Wireless Pers Commun
DOI 10.1007/s11277-009-9845-6

Backhaul and Routing Assignments with End-to-End QoS Constraints for Wireless Mesh Networks

Yean-Fu Wen · Frank Yeong-Sung Lin · Yeong-Cheng Tzeng · Cheng-Ta Lee

Abstract In scalable last-mile broadband networks such as wireless mesh networks (WMNs), quality-of-service (QoS) concerns are vital to multimedia applications such as video-conferencing and voice over IP (VoIP). Crucial decisions involve the number of backhauls that are to be deployed as well as the optimal assignment of paths and bandwidths. We focus on cost effectiveness and QoS requirements to develop a solution based on Lagrangean Relaxation and the subgradient method. Our approach satisfies QoS demands and minimizes costs more effectively than general algorithms, as demonstrated by our experimental results.

Keywords Backhaul assignment · WMNs · QoS · Routing · Optimization · Lagrangean relaxation method

1 Introduction

Wireless mesh network (WMN) topologies utilize wireless multi-hop communications to extend the range of traditional LANs and WLANs. As shown in Fig. 1, a WMN consists of

Y.-F. Wen (✉)
Department of Management Information Systems, National Chiayi University, Chiayi, Taiwan
e-mail: yeanfu@mail.ncyu.edu.tw; yeanfu@ieee.org

F. Y.-S. Lin
Department of Information Management, National Taiwan University, Taipei, Taiwan
e-mail: yslin@im.ntu.edu.tw

Y.-C. Tzeng
ZyXEL Communications Corp., Hsinchu, Taiwan
e-mail: r93030@im.ntu.edu.tw

C.-T. Lee
Department of Information Management, National Taiwan University and Lan Yang Institute of Technology, Taipei, Taiwan
e-mail: d90001@im.ntu.edu.tw

Published online: 18 October 2009

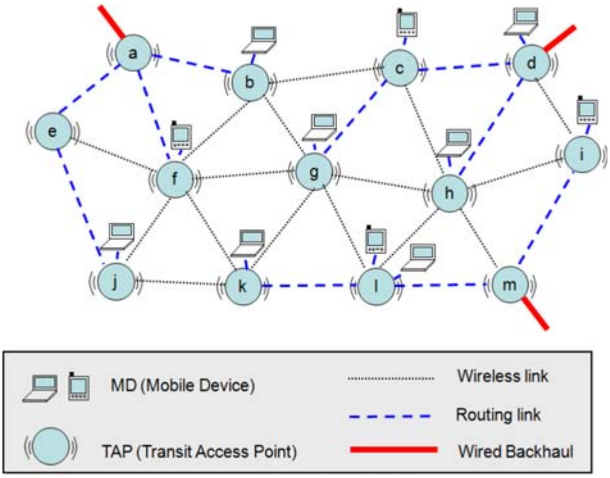


Fig. 1 A mesh network constructed with a BS-oriented and ad hoc structure connects the wired network via some backhauls and covers a large area via wireless links

Transient Access Points (TAPs) with at least one backhaul, an ingress/egress link routing to network infrastructure. It has been argued that TAPs, responsible for coordinating connection requests of clients and for relaying data traffic, can provide for a scalable architecture that enables wide deployment of high capacity wireless networks [1,16,31]. WMNs can achieve the performance of wired networks by using multi-radio mesh [17].

Despite the potentials of WMNs, there remain several difficulties to overcome. The deployment of high-capacity TAPs bears considerable expense [1,31], yet under quality-of-service (QoS) requirements, mobile devices (MDs) must be in close proximity to TAPs. In this paper we focus on cost effectiveness and coverage in addressing the following issues: (a) appropriate assignment of backhauls; (b) multiple-hop load-balanced routing; and (c) guarantees of QoS.

The objective of traditional base stations deployment is to address the coverage planning with one-hop transmission. The problem in which the number of transmitters placed is minimized while covering a specific area or at least q demand nodes is NP-hard. Several methods suitable for coverage planning can be found in the literature, mostly based on random search [11,12] and greedy algorithm [15,28]. Here, the backhaul assignment problem does not address geographic coverage (i.e., the given set of TAPs to serve mobile devices that have satisfied coverage condition), but these backhauls assignment satisfy the traffic requirement with multi-hop transmission.

The issue of fair allocation of resources in WMNs is an important topic central to this work. Fairness schemes, such as appear in [10] and [22], are concerned with the following two issues: (a) Temporal fairness: Within a collision area, the allocation of resources is controlled to ensure that the channel access time is fair for all nodes. (b) Spatial bias fairness: The fairness scheme assigns channel access time uniformly to flows by allocating more resources to nodes located further away from the destination.

We note that the above fairness concerns are not sufficient to ensure the level of QoS necessary for multimedia. Despite controlling for temporal fairness, a node that is located further from a backhaul potentially suffers a starvation problem that leads to lower throughput and may prevent the delivery of data within a reasonable amount of time. And although throughput does not decrease as a result of providing spatial bias fairness, it does not the amount of time for transmissions from more distant MDs to reach the backhaul. Therefore, it has been argued that routing protocols should be redesigned with respect to fairness concerns that are characteristic to WMNs [1,4,14,21,27]. For example, in WMNs, throughput for clients reliant on a single backhaul is dependent on distance from the backhaul. Clients on longer hop paths suffer from significantly lower throughput in contrast with those on shorter paths [1,10,18,31].

As in [30], we extend the fairness issue from the viewpoint of resource allocation and load balancing to achieve end-to-end delay fairness. The algorithms are designed to optimize load-balanced routing and resource allocation along the path for different length (i.e., different number of hops) flow. In [26], the authors sought to maximize network throughput and achieve fairness by cross-layer schemes that solved for joint rate allocation, routing, scheduling, power control, and channel assignment problems. Fairness was measured by a simplified max- min and proportional fairness model. An upper bound on the maximum throughput was obtained that made a good tradeoff between throughput and fairness. In this study, we additionally consider nodal capacity, nodal delay, delay jitter, and backhaul assignment and selection. We provide both lower and upper bounds on the backhaul assignment costs.

In order to enable multimedia applications such as video-conferencing and voice over IP (VoIP) in WMNs, the guarantees of Quality-of-Service (QoS) are very essential. This is because multimedia applications are very sensitive to delay and delay jitter. If the network is well designed and Internet gateways are optimally deployed, the transmission can satisfy QoS-guaranteed multimedia applications.

A number of multiple constraint routing problems were surveyed in [19], with a focus on quality of service (QoS) for wireless ad hoc networks. Observing that two or more QoS constraints is an NP-complete problem, they used approximate solutions that an energy function to translate multiple QoS weights as a mixed metric and then adopt simulated annealing approach to find a feasible path. Many other multiple constraints routing method include CEDAR (which denotes as Core-Extraction Distributed Ad hoc Routing) [25], a ticket-based distributed QoS routing scheme [6], a genetic algorithm (named GAMAN) [3], and a randomized algorithm are also introduced. In this paper, we adopt the multiple QoS constraints routing algorithms as one of major issue and include capacity, load-balancing, and backhaul assignment aspects.

Only a single TAP in the area can transmit data at a time in collision areas where links must share a channel. As [8,10,13,18,22] demonstrate, performance degrades sharply as the number of hops traversed increases. It would be better less concurrent transmissions in the network to avoid contention delay. Thus, the multi-channel configuration pattern, such as the availability of cognitive radio technology [2], that multiple non-overlapping channels, multiple radios per node, and directional antenna [17], is considered in this study. Even through the multi-channel is considered, the overshooting interference (i.e., the power range is cover to farer node) is also included. The interference probability function is referred to [5], which directly reflects on the link capacity. The calculated range is referred to the link-based “conflict graph” [13] to model the capacity constraint.

We model the problem as a mixed-integer nonlinear programming problem. As shown in Theorem 1, the problem is NP-hard that we adopt an LR-based approach, which achieves a near-optimal solution and provides a lower bound (LB) on backhaul assignment costs [9]. In order to solve the problems, heuristics were developed and the Lagrangean Relaxation (LR) method is applied. The difficulty of solving the original NP-hard problem is reduced by dualizing the set of constraints with fixed multipliers to the objective function. This new problem is further divided into several mutually independent subproblems with its own constraints. We take the sub-gradient method with finding the extreme points to solve the LR problem. This approach allows our algorithms to be used to solve each subproblem optimally within a smaller space [8,9].

Theorem 1 *The minimum cost backhaul assignment subject to end-to-end delay fairness issue is NP-hard.*

Proof Minimum Backhaul assignment problem is a kind of BS placement problem that many researches have been proved NP-hard [11,12]. In addition, the cost for each routing link is constrained to multiple constrains, include capacity, delay, delay jitter, and load-balancing constrains, that are associated into a classical multi-constrained path (MCP) problem is also proved NP-hard [29]. Thus, the combination of Backhaul assignment and MCP routing is a NP-hard problem. □

The remainder of this paper is organized as follows. A mathematical formulation for the WMNs design problem is first shaped in Sect. 2. Section 3 presents the constraints relaxation of the primal problem and the methods for solving a Lagrangean dual problem. Section 4 describes how to get primal feasible solutions and its heuristics of each problem. Section 5 is the computational experiments for each problem. Finally, in Sect. 6, the summary of this paper is presented.

2 Problem Description and Formulation

In this section, we firstly describe and definite the problem, and then formulate the problem as an integer nonlinear problem.

2.1 Problem Description

Initially, decisions are backhauls are to be deployed economically and how to assign MDs to appropriate TAPs. Then, each TAP selects an appropriate backhaul and finds a routing path to send/receive data to/from that the bandwidth is dynamically allocated along the path such that the end-to-end QoS requirements of each MD are satisfied.

To jointly consider the backhauls deployment and end-to-end performance of the entire network is NP-hard. Under the goal of minimum cost of backhaul deployment, we install a minimum number of backhauls to satisfy the end-to-end QoS requirements of all MDs. However, an inefficient backhauls deployment causes large budget requirement and waste the allocated resource (i.e., low utilization). Thus, there implies a tradeoff between deployment cost and end-to-end QoS requirements. The summaries of problem description are listed as follows.

Assumptions:

- The node, which attached with backhaul links, integrates both functions of AP and backhaul
- All flows are transmitted to Internet through backhauls
- There is no additional round trip time from the wired Internet

- MD to TAP and TAP to TAP transmission occur on orthogonal channels, which available based on cognitive channel technique [2]
- The average delay and jitter from one MD to any TAP can be formulated as a function of required traffic rate and link capacity
- The average delay and delay jitter from one TAP to another can be formulated as a function of link aggregated flow and capacity

Given:

- The set of all TAPs V, which is also the set of candidate backhauls B
- The set of all backhaul configurations K
- The set of all MDs N
- The set of all candidate paths P_{bs} from TAP s to backhaul b, where $s, b \in V$
- The link capacity $C_{(u,v)}$ (packets/sec), which is calculated by bit error rate $(1 - e_f)^{(E_u+E_v)}$, on link (u, v). Note that E_u and E_v denotes the number TAPs within interference range of node u and v, respectively
- The nodal capacity $\overline{C_v}$(packets/sec) of TAP v
- The air-interface capacity $\overline{\overline{C_v}}$ (packets/sec), which is calculated by bit error rate $(1 - e_f)^{E_s}$, of TAP v. Note that E_s denotes the number mobile nodes within interference range
- The cost function $\Phi_b(k)$ of building the wired line on backhaul b, which is a function of backhaul configuration k
- The capacity function $Q_b(k)$ of the wired line on backhaul b, which is a function of backhaul configuration k
- The data rate θ_n (packets/sec) required to be transmitted of MD n
- The link capacity r_{ns} (packets/sec) from MD n to TAP s
- The average delay function $\overline{F}_{ns}$ (θ_n, r_{ns}) from $n(\in N)$ to source TAP $s(\in V)$, which is a function of required data rate θ_n and link capacity r_{ns}
- The delay jitter function $\overline{M}_{ns}$ (θ_n, r_{ns}) from MD n to source TAP s, which is a function of required data rate θ_n and MD to TAP link capacity r_{ns}
- The cost of backhaul installation and configuration c_b
- The required data rate a_s (packets/sec) of each MD
- The QoS requirements including end-to-end mean delay T and delay jitter J
- Arbitrarily large numbers M_1, M_2, and M_3

Objective:

To minimize the total cost of backhaul deployment

Subject to the following constraints:

- Backhaul assignment: at least one backhaul must to be selected in order for it to receive and transmit data from/to Internet
- Routing: at least one path to the backhaul must be found for each node in order for it to transmit and receive data
- Link: the selected links of all routing paths to the backhaul
- TAP selection: each MD must select a TAP to transmit and receive its data to
- Link capacity: the allocated capacity for the flows on each link must be less than the capacity
- Nodal capacity: each TAP's or backhaul's total incoming data flow from others TAPs should not be large than its nodal capacity
- QoS constraints: the end-to-end delay and delay jitter of all source nodes must be less than the application requirements

To determine:
- Backhaul deployment and configuration
- The source TAP assignment of each MD
- The routing path from a TAP to a backhaul
- The links that are selected for the routing path
- The capacity allocated to the selected links of a TAP node
- The end-to-end delay and delay jilter of a node

2.2 Problem Formulation

Listed Table 1 are the decision variables corresponding to our problem formulation. According to the above description, the problem we addressed is to minimize the total cost of backhaul deployment in WMNs, while considering the end-to-end QoS requirements of each MD.
Objective function: The main objective of this problem is to minimize the cost of backhaul deployment that includes installation cost of upgrading existing TAPs to backhauls (first term in the bracket) and wired line cost of leasing wired lines on selected backhauls to the Internet (second term in the bracket).

$$\min \sum_{b \in V} \sum_{k \in K} \left(c_b + \Phi_b\left(k\right)\right) \eta_{bk} \tag{1}$$

subject to:
(a) Backhaul assignment constraints:
Constraint (2) confines that each candidate backhaul selects exactly only one configuration (when the backhaul is assigned) or none (when the backhaul is not assigned) in order to leave more resource (e.g., channel bands) and save money for other assigned backhaul.

$$\sum_{k \in K} \eta_{bk} \leq 1, \quad \forall b \in V \tag{2}$$

Constraint (3) confines that each TAP s can only select a candidate backhaul b as its gateway so that the number of assigned backhaul is minimized, as shown as (4).

$$\sum_{b \in V} z_{bs} = 1, \quad \forall s \in V \tag{3}$$

Table 1 Notation of decision variables

Notation	Description
η_{bk}	1 if TAP b is selected to associate a backhaul with configuration k; otherwise 0
z_{bs}	1 if TAP s connects to the wired network via backhaul b; otherwise 0
x_p	1 if path p from TAP s to backhaul b is selected; otherwise 0
$y_{s(u,v)}$	1 if link (u, v) is on the path adopted by TAP s; otherwise 0
κ_{ns}	1 if MD n associates to TAP s; otherwise 0
α_s	The data rate (packets/sec) required to be transmitted of TAP s
$\gamma_{s(u,v)}$	The bandwidth allocation of TAP s on link (u, v)
$f_{(u,v)}$	The aggregate flow on link (u, v)
$F_{(u,v)}(f_{(u,v)}, C_{(u,v)})$	The average delay on link (u, v), which is a function of aggregate flow $f_{(u,v)}$ and link capacity $C_{(u,v)}$
$M_{(u,v)}(f_{(u,v)}, C_{(u,v)})$	The delay jitter on link (u, v), which is a function of aggregate flow $f_{(u,v)}$ and link capacity $C_{(u,v)}$

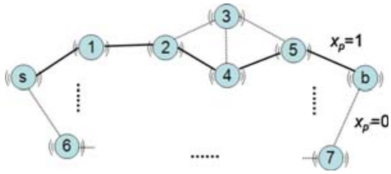


Fig. 2 All paths from s (i.e., TAP s) to b (i.e., backhaul b) are included in the set P_{sb}. The path above, shown by the *bold dash line* $\{s, 2, 4, 5, 6, b\}$, is the only active path in this set, so for that path $x_p = 1$. As there are no other active paths, for any other path p' (e.g., path $\{s, 3, \ldots, 7, b\}$), $x_{p'} = 0$

Constraint (4) confines that once a TAP s selects TAP b as its backhaul, the TAP b should be installed as a backhaul firstly. This backhaul also selects one of candidate configuration.

$$z_{bs} \leq \sum_{k \in K} \eta_{bk}, \quad \forall b, s \in V \tag{4}$$

Definition 1 Our objective is to build-up enough necessary bandwidth in the network under given QoS traffic requirement. Thus, equation (5) is added to confine that the total wired line capacity on backhauls (right-hand side) should be equal to or large than the total data rate required to all MDs (left-hand side). Therefore, all incoming flows from MDs can be transmitted to the Internet via backhauls. This also helps to determine the lower bound number of backhauls requirement.

$$\sum_{n \in N} \theta_n \leq \sum_{b \in V} \sum_{k \in K} Q_b(k) \cdot \eta_{bk} \tag{5}$$

□

(b) Routing constraints:

Constraint (6) confines that once TAP s selects candidate backhaul b as its gateway, a path from TAP s to candidate backhaul b must be found as shown an example in Fig. 2.

$$z_{bs} \leq \sum_{p \in P_{bs}} x_p, \quad \forall b, s \in V \tag{6}$$

Constraint (7) confines that each TAP s selects exactly one candidate backhaul as egress, and exactly one routing path to the selected egress. When decision variable $x_p = 1$, it indicates that the path $p \in P_{bs}$ is the path used to connect the TAP to the backhaul; $x_p = 0$ means that it is not used. Figure 2 shows an example of the decision variable.

$$\sum_{b \in V} \sum_{p \in P_{bs}} x_p = 1, \quad \forall s \in V \tag{7}$$

Constraint (8) confines that once the path p is selected and the link (u, v) is on the path, then the decision variable $y_{s(u,v)}$ must be equal to 1 (e.g., links $(s, 2), (2, 4), \ldots, (6, b)$ in Fig. 2).

$$\sum_{b \in V} \sum_{p \in P_{bs}} x_p \delta_{p(u,v)} \leq y_{s(u,v)}, \quad \forall s \in V; (u, v) \in L \tag{8}$$

(c) Link constraints:

Based on the graph theory, the constructed links to the routing paths are restricted. Constraints (9) and (10) are two complementary constraints which confine that each TAP, except the backhauls, has at least one outgoing link. If node u is not installed as a backhaul, the left-hand side of (9) is equal to 1 and let right-hand side has large to 1 (i.e., at least one outgoing link is selected to forward the MDs' data cover by node u) (e.g., node 2 in Fig. 2).

However, the right-hand side of (10) is equal to 0, then the number of outgoing link of node u (e.g., node b in Fig. 2) is limited to 0.

$$1-\sum_{k\in K}\eta_{uk}\leq\sum_{s\in V}\sum_{v\in V}y_{s(u,v)},\quad\forall u\in V \tag{9}$$

$$\sum_{s\in V}\sum_{v\in V}y_{s(u,v)}\leq M_1\left(1-\sum_{k\in K}\eta_{uk}\right),\quad\forall u\in V \tag{10}$$

Constraint (11) confines that the backhauls has at least one incoming link. The right-hand side value in (11) needs to select at least one incoming links when node v (e.g., node b in Fig. 2) is installed as a backhaul.

$$\sum_{k\in K}\eta_{vk}\leq\sum_{s\in V}\sum_{u\in V}y_{s(u,v)},\quad\forall v\in V \tag{11}$$

Definition 2 The upper bound number of incoming links is the number of source nodes minus 1 that except itself. This value is formulated as equation (12).

$$\sum_{s\in V}\sum_{u\in V}y_{s(u,v)}\leq|S|-1,\quad\forall v\in V \tag{12}$$

□

(d) TAP selection constraints:
Constraint (13) confines that each MD is assigned to exactly one TAP.

$$\sum_{s\in V}\kappa_{ns}=1,\quad\forall n\in N \tag{13}$$

Constraint (14) confines that the total incoming data rate from MDs admitted by TAP s should not be large than its air-interface capacity $\overline{\overline{C_s}}$.

$$\sum_{n\in N}\theta_n\kappa_{ns}\leq\overline{\overline{C_s}},\quad\forall s\in V \tag{14}$$

Constraint (15) confines the total incoming data rates from MDs admitted by TAP s should not be large than the data rate required to be transmitted by TAP s.

$$\sum_{n\in N}\theta_n\kappa_{ns}\leq a_s,\quad\forall s\in V \tag{15}$$

Constraint (16) confines the boundaries of data rate required to be transmitted by each TAP.

$$0\leq a_s\leq\overline{\overline{C_s}},\quad\forall s\in V \tag{16}$$

Constraint (17) confines that the total data rate required to be transmitted of all TAPs should be equal to or large than the total data rate required to be transmitted of all MDs.

$$\sum_{n\in N}\theta_n\leq\sum_{s\in V}a_s \tag{17}$$

(e) Link capacity constraints:
Constraints (18) and (19) are two complementary constraints which confine that the bandwidth allocation of TAP s on link (u,v) should be equal to the data rate required to

be transmitted of TAP s if link (u, v) is on the selected path of TAP s. Otherwise, the bandwidth allocation of TAP s on link (u, v) should be 0.

$$a_s - M_2(1 - y_{s(u,v)}) \leq \gamma_{s(u,v)}, \quad \forall s \in V; (u, v) \in L \tag{18}$$

$$\gamma_{s(u,v)} \leq a_s, \quad \forall s \in V; (u, v) \in L \tag{19}$$

Constraint (20) confines that total bandwidth allocation of all TAPs on link (u, v) is aggregate into link flow decision variable $f_{(u,v)}$. The reason we use least and equal symbol (i.e., $\leq$) to replace equal symbol (i.e., $=$) is for constraint relaxation, which will be shown in next section.

$$\sum_{s \in V} \gamma_{s(u,v)} \leq f_{(u,v)}, \quad \forall (u, v) \in L \tag{20}$$

Constraint (21) confines the limitation of aggregate flow on link (u, v). Namely, the aggregation flow of each link is not over the link capacity $C_{(u,v)}$.

$$0 \leq f_{(u,v)} \leq C_{(u,v)}, \quad \forall (u, v) \in L \tag{21}$$

(f) Nodal capacity constraints:
Constraint (22) confines that each TAP's total incoming data flow from others TAPs should not be large than its nodal capacity.

$$\sum_{s \in V} \sum_{u \in V} \gamma_{s(u,v)} \leq \overline{C}_v, \quad \forall v \in V \tag{22}$$

Constraint (23) confines that total incoming flow of all backhauls should not be large than total wired line capacities.

$$\sum_{s \in S} \sum_{u \in V} \gamma_{sub} + a_b - M_3 \left(1 - \sum_{k \in K} \eta_{bk}\right) \leq \sum_{k \in K} Q_b(k)\, \eta_{bk}, \quad \forall b \in V \tag{23}$$

(g) QoS constraints:
Constraint (24) confines the end-to-end average delay should be no longer than maximum allowable end-to-end average delay requirement.

$$\sum_{u \in V} \sum_{v \in V} y_{s(u,v)} F_{(u,v)} \left(f_{(u,v)}, C_{(u,v)}\right) + \kappa_{ns} \overline{F}_{ns} (\theta_n, r_{ns}) \leq T, \quad \forall n \in N, s \in V \tag{24}$$

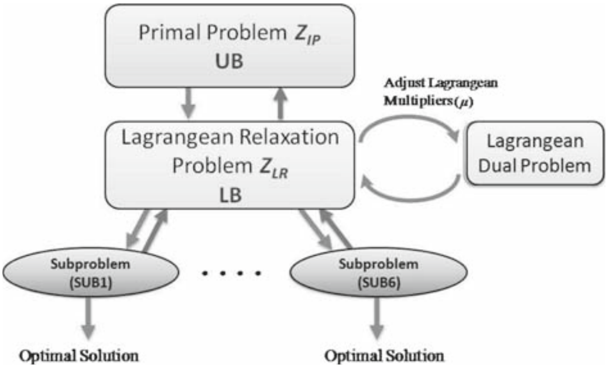
Constraint (25) confines the end-to-end delay jitter should be no longer than maximum allowable end-to-end delay jitter requirement. For simplification, we take $M/M/1$ model to calculate the intra-TAP mean delay. And we compute the delay from mobile host to TAP by the formulation proposed in [24]. We assume the delay time is exponential distribution. Therefore, the delay jitter is the square of the mean delay.

$$\sum_{u \in V} \sum_{v \in V} y_{s(u,v)} M_{(u,v)} \left(f_{(u,v)}, C_{(u,v)}\right) + \kappa_{ns} \overline{M}_{ns} (\theta_n, r_{ns}) \leq J, \quad \forall n \in N, s \in V \tag{25}$$

3 Solution Approach

The overall procedure to solve the network design problem is shown as in Fig. 3. They are composed of two procedures, constraints relaxation and subgradient optimization procedure. The relaxation of the primal problem is developed first which provides lower bound (LB) on

Fig. 3 Lagrangean relaxation procedures



the optimal solutions. Since we relax some constraints of the original problem, the boundary is used to design a heuristic approach to get a primal feasible solution. To solve the original problem near-optimally and minimize the gap between the primal problem and the Lagrangean dual problem, we improve the LB by solving the sub-problems optimally and using the subgradient method to adjust the multipliers per iteration. Then, subgradient optimization procedure is used for further improving these solutions by updating the Lagrangean multipliers.

3.1 Constraints Relaxation

We relax Constraints (4), (5), (6), (8), (9), (10), (11), (14), (15), (18), (19), (20), (23), (24), and (25) and multiply them by the multiplier vectors, which are the costs of decision variables in these constraints, that add to the objective function as follows:

$$
\begin{aligned}
& Z_{LR}\begin{pmatrix} \mu^1, \mu^2_{bs}, \mu^3_{bs}, \mu^4_{suv}, \mu^5_{u}, \mu^6_{u}, \mu^7_{v}, \mu^8_{s}, \\ \mu^9_{s}, \mu^{10}_{suv}, \mu^{11}_{suv}, \mu^{12}_{uv}, \mu^{13}_{b}, \mu^{14}_{ns}, \mu^{15}_{ns} \end{pmatrix} \\
& \quad = \min \sum_{b\in V}\sum_{k\in K} \left(c_b + \Phi_b\left(k\right)\right)\eta_{bk} \\
& \qquad + \mu^1 \left[\sum_{n\in N}\theta_n - \sum_{b\in V}\sum_{k\in K} Q_b\left(k\right)\cdot \eta_{bk}\right] \\
& \qquad + \sum_{b\in V}\sum_{s\in V}\mu^2_{bs}\left[z_{bs} - \sum_{k\in K}\eta_{bk}\right] \\
& \qquad + \sum_{b\in V}\sum_{s\in V}\mu^3_{bs}\left[z_{bs} - \sum_{p\in P_{bs}} x_p\right] \\
& \qquad + \sum_{s\in V}\sum_{u\in V}\sum_{v\in V}\mu^4_{suv}\left[\sum_{b\in V}\sum_{p\in P_{bs}} x_p\delta_{p(u,v)} - y_{s(u,v)}\right] \\
& \qquad + \sum_{u\in V}\mu^5_{u}\left[1 - \sum_{k\in K}\eta_{uk} - \sum_{s\in V}\sum_{v\in V} y_{s(u,v)}\right]
\end{aligned}
$$

$$
\begin{aligned}
&+\sum_{u\in V}\mu_u^6\left[\sum_{s\in V}\sum_{v\in V}y_{s(u,v)}-M_1\left(1-\sum_{k\in K}\eta_{uk}\right)\right]\\
&+\sum_{v\in V}\mu_v^7\left[\sum_{k\in K}\eta_{vk}-\sum_{s\in V}\sum_{u\in V}y_{s(u,v)}\right]\\
&+\sum_{s\in V}\mu_s^8\left[\sum_{n\in N}\theta_n\kappa_{ns}-\overline{\overline{C_s}}\right]\\
&+\sum_{s\in V}\mu_s^9\left[\sum_{n\in N}\theta_n\kappa_{ns}-a_s\right]\\
&+\sum_{s\in V}\sum_{u\in V}\sum_{v\in V}\mu_{suv}^{10}\left[a_s-M_2(1-y_{s(u,v)})-\gamma_{s(u,v)}\right]\\
&+\sum_{s\in V}\sum_{u\in V}\sum_{v\in V}\mu_{suv}^{11}\left[\gamma_{s(u,v)}-a_s\right]\\
&+\sum_{u\in V}\sum_{v\in V}\mu_{uv}^{12}\left[\sum_{s\in V}\gamma_{s(u,v)}-f_{(u,v)}\right]\\
&+\sum_{b\in V}\mu_b^{13}\left[\begin{array}{l}\sum_{s\in S}\sum_{u\in V}\gamma_{sub}+a_b-M_3\left(1-\sum_{k\in K}\eta_{bk}\right)\\-\sum_{k\in K}Q_b\left(k\right)\eta_{bk}\end{array}\right]\\
&\times\sum_{n\in N}\sum_{s\in V}\mu_{ns}^{14}\left[\begin{array}{l}\sum_{u\in V}\sum_{v\in V}y_{s(u,v)}F_{(u,v)}\left(f_{(u,v)},C_{(u,v)}\right)\\+\kappa_{ns}\overline{F}_{ns}\left(\theta_n,r_{ns}\right)-T\end{array}\right]\\
&+\sum_{n\in N}\sum_{s\in V}\mu_{ns}^{15}\left[\begin{array}{l}\sum_{u\in V}\sum_{v\in V}y_{s(u,v)}M_{(u,v)}\left(f_{(u,v)},C_{(u,v)}\right)\\+\kappa_{ns}\overline{M}_{ns}\left(\theta_n,r_{ns}\right)-J\end{array}\right]
\end{aligned}
\tag{LR}
$$

subject to: (2), (3), (7), (12), (13), (16), (17), (21), (22).
where μ^1, μ_{bs}^2, μ_{bs}^3, μ_{suv}^4, μ_u^5, μ_u^6, μ_v^7, μ_s^8, μ_s^9, μ_{suv}^{10}, μ_{suv}^{11}, μ_{uv}^{12}, μ_b^{13}, μ_{ns}^{14}, and μ_{ns}^{15} are the vectors of non-negative Lagrangean multipliers. To solve the LR, we decompose the problem into seven mutually independent and easily solvable optimization subproblems, which shown in Appendix.

3.2 The Dual Problem and the Subgradient Method

According to the weak Lagrangean duality theorem, for any multipliers $\mu^1, \mu_{bs}^2, \mu_{bs}^3, \mu_{suv}^4, \mu_u^5, \mu_u^6, \mu_v^7, \mu_s^8, \mu_s^9, \mu_{suv}^{10}, \mu_{suv}^{11}, \mu_{uv}^{12}, \mu_b^{13}, \mu_{ns}^{14}, \mu_{ns}^{15} \geq 0$, the objective value of $Z_{LR}\big(\mu^1, \mu_{bs}^2, \mu_{bs}^3, \mu_{suv}^4, \mu_u^5, \mu_u^6, \mu_v^7, \mu_s^8, \mu_s^9, \mu_{suv}^{10}, \mu_{suv}^{11}, \mu_{uv}^{12}, \mu_b^{13}, \mu_{ns}^{14}, \mu_{ns}^{15}\big)$ is a lower bound of Z_{IP}. Based in problem (LR), the following dual problem (D) is then constructed to calculate the tightest lower bound.

Lagrangea Dual Problem (D):

$$
Z_D=\max Z_{\mathrm{LR}}\left(\mu^1, \mu_{bs}^2, \mu_{bs}^3, \mu_{suv}^4, \mu_u^5, \mu_u^6, \mu_v^7, \mu_s^8, \mu_s^9, \mu_{suv}^{10}, \mu_{suv}^{11}, \mu_{uv}^{12}, \mu_b^{13}, \mu_{ns}^{14}, \mu_{ns}^{15}\right)
$$

subject to:

$$
\mu^1, \mu_{bs}^2, \mu_{bs}^3, \mu_{suv}^4, \mu_u^5, \mu_u^6, \mu_v^7, \mu_s^8, \mu_s^9, \mu_{suv}^{10}, \mu_{suv}^{11}, \mu_{uv}^{12}, \mu_b^{13}, \mu_{ns}^{14}, \mu_{ns}^{15} \geq 0
$$

There are several methods to solve the dual problem Z_D. One of the most popular methods is the subgradient method which employed here. Let the vector g be a subgradient of $Z_D(\mu^1, \mu^2_{bs}, \mu^3_{bs}, \mu^4_{suv}, \mu^5_u, \mu^6_u, \mu^7_v, \mu^8_s, \mu^9_s, \mu^{10}_{suv}, \mu^{11}_{suv}, \mu^{12}_{uv}, \mu^{13}_b, \mu^{14}_{ns}, \mu^{15}_{ns})$. Then, in iteration k of the subgradient optimization procedure, the multiplier vector $\pi^k = (\mu^1, \mu^2_{bs}, \mu^3_{bs}, \mu^4_{suv}, \mu^5_u, \mu^6_u, \mu^7_v, \mu^8_s, \mu^9_s, \mu^{10}_{suv}, \mu^{11}_{suv}, \mu^{12}_{uv}, \mu^{13}_b, \mu^{14}_{ns}, \mu^{15}_{ns})$. The step size t^k is determined by $t^k = \lambda\left(Z^h_{IP} - Z_D\left(\pi_k\right)\right) \Big/ \left\|g^k\right\|^2$. is the primal objective function value for a heuristic solution (an upper bound on Z_{IP}) and λ is a constant where $0 \leq \lambda \leq 2$.

Lemma 1 *The step size is important in the subgradient algorithm. If the step sizes are too small, the convergence speed of algorithm will be slow, and if they are too large, the multiplier may oscillate around the optimal solution and the algorithm will fail to converge. However, convergence is guaranteed when the step size t^k satisfies the following condition [23]:*

$$t^k \geq 0, \lim_{k\to\infty} t^k = 0 \quad \text{and} \quad \sum_{k=1}^{\infty} t^k = \infty$$

□

4 Getting Primal Feasible Solution

By applying LR method and the subgradient method to solve the complex problem, we can get a theoretical lower bound of the primal problem and some hints to get a feasible solution to the primal problem. Because some difficult constraints of the primal problem are relaxed by using LR method, we can not guarantee that the consolidated result of the LR problem is feasible to the primal problem. We have to ensure that it is a feasible solution, which is satisfied with all constraints of the primal problem, if not, we have to make some modifications.

4.1 Getting Primal Heuristic

We take the major decision variable, η_{bk} (i.e., whether to deploy backhaul b with configuration k), into consideration. According to η_{bk}, we can obtain which TAPs should be installed as backhauls in each LR iteration. We count the frequency that each TAP should be installed as a backhaul iteration by iteration. Because the maximal data rate that a backhaul can process limits to the sum of its nodal capacity and air-interface capacity. And the total maximal processing data rate of all backhauls should not be less than the total data rate required to be transmitted of all MDs. Therefore, we pick up TAPs to be installed as backhauls with frequency in ascending order, until all backhauls' maximal processing data rate do not less than the total data rate request from MDs to satisfy Definition 1.

After initiate the backhaul deployment, we should assign MDs to appropriate TAPs. Therefore, we can obtain the average data rate required to be transmitted of each TAP. Then, we run routing heuristic for TAPs to decide backhaul assignment and routing paths selection. Besides, the initiated backhaul deployment may not be feasible. Thus, we propose add backhaul heuristic to get the feasible solution. The procedures are shown as follows.

Step 1 Initiate backhaul deployment according to decision variable η_{bk}
Step 2 Run Assign_Mobile_Host_Heuristic
Step 3 Run Routing_Heuristic
Step 4 Go to **Step 5** if all TAPs can route to associate backhauls without violating end-to-end QoS requirements
  Step 4.1 Run Add_Backhaul_Heuristic
  Step 4.2 Go back to **Step 2**
Step 5 Calculate the total cost of backhaul deployment

4.2 Assign Mobile Device Heuristic

By decision variable κ_{ns}, we can decide how to assign MDs to associated TAPs. Some TAPs may violate the air-interface capacity due to admit too many MDs. For getting primal feasible solutions, the MD assignment should be adjusted.

If a MD within the access range of a TAP and a backhaul at the same time, the MD should try to access the backhaul first. Therefor, the MD can get into the Internet via the backhaul directly and does not experience the poor performance of wireless multi-hop transition. We describe the detail procedures as follows.

Step 1 Initiate MD assignment according to variable κ_{ns}
Step 2 Find a TAP that violates the air-interface capacity most seriously. If not found, **Stop**
Step 3 For each MD, we try to reassign to another TAP and calculate the coefficient of κ_{ns}. Then, we find the MD with smallest coefficient and reassign to the relative TAP. Repeat **Step 3** until this TAP does not violate air-interface capacity
Step 4 Repeat **Steps 2 and 3** until all TAPs do not violate air-interface capacity
Step 5 For each backhaul, we reassign the nearby MDs with smallest coefficient κ_{ns} of to the backhaul one at a time until cannot admit one MD without violating the air-interface capacity

4.3 Routing Heuristic

The basic idea of routing heuristic is that if the end-to-end QoS performance of one TAP is close to its QoS requirements, this TAP should route first. This means the TAP with tightest QoS has less flexibility in routing path selection. In the following, we show the detail procedures as follows.

Step 1 Set the arc weight, which is equal the link delay $F_{(u,v)}(f_{(u,v)}, C_{(u,v)})$ and delay jitter $M_{(u,v)}(f_{(u,v)}, C_{(u,v)})$, for each link to be the coefficient of variable x_p and run Dijkstra's algorithm to get the shortest path from each TAP
Step 2 Choose a path with the tightest QoS performance
Step 3 Repeat **Steps 1 and 2** until all TAP have a path to a backhaul

4.4 Add Backhaul Heuristic

The basic idea of this heuristic is that if a TAP locate at the place that many traffic flows may pass through, this TAP is at a proper location for installed as a backhaul. This means many other TAPs' data flow can reach to this TAP. We denote the times of reaching by other TAPs as "reachability". Therefore, we calculate each TAP's reachability, then we pick the highest reachability value for backhaul deploy. We show the detail procedures as follows.

Step 1 Initiate all TAPs' reachability counter to zero

Step 2 Find a TAP that admits data flow from MDs but not be assigned to any backhaul. If a TAP without assigned backhaul found, we run Dijkstra's algorithm to get the shortest path tree and check end-to-end QoS from root to any other TAP on this tree. Then, we increase reachability counter of the TAP without violating end-to-end QoS requirements. Repeat **Step 2** until all TAPs are assigned to associated backhauls

Step 3 Select the TAP of highest reachability counter, and installed it as a backhaul

---

Theorem 2 *The LR-based backhaul and routing assignment algorithm obtains the near-optimal solution in* $O(I|B||K||V|^3)$

Proof For each assigned backhaul b with configuration k, each source node refines its better TAP and backhaul with Dijkstra's algorithm, which is known to $O(|V|^2)$. To achieve near-optimal the LR-based algorithm is set to execute I iterations. Thus, overall these algorithm, our proposed algorithms find the solution to minimum cost backhaul deployment in $O(I|B||K||V|^3)$. □

5 Computational Experiments

In this section, we conduct several computational experiments to examine how good of the quality of our solution approach. In the mean time, for the purpose of evaluating the solution quality, we implement three simple algorithms for comparison.

5.1 Experiment Environment

We implement RA (i.e., the short name of Random Algorithm that is referred to [12]) and GA (i.e., the short name of Greedy Algorithm that is referred to [28] and [15]). Random Algorithm deploy backhaul and decide the sequence of paths selection randomly, while Greedy Algorithm chooses minimum deployment cost backhaul and minimum data flow first. We also implement MRFA (i.e., the short name of Minimum Resource First Algorithm which is a kind of modified greed algorithm) that chooses minimum usage of network resources first and use the same deploy backhaul manner as LR in order to conserve the righteousness for comparison with different sequence of paths selection.

The Lagrangean dual solution denotes as "Z_D" and primal feasible solution as "Z_{IP}". We use two metrics - "Gap" and "Improvement Ratio" to evaluate our solution quality. Where Gap is calculated by $\left|(Z_{IP} - \mathrm{Z_D})/\mathrm{Z_D}\right| \cdot 100\%$. and Improvement Ratio is calculated by $\left|(RA - Z_{IP})/Z_{IP}\right| \cdot 100\%$.

5.2 Experiment Results

In order to test the solution quality of our algorithm, we design several scenarios with different features: (a) grid network with different number of TAPs; (b) random network with different number of TAPs; (c) hexagonal network with different number of TAPs; and (d) random network with different data flow. The average traffic requirement within each TAP is given and follow uniform distribution for the previous 3 features, but the traffic flow is variance for the experiment (d).

1. Grid network with different number of TAPs
   The grid network is constructed the set of given nodes based on the square shape. Figure 4 shows the experiment results. In this experiment, the number of backhauls increases (i.e., the backhaul deployment cost increases) when the traffic requirement increases. The increasing ratio follows unobvious exponential curve since the number of TAPs increases exponentially. In addition, because the relationship between relative decision variable is complex, the experiment results show that the gap increases as the number of TAP nodes increases. The improvement ratio is much better than RA and outperforms GA by at least 9.5%.
2. Random Network with Different Number of TAPs
   In this experiment, the TAP nodes are random deployed in a square area. The area increases as the number of nodes increases, namely the average distance is the same with different size of networks. Figure 5 show the experiment results. The curve is not as smooth as previous experiment is caused by random deployment. The gap increases as the number of TAP nodes increases, too. However, the gap is here less than the grid network. The reason is that the number of links is larger and some nodes may closer. The improvement ratios are similar to grid networks.
3. Hexagonal network with different number of TAPs
   In this experiment, the TAP nodes are deployed based on the fixed point in a hexagonal shape. Figure 6 show the experiment results. The gap is close to grid network. However, the gap is small since the number of outgoing links is larger. The improvement ratios are similar to the above two network structures. When the number of TAPs is 9, the improvement ratio is 0 due to the number of candidate backhaul is less. However, when the number of nodes increases, the improvement ratio is at least 7.91%.

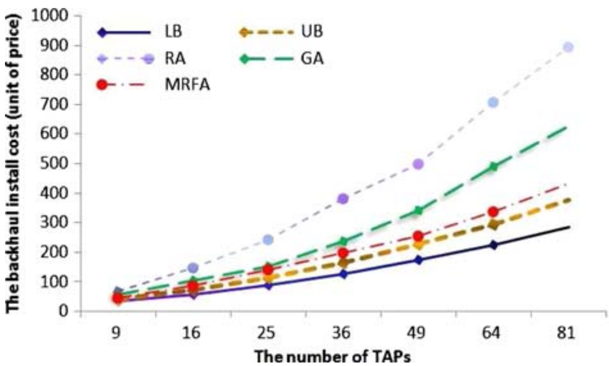


Fig. 4 Deployment cost of grid network with different number of TAPs

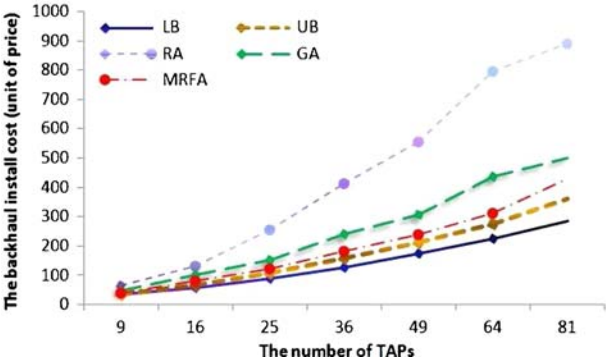


Fig. 5 Deployment cost of random network with different number of TAPs

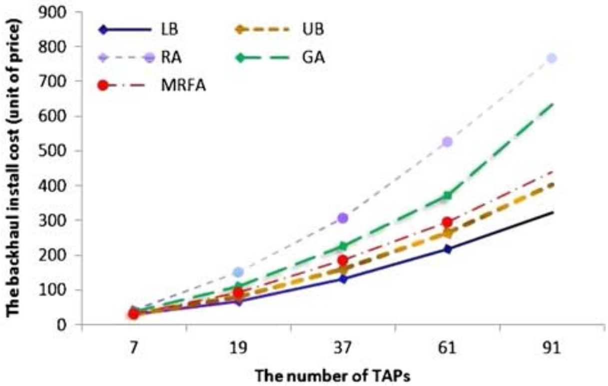


Fig. 6 Deployment cost of hexgonal network with different number of TAPs

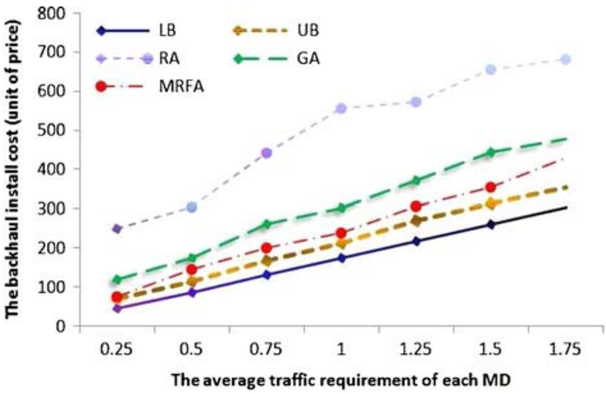


Fig. 7 Deployment cost of random network with different data flow (49 TAPs)

4. Random Network with Different Data Flow We extend the experiment (b) with different amount of data flows (λ) with 49 TAPs. Figure 7 shows the experiment results. When the traffic load increases, the gap decreases. The reason is that the number of selecting paths is less and limited by the QoS requirements. Overall, our proposed algorithm outperforms other comparing algorithms by 5.87, 20.6%.

According to the experiment results, shown in Figs. 4, 5, 6, and 7, we can find that the cost of backhaul deployment increases with the number of TAPs and MDs. And the LR-based algorithm always outperforms other algorithms. LR-based algorithm and MRFA adopt the concept of reachbility to deploy the backhauls. We can see these two algorithms performs well eminently by comparison with random deploy manor and greedy deploy manner. Therefore, we can take this deploy manner to deploy backhaul economically and effectively.

Although LR-based algorithm and MRFA use the same deploy manner, the proposed algorithm performs better than MRFA. Accordingly, we discover that the sequence of routing path selection has impact to the experimental outcomes. This is because the TAPs routing previously consume part of network resources and the following TAPs restricted to less network resources. Therefore, the more previous TAP has more flexible in paths selection.

The experiment results demonstrated that the proposed algorithms arrive at near optimal solutions with gaps of less than 31.26%. In comparison with some existing algorithm, the improvement ratio is more than 11.02% when the number of TAP nodes is larger than 49.

6 Conclusions

In this paper, we considered end-to-end QoS routing in WMNs with an emphasis on the problem of backhaul deployment. Using an LR-based approach, a number of algorithms were proposed in this paper: backhaul assignment, routing, TAP selection, and backhaul adjustment. These algorithms yield the minimum objective function value and also achieve balanced traffic loads of the backhauls, QoS requirements. The proposed algorithms were evaluated by comparisons with lower bounds that was obtained from an Lagrangean Dual problem. This research demonstrates that although the constraints involved in supporting fairness and QoS routing are significant, this NP-complete problem can be solved by decomposing it into sub-problems. Therefore, it provides insights for cost effective broadband coverage over wide areas.

In our future work, other heuristic-based approaches, such as simulated annealing approach and genetic algorithm, are adopted to address this complex problem.

Appendix

A The Solutions of Seven Subproblems

Each of the following seven sub-problems, generated from the Lagrangean problem (LR), is related to decision variables.

Subproblem SUB1 (related to decision variable η_{bk})

Objective function:

$$
\begin{aligned}
&Z_{SUB1}\left(\mu^1, \mu^2_{bs}, \mu^5_b, \mu^6_b, \mu^7_b, \mu^{13}_b\right) \\
&\quad = \min \sum_{b\in V}\sum_{k\in K}\Bigg[c_b + \Phi_b(k) - Q_b(k)\,\mu^1 - \sum_{s\in V}\mu^2_{bs} - \mu^5_b + M_1\mu^6_b + \mu^7_b \\
&\qquad\qquad + (M_3 - Q_b(k))\,\mu^{13}_b\Bigg]\eta_{bk}
\end{aligned}
$$

subject to: (2).

Sub-problem (SUB1) can be further decomposed into $|V|$ independent subproblems. For each candidate backhaul b, the objective function is $\min \sum_{k\in K} (c_b + \Phi_b(k) - Q_b(k)\mu^1 - \sum_{s\in V} \mu_{bs}^2 - \mu_b^5 + M_1\mu_b^6 + \mu_b^7 + (M_3 - Q_b(k))\mu_b^{13})\eta_{bk}$, we calculate the coefficient $[c_b + \Phi_b(k) - Q_b(k)\mu^1 - \sum_{s\in V} \mu_{bs}^2 - \mu_b^5 + M_1\mu_b^6 + \mu_b^7 + (M_3 - Q_b(k))\mu_b^{13}]$ for each configuration k. Then, we find the smallest coefficient for all configuration k of candidate bachhaul b. If the smallest coefficient is negative, the corresponding η_{bk} is set to be 1; otherwise set all configuration k to be 0.

Subproblem SUB2 (related to decision variable z_{bs})

Objective function:

$$Z_{SUB2}\left(\mu_{bs}^2, \mu_{bs}^3\right) = \min \sum_{s\in V}\sum_{b\in V} [\mu_{bs}^2 + \mu_{bs}^3] z_{bs}$$

subject to: (3).

This problem can be further decomposed into $|V|$ independent subproblems. For each source TAP s, the objective function is $\min \sum_{b\in V} \left[\mu_{bs}^2 + \mu_{bs}^3\right] z_{bs}$. The algorithm to solve the decomposed subproblem is stated as follows:

Step 1 Compute the coefficient $\left(\mu_{bs}^2 + \mu_{bs}^3\right)$ of z_{bs} for each candidate backhaul b, and sort it in ascending order

Step 2 Select the first order coefficient and assign the corresponding decision variable z_{bs} to 1; otherwise 0

Subproblem SUB3 (related to decision variable x_p)

Objective function:

$$Z_{SUB3}\left(\mu_{bs}^3, \mu_{suv}^4\right) = \min \sum_{s\in V}\sum_{b\in V}\sum_{p\in P_{bs}} \left[-\mu_{bs}^3 + \sum_{u\in V}\sum_{v\in V} \mu_{suv}^4 \delta_{p(u,v)}\right] x_p$$

subject to: (7).

This problem can be further decomposed into $|V|$ independent shortest path problems with non-negative arc weights. Each shortest path problem can be easily solved by the Dijkstra's algorithm. If the coefficient of x_p is negative, then set x_p to 1, otherwise 0.

Subproblem SUB4 (related to decision variable a_s)

Objective function:

$$Z_{(SUB4)}\left(\mu_s^9, \mu_{suv}^{10}, \mu_{suv}^{11}, \mu_s^{13}\right) = \min \sum_{s\in V} \left[-\mu_s^9 + \sum_{u\in V}\sum_{v\in V} \mu_{suv}^{10} - \sum_{u\in V}\sum_{v\in V} \mu_{suv}^{11} + \mu_s^{13}\right] a_s$$

subject to: (16), (17).

The proposed algorithm for solving (SUB4) is described as follows:

Step 1 Reset all a_s to 0

Step 2 For each TAP s, we compute the coefficient $\left(-\mu_s^9 + \sum_{u\in V}\sum_{v\in V} \mu_{suv}^{10} - \sum_{u\in V}\sum_{v\in V} \mu_{suv}^{11} + \mu_s^{13}\right)$ for each a_s

Step 3 Find the unset a_s with smallest coefficient. If found, then set it to $\overline{\overline{C_s}}$, else **stop**

Step 4 Repeat **Step 3** until the total data rate required to be transmitted of all TAPs is equal to or large than the total incoming flow of all MDs

---

Subproblem SUB5 (related to decision variable $\gamma_{s(u,v)}$)
Objective function:

$$
\begin{aligned}
&Z_{SUB5}\left(\mu_{suv}^{10}, \mu_{suv}^{11}, \mu_{uv}^{12}, \mu_{v}^{13}\right) \\
&\quad = \min \sum_{s\in V}\sum_{u\in V}\sum_{v\in V}\left[-\mu_{suv}^{10} + \mu_{suv}^{11} + \mu_{uv}^{12} + \mu_{v}^{13}\right]\gamma_{s(u,v)}
\end{aligned}
$$

subject to: (22).

This problem can be further decomposed into $|V|$ independent subproblems. For each source TAP s, The proposed algorithm for solving (SUB5) is described as follows:

Step 1 For each TAP v, we compute the coefficient $\left(-\mu_{suv}^{10} + \mu_{suv}^{11} + \mu_{uv}^{12} + \mu_{v}^{13}\right)$ for each $\gamma_{s(u,v)}$

Step 2 For all incoming links of TAP v, we find the smallest coefficient. If the total incoming flow of TAP v does not exceed the nodal capacity $\overline{C}_v$ and the smallest coefficient is negative then we set the corresponding $\gamma_{s(u,v)}$ to 1. Repeat **Step 2** for all TAP v

Step 3 Set the other incoming flow $\gamma_{s(u,v)}$ to 0

Subproblem SUB6 (related to decision variable $y_{s(u,v)}$ and $f_{(u,v)}$)
Objective function:

$$
\begin{aligned}
&Z_{SUB6}\left(\mu_{suv}^{4}, \mu_{u}^{5}, \mu_{u}^{6}, \mu_{v}^{7}, \mu_{suv}^{10}, \mu_{ns}^{14}, \mu_{ns}^{15}, \mu_{uv}^{12}\right) \\
&= \min \sum_{u\in V}\sum_{v\in V}\left[\left[\sum_{s\in V}\left[-\mu_{suv}^{4} - \mu_{u}^{5} + \mu_{u}^{6} - \mu_{v}^{7} + M_2\mu_{suv}^{10}\right.\right.\right. \\
&\quad \left.\left. + \sum_{n\in N}\mu_{ns}^{14}F_{(u,v)}\left(f_{(u,v)}, C_{(u,v)}\right) + \sum_{n\in N}\mu_{ns}^{15}M_{(u,v)}\left(f_{(u,v)}, C_{(u,v)}\right)\right]y_{s(u,v)} - \mu_{uv}^{12}f_{(u,v)}\right] \\
&= \min \left[\left[\sum_{s\in V} -\mu_{suv}^{4} - \mu_{u}^{5} + \mu_{u}^{6} - \mu_{v}^{7} + M_2\mu_{suv}^{10} + \sum_{n\in N}\mu_{ns}^{14}F_{(u,v)}\left(f_{(u,v)}, C_{(u,v)}\right)\right.\right. \\
&\quad \left. + \sum_{n\in N}\mu_{ns}^{15}M_{(u,v)}\left(f_{(u,v)}, C_{(u,v)}\right)\right]y_{s(u,v)} - \mu_{uv}^{12}f_{(u,v)}\Bigg]
\end{aligned}
$$

subject to: (12) and (21).

This subproblem is complicated due to the coupling of $y_{s(u,v)}$ and $f_{(u,v)}$. It can be further decomposed into $|V \times V|$ independent subproblems. For each link (u, v).

For each (SUB6) can be solved analytically [32,7] by the algorithm stated as follows:

Step 1 Solve $y_{s(u,v)} f_{(u,v)} = -\sum_{b\in V}\mu^3_{suv} - \mu^4_u + \mu^5_u - \mu^6_v + M_2\mu^7_{suv} + \mu^{10}_s F_{(u,v)}\left(f_{(u,v)}, C_{(u,v)}\right) + \mu^{11}_s M_{(u,v)}\left(f_{(u,v)}, C_{(u,v)}\right) = 0$ for each TAP s, call them the break points of $f_{(u,v)}$

Step 2 Sorting these break points and denoted as $f^1_{(u,v)}, f^2_{(u,v)}, \ldots, f^n_{(u,v)}$

Step 3 At each interval, $f^i_{(u,v)} \le f_{(u,v)} \le f^{i+1}_{(u,v)}$, $y_{s(u,v)}\left(f_{(u,v)}\right)$ is 1 if $-\sum_{b\in V}\mu^3_{suv} - \mu^4_u + \mu^5_u - \mu^6_v + M_2\mu^7_{suv} + \mu^{10}_s F_{(u,v)}\left(f_{(u,v)}, C_{(u,v)}\right) + \mu^{11}_s M_{(u,v)}\left(f_{(u,v)}, C_{(u,v)}\right) \le 0$ and otherwise 0

Step 4 Within the interval, $f^i_{(u,v)} \le f_{(u,v)} \le f^{i+1}_{(u,v)}$, we can take calculus to find the local minimal

Step 5 The global minimum point can be found by comparing these local minimum points.

Subproblem SUB7 (related to decision variable κ_{ns})
Objective function:

$$Z_{SUB7}\left(\mu^8_s, \mu^9_s, \mu^{14}_{ns}, \mu^{15}_{ns}\right) = \min \sum_{n\in N}\sum_{s\in V}\left[\begin{array}{l}\theta_n\mu^8_s + \theta_n\mu^9_s + \mu^{14}_{ns}\overline{F}_{ns}(\theta_n, r_{ns}) + \\ \mu^{15}_{ns}\overline{M}_{ns}(\theta_n, r_{ns})\end{array}\right]\kappa_{ns}$$

subject to: (13).

This problem can be further decomposed into $|N|$ independent subproblems. For each MD n, the algorithm to solve the decomposed subproblem is stated as follows:

Step 1 Compute the coefficient $\left(\theta_n\mu^8_s + \theta_n\mu^9_s + \mu^{14}_{ns}\overline{F}_{ns}(\theta_n, r_{ns}) + \mu^{15}_{ns}\overline{M}_{ns}(\theta_n, r_{ns})\right)$ for each TAP s

Step 2 Find the smallest coefficient, then set the corresponding decision variable κ_{ns} to 1 and others to 0

References

1. Akyildiz, I. F., Wang, X., & Wang, W. (2005). Wireless mesh networks: A survey. *Elsevier Computer Networks, 47*(4), 445–487.
2. Akyildiz, I. F., Lee, W. Y., Vuran, M. C., & Mohanty, S. (2006). NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey. *Computer Networks, 50*(13), 2127–2159.
3. Barolli, L., Koyama, A., Suganuma, T., & Shiratori, N. (2003). GAMAN: A GA based QoS routing method for mobile ad-hoc networks. *Journal Of Interconnection Networks, 4*(3), 251–270.
4. Bruno, R., Conti, M., & Gregori, E. (2005). Mesh networks: Commodity multihop ad hoc networks. *IEEE Communications Magazine, 43*(3), 123–131.
5. Burton, M. (2002). Channel overlap factor calculations for 802.11b networks, White Paper/Cirond Technologies Inc. http://www.cirond.com/whitepapers/fourpoint.pdf.
6. Chen, S., & Nahrstedt, K. (1999). Distributed Quality-of-Service routing in ad-hoc networks. *IEEE Journal of Selected Areas in Communications, 17*(8), 1–18.
7. Cheng, K. T., & Lin, F. Y. S. (1995). Minimax end-to-end delay routing and capacity assignment for virtual circuit networks, *Proceeding of IEEE Globecom*, pp. 2134–2138.
8. Fisher, M. (1973). Optimal solution of scheduling problems using lagrange multipliers: Part I. *Operations Research, 21*(5), 1114–1128.

9. Fisher, M. (1981). The lagrangean relaxation method for solving integer programming problems. *Management Science, 27*(1), 1–18.
10. Gambiroza, V., Sadeghi, B., & Knightly, E. (2004). End-to-end performance and fairness in multihop wireless backhaul networks, *Proceeding of ACM MobiCom*, pp. 287–301.
11. Glaber, C. (1998). On the approximability of problems for cellular networks. Technical Report No. 208, University of Wuerzburg.
12. He, J., Verstak, A., Watson, L. T., Stinson, C. A., Ramakrishnan, N., Shaffer, C. A., Rappaport, T. S., Anderson, C. R., Bae, K., Jiang, J., & Tranter, W. H. (2002). Globally optimal transmitter placement for indoor wireless communication systems. *IEEE Transactions on Wireless Communications, 3*(6), 1906–1911.
13. Jain, K., Padhye, J., Padmanabhan, V. N., & Qiu, L. (2003). Impact of interference on multi-hop wireless network performance, *Proceeding of ACM MobiCom*, pp. 66–80.
14. Jun, J., & Sichitiu, M. (2003). The nominal capacity of wireless mesh networks. *IEEE Wireless Communications, 10*(5), 8–14.
15. Kamenetsky, M., & Unbehaun, M. (2002). Coverage planning for outdoor wireless LAN systems, *Proceeding of broadband communications; access, transmission, networking. International Zurich seminar on*, pp. 49-1–49-6.
16. Karrer, R., Sabharwal, A., & Knightly, E. (2004). Enabling large-scale wireless broadband: The case for TAPs. *ACM SIGCOMM Computer Communication Review, 34*(1), 27–32.
17. Konhäuser, W. (2006). Broadband wireless access solutions–progressive challenges and potential value of next generation mobile networks. *Wireless Personal Communications, 37*(3-4), 243–259.
18. Li, B. (2005). End-to-end fair bandwidth allocation in multi-hop wireless ad hoc networks. *Proceeding of IEEE ICDCS*, pp. 471–480.
19. Liu, L. G., & Feng, G. Z. (2007). Simulated annealing based multi-constrained QoS routing in mobile ad hoc networks. *Wireless Personal Communications, 41*(3), 393–405.
20. Maki, I., Shimonishi, H., Murase, T., & Murata, M. (2006). Hierarchically aggregated fair queueing (HAFQ) for per-flow fair bandwidth allocation. *IEICE Transaction on Communication, E89-B*(2), 427–435.
21. Padhye, J., Draves, R., & Zill, B. (2004). Routing in multi-radio, multi-hop wireless mesh networks. *Proceeding of ACM MobiCom*, pp. 114–128.
22. Salem, N. B., & Hubaux, J. P. (2005). A fair scheduling for wireless mesh networks, *Proceeding of IEEE WiMesh*, Santa Clara, CA.
23. Sherali, H. D., & Gyunghyun, C. (1996). Recovery of primal solutions when using subgradient optimization methods to solve lagrangian duals of linear programs. *Operations Research Letters, 19*(3), 105–113.
24. Sheu, S. T., & Chen, J. H. (2002). MRR^{2}RP: The multi-rate and multi-range routing protocol for ad hoc wireless networks, *Proceeding of IEEE ISCC*, pp. 435–440.
25. Sivakumar, R., Sinha, P., & Bharghavan, V. (1999). CEDAR: A core-extraction distributed ad hoc routing algorithm. *IEEE Journal on Selected Areas in Communications, 17*(8), 1454–1465.
26. Tang, J., Xue, G. L., & Zhang, W. Y. (2007). Cross-layer optimization for end-to-end rate allocation in multi-radio wireless mesh networks, *Wireless Networks* (in press).
27. Tsai, T. J., & Chen, J. W. (2005). IEEE 802.11 MAC protocol over wireless mesh networks: Problems and perspectives, *Proceeding of IEEE AINA*, pp. 60–63.
28. Tutschku, K., Mathar, R., & Niessen, T. (1999). Interference minimization in wireless communication systems by optimal cell site selection, *Proceeding of European personal mobile communication conference*, Paris, France.
29. Wang, Z., & Crowcroft, J. (1996). Quality-of-service routing for supporting multimedia applications. *IEEE Journal on Selected Areas in Communications, 14*(7), 1228–1234.
30. Wen, Y. F., & Lin, F. Y. S. (2007). Fair bandwidth allocation and end-to-end delay routing algorithms in wireless mesh networks. *IEICE Transactions on Communications, E90-B*(5), 1042–1051.
31. Wen, Y. F., & Lin, F. Y. S. (2006). The top load balancing forest routing in mesh networks, *Proceeding of IEEE CCNC*, pp. 468–472.
32. Yen, H. H., & Lin, F. Y. S. (2001). Near-optimal delay constrained routing in virtual circuit networks. *Proceeding of IEEE Infocom*, pp. 750–756.

Author Biographies

Yean-Fu Wen received the M.S. degree from the Department of Information Management, National Taiwan University of Science Technology, Taiwan, ROC, in 1998. He finished his Ph.D. study and got a doctoral degree from the Department of Information Management, National Taiwan University in July 2007. In August 2008, he joined the Department of Management Information Systems, National Chiayi University as an assistant professor. His research interests include network planning, performance optimization, and cross-layer technology in next-generation wireless networks. He is a member of IEEE.

Frank Yeong-Sung Lin received his B.S. degree in electrical engineering from the Electrical Engineering Department, National Taiwan University in 1983, and his Ph.D. degree in electrical engineering from the Electrical Engineering Department, University of Southern California in 1991. After graduating from the USC, he joined Telcordia Technologies (formerly Bell Communications Research, abbreviated as Bellcore) in New Jersey, USA, where he was responsible for developing network planning and capacity management algorithms. In 1994, Prof. Lin joined the faculty of the Electronic Engineering Department, National Taiwan University of Science and Technology. Since 1996, he has been with the faculty of the Information Management Department, National Taiwan University. His research interests include network optimization, network planning, network survivability, performance evaluation, high-speed networks, distributed algorithms, content-based information retrieval/filtering, biometrics and network/ information security.

Yeong-Cheng Tzeng received the M.S. degree from the Department of Information Management, National Taiwan University, Taiwan, ROC, in 2006. He joined the Division of Research and Development, ZyXEL Communications Corp. as an Engineer in the same year. His research interests include performance optimization and cross-layer technology in WiMax networks.

Cheng-Ta Lee received the M.S. degree from the Department of Information Management at the National Taiwan University of Science and Technology in 1997. Now, he is a Ph.D. candidate in the Department of Information Management at the National Taiwan University. Since September 1997, he is a Lecturer in the Department of International Business Management at the Lan Yang Institute of Technology. His research interests include performance optimization and wireless sensor networks.