

Robust gateway placement in wireless mesh networks[☆]

Kagan Gokbayrak

Department of Industrial Engineering, Bilkent University, 06800 Bilkent, Ankara, Turkey



ARTICLE INFO

Article history:

Received 10 April 2017

Revised 19 April 2018

Accepted 24 April 2018

Available online 26 April 2018

Keywords:

Optimization

Network design

Robust gateway placement

Wireless mesh networks

Local search

ABSTRACT

Wireless mesh networks (WMNs) are communication networks that provide wireless Internet access over areas with limited infrastructure. Each node in a WMN serves several clients in its coverage area and transfers their traffic over wireless media to a few gateway nodes that have wired connections to the Internet. In this paper, we consider the Internet gateway placement (IGP) problem along with operational problems such as routing and wireless transmission capacity allocation. To eliminate wireless interference, we adopt the spatial reuse time division multiple access (TDMA) method, in which wireless transmissions are scheduled to occur at different time slots. We also employ destination-based single path routing for its ease of implementation. We present two mixed integer linear programming (MILP) formulations, both of which jointly determine the minimum number of gateway nodes needed to support forecasted demand, the locations of these gateway nodes, the routing trees, and the time slot allocations to wireless links. These formulations differ in the flow constraints. We also present a set of valid inequalities for the formulation with the multi-commodity flow constraints. In most cases, the solution to the IGP problem is not unique. Therefore, we also introduce a local search algorithm to select the most robust solution against any demand forecast errors. On example networks, we compare the proposed formulations with and without the valid inequalities in terms of the exact solution performances and the linear programming (LP) relaxations. We also demonstrate our local search algorithm to improve robustness against forecast errors on these example networks.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

Wireless mesh networks (WMNs) form the final portion of Internet access networks that reach the end-users' premises. They consist of radio nodes organized in a mesh topology that communicate via a wireless media in a multihop fashion. Each node provides access to clients in its coverage area. A few of these nodes (we call them gateways) have wired connections to the Internet, and all nodes of the network connect to the Internet via these gateway nodes. As WMNs eliminate the need for a wired connection to each node, they have emerged as a low-cost access technology, especially for developing areas.

In contrast to battery-operated *ad hoc* network nodes, WMN nodes are always plugged into an electric outlet; therefore, energy consumption is not a concern. In addition, we assume that all nodes of the network stay operational at all times, i.e., they are reliable nodes. We design the network operation in a deterministic fashion to support a set of forecasted node demands.

The main concern in wireless communications is the interference due to simultaneous nearby transmissions over the same frequency band (channel). Since wireless signals are transmitted in a shared medium (air), they can be received by unintended nodes and cause interference at them. According to the physical interference model in Gupta and Kumar (2000), a node can successfully decode a signal if the signal-to-interference-plus-noise ratio (SINR) exceeds a certain threshold value. Note that we cannot eliminate ambient noise but interfering transmissions can be either scheduled at different time slots or assigned to different channels. In this paper, we adopt the time division multiple access (TDMA) scheme, in which a time frame is divided into equal time slots that are allocated to wireless transmissions, and the same frame structure is repeated periodically. Since the signal levels decrease with distance, it is possible to use the same time slot for multiple transmissions if the transmissions are located far enough from each other, which is called the spatial reuse of wireless resources. In this paper, we benefit from the spatial reuse and allocate the same time slot to multiple non-interfering transmissions as in Nelson and Kleinrock (1985), Djukic and Valaee (2007), Papadaki and Friderikos (2008), Quintas and Friderikos (2012), and Chen and Lea (2013).

[☆] This work was supported in part by TÜBİTAK (Turkish Scientific and Technological Research Council) grant 110M312.

E-mail address: kgokbayr@bilkent.edu.tr

Unlike Amaldi et al. (2008) and Benyamina et al. (2009), where the locations of the WMN nodes are to be selected to provide full coverage to mesh clients, we assume that the node locations are given. The links of the network are determined based on the ambient noise level: we assume that a link between nodes i and j exists if the SINR ratio at j exceeds a certain threshold value when all other nodes are silent and node i is transmitting at its maximum power level. Note that if the ambient noise levels at all nodes were zero, then we would have a fully connected network. Under appropriate transmission power settings, it may be possible to activate multiple distant links simultaneously. The set of all non-interfering transmission sets and the corresponding transmission power levels are assumed to be pre-determined via a method that we describe in Gokbayrak and Yildirim (2017).

To eliminate the need to reorder packets at the destination, we assume that traffic is routed towards each gateway on a single path. Specifically, for its ease of implementation (as described in Hsiao et al., 2001), we employ the destination-based routing, i.e., all traffic at a non-gateway node is routed to the same next node regardless of the source of the traffic. Destination-based routing operates with routing trees, one tree for each gateway node.

Our objective is to determine the minimum number of gateways to satisfy forecasted demand, which is mentioned in Targon et al. (2010) to be NP-hard. Efficient use of wireless capacity is crucial for this task. We formulate two mixed integer linear programming (MILP) models that jointly determine routing, time slot allocations, and gateway locations to achieve this efficiency. These models differ in the flow constraints. We also introduce a set of valid inequalities for the formulation with the multi-commodity flow constraints. We perform computational experiments on WMNs of different sizes to compare these formulations with and without the valid inequalities in terms of the exact solution performances, and the LP relaxations.

Each MILP model yields the minimum number of gateways together with a set of gateways of that cardinality, the routing trees rooted at these gateways, and the TDMA time slot allocations to satisfy a given set of node demands. While the minimum number is unique, there may be many different sets of gateways of the same cardinality with their routing trees and slot allocations that may satisfy the demand. Some of these solutions can barely satisfy demand while others can accommodate some additional unforecasted demand. Since the forecasting process is not always accurate and since the demand is expected to increase over time, to postpone the deployment of an additional gateway, we search for the gateway set of the same cardinality that can satisfy the node demands multiplied by the highest multiplier (as the higher the multiplier is the more robust the gateway set will be.) Note that, as the major component of the gateway deployment cost is due to cable installations, we assume that the gateways cannot be relocated to accommodate variations from the forecasts. Hence, rather than implementing the MILP model solution directly, we employ a k -opt local search algorithm to obtain a more robust solution against forecast errors.

The rest of this paper is organized as follows: We discuss the related literature in Section 2 and present the mathematical formulations for the IGP problem along with some valid inequalities in Section 3. In Section 4, we present the local search algorithm to obtain more robust solutions. We present our computational studies in Section 5. In this section, we solve the two alternative formulations of the problem with and without the valid inequalities for some example networks and evaluate their performances by comparing the resulting solutions and the solution times. We also compare the formulations in terms of their LP relaxations. More robust solutions are also obtained for these networks. Our example networks in this section are of grid topology; however, our meth-

ods do not assume any specific topology structure. In Section 6, we conclude the paper with some remarks.

2. Related work

Since the turn of the century, there has been an extensive research effort to analyze and design WMNs. The survey paper by Akyildiz et al. (2005) is a good source of information on WMNs. A more recent survey paper by Pathak and Dutta (2011) identified fundamental WMN design problems and surveyed recent research with a special focus on joint design efforts. This paper is an addition to that line of work as it jointly considers the design problems of gateway selection, routing, and time slot allocation. We report below a few studies on the IGP problem.

The IGP problem consists of determining the minimum number of gateway nodes and their locations to satisfy a given set of node demands. The problem can be viewed as a version of the capacitated facility location /network design problem (CFLNDP) introduced by Melkote and Daskin (2001a,b) with additional routing and link capacity constraints as in the network loading problem (NLP) defined in Magnanti et al. (1993). It can also be classified as a user-facility type general network design problem as in Contreras and Fernández (2012).

Chandra et al. (2004) formulated a network flow problem for gateway placement under multipath routing, and stated that it is NP-hard to find the minimum number of gateways to satisfy demand. They presented max-flow-min-cut-based greedy approximation algorithms for its solution. Aoun et al. (2006) and Bejerano (2004) considered quality-of-service requirements on delay and bandwidth while minimizing the number of gateways. They employed tree-based routing and proposed to divide the network into disjoint clusters covering all nodes. To meet delay requirements, an upper bound on either the cluster radius or the depth of the routing tree was considered. They presented the problem as variants of the capacitated facility location problem (CFLP) with additional constraints. While Aoun et al. tackled the combined problem, Bejerano split the problem into two sub-problems that are both NP-hard: The first problem was to form a small number of clusters that met the delay requirements and the second problem was to determine a gateway and a routing tree for each cluster such that the bandwidth requirement was met. Each cluster where the bandwidth requirement was not satisfied was further subdivided. Heuristic methods that run in polynomial time and yield near-optimal results were presented in both papers. Prasad and Wu (2006) considered a different objective function, one that minimizes the sum of the total interface costs at the gateways and the total hop counts of the nodes to reach their gateways. They proposed a recursive heuristic method that starts at a feasible solution and decreases cost at each iteration by a local search. He et al. (2008) also considered the gateway placement as a CFLP with additional constraints. They employed the objectives of minimizing both the number of gateways and the average number of hops to reach the gateways, and proposed two heuristic solution methods. A similar problem is considered in Seyedzadegan et al. (2013). Note that none of the papers above considered wireless interference in their models. Benyamina et al. (2009) considered a multiobjective problem that minimized both the gateway deployment cost and the congestion at the gateways. Multipath routing and the protocol interference introduced by Gupta and Kumar (2000) was employed in this work. To guarantee a bounded communication delay, they presented a tree-independent clustering approach, where each cluster is bounded by a maximal number of hops between its nodes. This approach was coupled with the solution of the multiobjective problem. Wu et al. (2009) formulated the gateway placement problem as a CFLP with the objective to minimize three factors: the

number of gateways, the average hop count to a gateway, and the variance of traffic at the gateways. To achieve better load balancing, each node was attached to multiple gateways and its traffic was split among them. They proposed a two-stage load-balanced gateway placement algorithm. Li et al. (2008) addressed the problem of maximizing traffic delivered to the Internet via a given number of gateway nodes. Assuming that the traffic of each node is infinitely divisible between multiple paths towards the gateways, they proposed a greedy algorithm for interference-free link scheduling. This algorithm was used as the evaluation tool for a grid-based gateway placement scheme.

Targon et al. (2010) were the first to present a mathematical formulation for the joint routing, scheduling, and gateway placement problem. While they proposed a similar formulation to ours, there are differences in terms of the objective function, the routing method, and the interference model employed: Their objective was to minimize the total gateway costs of a WMN where gateway interfaces of different capacities and costs were available. They employed single-path routing, but the formulation did not enforce a routing tree structure; hence, it was possible to leave a node on different links based on the source of traffic. Finally, they employed the binary conflict model to simplify the interference constraints. The more complicated physical interference model by Gupta and Kumar (2000) was employed by Capone et al. (2010) that considered the problem of minimizing the number of time slots needed to transfer data packets between node pairs. A column generation method was proposed to form non-interfering transmission sets, called configurations, that satisfy the physical interference constraints, and a simplified mathematical model was formulated based on these configurations. Our problem formulations in this paper are also based on these configurations. We assume, however, that these configurations are predetermined via the method presented in Gokbayrak and Yildirim (2017), which we discuss below.

Gokbayrak and Yildirim (2013) presented a mathematical formulation for the joint routing, transmission slot assignment, power control, and gateway placement problem under physical interference and static power control assumptions. The objective was to design a network with a given number of gateways that would meet the demand for the longest time, assuming that the demand would increase at the same percentage rate at all nodes. A set of valid inequalities based on interfering link pairs was derived that improved the solution performance of the mathematical model. Gokbayrak and Yildirim also presented a heuristic method to select the gateways and afterwards the original problem was solved with the gateways fixed. In Gokbayrak and Yildirim (2017), the authors modified the assumptions to employ dynamic power control. Instead of the column generation method proposed in Capone et al. (2010), Gokbayrak and Yildirim derived rules to determine non-interfering links (configurations) and their transmission power settings; consequently, the mathematical formulation was simplified by removing the power control and wireless interference components. The resulting NP-hard problem was formulated using non-interfering link sets (configurations) rather than individual links. A different set of valid inequalities was presented that improved the solution times. A three-step heuristic solution method for this NP-hard problem was also proposed: In the first step, the gateway locations are selected by solving a relaxed and simplified version of the problem. Then, in the second step, these gateways are fixed while determining the routing trees in a relaxed version. Finally, in the third step, a time slot allocation problem is solved to form the routing trees from the second step. Solving the routing and the time slot allocation problems separately for a given set of gateways improved the solution times by orders of magnitude in exchange for a slight decrease in the solution performances.

In this paper, we assume that non-interfering link sets and their corresponding power settings are already determined as in Gokbayrak and Yildirim (2017), hence no power control or interference constraints are included in the mathematical models. While Gokbayrak and Yildirim (2013, 2017) considered service level maximization problems for given number of gateways, in this paper, we consider the IGP problem to satisfy forecasted demand with the minimum number of gateway nodes. We also address the robustness issue that is ignored by all the IGP papers mentioned above. Since the actual demand may be different than the forecasted demand, rather than providing a solution for satisfying the forecasted demand, we search for the solution that is most resistant to under-forecast errors: Once the minimum number of gateways is determined, we consider a relaxed version of the service level maximization problem in Gokbayrak and Yildirim (2017) to obtain a more robust solution among the several solutions of the IGP problem. We employ a local search method to solve this relaxed joint gateway location and routing problem.

Our problem can be considered as a variant of the CFLNDP where link activation and flow costs are set to zero. Even though we do not have costs for activating links, we cannot activate any arbitrary combination of them as that may cause interference. Hence, the need for the network design arises. The flow capacities of links, on the other hand, are integer multiples of a capacity unit as in the NLP: We have a finite number of TDMA time slots that we allocate to links. The number of time slots assigned to a link determines its flow capacity. Note that we can assign the same time slot to multiple non-interfering links whereas NLP allows only one-to-one assignments. Finally, we also have single-path routing constraints, so our problem is considerably different than the ones considered in Melkote and Daskin (2001a) and Magnanti et al. (1993).

In the next section, we present two alternative mixed integer linear programming (MILP) formulations for the IGP problem. We also present a set of valid inequalities for the second formulation. The local search algorithm to obtain the most robust solution will be presented in the following section.

3. Problem formulation

We assume that the set of nodes N for the WMN, their locations, and their forecasted traffic demands are given. The links between these nodes are formed according to the physical interference model proposed in Gupta and Kumar (2000): Let P_i^{max} and l_{ij} represent the maximum transmission power of node i and the path loss between nodes i and j , respectively. The path loss model presented in Erceg et al. (1999) states that l_{ij} is inversely proportional to an exponential power of the distance between the transmitter node i and the receiver node j . This exponent depends on the terrain and the height of the antennas, and typically ranges around 3. Let $g_{ij} = P_i^{max} l_{ij}$ represent the signal power received at node j from node i when node i is transmitting at its maximum power level. In addition, let $\rho_i \in [0, 1]$ represent the ratio of the transmission power at node i to its maximum power, and η_j represent the ambient noise level at node j when all nodes are silent. According to the physical interference model, a link (i, j) of data rate r between nodes i and j can be activated if the following SINR exceeds γ_r , the threshold value for the data rate r :

$$\frac{\rho_i g_{ij}}{\eta_j + \sum_{n \in N \setminus \{i, j\}} \rho_n g_{nj}} \geq \gamma_r. \quad (1)$$

Note that signals received from other nodes cause interference, the summation term in the denominator, decreasing the ratio.

For simplicity, we assume that all wireless links are operated at the same data rate r , and that the ambient noise levels at all nodes are the same so that we drop the subscripts from γ and η .

Table 1
Problem parameters.

N	Set of nodes
E	Set of wireless links
S	Set of configurations
d_i	Uplink traffic demand of node i , $i \in N$
r	Data rate of a wireless link
a	Data rate of a wired link at a gateway node
T	Number of time slots in a TDMA frame

Consequently, we form the set of wireless links E as

$$E = \{(i, j) | i, j \in N, g_{ij} \geq \gamma\}$$

and define WMN as the directed graph (N, E) . Note that when node i is transmitting at its maximum power $\rho_i = 1$ and all other nodes are silent, i.e., $\rho_k = 0$ for all $k \neq i$, if (1) holds for a node j , i.e., $g_{ij} \geq \gamma$, then it is possible to activate the link (i, j) with such transmission power ratios. Since the wireless signal level decreases with distance, we can state that each node can form links with nodes within a coverage distance.

In addition to activating single links, under proper transmission power settings (as shown in Gokbayrak and Yildirim, 2017), several distant links in E can be activated simultaneously. Similar to Capone et al. (2010), we call these compatible sets of links that can be activated simultaneously as configurations and we denote the set of all configurations by $S \subset 2^E$. Note that since every link in E can be activated alone, single element subsets of E will be configurations. Moreover, for every configuration $\sigma \in S$ with multiple links, we also have all non-empty subsets of σ also in S .

The nodes of the WMN collect data flows both from the clients in their coverage areas and from the neighboring nodes, and deliver them in a multihop fashion to the Internet via gateway nodes. The traffic of each node is assumed to be routed to a gateway on a single path such that routing trees are formed, one for each gateway. We assume that each gateway node has a wired link of data rate a connecting to the Internet. To provide Internet access to all nodes of a WMN with G gateways and $|N|$ nodes, we need to form $|N| - G$ outgoing wireless links, because every node that is not a gateway has to send its traffic over a wireless link to a neighboring node. These links will constitute a network formed of trees whose roots are at the gateway nodes. As it is not possible to activate all $|N| - G$ wireless links of the trees simultaneously due to interference, we employ the spatial reuse TDMA method proposed in Nelson and Kleinrock (1985), in which each time frame is divided into T equal duration time slots that are allocated to configurations in S to deliver uplink traffic from the nodes to the gateways. Note that these configurations may have common links. The effective data rate of a link that is active in t time slots of a frame is calculated as tr/T . For G gateways, we need to select at most T configurations from S such that a) their union will form a routing forest of G trees with a total of $|N| - G$ wireless links, b) each one of these links should be allocated a sufficient number of time slots to carry its traffic flow.

Our objective is to determine the minimum number of gateways and their locations as well as the routing trees and the TDMA time slot allocations. For future reference, we list all the parameters of the problem in Table 1.

We present next two formulations for the IGP problem.

3.1. First formulation

Our first formulation is the one with single-commodity flow constraints. Let us start by defining the following decision variables that we employ in this formulation:

f_{ij} : the traffic flow on link $(i, j) \in E$;

ϕ_i : total traffic flow exiting the network at node $i \in N$;

x_σ : the number of time slots allocated to configuration $\sigma \in S$;

u_{ij} : the number of time slots in which link $(i, j) \in E$ is active;

$$z_{ij} = \begin{cases} 1, & \text{if link } (i, j) \text{ is utilized,} \\ 0, & \text{otherwise,} \end{cases}$$

$(i, j) \in E$; and

$$y_i = \begin{cases} 1, & \text{if node } i \text{ is a gateway,} \\ 0, & \text{otherwise,} \end{cases}$$

$i \in N$.

Next, we present our first MILP formulation, denoted by (MING-A):

$$\min \sum_{i \in N} y_i \quad (2)$$

subject to

$$d_i + \sum_{j: (j,i) \in E} f_{ji} = \phi_i + \sum_{j: (i,j) \in E} f_{ij}, \quad i \in N; \quad (3)$$

$$f_{ij} \leq \frac{r}{T} u_{ij}, \quad (i, j) \in E; \quad (4)$$

$$\phi_i \leq ay_i, \quad i \in N; \quad (5)$$

$$\sum_{j: (i,j) \in E} z_{ij} = 1 - y_i, \quad i \in N; \quad (6)$$

$$u_{ij} = \sum_{\sigma \in S: (i,j) \in \sigma} x_\sigma, \quad (i, j) \in E; \quad (7)$$

$$u_{ij} \leq Tz_{ij}, \quad (i, j) \in E; \quad (8)$$

$$\sum_{\sigma \in S} x_\sigma = T; \quad (9)$$

$$x_\sigma \in \{0, 1, 2, \dots, T\}, \quad \sigma \in S; \quad (10)$$

$$y_i \in \{0, 1\}, \quad i \in N; \quad (11)$$

$$z_{ij} \in \{0, 1\}, \quad (i, j) \in E; \quad (12)$$

$$u_{ij} \geq 0, \quad (i, j) \in E; \quad (13)$$

$$f_{ij} \geq 0, \quad (i, j) \in E; \quad (14)$$

$$\phi_i \geq 0, \quad i \in N. \quad (15)$$

Our objective in (2) is to minimize the number of gateways. The flow balance constraints are given in (3). The wireless link capacity constraints in (4) state that the wireless link capacities are proportional to the number of time slots in which the link is activated in each frame. The wired link capacity constraints in (5), on the other hand, ensure that traffic flow exits network only at the gateway nodes. For each node that is not a gateway, constraint (6) ensures a single outgoing link towards another node. The equality (7) defines the decision variables u_{ij} . If a link $(i, j) \in E$ is not on a routing tree, i.e., if $z_{ij} = 0$, then constraints (7) and (8) prevent all configurations including that link from being activated. The total number

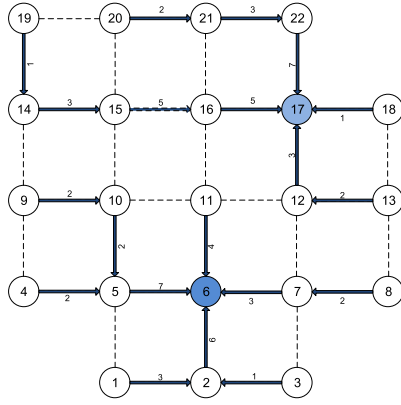


Fig. 1. A solution example.

Table 2
Size of model (MING-A).

Model	(MING-A)
# of constraints	$5 E + 4 N + 1$
# of continuous variables	$2 E + N $
# of binary variables	$ E + N $
# of general integer variables	$ S $

of TDMA time slots to be allocated to configurations is given in (9) as T . Constraints (10) to (15) define the integer, the binary, and the continuous decision variables.

The tasks performed by this model are to determine the minimum number of gateways to satisfy demand and their locations, to form the routing trees, and to allocate time slots to configurations in order to meet the demand. A solution example is given in Fig. 1 for a 22-node grid network. In this example, the arrows indicate the links (i, j) for which $z_{ij} = 1$, the numbers written next to these arrows are the u_{ij} values for these links, and the painted nodes indicate the gateway nodes for which $y_i = 1$. Note that Nodes 1–11 are served by the gateway at Node 6 to which they are connected on a tree topology. Similarly, Nodes 12–22 are served by the gateway at Node 17 to which they are connected also on a tree topology. Even though the network is partitioned into two trees, the whole network has to be solved at once as the links share the common resource that is the time slots in a TDMA frame.

Table 2 presents the size of the model in terms of the number of constraints and the decision variables.

From the flow balance (3) and the wired link capacity (5) constraints, we obtain

$$\sum_{i \in N} y_i \geq \frac{\sum_{i \in N} d_i}{a}, \quad (16)$$

a lower bound for the linear programming (LP) relaxation of (MING-A).

3.2. Second formulation

Our second formulation is the one with multi-commodity flow constraints. Let us start by defining the following binary decision variables that we employ in this formulation:

$$f_{ij}^k = \begin{cases} 1, & \text{if link } (i, j) \text{ is utilized to carry traffic from node } k, \\ 0, & \text{otherwise,} \end{cases} \quad (i, j) \in E, k \in N; \text{ and}$$

$$\phi_i^k = \begin{cases} 1, & \text{if node } k\text{'s traffic exits network at node } i, \\ 0, & \text{otherwise,} \end{cases} \quad i \in N, k \in N.$$

The decision variables y_i , ϕ_i , z_{ij} , f_{ij}^k of the first formulation can be written in terms of the binary variables of the second formulation as follows:

$$y_i = \phi_i^i, \quad i \in N; \quad (17)$$

$$\phi_i = \sum_{k \in N} \phi_i^k d_k, \quad i \in N; \quad (18)$$

$$z_{ij} = f_{ij}^i, \quad (i, j) \in E; \quad (19)$$

$$f_{ij} = \sum_{k \in N} f_{ij}^k d_k, \quad (i, j) \in E. \quad (20)$$

In addition to the decision variables x_σ and u_{ij} from the first formulation, we employ the newly defined binary decision variables f_{ij}^k and ϕ_i^k in our second MILP formulation, denoted by (MING-B), that we present below:

$$\min \sum_{i \in N} \phi_i^i \quad (21)$$

subject to (7), (9), (10), (13), and

$$\sum_{j: (j, k) \in E} f_{jk}^k = 0, \quad k \in N; \quad (22)$$

$$\sum_{j: (j, i) \in E} f_{ji}^k = \phi_i^k + \sum_{j: (i, j) \in E} f_{ij}^k, \quad i, k \in N : i \neq k; \quad (23)$$

$$1 = \phi_k^k + \sum_{j: (k, j) \in E} f_{kj}^k, \quad k \in N; \quad (24)$$

$$\sum_{k \in N} f_{ij}^k d_k \leq \frac{r}{T} u_{ij}, \quad (i, j) \in E; \quad (25)$$

$$\sum_{k \in N} \phi_i^k d_k \leq a \phi_i^i, \quad i \in N; \quad (26)$$

$$u_{ij} \leq T f_{ij}^i, \quad (i, j) \in E; \quad (27)$$

$$f_{ij}^k \in \{0, 1\}, \quad (i, j) \in E, k \in N; \quad (28)$$

$$\phi_i^k \in \{0, 1\}, \quad i \in N, k \in N. \quad (29)$$

In this formulation, the minimization objective (21) is derived from the objective of (MING-A) in (2) and from (17). Constraints (22) state that the traffic of a node does not return back to that node. Constraints (23) are the intermediate node flow balance constraints. Constraints (24) are derived from the single path constraints (6) and from (17) and (19), and they are the flow balance constraints at the source nodes. Constraints (25) are the wireless link capacity constraints derived from the constraints (4) and from

Table 3
Size of model (MING-B).

Model	(MING-B)
# of constraints	$4 E + 2 N + N ^2 + 1$
# of continuous variables	$ E $
# of binary variables	$ E N + N ^2$
# of general integer variables	$ S $

(20). The wired link capacity constraints (26), on the other hand, are derived from the constraints (5) and from (17) and (18). If a node i is not employing its link (i, j) on its path to the gateway, then that link is never activated as in constraints (27) that are derived from the constraints (8) and from (19). Note that this set of constraints is added to enforce a tree structure along with the set of constraints (25). Finally, (28) and (29) define the binary flow variables.

If we compare the constraints of (MING-A) and (MING-B), we observe that $|N|$ flow balance constraints of (MING-A) in (3) are replaced by $|N|^2$ flow balance constraints in (22) and (23). The rest of the constraints of (MING-A) are preserved but rewritten in terms of the newly defined binary variables. Similarly, $|N| + |E|$ binary and $|N| + |E|$ continuous variables of (MING-A) are replaced by $|E||N| + |N|^2$ binary variables in (MING-B). Table 3 presents the size of the model in terms of the number of constraints and the decision variables.

From the flow balance equations (22)–(24) and the wired link capacity constraints (26), we again obtain the inequality (16) as a lower bound for the LP relaxation of (MING-B).

To improve the (MING-B) formulation, we propose the following set of valid inequalities:

$$\phi_i^k \leq \phi_i^j, \quad i, k \in N. \quad (30)$$

The set of valid inequalities (30) states that the traffic of each node can exit network only at a gateway node. Note that for all nodes $k \in N$ with a positive traffic demand $d_k > 0$, we can make the same statement by the wired link capacity constraints in (26).

In Section 5, we solve the gateway placement problem for some example networks using our alternative formulations and the valid inequalities. LP relaxations are also reported for comparison purposes.

4. Selecting among alternative solutions

The solution to the IGP problem provides a gateway set of minimum cardinality that can satisfy forecasted node demands along with the corresponding routing trees and time slot allocations. In general, there may be multiple solutions to the IGP problem, i.e., many different gateway sets of the minimum cardinality can satisfy demand. In this section, we search for the solution among them that minimizes the maximum utilization on the links. (We define the utilization on any link (i, j) as $\frac{rf_{ij}}{ru_{ij}}$, the ratio of total flow on the link to its allocated capacity.) Let us denote the maximum utilization on any link by ρ_{max} . Then, that gateway set along with the routing trees and the time slot allocations will suffice even when the node demands are realized as up to $1/\rho_{max}$ times the forecasted amount. Note that this min-max problem was also considered in Gokbayrak and Yildirim (2017) for a given number of gateways. The heuristic algorithm proposed therein started from a random gateway set and applied a local search method to solve a relaxed version that considered only the gateway location component. In this paper, we employ another relaxed version considering both the gateway location and the routing components, and propose a local search algorithm that starts from the solution to the IGP problem.

Table 4
Size of model (MAXW).

Model	(MAXW)
# of constraints	$6 E + 4 N + 2$
# of continuous variables	$2 E + N $
# of binary variables	$ E + N $
# of general integer variables	$ S $

For the sake of completeness, we start by presenting the optimization problem proposed in Gokbayrak and Yildirim (2017) to minimize the maximum utilization. Let w denote the demand multiplier that we want to maximize while satisfying the multiplied demand. The maximum utilization on any link will then be $\rho_{max} = 1/w$ under the original node demands. The problem of maximizing w , (MAXW), considered in Gokbayrak and Yildirim (2017) is defined as follows:

$$\max w \quad (31)$$

subject to (4)–(15), and

$$wd_i + \sum_{j:(j,i) \in E} f_{ji} = \phi_i + \sum_{j:(i,j) \in E} f_{ij}, \quad i \in N; \quad (32)$$

$$u_{ij} \geq \left\lceil \frac{d_i T}{r} \right\rceil z_{ij}, \quad (i, j) \in E; \quad (33)$$

$$\sum_{i \in N} y_i = g, \quad (34)$$

where g is a given number of gateways. Note that (MAXW) employs single-commodity flow constraints like (MING-A) because multi-commodity flow formulations like (MING-B) already have capacity constraints like (25) and (26) that have decision variables multiplied with demand values. Multiplying with another decision variable, w , would make those constraints nonlinear.

Table 4 presents the size of the optimization model in terms of the number of constraints and the decision variables.

To obtain more robust gateway sets, we can set the parameter g as the objective value of the IGP solution and solve (MAXW). Note that, for small networks, we can solve (MAXW) exactly. For large networks, however, we propose to solve the following relaxed version of (MAXW) employing the k-opt (also known as k-change) local search algorithm that starts from the solution to the IGP problem.

4.1. A relaxed version of (MAXW)

In all the formulations we presented so far, the capacity on each wireless link is constrained to be an integer multiple of r/T . Due to this constraint, we end up with unused capacities on several wireless links. To remove this constraint, we propose to relax the general integer variables x_{σ} in (MAXW) to become continuous and to re-scale x_{σ} and u_{ij} variables by dividing by T . Afterwards, the wireless link capacity constraints in (4) become

$$f_{ij} \leq ru_{ij}, \quad (i, j) \in E. \quad (35)$$

These constraints should be active (satisfied with equality) in the optimal solution, because any excess capacity on a wireless link can be distributed among others to increase w . Hence, we replace the flow variables f_{ij} in (MAXW) with ru_{ij} and obtain the following relaxed optimization model, denoted by (MAXWR):

$$\max w \quad (36)$$

subject to (5)–(7), (11)–(13), (15), (34) and

$$wd_i + r \sum_{j:(j,i) \in E} u_{ji} = \phi_i + r \sum_{j:(i,j) \in E} u_{ij}, \quad i \in N; \quad (37)$$

$$u_{ij} \leq z_{ij}, \quad (i, j) \in E; \quad (38)$$

$$u_{ij} \geq \frac{d_i}{r} z_{ij}, \quad (i, j) \in E; \quad (39)$$

$$\sum_{\sigma \in S} x_{\sigma} = 1; \quad (40)$$

$$x_{\sigma} \geq 0, \quad \sigma \in S. \quad (41)$$

Note that this model is independent of the number of time slots T , and it provides both a set of gateway nodes and the routing trees rooted at these gateway nodes. The objective value of the solution to (MAXWR) is an upper bound on the objective value of the solution to (MAXW). We propose to solve the easier problem (MAXWR) as a heuristic method to obtain a more robust gateway set.

Fischetti and Lodi (2003) introduce the notion of “local branching”. In this scheme, given a binary vector $y \in \{0, 1\}^N$ and a positive integer $k \leq N$, the authors define a k -opt neighborhood as the set of all binary vectors which differ from y in at most k components. Local branching is proposed as a branching criterion, in which, given an incumbent solution y , the solution space is partitioned into two sets with respect to solutions that are in the k -opt neighborhood of y and the remaining solutions. As we have to satisfy the constraint (34), the number of changes in y from zero to one should equal to the number of changes from one to zero. Hence, we adopt a different k -opt neighborhood definition for the binary vector y in (MAXWR) corresponding to gateway locations defined as

$$\mathcal{N}_k(y) = \left\{ \bar{y} \mid \sum_{i: y_i=1} \bar{y}_i \geq g - k, \sum_{i=1}^N \bar{y}_i = g \right\}. \quad (42)$$

For large networks, we apply the k -opt local search algorithm to solve (MAXWR): We start with the gateway set G obtained from the IGP solution. At each iteration of this local search algorithm, we add the constraint

$$\sum_{i \in G} y_i \geq g - k \quad (43)$$

to (MAXWR) and solve. If the objective value increases, we update the gateway set G and repeat until we observe no increase in the objective value.

In the next section, we demonstrate how much more traffic can be accommodated if the additional steps presented in this section are carried out.

5. Computational results

In this section, we consider three networks denoted by A, B, and C. These networks are of grid topology consisting of node configurations of 5×8 , 7×7 , and 9×7 , respectively, as shown in Figs. 2–4. The vertical and horizontal distances between any two neighboring nodes are assumed to be 1 km. Note that Network A and Network B were also considered in Gokbayrak and Yildirim (2013) as example networks. In this paper, we also consider the larger example network, Network C.

In these networks, each gateway connection is assumed to have a data rate of $a = 45$ Mbps as of a T3 cable connection, and each link has a data rate of $r = 24$ Mbps corresponding to the SINR threshold value of $\gamma = 50$. The ambient noise level η is assumed to be 1.5×10^{-10} mW at all nodes. Under the assumptions that each node has a maximum transmission power of 15 mW (measured at

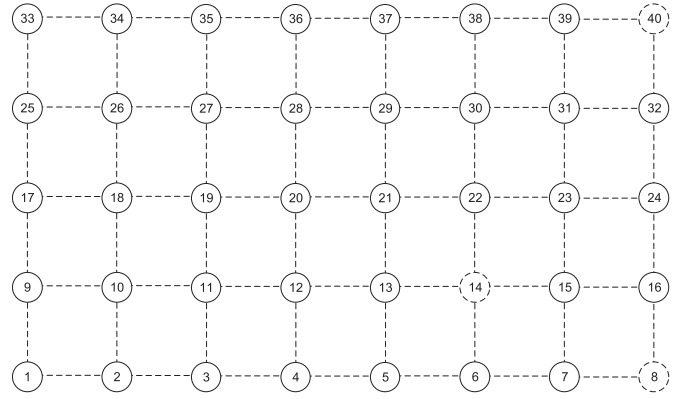


Fig. 2. Topology of Network A.

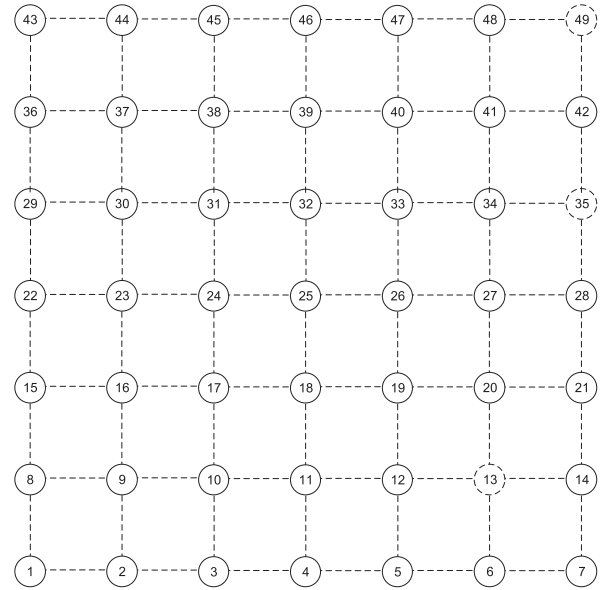


Fig. 3. Topology of Network B.

Table 5
Parameter values.

SINR threshold value	$\gamma = 50$
Ambient noise level	$\eta = 1.5 \times 10^{-10}$ mW
Maximum transmission power	15 mW
Number of TDMA time slots	$T \in \{64, 128\}$
Data rate of a wireless link	$r = 24$ Mbps
Data rate of a wired link at a gateway node	$a = 45$ Mbps

Table 6
Characteristics of networks.

Network	$ N $	$ E $	$ S $	One link	Two links	Three links	Four links
A	40	134	1430	134	1296	–	–
B	49	168	2716	168	2436	112	–
C	63	220	19,504	220	6960	11,932	392

a meter away from the transmitter) and that the signal power is inversely proportional to the cube of the distance in meters, this setup allows wireless links that are at most 1.26 km long. Hence, the network topologies are formed as grids. The number of time slots in each TDMA frame, T , is assumed to take values from the set $\{64, 128\}$. We summarize our parameters in Table 5.

We assume that the set of configurations for each network is available (determined via the calculation method in Gokbayrak and Yildirim (2017)). In Table 6, we present the number of nodes $|N|$,

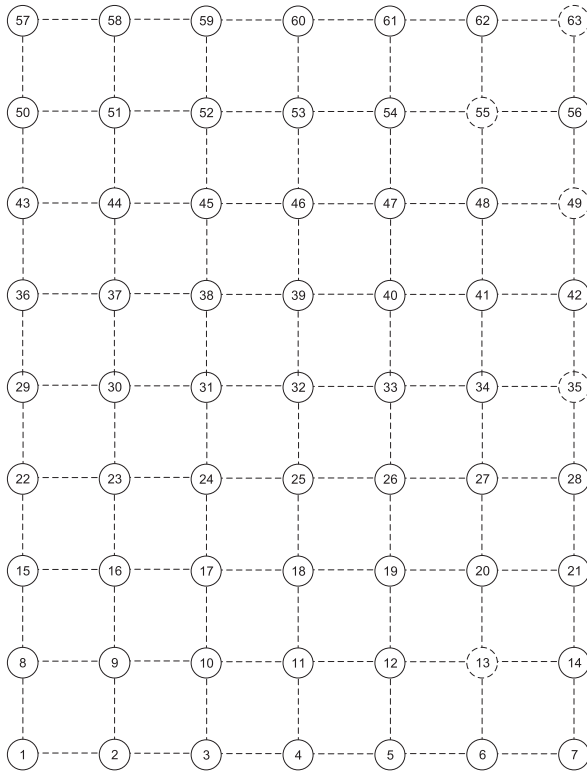


Fig. 4. Topology of Network C.

the number of links $|E|$, and the number of configurations $|S|$ for each network. At most two links in Network A, at most three links in Network B, and at most four links in Network C can be activated simultaneously. We also list in Table 6 the number of configurations composed of one link, two links, three links, and four links, separately.

All the models in this section are solved on a workstation with dual 2.4 Ghz Intel Xeon E5-2630v3 CPUs and 64GB RAM running CPLEX 12.7 solver. We impose a time limit of twenty hours (72,000 s) of CPU time, unless stated otherwise.

In the following subsection, we compare our alternative formulations for the IGP problem.

Table 7
Number of unsolved instances within the time limit.

Network	(MING-A)		(MING-B)		(MING-B) w/ (30)	
	$T = 64$	$T = 128$	$T = 64$	$T = 128$	$T = 64$	$T = 128$
A	6	4	0	0	0	0
B	18	11	0	0	0	0
C	30	30	6	6	7	6

5.1. Comparison of (MING-A) and (MING-B) formulations

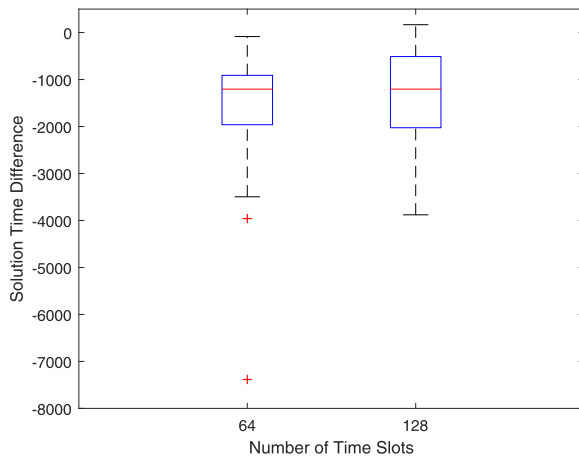
In this subsection, we solve the IGP problem for each network under thirty randomly generated node demand sets. The node demands are sampled from the set $\{0.1, 0.2, \dots, 1.0\}$ uniformly. For each demand set, we select T as 64 and 128 to generate a total of sixty instances for each network.

In Table 7, we report the number of instances for each formulation that cannot be solved within the time limit. While (MING-B) with and without the valid inequalities could be solved for all instances of Networks A and B, (MING-A) could not be solved to optimality for ten instances of Network A and twenty-nine instances of Network B within the time limit. Moreover, (MING-A) could not be solved to optimality within the time limit for any instance of Network C. Therefore, we declare (MING-A) as the inferior formulation among the three alternatives.

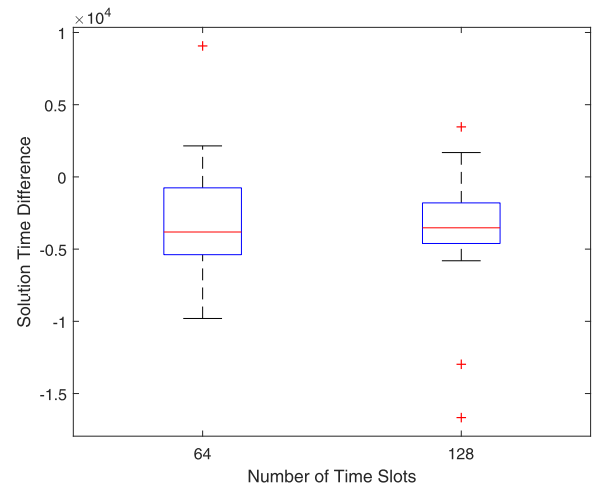
To evaluate the benefit of the valid inequalities (30) on instances of Networks A and B, we determine the solution time differences in CPU seconds of (MING-B) with and without the valid inequalities. We report these differences, which are calculated by subtracting the solution time of (MING-B) from the solution time of (MING-B) with (30) for each instance, as boxplots in Fig. 5. On each box, the central mark indicates the median, the bottom and the top edges of the box indicate the first and the third quartiles, respectively. The whiskers extend to the most extreme data points that are not considered outliers, and the outliers are plotted individually using the '+' symbol. Note that data points that are more than 1.5 times the inter-quartile range above the third quartile or below the first quartile are considered outliers.

Fig. 5 shows that for most of the instances of Networks A and B, (MING-B) with (30) obtained optimal solutions in shorter times.

As for Network C, where both (MING-B) formulations had instances that are not solved to optimality within the time limit, we determined the solution time differences under no time limit. We



(a) Network A



(b) Network B

Fig. 5. Solution time differences in CPU seconds.

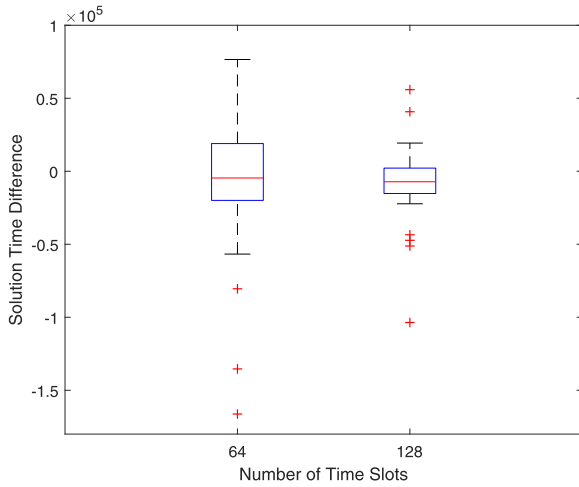


Fig. 6. Network C solution time differences in CPU seconds.

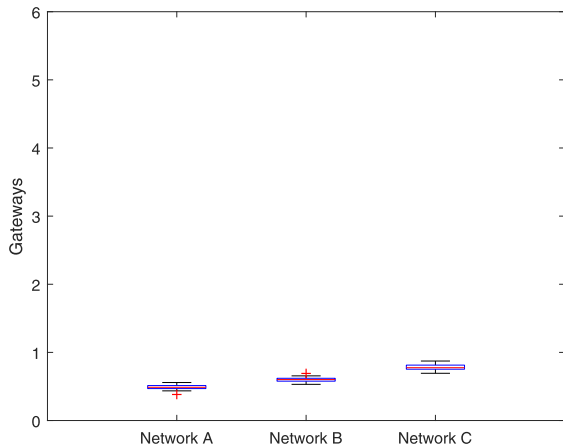
report these differences that are calculated the same way as explained above as boxplots in Fig. 6.

Fig. 6 suggests that the benefit from the valid inequalities (30) diminishes as the number of nodes in the network increases. A possible explanation for this observation is that the increase in the problem size due to additional $|N|^2$ constraints outweigh the benefit from these valid inequalities when the number of nodes $|N|$ is large.

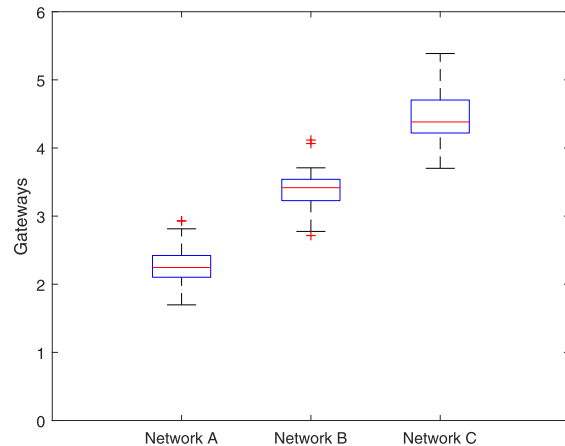
In Fig. 7, we report the LP relaxation objectives of the instances of Networks A, B, and C for the formulation (MING-B) with and without the valid inequalities (30) as boxplots. (Note that the LP relaxation objectives of (MING-A) are equal to the LP relaxation objectives of (MING-B) without the valid inequalities for all these instances.) As the number of time slots in a frame does not affect the LP relaxation objectives, we do not specify a T value in this figure.

The valid inequalities in (30) increase the LP relaxation objectives drastically for (MING-B) suggesting that the addition of these valid inequalities results with a stronger formulation.

We select the (MING-B) with the valid inequalities (30) as the best formulation based on its overall solution performance. Next, we select a set of node demands for each network as presented in Table 8, and show how robust gateway placement is carried out for these examples.



(a) Without Valid Inequalities



(b) With Valid Inequalities

Fig. 7. (MING-B) linear programming relaxation objectives.

5.2. Robust gateway placement

5.2.1. Determining the minimum number of gateways

In this subsection, we solve the IGP problem to determine the minimum number of gateways needed to support the node demands presented in Table 8 and to obtain an initial gateway set for the k -opt local search algorithm that follows. For Networks A–C, we solve (MING-B) with the valid inequalities (30). The gateway nodes in the solutions and the solution times in CPU seconds are presented in Table 9.

While Network A needs two gateways, Network B needs four gateways for both $T = 64$ and $T = 128$. Network C, on the other hand, needs seven gateways for $T = 64$ and six gateways for $T = 128$.

In the next subsection, we employ the k -opt local search algorithm to improve the objective values of (MAXWR) starting from the gateway sets determined by (MING-B) with (30).

5.2.2. Obtaining more robust gateway sets

Before we proceed with our local search algorithm, for referencing purposes, we present the (MAXWR) solutions under our time limit for the three networks in Table 10. This table presents the best integer solutions, the best bounds, and the solution times in CPU seconds. Note that we have two rows for Network C because we obtained two different numbers of gateways for $T = 64$ and $T = 128$. While (MAXWR) can be solved quickly for Networks A and B, the time limit was too short to obtain an optimal solution for Network C. Even though the k -opt local search algorithm is needed only for Network C, we apply it to all networks for illustrative purposes.

In our local search algorithm, selecting a larger k value increases the search space and slows down the solution process. Therefore, to obtain results quickly, we initially employ $k = 1$. We proceed with larger k values when the local search algorithm gets stuck. We can incrementally increase the k value up to $g - 1$; however, the solution times for iterations are increased an order of magnitude as we increment the k value. To keep the solution times within reasonable limits, we suggest not to increase k beyond two.

5.2.2.1. Network A. Starting from the initial gateway set for $T = 64$ from Table 9, we apply the 1-opt local search algorithm. In Table 11, we present the gateway nodes, the (MAXWR) objective values, k values, and the solution times in CPU seconds for each iteration of the local search algorithm. The objective value for the initial gateway set is also presented in iteration 0.

Table 8
Node demands in Mbps.

Nodes	Network A	Network B	Network C
1	0.6	0.6	0.4
2	0.6	0.6	0.3
3	0.1	0.1	0.5
4	0.4	0.4	0.1
5	0.7	0.7	0.2
6	0.4	0.4	1.0
7	0.3	0.3	1.0
8	0.3	0.3	0.6
9	0.3	0.3	0.1
10	0.1	0.9	0.3
11	0.8	0.8	0.4
12	0.2	0.2	0.9
13	0.3	0.3	0.1
14	0.4	0.4	0.1
15	0.3	0.3	0.2
16	0.2	0.2	0.7
17	0.6	0.6	0.8
18	0.2	0.2	0.7
19	0.2	0.2	0.5
20	0.3	0.3	0.6
21	0.3	0.3	0.3
22	0.9	0.9	0.8
23	0.6	0.6	0.2
24	0.6	0.6	0.7
25	0.1	0.9	0.2
26	0.4	0.4	0.4
27	0.7	0.7	0.7
28	0.4	0.4	0.8
29	0.3	0.3	0.1
30	0.3	0.3	1.0
31	0.3	0.3	0.8
32	0.1	0.1	0.5
33	0.5	0.8	0.5
34	0.2	0.2	0.5
35	0.3	0.3	0.4
36	0.4	0.4	0.6
37	0.3	0.3	0.6
38	0.2	0.2	0.9
39	0.6	0.6	0.8
40	0.1	0.1	0.7
41		0.6	0.4
42		0.6	0.9
43		0.9	0.6
44		0.4	0.4
45		0.7	1.0
46		0.4	0.9
47		0.3	0.6
48		0.3	0.7
49		0.3	0.6
50			0.3
51			0.4
52			0.5
53			0.3
54			0.9
55			0.2
56			0.3
57			0.2
58			0.3
59			0.5
60			0.4
61			1.0
62			0.5
63			0.2

Table 9
Solution times and gateway nodes obtained by (MING-B) with (30).

Network	$T = 64$		$T = 128$	
	Gateways	Time	Gateways	Time
A	16, 18	270	22, 25	298
B	4, 23, 27, 45	986	10, 27, 34, 43	24,066
C	12, 15, 35, 38, 47, 51, 61	24,422	6, 16, 35, 38, 54, 58	13,362

Table 10
(MAXWR) solutions.

Network	g	Solution	Bound	Time	Gateways
A	2	1.3260	1.3260	28	18, 23
B	4	1.2903	1.2903	2893	4, 23, 27, 45
C	6	1.1321	2.8252	72,000	13, 16, 31, 42, 44, 61
C	7	1.2355	3.2352	72,000	6, 18, 22, 32, 42, 45, 61

Table 11
Local search iterations for Network A with $T = 64$.

Iteration	Gateways	Objective	k	Time
0	16, 18	1.1823	0	1
1	18, 23	1.3260	1	5
2	18, 23	1.3260	1	5

Table 12
Local search iterations for Network A with $T = 128$.

Iteration	Gateways	Objective	k	Time
0	22, 25	1.2834	0	1
1	22, 26	1.2903	1	6
2	23, 26	1.3260	1	7
3	23, 26	1.3260	1	8

Table 13
(MAXW) objectives for Network A.

k	Gateways	$T = 64$	$T = 128$
0	16, 18	1.0547	–
1	18, 23	1.1250	–
0	22, 25	–	1.1932
1	23, 26	–	1.2500

Similarly, in Table 12, starting from the initial gateway set for $T = 128$ from Table 9, we apply the 1-opt local search algorithm to improve the objective value of (MAXWR) for this network.

Tables 11 and 12 show that two different gateway sets are obtained by the 1-opt local search algorithm starting from different initial sets. Based on the solution for Network A in Table 10, we can claim that both are optimal solutions for (MAXWR). The time it takes to obtain these gateway sets by the local search algorithm is comparable to the time it takes to solve (MAXWR) directly; therefore, we do not recommend the local search algorithm for small networks.

In Table 13, we present the (MAXW) objectives for the initial and the final gateway sets for each k value from Tables 11 and 12. Note that all these gateway sets can satisfy forecasted node demands as their (MAXW) objective values exceed one. However, as far as robustness is concerned, our proposed method obtained better gateway sets.

We could also solve (MAXW) for Network A for both $T = 64$ and $T = 128$, and confirmed that the 1-opt method returned the optimal gateway sets for (MAXW).

5.2.2.2. Network B. Starting from the initial gateway set for $T = 64$ from Table 9, we apply the k -opt local search algorithm with increasing k values in Table 14, which is organized the same way as Tables 11 and 12. We also present entries with $k = 3$ to show the drastic increase in the solution times.

Similarly, starting from the initial gateway set for $T = 128$, in Table 15, we apply the k -opt local search algorithm with increasing k values to improve the objective value of (MAXWR).

As shown in Tables 14 and 15, 3-opt steps may take too long to include in a heuristic algorithm. As Table 15 illustrates, the local

Table 14Local search iterations for Network B with $T = 64$.

Iteration	Gateways	Objective	k	Time
0	4, 23, 27, 45	1.2903	0	1
1	4, 23, 27, 45	1.2903	1	25
2	4, 23, 27, 45	1.2903	2	192
3	4, 23, 27, 45	1.2903	3	2309

Table 15Local search iterations for Network B with $T = 128$.

Iteration	Gateways	Objective	k	Time
0	10, 27, 34, 43	1.1650	0	1
1	10, 24, 34, 43	1.2403	1	21
2	10, 24, 34, 43	1.2403	1	24
3	11, 16, 34, 43	1.2533	2	210
4	11, 22, 34, 45	1.2698	2	213
5	11, 22, 35, 45	1.2732	2	212
6	11, 22, 35, 45	1.2732	2	222
7	4, 23, 27, 45	1.2903	3	1984
8	4, 23, 27, 45	1.2903	3	2439

Table 16

(MAXW) objectives for Network B.

k	Gateways	$T = 64$	$T = 128$
0	4, 23, 27, 45	1.1538	–
0	10, 27, 34, 43	–	1.0714
1	10, 24, 34, 43	–	1.1413
2	11, 22, 35, 45	–	1.1932
3	4, 23, 27, 45	–	1.2054

search algorithm may take longer than solving (MAXWR) directly if $k = 3$ iterations are included. On the other hand, 3-opt steps obtained a better solution for Network B with $T = 128$ than 2-opt steps in return for longer iteration times.

In Table 16, we present the (MAXW) objectives for the initial and the final gateway sets for each k value from Tables 14 and 15. Note that all these gateway sets can satisfy forecasted node demands as their (MAXW) objective values exceed one. However, as far as robustness is concerned, our local search algorithm obtained better gateway sets.

We could also solve (MAXW) for Network B for both $T = 64$ and $T = 128$, and confirmed that the gateway set {4, 23, 27, 45} is an optimal gateway set for this network for both T values.

5.2.2.3. Network C. Starting from the initial gateway set for $T = 64$ from Table 9, we apply the k -opt local search algorithm with increasing k values in Table 17, which is organized the same way as Tables 11, 12, 14, and 15. We also present entries with $k = 3$ to show the drastic increase in the iteration times.

Similarly, starting from the initial gateway set for $T = 128$ from Table 9, we apply the k -opt local search algorithm with increasing k values in Table 18.

Note that, for both cases of T , 1-opt local search algorithm could determine better solutions to (MAXWR) than the ones presented in Table 10 that are obtained under the time limit of 72,000 s.

In Table 19, we present (MAXW) objectives for the initial and the final gateway sets for each k value from Tables 17 and 18. Even when we specify the gateway nodes, we cannot solve (MAXW) to optimality within the time limit for this network. Hence, we solved them without any time limits so that we can comment on the improvements.

For $T = 64$, 1-opt improved both (MAXW) and (MAXWR) objectives, whereas 2-opt did not provide any additional improvements on the (MAXW) objective even though it further improved the

Table 17Local search iterations for Network C with $T = 64$.

Iteration	Gateways	Objective	k	Time
0	12, 15, 35, 38, 47, 51, 61	1.1940	0	4
1	12, 15, 31, 35, 47, 51, 61	1.2427	1	259
2	12, 15, 31, 35, 44, 47, 61	1.2598	1	293
3	12, 15, 31, 35, 44, 47, 61	1.2598	1	211
4	12, 15, 31, 35, 43, 46, 61	1.2721	2	2323
5	12, 15, 31, 35, 43, 46, 61	1.2721	2	2742
2	12, 15, 31, 35, 43, 46, 61	1.2721	3	36111

Table 18Local search iterations for Network C with $T = 128$.

Iteration	Gateways	Objective	k	Time
0	6, 16, 35, 38, 54, 58	1.1307	0	4
1	6, 16, 35, 38, 50, 54	1.1344	1	224
2	6, 16, 35, 38, 50, 54	1.1344	1	245
3	6, 16, 35, 38, 50, 54	1.1344	2	2194
4	6, 16, 35, 38, 50, 54	1.1344	3	29671

Table 19

(MAXW) objectives for Network C.

k	Gateways	$T = 64$	$T = 128$
0	12, 15, 35, 38, 47, 51, 61	1.0227	–
1	12, 15, 31, 35, 44, 47, 61	1.0714	–
2	12, 15, 31, 35, 43, 46, 61	1.0714	–
0	6, 16, 35, 38, 54, 58	–	1.0417
1	6, 16, 35, 38, 50, 54	–	1.0417

(MAXWR) objective. For $T = 128$, even though 1-opt improved the (MAXWR) objective slightly, (MAXW) objective was not improved. Consequently, we suspected that we may have already reached the optimal solutions for (MAXW). To check the optimality of our final gateway sets, we multiplied the demand values for Network C in Table 8 with the best feasible objective values in Table 19 plus an ϵ , which we take as 0.0001, and solved (MING-B) with (30) for both instances. Then, it turned out that we needed 8 gateways to support the demand multiplied by 1.0715 under $T = 64$, and 7 gateways to support the demand multiplied by 1.0418 under $T = 128$. Hence, we conclude that the final gateway sets are indeed the optimal gateway sets for (MAXW) under $T = 64$ and $T = 128$.

6. Conclusions

The IGP problem determines a set of gateways of minimum cardinality that can satisfy node demands. It also determines the routing trees and the TDMA time slot allocations to configurations. We presented two alternative formulations for the IGP problem and proposed a set of valid inequalities. These valid inequalities increased the LP relaxation objective values considerably.

There may be several different gateway sets of the same cardinality that can satisfy demand. While some of these gateway sets can barely satisfy demand, others may accommodate extra demand. Since the IGP problem is generally solved based on forecasted demand, it is important to determine the gateway set of the same cardinality that is the most robust to any forecasting errors. We introduced an additional step to obtain more robust gateway sets and demonstrated its benefit by computational results. This additional step employed a k -opt local search algorithm starting from the IGP solution. While the 1-opt local search algorithm could obtain quite satisfactory results, higher k values could obtain better results at the expense of longer solution times that are increased by orders of magnitude as k increases.

An alternative objective to employ while selecting among alternative gateway sets is to deliver forecasted traffic to Internet in the shortest time, and that is the topic of our ongoing research.

Acknowledgment

We gratefully acknowledge the insightful comments and suggestions of three anonymous reviewers and the editors, which led to several improvements in our manuscript.

References

- Akyildiz, I.F., Wang, X., Wang, W., 2005. Wireless mesh networks: a survey. *Comput. Netw.* 47 (4), 445–487.
- Amaldi, E., Capone, A., Cesana, M., Filippini, I., Malucelli, F., 2008. Optimization models and methods for planning wireless mesh networks. *Comput. Netw.* 52 (11), 2159–2171.
- Aoun, B., Boutaba, R., Iraqi, Y., Kenward, G., 2006. Gateway placement optimization in wireless mesh networks with QoS constraints. *IEEE J. Sel. Areas Commun.* 24 (11), 2127–2136.
- Bejerano, Y., 2004. Efficient integration of multihop wireless and wired networks with QoS constraints. *IEEE/ACM Trans. Netw.* 12 (6), 1064–1078.
- Benyamina, D., Hafid, A., Gendreau, M., 2009. Optimal placement of gateways in multi-hop wireless mesh networks: a clustering-based approach. In: *Proceedings of 34th IEEE Conference on Local Computer Networks*, pp. 625–632.
- Capone, A., Carello, G., Filippini, I., Gualandi, S., Malucelli, F., 2010. Routing, scheduling and channel assignment in wireless mesh networks: optimization models and algorithms. *Ad Hoc Netw.* 8, 545–563.
- Chandra, R., Qiu, L., Jain, K., Mahdian, M., 2004. Optimizing the placement of internet TAPs in wireless neighborhood networks. In: *Proceedings of the 12th IEEE International Conference on Network Protocols*, pp. 271–282.
- Chen, W., Lea, C.-T., 2013. A node-based time slot assignment algorithm for STDMA wireless mesh networks. *IEEE Trans. Veh. Technol.* 62 (1), 272–283.
- Contreras, I., Fernández, E., 2012. General network design: a unified view of combined location and network design problems. *Eur. J. Oper. Res.* 219 (3), 680–697.
- Djukic, P., Valaee, S., 2007. Link scheduling for minimum delay in spatial re-use TDMA. In: *INFOCOM 2007. 26th IEEE International Conference on Computer Communications*. IEEE, pp. 28–36.
- Erceg, V., Greenstein, L.J., Tjandra, S.Y., Parkoff, S.R., Gupta, A., Kulic, B., Julius, A.A., Bianchi, R., 1999. An empirically based path loss model for wireless channels in suburban environments. *IEEE J. Sel. Areas Commun.* 17 (7), 1205–1211.
- Fischetti, M., Lodi, A., 2003. Local branching. *Math. Program.* 98 (1–3), 23–47.
- Gokbayrak, K., Yildirim, E.A., 2013. Joint gateway selection, transmission slot assignment, routing and power control for wireless mesh networks. *Comput. Oper. Res.* 40 (7), 1671–1679.
- Gokbayrak, K., Yildirim, E.A., 2017. Exact and heuristic approaches based on noninterfering transmissions for joint gateway selection, time slot allocation, routing and power control for wireless mesh networks. *Comput. Oper. Res.* 81, 102–118.
- Gupta, P., Kumar, P.R., 2000. The capacity of wireless networks. *IEEE Trans. Inf. Theory* 46 (2), 388–404.
- He, B., Xie, B., Agrawal, D.P., 2008. Optimizing deployment of internet gateway in wireless mesh network. *Comput. Commun.* 31, 1259–1275.
- Hsiao, P.-H., Hwang, A., Kung, H.T., Vlah, D., 2001. Load-balancing routing for wireless access networks. In: *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings of the IEEE*, 2. IEEE, pp. 986–995.
- Li, F., Wang, Y., Li, X.-Y., Nusairat, A., Wu, Y., 2008. Gateway placement for throughput optimization in wireless mesh networks. *Mobile Netw. Appl.* 13, 198–211.
- Magnanti, T.L., Mirchandani, P., Vachani, R., 1993. The convex hull of two core capacitated network design problems. *Math. Program.* 60 (1), 233–250.
- Melkote, S., Daskin, M.S., 2001. Capacitated facility location/network design problems. *Eur. J. Oper. Res.* 129 (3), 481–495.
- Melkote, S., Daskin, M.S., 2001. An integrated model of facility location and transportation network design. *Transp. Res. Part A: Policy Pract.* 35 (6), 515–538.
- Nelson, R., Kleinrock, L., 1985. Spatial TDMA: a collision-free multihop channel access protocol. *IEEE Trans. Commun.* 33 (9), 934–944.
- Papadaki, K., Friderikos, V., 2008. Approximate dynamic programming for link scheduling in wireless mesh networks. *Comput. Oper. Res.* 35 (12), 3848–3859.
- Pathak, P.H., Dutta, R., 2011. A survey of network design problems and joint design approaches in wireless mesh networks. *IEEE Commun. Surv. Tutorials* 13–3, 396–428.
- Prasad, R., Wu, H., 2006. Gateway deployment optimization in cellular wi-Fi mesh networks. *J. Netw.* 1–3, 31–39.
- Quintas, D., Friderikos, V., 2012. Energy efficient spatial TDMA scheduling in wireless networks. *Comput. Oper. Res.* 39 (9), 2091–2099.
- Seyedzadegan, M., Othman, M., Ali, B.M., Subramaniam, S., 2013. Zero-degree algorithm for internet gateway placement in backbone wireless mesh networks. *J. Netw. Comput. Appl.* 36 (6), 1705–1723.
- Targon, V., Sanso, B., Capone, A., 2010. The joint gateway placement and spatial reuse problem in wireless mesh networks. *Comput. Netw.* 54 (2), 231–240.
- Wu, W., Luo, J., Yang, M., 2009. Gateway placement optimization for load balancing in wireless mesh networks. In: *Proceedings of 13th International Conference on Computer Supported Cooperative Work in Design*, pp. 408–413.