Point-to-Point Traffic in Wireless Mesh Networks

by

Jiajia Han

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2010

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Although most work on Wireless Mesh Networks (WMNs) has focused on traffic flowing in and out of the network via gateways, traffic within the WMN may also be significant in many environments. This point-to-point (P2P) traffic between the nodes within WMN can be handled in different ways, particularly in WMNs containing multiple gateways. The approach used affects the performance of both the P2P flows and other traffic in the network.

This work studies the impact of handling P2P traffic in the presence of gateways and gateway traffic. Through mathematical analysis of the resulting traffic patterns, along with extensive simulations, the need to route P2P traffic appropriately is demonstrated. While direct routing yields considerable performance improvements in small networks, it can actually decrease capacity in larger networks.

Consequently, we extend the Hybrid Wireless Mesh Protocol (HWMP) proposed in IEEE 802.11s by adding two new message types to obtain more information useful for choosing the best route. Through simulations on different networks, HWMP shows better average delivery ratio and end-to-end delay than the original gateway-based and the P2P routing mechanisms in the simulation settings.

Acknowledgements

I would like to thank Brent Ishibashi for his help in my initial study of the simulation environment and countless discussions in this regard. I wish to give my special thanks to Prof. Raouf Boutaba for his kindness and support over the past three years, first as an undergraduate research assistant and then as a Master's student.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Wireless mesh networks (WMNs) have become a popular approach to providing wireless network access. By using wireless interconnection of its infrastructure, the WMN allows WLAN-like services to be deployed over wider area. By only connecting a small number of gateways (GWs) directly to the Internet, networks can be installed quickly and inexpensively.

A WMN is an edge network to which users connect. Most works have assumed that traffic will travel between these users and hosts elsewhere on the Internet, passing through a GW to transit in and out of the WMN. However, a second type of traffic also exists — two users connected through the same WMN may wish to communicate. A significant volume of this point-to-point (P2P) traffic may exist in many scenarios, and therefore must be handled effectively.

The nature of the WMN and the presence of gateways create a variety of ways that this P2P traffic can be handled. A gateway represents a potential relay point — traffic sent in this way mimics the way all other traffic is handled. In a multiple gateway network, this could require the use of the wired network to transfer traffic between GWs. However, the mesh topology may also allow for a more direct path to be taken. These three approaches are depicted in Fig. 1.1.

The different approaches have different effects on the forwarding load of the network, and therefore the performance. With gateway traffic creating a known resource bottleneck around the gateway[27], it is important to understand these effects and the relationships that exist between gateway and P2P traffic. Therefore, this work investigates these issues through analysis and simulation, studying a wide variety of scenarios, including different topologies, traffic loads, and routing approaches.

This work demonstrates the need for having an appropriate strategy for handling P2P traffic. There is a clear performance advantage to be gained by directly supporting P2P

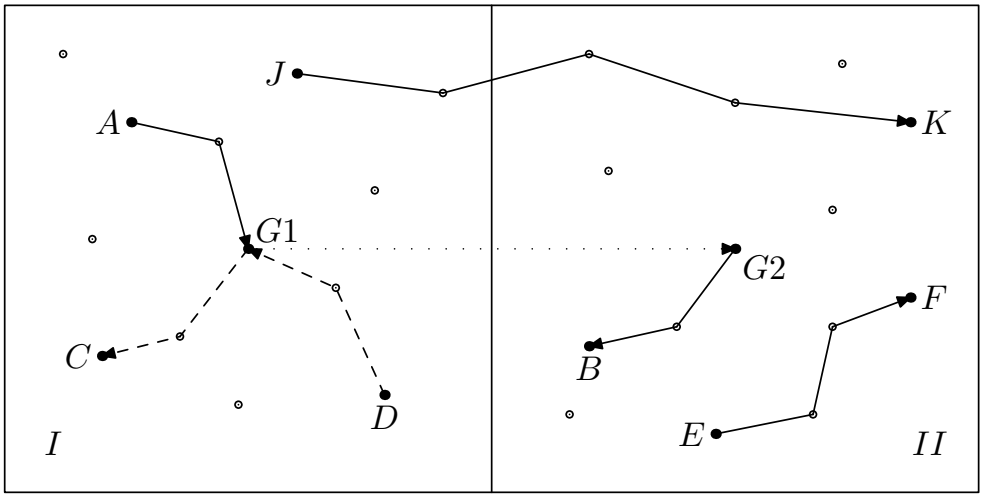


Figure 1.1: Handling Peer-to-Peer Traffic in a WMN: I. Via GW (D-C) II. Via different GW (A-B) III. Direct (E-F) IV. Direct across different clusters (J-K)

traffic, however all traffic cannot be treated identically. Doing so in a large network actually decreases the achievable throughput. Therefore we propose a P2P-Aware Wireless Mesh Protocol (PAWMP), and through simulations, demonstrate its effectiveness.

1.1 Wireless Mesh Networks Definitions

The building blocks of WMN consist of Mesh Points (MP), Mesh Portals (MPP), Mesh Access Points (MAP), and wireless clients. In practice, a WMN product usually contains more than one element in it. For example, a WMN router behaves both as a Mesh Point and a Mesh Access Point.

The Mesh Access Points behave the same as traditional IEEE 802.11 access points to the wireless clients. The users use their Wi-Fi-enabled device, e.g., laptop, PDA, or cellphones, to connect to the access points and browse websites, transfer files, etc. using the connectivity provided by the WMN.

The Mesh Points are the wireless routers in the WMN. When an end user want to browse a website, the MPs will route the packets from the MAP to MPP, often through a multi-hop wireless path. The performance of a WMN largely depends on the efficiency of MPs routing. Many of the work in WMN tried to improved the routing of MPs.

The Mesh Portals are the gateways of the WMN. In a WMN, only MPPs are connected to the Internet, the other nodes have to connect to the MPPs to get Internet access. As

such, they form a natural bottleneck for all the traffic in and out of the WMN. To improve performance and reliability, there may be several MPPs in a WMN.

1.2 Problem Description

Previous works on WMNs have addressed P2P-style communication[8, 29]. In fact, this is expected based on WMN's roots in mobile ad hoc networks (MANETs). However, the networks in these works typically do not contain a gateway. This creates an important difference with our definition of a WMN.

We consider a WMN to be a network infrastructure providing access service to user devices. The WMN is made up of a collection of nodes or access points, out of which a limited number are provided with a wired network connection, creating gateways. Using wireless communication, the Access Points (APs) and GWs interconnect to form a mesh, allowing traffic to be forwarded from node to node through the WMN. Traffic destined for a host on the Internet (outside the WMN) can be sent via a gateway.

For WMNs with gateways, past studies have shown that the capacity of the network is limited by the wireless region surrounding the gateway[12]. With all traffic in the network passing through this area, the shared nature of the wireless medium and the multihop nature of traffic forwarding, combine to create heavy demand on these links. [11] showed that the maximally-loaded clique could be found, in order to determine an upper bound for the throughput that the network can achieve. No additional traffic can pass through this bottleneck.

All traffic in these works flow through the gateway. For many applications, this is a reasonable assumption — web-browsing, for example, most probably occurs between a WMN user and a web server located on the Internet. However, this is not necessarily true for other applications. The WMN is inherently (due to wireless propagation characteristics) geographical in nature. Person to person communication may exhibit similar locality. In addition, the performance advantages of communicating with nearby nodes can be beneficial to certain applications, such as peer-to-peer networking[23].

To further illustrate the potential significance of P2P traffic in a WMN, the following four use cases are presented. The relative importance of P2P traffic varies across the scenarios, although some traffic can be expected in all cases.

- Use Case 1: WMN Service throughout an Airport — Consider the use of a WMN to provide service coverage in an airport. One would expect that Internet traffic would comprise the vast majority of traffic volume. However, there are a number of situations where some P2P traffic could be generated: locating a fellow passenger,

meeting up with transportation, or contacting various airport services. Although the volume of this traffic would likely be much smaller than the Internet traffic, it could still be significant.

- Use Case 2: WMN Service on a Campus — For a WMN deployed across a company or university campus, there is potentially a larger opportunity for P2P traffic. Again, Internet traffic will be important, but locality suggests that people within the campus may have more reason to interact with each other, while connected to the mesh. Campus services may also have a presence on the mesh network. Therefore, this scenario creates an increased opportunity for P2P traffic.

- Use Case 3: WMN Service to a Remote Village — Consider a WMN deployed to deliver wireless services to a remote village. In this scenario, Internet bandwidth may be severely limited. However, if the WMN has the ability to provide additional communication services in a location where there are little or no existing services, P2P traffic may in fact comprise the majority of traffic within the WMN. In this scenario, the P2P services may be developed because there is P2P bandwidth available, bandwidth that can not be otherwise used due to the limitations of the gateway link.

- Use Case 4: Emergency Services — Emergency services have often been cited as a potential use for WMNs. Communication between emergency personnel during a crisis is extremely important. This includes a high level of interaction and coordination between personnel on site. Although this type of use may be temporary in nature, it may also create extremely heavy demand. Under these conditions, it is important that traffic be handled as efficiently as possible, in order to ensure that capacity is maximized and QoS is ensured.

1.2.1 Handling P2P Traffic

Handling traffic to and from a gateway is relatively straightforward. In many WMN works, each AP is assumed to associate with a particular GW. In this case, each gateway handles traffic for a cluster of nodes[20]. The capacity of the network is largely determined by the locations of gateways, although routing may also have some effect[10].

Consider if a source AP AP_S has traffic to send to an Internet host H. The AP sends its data to its gateway $GW(AP_S)$. This path may include any number of intermediate APs who relay the data. The gateway in turn forwards to H via its wired connection. The reverse would occur if H wanted to send to a destination AP_D.

P2P traffic can be handled in a similar manner. Consider traffic between AP_S and AP_D. AP_S can send to $GW(AP_S)$, who can in turn send to AP_D. This can occur if the WMN consists of a single cluster, or if both APs are in the same cluster ($GW(AP_S) = GW(AP_D)$).

If the APs attach to different clusters, then the $GW(AP_S)$ can send to $GW(AP_D)$ via the wired network. In many WMN architectures, a wired distribution network may ensure that this can be done with minimal cost.

This approach is simple, and makes use of the routing facilities already existing for GW traffic. In fact, within the WMN, the traffic pattern is equivalent to AP_S sending to H and H sending to AP_D. The only differences are on the GW link and through the Internet. This approach occurs in WMN solutions that tunnel all traffic to the GW.

The alternative is that traffic can be sent directly between AP_S and AP_D. This is the same as in a MANET, or a WMN that does not include GWs. The routing protocol determines the path that is taken between source and destination, which may direct, or may include intermediate APs. For most source-destination pairs, the direct route is shorter than the GW path. In fact, one can even consider the case where $AP_S = AP_D$, where both users are attached to the same AP — in this case, the direct route requires no forwarding at all.

In a single gateway network the direct route is at worst the same distance (or in fact the same route) as the path via the GW. In a network with multiple gateways, different cases must be considered. If the source and destination APs use the same GW, then the situation is the same as in the single GW scenario. However, if they use different GWs, then two possibilities exist: the direct path can be used, or the gateways can be used, along with the wired network.

It is clear that as the network grows, the average length of direct paths will also grow. For distant pairs, the path through the GWs may represent a considerable savings in wireless resource use. However, this path inherently adds traffic to the already congested bottleneck regions. Direct routes could potentially avoid these areas, allowing their traffic to be carried by the network with additional cost.

1.3 Contribution

Our contributions to the networking research in wireless mesh networks are stated as follows.

- *An analytical study of the WMN.* We use an analytical model to compute the bottleneck collision domains in grid-topology networks, which ultimately limits the maximum capacity of WMNs. The results from the analytical model give us insights into the performance of different routing mechanisms in WMN with point-to-point traffic.
- *Comparison of different routing mechanisms in WMN with P2P traffic.* Through extensive simulations, we compare the performance of different routing mechanisms

under different topologies, varying number of gateways, and different traffic loads. We show the disadvantages of the current routing mechanisms in WMN with P2P traffic.

- *A new routing protocol.* We devise a P2P Aware Wireless Mesh Protocol(PAWMP) which chooses better paths for point-to-point traffic and we show its performance through simulations.

1.4 Thesis Organization

The rest of the thesis is organized as follows. We survey related works in Chapter 2. Then an analytical model is presented in Chapter 3. Simulation studies are presented in Chapter 4. After that, we describe the PAWMP protocol in Chapter 5. Finally, conclusions and future work are presented in Chapter 6.

Chapter 2

Related Work

2.1 Capacity Analysis of Wireless Networks

In [8], the authors analyzed the capacity of wireless networks on two types of networks, *Arbitrary Networks* and *Random Networks*. They computed the capacity using two models, the *Protocol Model*, where the destination can receive the message if all the nodes nearer than the source are not transmitting, and the *Physical Model*, where a minimum signal-to-interference ratio (SIR) is necessary for successful receptions. The authors showed that when n randomly located nodes, each capable of transmitting at W bits per second and using a fixed range, the throughput $\lambda(n)$ obtainable by each node for a randomly chosen destination is $\theta(\frac{W}{\sqrt{n \log n}})$ bits per second under the Protocol Model. Similar results also holds under the physical model. The authors showed that it is the need for each node in the entire domain to share whatever portion of the channel it is utilizing with nodes in its local neighborhood that is the reason for the constraint in capacity. The authors suggested that networks connecting smaller numbers of users, or featuring connections mostly with nearby neighbors, are more likely to find acceptance, since the throughput provided to each user decreases to zero as the number of users is increased.

Based on the results of [8], several other works analyzed some specific wireless network scenarios. For example, the work in [17, 1] analyzed how capacity can be improved by adding K base stations into a network. [17] showed that when the number of base stations m is less that $\sqrt{n}$, the per-node capacity is $\Theta(\sqrt{1/(n \lg \frac{n}{m^2})}\ )$. In turn, when the number of base stations is larger than $\sqrt{n}$, the per-node capacity is linear to m, i.e., $\Theta(m/n)$. [15] showed that the per-node capacity of a hybrid network can achieve $\Theta(1/\lg n)$. [1] further demonstrated that with power control on each node, the per-node capacity can achieve $\Theta(1)$ for a fraction f of nodes.

For wireless mesh networks with one gateway, [12] and [28] both showed that the per-

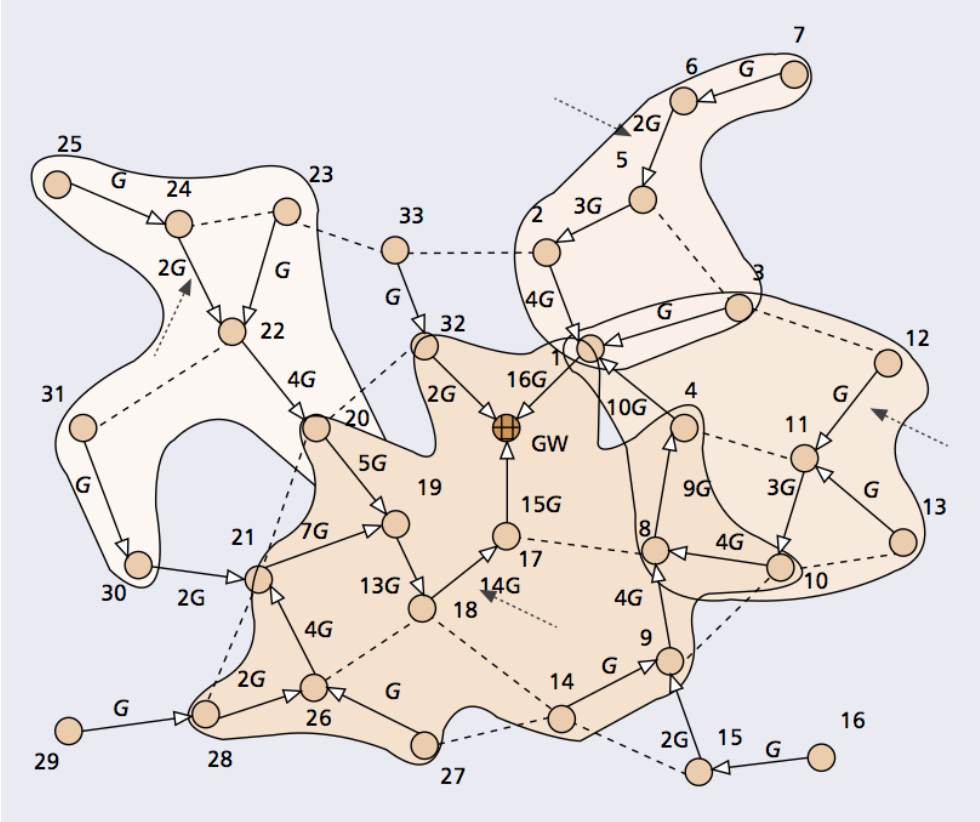


Figure 2.1: A WMN with load on the links

node capacity is $O(1/n)$, [28] using the model from [8] and [12] introduced the concept of bottleneck collision domain, which throttles the throughput of the entire network. For example, in Fig. 2.1, each node has G unit of traffic destined to the gateway in the middle labeled with GW. The number xG besides each link is the amount of traffic that this link has to carry. We can see that the links around the gateway should carry a total of $33G$. Thus if each node sends the same amount of traffic and the capacity of each link is C, then the effective capacity for each node is only $\frac{1}{33}C$.

2.2 Routing in Wireless Mesh Networks

Over the past few years, a number of routing algorithms have been proposed under the WMN settings. In this section, we introduce some prominent routing algorithms in different categories.

There are two main differences between WMN and Mobile Ad-Hoc Network (MANET):

mobility and infrastructure. That is, a node in a WMN has minimum mobility, if any, and there are gateways in WMN providing Internet access.

In MANET, there are several proactive[22] and reactive[24, 18] routing algorithms. In theory, WMN is a special case of MANET, since the WMN is also a multi-hop wireless network and the gateways can be treated as nodes with large amount of traffic. So the MANET routing algorithms can be applied directly to WMN. And there are indeed several WMN routing algorithms proposed in this manner, including SrcRR[6], Link Quality Source Routing (LQSR) and Multi-Radio LQSR (MR-LQSR)[7].

In [17], a k-nearest cell routing and a probabilistic routing strategies are analysed. In the k-nearest cell routing strategy, the route to destination nodes within the k-th neighboring cell will use the ad hoc mode and routes to other nodes will use the access point mode. [15] and [1] didn't consider a specific routing algorithm. In turn, they try all the possible routing paths and find a feasible scheduling for all transmissions. This, however, assumes that the nodes only use the access point when they are one hop away from it.

In [16], the authors compared the tree based routing(TBR), namely, the HWMP routing protocol in 802.11s, with their proposed root driven routing(RDR). The idea of RDR is to let the root(usually the gateway) know the topology of the whole network and the corresponding link metrics. Then during the routing phase, the root will compute the peer-to-peer paths for the nodes within the network and tell the nodes to route using these paths. Their results showed that the RDR protocol outperforms the TBR protocol with much lower average end-to-end delay and much higher packet delivery ratio for the intra-mesh traffic. However, they only consider the one gateway case.

In [9], the author proposed the Zone Routing Protocol(ZRP), in which a clustered set of nodes choose a cluster head for communication among different clusters, while the traffic within the same cluster uses the peer to peer manner. The idea of this routing protocol is similar to that of the wireless mesh networks, that is, there is a gateway to the outside world. In ZRP, through the use of the zone radius parameter, the routing scheme exhibits the adjustable hybrid behavior of proactive and reactive routing schemes. In our work, the multiple gateways case show similar properties to this routing scheme.

2.3 Address Allocation in Wireless Mesh Networks

In this section we describe the address allocation algorithms proposed in the literature.

In [30], the authors proposed the Dynamic WMN Configuration Protocol (DWCP). DWCP follows the hierarchy of the WMN architecture. At the top level, DWCP subdivides the class of mesh gateways into three sub-classes: one master gateway (MGW),

pre-configured gateways (PGWs) and self-configured gateways (SGWs). The auto configuration process starts at the special MGW. The MGW assigns all newly arriving mesh gateways smaller disjoint portions of its address pool. PGWs are pre-configured with external address of the MGW and can thus instantly request a gateway address pool from the MGW. SGWs should wait for the broadcast of the MGW to be able to request address pools from the MGW.

The mesh routers are split up into those with an address pool and without an address pool. A mesh router with an address pool runs a standard DHCP server and assigns IP addresses to other requesting mesh routers and mesh clients. Mesh routers without an address pool start a DHCP *relay* and forward DHCP requests to mesh routers with an address pool. The division of the mesh routers is made to keep fragmentation and dissipation of the IP address space as low as possible.

In [3], the authors analyzed the deployment of WMNs using current residential routers. In this context, they investigated the ability of Passive Auto-Configuration for Mobile Ad-hoc Networks (PACMAN) – a mechanism developed for IP auto-configuration in ad-hoc networks – to satisfy the requirements that an IP address auto-configuration protocol for community WMNs should meet.

PACMAN [26] is a fully distributed address auto-configuration mechanism for ad-hoc networks that aims to guarantee unique IP addresses in the network even in the presence of network merging. The basic idea is that a router joining the mesh network assigns an address to itself by randomly picking one from the set of yet unassigned addresses according to the router's local knowledge, and relying on the passive duplicate address detection (PDAD) concept to detect conflicts originating from this optimistic address assignment or from network merging. The mesh router may learn about already assigned addresses by monitoring the routing protocol traffic or by requesting a list of addresses that are known to be assigned in the network from a neighbour router.

PACMAN is used for the address assignment to wireless routers in WMN, and then the wireless routers will use NAT to allocate addresses for mesh clients. The latter step is not part of PACMAN.

In [4], the authors built the Roofnet. Roofnet uses 802.11b to connect the mesh routers, and 802.11g to connect the mesh clients. Each mesh router has an IP address of the form 10.x.x.x, where the low three bytes come from the lowest three bytes of its wireless interface's MAC address.

By default, every mesh router uses the IP address 192.168.0.1 on its port to the mesh clients, and runs a DHCP server that answers queries heard on that interface. Each mesh router runs a NAT that makes packets from hosts connected by Ethernet appear to come from the mesh router's 10.x.x.x address.

Any mesh router can be configured as a gateway to the Internet if it has, for example, a DSL or cable modem attached to the router's Ethernet port. The gateway router periodically flood advertisements to all Roofnet routers saying that it is a gateway. Each non-gateway node chooses a gateway through which to route its Internet traffic. A router switches gateways only if its current gateway is unreachable. This avoids unnecessarily breaking of existing TCP connections by switching gateway NATs, but may cause the node to continue to use a gateway to which it has a low-quality route when better gateways are available.

Chapter 3

Analytical Model

In order to better understand the problem, an analysis of two simple network topologies is performed. The throughput capacity of the network can be estimated using the concept of a bottleneck collision domain (BCD)[12]. Each traffic flow in the network is defined in terms of a throughput variable T. Based on this traffic requirement, the load on each link is determined, again in terms of T. However, the shared nature of the wireless medium means the traffic over multiple links must share the same wireless capacity. A collision domain links must share the same wireless capacity. A collision domain exists for each link l, consisting of all of the other links that interfere with l. The total traffic load on these links (L_{BCD} provides an estimate of how the capacity must be shared). The maximally loaded collision domain in the network limits the throughput of the network. This bottleneck can be used to estimate the throughput variable, based on the relationship $T = \frac{C}{L_{BCD}}$, where C is the capacity of the channel.

3.1 Linear Network

We will first evaluate a 1-dimensional network, where APs are arranged linearly and spaced equal distances apart. A major advantage of this network is that there is a single path between any two nodes within the mesh — therefor, for any traffic, we can be sure of the exact load created on any individual link. The network consists of a set of N access points, arranged in a linear configuration from node 1 to N. Spacing between the APs allows for communication with their immediate neighbors only — that is, AP_i can send and receive from nodes AP_{i-1} and AP_{i+1}. Therefore, there are $N-1$ links, with link l_j connecting nodes AP_j and AP_{j+1}.

A simple interference model will be used for this analysis, where a link interferes with any link within two hops. For this scenario, this corresponds to a model where interference

range is related to transmission range by a constant $\delta = 2$. In the linear network we used, this can be expressed as: l_i interferes with l_j if, and only if, $|i - j| \leq 2$.

Formation of collision domains using this model is straightforward. For link k, links $k-2$, $k-1$, $k+1$, and $k+2$ interfere with it ($k = 3 \ldots N-3$). Therefore, if L_i is the load on link l_i, then the load on link i's collision domain is $L_{CD}(i) = \sum_{k=-2}^{2} L_k$. L_{BCD} is defined as $\max(L_{CD}(i))$.

3.1.1 Gateway Traffic

First, we consider the scenario when all the traffic is gateway traffic. Assume that each AP sends the same amount of traffic (T_{GW}) to an Internet host H. We make this assumption to ensure the fairness among the nodes in the network. Without the assumption, the nodes near the gateway would probably get very high throughput while other nodes would barely get any, making the network useless to other nodes. Based on the bidirectional nature of links and interference, sending and receiving is equivalent, so only one direction must be considered. The gateway is located at AP_G. For multiple gateway networks, we will assume that clusters can be created, with APs associating with a GW. Therefore, for GW traffic, clusters are independent, each operating as a smaller, single-GW network.

The load on each link must be determined. In the linear network, there are $G-1$ APs and links to the left of the GW, and $N-G$ to the right. First, consider the links to the left of GW (l_i, $i = 1 \ldots G-1$). Traffic originating on the left side of the GW flows to the right. The link closest to the GW (l_{G-1}) must carry all traffic coming from this side — a total of $L_{G-1} = (G-1) \cdot T_{GW}$. Moving to the left, each subsequent link forwards for one less AP. Therefore, L_i can be expressed as $i \cdot T_{GW}$, for $i = 1 \ldots G-1$. Similar reasoning for the links to the right side of the GW gives $L_i = (N-i) \cdot T_{GW}$, for $i = G \ldots N-1$.

From this, L_{CD} and L_{BCD} can be determined. As L_i decreases further away from the GW, the bottleneck is clearly around the GW. Depending on the position (how centered the GW is), the BCD may include links on one or both sides of the GW. The extremes occur when the gateway is located at the ends of the network, or when it is centered. Located at an end ($G = 1$ or $G = N$), L_{BCD} is a maximum, at $5(N-3)T_{GW}$. This yields the minimum capacity for $T_{GW} = \frac{C}{5(N-3)}$. Centered gateway gives the maximum capacity, with $G = \frac{N-1}{2}$ (for odd N), $L_{BCD} = \frac{5N-13}{2} \cdot T_{GW}$, giving $T_{GW} = \frac{2C}{5N-13}$. The total capacity of the network is therefore $\frac{2NC}{5N-13}$.

3.1.2 P2P Traffic via the Gateway

A similar analysis applies to P2P traffic via the GW. Consider if each AP generates a traffic flow of volume T_{PG} to every AP in the network. This assumes that all the pairs

in the network has the same probability to communicate. The reason we choose this is to not favor the neighboring pairs, which is not always easy to get in practical usage. Forwarding of each flow consists of the source sending to the GW, and the GW sending to the destination. Therefore, each AP is the source for N flows to the GW, and the destination for N flows from the GW.

For the maximum capacity scenario (again, when the GW is centered), L_{BCD} is $N(5N-13)T_{PG}$, giving a throughput capacity of $T_{PG} = \frac{G}{N(5N-13)}$. With N^2 total flows, the network capacity is then $\frac{NC}{5N-13}$. Obviously, this is one-half the capacity of the network carrying only gateway traffic due to the two part relay process.

3.1.3 Direct P2P Traffic

Here we have each node send P2P traffic to every other node in the WMN. Each flow will have throughput T_{PD}. As in the relayed traffic scenario, this can be broken down into parts that are analogous to the GW scenario. Each AP sends to every other AP, therefore for any destination d, there are N flows to it. This is equivalent to sending to a GW at position d. To include all the traffic, this is repeated for all possible destinations.

The load on each link is therefore the sum of the loads using each node as a GW. Recall that if the GW is at G, then the link is $i \cdot T_{GW}$, for $i = 1 \ldots G-1$, and $L_i = (N-i)T_{GW}$, for $i = G \ldots N-1$. Link i is on the left of G $N-i$ times, and on the right i times. Therefore, the total link load is $L_i = [(N-i) \cdot i + i \cdot (N-i)]T_{PD} = 2i(N-i)T_{PD}$.

Again, this load is maximized at the center of the network, and is monotonically decreasing to the left and right of this point. Therefore, the calculation of L_{BCD} should include the 5 middle links — $i = \frac{N-5}{2}, \frac{N-3}{2}, \frac{N-1}{2}, \frac{N+1}{2}, \frac{N+3}{2}$. Loads on these links are $\frac{N^2-25}{2}T_{PD}$, $\frac{N^2-9}{2}T_{PD}$, $\frac{N^2-1}{2}T_{PD}$, $\frac{N^2-1}{2}T_{PD}$, and $\frac{N^2-9}{2}T_{PD}$ respectively. Summing these, we get $L_{BCD} = 5\frac{N^2-45}{2}T_{PD}$, for a throughput of $T_{PD} = \frac{2C}{N^2-45}$. Total network capacity is therefore $\frac{2N^2C}{N^2-45}$.

These results demonstrate that there is a clear advantage to directly supporting P2P traffic, rather than relying on GW-based forwarding. If the limit is taken as $N \to \infty$, the capacity of the direct approach is ten times greater than that of relayed traffic.

3.2 Grid Network

Using the same basic assumptions, the simple linear network is now replaced with a grid topology. The single GW network consists of a $n \times n$ array of APs, with the GW located at the center. It might not be exactly true in real life. however, in real life deployment,

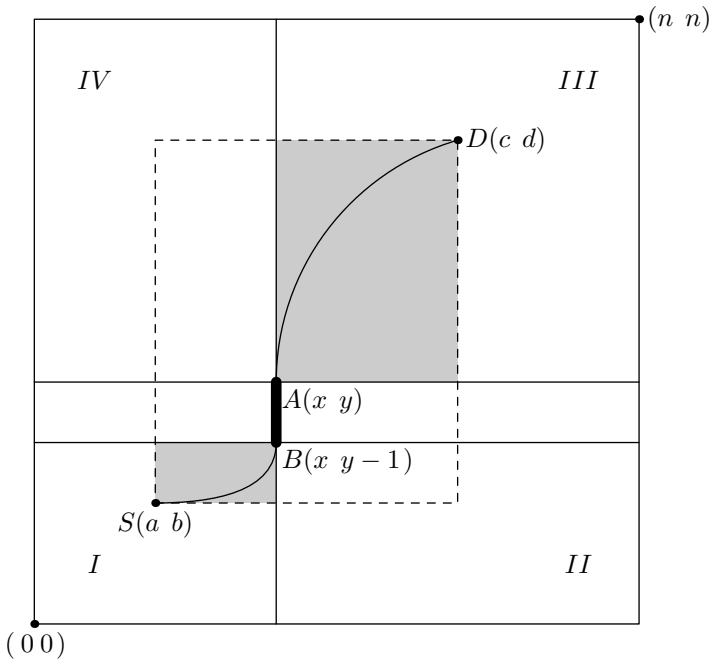


Figure 3.1: Link usage in a 2-dimensional network

the gateway is desirable to be placed nearer to the center to get better ceverage for more nodes. For reference purposes, a two-dimensional coordinate system will be adopted with the origin in the lower left corner.

Unlike the linear network, routing in the grid mesh is not deterministic — even while limiting to shortest (hop count) paths, multiple paths exist. Assuming that only hop count is used as a routing metric, any of these equal-length paths may be chosen. Therefore, we adopt a probabilistic approach for determining the expected load on individual links. Some routing algorithms will take into account more information in choosing the actual routing path. No matter what information they choose, the goal is to find the "best" path to the destination. "best" can be minimum hop count, maximum throughput, shortest delay, etc. In our analytical model here, all these criteria boil down to the link usage. So here the shortest path routing represents the "best" path routing which in practice may be other more sophisticated routing.

Consider a vertical link l_{AB}, as depicted in Fig. 3.1. The same approach can also be applied to a horizontally oriented link. We wish to determine the expected number of paths that flow through this link. As only shortest paths will be used, the source and destination APs must be in diagonal regions to each other — i.e. I and III, or II and IV.

The case where AP_S is in region I and AP_D is in region III will be used as an example.

AP_S is at coordinate $S(a,b)$, and AP_D is at coordinate $D(c,d)$. The probability that these two nodes use l_{AB} as a link along the path can be expressed as:

$$P = \binom{x-a+y-1-b}{x-a} \cdot \frac{\binom{c-x+d-y}{c-x}}{\binom{c-a+d-b}{c-a}} \tag{3.1}$$

Obviously similar probabilities can be obtained for traffic between other regions.

3.2.1 P2P via the Gateway

The previous probabilities can be applied to finding the expected link loads resulting from handling P2P via the gateway. Again, the P2P flow can be broken into two sections — to and from the gateway. There is flow from AP_S to G and from G to AP_D. However, the analysis for each of these cases is the same.

Consider the flow from AP_S to intermediate destination $G(p,q)$. The expectation that the flow will pass through l_{AB} is:

$$E_1 = \binom{x-a+y-1-b}{x-a} \cdot \frac{\binom{p-x+q-y}{p-x}}{\binom{p-a+q-b}{p-a}}$$

Note that if l_{AB} is in line with the GW (i.e. $x = p$) then the contribution from different network regions must be summed.

Each source node sends to N different destinations, each generating the above flow with the above expectation. The same expectation results for each destination in region I, where traffic flows in the reverse direction. Therefore, the total expectation for number of flows passing through l_{AB} is:

$$E(x,y) = 2 \cdot (N - N_G - 1) \sum_{\substack{1 \le a \le x, \\ 1 \le b \le y-1}} \frac{\binom{x-a+y-1-b}{x-a}\binom{p-x+q-y}{p-x}}{\binom{p-a+q-b}{p-a}}$$

Similar expressions can be obtained for expected link usage in the other regions of the network.

3.2.2 Direct P2P Traffic

The result from Eq. 3.1 can be directly applied to the P2P case. AP_S could be in any of the four regions created by the link l_{AB}. Given, that location, the destination must be in

the diagonally opposite region in order to use the link. Therefore, probabilities based on the four possible locations can be used to calculate the expected load.

The expected load contributed is based on the probabilities over all possible source-destination pairs. The expressions are region dependent. For traffic from region I to region III:

$$E_1 = \sum_{\substack{x \le a \le n, 1 \le b \le y-1, \\ x \le c \le n, y \le d \le n}} \frac{\binom{x-a+y-1-b}{x-a}\binom{c-x+d-y}{c-x}}{\binom{c-a+d-b}{c-a}}$$

Similarly, expectations for other regions are found:

$$II \to IV \quad : \quad E_2 = \sum_{\substack{x \le a \le n, 1 \le b \le y-1, \\ 1 \le c \le x, y \le d \le n \\ a \ne c}} \frac{\binom{a-x+y-1-b}{a-x}\binom{x-c+d-y}{x-c}}{\binom{a-c+d-b}{a-c}}$$

$$III \to I \quad : \quad E_3 = \sum_{\substack{x \le a \le n, y \le b \le n, \\ 1 \le c \le x, 1 \le d \le y-1}} \frac{\binom{a-x+b-y}{a-x}\binom{x-c+y-1-d}{x-c}}{\binom{a-c+b-d}{a-c}}$$

$$IV \to II \quad : \quad E_4 = \sum_{\substack{1 \le a \le x, y \le b \le n, \\ x \le c \le n, 1 \le d \le y-1 \\ a \ne c}} \frac{\binom{x-a+b-y}{x-a}\binom{c-x+y-1-d}{c-x}}{\binom{c-a+b-d}{c-a}}$$

Note that in the case where $a = c = x$, traffic should not be double counted. That is, for the case where both source and destination are on the border between regions, they should be counted in one of the regions only (e.g. I-III, but not II-IV). The total expectation for a link is then the sum of these four cases:

$$E(x, y) = E_1 + E_2 + E_3 + E_4$$

3.2.3 Comparison of Results

The expressions derived in the previous subsections provide link usages for the two approaches. Unfortunately, they do not provide succinct formulae that can be easily compared. Instead, they will be evaluated for specific scenarios, so that the results can be compared.

Similar to the work in linear networks, we consider the load within a collision domain. More links are affected in the grid scenario — as shown in Fig. 3.2, using the same interference model leads to a collision domain consisting of 23 total links.

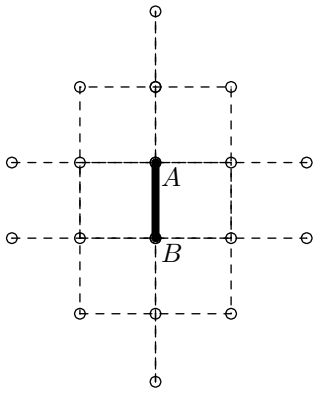
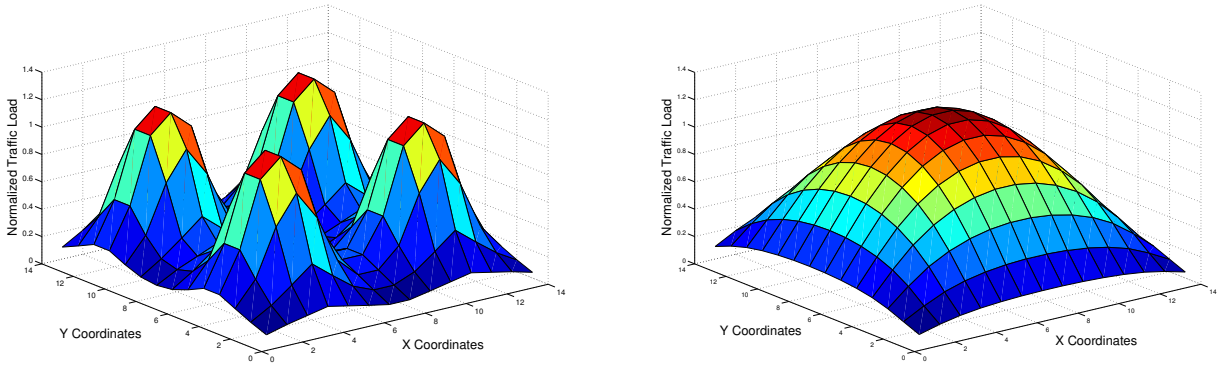


Figure 3.2: Collision domain in a Grid WMN



(a) Gateway Model

(b) Pure P2P Model

Figure 3.3: Link Usage Distribution in a 14×14, 4-GW Network

Fig. 3.3 shows the collision domain loads of traffic in a 14×14 network with 4 gateways. When traffic is forwarded via the GW, each GW is responsible for 7×7 cluster of APs. The traffic loads are normalized to 1 unit of total network (end-to-end) throughput. In the scenario using the gateways, it is clear that the bottlenecks occur around each of the GWs. Although each GW handles only a quarter of the total network throughput, a collision domain in these areas must carry traffic in excess of the total throughput. For peer-to-peer traffic, a single bottleneck occurs, although again exceeding the total network throughput.

No. Gateways	Cluster size	Direct	Gateway
1	3	2.5714	3.0000
	5	2.4117	4.0833
	7	1.9738	4.4083
	9	1.6299	4.5345
2	3	1.4686	1.7500
	5	1.2819	2.0417
	7	1.0289	2.2042
	9	0.8423	2.2673
4	3	2.1888	1.1562
	5	1.4950	1.0208
	7	1.1170	1.1021
	9	0.8890	1.1336

Table 3.1: Bottleneck Collision Domain Loads

We have evaluated a variety of different networks using this analysis. These include 1, 2, and 4 GW scenarios, with GW-clusters ranging from 3×3 to 9×9. Table 3.1 presents L_{BCD} for these scenarios. Recall that a lower load implies that a greater capacity can be achieved.

Several characteristics are visible from this data. As would be expected, for gateway traffic capacity decreases as cluster size grows, and increases as more gateways are added. For P2P traffic, throughput actually increases as the network grows. Comparatively, the direct P2P routing outperforms the GW approach in most cases. However, as the number of GWs increases, sending traffic via the GW has advantages, at least in small networks. As the cluster size grows, direct P2P again outperforms GW relaying.

Chapter 4

Simulation Study

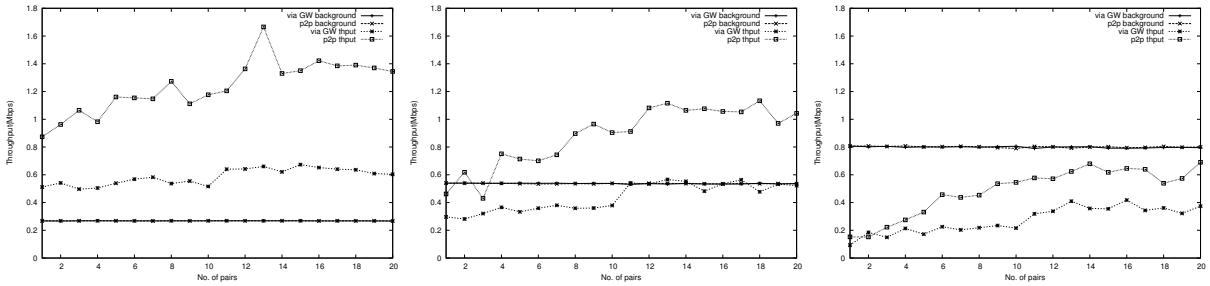
The analysis presented in Chapter 3 gives insights into the expected behavior of the system. Simulation is used to extend this understanding. The performance of a IEEE 802.11-based WMN is studied with ns-2 [19], in order to demonstrate the benefits of handling P2P traffic effectively.

The following methodology has been used. First, a capacity for gateway traffic has been established for each network topology. Gateway traffic is generated by creating a constant bit-rate source from each AP to the GW. Using constant-sized packets, the interval between arrivals is increased until a peak capacity is obtained, with the condition that packet loss remains low (below 5%) for every flow. After averaging this value over multiple runs, this peak rate is used as the 100%-level for gateway traffic.

After this rate has been established for each topology, both GW and P2P traffic can be considered. GW traffic is used as an underlying background traffic. Different levels of background traffic are used — 25, 50 and 75 percent of the peak gateway rate. Different number of P2P flows are also created, with source and destination nodes chosen randomly from the WMN APs. To determine the capacity for P2P traffic, the traffic rate is increased until losses of any flow surpasses the threshold value.

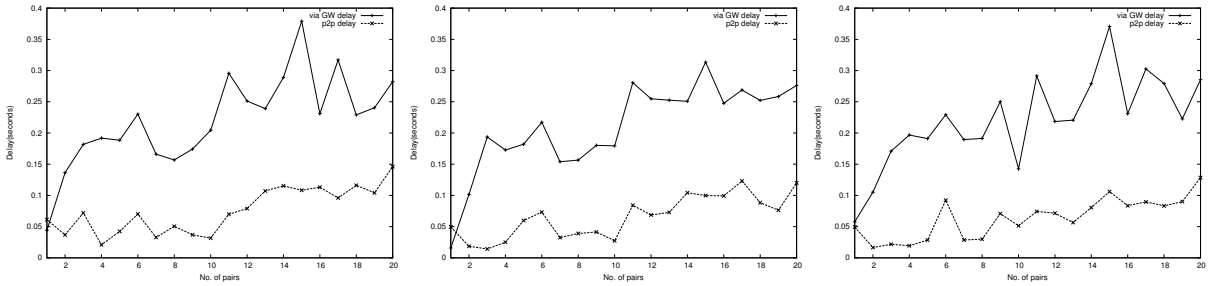
4.1 Single Gateway Networks

First, the performance of handling P2P traffic directly versus via the gateway is compared in a single gateway network. For these simulations $n \times n$-grid topologies are considered. We have restricted the choice of n to cases where n is odd. This allows the GW to be located at the exact center of the network, which should provide the optimal performance. Routing in the network is performed via AODV, except in the comparison of routing protocols.



(a) 25% background traffic (b) 50% background traffic (c) 75% background traffic

Figure 4.1: 7×7 grid throughput: GW and P2P Throughput with varying background traffic levels



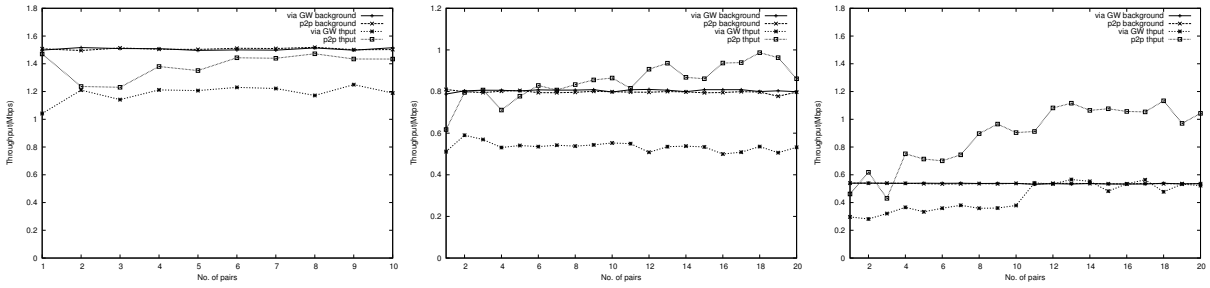
(a) 25% background traffic (b) 50% background traffic (c) 75% background traffic

Figure 4.2: 7×7 grid delay: Delay for P2P Flows with varying background traffic levels

Fig. 4.1 shows the total throughput achieved in a network of 7×7 network. Results are shown for each level of background traffic. By the nature of the experiment, the background throughput is constant for both the P2P direct and GW approaches. The values (approx. 0.27, 0.53 and 0.8) represent one-quarter, one-half and three-quarters of the network's capacity for GW traffic.

For the P2P traffic, the direct approach yields greater throughput than the GW approach for all cases. The greatest gains come in the scenarios with lower background traffic, where more capacity is available for the P2P traffic. The total throughput increases slightly as the number of flows (source- destination pairs) increases, as this has the effect of spreading (on average) the P2P load throughout the network, avoiding the likelihood of a single flow limiting the P2P capacity. The delay results for the same scenarios are shown in Fig. 4.2. Again, it is clear that the direct approach yields a much lower end-to-end delay than the GW approach. There is also more variation in the average delay, as all traffic is affected by the bottleneck congestion, resulting in a higher potential for long queuing delays.

In Fig. 4.3, throughput is again considered, this time in terms of differences in the



(a) 3×3 grid (b) 5×5 grid (c) 7×7 grid

Figure 4.3: GW and P2P Throughput in different sized networks

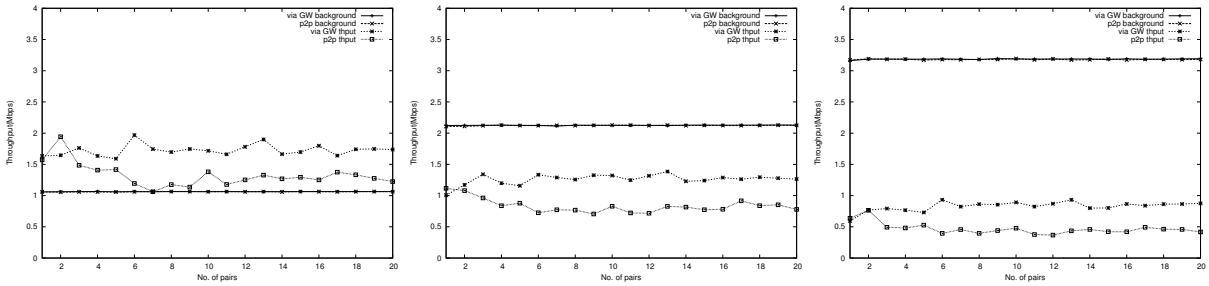
size of the network. The results for 3×3, 5×5, and 7×7 networks are shown, all at 50% background traffic. Again, it is clear that the direct approach has a clear advantage over the GW approach in all scenarios. However, the magnitude of this advantage is quite small. It is much larger in the 5×5, but only slightly larger in the largest network. From this, we can see that the size of the network impacts on the efficiency of the approach. For very small networks, there is little advantage to be gained, and all nodes are within interference range of each other. For large networks, there is a decreasing advantage over the gateway approach.

4.2 Multiple Gateway Networks

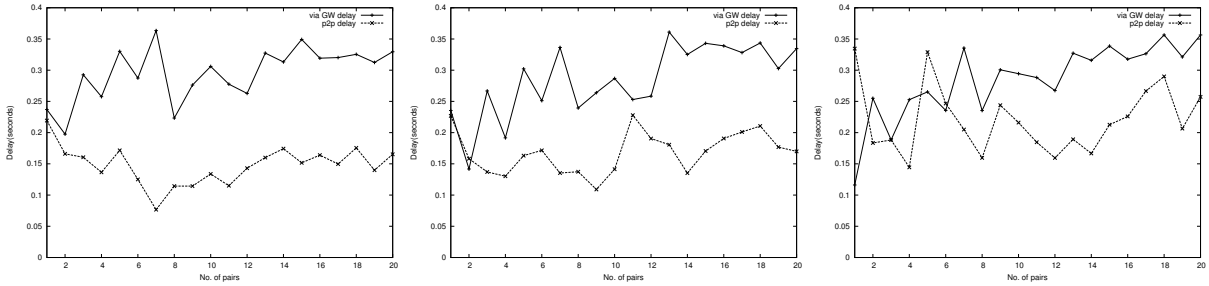
Using the same approach, networks with more than one GW will now be investigated. A 2-GW scenario is created by joining two adjacent 1-GW networks. The resulting network is a $n \times 2n$ grid.

Fig. 4.4 shows the throughput and delay for a 3×6 grid. In this scenario, we can see that the direct approach maintains its advantage. However, as the network size increases, this advantage appears to disappear. In a 5×10 grid, the throughput of the direct P2P approach is almost identical to that of the GW-relayed traffic, as seen in Fig. 4.5.

It appears that as the overall network gets larger, the increased path lengths of some of the paths eventually overwhelms any savings for nearby source-destination pairs. These longer paths are likely to cross congested areas of the network, so that no benefit is seen. In fact, when the size of the network is increased further, sending all traffic via the gateways becomes the preferred approach. Fig. 4.6 demonstrates this in a 4-GW 10×10 network.

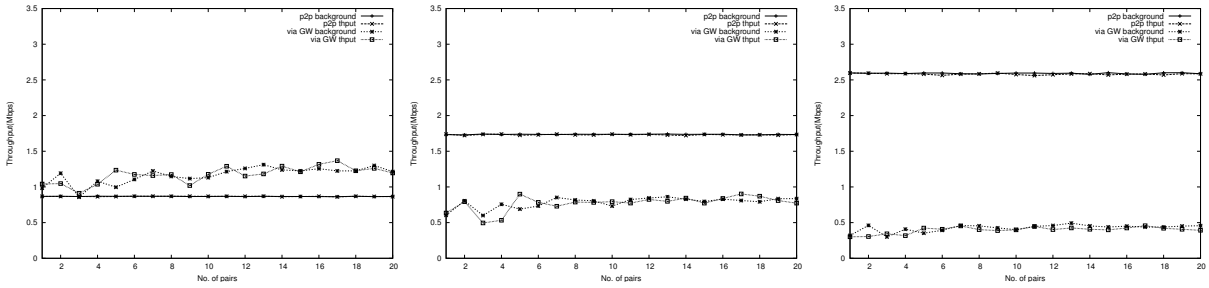


(a) Throughput: 25% background (b) Throughput: 50% background (c) Throughput: 75% background



(d) Delay: 25% background (e) Delay: 50% background (f) Delay: 75% background

Figure 4.4: Throughput and Delay in a 3 × 6, 2-GW Network

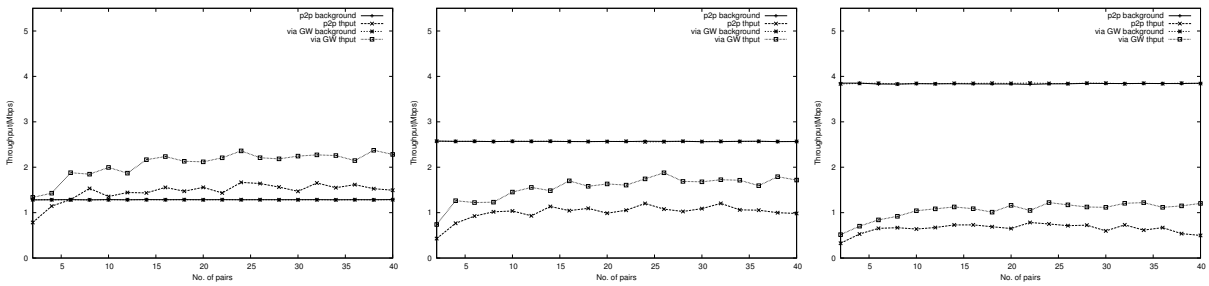


(a) 25% background (b) 50% background (c) 75% background

Figure 4.5: Throughput in a 5 × 10, 2-GW Network

4.3 Performance on Random Topologies

We generate 10 random networks, each of size 50. The nodes are randomly placed in an 1400x1400 area, which is the same area used by the 7x7 grid. While generating the network, we make sure that the network is connected. The gateways are randomly chosen from the network. For each topology, we measure the end-to-end delay and delivery ratio under different number of traffic flows. For each traffic flow number, we randomly generate



(a) 25% background (b) 50% background (c) 75% background

Figure 4.6: Throughput in a 10×10, 4-GW Network

50 runs of different sources and destinations. And then average the end-to-end delay and delivery ratio over all the 10 runs and 10 topologies. The data rate in this scenario is 80 kbps.

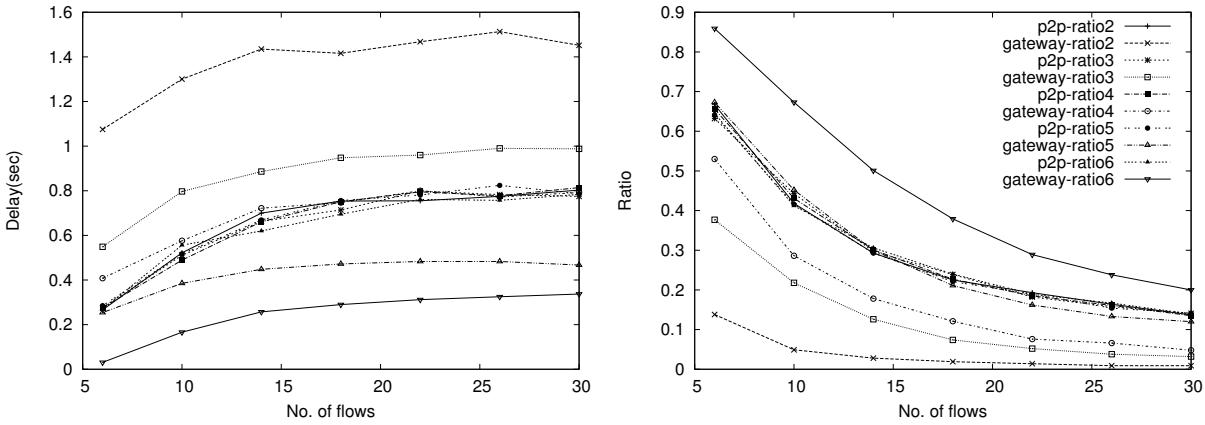
To measure the performance of these delivery mechanisms, we mainly use two metrics: end-to-end delay and delivery ratio. We choose different number of flows in the network and generate CBR traffic on these flows. We compare the performance of two different mechanisms: pure peer-to-peer mechanism and via the gateway mechanism.

In the pure peer to peer mechanism, all the nodes send the traffic to the destination directly, using the path found by AODV. The routing metric used is the hop count. In the via the gateway mechanism, all the traffic will pass through the gateway, like the one introduced before. Each non-gateway node will associate with one gateway, which is the shortest one to it, in terms of hop count.

Fig. 4.7 shows the results for end-to-end delay and delivery ratio in random networks of size 50. We run the simulation on 10 different topologies and with different rounds of random traffics. In Part (a), we can see that the end-to-end delay of via the gateway mechanism decreases as the number of gateways increases, while the end-to-end delay of the pure peer to peer mechanism is very close to each other under different gateways, since the route does not depend on the gateway. An interesting observation is that the delay of the pure peer to peer mechanism is the same as the via the gateway mechanism with 4 gateways.

In Part (b), the delivery ratio of via the gateway mechanism increases as the number of gateways increases, while the peer-to-peer mechanism stays the same for different gateways. This time, the performance of via the gateway mechanism is the same as the peer to peer mechanism with 5 gateways.

These results not only compare the performance of the pure peer-to-peer mechanism and via the gateway mechanism, but also give some implications about how we should



(a) Delay (b) Delivery ratio

Figure 4.7: Delay and Delivery ratio in Random Networks of Size 50

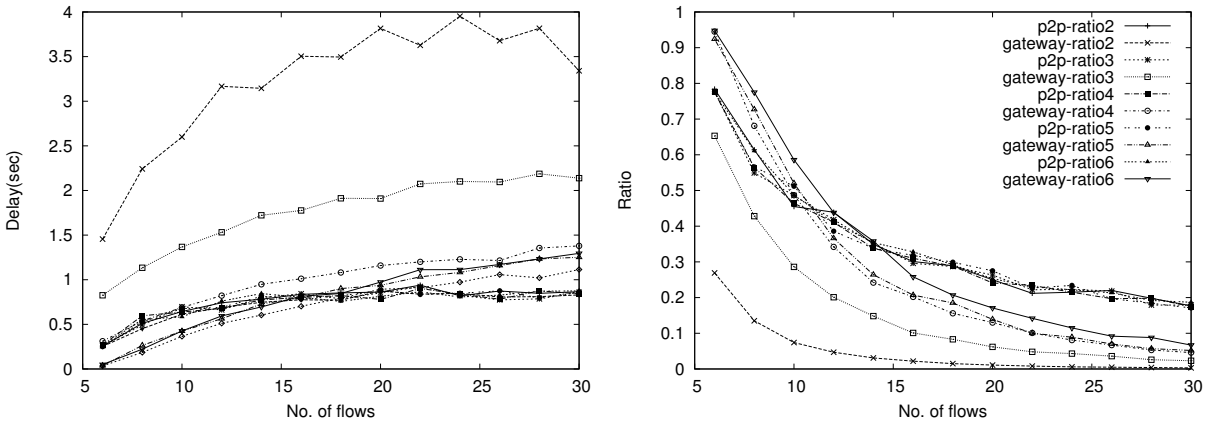
deploy gateway nodes in wireless mesh networks to satisfy some traffic requirements.

4.4 Performance on Known Topologies

After demonstrating the idea on random topologies, we evaluate this approach in a real life network topology: the Chaska wireless networks. There are 195 nodes in the Chaska network. Due to the simulation limitation and also to compare with the previous results, we choose a 50 nodes subset of the Chaska network which is in the middle of the original Chaska network. The gateways are randomly chosen in the network. The traffic patterns and routing mechanisms are the same as those in the random toplogy case.

Fig. 4.8 (a) shows that the direct delivery mechanism has a similar performance with the gateway mechanism with 4 gateways. In part (b) the delivery ratio performance of direct delivery mechanism is similar with gateway mechanism with 5 gateways.

These results show that in a real life network topology, due to the potential limitations inherent to the locations and other restrictions which may prevent the "optimal" or fair placement of wireless routers, the direct mechanism still has advantages over the pure gateway mechanism in terms of end-to-end delay and delivery ratio.



(a) Delay (b) Delivery ratio

Figure 4.8: Delay and Delivery ratio in Chaska Networks of Size 50

Chapter 5

P2P Aware Wireless Mesh Protocol

The analysis in Chapter 3 and simulation in Chapter 4 show that the direct delivery mechanism will have a better performance than in the gateway mechanism for small networks. When the network becomes larger, the benefits are less significant. The reason is that the gateway mechanism forms a bottleneck at the gateway, and as the network goes larger, the direct delivery mechanism also tends to form a "hot spot" region in the center of the network.

To make use of the peer-to-peer potential and also avoid the "hot spot" region incurred from direct delivery, we propose the Peer-to-Peer Aware Wireless Mesh Protocol (PAWMP). The basic idea of PAWMP is to route short-path packets using the P2P direct delivery mechanism and long-path packets using the via gateway mechanism. The challenge in designing PAWMP lies in how to identify short-path packets and how to find a good path.

The IETF draft standard proposed for WMNs is the 802.11s[2]. It uses the Hybrid Wireless Mesh Protocol (HWMP) [21] as the layer-2 (link layer) routing protocol. To make PAWMP more general and compatible with the proposed standard, we build PAWMP as an extension of HWMP.

In this chapter, we first give an overview of HWMP. Then we describe our extensions, namely the message format, and the routing procedures. We also discuss the overhead introduced by PAWMP and evaluate its performance through simulations.

5.1 HWMP Overview

The Hybrid Wireless Mesh Protocol is a mesh routing protocol that combines the flexibility of on demand routing with proactive topology tree extensions. 802.11s, the draft IEEE

802.11 amendment for mesh networking, defines the HWMP as the default mandatory routing protocol.

HWMP is mainly based on the Ad Hoc On Demand (AODV) routing protocol, with the message primitives borrowed from AODV and adapted for Layer-2 (link layer) address-based routing and link metric awareness.

HWMP supports two modes of operations depending on the configuration. These modes are:

- On demand mode: allows MPs to communicate using peer-to-peer routes and is used when there are no gateways in the network.
- Proactive tree building mode: performed by the gateway, which builds a routing spanning tree by broadcasting route building messages.

These modes are not exclusive and may be used concurrently.

HWMP control messages are the Route Request (RREQ), Route Reply (RREP), Route Error (RERR) and Root Announcement (RANN). The metric cost of the links determines which routes HWMP builds.

5.1.1 HWMP Message Format

RREQ Message

Octets: 1	1	1	1	1	4	6	4	4	...
Element ID:	len	Flags	Hopcount	TTL	REQ ID	Src	SSN	Lifetime	

4	1			6	4		1			6	4
Metric	Per Dest Flags			Dest #1	DSN #1	...	Per Dest Flags			Dest #N	DSN #N
	DO #1	RF #1	RES				DO #N	RF #N	RES		

where *Src* is the *Source Address*, *SSN* is the *Source Sequence Number*, *DO* is the *Destination Only* flag, *RF* is the *Reply-and-Forward flag*, and *RES* is reserved.

RREP Message

Octets: 1	1	1	1	1	6	4	4	...
Element ID:	len	Flags	Hopcount	TTL	Dest	DSN	Lifetime	

4	6	4	1	6	4		6	4
Metric	Source Address #1	SSN	MP count	MP Addr #1	MP DSN #1	$\cdots$	MP Addr #N	MP DSN #N

where *Dest* is the *Destination Address*, *DSN* is the *Destination Sequence Number*, *MP* is the *dependent Mesh points*.

RANN Message

Octets:1	1	1	1	1	6	4	4	4
Element ID	len	Flags	Hopcount	TTL	Src	DSN	Lifetime	Metric

where *TTL* is the *Time to Live* value, *DSN* is the *Destination Sequence Number*.

5.1.2 On Demand Routing Mode

If a source MP needs to find a route using the on demand routing mode, it broadcasts a RREQ with the destination MP specified in the destination list and the metric field initialized to 0.

When a MP receives a RREQ it creates a route to the source or updates its current route if the RREQ contains a greater sequence number, or the sequence number is the same as the current route but the RREQ offers a better metric than the current route. If a new route is created or an existing route updated, the RREQ is re-broadcast. Each MP may receive multiple copies of the same RREQ that originated in the source, each RREQ traversing a unique path from the source to the MP.

Whenever a MP forwards a RREQ, the metric field in the RREQ will be updated to reflect the cumulative metric of the route to the RREQ's source. After creating or updating a route to the source, the destination MP sends a unicast RREP back to the source.

5.1.3 Proactive Tree Building Mode

There are two mechanisms for proactively disseminating routing information for reaching the root MP. The first method uses a *proactive* Route Request (RREQ) message and is intended to create routes between the root and all MPs in the network proactively. The second method uses a Root Announcement (RANN) message and is intended to distribute

route information for reaching the root but the actual routes to the root can be built on demand.

A MP configured as the root (gateway) would send either proactive RREQ or RANN messages periodically.

Proactive RREQ mechanism

The RREQ tree building process begins with a proactive *Route Request* message sent by the root MP, with the destination address set to broadcast address. So each node that receive the RREQ will build a reverse path to the root, send a RREP back to the root, and also re-broadcast this proactive RREQ message. In the end, all the nodes within the network will periodically update the path to the root, and the root also has the route information to all the nodes.

RANN mechanism

The RANN periodically floods a RANN message into the network, which contains the route metrics to the root.

Upon reception of the RANN message, each MP that has to create or refresh a route to the root will send a RREQ to the root. The root then responds with a RREP to every RREQ message. In this way, the root builds a reverse path to the MP, and the MP finds a path to the root.

5.2 PAWMP Extensions

In our Peer-to-Peer Aware Wireless Mesh Protocol (PAWMP), the on demand mode works the same way as the HWMP. We add two message types in the proactive mode, Discovery Request (DREQ), and Discovery Reply (DREP). Before sending the payload to the root, the MPs first send DREQ packets to discover whether the destination is within the same wireless mesh network. And if the answer is yes according to the reply message DREP, the MPs will send the packets to the destination using the on demand mode. Otherwise they will send the packets to the gateway.

5.2.1 DREQ Format

The proposed DREQ message has the following format:

Octets: 1	1	1	1	1	6	4	4	4	...
Element ID:	len	Flags	Hopcount	TTL	Src	SSN	Lifetime	Metric	

6	4	4	...
Root MAC Address	RSN	Dest IP	NAT Reserved

- *Root MAC Address*: The layer 2 address of the root
- *RSN*: Root Sequence Number
- *Dest IP:Port*: The destination IP address (4 octets) and port number (2 octetes).
- *NAT Reserved*: when the NAT box is used at the gateway for all outgoing and incoming traffic in the wireless mesh network, more information is needed for the protocol to work. This will be explained subsequently.

5.2.2 DREP Format

The proposed DREP message will have the following format:

Octets: 1	1	1	1	1	6	4	4	4	...
Element ID:	len	Flags	Hopcount	TTL	Root	RSN	Lifetime	Metric	

6	4	4	6	...
Source MAC Address	SSN	Dest MAC Address	Original Dest IP	NAT Reserved

- *Source MAC Address*: The MP that initiates this DREQ/DREP process.
- *SSN*: The Sequence Number of the originator.
- *Dest MAC Address*: The MAC address of the destination within the network. If the destination is an Internet host outside the wireless mesh network, this value will be the MAC address of the root.
- *Original Dest IP:Port*: This will help the source MP to determine the mapping between the layer 2 address (MAC) and layer 3 addresses (IP).
- *NAT Reserved*: This will be explained in Section 5.2.3.

5.2.3 PAWMP Format with NAT

When a NAT box is used at the gateway, the messages will be different, since all the MPs in the network will have the same IP address. Therefore the destination IP address is not sufficient to locate the destination in the network. We add port information into the DREQ/DREP message. Also, there is a IP address and port manipulation at the NAT box, so we have to do a similar conversion at each MP. The detail is as follows.

DREQ with NAT

For the NAT reserved field in section 5.2.1, we append the following fields:

...

2	4	2
Dest Port	Source IP	Source Port

The Source IP and port are contained in the message for the root to match or create a new port mapping for the source-end of the socket. That is, both source and destination have the same public IP address $ADDR_R$ (the address allocated by the root). The two ends of the socket link between the source and destination will both be mapped by the root NAT box, with the format $\{ADDR_R : Port_S\}$ and $\{ADDR_R : Port_D\}$. The source IP and port contained in the message are used to be allocated a new port by the root NAT box, that is, a new entry $\{ADDR_R : Port_S\}$. The same applies for the destination.

DREP with NAT

For the NAT reserved field in section 5.2.2, we append the following fields:

...

2	4	2	4	2
Original Dest Port	New Dest IP	New Dest Port	Mapped Source IP	Mapped Source Port

- *Original Dest Port*: used to identify the original flow
- *New Dest IP*: the private IP of the destination within the network
- *New Dest Port*: the port corresponding to the private IP. These two fields are used to modify the corresponding field in the IP packets. Emulating a local NAT box function.
- *Mapped Source IP*: the public IP allocated to the source.
- *Mapped Source Port*: the corresponding port.

After receiving this DREP, the source will apply a NAT-box-like conversion on all the IP packets corresponding to the flow identified by the dest IP and port. In this way, the traffic originally destined to the root will be re-routed in a peer-to-peer manner in the network.

5.2.4 PAWMP Procedure

In a wireless mesh network with several gateways, each gateway has its own Internet connection. A node joining the network will be associated with a specific gateway. Thus, the overall wireless mesh network is divided into different clusters, where a cluster is a gateway with its associating nodes.

The IP addresses of mesh points are assigned by the associated gateways. Depending on different address allocation mechanisms, i.e., whether Network Address Translation (NAT) is used or not, the method to check whether the destination is within the same cluster may be different. In the NAT case, if the destination has the same global IP address as the gateway or the source mesh point, then they are within the same cluster. Otherwise, e.g., if all the mesh points have global IP addresses and each cluster forms a subnet, the source mesh points are able to check whether the destination IP address is in the same subnet as the source.

Gateway Association

In the initial gateway association stage, each gateway broadcasts RANN messages into the wireless mesh network. The RANN message contains the MAC address, IP address, and subnet mask information related to the configuration of the gateway. The nodes that receive the RANN message store the gateway information and compare with the existing information they have, if any. The algorithm for choosing the gateway is based on the routing cost to the gateway, which is contained in the RANN message, and the current load of the gateway.

If the RANN message is the most current information received from that gateway, the node rebroadcasts this RANN message. It will also add the cost of the last link to the total routing cost field, and append itself to the path to the gateway. In this way, each node receiving the RANN message will be able to know the neighbors towards the gateway.

Discovery Query and Reply Procedure

Before sending data packets, each node has to choose an appropriate route which is not necessarily the routing path. The destination can be in one of the four categories: outside

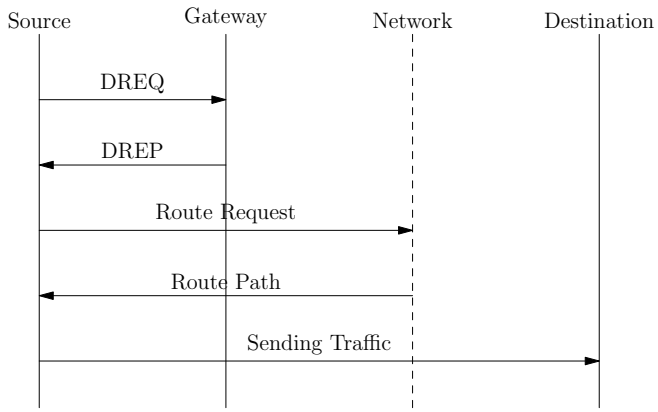


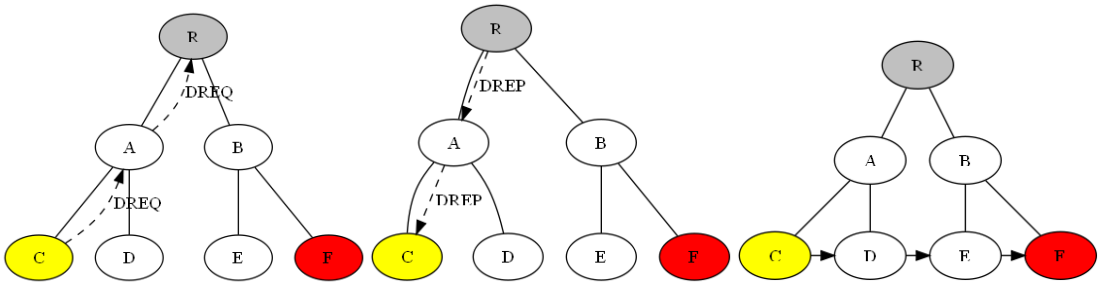
Figure 5.1: Sequence diagram of PAWMP procedure

the wireless mesh network; in the current cluster; in neighboring clusters and "near" the current node; in a different cluster but "far" from the current node. The "near" and "far" will be defined more formally in the following sections. According to the experience gained from the previous analysis and simulation, our goal here is to use the point-to-point path for the destination in the same cluster and "nearby" nodes. For the other cases, we send the packets to the gateway which will then forward the packets to the destination.

Fig. 5.1 is the sequence diagram of the query and reply procedure. After receiving the response from the gateway, the source initiates a route request for the destination in the network. When the source has enough routing information to the destination, it sends the data traffic to the destination directly.

To check whether the destination belongs to the same cluster, if using global IP address, the source can look at the IP address of the destination and subnet mask of the gateway. If the IP address falls into the range of the subnet, the source will know that the destination is within the same cluster as itself. When using the NAT mechanism, the source can check whether the destination has the same global IP address as the gateway.

In the most commonly used transport layer protocols TCP and UDP, to identify a connection or flow, we also need the port number in addition to the IP address. Therefore, we must also make sure that the port number is correct. In the global IP case, there is no problem since the (IP, port) pair will not change during the forwarding. However, in the NAT case, the gateway will have to map the (private_IP, local_port) to the (global_IP, gateway_port) since the source node only has the (global_IP, gateway_port).



(a) MP C send a DREQ to the Root R

(b) Root R reply to C with a DREP

(c) MP C sends packets directly to MP F

Figure 5.2: The process when C wants to send traffic to F

Check Vicinity

In this part, we explain how to check whether the nodes are close to each other. We want to find an efficient way to estimate the vicinity of a node rather than a precise information.

It is possible to obtain the distance between any two nodes through flooding. However, the overhead of this approach is not tolerable, since the check will be executed for every flow. Instead, we consider the nodes in neighboring clusters as "near" and we use the peer-to-peer delivery between them.

During the gateway association process, a node usually receives the broadcast from more than one gateway and chooses the best one as its default gateway. Here we use the remaining gateway candidates to help find the neighboring clusters. If the destination address belongs to the subnet of a candidate gateway, then the source can decide that this node is "near" and send the packets using the peer-to-peer delivery mechanism. Note that in this case, the DREQ packets will be sent to the candidate gateway, not the gateway the source is associated with.

5.2.5 An Example

In Fig. 5.2, we show an example where node R is the root (gateway), C is the source, and F is the destination. Before sending packets directly to the gateway, MP C first sends the DREQ packets to R, querying whether F is within the same network. After it receives a positive answer through DREP, C sends the packets directly to the destination F.

5.3 PAWMP Overhead

The overhead of this protocol concerns the communication, the local processing and storage.

5.3.1 Throughput Overhead

The throughput overhead can be expressed as follows.

$$Throughput\ Overhead = \frac{D}{D+P}$$

where D is the total volume of the DREQ and DREP packets, and P is the size of the payload traffic.

We only generate DREQ and DREP packets when a new flow comes. So the D can be expressed as

$$D = s_D \cdot \frac{P}{s_{flow}}$$

where s_D is the average size of the DREQ plus DREP packets, and s_{flow} is the average flow size.

Thus the throughput overhead can be computed using the formula:

$$Throughput\ Overhead = \frac{s_D}{s_D + s_{flow}} \tag{5.1}$$

For each flow, theoretically we only need one pair of DREQ and DREP. And based on the formats in Section 5.2, s_D in this case is less than 100 bytes. In practice, we can also update the DREQ and DREP during every interval d. This will maintain the freshness of the address mapping, which is especially good for long flows. For this scenario, the overhead can be expressed as

$$D = s_D \cdot \frac{t_{flow}}{d} \cdot \frac{P}{s_{flow}} = s_D \cdot \frac{P}{d \cdot r_{flow}}$$
$$Throughput\ Overhead = \frac{s_D}{s_D + d \cdot r_{flow}} \tag{5.2}$$

where d is the time interval to refresh the discovery information, t_{flow} is the average flow duration, and r_{flow} is the average flow rate.

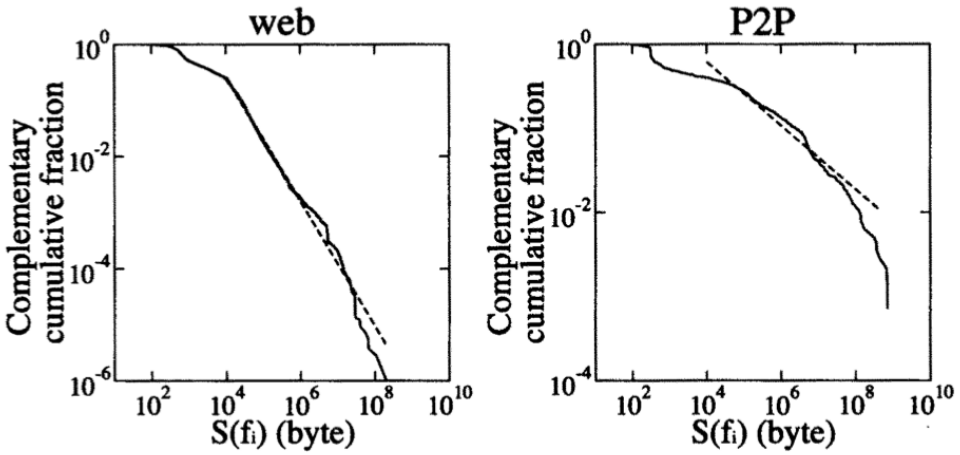


Figure 5.3: Complementary cumulative distribution of flow size

According to the estimate of daily Internet usage, for web browsing, the average flow size (the html file size) is in the order of tens of kilobytes; for p2p file sharing, the average flow size (the file segment size) is in the order of several hundred kilobytes or several megabytes; and for voice over IP (VoIP) and video streaming, the flow size is in the order of 10's or 100's megabytes respectively.

Fig. 5.3 from [25] gives the complementary cumulative distribution of flow size in web and P2P traffic. The authors of [25] showed that the average flow size for web and P2P are $20.58k$ and $5.81M$ bytes respectively, and the average flow rates are $72.46k$ and $12.94k$ respectively. Thus the throughput overhead is less than 0.5% if the DREQ and DREP are sent only once. In the second scenario, the discovery messages are updated every d seconds. d should be in the order of several seconds, similar to the routers update frequency. So even when d is as small as 1, the throughput overhead is still less than 1%.

5.3.2 Delay Overhead

The initial DREQ and DREP will add delay to the first flow packet. That is, the flow has to wait for the round trip of the DREQ and DREP. As we will show through simulations, the end-to-end delay is in the order of few seconds. This means that PAWMP will add a delay of few seconds before transmitting the first packet of each flow. And it will not affect the following packets. Actually, since PAWMP will have better delay performance over the direct or gateway mechanism, the latter packets should have lower delay.

5.3.3 Local Processing and Storage

To distinguish whether the destination is within the same network, each node will have to maintain a list of IP addresses that is located within the network. Due to the nature of the wireless mesh network, normally the number of wired gateways is limited compared to the number of MPs. So the memory space required for the IP addresses is sub-linear to the number of routers in the network.

For the NAT-box case, each router should also maintain an additional local address table which otherwise only resides at the gateway. This is similar to [4, 30] where NAT at the router was used to provide access for mesh clients. Accordingly, PAWMP at most doubles the NAT table size, and that happens when all traffic flows are within the WMN and need a conversion entry from the gateway. However, even that is not a big issue. The NAT conversion for the packets is the same as the local NAT conversion.

5.4 PAWMP Evaluation

In this section, we evaluate the new routing protocol under both the Chaska network toplogy and random network topologies in ns-2. The simulation shows that our protocol performs better than the direct delivery mechanism and via the gateway mechanism.

5.4.1 Performance on Ideal Grid Network

We first run our proposed protocol in a 10×10 grid network with 4 gateways. Different amounts of background traffic are added in the simulations. Fig. 5.5 shows the throughput performance comparison of PAWMP, direct delivery mechanism and via the gateway mechanism.

The background traffic varies from 25% to 75%. In all the three cases, PAWMP has better performance than the other two. Also, the direct delivery mechanism also has a slightly better performance than via the gateway mechanism.

Fig. 5.5a also shows that when the background traffic is low, i.e., when the network is not congested, PAWMP achieves more gain than via the gateway mechanism. As the background traffic increases, more resources are occupied in the network and thus less opportunity left for improvements.

Another phenomenon is that when the number of point-to-point flows are few (less than 10), the performance of the three mechanisms are similar. The reason is that the fewer the flows, the less probable that the network will be congested. Therefore most of the flows will get the opportunity to pass through.

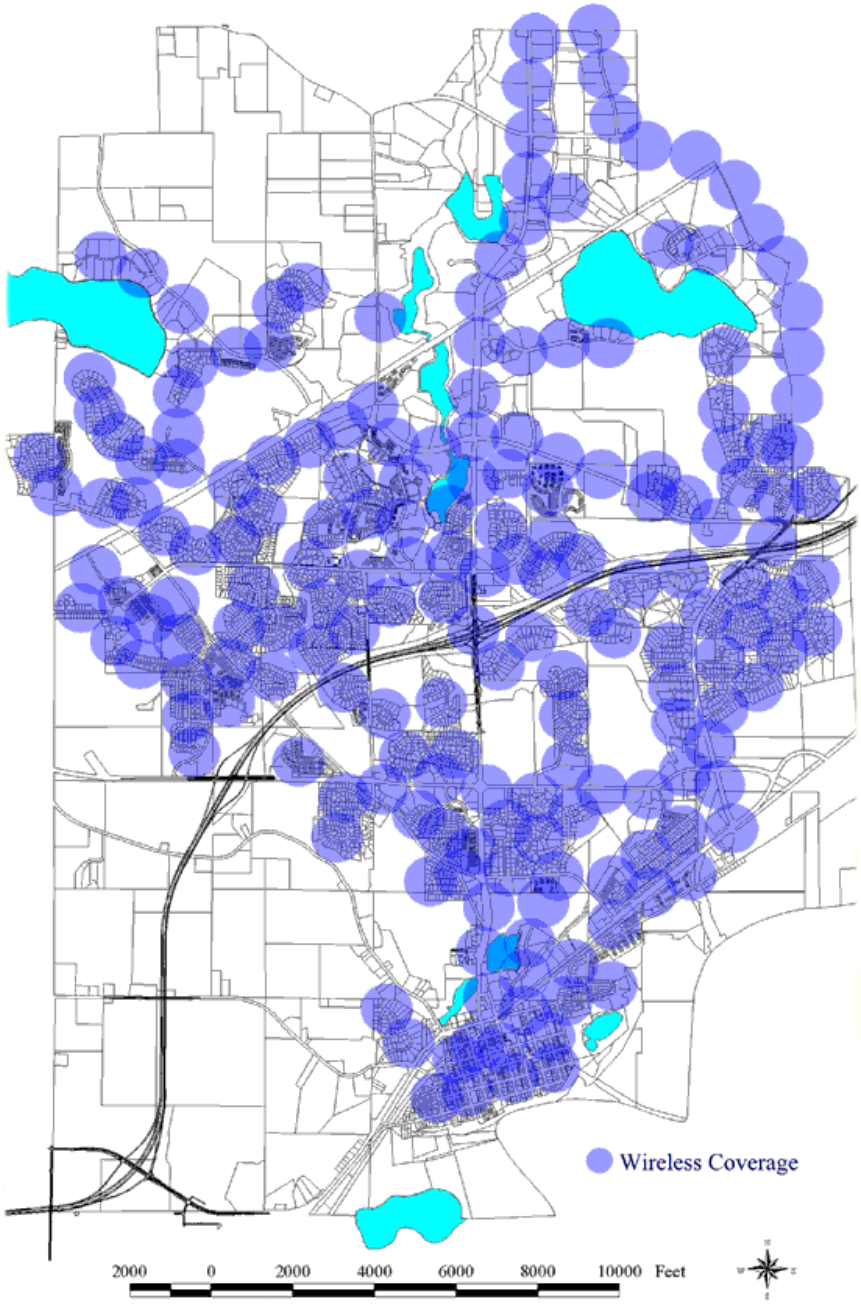
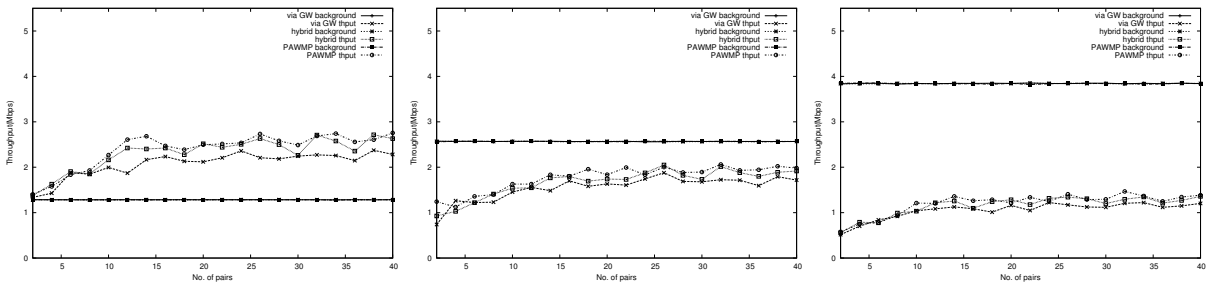


Figure 5.4: Chaska Wireless Network Map



(a) 25% background (b) 50% background (c) 75% background

Figure 5.5: PAWMP Throughput Comparison in a 10×10, 4-GW Network

5.4.2 Performance on Random Networks

In the second set of simulations, we run our protocol on 10 random networks of size 50, each with 4 gateways and under 7 different data rates. The results are shown in Fig. 5.6.

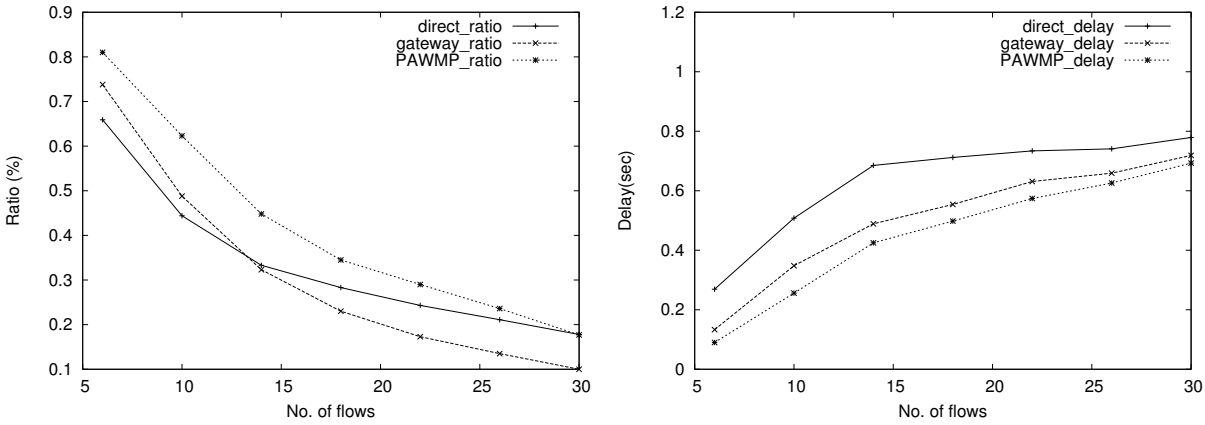
In Fig. 5.6a, we see that the delivery ratio of PAWMP is better than the direct and via the gateway mechanism. As the number of flows increases, the performance difference between PAWMP and direct delivery decreases.

In Fig. 5.6a, we can see that the delivery ratio of via the gateway is better than the direct mechanism for small number of flows (less than 14). However, the direct mechanism outperforms via the gateway with larger numbers of traffic flows. The reason is that when the number of flows is small and the network is not too loaded, the gateway mechanism has the advantage of using shorter paths for inter-cluster flows, the impact of intra-cluster flows being not significant. However, as the number of flows increases and gateway becomes congested, all the flows are affected. That is why the performance of via the gateway mechanism degrades faster than the direct mechanism.

On the contrary, PAWMP does not have the problem of long intra-cluster paths. Therefore it performs well for small number of flows. When the number of flows is large and the network is very loaded, PAWMP has slightly better performance than the direct mechanism. The reason is that most of the links are occupied and it is hard to exploit any more potential.

5.4.3 Performance on Chaska Wireless Network

The simulation settings for the Chaska network are the same as those in random networks. Fig. 5.7 shows the performance comparison of PAWMP, direct delivery mechanism and via the gateway mechanism.



(a) Delivery Ratio Comparison in Random Networks (b) Delay Comparison in Random Networks

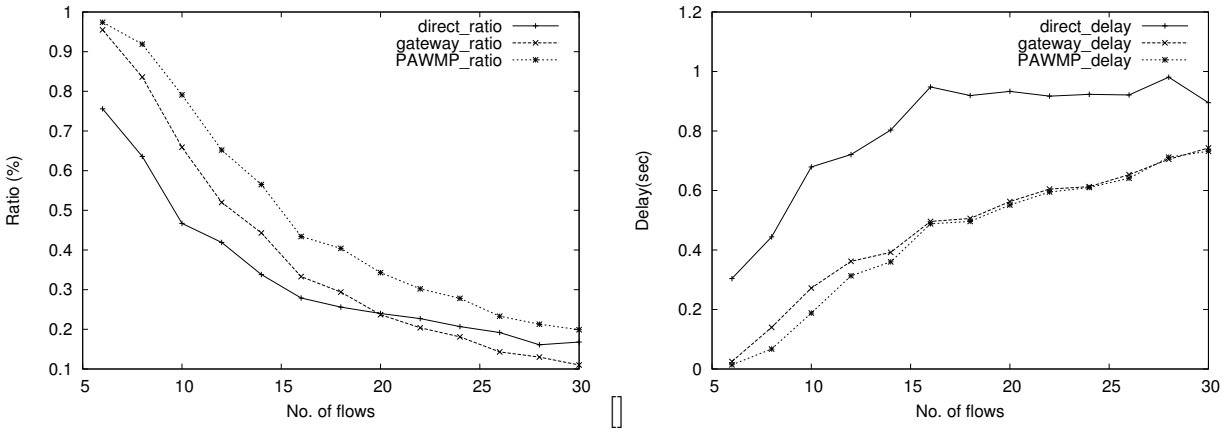
Figure 5.6: PAWMP Performance Comparison in Random Networks

Fig. 5.7a is the delivery ratio comparison result. We can see a similar behavior as that in random networks. The delivery ratio of PAWMP is about 10% better than the other two mechanisms. Via the gateway performs better than direct delivery when the number of flows is less than 20.

Fig. 5.7b is the delay comparison result. Direct delivery has longer delay than PAWMP and via the gateway. The reason is that in multiple gateway scenarios, direct delivery mechanism has longer paths on average, and the delay of PAWMP is only slightly better than that of via the gateway mechanism. The benefit comes from the direct peer-to-peer paths.

5.4.4 Summary

Through simulations on different topologies and different traffic patterns, we showed that PAWMP has a better average performance in throughput, delivery ratio and end-to-end delay compared to via the gateway and direct delivery mechanisms in the simulation setting. However, the benefit gain also depends on the volume of background traffic and the total number of traffic flows in the network. The gain is larger when the network is not congested. Even under extreme conditions, PAWMP still has a slightly better average throughput and delivery ratio. The average end-to-end delay is always better than direct delivery mechanism and via the gateway mechanism.



(a) Delivery Ratio Comparison in Chaska Network

(b) Delay Comparison in Chaska Network

Figure 5.7: PAWMP Performance Comparison in Chaska Network

Chapter 6

Conclusions and Future Work

As an access network, it remains likely that a large percentage of the traffic in a WMN will be directed through the gateways to/from the Internet. However, point-to-point communication between users in the mesh network may also be significant in many scenarios. For example, in neighbourhoods or on campuses, many users are likely to communicate with each other. It is therefore important for this traffic to be handled appropriately.

In this work, P2P traffic has been considered in the presence of GW traffic. In handling P2P traffic, we must always be aware of the effects it has on network links – the bottleneck created around the gateway affects all traffic in the network. However, the appropriate handling of P2P traffic can relieve some of the pressure on these congested links. By doing so, the capacity of the network can be improved. By contrast, handling all P2P traffic via the gateway adds traffic and further constricts the bottleneck.

When multiple gateway networks are considered, a balance must be achieved. While the direct handling of P2P traffic can still yield a benefit in some situations, in others the added cost created by long forwarding paths can be detrimental. Simulations showed that the pure strategy of sending all P2P traffic directly resulted in a net decrease in network throughput.

Based on these observations, we proposed PAWMP, which make use of local potential of direct delivery mechanism as well as avoiding long P2P forwarding paths and congested areas created by P2P paths. Through simulations on different topologies, we showed that the average performance of PAWMP is better than the pure direct delivery mechanism or via the gateway mechanism in the simulation setting.

Future improvements of this work may involve new wireless technologies or new routing schemes. For example, multi-radio multi-channel routers [7] are used in some WMN testbeds. The added radios/channels can be used to make the DREQ/DREP request faster and interfering less with the data channels. Network coding[14, 5, 13] is also a popular

trend in WMN routing. In network coding, in addition to forwarding packets, routers mix packets from different sources to increase the information content of each transmission. In WMN, there is a considerable amount of traffic from and to the gateway. Finding a coding scheme to encode and decode DREQ and DREP messages efficiently can reduce the overhead and improve the overall performance.

References

[1] Ashish Agarwal and P. R. Kumar. Capacity bounds for ad hoc and hybrid wireless networks. *SIGCOMM Comput. Commun. Rev.*, 34(3):71–81, 2004. 7, 9

[2] M. Bahr. Update on the hybrid wireless mesh protocol of ieee 802.11s. pages 1 –6, oct. 2007. 27

[3] Carlos J. Bernardos, Maria Calderon, Ignacio Soto, Ana Beatriz Solana, and Kilian Weniger. Building an ip-based community wireless mesh network: Assessment of pacman as an ip address autoconfiguration protocol. *Comput. Netw.*, 54(2):291–303, 2010. 10

[4] John Bicket, Daniel Aguayo, Sanjit Biswas, and Robert Morris. Architecture and evaluation of an unplanned 802.11b mesh network. In *MobiCom '05: Proceedings of the 11th annual international conference on Mobile computing and networking*, pages 31–42, New York, NY, USA, 2005. ACM. 10, 38

[5] Szymon Chachulski, Michael Jennings, Sachin Katti, and Dina Katabi. Trading structure for randomness in wireless opportunistic routing. In *SIGCOMM '07: Proceedings of the 2007 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 169–180, New York, NY, USA, 2007. ACM. 43

[6] Douglas S. J. De Couto, Daniel Aguayo, John Bicket, and Robert Morris. A high-throughput path metric for multi-hop wireless routing. In *MobiCom '03: Proceedings of the 9th annual international conference on Mobile computing and networking*, pages 134–146, New York, NY, USA, 2003. ACM. 9

[7] Richard Draves, Jitendra Padhye, and Brian Zill. Routing in multi-radio, multi-hop wireless mesh networks. In *MobiCom '04: Proceedings of the 10th annual international conference on Mobile computing and networking*, pages 114–128, New York, NY, USA, 2004. ACM. 9, 43

[8] P. Gupta and P.R. Kumar. The capacity of wireless networks. *Information Theory, IEEE Transactions on*, 46(2):388–404, Mar 2000. 3, 7, 8

[9] Z.J. Haas. A new routing protocol for the reconfigurable wireless networks. In *Universal Personal Communications Record, 1997. Conference Record., 1997 IEEE 6th International Conference on*, volume 2, pages 562 –566 vol.2, oct 1997. 9

[10] Bing He, Bin Xie, and D.P. Agrawal. Optimizing the internet gateway deployment in a wireless mesh network. In *Mobile Adhoc and Sensor Systems, 2007. MASS 2007. IEEE Internatonal Conference on*, pages 1 –9, oct. 2007. 4

[11] Kamal Jain, Jitendra Padhye, Venkata N. Padmanabhan, and Lili Qiu. Impact of interference on multi-hop wireless network performance. *Wirel. Netw.*, 11:471–487, July 2005. 3

[12] Jangeun Jun and M.L. Sichitiu. The nominal capacity of wireless mesh networks. *Wireless Communications, IEEE*, 10(5):8–14, Oct 2003. 3, 7, 8, 12

[13] Sachin Katti, Dina Katabi, Hari Balakrishnan, and Muriel Medard. Symbol-level network coding for wireless mesh networks. In *ACM SIGCOMM*, Seattle, WA, August 2008. 43

[14] Sachin Katti, Hariharan Rahul, Wenjun Hu, Dina Katabi, Muriel Médard, and Jon Crowcroft. Xors in the air: practical wireless network coding. *SIGCOMM Comput. Commun. Rev.*, 36(4):243–254, 2006. 43

[15] Ulaş C. Kozat and Leandros Tassiulas. Throughput capacity of random ad hoc networks with infrastructure support. In *MobiCom '03: Proceedings of the 9th annual international conference on Mobile computing and networking*, pages 55–65, New York, NY, USA, 2003. ACM. 7, 9

[16] Azman Osman Lim, Xudong Wang, Youiti Kado, and Zhang Bing. A hybrid centralized routing protocol for 802.11s wmns. *Mobile Networks and Applications*, 13:117–131, April 2008. 9

[17] B. Liu, Z. Liu, and D. Towsley. On the capacity of hybrid wireless networks. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE*, volume 2, pages 1543–1552 vol.2, March-3 April 2003. 7, 9

[18] M. Marina and S. Das. On-demand multi path distance vector routing in ad hoc networks. *Network Protocols, IEEE International Conference on*, 0:0014, 2001. 9

[19] McCanne and S. Floyd. Network simulator. http://www.isi.edu/nsnam/ns/. 20

[20] Deepti Nandiraju, Lakshmi Santhanam, Nagesh Nandiraju, and Dharma P. Agrawal. Achieving load balancing in wireless mesh networks through multiple gateways. In *Mobile Adhoc and Sensor Systems (MASS), 2006 IEEE International Conference on*, pages 807 –812, oct. 2006. 4

[21] IEEE P802.11s/D1.06. Draft amendment to standard ieee 802.11: Mesh networking. IEEE, May 2007. 27

[22] Charles E. Perkins and Pravin Bhagwat. Highly dynamic destination-sequenced distance-vector routing (dsdv) for mobile computers. In *SIGCOMM '94: Proceedings of the conference on Communications architectures, protocols and applications*, pages 234–244, New York, NY, USA, 1994. ACM. 9

[23] Sylvia Ratnasamy, Paul Francis, Mark Handley, Richard Karp, and Scott Schenker. A scalable content-addressable network. In *SIGCOMM '01: Proceedings of the 2001 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 161–172, New York, NY, USA, 2001. ACM. 3

[24] E.M. Royer and C.E. Perkins. An implementation study of the aodv routing protocol. In *Wireless Communications and Networking Conference, 2000. WCNC. 2000 IEEE*, volume 3, pages 1003 –1008 vol.3, 2000. 9

[25] Shigeki Goto Tatsuya Mori, Masato Uchida. Flow analysis of internet traffic: World wide web versus peer-to-peer. *Systems and Computers in Japan*, 36(11):70–81, 2005. 37

[26] K. Weniger. Pacman: passive autoconfiguration for mobile ad hoc networks. *Selected Areas in Communications, IEEE Journal on*, 23(3):507 – 519, march 2005. 10

[27] Xiaobing Wu, Jiangchuan Liu, and Guihai Chen. Analysis of bottleneck delay and throughput in wireless mesh networks. In *Mobile Adhoc and Sensor Systems (MASS), 2006 IEEE International Conference on*, pages 765 –770, oct. 2006. 1

[28] Jun Zhang and Xiaohua Jia. Capacity analysis of wireless mesh networks with omni or directional antennas. In *INFOCOM 2009, IEEE*, pages 2881 –2885, april 2009. 7, 8

[29] Jing Zhao, Ping Zhang, and Guohong Cao. On cooperative caching in wireless p2p networks. In *Distributed Computing Systems, 2008. ICDCS '08. The 28th International Conference on*, pages 731 –739, june 2008. 3

[30] A. Zimmermann, V. Bohme, and A. Hannemann. IP address assignment in wireless mesh networks. In *Local Computer Networks, 2008. LCN 2008. 33rd IEEE Conference on*, pages 44 –51, oct. 2008. 9, 38