a model is presented to find the locations of RSs that extend the range of a BS in a WiMAX network. This work defines preset topologies and finds the RS placement for these topologies; in comparison, our model in this paper and in [5] can work with any topology. This work also considers RS location planning for sector-based topology. Each sector uses a frequency that is different from adjacent sectors to reduce interference.

In [16], an RS placement model is presented. This work is based on cooperative transmission between the source node and the relay node to provide a better signal to the destination node. They consider the decode-and-forward scheme and the compress-and-forward scheme for cooperative transmission. This model is different from our work since it considers the placement of a single RS to serve multiple MSs.

In [17], the problem of joint BS and RS deployment is considered and an optimization model is presented. Due to the large size of the problem, the model takes a long time to solve. Thus, the authors also present an efficient heuristic algorithm to find the problem suboptimal solution.

In [18], the problem of RS placement in the WiMAX network is considered. The location of the RSs and the bandwidth allocation to users are found. This work assumes that users' demands could change due to fluctuations in traffic demands and due to mobility. Thus, the optimization of the RS locations is found on a long-term basis and the bandwidth allocation to users is found on a short-term basis.

Chang *et al.* [19] consider using relays for the purpose of capacity enhancements as follows. There is a BS, an area that can be totally covered by the BS, and a given number of relays. This work decides where to place the relays to maximize the system capacity.

In [20], the following paradigm is considered for the placement of RSs in WiMAX networks. The number and locations of BSs are given. The goal of the problem is to place RSs that use the transparent mode. In this mode, the RSs do not transmit control information; the control information are only transmitted by the BS. The RSs are thus in range of the BS and the goal of the RS placement is capacity enhancement. Other approaches used in BSs and RSs placement are presented in [21], [22], with the goal of enhancing the overall network capacity.

c) Planning Locations in WSN

There are approaches in the literature that provide relay location planning with fault tolerance. But these approaches have been designed for WSNs and not for WiMAX networks.

In [23], an Integer Linear Program model is presented for placing relays in sensor networks to provide fault tolerance in case some nodes fail. The main issue was connectivity, regardless of bandwidth requirements, which implies that all relay nodes may be operational all the time. Bari *et al.* present an extension of their work in [24], which takes into consideration the routing strategy to reduce battery consumption. Other approaches on fault-tolerant placement of relay nodes are given in [25]–[32].

The approaches presented in this section consider a multitude of issues and configurations for traffic relaying used in broadband and WSNs. There are multiple approaches for planning the placement of nodes that is discussed in these networks, some of them, specifically in WSNs, address the planning of relay nodes locations with fault tolerance. However, none of this work is applicable to broadband wireless HetNets. To the best of our knowledge, there is no approach that provides the planning of SCs locations operating within the same frequency band (in-band mode) in HetNets with fault-tolerance. Moreover, neither the approaches introduced for WSNs that discussed fault tolerance nor our previous work in [5] considered the interference caused due to the in-band transmission mode. Using the SON framework, the proposed approach can be implemented as a self-healing functionality to compensate for HetNet failures. However, the offered service in case of failures is downgraded to the backup rates. Hence, our paper is the first to propose such solutions that map from nonlinear to linear interference using state space transformation as an approach to perform HetNet SCs recovery in an in-band transmission mode and to guarantee the business continuity even in case of partial failures by using the proposed fault tolerance planning.

III. Network Model

This section presents the network model that we consider in this paper.

A. Small Cells Offloading Modes

The 4G/5G HetNets defines two modes of macrocells to SCs offloading operation modes: a transparent mode and a nontransparent mode. In the transparent mode, the users (MSs/UE) are unaware of the presence of a SC. The SC does not transmit control information (such as down-link map and up-link map). These are transmitted by the BS/eNB. Thus, all the MSs/UE are within range of the BS. However, the SCs are used in the transparent mode for the purpose of capacity enhancement.

In the nontransparent mode, the SCs perform all functions needed for a standalone cell and transmits control information as well as data to the MSs/UE it serves. Multi-hop routes are allowed in the nontransparent mode. The goal for using nontransparent SCs is to extend the range of the network and to also enhance the capacity. Currently this mode is widely used for SCs deployment and is the mode considered in our study.

B. Duplexing Mode

When SCs are used, transmissions from two stations that are in range should be duplexed either in the frequency domain (FDD) or in the time domain (TDD) to avoid interference. The 4G standards allow the use of different frequencies for SCs serving the same BS. Thus, we make the assumption that the SCs duplex their transmission using the frequency division duplex (FDD) mode. For example, on a two-hop route eNB-SC-UE, we can have a transmission of rate r on the eNB-SC hop and another transmission of the same rate on the SC-UE hop. This happens if the two hops are using different frequencies. With time division duplex (TDD), the two hops will alternate in transmission using the same frequency channel. However, each hop will have a larger bandwidth since the bandwidth is not divided anymore. We use FDD for simplicity, but our model is logically equivalent to TDD. For more generalization of the studied problem, we also assumed the utilization of nonorthogonal physical layer multiplexing approaches (e.g, CDMA, FDMA). The proposed model can benefit from adopting frequency partitioning and reuse techniques whereby the same channels can be reallocated to different small cells if they are geographically distributed so that the intercell interference between them does not negatively impact their transmission rates.

TABLE I
OFDMA Rates (in Mbps) for Various Modulation Schemes Using 7 mhz Bandwidth

QPSK	QPSK	16-QAM	16-QAM	64-QAM	64-QAM
$1/2$	$3/4$	$1/2$	$3/4$	$2/3$	$3/4$
5.82	8.73	11.64	17.45	23.27	26.18

C. Link Capacity

Our model allocates a rate on each link that is used in the produced topology. The allocated rate on a link is bounded by the maximum capacity of the link. The maximum capacity of a wireless link can be modeled with the Shannon–Hartley equation as given in [33]. It is given by the equation: $\mathsf{C} = \mathsf{B}.\log_2(1 + \text{SINR})$, where C is the capacity in bit/s, B is the channel bandwidth in Hz.

The signal to interference plus noise ratio can be calculated as $\text{SINR} = S/\left[N_0 + I\right]$, where S is the received signal power, N_0 is the noise power, and I is the signal power received from all interferers, j. The capacity changes with the distance since the SINR degrades when the distance increases. The SINR can be expressed as

$$\text{SINR}_i = \frac{\beta p_i}{(d)^{\alpha}\left[N_0 + \sum_{j \neq i} p_j \Big/ (d)^{\alpha}\right]} \tag{1}$$

where p_i is the signal transmission power, d is the Euclidean distance between the transmitter and receiver, $\alpha > 2$ is the path loss exponent, and β is the antenna gain.

Other factors also affect the link capacity, such as the coding and modulation schemes. When a high SINR is measured on the link, coding, and modulation schemes with high rates are used. However, when the SINR is low, robust coding and modulation schemes are preferred to limit the bit error rate (BER), although they provide low data rates. Table I shows the achievable bit rates for the Orthogonal Frequency-Division Multiple Access (OFDMA) physical layer as given in the standard [34]. Quadrature Phase Shift Keying (QPSK) is more robust but achieves a small rate. On the other hand, 64-QAM is less robust but achieves a high rate.

The factors that affect a link's capacity can be combined in an equation. For any link i, the maximum rate is: $m_i = \Gamma(\text{SINR}_i,\ \chi,\ \text{Cod},\ \text{Mod})$, where χ is the upper-bound on the

TABLE II
ABBREVIATIONS

Parameters	Description
R	Set of candidate sites for SC
T	Set of TPs that represent the user traffic
B	Base Station
L_i^x	Set of links that can interfere with link i
F_n^x	Sets n of active interference links to link i
$r_i,\ rb_i$	Rate requirement, backup rate requirement
k	Backup topology number
S	Total number of SCs
T	Total number of SCs in backup topology
Variables	Description
dR, dBR, dBTdRT, dRR	Decision variables
fBR, fBT, fRT, fRR	Traffic flow
nBR, nBT, nRT, nRR	Number of subcarriers
mBR, mBT, mRT, mRR	Maximum link capacity
cBR, cBT, cRT, cRR	Actual link capacity
dnBR, dnBT, dnRT, dnRR, Ix_j, Yx_{Fxn}	Binary variables
vBR, vBT, vRT, vRR wBR, wBT, wRT, wRR	Variables
$X_i, Z_{\mathrm{ij}}, X_i^k, Z_{\mathrm{ij}}^k$	Auxiliary variable

BER, Cod is the coding scheme, and Mod is the modulation scheme. Γ is the function that maps all the three parameters to the maximum rate.

Any definition of the function Γ can work with our model. However, for simplicity, we assume that the maximum rate changes with distance. In real-life scenarios, there is usually a field survey that precedes the network deployment [35]–[37]. The link rates are selected based on the links characteristics, such as the SINR, fading, the specifics of the terrain, and interference with other wireless systems.

D. Definition of Fault Tolerance

The planning model we present in this paper allows the failure of an SC without interrupting service to the users, albeit at a reduced bit rate, hence tolerating equipment failure.

We assume that only one SC will fail at a given time. This is a reasonable assumption since usually in the time it takes the SC to be repaired, there is a very small probability that another SC will fail. This is true since the number of operator supported SCs (e.g, Pico Cells) supporting a BS/eNB will typically be a small number of SCs. This assumption will keep the cost of SCs small, since tolerating the failure of two or more SCs at the same time requires installing many extra SCs, which is not a cost-effective approach.

For every set of customers represented by a TP, a tuple $\{r_i,\ rb_i\}$ defines the requested service rates. When all the SCs are operational, the full rate for a TP_i, given by r_i is provided. However, when there is an SC failure, a reduced rate that is the backup rate rb_i, is provided, with $rb_i \leq r_i$. Users who request the same service rate, even in the case of an SC failure, will have $rb_i = r_i$.

IV. OPTIMIZATION MODEL: THE OUT-BAND MODE

This section presents the optimization model for the SC planning problem with fault tolerance in the out-band mode. The model takes the following as input.

TABLE III
LINK RATES OUT-BAND MODE

Distance (unit)	Link Rate (Mb/s)
if distance $<= 1$	rate $= 10$
else if distance $<= 2$	rate $= 5$
else if distance $<= 3$	rate $= 2$
else if distance $<= 4$	rate $= 1$
else	rate $= 0$

1) the possible sites where an SC can be installed;
2) the locations of the TP that represent the users' traffic;
3) the rates (full and reduced) in Mb/s of each TP. The full rate is provided when all the SCs are operational, and the reduced rate is guaranteed when there is an SC failure; and
4) the model takes as an input the maximum rate on any link: eNB-SC, SC-TP, SC-SC, and SC-TP, which depends on the link characteristics such as distance, SINR, and bandwidth.

Table III helps to clarify the notations in this paper.

The output of our model is the full-rate (main) topology and the reduced-rate (backup) topologies. The full-rate topology is defined by the number of SCs used, their positions, the links used, the rate on each link, and finally, the connection node for each TP (either the eNB or an SC). Each of the backup topologies corresponds to a failure in one of the SCs used in the main topology. For example, if the main topology uses SC_1, SC_3, and SC_8, then there will be three backup topologies that are used in case any of these SCs fails.

For any TP (TP_i), the full rate is designated by r_i and the reduced backup rate is designated by rb_i, which is the minimum acceptable rate in the case of failure.

Let $R = \{SC_0, \ldots, SC_{N-1}\}$ be the set of candidate sites for SC with cardinality $|R| = N$. Similarly, let $T = \{TP_0, \ldots, TP_{M-1}\}$ be the set of TPs that represent the user traffic with cardinality $|T| = M$.

A. Decision Variables

The following decision variables define the full-rate topology

$$dR_i = \begin{cases} 1; & \text{a SC is deployed in site } SC_i \\ 0; & \text{otherwise } (i \in\ R) \end{cases}$$

$$\mathrm{dBR}_i = \begin{cases} 1; & \text{a link is used between the eNB and } SC_i \\ 0; & \text{otherwise} (i \in\ R) \end{cases}$$

$$\mathrm{dBT}_i = \begin{cases} 1; & TP_i \text{is assigned to the eNB} \\ 0; & \text{otherwise } (i \in\ T) \end{cases}$$

$$\mathrm{dRR}_{\mathrm{ij}} = \begin{cases} 1; & \text{a link is used between the } SC_i \text{and } SC_j \\ 0; & \text{otherwise } (i, j \in\ R) \end{cases}$$

$$\mathrm{dRT}_{\mathrm{ij}} = \begin{cases} 1; & \mathrm{T}P_j \text{ is assigned to the } SC_i \\ 0; & \text{otherwise } (i \in\ R, j \in\ T). \end{cases}$$

We also define variables that are similar to the above to specify the backup topologies. These variables are: $\mathrm{dR}_i^k, \mathrm{dBR}_i^k, \mathrm{dBT}_i^k, \mathrm{dRR}_{\mathrm{ij}}^k$, and $\mathrm{dRT}_{\mathrm{ij}}^k$. The term k indicates the backup topology number used when SC_i^k has failed. For example, when $k = 3$, these variables define the backup topology that is used when SC^3 fails.

We also define decision variables that designate the assigned flow (in Mb/s) on each link. Although the previous variables were binary, the flow variables take continuous values. In the full-rate topology, the variables fBR_i and $\mathrm{fRR}_{\mathrm{ij}}$ designate the flow on the links from eNB to SC_i and SC_i to SC_j, respectively, where i and j are indexes of SCs $(i,\ j\ \in\ R)$.

Similarly, in the backup topology, the variables fBR_i^k and $\mathrm{fRR}_{\mathrm{ij}}^k$ designate the flow on the links from eNB to SC_i^k and SC_i^k to SC_j^k, respectively, where $i,\ j$ are indexes of SCs, k is the index of the backup topology when SC^k fails, and $(i,\ j,\ k\ \in\ R)$.

B. Topology Constraints

The following constraints define the topology of the SCs domain. They ensure that when a link is used in the solution, the two end nodes of the link exist (i.e., the SCs are selected). They also ensure that a TP is connected either directly to the eNB or to only one SC; we use this condition to not add complexities to the UEs.

First, when there is a link between the eNB and SC_i, there should be a SC deployed at site SC_i. This is ensured by the following constraints in the full-rate and the backup topologies

$$\mathrm{dBR}_i \le dR_i \quad \forall\, i\ \in\ R \tag{2}$$

$$\mathrm{dBR}_i^k \le dR_i^k\ \ i \ne k \quad \forall\, i,\ k\ \in\ R. \tag{3}$$

When there is a link between SC_i and SC_j, two SCs should be installed at sites SC_i and SC_j. This is ensured by the following constraints:

$$\mathrm{dRR}_{\mathrm{ij}} \le \frac{dR_i + dR_j}{2} \quad \forall\, i,\ j\ \in\ R \tag{4}$$

$$\mathrm{dRR}_{\mathrm{ij}}^k \le \frac{dR_i^k + dR_j^k}{2}\ \ i \ne k,\ j \ne k \quad \forall\, i,\ j,\ k\ \in\ R. \tag{5}$$

When there is a link between SC_i and TP_j, a SC should be deployed at site SC_i. This is ensured by the following constraints:

$$\mathrm{dRT}_{\mathrm{ij}} \le dR_i \quad \forall\, i\ \in\ R \quad \forall\, j\ \in\ T \tag{6}$$

$$\mathrm{dRT}_{\mathrm{ij}}^k \le dR_i^k\ \ i \ne k,\ j \ne k \quad \forall\, i,\ k\ \in\ R \quad \forall\, j\ \in\ T. \tag{7}$$

The following constraints send all the traffic of a TP either through a direct link with the eNB or through a single SC:

$$\mathrm{dBT}_i + \sum_{j\,\in\,R} \mathrm{dRT}_{\mathrm{ji}} = 1 \quad \forall\, i\ \in\ T \tag{8}$$

$$\mathrm{dBT}_i^k + \sum_{j\,\in\,R,\ j \ne k} \mathrm{dRT}_{\mathrm{ji}}^k = 1 \quad \forall\, i\ \in\ T. \tag{9}$$

C. Flow Constraints

The flow constraints ensure that the amount of data that is transported is balanced and sufficient for the demands of all the TPs.

1) Flow Balance at the BS: In the main topology, the total traffic going out of the eNB should be equal to the sum of the full rates r_i of all the TPs. This condition is ensured by the following equation:

$$\sum_{i\,\in\,R} \mathrm{fBR}_i \cdot \mathrm{dBR}_i + \sum_{j\,\in\,T,\ \mathrm{mBT}_j \ge r_j} r_j \cdot \mathrm{dBT}_j = \sum_{j\,\in\,T} r_j \tag{10}$$

where (mBT_j, mRT_j) are the upper bounds of the rates on the links for the main topology which are input parameters to the problem and are calculated in Section V-A

At a backup topology, the rate provided to TP_i is greater than or equal to rb_i. Then, this condition is used

$$\sum_{i\,\in\,R,\ i \ne k} \mathrm{fBR}_i^k \cdot \mathrm{dBR}_i^k + \sum_{j\,\in\,T,\ \mathrm{mBT}_j \ge \mathrm{rb}_j} \mathrm{rb}_j \cdot \mathrm{dBT}_j^k = \sum_{j\,\in\,T} \mathrm{rb}_j. \tag{11}$$

We are interested in keeping the system linear. Thus, we use the following transformation and substitute in (10):

$$X_i = \mathrm{fBR}_i \cdot \mathrm{dBR}_i$$

where X_i is an auxiliary variable.

Equation (10) therefore becomes

$$\sum_{i\,\in\,R} X_i + \sum_{j\,\in\,T,\ \mathrm{mBT}_j \ge r_j} r_j \cdot \mathrm{dBT}_j = \sum_{j\,\in\,T} r_j. \tag{12}$$

X_i can be evaluated using the following set of linear constraints, where Q is a large number such that $Q > \max(\mathrm{fBR}_i)\ \forall\, i \,\epsilon\, R$:

$$X_i \ge Q \cdot \mathrm{dBR}_i - Q + \mathrm{fBR}_i \quad \forall\, i\ \in\ R \tag{13}$$

$$X_i \le \mathrm{fBR}_i \quad \forall\, i\ \in\ R \tag{14}$$

$$X_i \ge 0 \quad \forall\, i\ \in\ R \tag{15}$$

$$X_i \le Q \cdot \mathrm{dBR}_i \quad \forall\, i\ \in\ R. \tag{16}$$

Similarly, we use the following transformation for (11):

$$X_i^k = \mathrm{fBR}_i^k \cdot \mathrm{dBR}_i^k\ \ i \ne k \quad \forall\, i,\ k\ \in\ R. \tag{17}$$

Hence, (11) becomes

$$\sum_{i\,\in\,R,\ i \ne k} X_i^k + \sum_{j\,\in\,T,\ \mathrm{mBT}_j \ge \mathrm{rb}_j} \mathrm{rb}_j \cdot \mathrm{dBT}_j^k = \sum_{j\,\in\,T} \mathrm{rb}_j. \tag{18}$$

X_i^k is evaluated like X_i was evaluated in (13)–(16).

2) Flow Balance at a SC: At any SC, the amount of traffic that is coming from the eNB and from upstream SCs is equal to the amount of traffic that is going to downstream SCs and to TPs that are directly connected to the SC. This is ensured by the following constraint:

$$\mathrm{fBR}_i \cdot \mathrm{dBR}_i + \sum_{j\,\in\,R} \mathrm{fRR}_{\mathrm{ji}} \cdot \mathrm{dRR}_{\mathrm{ji}} = \sum_{j\,\in\,R} \mathrm{fRR}_{\mathrm{ij}} \cdot \mathrm{dRR}_{\mathrm{ij}} + \sum_{y\,\in\,T,\ \mathrm{mRT}_j \ge r_y} r_y \cdot \mathrm{dRT}_{\mathrm{iy}} \quad \forall\, i\ \in\ R. \tag{19}$$

The equation above is made linear by using the transform $Z_{\mathrm{ij}} = \mathrm{fRR}_{\mathrm{ij}} \cdot \mathrm{dRR}_{\mathrm{ij}}$ and becomes

$$X_i + \sum_{j\,\in\,R} Z_{\mathrm{ji}} = \sum_{j\,\in\,R} Z_{\mathrm{ij}} + \sum_{y\,\in\,T,\ \mathrm{mRT}_{\mathrm{iy}} \ge r_y} r_y \cdot \mathrm{dRT}_{\mathrm{iy}} \quad \forall\, i\ \in\ R. \tag{20}$$

For the backup topologies, the flow conservation at the SC is ensured by the following equation:

$$\mathrm{fBR}_i^k \cdot \mathrm{dBR}_i^k + \sum_{j\,\in\,R,\;j\neq k} \mathrm{fRR}_{\mathrm{ji}}^k \cdot \mathrm{dRR}_{\mathrm{ji}}^k$$
$$= \sum_{j\,\in\,R,\;j\neq k} \mathrm{fRR}_{\mathrm{ij}}^k \cdot \mathrm{dRR}_{\mathrm{ij}}^k + \sum_{y\,\in\,T,\;\mathrm{mBT}_{\mathrm{iy}}\geq \mathrm{rb}_y} \mathrm{rb}_y \cdot \mathrm{dRT}_{\mathrm{iy}}^k$$
$$i \neq k \quad \forall\, i,\, k \,\in\, R. \tag{21}$$

The equation above is made linear by using the transform $Z_{\mathrm{ij}}^k = \mathrm{fRR}_{\mathrm{ij}}^k.\mathrm{dRR}_{\mathrm{ij}}^k$ and it becomes

$$X_i^k + \sum_{j\,\in\,R,\;j\neq k} Z_{\mathrm{ji}}^k = \sum_{j\,\in\,R,\;j\neq k} Z_{\mathrm{ij}}^k$$
$$+ \sum_{y\,\in\,T,\;\mathrm{mBT}_{\mathrm{iy}}\geq \mathrm{rb}_y} \mathrm{rb}_y \cdot \mathrm{dRT}_{\mathrm{iy}}^b \qquad i \neq k \quad \forall\, i,\, k \,\in\, R. \tag{22}$$

Z_{ij} and Z_{ij}^k are evaluated similar to how X_i was evaluated in (13)–(16).

3) Flow Balance at a TP: In the main topology, the amount of traffic between the eNB or the SC and the TP should be equal to the full-rate, r_i of the TP. This is ensured by the following constraint:

$$\sum_{i=i,\;\mathrm{mBT}_i\geq r_i}^{i} r_i \cdot \mathrm{dBT}_i + \sum_{j\,\in\,R,\;\mathrm{mRT}_{\mathrm{ji}}\geq r_i} r_i \cdot \mathrm{dRT}_{\mathrm{ji}} = r_i$$
$$\forall\, i \,\in\, T. \tag{23}$$

For the backup topologies, the amount of traffic at the TP should be equal to rb_i. This is ensured by the following constraint:

$$\sum_{i=i,\;\mathrm{mBT}_i\geq \mathrm{rb}_i}^{i} \mathrm{rb}_i \cdot \mathrm{dBT}_i^k + \sum_{j\,\in R,\,j\neq k,\,\mathrm{mRT}_{\mathrm{ji}}\geq \mathrm{rb}_i} \mathrm{rb}_i \cdot \mathrm{dRT}_{\mathrm{ji}} = \mathrm{rb}_i$$
$$\forall\, i \,\in\, T\,,\; k \,\in\, R. \tag{24}$$

V. OPTIMIZATION MODEL

A. In-Band Model

The model that we have so far assumes the out-band mode, and in this case the maximum capacity of link BR_i, RR_i, BT_i, and RT_i are mBR_i, mRR_i, mBT_i, and mRT_i respectively.

To accommodate the in-band mode, the capacity on link i depends on the activity of other links j, which may interfere with link i. We will need to set the capacity on link i so that it corresponds to the capacity that is subject to interference with other active links in the system. Rather than using a nonlinear formulation, we use a linear formulation that comes at the cost of an expanded space state.

The basic idea of the transformation is to precompute the capacity of the target link i for all possible cases of interference. This can be done offline, and outside the optimization formulation. Then, for each of the interference cases, we have a binary variable that is equal to 1 if this case is valid. Multiple interference cases may occur at the same time, e.g., if two links interfere with the target link, then there are three cases of interference, one for each link and the third for both links. Then, the optimization problem by determining the interference cases can select the corresponding capacity as the minimum capacity for all valid interference cases. The expansion in the state space is the result of the use of the binary variable corresponding to the interference cases.

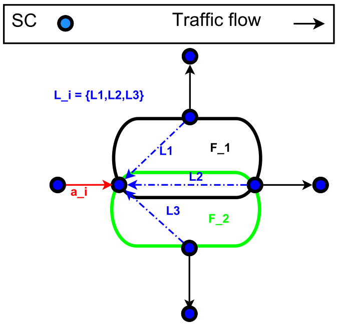


Fig. 1. Example for link interference.

It is worth mentioning that the conversion of the interference nonlinear characteristics to linear is developed without changing the parameters of the original problem (e.g., no. of subcarriers allocated for an SC). This is done by using a binary linear formulation in which the capacity of link i is defined as follows.

1) Assume that the maximum number of other links that can interfere with link i is a_i, and the set of such links is $Lx_i = l_1, l_2, \ldots, l_{a_i}$ where $x \in \{\mathrm{BR},\ \mathrm{RR},\ \mathrm{BT},\ \mathrm{RT}\}$.
2) The capacity of link i, given that links in the nth subset $\mathrm{Fx}_n \subseteq \mathrm{Lx}_i$ are active, including the empty subset, is given by $\mathrm{cx}_{\mathrm{Fx_n}}$ and is precomputed.

An example is shown in Fig. 1 for the interference sets and subsets to explain the interference that a link may suffer.

To find out which combination of links are active, a binary variable Ix_j is defined as being equal to 1 if link j is active. The number of active links in each Fx_n is evaluated as follows:

$$\mathrm{Ax}_{\mathrm{Fx_n}} = \sum_{j\,\in\,F_n^x} \mathrm{Ix}_j \quad \forall\, \mathrm{Fx}_n \subseteq \mathrm{Lx}_i. \tag{25}$$

Then, to find the combination that has all of its member links active. We define a binary variables $\mathrm{Yx}_{\mathrm{Fx_n}}$, which will be equal to 1 only if all links in the subset Fx_n are active. $\mathrm{Yx}_{\mathrm{Fx_n}}$ can be evaluated using the following constraints:

$$\mathrm{Yx}_{\mathrm{Fx_n}} \geq \mathrm{Ax}_{\mathrm{Fx_n}} - |\mathrm{Fx}_n| + 1 \tag{26}$$

$$\mathrm{Yx}_{\mathrm{Fx_n}} \leq \frac{\mathrm{Ax}_{\mathrm{Fx_n}} + \delta}{|\mathrm{Fx}_n| + \delta} \tag{27}$$

where δ is a small number. The addition of δ to both the numerator and denominator is to include the case of empty subset, in which case both $|\mathrm{Fx}_n|$ and $\mathrm{Ax}_{\mathrm{Fx_n}}$ are zeros.

Therefore, the maximum capacity of a link can be evaluated as the minimum for all cases in which $\mathrm{Yx}_{\mathrm{Fx_n}} = 1$. The upper bounds on the rates on this link ($\mathrm{mBR}^{i}_{\mathrm{FBR_n^i}}$, $\mathrm{mRR}^{\mathrm{ij}}_{\mathrm{FRR_n^{ij}}}$, $\mathrm{mBT}^{i}_{\mathrm{FBT_n^i}}$, $\mathrm{mRT}^{\mathrm{ij}}_{\mathrm{FRT_n^{ij}}}$) for main topology and ($\mathrm{mBR}^{\mathrm{ik}}_{\mathrm{FBR_n^{ik}}}$, $\mathrm{mRR}^{\mathrm{ijk}}_{\mathrm{FRR_n^{ijk}}}$, $\mathrm{mBT}^{\mathrm{ik}}_{\mathrm{FBT_n^{ik}}}$, $\mathrm{mRT}^{\mathrm{ijk}}_{\mathrm{FRT_n^{ijk}}}$) for backup topologies are input parameters to the problem and are calculated as follows:

$$\mathrm{mBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} = \mathrm{cBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} \cdot \mathrm{YBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} + (1 - \mathrm{YBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}}) \cdot M \tag{28}$$

$$\mathrm{mRR}^{\mathrm{ij}}_{\mathrm{FRR}^{\mathrm{ij}}_{\mathrm{n}}} = \mathrm{cRR}^{\mathrm{ij}}_{\mathrm{FRR}^{\mathrm{ij}}_{\mathrm{n}}} \cdot \mathrm{YRR}^{\mathrm{ij}}_{\mathrm{FRR}^{\mathrm{ij}}_{\mathrm{n}}} + (1 - \mathrm{YRR}^{\mathrm{ij}}_{\mathrm{FRR}^{\mathrm{ij}}_{\mathrm{n}}}) \cdot M \tag{29}$$

$$\mathrm{mBT}^{i}_{\mathrm{FBT}^{\mathrm{i}}_{\mathrm{n}}} = \mathrm{cBT}^{i}_{\mathrm{FBT}^{\mathrm{i}}_{\mathrm{n}}} \cdot \mathrm{YBT}^{i}_{\mathrm{FBT}^{\mathrm{i}}_{\mathrm{n}}} + (1 - \mathrm{YBT}^{i}_{\mathrm{FBT}^{\mathrm{i}}_{\mathrm{n}}}) \cdot M \tag{30}$$

$$\mathrm{mRT}^{\mathrm{ij}}_{\mathrm{FRT}^{\mathrm{ij}}_{\mathrm{n}}} = \mathrm{cRT}^{\mathrm{ij}}_{\mathrm{FRT}^{\mathrm{ij}}_{\mathrm{n}}} \cdot \mathrm{YRT}^{\mathrm{ij}}_{\mathrm{FRT}^{\mathrm{ij}}_{\mathrm{n}}} + (1 - \mathrm{YRT}^{\mathrm{ij}}_{\mathrm{FRT}^{\mathrm{ij}}_{\mathrm{n}}}) \cdot M \tag{31}$$

$$\mathrm{mBR}^{\mathrm{ik}}_{\mathrm{FBR}^{\mathrm{ik}}_{\mathrm{n}}} = \mathrm{cBR}^{\mathrm{ik}}_{\mathrm{FBR}^{\mathrm{ik}}_{\mathrm{n}}} \cdot \mathrm{YBR}^{\mathrm{ik}}_{\mathrm{FBR}^{\mathrm{ik}}_{\mathrm{n}}} + (1 - \mathrm{YBR}^{\mathrm{ik}}_{\mathrm{FBR}^{\mathrm{ik}}_{\mathrm{n}}}) \cdot M \tag{32}$$

$$\mathrm{mRR}^{\mathrm{ijk}}_{\mathrm{FRR}^{\mathrm{ijk}}_{\mathrm{n}}} = \mathrm{cRR}^{\mathrm{ijk}}_{\mathrm{FRR}^{\mathrm{ijk}}_{\mathrm{n}}} \cdot \mathrm{YRR}^{\mathrm{ijk}}_{\mathrm{FRR}^{\mathrm{ijk}}_{\mathrm{n}}} + (1 - \mathrm{YRR}^{\mathrm{ijk}}_{\mathrm{FRR}^{\mathrm{ijk}}_{\mathrm{n}}}) \cdot M \tag{33}$$

$$\mathrm{mBT}^{\mathrm{ik}}_{\mathrm{FBT}^{\mathrm{ik}}_{\mathrm{n}}} = \mathrm{cBT}^{\mathrm{ik}}_{\mathrm{FBT}^{\mathrm{ik}}_{\mathrm{n}}} \cdot \mathrm{YBT}^{\mathrm{ik}}_{\mathrm{FBT}^{\mathrm{ik}}_{\mathrm{n}}} + (1 - \mathrm{YBT}^{\mathrm{ik}}_{\mathrm{FBT}^{\mathrm{ik}}_{\mathrm{n}}}) \cdot M \tag{34}$$

$$\mathrm{mRT}^{\mathrm{ijk}}_{\mathrm{FRT}^{\mathrm{ijk}}_{\mathrm{n}}} = \mathrm{cRT}^{\mathrm{ijk}}_{\mathrm{FRT}^{\mathrm{ijk}}_{\mathrm{n}}} \cdot \mathrm{YRR}^{\mathrm{ijk}}_{\mathrm{FRT}^{\mathrm{ijk}}_{\mathrm{n}}} + (1 - \mathrm{YRT}^{\mathrm{ijk}}_{\mathrm{FRT}^{\mathrm{ijk}}_{\mathrm{n}}}) \cdot M \tag{35}$$

where M is a very large number. In (28), if subset FBR^{i}_{n} is active, hence $\mathrm{YBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}}$ is 1, then the capacity of link i is equal to $\mathrm{cBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}}$. Otherwise, the effect of subset FBR^{i}_{n} is excluded by having this capacity equal to a very large number, M. Equations (29)–(35) follow similar reasoning.

The maximum flow that can be transmitted on a link i is limited by the transmission power, the link distance, and the coding and modulation schemes. The maximum rate that can be assigned on the link from the eNB to SC_i, fBR_i, is limited by $\mathrm{mBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}}$, where $\mathrm{mBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}}$ is the maximum rate on this link. A similar notation is used for all the other links, and the constraints that ensure the upper bound are the following:

$$C \cdot \mathrm{fBR}_i \leq \mathrm{nBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} \cdot \mathrm{mBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} \quad \forall\, \mathrm{FBR}^{i}_{n} \subseteq \mathrm{LBR}_i \tag{36}$$

$$C \cdot \mathrm{fRR}_{\mathrm{ij}} \leq \mathrm{nRR}^{\mathrm{ij}}_{\mathrm{FRR}^{\mathrm{ij}}_{\mathrm{n}}} \cdot \mathrm{mRR}^{\mathrm{ij}}_{\mathrm{FRR}^{\mathrm{ij}}_{\mathrm{n}}} \quad \forall\, \mathrm{FRR}^{\mathrm{ij}}_{n} \subseteq \mathrm{LRR}_{\mathrm{ij}} \tag{37}$$

$$C \cdot \mathrm{fBR}^{k}_{i} \leq \mathrm{nBR}^{\mathrm{ik}}_{\mathrm{FBR}^{\mathrm{ik}}_{\mathrm{n}}} \cdot \mathrm{mBR}^{\mathrm{ik}}_{\mathrm{FBR}^{\mathrm{ik}}_{\mathrm{n}}} \quad \forall\, \mathrm{FBR}^{\mathrm{ik}}_{n} \subseteq \mathrm{LBR}^{k}_{i} \tag{38}$$

$$C \cdot \mathrm{fRR}^{k}_{\mathrm{ij}} \leq \mathrm{nRR}^{\mathrm{ijk}}_{\mathrm{FRR}^{\mathrm{ijk}}_{\mathrm{n}}} \cdot \mathrm{mRR}^{\mathrm{ijk}}_{\mathrm{FRR}^{\mathrm{ijk}}_{\mathrm{n}}} \quad \forall\, \mathrm{FRR}^{\mathrm{ijk}}_{n} \subseteq \mathrm{LRR}^{k}_{\mathrm{ij}} \tag{39}$$

$$C \cdot \mathrm{fBT}_i \leq \mathrm{nBT}^{i}_{\mathrm{FBT}^{\mathrm{i}}_{\mathrm{n}}} \cdot \mathrm{mBT}^{i}_{\mathrm{FBT}^{\mathrm{i}}_{\mathrm{n}}} \quad \forall\, \mathrm{FBT}^{i}_{n} \subseteq \mathrm{LBT}_i \tag{40}$$

$$C \cdot \mathrm{fRT}_{\mathrm{ij}} \leq \mathrm{nRT}^{\mathrm{ij}}_{\mathrm{FRT}^{\mathrm{ij}}_{\mathrm{n}}} \cdot \mathrm{mRT}^{\mathrm{ij}}_{\mathrm{FRT}^{\mathrm{ij}}_{\mathrm{n}}} \quad \forall\, \mathrm{FRT}^{\mathrm{ij}}_{n} \subseteq \mathrm{LRT}_{\mathrm{ij}} \tag{41}$$

$$C \cdot \mathrm{fBT}^{k}_{i} \leq \mathrm{nBT}^{\mathrm{ik}}_{\mathrm{FBT}^{\mathrm{ik}}_{\mathrm{n}}} \cdot \mathrm{mBT}^{\mathrm{ik}}_{\mathrm{FBT}^{\mathrm{ik}}_{\mathrm{n}}} \quad \forall\, \mathrm{FBT}^{\mathrm{ik}}_{n} \subseteq \mathrm{LBT}^{k}_{i} \tag{42}$$

$$C \cdot \mathrm{fRT}^{k}_{\mathrm{ij}} \leq \mathrm{nRT}^{\mathrm{ijk}}_{\mathrm{FRT}^{\mathrm{ijk}}_{\mathrm{n}}} \cdot \mathrm{mRT}^{\mathrm{ijk}}_{\mathrm{FRT}^{\mathrm{ijk}}_{\mathrm{n}}} \; \forall\, \mathrm{FRT}^{\mathrm{ijk}}_{n} \subseteq \mathrm{LRT}^{k}_{\mathrm{ij}} \tag{43}$$

$$i \neq k,\ j \neq k,\ i, j, k \in R$$

where the variable $\mathrm{nBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} \in \{0, 1, \ldots, C\}$ corresponds to the number of subcarriers allocated to any SC_i out of a total of C subcarriers, which is an input parameter, and C is assumed to be an integral power of 2 (e.g., $C = 512$)[1]. To avoid further complexity of the modeled problem, we assumed that the orthogonality among subcarriers is maintained and that there is no intercarrier interference between transceivers of different links. The constraint in (36) is a nonlinear constraint with $\mathrm{nBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}}$ a discrete variable and $\mathrm{mBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}}$ a continuous variable. The following equations are used to transform it to a linear form. A new variable $\mathrm{vBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} = \mathrm{nBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} * \mathrm{mBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}}$ is defined and a binary expansion [38] for $\mathrm{nBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}}$ is performed as

$$\mathrm{nBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} = \sum_{r=0}^{\rho} 2^r \cdot \mathrm{dnBR}^{\mathrm{ir}}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} \quad \forall\, \mathrm{FBR}^{i}_{n} \subseteq \mathrm{LBR}_i,\ i \in R \tag{44}$$

$$\rho = \log_2(C) \tag{45}$$

$$\mathrm{dnBR}^{\mathrm{ir}}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} \in \{0, 1\} \tag{46}$$

Then $\mathrm{vBR}^{i}_{F^{\mathrm{BR}}_{n}}$ can be rewritten as

$$\mathrm{vBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} = \sum_{r=0}^{\rho} 2^r \cdot \mathrm{wBR}^{\mathrm{ir}}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} \quad \forall\, \mathrm{FBR}^{i}_{n} \subseteq \mathrm{LBR}_i,\ i \in R \tag{47}$$

where

$$\mathrm{wBR}^{\mathrm{ir}}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} = \mathrm{dnBR}^{\mathrm{ir}}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} \cdot \mathrm{mBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}} \quad \forall\, \mathrm{FBR}^{i}_{n} \subseteq \mathrm{LBR}_i,\ i \in R. \tag{48}$$

Now the constraint in (36) is converted into nonlinear constraint but with $\mathrm{dnBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}}$ a binary variable and $\mathrm{mBR}^{i}_{\mathrm{FBR}^{\mathrm{i}}_{\mathrm{n}}}$ a continuous variable, which can be linearized using the same approach used in (13)–(16).

Similarly, the same conversion is used for variables ($\mathrm{vBT}^{i}_{\mathrm{FBT}^{\mathrm{i}}_{\mathrm{n}}}$, $\mathrm{vRR}^{\mathrm{ij}}_{\mathrm{FRR}^{\mathrm{ij}}_{\mathrm{n}}}$, $\mathrm{vRT}^{\mathrm{ij}}_{\mathrm{FRT}^{\mathrm{ij}}_{\mathrm{n}}}$) and the backup topologies variables ($\mathrm{vBR}^{\mathrm{ik}}_{\mathrm{FBT}^{\mathrm{ik}}_{\mathrm{n}}}$, $\mathrm{vBT}^{\mathrm{ik}}_{\mathrm{FBT}^{\mathrm{ik}}_{\mathrm{n}}}$, $\mathrm{vRR}^{\mathrm{ijk}}_{\mathrm{FRR}^{\mathrm{ijk}}_{\mathrm{n}}}$, $\mathrm{vRT}^{\mathrm{ijk}}_{\mathrm{FRT}^{\mathrm{ijk}}_{\mathrm{n}}}$). Also, when an SC fails, the rate on all the links incident on it is zero.

B. Out-Band Model

For out-band mode, the interference between the links are not considered, and the maximum rate that can be assigned on the link from the BS to SC_i, fBR_i, is limited by $\mathrm{fBR}_i \leq \mathrm{mBR}_i$. Similarly for all the other links, the constraints that ensure the upper bound are calculated according to

$$\mathrm{fBR}_i \leq \mathrm{mBR}_i \quad i \in R$$

$$\mathrm{fRR}_{\mathrm{ij}} \leq \mathrm{mRR}_{\mathrm{ij}} \quad i, j \in R$$

$$\mathrm{fBR}^{k}_{i} \leq \mathrm{mBR}_i \quad i \neq k,\ i, k \in R$$

$$\mathrm{fRR}^{k}_{\mathrm{ij}} \leq \mathrm{mRR}_{\mathrm{ij}} \quad i \neq k,\ j \neq k,\ i, j, k \in R.$$

---

[1] Without loss of generality, and to reduce the model complexity, we consider C to be a power of 2.

C. Objective Function

The primary objective of our solution is to minimize the total number of SCs used. This will minimize the cost of SC installation. We define the variable S_i, which indicates if an SC is installed at site SC_i either in the main topology or in any other topology

$$S_i \geq \mathrm{dR}_i \qquad i \in R \tag{49}$$

$$S_i \geq \mathrm{dR}_i^k \qquad i \neq k,\ i,\ k \in R. \tag{50}$$

To minimize the total number of SCs that are installed, $\sum_{i \epsilon R} S_i$ should be minimized.

We also aim at reducing the number of SCs used in backup topologies. Minimizing the number of SCs used in every topology allows us to remove lengthy paths. For example, if a TP can connect to the eNB by going through one SC only, it is better not to use two SCs for this TP. Thus, minimizing the number of SCs in the backup topologies will make a TP use the minimum necessary number of SCs it needs to connect to the eNB. The variable $\boldsymbol{T}_k$ designates the number of SCs used in the backup topology when SC k fails. We have the following constraint:

$$\mathsf{T}_k \geq \sum_{i \in R,\, i \neq k} \mathrm{dR}_i^k \quad \forall\, k \in R. \tag{51}$$

For a similar reason to the above, we aim to minimize the number of SCs that is used in the main topology, designated by the term V. We have the constraint

$$V \geq \sum_{i \in R} \mathrm{dR}_{i}. \tag{52}$$

The term Obj combines the terms above. The main term from the above is $\sum_{i\ \epsilon\ R} S_i$, since it gives the number of SCs that should be installed. It should be given a higher weight than the other terms. The maximum value of $\sum_{k\ \epsilon\ R} \boldsymbol{T}_k$ is N^2 and the maximum value of V is N. Thus, we give the weight $N^2 + N$ to the term with S_i

$$\mathrm{Obj} = (N^2 + N) \cdot \sum_{i \in R} S_i + \sum_{k \in R} \boldsymbol{T}_k + V. \tag{53}$$

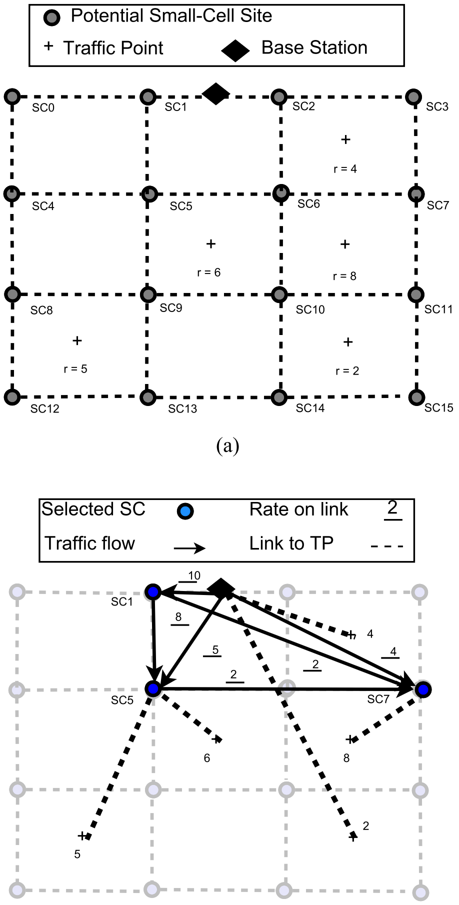
Then, the objective function is

$$\text{Minimize Obj}. \tag{54}$$

Implementing a system with multihop relays increases the operational complexity, and also solving the problem with wireless back-hauling and multihop topology adds more complexity to the problem. However, using more than two hops has advantages in terms of performance, such as range of coverage and bit rate. This is why we opted to develop a generic model that can accommodate any number of hops. To accommodate the case of only two hops, we can force all the nRR, fRR, dRR, mRR, and wRR to zero.

VI. NUMERICAL RESULTS

This section introduces numerical results based on the planning models presented above. First, we show examples of planning without fault tolerance in a WiMAX network. In these examples, when an SC fails, there is no guarantee of service to the TPs. Second, we show examples with fault tolerance, where service will be guaranteed even in case of failure of a SC. We consider both the out-band and in-band operation in solving the location planning MILP problem in a WiMAX HetNet. By using CPLEX, which runs on a modern multicore machine, we obtained solution times in terms of hours for realistic scenarios that shows a reasonable computation time for practical cases. The solution of the optimization problem for the proposed model is feasible as long as proper design parameters are selected to support the required resources for a certain network design scenario.



(a)

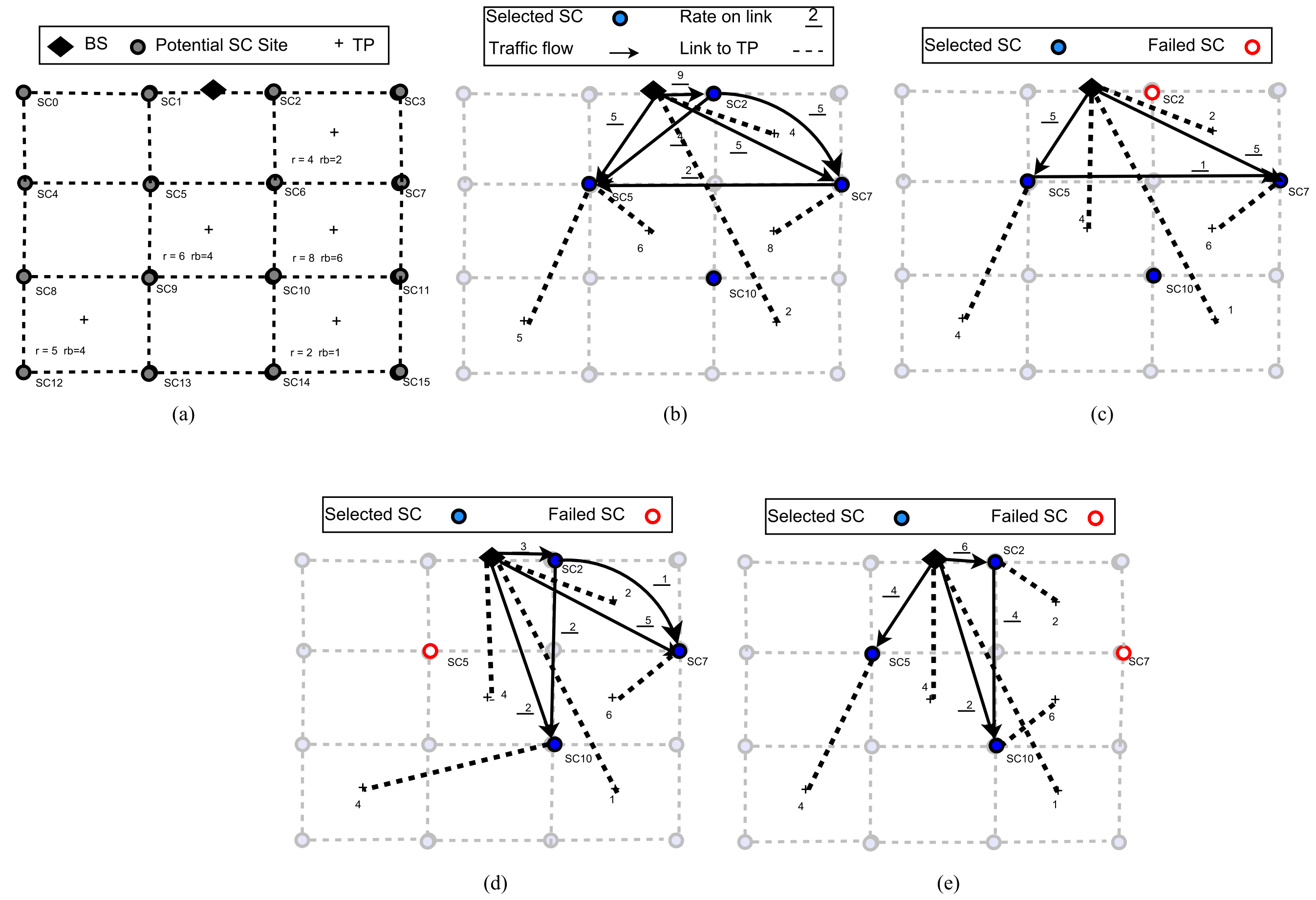
(b)

Fig. 2. Planning SC locations without fault-tolerance for out-band model. (a) Problem (main rates and backup rates in Mb/s). (b) Solution.

A. Planning SC Locations For Out-Band Mode

1) Planning Without Fault Tolerance: This section presents the initial results of planning the SC locations without fault tolerance. In this section, where no fault tolerance is considered, the variables and constraints in the model that are used for fault tolerance are omitted. These are all the variables that have an index k, which are: dR_i^k , dBR_i^k , dBT_i^k , dRR_i^k , fBR_i^k , fBR_i^k, and $\mathrm{fRR}_{\mathrm{ij}}^k$.

Also, the objective function will change. For the case in which we do not have fault tolerance, the objective is simply to minimize the total number of SCs used. Then, the objective function



(a) (b) (c)

(d) (e)

Fig. 3. Planning SC locations with fault-tolerance for out-band model. (a) Problem (main rates and backup rates in Mb/s). (b) Solution: The main topology. (c) Solution: Backup topology when SC2 fails. (d) Solution: Backup topology when SC5 fails. (e) Solution: Backup topology when SC7 fails.

is

$$\min \sum_{i \in R} \mathrm{dR}_i . \tag{55}$$

a) Theoretical Model: The following is a SC planning example that uses our solution. The problem is shown in Fig. 2(a). The planning area is made discrete by the use of a square grid. The BS location is on the top line of the grid, as shown in the figure. We select this setting since we consider that the BS is at the edge of the connected area. The area below the BS does not have the connection, and we plan to connect this area through the BS. Without loss of generality, we can use any topology with our model.

The potential sites for a SC are the corners of a grid square. In the 4 × 4 grid, the SC sites are numbered 0 to 15, as shown in the figure. The possible site of a TP is in the center of a square. The TPs are numbered 0 to 9. In the figure, the TP numbers are $TP_{(2,4,5,6,8)}$. The number shown in the figure next to each TP is its traffic demand in Mb/s.

TABLE IV
SYSTEM PARAMETERS

Description	Value
Band Width	5 MHz
Transmitter Power	46 dBm / (39.81 W)
path loss exponent (α)	2
Receiver Noise	– 104 dBm
Coverage area	12 KM × 12 KM
Number of SC sites	8

The maximum rate on the links is shown in Table III. The distance unit is the side length of a square in the grid. The table shows the feasible rate for the corresponding distance interval (per the model in Section III). The solution to this planning example is shown in Fig. 2(b). The shaded SC sites are the ones that have been selected. Three SCs are needed for this problem, which are SC_1, SC_5, and SC_7. The solid line links are the BS-SC and SC-SC links. The dashed lines are the BS-TP and SC-TP links. The underlined numbers are the link rates allocated by the solution. The rates of the dotted links are equal to the corresponding TPs' rates. The arrows on the links show the flow of traffic in the down-link to facilitate interpreting the results. However, the traffic may go in the up-link or down-link direction.

We note the following observations from this example.

1) The distance from the TP to the BS does not necessarily indicate a direct or relayed connection. For example, the TP with demand of 2 Mb/s is the farthest from the BS.

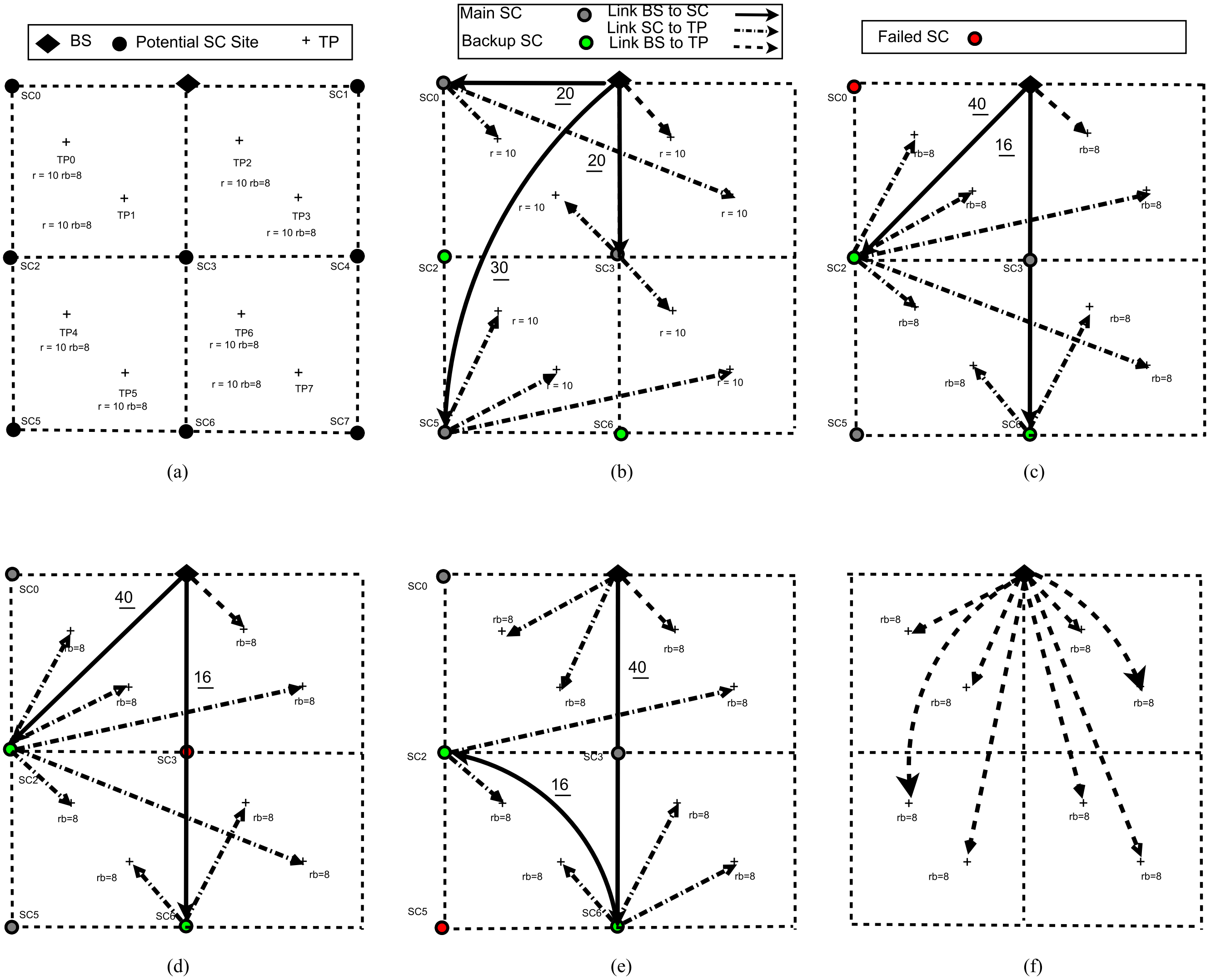


Fig. 4. Planning SC locations with fault tolerance for in-band model. (a) Problem (main rates and backup rates in Mb/s). (b) Solution: The main topology. (c) Solution: Backup topology when SC0 fails. (d) Solution: Backup topology when SC3 fails. (e) Solution: backup topology when SC5 fails. (f) Solution: Main topology for out-band mode (no interference).

However, its demand is relatively low, which can be satisfied by a single link. On the other hand, TPs that are closer to the BS have higher demands and require the use of a SC.

2) Second, the SC-to-SC links help in reducing the number of SCs in the HetNet. In our example, there is more traffic to the right of the BS ($4 + 8 + 2 = 14$) than the left (5) and the middle (6). Thus, in the solution, the diagonal and horizontal links, both with rate of 2 Mb/s, between SCs (1,7) and (5,7), respectively, relay the traffic from the right side to the less congested left side. If this was not the case, more SCs would be needed on the right side.

2) Planning With Fault Tolerance: In this part, we present planning results with fault tolerance. The problem input is shown in Fig. 3(a), and it has the same TP locations and rates as the example in Fig. 2(a). In this case, there are also backup rates for each TP, which are smaller than or equal to the the main rate.

Fig. 3(b) shows the main topology that supports the main rates of the TPs. This topology, similar to that in Fig. 2(b), supports the same normal operation rates and also uses three SCs. However, unlike the topology in Fig. 2(b), it uses SC_2, SC_5, and SC_7. Moreover, it also requires the installation of an additional relay at site SC_{10}. Even though SC_{10} is not used in the main topology, it is required in case one of the three used SCs fails.

In Fig. 3(b), the TPs with rates of 4 and 2 Mb/s connect directly to the BS since their direct link can support the required rate. This is similar to Fig. 2(b). Each of the other TPs connects to the SC that is closest to it.

Fig. 3(c) shows the backup topology that is used when SC_2 fails. In this topology, the backup rates are supported, which are smaller than the main rates in this example. Due to the lower rates, now three TPs are able to have a direct connection to the BS (compared to two in the main topology). The other two TPs connect through SCs. In this topology, SC_{10} is also not used

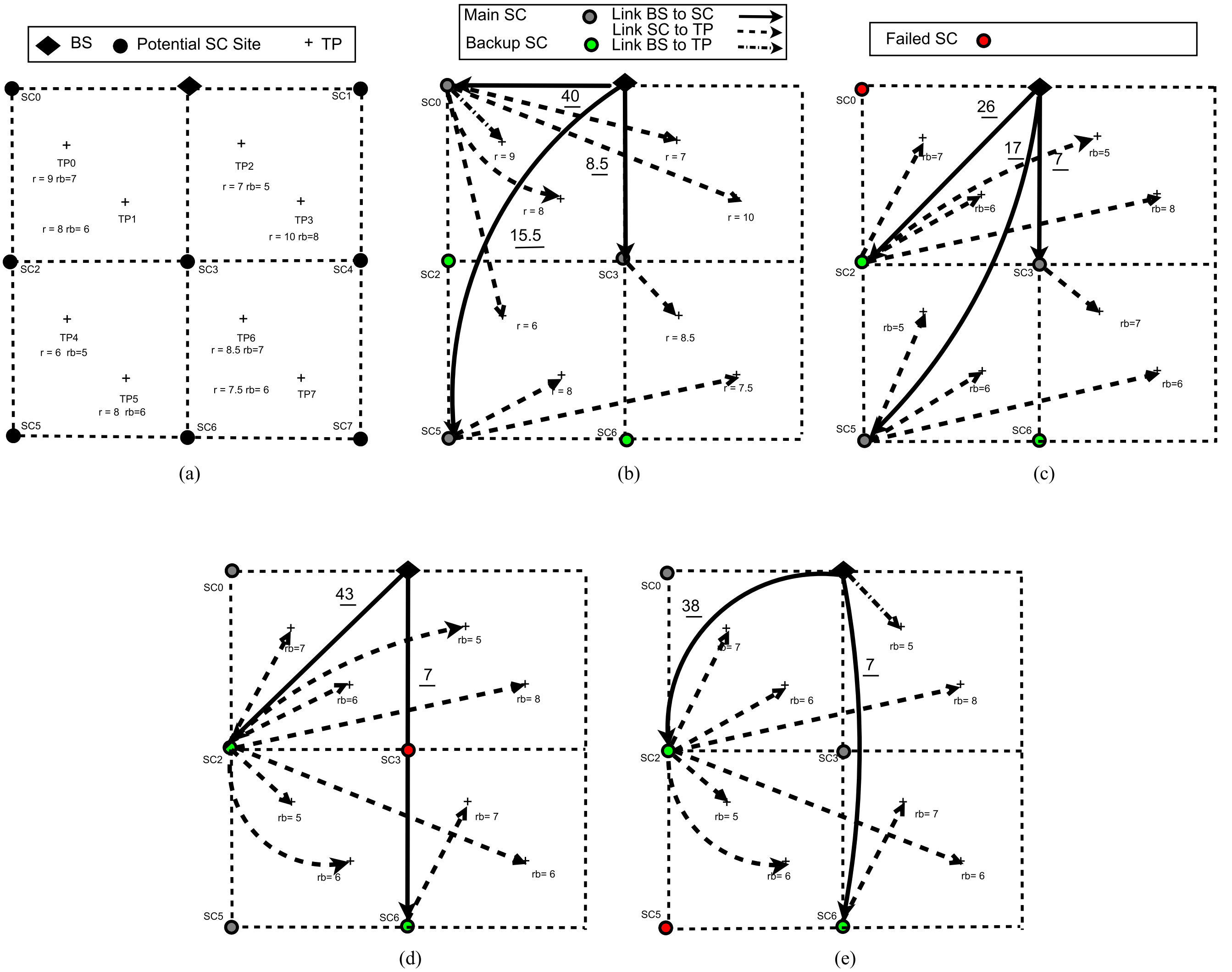


Fig. 5. Planning SC locations with fault tolerance for in-band model. (a) Problem (main rates and backup rates in Mb/s). (b) Solution: The main topology. (c) Solution: Backup topology when SC0 fails. (d) Solution: Backup topology when SC3 fails. (e) Solution: Backup topology when SC5 fails.

since SC_5 and SC_7 are able to support the TPs' demands, which makes the recovery from SC_2 failure faster, since SC_{10} need not be used.

The topology in Fig. 3(d) is the case in which SC_5 fails. In this case, three SCs are needed to support the TPs. Notice that in the previous topology, SC_5 was strategically located between the BS and the TP in the lower-left corner. Since SC_5 has failed, there is no SC that can satisfy this TP. Thus, two SCs are used to connect this TP.

In Fig. 3(e), the topology that is used when SC_7 fails is shown. Now the TP in the lower-left corner is able to connect via SC_5. However, the TP with demand 6 Mb/s, which was 9 previously relying on SC_7 cannot connect with only one SC. Then, SC_2 and SC_{10} convey the traffic of this TP in this case. However, the TP connects only to SC_{10}, and SC_{10} connects to both the BS and SC_2 to receive the data. Finally, when SC_{10} fails, we can continue to use the main topology as in Fig. 3(b) since this topology does not use SC_{10}.

We compared our proposed solution to provide fault tolerance to the one without fault tolerance. The results in Fig. 3(e) shows that with one extra SC for achieving fault tolerance, the transmission rates achieved is about 68% of the rates achieved in the main topology. Fig. 3(b) shows that in case of SC_5 failure and without fault tolerance planning only 44% of the rate can be provided and 40% of the TPs will have no service at all. This transmission performance can definitely guarantee business continuity in case of failures. However, this comes on an increase of 25% of the capital cost for acquiring the additional SC.

B. Planning SC Locations for In-Band Model

For in-band mode, the interference between different links is taken into consideration and the maximum rate that can be assigned on any link is calculated according to the interference model listed in Section V-A. In the following planning case, we are trying to present the effect of the interference consideration on the allocation of SCs using the planning with fault tolerance model and compare it with out-band model.

Planning results for the in-band with fault tolerance model are presented to show the interference effect on the planning process. Interference is modeled such that the transmission from each BS to another SC or TP is interfered by transmission from other SCs within 1 unit grid, similarly the transmission from each SC to an SC or TB may suffer from interference by the SCs within the 1-by-1 grid distance.

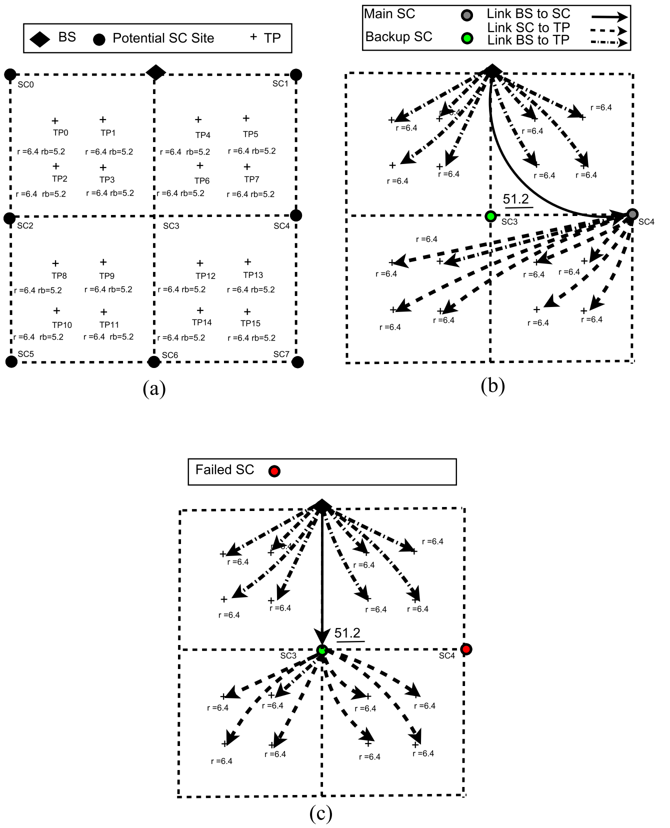
Two scenarios using different numbers of TPs are presented to examine the network load conditions. The rate requirements change in each scenario to show the effect of the load conditions on the SCs allocation. The parameters used in the simulation are shown in Table IV.

1) Scenario 1 (8 TPs):

a) Homogeneous Rate Requirements: Fig. 4(a) shows the locations of the TPs and SCs. In this in-band case, eight SCs and eight TPs are used. For the case of in-band mode, the different maximum link rates are calculated according to (28)–(43) and the SINR is calculated according to (1). The main (r_i) and backup rate (rb_i) requirement for each TP are shown in Fig. 4(a), where backup rates are smaller than the the main topology required rates.

Fig. 4(b) shows the main topology that supports the main rates to the TPs. This topology supports the equal homogeneous rates for all TPs (10 Mb/s), and uses three SC (SC_0, SC_3, SC_5) to satisfy all the TPs rate requirements. In Fig. 4(c,d,e) the backup topology for the in-band mode is presented, in case of any SC failure in the main topology, the network will implement the backup solution using two SCs (SC_2, SC_6). Both SCs will operate to support the backup rates (8 Mb/s) to the TPs. The solution shows that the BS supports TP_2, SC_2 supports $\mathrm{TP}_{(0,1,3,4,7)}$ and SC_6 supports $\mathrm{TP}_{(5,6)}$ in both backup plans for (SC_0 and SC_3). However in backup plan for SC_5, $\mathrm{TP}_{(0,1,2)}$ are supported by the BS. SC_6 supports $\mathrm{TP}_{(5,6,7)}$ and SC_2 supports $\mathrm{TP}_{(3,4)}$. The increase in the number of SCs and the diversity of their locations and which TPs they serve are due to the consideration of the interference caused by the in-band mode. Results in Fig. 3(f) for the main topology without interference considerations shows that the TPs rate requirements are all satisfied by direct links from the BS. The reason is that the BS to TPs links maximum rates are capable of delivering the TPs rate requirements without any relaying. This comparison shows that modeling the problem with interference constraints requires SCs implementation, but in case of ignoring the interference in the model no relaying is required to support the TPs with same rate requirements. The comparison clearly shows the importance of considering the interference effect in planning the SCs placement.

b) Heterogeneous Rate Requirements: Fig. 5(a) shows the main and backup rate requirement for each TP. The rate requirements for the TPs in this heterogeneous case are different, and the same number of SCs are needed to satisfy a smaller total rate requirements (64 Versus 80 Mb/s) than that of the homogeneous case. The results for the main topology in Fig. 5(b) show that SC_0 supports $\mathrm{TP}_{(0,1,2,3,4)}$, SC_3 supports TP_6, and SC_5 supports $\mathrm{TP}_{(5,7)}$. Fig. 5(c) shows a the backup plan when SC_0 fails, only backup SC_2 is activated to support $\mathrm{TP}_{(0,1,2,3)}$, SC_3 still supports TP_6, and SC_5 is supporting $\mathrm{TP}_{(4,5,7)}$. Results in Fig. 5(d) show the backup plan when SC_3 fails, both backup $\mathrm{SC}_{(2,6)}$ are activated where SC_2 supports $\mathrm{TP}_{(0,1,2,3,4,5,7)}$ and SC_6 is supporting TP_6. Fig. 5(e) shows the backup plan when SC_5 fails, also both backup $\mathrm{SC}_{(2,6)}$ are activated but SC_2 supports $\mathrm{TP}_{(0,1,3,4,5,7)}$, BS supports TP_2, and SC_3 still supports TP_6. It is also noticed that not all SCs from either the main or backup SCs are used in all backup plans since part of the objective is to use the minimum amount of SCs in any individual plan.



(a) (b)

(c)

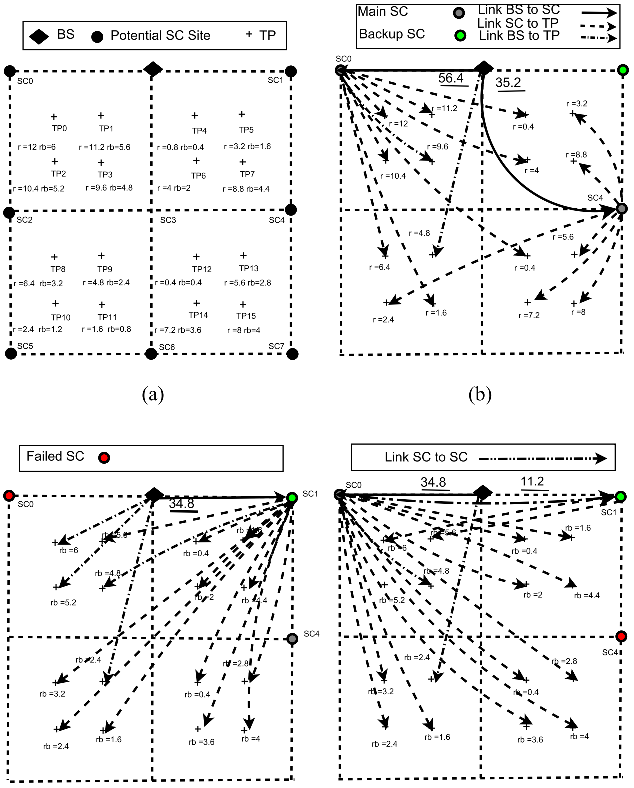
Fig. 6. Planning SC locations with fault tolerance for in-band model. (a) Problem (main rates and backup rates in Mb/s). (b) Solution: The main topology. (c) Solution: Backup topology when SC4 fails.

2) Scenario 2 (16 TPs):

a) Homogeneous Rate Requirements: In this scenario, the number of TPs is increased to 16 TPs to show more insight about the distribution of the network load in the network. The case shown in Fig. 6(a) presents the locations and homogeneous rate requirements for all TPs. Only two SCs are needed in this scenario, Fig. 6(b) shows the main topology where TPs 0 to 7 are supported by the BS and TPs 8 to 15 are supported by TP_4. Once TP_4 has failed, the backup topology activates SC_3, which supports the TPs 8 to 15 and the BS keeps supporting TPs 0 to 7.

The results show the capability of the BS to support the TPs when their rate requirements decrease from 10 Mb/s in the first scenario to 6.4 Mb/s in this scenario. The reason is that the maximum link capacities are able to support the required rates to the upper SCs without any relaying.

b) Heterogeneous Rate Requirements: Finally, the heterogeneous case of the 16 TPs, in which the TPs have different rate requirements, as shown in Fig. 7(a). The main topology shown in Fig. 7(b) requires two SCs to satisfy the TPs. SC_0 supports $\mathrm{TP}_{(0,1,2,3,4,6,8,11,12)}$, and SC_0 supports the rest of the TPs. Fig. 7(c) and 7(d) show the backup solution in case of SC_0



(a) (b)

(c) (d)

Fig. 7. Planning SC locations with fault tolerance for in-band model. (a) Problem (main rates and backup rates in Mb/s). (b) Solution: The main topology. (c) Solution: Backup topology when SC0 fails. (d) Solution: Backup topology when SC4 fails .

or SC_4 failure, respectively. In case of SC_4 failure, multihop relaying occurs from SC_0 to SC_1 for 11.2 Mb/s to be relayed to $TP_{(0,2)}$. This result shows that for this scenario, although the TPs total required rate is less than the homogeneous case (96.4 Versus 102.4 Mb/s), the planning still needs more SCs to serve the TPs. This increase in the number of the SCs is clearly due to the heterogeneity in the rate requirements that causes more interference and requires more relaying.

VII. CONCLUSION

In this paper, we considered the problem of planning the SC locations in the WiMAX network in a fault-tolerant manner. To the best of our knowledge, this is the first work that provides fault tolerance in planning SC locations in WiMAX. We provided a MILP that formulates the planning problem. The allocation problem is studied in both the out-bound and in-bound relaying modes. To address the nonlinearity in the problem formulation, a mapping form nonlinear to linear formulation is performed. The mapping utilized a binary conversion methodology and traded the nonlinearity by an increase in the state space size of the problem. We solved the problem with CPLEX and obtained numerical results that show how our model produces the main topology and the backup topologies of a network. Finally, we considered the existence of obstacles in the planning field, such as a large structure or a natural obstacle. We showed how our model can deal with these obstacles and plan the network around them effectively.

REFERENCES

[1] Z. Abichar and J. Chang, "A medium access control scheme for wireless lans with constant-time contention," *IEEE Trans. Mobile Comput.*, vol. 10, no. 2, pp. 191–204, Feb. 2011.

[2] *IEEE Standard for Air Interface for Broadband Wireless Access Systems*, IEEE 802.16–2012, 2012.

[3] *3GPP TS 132541 version 11 Release 11,Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management; Self-Organizing Networks (SON); Self-healing concepts and requirements*, 3GPP, 2012.

[4] E. Chu, I. Bang, S. H. Kim, and D. K. Sung, "Self-organizing and self-healing mechanisms in cooperative small-cell networks," in *Proc. 2013 IEEE 24th Int. Symp. Pers. Indoor Mobile Radio Commun.*, Sep. 2013, pp. 1576–1581.

[5] Z. Abichar, A. Kamal, and J. Chang, "Planning of relay station locations in IEEE 802.16 (WiMAX) networks," in *Proc. 2010 IEEE Wireless Commun. Netw. Conf.*, Apr. 2010, pp. 1–6.

[6] M. Peng, Y. Li, Z. Zhao, and C. Wang, "System architecture and key technologies for 5g heterogeneous cloud radio access networks," *IEEE Netw.*, vol. 29, no. 2, pp. 6–14, Mar./Apr. 2015.

[7] M. Eguizabal and A. Hernandez, "Resource allocation and interference management strategies for inband relaying in LTE-A," *Telecommun. Syst.*, vol. 61, pp. 839–860, 2016. [Online]. Available: http://dx.doi.org/10.1007/s11235–015-0040-7

[8] M. Minelli, M. Ma, M. Coupechoux, and P. Godlewski, "Scheduling impact on the performance of relay-enhanced LTE-A networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 2496–2508, Apr. 2016.

[9] J.-Y. Chang and Y.-S. Lin, "A clustering deployment scheme for base stations and relay stations in multi-hop relay networks," *J. Comput. Elect. Eng.*, vol. 40, no. 2, pp. 407–420, 2014.

[10] B. Lin and P.-H. Ho, "Dimensioning and location planning of broadband wireless networks under multi-level cooperative relaying," *IEEE Trans. Wireless Commun.*, vol. 8, no. 11, pp. 5682–5691, Nov. 2009.

[11] B. Lin, M. Mehrjoo, P.-H. Ho, L.-L. Xie, and X. Shen, "Capacity enhancement with relay station placement in wireless cooperative networks," in *Proc. 2009 IEEE Wireless Commun. Netw. Conf.*, Apr. 2009, pp. 1–6.

[12] B. Lin, P.-H. Ho, L.-L. Xie, X. Shen, and J. Tapolcai, "Optimal relay station placement in broadband wireless access networks," *IEEE Trans. Mobile Comput.*, vol. 9, no. 2, pp. 259–269, Feb. 2010.

[13] Y. Yu, S. Murphy, and L. Murphy, "Planning base station and relay station locations in IEEE 802.16j multi-hop relay networks," in *Proc. 2008 5th IEEE Consum. Commun. Netw. Conf.*, Jan. 2008, pp. 922–926.

[14] Y. Yu, S. Murphy, and L. Murphy, "A clustering approach to planning base station and relay station locations in IEEE 802.16j multi-hop relay networks," in *Proc. 2008 IEEE Int. Conf. Commun.*, May 2008, pp. 2586–2591.

[15] S.-J. Kim, S.-Y. Kim, H.-W. Lee, S.-W. Ryu, H.-W. Lee, and C.-H. Cho, "Multi-hop relay based coverage extension in the IEEE802.16j based mobile WiMAX systems," in *Proc. 2008 4th Int. Conf. Netw. Comput. Adv. Inf. Manage.*, Sep. 2008, vol. 1, pp. 516–522.

[16] B. Lin, P.-H. Ho, L.-L. Xie, and X. Shen, "Optimal relay station placement in IEEE 802.16J networks," in *Proc. 2007 Int. Conf. Wireless Commun. Mobile Comput.*, 2007, pp. 25–30.

[17] H.-C. Lu and W. Liao, "Joint base station and relay station placement for IEEE 802.16j networks," in *Proc. 2009 IEEE Glob. Telecommun. Conf.*, Nov 2009, pp. 1–5.

[18] D. Niyato, E. Hossain, D. I. Kim, and Z. Han, "Relay-centric radio resource management and network planning in IEEE 802.16j mobile multihop relay networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 6115–6125, Dec. 2009.

[19] C. Y. Chang, C. T. Chang, M. H. Li, and C. H. Chang, "A novel relay placement mechanism for capacity enhancement in IEEE 802.16j WiMAX networks," in *Proc. 2009 IEEE Int. Conf. Commun.*, Jun. 2009, pp. 1–5.

[20] Y. Yu, S. Murphy, and L. Murphy, "Interference aware relay station location planning for IEEE 802.16J mobile multi-hop relay network," in *Proc. 4th ACM Workshop Perform. Monit. Meas. Heterogeneous Wireless Wired Netw.*, 2009, pp. 201–208.

[21] Y. Yu, S. Murphy, and L. Murphy, "Planning base station and relay station locations for IEEE 802.16j network with capacity constraints," in *Proc. 2010 7th IEEE Consum. Commun. Netw. Conf.*, Jan. 2010, pp. 1–5.

[22] C.-Y. Chang and M.-H. Li, "A placement mechanism for relay stations in 802.16j WiMAX networks," *Wireless Netw.*, vol. 20, no. 2, pp. 227–243, 2014.

[23] A. Bari, A. Jaekel, and S. Bandyopadhyay, "Optimal placement of relay nodes in two-tiered, fault tolerant sensor networks," in *Proc. 2007 12th IEEE Symp. Comput. Commun.*, Jul. 2007, pp. 159–164.

[24] A. Bari, Y. Xu, and A. Jaekel, "Integrated placement and routing of relay nodes for fault-tolerant hierarchical sensor networks," in *Proc. 2008 17th Int. Conf. Comput. Commun. Netw.*, Aug. 2008, pp. 1–6.

[25] S. Misra, S. D. Hong, G. Xue, and J. Tang, "Constrained relay node placement in wireless sensor networks: Formulation and approximations," *IEEE/ACM Trans. Netw.*, vol. 18, no. 2, pp. 434–447, Apr. 2010.

[26] J. Bredin, E. Demaine, M. Hajiaghayi, and D. Rus, "Deploying sensor networks with guaranteed fault tolerance," *IEEE/ACM Trans. Netw.*, vol. 18, no. 1, pp. 216–228, Feb. 2010.

[27] X. Han, X. Cao, E. Lloyd, and C.-C. Shen, "Fault-tolerant relay node placement in heterogeneous wireless sensor networks," in *Proc. 26th IEEE Int. Conf. Comput. Commun.*, May 2007, pp. 1667–1675.

[28] W. Zhang, G. Xue, and S. Misra, "Fault-tolerant relay node placement in wireless sensor networks: Problems and algorithms," in *Proc. 26th IEEE Int. Conf. Comput. Commun.*, May 2007, pp. 1649–1657.

[29] P. Meena, D. Gurjar, A. Singh, and S. Verma, "Optimal positioning of base station in wireless sensor networks: A survey," in *Intelligent Computing, Networking, and Informatics, (ser. Advances in Intelligent Systems and Computing)*, D. P. Mohapatra and S. Patnaik, Eds. Delhi, India: Springer India, 2014, vol. 243, pp. 1135–1143.

[30] A. Samson Arun Raj, K. Ramalakshmi, and C. Priyadharsini, "A survey on classification of fault tolerance techniques available in wireless sensor network," in *Int. J. Eng. Res. Technol.*, vol. 3, no. 1, pp. 688–691, 2014.

[31] M. Nivedita and G. Raja, "Efficient relay station placement strategy for broadband wireless networks - 4G," in *Proc. Int. Conf. Recent Trends Inf. Technol.*, Apr. 2012, pp. 282–286.

[32] H. Wang, X. Yin, C. Chen, and X. Wang, "DPRP: Dual-path relay placement in WiMAX mesh networks," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2013, pp. 1597–1602.

[33] W. Stallings, *Wireless Communications and Networks*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2001.

[34] *IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, IEEE Std 802.16–2004, 2004.

[35] T. Tsourakis and K. Voudouris, "WiMax network planning and system's performance evaluation," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2007, pp. 1948–1953.

[36] M. Molina-Garcia and J. Alonso, "Planning and sizing tool for WiMAX networks," in *Proc. 2007 IEEE Radio Wireless Symp.*, Jan. 2007, pp. 403–406.

[37] J. Garcia-Fragoso and G. Galvan-Tejada, "Cell planning based on the WiMax standard for home access: A practical case," in *Proc. 2005 2nd Int. Conf. Elect. Electron. Eng.*, Sep. 2005, pp. 89–92.

[38] P. Rubin, "Integer variables and quadratic terms," Jul. 2013. [Online]. Available: http://orinanobworld.blogspot.com/2013/07/integer-variables-and-quadratic-terms.html

Tamer Omar received the B.S. degree in electrical engineering from Ain Shams University, Cairo, Egypt, in 1999, the MBA degree with emphasis on MIS from the Arab Academy for Science and Technology, Alexandria, Egypt, in 2004, and the Ph.D. degree from the Electrical Engineering Department, Iowa State University, Ames, IA, USA, in 2014.

He is an Assistant professor in the Department of Technology Systems, East Carolina University, Greenville, NC, USA. His research interests include wireless networks architecture, resources allocation in wireless networks, heterogeneous networks, self-organized networks, big data implementation and analysis, RDBMS, and decision support systems. Dr. Omar has six years of experience in academia and more than 10 years of industrial experience in different ICT positions.

Zakhia Abichar received the B.S. and Ph.D. degrees in computer engineering from Iowa State University, Ames, IA, USA.

He is a Lecturer in the Department of Electrical and Computer Engineering, University of Central Florida, Orlando, FL, USA. His teaching experience is in the undergraduate computer engineering classes, and his research experience is in the area of wireless networks and mobile computing. His research interests include STEM education and developing effective teaching techniques and contents to spread computing learning beyond higher education institutions. He also has experience in engineering assessment and has worked on undergraduate computer engineering program assessment.

Ahmed E. Kamal (S'82–M'87–SM'91–F'12) received the B.Sc. (distinction with honors) and the M.Sc. degrees both from Cairo University, Cairo, Egypt, in 1978 and 1980, respectively, and the M.A.Sc. and Ph.D. degrees from the University of Toronto, Toronto, ON, Canada, in 1982 and 1986, respectively, all in electrical engineering.

He is a Professor in the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA. His research interests include wireless networks, cognitive radio networks, optical networks, wireless sensor networks, Internet of Things, and performance evaluation.

Dr. Kamal received the 1993 IEE Hartree Premium for papers published in Computers and Control in IEE Proceedings, and the Best Paper Award of the IEEE Globecom 2008 Symposium on Ad Hoc and Sensors Networks Symposium. He chaired or co-chaired Technical Program Committees of several IEEE-sponsored conferences, including the Optical Networks and Systems Symposia of the IEEE Globecom 2007 and 2010, the Cognitive Radio and Networks Symposia of the IEEE Globecom 2012 and 2014, and the Access Systems and Networks track of the IEEE International Conference on Communications 2016. He is also the Chair of the IEEE Communications Society Technical Committee on Transmission, Access and Optical Systems for 2015 and 2016. He is on the editorial boards of the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, the *Computer Networks* journal, and the *Optical Switching and Networking* journal. He served as an IEEE Communications Society Distinguished Lecturer in 2013 and 2014. He is a senior member of the Association of Computing Machinery.

J. Morris Chang (SM'08) received the Ph.D. degree from North Carolina State University, Raleigh, NC, USA, in 1993.

He is a Professor in the Department of Electrical Engineering, University of South Florida, Tampa, FL, USA. His past industrial experiences include positions at Texas Instruments, Microelectronic Center of North Carolina, and AT&T Bell Labs. In the last five years, his research projects on cyber security have been funded by DARPA. His research interests include cyber security, wireless networks, and energy-efficient computer systems.

Dr. Chang received the University Excellence in Teaching Award at Illinois Institute of Technology, in 1999. He is currently a Handling Editor of the *Journal of Microprocessors and Microsystems* and the Associate Editor-in-Chief of the IEEE IT PROFESSIONAL.

Mohammed Abdullah Alnuem received the M.Sc. degree in distributed systems and networks from the University of Bradford, Bradford, U.K., in 2005, and the Ph.D. degree in mobile computing and networks from the School of Informatics, University of Bradford, in 2009.

He is an Associate Professor in the College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia. His research interests include computer networks (wired and wireless), mobile ad hoc and sensor networks, cloud computing, and distributed systems. He has published a number of research papers in peer-reviewed conferences and journals. His research is financially supported by several grants.