Journal of Universal Computer Science, vol. 17, no. 14 (2011), 2064-2083
submitted: 25/3/11, accepted: 28/9/11, appeared: 1/10/11 © J.UCS

Optimization of Gateway Deployment with Load Balancing and Interference Minimization in Wireless Mesh Networks

Junzhou Luo
(School of Computer Science and Engineering, Southeast University
Nanjing, P.R. China
jluo@seu.edu.cn)

Wenjia Wu
(School of Computer Science and Engineering, Southeast University
Nanjing, P.R. China
wjwu@seu.edu.cn)

Ming Yang
(School of Computer Science and Engineering, Southeast University
Nanjing, P.R. China
yangming2002@seu.edu.cn)

Abstract: In a wireless mesh network (WMN), gateways act as the bridges between the mesh backbone and the Internet, and significantly affect the performance of the whole network. Hence, how to determine the optimal number and positions of gateways, i.e., gateway deployment, is one of the most important and challenging topics in practical and theoretical research on designing a WMN. Although several approaches have been proposed to address this problem, few of them take load balancing and interference minimization into account. In this paper, we study the **L**oad-balancing and **I**nterference-minimization **G**ateway **D**eployment **P**roblem (LIGDP), which aims to achieve four objectives, i.e. minimizing deployment cost, minimizing MR-GW path length, balancing gateway load and minimizing link interference. We formulate it as a multi-objective integer linear program (ILP) issue first, and then propose an efficient gateway deployment approach, called **LIGDP Heuristic**. The approach joints two heuristic algorithms, i.e., MSC-based location algorithm (MLA) and load-aware and interference-aware association algorithm (LIAA), to determine gateway positions and construct GW-rooted trees. Simulation results not only show that the trade-off between deployment cost and network performance can be achieved by adjusting R-hop, GW throughput and MR throughput constraints, but also demonstrate that, compared with other existing approaches, **LIGDP Heuristic** performs better on MR-GW path, load balancing and interference minimization without deploying more gateways.

Key Words: wireless mesh networks, gateway deployment, load balancing, interference minimization

Category: C.2.1, C.2.5, C.2.6, G.1.6

1 Introduction

As a promising technology to provide ubiquitous Internet access, wireless mesh networks (WMNs) [Akyildiz et al. 2005, Bruno et al. 2005, Zhang et al. 2006] have received more and more attention recently. This is due to their attractive characteristics, including low infrastructure cost, large coverage area, flexible deployment, easy maintenance and robustness. The most commonly used mesh architecture is referred to as two-tier WMNs, also called infrastructure WMNs [Akyildiz et al. 2005]. As shown in Figure 1, an infrastructure WMN consists of a backbone tier and an access tier. In the WMN, nodes are comprised of static mesh routers (MRs) and mobile mesh clients (MCs). A small fraction of MRs, called gateways (GWs), are configured with wired links, which act as the bridges between the WMN and the Internet. MRs, constituting the backbone tier, not only provide wireless connections for MCs in their respective domains as access points, but also relay traffic for other MRs as routers. MCs are mobile clients in the access tier, e.g., desktops, laptops, PDA and phones. To achieve Internet access, each MC associates with one of the nearest MRs, and connects to a gateway through wireless multi-hop forwarding.

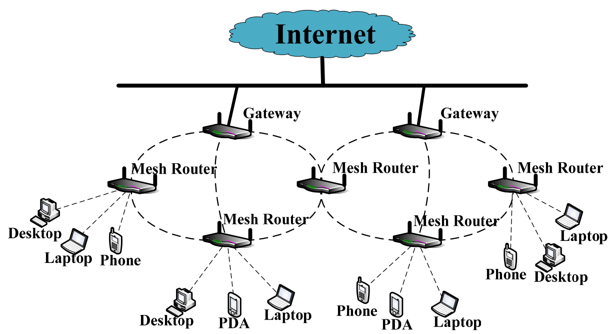

Figure 1: Architecture of Infrastructure WMN

In a WMN, relatively higher bandwidth and various quality of service (QoS) are required by mobile users. Furthermore, the bandwidth and QoS demands are keeping rising. Hence, it is urgent and challenging to improve the performance of WMNs, especially network throughput. There have been a lot of research works about WMNs in recent years, and most of them [Alicherry et al. 2005, Raniwala et al. 2005, Draves et al. 2004] focus on routing, channel assignment, link schedule, etc. In these works, network performance is improved in the case of a deployed WMN. Moreover, network performance can also be optimized in the phase of WMN design, e.g., optimizing the number and positions of network nodes. Nevertheless, the research in this respect is still in its infancy and full of challenges.

Gateway deployment, as an indispensable part of WMN design, is to determine the optimal number and positions of gateways in a WMN. Since most users are primarily interested in Internet access, gateways that connect the WM-N with the wired Internet are inevitably become critical nodes. In other words, the number of gateways and their positions will have a great impact on network performance. Adding additional gateways can share the load of congested gateways and improve network throughput. Meanwhile, choosing proper gateway positions can also optimize network topology and traffic distribution. However, due to expensive construction of wired links in gateways, deploying more gateways will significantly increase the cost. Hence, there is a trade-off between deployment cost and network performance, and it is important to minimize the number of gateways while satisfying network performance requirements.

As discussed above, gateway deployment is a practical and meaningful problem, and should be well addressed. Several research efforts [Aoun et al. 2006, Bejerano 2004, Chandra et al. 2004, Durocher et al. 2008, He et al. 2008, Li et al. 2008, Marco, 2008, Zhou et al. 2007, Zeng et al. 2009] have been made to place gateways strategically in WMNs. These approaches aim to minimize the number of gateways with various network parameters considered, such as traffic demand, network throughput, node capacity, link bandwidth, and path length. Unfortunately, few of them consider load balance and interference minimization. Although Zeng et al. [Zeng et al. 2009] have proposed two algorithms to address the load-balancing gateway deployment problem, their solutions doesn't take interference minimization into account.

Traffic aggregation make gateways become the bottlenecks of WMNs. Imbalance of gateway load will lead to heavy congestion of partial gateways, and dramatically influence the network throughput. Interference is fundamental to wireless networks, due to the broadcast nature of wireless medium [Lukovszki et al. 2006]. In WMNs, if gateways are deployed densely, excessive interference among them will significantly affect network performance. Hence, it is necessary to consider load balancing and interference minimization when determining the positions of gateways. In this paper, we address the **L**oad-balancing and **I**nterference-minimization **G**ateway **D**eployment **P**roblem (LIGDP), i.e., considering load balancing and interference minimization in the process of gateway deployment. We formulate the LIGDP problem as a multi-objective integer linear program (ILP) issue first, and then propose an efficient gateway deployment approach to address the problem. Finally, we evaluate the performance of the proposed approach via simulations. In summary, this paper has four-fold main contributions:

(1) We find that existing gateway deployment approaches may lead to imbalance of gateway load and heavy interference, and define the LIGDP problem that considers load balancing and interference minimization.

(2) We describe the LIGDP problem as a multi-objective optimization problem,

and formulate it using ILP. Four objective functions are defined, in terms of deployment cost, MR-GW path length, load balancing and interference minimization.

(3) We propose a novel gateway deployment approach, called **LIGDP Heuristic**. It includes two heuristic algorithms, i.e., MSC-based location algorithm, and load-aware and interference-aware association algorithm.

(4) We evaluate the proposed **LIGDP Heuristic** on the number of gateways, the average length of MR-GW paths, the standard deviation of gateway load and the average interference of active links. We discuss the impact of R-hop, GW throughput and MR throughput constraints, and compare **LIGDP Heuristic** with other existing approaches.

The rest of the paper is organized as follows. Section 2 reviews the existing gateway deployment approaches for WMNs. In Section 3, we introduce system model, describe the LIGDP problem and formulate the problem as an ILP issue. In Section 4, we propose an efficient heuristic approach, called **LIGDP Heuristic**. In Section 5, the performance of the proposed approach is evaluated via simulations. Finally, Section 6 concludes the paper and puts forward some possible future directions.

2 Related Work

WMNs have been an attractive research area in the last few years. Akyildize et al. [Akyildiz et al. 2005] presented a survey of WMNs, and discussed some open research issues. Among these research issues, gateway deployment is one of the most important challenges, and has become a hot research topic. Some researchers have begun to study this problem and proposed several approaches.

Gateway deployment is a combinatorial optimization problem, and has been proved to be NP-hard [He et al. 2008]. In the literature, existing approaches can be mainly classified into two categories: heuristics and meta-heuristics (in particular genetic algorithms).

Chandra et al. [Chandra et al. 2004] addressed the gateway deployment problem aimed at minimizing the number of gateways while satisfying the bandwidth requirements of all clients. The problem was formulated as a network flow problem, and a max-flow min-cut based algorithm was developed for gateway selection. They defined three wireless link models, and developed heuristic algorithms for each model. Bejerano [Bejerano 2004] addressed the gateway deployment problem as a variant of the capacitated facility location problem (CFLP), and proposed a clustering algorithm. Each gateway served a cluster of its nearby MRs, and a spanning tree rooted at the gateway (cluster head) was used for message delivery. Bejerano's approach [Bejerano 2004] took two steps. The first step was to find a minimal number of disjoint clusters containing all the nodes subject to an upper bound on clusters' radius. The second step was to place a spanning tree in each cluster and subdivide the clusters which violate the relay load or cluster size constraints. Aoun et al. [Aoun et al. 2006] and He et al. [He et al. 2008] improved Bejerano's work [Bejerano 2004]. They also adopted a clustered view, but their approaches were different. Aoun et al. [Aoun et al. 2006] proposed a recursive gateway selection approach. A one-hop dominating-set was first formed from the original network graph and this result was used as the input to the next recursion. The greedy dominating-set searching operation continued until the cluster radius reached R, which was the predefined upper bound of cluster radius. He et al. [He et al. 2008] also proposed heuristic approaches to address the gateway deployment problem. The optimization objectives included not only minimizing the number of gateways, but also minimizing the average length (hop count) of MR-GW paths. They developed two heuristic algorithms: degree based greedy dominating tree set partitioning (Degree based GDTSP) and weight based greedy dominating tree set partitioning (Weight based GDTSP).

In addition, the gateway deployment problem can also be addressed by genetic algorithms. Marco [Marco, 2008] proposed a genetic algorithm for gateway deployment optimization. Zeng et al. [Zeng et al. 2009] focused on the optimization of the cost and load balance in the gateway deployment strategy, and proposed a heuristic algorithm and a genetic algorithm respectively. The genetic algorithm had the advantage of global search for multiple goals, and got a better solution at the price of computing complexity.

It is clear that gateway deployment is a hot research topic currently, and several approaches have been proposed. However, there are still some deficiencies in the existing approaches. Most of them simplify the problem without link interference considered. It is necessary to consider load balancing in the phase of gateway deployment, but this is not included in most existing research. There have been several works on load-balancing routing in WMNs [Hsiao et al. 2001, Bejerano et al. 2007, Hundewale et al. 2007, Huang et al. 2009]. Hsiao et al. [Hsiao et al. 2001] introduced two kinds of load balanced trees for wireless access networks which were fully load-balanced trees and top load-balanced trees. They also provided an algorithm to readjust or rebalance an unbalanced tree. Bejerano et al. [Bejerano et al. 2007] studied the load balance problem in WMNs and proposed an approximation algorithm to balance load of network nodes while maximizing network utilization. However, these works are under the circumstances that network topology has been constructed in advance.

In this paper, we aim to address the LIGDP problem that taking load balancing and interference minimization into account. We first formulate this problem using the ILP method. Then we propose an efficient approach called **LIGDP Heuristic**. Finally, the performance of **LIGDP Heuristic** is evaluated.

3 LIGDP Problem

As shown in Figure 1, the WMN consists of a backbone tier for interconnection between MRs and an access tier for connection between MRs and MCs. To address the LIGDP problem, we only focus on the topology of the backbone tier, and ignore MCs in the access tier for simplicity. In this section, we first describe the assumptions and notations. Then, we describe the LIGDP problem. Finally, we formulate this problem as an ILP issue. For the sake of clarity, Table 1 summarizes the notations used in the following analysis.

Table 1: Summary of Notations

Symbol	Definition	Symbol	Definition
$G(M, L)$	WMN backbone	M	Set of MRs
L	Set of wireless links	n	Number of MRs
k	Number of gateways	m_i, m_j	MR node
M_g	Set of gateways	$l_{i,j}$	Link between two MRs
$d_{i,j}$	Distance between two MRs	(i_x, i_y)	Coordinate of MR m_i
r_t	Radius of transmission range	$N(m_i)$	MR's Neighborhood
$P_{i,j}$	Transmission path	$h_{i,j}$	Path length of $P_{i,j}$
R	Maximum path length	f_i	Traffic aggregated by MR m_i
$N_R(m_i)$	R-hop neighborhood	$i_{i,j,p,q}$	Interference between two links
C_m	MR throughput capacity	r_s	Radius of interference range
C_g	Gateway throughput capacity	AL	Set of active links
$N_s(m_i)$	Interference neighborhood	$I(l_{i,j})$	Link's interference domain

3.1 Assumptions and Definitions

The WMN backbone is viewed as an undirected graph $G(M, L)$ with MR set M and link set L, and the number of MRs is denoted as n. $M = \{m_1, \ldots, m_n\}$ is the set of MRs, which are deployed on a two dimensional Euclidean plane. The position of MR m_i is static after deployment, and denoted by (i_x, i_y), where i_x is its x-coordinate and i_y is its y-coordinate. Let $d_{i,j}$ be the Euclidean distance between MRs m_i and m_j, i.e., $d_{i,j} = \sqrt{(i_x - j_x)^2 + (i_y - j_y)^2}$. We assume that each MR m_i has a circular transmission range with radius r_t, and its neighborhood, denoted by $N(m_i)$ (including m_i), is the set of nodes that reside within its transmission range, i.e., $d_{i,j} < r_t$ for each $m_j \in N(m_i)$. A bidirectional wireless link exists between MR m_i and its neighbor m_j, which is represented as an edge $l_{i,j}$, and $l_{i,j} \in L$. MR m_i can communicate with MR m_j ($m_j \notin N(m_i)$) through multi-hop wireless forwarding. The transmission path is denoted by $P_{i,j}$, and $h_{i,j}$ is the length (hop count) of this path. Among n MRs, there are several MRs deployed as gateways. The set of gateways is represented by M_g, and their number is denoted as k.

We consider the delay and bandwidth in a WMN. In multi-hop transmission, delay occurs at each hop due to contention for wireless channel, packet queuing and processing. A few hops will cause significant delay. Meanwhile, as proposed by Li et al. [Li et al. 2001], when path length increases, the bandwidth obtainable for an originating node linearly decreases. Hence, MR-GW path length is bounded by R in our work. The R-hop neighborhood of MR m_i, denoted by $N_R(m_i)$, is the set of nodes that reside within R-hop range, i.e., $h_{i,j} \leq R$ for each $m_j \in N_R(m_i)$. In order to guarantee the efficient utilization of bandwidth, we bound the gateway throughput and MR throughput by C_g and C_m $(C_g > C_m)$ based on the capacity of wireless links.

In a WMN, each MR m_i aggregates traffic from MCs, and its traffic is denoted as f_i. Since most users in WMNs are primarily interested in Internet access, traffic patterns are different from a pure multi-hop wireless network. The traffic in a WMN is not randomly generated between arbitrary pairs of MRs, but mostly directed to/from gateways that connect the WMN to the wired Internet. So in this paper, we only consider the traffic directed to/from gateways. We can construct GW-rooted relay trees [He et al. 2008] for traffic forwarding in wireless mesh backbone, as illustrated in Figure 2. In the example of a WMN topology, MRs m_7 and m_{15} are configured as gateways, and two spanning trees rooted at m_7 and m_{15} are generated. Each MR associates with a gateway through attaching to a GW-rooted tree. MR-GW path is the path between a MR and its associated gateway in the GW-rooted tree, e.g., the MR-GW path of MR m_3 is the path between m_3 and its root m_7 in the m_7-rooted tree. Some of the links in L are included in GW-rooted trees, and the links employed for data transmission are active links, denoted by AL.

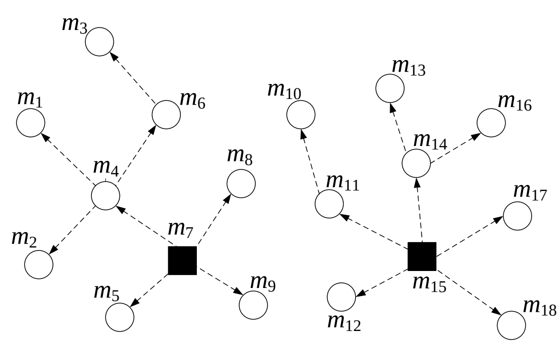

Figure 2: GW-rooted Tree Topology

In wireless networks, interference prevents wireless devices from correctly receiving packages, and can be characterized by a physical interference model or a protocol interference model [Gupta et al. 2000]. In the physical model, the packets transmitted from node m_i to node m_j are well received if the Signal to Interference and Noise Ratio (SINR) at node m_j (the receiver) is above a certain threshold. In the protocol model, the interference relationship between two links is determined by the specified interference range. The physical model is not suitable for large-scale WMNs due to its high computational complexity, and we can properly set the interference range so that a physical interference model can be accurately transformed to a protocol interference model [Shi et al. 2009]. Hence, we adopt the protocol interference model in this work. We assume that the interference range of every MR is its carrier sensing range, and use r_s to denote the radius of the circular range. $N_s(m_i)$ represents the interference neighborhood of MR m_i, which is the set of nodes that can interfere with node m_i. A transmission over link $l_{i,j}$ is successful when the nodes in $(N_s(m_i) \cup N_s(m_j)) \setminus \{m_i, m_j\}$ are silent during this transmission. Therefore, the interference domain of link $l_{i,j}$, denoted as $I(l_{i,j})$, is the set of links, of which one node or both two nodes are in $N_s(m_i) \cup N_s(m_j)$. We use a binary variable $i_{i,j,p,q}$ to represent the interference relation between two links $l(i,j)$ and $l(p,q)$. The variable $i_{i,j,p,q} = 1$ if $l_{p,q} \in I(l_{i,j})$, and $i_{i,j,p,q} = 0$ otherwise.

3.2 Problem Description

Gateway deployment, in its simplest form, is to select k nodes (k is not fixed before) from a WMN with n MRs ($k \leq n$) to serve as gateways and construct the corresponding GW-rooted trees, while optimization objectives are achieved and constraints are satisfied. In this paper, we plan to address the LIGDP problem, i.e., the gateway deployment problem with load balancing and interference minimization considered. It is a multi-objective optimization problem, which has multiple optimization objectives and several constraints.

Based on our network model and analysis of this problem, four objectives are considered, i.e., minimizing deployment cost, minimizing MR-GW path length, balancing gateway load and minimizing link interference. The four optimization objectives are defined as follows:

- **Minimizing deployment cost**: Because of expensive deployment of wired links in gateways, the deployment of additional gateways will increase the investment cost significantly. Hence, it is necessary to minimize the deployment cost, while satisfying the performance requirements.
- **Minimizing MR-GW path length**: As discussed above, in multi-hop forwarding, MR-GW path should be shortened to improve network performance, such as available bandwidth improvement and transmission delay reduction. Hence, minimization of MR-GW path length is another objective in the problem.
- **Balancing gateway load**: All Internet traffic passes through gateways, and may cause one or more network bottlenecks due to gateway congestion. Nevertheless, the bottlenecks can by avoided through balancing gateway load. Therefore, we plan to achieve load balancing among gateways in the phase of gateway deployment.

- **Minimizing link interference**: Interference significantly affects the efficiency of resource utilization in wireless networks. So interference among links should be avoided as much as possible.

While k gateways are selected from n MRs, it should satisfy several constraints. We model the problem with the following constraints.

- **R-hop constraint**: As discussed above, the length of MR-GW path should be bounded by R hops.
- **Full association constraint**: Adequate gateways should be located such that all MRs can achieve MR-GW association. This means that each MR should have at least one gateway in its R-hop range so that all MRs have the Internet accessibility by single hop or multiple hops.
- **GW throughput constraint**: Each gateway has the upper bound of the throughput, denoted as C_g. Therefore, the amount of MRs associated with a gateway is limited, and the traffic generated from its associated MRs cannot exceed its maximum throughput capacity.
- **MR throughput constraint**: Each MR has its throughput capacity, denoted as C_m. Therefore, the traffic (local traffic and relaying traffic) passing through a MR cannot exceed its maximum throughput capacity.
- **GW-rooted relay tree constraint**: A gateway and its associating MRs maintain a tree structure. The traffic is forward by the edges (wireless links) in the tree, and directed to/from the root (gateway).

3.3 ILP Formulation

In this section, we formulate the LIGDP problem as an ILP issue with four optimization objectives and several constraints. In the ILP formulation, the gateway location variable

$$x_i = \begin{cases} 1, & \text{MR } m_i \text{ is selected as a gateway} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

, the MR-GW association variable

$$y_{i,j} = \begin{cases} 1, & \text{MR } m_j \text{ associates with GW } m_i \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

, the MR ancestor variable

$$z_{i,j} = \begin{cases} 1, & \text{MR } m_i \text{ is ancestor node of MR } m_j \text{ in the GW-rooted tree} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

and the MR parent variable

$$e_{i,j} = \begin{cases} 1, & \text{MR } m_i \text{ is parent node of MR } m_j \text{ in the GW-rooted tree} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

are defined for all $i, j = 1, 2, \ldots, n$.

Our objective functions are formulated as follows:

$$\min \sum_{i=1}^{n} x_i \tag{5}$$

$$\min \frac{1}{n} \cdot \sum_{i=1}^{n} \sum_{j=1}^{n} y_{i,j} \cdot h_{i,j} \tag{6}$$

$$\min \max_{i=1,2,\ldots,n} \sum_{j=1}^{n} y_{i,j} \cdot f_i \tag{7}$$

$$\min \max_{e_{i,j}=1} \sum_{p=1}^{n} \sum_{q=1}^{n} i_{i,j,p,q} \cdot e_{p,q} \tag{8}$$

subject to

$$y_{i,j} \leq x_i \tag{9}$$

$$y_{i,j} \cdot h_{i,j} \leq R \tag{10}$$

$$\sum_{i=1}^{n} y_{i,j} = 1 \quad \text{for all } j = 1, 2, \ldots, n \tag{11}$$

$$\sum_{j=1}^{n} y_{i,j} \cdot f_j \leq C_g \quad \text{for all } i = 1, 2, \ldots, n \tag{12}$$

$$(1 - y_i) \cdot \sum_{j=1}^{n} z_{i,j} \cdot f_j \leq C_m \quad \text{for all } i = 1, 2, \ldots, n \tag{13}$$

$$z_{k,j} \cdot y_{i,j} \leq y_{i,k} \quad \text{for all } i, j, k = 1, 2, \ldots, n \tag{14}$$

$$e_{i,j} \cdot h_{i,j} \leq z_{i,j} \quad \text{for all } i, j = 1, 2, \ldots, n \tag{15}$$

where

$$x_i, y_{i,j}, z_{i,j}, e_{i,j} = 0, 1 \quad \text{for all } i, j = 1, 2, \ldots, n \tag{16}$$

Formula (5) represents the objective of minimizing deployment cost, i.e., minimizing the number of gateways. Formula (6) shows the objective of minimizing the MR-GW path length. Formula (7) denotes that the objective of balancing gateway load is achieved by minimizing the load of congested gateways. Formula (8) indicates the objective of minimizing link interference. The above four objectives are subjected to the following constraints. Inequation (9) denotes that m_j associates with m_i only when m_i is selected as a gateway. Inequation (10) shows that MR-GW association is limited by R-hop constraint. Equation (11) guarantees that each MR associates with a gateway. Inequation (12) and (13) provide an upper bound on gateway throughput and MR throughput respectively. Inequation (14) shows that if MR m_k is ancestor node of m_j, m_k and m_j must associates with the same gateway. Inequation (15) indicates that m_i is parent node of m_j only when m_i is ancestor node of m_j and $h_{i,j} = 1$.

As discussed above, the LIGDP problem is formulated as a multi-objective ILP issue. We can transfer multiple objectives to a single objective by forming a weighted combination of different objectives, and solve this problem by ILP solvers [He et al. 2008], such as MATLAB, CPLEX and LINGO. However, the running time of ILP solvers is too long for large-scale networks because of large solution space. Hence, we propose an efficient heuristic approach to address this problem in the next section.

4 LIGDP Heuristic

When n MRs are deployed in a WMN, positions of MRs are the candidates for gateway deployment. The optimal solution of gateway deployment can be found by a brute-force search. In the worst case, the brute-force search totally results in $2^n - 1$ choices, and at each choice it involves the evaluation of optimization objectives and the validation of above-listed constraints. Although the brute-force approach is simple, its complexity suffers from exponential growth. Thus, it is impractical when n increases.

In this paper, we propose an efficient approach, called **LIGDP Heuristic**, that is able to find an approximating optimal solution of gateway deployment in a WMN. To address the LIGDP problem, it requires solving two subproblems: gateway location and MR-GW association. **LIGDP Heuristic** involves two algorithms to solve the two subproblems respectively. The first one is MSC-based location algorithm (MLA) and the second one is load-aware and interference-aware association algorithm (LIAA). The MLA algorithm is to select minimum gateway positions while satisfying all the constraints. Then the LIAA algorithm is proposed to achieve MR-GW association and construct GW-rooted trees.

4.1 MSC-based Location Algorithm

In the minimum set cover (MSC) problem [Chvatal 1979], let S be a collection of subsets of a given universe U and S covers U, i.e., $\bigcup\limits_{s \in S} s = U$. The MSC problem is to find the minimum cardinality subset of S, denoted as $msc(S)$, to covers U, i.e. $\bigcup\limits_{s \in msc(S), msc(S) \subset S} s = U$.

We model a WMN backbone by an undirected graph $G(M, L)$, as described in Section 3. Gateway location is to select the minimum number of gateways that guarantee the full MR-GW association and satisfy all other constraints. MR m_j can associate with GW m_i if MR m_j is in the R-hop range of GW m_i. Hence, this is similar to looking for the MSC of S, while $S = \{N_R(m_1), N_R(m_2), \ldots, N_R(m_n)\}$ and $U = \{m_1, m_2, \ldots, m_n\}$. However, due to gateway throughput and MR throughput constraints, not all MRs in the R-hop of a gateway can associate with the gateway. This limitations make gateway location more complex than MSC.

Based on the above analysis, we propose a MSC-based location algorithm (MLA), employing greedy heuristic for MSC [Chvatal 1979]. Here, we use the term "cover" to represent that MR m_j is in the R-hop range of GW m_i, i.e., GW m_i covers MR m_j. MLA algorithm determines gateway positions iteratively, and in each iteration, it choose the MR with maximum weight as a gateway. In the

Algorithm 1 MSC-based Location Algorithm (MLA)

$M_g \leftarrow \emptyset, U \leftarrow M$
while $U \neq \emptyset$ **do**
    for all m_i such that $m_i \in U$ **do**
        weight calculation for $W_m(m_i)$
    end for
    select $m_g \in U$ with the maximum weight
    $M_g \leftarrow M_g \cup \{m_g\}$
    $U \leftarrow U \setminus \{m_g\}$
    determining $D(m_g)$ from $N_R(m_g)$ with R-hop, C_g and C_m constraints
    for all m_j such that $m_j \in D(m_g)$ **do**
        if $h_{g,j} = R$ **then**
            if $|M_g \cap N_R(m_j)| \geq 2$ **then**
                $U \leftarrow U \setminus \{m_j\}$
            end if
        else
            $U \leftarrow U \setminus \{m_j\}$
        end if
    end for
end while

algorithm, the weight calculation of MR m_i is defined as follows:

$$W_m(m_i) = \sum_{m_j \in N_R(m_i)} \frac{1}{(h_{i,j}+1) \cdot 2^{\lambda_j}} \tag{17}$$

where $\lambda_j = \sum_{m_k \in M_g, h_{k,j} \leq R} \frac{1}{h_{k,j}+1}$. The weight of MR m_i is not simply the number of nodes within its R-hop neighborhood, as presented in [Bejerano 2004]. Instead, it is a weighted sum, which is not only inversely proportional to the path length between MR m_i and MR m_j, but also inversely proportional to the number of selected gateways covering node m_j, where node m_j is within the R-hop neighborhood of node m_i. It means that MRs farther away will have a lower contribution to the value of $W_m(m_i)$, since it has negative impact on the throughput of network. Meanwhile, nodes covered by more gateways also contribute less to the value of $W_m(m_i)$, because nodes covered by fewer gateways should be covered preferentially.

The MLA algorithm is illustrated in Algorithm 1. Let M_g be the set of selected gateways, let U be the set of uncovered MRs and let $D(m_g)$ be the coverage domain of GW m_g. Initially, $M_g \leftarrow \emptyset$, $U \leftarrow M$. In each iteration, the algorithm selects the node $m_g \in U$ with the maximum weight value as a gateway and adds it to M_g. GW m_g and its covered MRs are removed from U. In order to mitigate imbalance of MRs covered by each gateway, the boundary MRs in the

R-hop range of GW m_g will stay in U if they are only covered by one gateway. This step is done until $U = \emptyset$.

4.2 Load-aware and Interference-aware Association Algorithm

Algorithm 2 Load-aware and Interference-aware Association Algorithm (LIAA)

$AL \leftarrow \emptyset, CL \leftarrow L, CM \leftarrow M \setminus M_g$, $t \leftarrow 0$, $tree(i) \leftarrow null$ for all $i = 1, 2, ..., n$
for $i = 1$; $i \leq n$; $i++$ **do**
  if $m_i \in M_g$ **then**
    $t \leftarrow t + 1$, $M_t \leftarrow \{m_i\}$, $L_t \leftarrow \emptyset$, $LD_t \leftarrow f_i$, $rt_t \leftarrow i$, $tree(i) \leftarrow t$
  end if
end for
while $|AL| < n - |M_g|$ **do**
  $weight \leftarrow 1$, $optLink \leftarrow null$
  for all $l_{i,j}$ such that $l_{i,j} \in CL$ **do**
    if $m_i \notin CM$ *and* $m_j \notin CM$ **then**
      $CL \leftarrow CL \setminus \{l_{i,j}\}$
    end if
    if $(m_i \notin CM$ *and* $m_j \in CM)$ *or* $(m_i \in CM$ *and* $m_j \notin CM)$ **then**
      $t \leftarrow tree(i)$ or $tree(j)$
      if Satisfying R-hop, C_m, and C_g constraints **then**
        weight calculation for $W_l(l_{i,j}, t)$
        if $W_l(l_{i,j}, t) < weight$ **then**
          $weight \leftarrow W_l(l_{i,j}, t)$, $optLink \leftarrow l_{i,j}$
        end if
      else
        $CL \leftarrow CL \setminus \{l_{i,j}\}$
      end if
    end if
  end for
  if $optLink \neq null$ **then**
    $AL \leftarrow AL \cup \{optLink\}, CL \leftarrow CL \setminus \{optLink\}$
    $m_a \leftarrow$ one end of $optLink$, $m_b \leftarrow$ the other end of $optLink$
    if $m_a \notin CM$ **then**
      $t \leftarrow tree(a)$, $CM \leftarrow CM \setminus \{m_b\}$
      $tree(b) \leftarrow t$, $M_t \leftarrow M_t \cup \{m_b\}$, $L_t \leftarrow L_t \cup \{l_{a,b}\}$, $LD_t \leftarrow LD_t + f_b$
    else
      $t \leftarrow tree(b)$, $CM \leftarrow CM \setminus \{m_a\}$
      $tree(a) \leftarrow t$, $M_t \leftarrow M_t \cup \{m_a\}$, $L_t \leftarrow L_t \cup \{l_{a,b}\}$, $LD_t \leftarrow LD_t + f_a$
    end if
  end if
end while

The MR-GW association procedure shown in Algorithm 2 allocates MRs to GW-rooted trees iteratively. Let AL be the set of active links, CL be the set of potential active links, CM be the set of MRs out of GW-rooted trees. We use notation t to be the unique identity of trees, and notation $tree(i)$ to represent the tree that MR m_i belongs to. For tree t, notations M_t, L_t, LD_t and rt_t denote its node set, link set, load and root respectively. In each iteration, among all the links adjacent to GW-rooted trees, i.e., only one end (MR) of these links is in a GW-rooted tree, the link with the minimum weight is chosen. The other end (MR) of the selected link is added to the GW-rooted tree, and the link becomes an active link. This step is done until $n-|M_g|$ active links are selected. In LIAA algorithm, the weight calculation of link $l_{i,j}$ is defined as follows:

$$W_l(l_{i,j},t)=\frac{h_{rt_t,i}+1}{R}\cdot\frac{|I(l_{i,j})\cap AL|+1}{|AL|+1}\cdot\frac{(Ld_t-\min\limits_{p=1,2,\ldots,k}Ld_p)+1}{(\max\limits_{p=1,2,\ldots,k}Ld_p-\min\limits_{p=1,2,\ldots,k}Ld_p)+1} \quad (18)$$

where rt_t is the root of tree t and Ld_t is its current load. For link $l_{i,j}$, end m_i has been in the tree rooted by $root_t$, and end m_j is out of any gateway-rooted tree. When we choose active links, three factors are considered, i.e., MR-GW path length, link interference and gateway load. So, the weight calculation is comprised of three corresponding parts. The first part $\frac{h_{rt_t,i}+1}{R}$ indicates that shorter MR-GW path will lower the weight value. The second part $\frac{|I(l_{i,j})\cap AL|+1}{|AL|+1}$ implies that the link causing less interference will be chosen preferentially. The third part $\frac{(Ld_t-\min\limits_{p=1,2,\ldots,k}Ld_p)+1}{(\max\limits_{p=1,2,\ldots,k}Ld_p-\min\limits_{p=1,2,\ldots,k}Ld_p)+1}$ prevents links adjacent to the overloaded trees being selected.

4.3 Complexity Analysis

In our proposed **LIGDP Heuristic**, we consider a graph $G(M,L)$ with n nodes. Firstly, we need to obtain $N_R(m_i)$ for each $m_i\in M$. It can be implemented by finding the shortest path between each node pair. We use Floyd's algorithm [Cormen et al. 2001], and its time complexity is $O(n^3)$. We also need to obtain $I(l_{i,j})$ for each $l_{i,j}\in L$, and its time complexity is $O(|L|^2)$. Then, the gateway location algorithm MLA has at most n iterations. In each iteration, at most n nodes have its weight calculated, and the time complexity of calculating a node weight is $O(n)$. The time complexity of gateway location is $O(n^3)$. Finally, the MR-GW association algorithm LIAA has $n-|M_G|$ iterations. In each iteration, at most $|L|$ links have its weight calculated, and the time complexity of calculation a node weight is $O(|M_G|)$. The time complexity of MR-GW association is $O(n\cdot|L|\cdot|M_G|)$. Because $|L|\leq n^2$ and $|M_G|<n$, **LIGDP Heuristic** is with the polynomial time complexity of $O(n^4)$.

5 Performance Evaluation

We implement the proposed approach **LIGDP Heuristic** (MLA and LIAA algorithms) using Matlab, and evaluate its performance through simulation experiment. Our experiment consists of two parts: in the first part, we discuss the impact of R-hop, GW throughput and MR throughput constraints; in the second part, we compare the proposed approach with other existing approaches.

5.1 Performance Metrics

We evaluate the four objectives as discussed in Section 3, i.e., minimizing deployment cost, minimizing MR-GW path length, balancing gateway load and minimizing link interference. In the following, four corresponding metrics are defined to measure these objectives.

- **Number of gateways**: We use number of gateways to reflect the deployment cost. As the larger number of deployed gateways for a given network is, the higher will be deployment cost.
- **Average length of MR-GW paths**: We use average length of MR-GW paths to reflect the performance of MR-GW paths. Short path indicates high bandwidth and low latency.
- **Standard deviation of gateway load**: We introduce standard deviation of gateway load, i.e.,

$$\sqrt{\frac{\sum_{m_i \in M_g} (load(m_i) - avgLoad)^2}{(|M_g| - 1)}} \tag{19}$$

  , as the metric for load balance among gateways. $avgLoad$ is the average of gateway load. A low standard deviation indicates that the gateway load tend to be very close to the average, whereas high standard deviation indicates the imbalance of gateway load.
- **Average interference of active links**: We use average interference of active links, i.e.,

$$\frac{1}{|AL|} \cdot \sum_{l_{i,j} \in AL} |I(l_{i,j}) \cap AL| \tag{20}$$

  , to measure link interference in the network. AL is the set of active links, and $I(l_{i,j})$ is the set of links interfered by link $l_{i,j}$.

5.2 Parameter Configurations

Parameter configurations of the simulation experiment are shown in Table 2. We simulate WMN backbones with different scales. MRs are randomly and independently distributed in a square network domain. Each MR has 1 unit of local traffic demand. The minimum distance separating any pair of nodes is set to 150 meters because placing MRs very close to each other is not common in practice [Aoun et al. 2006]. When network size increases, additional MRs will join the backbone to maintain the node density.

5.3 Impact of *R*-hop, GW Throughput and MR Throughput Constraints

We consider a network of 600 MRs distributed over a 4900 m $\times$ 4900 m square domain. In order to see the impact of R-hop, GW throughput and MR throughput constraints, we evaluate the performance of **LIGDP Heuristic** with different constraints. The R-hop, GW throughput and MR throughput constraints (R,

C_g, C_m) are set to be (2, 16, 4), (2, 20, 5), (3, 24, 6), (3, 28, 7), (4, 32, 8) and (4, 36, 9) respectively, and four performance metrics are used, i.e., number of gateways, average length of MR-GW paths, standard deviation of gateway load, and average interference of active links. We run the algorithm with different constraints 20 times and use the average results as the final results. Every time, 600 MRs are randomly redistributed in the network domain.

Table 2: Common simulation parameters

Parameter	Value
Network size	100 MRs over a 2000 m × 2000 m area
	200 MRs over a 2800 m × 2800 m area
	300 MRs over a 3500 m × 3500 m area
	400 MRs over a 4000 m × 4000 m area
	500 MRs over a 4500 m × 4500 m area
	600 MRs over a 4900 m × 4900 m area
Transmission-range radius r_t	250 m
Carrier-Sensing-range radius r_s	500 m
Local traffic demand of each MR f_i	1
Hop-constraint R	2, 3 or 4
Upper bound of MR throughput C_m	4, 5, 6, 7, 8 or 9
Upper bound of GW throughput C_g	16, 20, 24, 28, 32 or 36

Table 3: Varying R-hop, GW throughput and MR throughput constraints

Constraints (R, C_g, C_m)	**Number of GWs**	**Length of MR-GW paths**	**Deviation of GW load**	**Interference of active links**
(2, 16, 4)	88.5000	1.0651	1.6699	20.6753
(2, 20, 5)	86.5500	1.0757	1.7650	20.7508
(3, 24, 6)	58.6000	1.4495	2.2790	21.1304
(3, 28, 7)	53.4500	1.5127	2.4530	21.2218
(4, 32, 8)	47.2500	1.7142	3.1733	21.6674
(4, 36, 9)	42.1000	1.8066	3.1834	21.8473

Table 3 reports the simulation results in term of the number of gateways, the average length of MR-GW paths, the standard deviation of gateway load and the average interference of active links. Six rows refers to the results under different constraints respectively. The main observation made from the results in Table 3 is that when we relaxes these constraints gradually, the number of deployed gateways will decrease, but the average length of MR-GW paths, the standard deviation of gateway load and the average interference of active links will increase at the same time. It is because the coverage range of gateways is extended, and the number of MRs severed by each gateway increases. Hence, we find that there exist certain contradictions between minimizing deployment cost and improving network performance. They cannot be achieved simultaneously. Meantime, we also find that constraints R, C_g and C_m are three configurable parameters which can provide a trade-off between minimizing deployment cost and improving network performance.

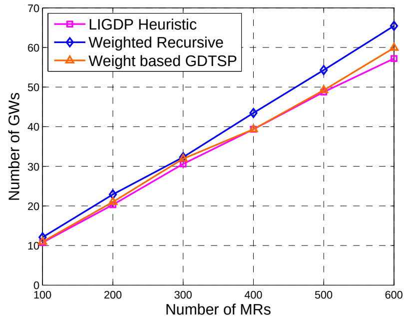

Figure 3: Comparison on Number of Gateways

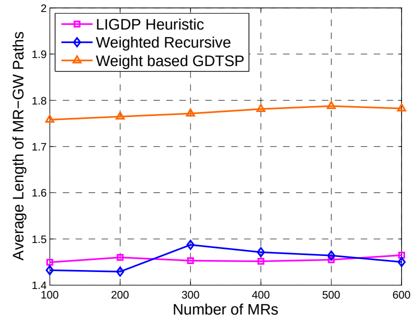

Figure 4: Comparison on Average Length of MR-GW paths

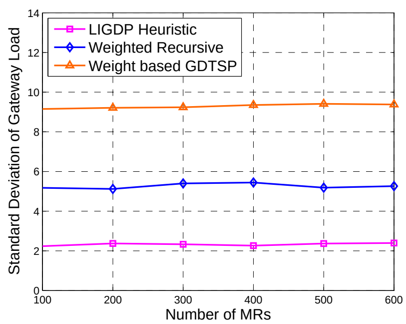

Figure 5: Comparison on Standard Deviation of Gateway Load

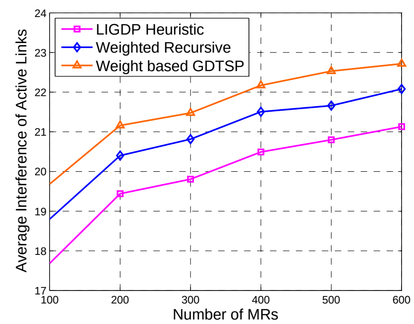

Figure 6: Comparison on Average Interference of Active Links

5.4 Comparison with Other Existing Approaches

In this part of the simulation experiment, our approach **LIGDP Heuristic** is compared with two previous works proposed by Aoun et al. [Aoun et al. 2006] and He et al. [He et al. 2008]. The approach proposed by Aoun et al. [Aoun et al. 2006] is a weighted recursive algorithm, and the approach proposed by He et al. [He et al. 2008] is a weight based GDTSP algorithm. We fix the following parameters, that is, $R = 3$, $C_g = 24$ and $C_m = 6$, and vary network size from 100 MRs to 600 MRs, i.e., 100, 200, 300, 400, 500 and 600 MRs respectively, as shown in Table 2. For every network size, we runs the three algorithms 20 times and uses the average results as the final results. Every time, MRs are randomly redistributed in the network domain.

When the network size increasing, we expect to observe that **LIGDP Heuristic** leads to more number of gateways. Figure 3 demonstrates this increment and compares it with weighted recursive algorithm and weight based GDTSP algorithm. It can be seen from the comparison that our approach in terms of number of gateways performs better than weighted recursive algorithm and almost the same as weight based GDTSP algorithm.

Figure 4 shows the comparison on average length of MR-GW paths. Our results is less than that of weight based GDTSP algorithm, because we consider minimizing MR-GW path length in the process of constructing GW-rooted trees. The results of weighted recursive algorithm is closed to ours, because its solution leads to more gateways that decrease the length of MR-GW paths. This indicates that our approach have less number of gateways while maintaining short length of MR-GW paths.

As shown in Figure 5, our approach achieves the lowest standard deviation of gateway load. This result is expected; our approach prevents MRs associating congested gateways, while the other two approaches do not take gateway congestion into account. Hence, our approach performs better on gateway load balancing than the other two approaches.

Figure 6 shows that the average interference of active links increases when the network size increases. That is because when the number of MRs increases, more links will be utilized to transmit packets. Figure 6 also shows that our approach results in less interference than the other two approaches and achieves better performance on interference minimization.

6 Conclusions and Future Work

In this paper, we propose a gateway deployment approach **LIGDP Heuristic** that minimizes deployment cost, while satisfying performance requirements. **LIGDP Heuristic** differs from all existing literature because it takes load balancing and interference minimization into account. It combines two algorithms, i.e, MSC-based location algorithm (MLA) and load-aware and interference-aware association algorithm (LIAA), to determine gateway positions and construct GW-rooted trees. We evaluate the performance of the proposed approach in terms of the number of gateways, the average length of MR-GW paths, the standard deviation of gateway load and the average interference of active links. Through performance evaluation, we have concluded that there exist contradictions between network performance and deployment cost, and a trade-off can be achieved by adjusting R-hop, GW throughput and MR throughput constraints. Subsequently, compared with other existing approaches, we demonstrate that **LIGDP Heuristic** performs better on MR-GW path, load balancing and interference minimization without deploying more gateways.

In our ongoing work, we are continuing to investigate the gateway deployment problem. We plan to extend our study into multi-channel multi-radio WMNs and address this problem through jointing gateway location, routing, channel assignment and bandwidth allocation. We also plan to achieve gateway-failure tolerance in the phase of gateway deployment to improve network reliability.

Acknowledgment

This work is supported by National Key Basic Research Program of China under Grants No. 2010CB328104, National Natural Science Foundation of China under Grants No. 60903161, No. 60903162, No. 61003257, No. 61070161 and No. 61070158, China National Key Technology R&D Program under Grants No. 2010BAI88B03, China Specialized Research Fund for the Doctoral Program of Higher Education under Grants No. 200802860031, Jiangsu Provincial Natural Science Foundation of China under Grants No. BK2008030, China National S&T Major Project under Grants No. 2009ZX03004-004-04, Jiangsu Provincial Scientific and Technological Achievements Transfer Fund under Grant No. BA2008010, State Key Laboratory of Information Security, Jiangsu Provincial Key Laboratory of Network and Information Security under Grants No. BM2003201, and Key Laboratory of Computer Network and Information Integration of Ministry of Education of China under Grants No. 93K-9.

References

[Akyildiz et al. 2005] Akyildiz, I. F., Wang, X., Wang, W.: Wireless mesh networks: a survey, Computer Networks, 47 (4), 2005, 445-487.

[Alicherry et al. 2005] Alicherry, M., Bhatia, R., Li, L.: Joint channel assignment and routing for throughput optimization in multi-radio wireless mesh networks, Proceedings of the 11th Annual International Conference on Mobile Computing and Networking (MobiCom'05), ACM, Cologne, 2005, 58-72.

[Aoun et al. 2006] Aoun, B., Boutaba, R., Iraqi, Y., Kenward, G.: Gateway Placement Optimization in Wireless Mesh Networks With QoS Constraints, IEEE Journal on Selected Areas in Communications, 24 (11), 2006, 2127-2136.

[Bejerano 2004] Bejerano, Y.: Efficient integration of multihop wireless and wired networks with QoS constraints, IEEE/ACM Transactions on Networking, 12 (6), 2004, 1064-1078.

[Bejerano et al. 2007] Bejerano, Y., Han, S. J., Kumar A.: Efficient load-balancing routing for wireless mesh networks, Computer Networks, 51 (10), 2007, 2450-2466.

[Bruno et al. 2005] Bruno, R., Conti, M., Gregori, E.: Mesh networks: commodity multihop ad hoc networks, IEEE Communication Magazine, 43 (3) 2005, 123-131.

[Chandra et al. 2004] Chandra, R., Qiu, L., Jain, K., Mahdian, M.: Optimizing the placement of Internet TAPs in wireless neighborhood networks, Proceedings of the 12th IEEE International Conference on Network Protocols (ICNP'04), IEEE, Berlin, 2004, 271-282.

[Chvatal 1979] Chvatal, V.: A greedy heuristic for the set-covering problem, Mathematicof Operations Research, 4 (3), 1979, 233-235.

[Cormen et al. 2001] Cormen, T. H., Leiserson, C. E., Rivest, R. L., Stein, C.: Introduction to Algorithms, Second Edition, MIT Press and McGraw-Hill, Cambridge, Mass, 2001.

[Draves et al. 2004] Draves R., Padhye, J., Zill, B.: Routing in multi-radio, multi-hop wireless mesh networks, Proceedings of the 10th Annual International Conference on Mobile Computing and Networking (MobiCom'04), ACM, Philadelphia, 2004, 114-128.

[Durocher et al. 2008] Durocher, S., Jampani, K. R., Lubiw, A., Narayanan, L.: Modelling Gateway Placement in Wireless Networks: Geometric k-centres of Unit Disc Graphs, Proceedings of the 5th international workshop on Foundations of mobile computing, ACM, Toronto, 2008, 79-86.

[Gupta et al. 2000] Gupta, P., Kumar, P.R.: The capacity of wireless networks, IEEE Transactions on Information Theory, 46 (2), 2000, 388-404.

[He et al. 2008] He, B., Xie, B., Agrawal, D. P.: Optimizing deployment of Internet gateway in Wireless Mesh Networks, Computer Communications, 31 (7), 2008, 1259-1275.

[Hsiao et al. 2001] Hsiao, P., Hwang, A., Kung, H., Vlah, D.: Load-balancing routing for wireless access networks, Proceedings of the 20th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'01), IEEE, Anchorage, 2001, 986-995.

[Hu et al. 2007] Hu, X., Lee, M. J.: An efficient multipath structure for concurrent data transport in wireless mesh networks, Computer Communications, 30 (17), 2007, 3358-3367.

[Huang et al. 2009] Huang, C. J., Chen, I. F., Hu, K. W., Shen, H. Y.: A Load Balancing and Congestion-Avoidance Routing Mechanism for Real-Time Traffic over Vehicular Networks, Journal of Universal Computer Science, 15 (13), 2009, 2506-2527.

[Hundewale et al. 2007] Hundewale, N., Jung, S., Zelikovsky, A.: Energy Efficient Node Caching and Load Balancing Enhancement of Reactive Ad Hoc Routing Protocols, Journal of Universal Computer Science, 13 (1), 2007, 110-132.

[Li et al. 2001] Li, J., Blake, C., Couto, D. D., Lee, H. I., Morris, R.: Capacity of Ad Hoc Wireless Networks, Proceeding of the 7th Annual International Conference on Mobile Computing and Networking (MobiCom'01), ACM, Rome, 2001, 61-69.

[Li et al. 2008] Li, F., Wang, Y., Li, X. Y., Nusairat, A., Wu, Y.: Gateway Placement for Throughput Optimization in Wireless Mesh Networks, Mobile Networks and Applications, 13 (1-2), 2008, 198-211.

[Lukovszki et al. 2006] Lukovszki, T., Schindelhauer, C., Volbert, K.: Resource Efficient Maintenance of Wireless Network Topologies, Journal of Universal Computer Science, 12 (9), 2006, 1292-1311.

[Marco, 2008] Marco, G. D.: GAMesh: Automatic Placement of Wireless Mesh Nodes Via Genetic Algorithms, Lecture Notes in Computer Science, 5186, 2008, 2-11.

[Raniwala et al. 2005] Raniwala, A., Tzi-cker, C.: Architecture and algorithms for an IEEE 802.11-based multi-channel wireless mesh network, Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'05), IEEE, Miami, 2005, 2223-2234.

[Shi et al. 2009] Shi, Y., Hou, Y. T., Liu, J., Kompella, S.: How to correctly use the protocol interference model for multi-hop wireless networks, Proceedings of the International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc'09), ACM, New Orleans, 2009, 239-248

[Zeng et al. 2009] Zeng, F., Chen, Z.: Cost-sensitive and load-balancing gateway placement in wireless mesh networks with QoS constraints, Journal of Computer Science and Technology, 24 (4), 2009, 775-785.

[Zhang et al. 2006] Zhang, Y., Luo, J., Hu, H.: Wireless Mesh Networking: Architectures, Protocols and Standards, Aerbach Publications, Boca Raton, 2006.

[Zhou et al. 2007] Zhou, P., Manoj, B. S., Rao, R.: A Gateway Placement Algorithm in Wireless Mesh Networks, Proceedings of the 3rd international conference on Wireless internet, ACM, Austin, 2007, 1-9.