Title	Optimal latency balancing algorithm for multiple portal in wireless mesh networks
Author(s)	Lim, Azman Osman; Chen, Zuan; Tan, Yasuo
Citation	2013 2nd IEEE/CIC International Conference on Communications in China (ICCC): 252-258
Issue Date	2013-08
Type	Conference Paper
Text version	author
URL	http://hdl.handle.net/10119/11611
Rights	This is the author's version of the work. Copyright (C) 2013 IEEE. 2013 2nd IEEE/CIC International Conference on Communications in China (ICCC), 2013, 252-258. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
Description	

Optimal Latency Balancing Algorithm for Multiple Portal in Wireless Mesh Networks

(Invited Paper)

Azman Osman Lim, Zuan Chen, and Yasuo Tan
School of Information Science
Japan Advanced Institute of Science and Technology (JAIST)
1-1 Asahidai, Nomi City, Ishikawa 923-1292, JAPAN
Email: {aolim, chenzuan, ytan}@jaist.ac.jp

***Abstract*—Wireless mesh networks (WMNs) are the key technology of wireless access network for wireless Internet application. WMNs exploit mesh gateways for delivering traffic to or from the end-users with a multihop wireless backbone. Because the traffic volume is expected too high for single mesh gateway, the adoption of multiple mesh gateways is very essential. Hence, a traffic balancing algorithm becomes vital to be incorporated into the routing protocol that used in the WMNs. In this paper, we propose an optimal latency balancing (OLB) algorithm, which is used to balance the traffic load among the mesh gateways and to improve the latency of packet sending. The proposed OLB algorithm effectively finds the potential traffic flows to be switchable until the inter-domain traffic volume is balanced. Simulation results reveal that the performance of routing protocol with the proposed OLB algorithm is improved in terms of latency, hop count, network throughput, and packet delivery ratio.**

I. INTRODUCTION

Comprehensive wireless mesh networks (WMNs) architecture that builds within the MAC layer of IEEE802.11 devices is specified in the IEEE802.11s standard [1]. This specified WMNs standard has frequently emerged in our society and been discussed as an excellent one of the next generation wireless technology. WMNs are a multihop IEEE802.11 wireless local area network (WLAN) which provides low-cost and convenient deployments to the users and possesses self-organization and self-configuration features. In IEEE802.11s standard, several aspects of the basic specification are introduced. For example, (i) it specifies new frame formats and information elements; (ii) it presents the path selection and forwarding procedures; (iii) it supports non-mesh stations by means of proxy operations; (iv) it adopts a new security framework to the mesh architecture; and (v) it states the peer node discovery and the management of the established link.

Generally, WMNs are comprised of mesh routers and mesh stations [2]. Mesh routers are mostly static (or quasi-static) in nature and are interconnected by wireless links. Mesh routers serve as an infrastructure wireless backbone, providing connectivity to mesh stations. Typically, mesh routers have direct connectivity to a fixed infrastructure through a wired-connected mesh router, which is also called *mesh gateway*. One the other hand, WMNs are mainly targeted for residential, office, public safety, military, campus, community, small to medium businesses, public access, emergency, municipality, and rural networks. Most of the aforementioned networks needs to connect to the infrastructure networks (e.g., a wired network for Internet applications). As a result, most of applications that is using WMNs technology contains high traffic volume and needs to be directed to or from the mesh gateway. Thus, traffic concentration occurs at the mesh gateway, which is noticeably overloaded with forwarding traffic. When the network size of WMNs grows and becomes significantly huge, the traffic volume around the mesh gateway becomes so heavy that the overall network performance can be degraded sharply and the entire network can slow down or even stall. One of the feasible solution is to employ multiple mesh gateways. When the WMNs operate multiple mesh gateways, one vital consideration is how to associate the mesh routers with a particular mesh gateway. In this paper, we refer to the set of mesh routers served by a mesh gateway as its *domain*. Because the traffic volume in the WMNs is expected to be very huge, performing traffic balancing in the multiple mesh gateways is important. Moreover, a routing protocol that is specified in the WMNs can only provide the *nearest mesh gateway* solution for each mesh router to its nearest mesh gateway by using the shortest hop count metric.

In this paper, we address the aforementioned problem by proposing a novel algorithm, called optimal latency balancing (OLB) algorithm, which enables to perform traffic balancing among the multiple mesh gateways. In this paper, our aim is to propose and evaluate the proposed OLB algorithm over the Hybrid Wireless Mesh Protocol (HWMP) that is specified in the IEEE802.11s for the WMNs environment with multiple mesh gateways. Our contribution is divided into three folds. First, we model the traffic balancing problem based on the latency metric and define this problem can be optimized by our proposed OLB algorithm. Second, we use the depth-first search (DFS) method to fasten the domain weight computation at the mesh gateway with less memory. Third, we examine the performance of the OLB algorithm for the multiple mesh gateways in the WMNs environment, that specified in the IEEE802.11s standard. Our proposed OLB algorithm executes periodically and adapts to the dynamic network change. It can use to balance the traffic load among the mesh gateways and improve the latency of packet sending. It also effectively can find the potential traffic flows to be switchable until the inter-domain traffic volume is balanced. As a result of applying the OLB algorithm, latency, network throughput and packet delivery ratio are improved.

The remainder of this paper is organized as follows. In Section II, we review related work. The research background of WMN, HWMP, airtime link metric are given in Section III. Section IV introduces preliminaries and problem formulation. The proposed OLB algorithm and its operation are described in Section V. The simulation studies are presented in Section VI. Finally, some concluding remarks are drawn in Section VII.

II. RELATED WORK

A few research works in literature are concerning the problem of traffic balancing for multiple mesh gateways in the WMNs. To deal with the issue of multiple mesh gateways, the routing protocol should be able to handle the direction of the traffic volume that is going through the multiple mesh gateways. At the mesh gateway viewpoint, we define an uplink traffic is the traffic that is flowing outside the WMNs. Whereas we represent a downlink traffic is the traffic that is flowing inside the WMNs. Most of the research contributions focusing on the load balancing for multiple mesh gateways in the aspect of uplink traffic, in which the mesh routers and mesh stations have to find the serving mesh gateway to send out their traffic and the serving mesh gateway is tried to balance the traffic load based on two aspects; routing metric and balancing algorithm.

Past work has focused on topological load-independent metrics and topological load-dependent metrics. Examples of the topological load-independent metrics are hop count, bandwidth, expected transmission count (ETX) that exploits the total number of transmissions needed to transmit a packet based on packet loss [3], and both expected transmission time (ETT) and weighted cumulative expected transmission time (WCETT) that are measured by the size of the packet times the data rate [4]. Meanwhile, examples of the topological load-dependent metrics are metric of interference and channel-switching (MIC) [5] and load-aware expected transmission time (LAETT) [6]. MIC composes of two metrics. One is interference-aware resource usage (IRU), which captures inter-flow interference, and another one is channel switching cost (CSC), which captures intra-flow interference. However, LAETT captures both load traffic and link quality.

The problem with routing metric for load balancing is that not all the flows is well-balanced among the serving mesh gateways. This creates an inter-domain unfairness problem. To deal with this problem, the balancing algorithm is intensively studied these days to ensure each mesh gateway gets a fair share of handling the traffic volume that is directed to or from the WMNs. Examples of research works are as follow. Tao et al. [7] show how traffic balancing (TB) method solves the uneven traffic load problem appearing in WMNs. With the right placement of the mesh gateway, TB method can improve the performance of routing protocol dramatically. Maurina et al. [8] introduce a tree-based with multi-gateway association (TBMGA), a novel routing protocol that elegantly balances the load among the different Internet gateways in the WMNs. TBMGA combines the flexibility of layer-2 routing with the self-configuring and self-healing capabilities of mobile ad hoc network routing. The selection of gateway is depending on a global metric estimated based on the average queue length and expected availability at the Internet gateway. Galvez et al. [9] study a feedback-based adaptive online algorithm for multi-gateway load-balancing in the WMNs. The algorithm is called gateway load-balancer (GWLB), which executes periodically, adapting to network conditions. GWLB takes into account the elastic nature of TCP flows, as well as flow interactions when switching nodes between domains, in order to prevent severe interference. Ashraf et al. [10] propose expected link performance-gateway selection (ELP-GS), which is used to handle gateway and route selection for gateway-oriented traffic. Along with ELP-GS, they also propose a gateway discovery protocol to help evaluate the extended ELP metric at finding the ‘best’ available route between mesh routers for peer-to-peer traffic. Owing to selection of ‘best’ routes to the gateway along with the least loaded gateway, overall ELP-GS provides better throughput, lower end-to-end delay and lower delay-deviation.

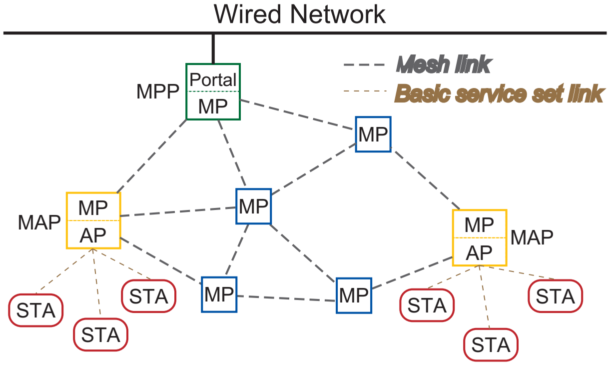

Fig. 1. Network architecture of IEEE802.11s standard.

III. RESEARCH BACKGROUND

A. WMN Architecture of IEEE802.11s Standard

Before introducing the network architecture of IEEE802.11s standard, we first explain the original IEEE802.11 WLAN architecture as described in [11]. The basic service set (BSS) is the basic block of an IEEE802.11. Each of BSS has some wireless stations (STAs) as members. The most basic type of IEEE802.11 is the independent BSS (IBSS). In IBSS, STAs directly communicate with each other without connecting to access point. This type of network is often referred to as an ad hoc network. An infrastructure BSS also forms a component of IEEE802.11. When a STA acts as an access point, it enables access to the distribution system (DS) which handles address to destination mapping and seamless integration of multiple BSSs. The DS and infrastructure BSSs allows IEEE802.11 to create a wireless network of arbitrary size, which is called an extended service set (ESS) networks. An ESS is the union of the infrastructure BSSs connected by a DS.

To address the need of wireless mesh networking in IEEE802.11, a mesh STA is not a member of an IBSS or an infrastructure BSS. Consequently, mesh STAs do not communicate with non-mesh STAs. However instead of existing independently, a mesh BSS (MBSS) also accesses the DS. The MBSS interconnects with other BSSs through the DS. Then, mesh STAs can communicate with non-mesh STAs. Such an enhanced architecture is known as a WMN. This WMN architecture with mesh stations (MSs) is specified in the IEEE 802.11s [1]. In the WMN architecture, different entities play different roles according to the functionalities they provide as illustrated in Fig. 1. Basically, the WMN architecture consists of three entities; a mesh point (MP), a mesh portal point (MPP), and a mesh access point (MAP). The ad hoc link formation formed by the MPs, which provides the backbone for the WMN infrastructure, whereas the MPP works as a repeater or a gateway. The MP that supports the associated MSs is usually called as MAP. The MPs can be either stationary or mobile. However, the MPP and MAPs are mostly immobile. MSs are usually regular devices that do not contribute to the WMN services, such as routing and forwarding for multi-hopping packets. Therefore, MSs simply connect to one of the MAPs in order to access the network resources. In this paper, we use the terms of MPP and MAP to represent mesh gateway and mesh router, respectively.

B. Hybrid Wireless Mesh Protocol

A path selection protocol, Hybrid Wireless Mesh Protocol (HWMP) is specified in IEEE802.11s mesh networking [1]. In HWMP, on-demand routing protocol is adopted for MPPs or MAPs or MPs that experience a changing environment, while proactive tree-based routing protocol is an efficient choice for MPPs or MAPs or MPs in a fixed network topology. The on-demand routing protocol is specified based on radio-metric ad hoc on-demand distance vector (RM-AODV) routing. The basic features of AODV are adopted, but extensions are made for IEEE802.11s. However, the proactive tree-based routing is applied when a MPP is configured in the WMNs. With this MPP, a distance vector tree can be built and maintained for MAPs/MPs, which can avoid unnecessary routing overhead for routing path discovery and recovery. It should be noted that the on-demand routing and tree-based routing can run simultaneously. Four information elements are specified for HWMP: root announcement (RANN), path request (PREQ), path reply (PREP), and path error (PERR). Except for PERR, all other information elements of HWMP contain three important fields: destination sequence number, time-to-live, and metric.

Before a MAP or a MP could send its traffic to another MAP or MP inside the WMNs, a path selection protocol either on-demand mode or proactive tree-based mode can be used. In the path selection on-demand mode, source S wanting to send data to destination D broadcasts a PREQ frame indicating the MAC address of D. All the MAPs or the MPs receiving the PREQ create or update their path to S, but only if the PREQ contains a sequence number greater than the current path or the same sequence number and a better metric. Every MAPs or MPs, before re-broadcasting the PREQ, must update the metric field to reflect the cumulative metric of the path to S. Once D receives the PREQ, it sends S a unicast PREP. If D receives further PREQs with a better metric (and the same or greater sequence number), it sends a new PREP along the updated path. Intermediate MAPs or MPs shall then forward the PREP(s) to S along the best path (stored during the PREQ flooding phase), and, when the PREP reaches S, the path is set up and can be used for a bi-directional exchange of data. If more than one PREP is received, the PREPs following the first are processed only if their information is not stale and announces a better metric (the same rules of PREQ apply). Note that the metric values carried by PREQ and PREP frames refer to two different paths: PREQs measure the reverse path, i.e., from D to S, whereas PREPs measure the forward $(S-D)$ path. This is because the value inserted by each MAP or MP refers to the metric it measures towards the MAP or the MP from which it received the PREQ or PREP. Hence, depending on the metric computation strategy, it may occur that the forward and reverse paths do not coincide.

One the other hand, the topology tree can also be constructed in order to link all the participating MAPs or MPs using a path selection proactive tree-based mode. This topology tree formation begins when the MPP starts to periodically broadcast RANN message by increasing the sequence number in every announcement. Upon receiving the RANN message, the MAP or the MP caches the originated the MAP or the MP address of the corresponding RANN as a potential parent, and rebroadcasts the RANN with an updated cumulative metric. There are two fundamental approaches for a child MAP or a child MP to select its parent. First, after waiting for a pre-defined time of a few seconds for other arriving RANNs from all possible parents, the child MAP or the child MP selects one parent with the best-metric for its path to the MPP from all the possible parents. Alternatively, the second approach is that the child MAP or the child MP does not wait for the pre-defined time for other arriving RANNs from all possible parents, whereas the child MAP or the child MP immediately selects the corresponding parent that sent the first received RANN message. After selecting the parent, the child MAP or the child MP updates its route table in which, for instance, the latest message sequence number. Then, the child MAP or the child MP that has the known path to the MPP also registers itself by sending a PREP message with the MPP as the destination address in the PREP message field. Each intermediate MAP or MP that received the PREP forwards the PREP to its selected parent and updates the MAP or the MP it was received from as the next-hop child to reach the source MAP or MP in its route table. After receiving all the PREPs, the MPP learns all participating MAPs or MPs and builds a tree topology to reach any MAPs or MPs in the WMNs. If a MAP or a MP does not hear the RANN for a pre-defined period, the MAP or the MP does not participate in tree-building until hearing a valid RANN again.

Each MAP or MP in the tree-topology network maintains its own route table, which has entries for recent route towards the destination MAP or MP. In the route table of each MAP or MP, the contents are destination, next hop, link metric, sequence number, time stamp, and node flag. The field of link metric represents that the metric that is associated with the airtime link metric. The field of sequence number represents the most recent information of an entry. The field of time stamp represents that the time for an entry is stored and it is used to monitor the expiration of an entry. Each time the tree route is used, its associated time stamp is updated. If the route is not used within the specified time, route table timeout must be at least the maximum of three times of $RANN\ announcement\ interval$, it is deleted. The field of node flag represents that the destination of entry is either a MPP or a MAP or a MP.

C. Airtime Link Metric

A link metric is a measured unit to a link and a path metric is a value which is assigned to a path, combining by all the link metrics in the path, used by the routing algorithm to select the optimized routes for a specified objective. The optimization objectives can be minimizing delay, minimizing energy consumption, maximizing throughput, etc. The path metric is the combination of the link metrics in the whole path and the method for combining the link metric and the path metric can be defined in varies ways which are depended by the actual situations. The usual functions are summation, multiplication and statistical measures (e.g., minimum, maximum, average, etc). Airtime link metric (ALM) is a default link metric for path selection routing protocol in the IEEE802.11s mesh networking. The extensibility framework allows this metric to be overridden by any path metric as specified in the mesh profile. Airtime reflects the amount of channel resources consumed by transmitting the frame over a particular link. This measure is approximate and designed for ease of implementation and interoperability. The airtime for each link is calculated as follows.

$$c_a = (O + \frac{B_t}{r})\frac{1}{1-e_f} \tag{1}$$

where O is channel access overhead, which includes frame headers, training sequences, access protocol frame, etc. B_t is number of bits in test frame. r is the transmission rate while e_f represents the frame error rate. Given the IEEE802.11a hardware specification, example value of O and B_t is $185\,\mu s$ and 8192 bits, respectively. This ALM is very similar to the ETT metric [4] and essentially captures physical interference. In particular, the ALM is a newly specified metric that tries to capture the quality of the links as a function of the estimated frame loss probability.

IV. PRELIMINARIES

A. Network Model and Assumptions

A MAP is a MP that works as an access point to receive or send the Internet traffic for the associated MSs. For the sake of simplicity, the abbreviations MAP and MP refer to the 'MAP' and are indistinguishably interchangeable throughout this paper. Let $\mathcal{N}$ be the set of MAPs of the WMNs, we assume that the active MAPs, $n = \{1, 2, \ldots, \mathcal{N}\} \in \mathcal{N}$ are distributed in a two-dimensional plane and have equal communication range. The network topology is modeled as a connected graph, $G(\mathcal{N}, E)$. An edge $\{u, v\} \in E$ if the only if the distance between u and v is within each other's communication range. The latency between two adjacent u and v can be defined as $\lambda(u, v)$. In this paper, we assume that $\lambda(u, v) = \lambda(v, u)$. The minimum latency between any source-destination in the network is the λ-cost of the shortest path connecting them that defined with the ALM metric. The Dijkstra algorithm can be utilized to construct the shortest path rooted by the source. Routes inside the WMNs are given by the proactive tree-based path selection protocol.

Since a MS is logically and wirelessly connected to a MAP, a user through the MS initiates connections to the Internet, which can generate downlink flows or uplink flows in between MAP and MPP. Let $\mathcal{G}$ be the set of MPPs in the WMNs. Let $\mathcal{F}$ be the set of downlink flows or uplink flows. A flow can be routed through different paths from each MAP to the MPP and vice versa. Let $f = \{1, 2, \ldots, \mathcal{F}\} \in \mathcal{F}$ is the set of active flows. Let $\mathcal{D}_g$ be the set of domains that containing the set of active flows served by the set of active MPPs, $g = \{1, 2, \ldots, \mathcal{G}\} \in \mathcal{G}$. Each domain $\mathcal{D}$ can generate a domain weight, called $\mathcal{W}$. Let $\mathcal{W}_g$ be the set of domain weights that containing the sum of the path routing weights for each source-destination paths of the corresponding domain $\mathcal{D}_g$. Therefore, the domain weight can be written as

$$\mathcal{W}_g = \sum_{i \in \mathcal{F}; i=1}^{f_g} w_i \tag{2}$$

This path routing weight (w_i) is decided by using the routing metric. In this paper, the path routing weight is decided by the latency metric that defined with the ALM metric. Thus, the path routing weight of the ith flow can be expressed as

$$w_i = \lambda_i(S, D) \tag{3}$$

where S denotes source and D denotes destination. If we consider two active MPPs and fifty active MAPs in the WMNs, then g is equal to 2 and n is equal to 50. If all the MAPs have a flow to the MPPs, then f is equal to 50 flows. The proactive tree-based path selection protocol will divide all the flows into two domains served by two MPPs, respectively. This also means that $\mathcal{D}_1$ and $\mathcal{D}_2$ contain $\mathcal{W}_1$ and $\mathcal{W}_2$, respectively.

B. Problem Statement

Internet traffic directed to or from a MAP will be served by a MPP; the traffic is routed from the MPP to the MAP using a minimum cost path (in terms of the used path routing metric) and vice versa. The traffic-balancing problem requires to choose the right serving MPP for every downlink flow or uplink flow. This problem concerns Internet traffic at the serving MPP, because it accounts for most of the network load. One of the possible solutions for the HWMP protocol to deal with multiple portal is to use the *Nearest MPP* (NMPP) solution, which assigns each MAP to its nearest serving MPP based on the shortest cost path that obtained from ALM metric. The NMPP solution can easily lead to traffic imbalance. To deal with this problem, we first formulate that $\mathcal{P}$ is the traffic balancing problem for multiple portal in the WMNs. $\mathcal{P}$ is also an optimization problem. A solution of $\mathcal{P}$ must choose the serving MPPs can be balanced with the most minimum value of $\mathcal{C}$ variable. Let $\mathcal{C} = min\{diff(\mathcal{W}_1, \mathcal{W}_2, \ldots, \mathcal{W}_\mathcal{G})\}$ be the achievable value. Given the above considerations, the problem can be formulated as:

$$\mathcal{P} : minimize \quad diff(\mathcal{W}_1, \mathcal{W}_2, \ldots, \mathcal{W}_\mathcal{G}) \tag{4}$$

$$subject\ to \quad \mathcal{C} \geq 0 \tag{5}$$

The $\mathcal{P}$ is an NP-hard problem so any deterministic algorithm cannot solve this problem in polynomial time. Therefore, an exact algorithm for solving this problem can be only used for small scale networks because the execution time of such kind of algorithms exponentially increases with the dimensions. For large scale networks (i.e., about hundreds or thousands of vertices), the only way to achieve quality results is to use some heuristic methods. The next section describes our proposed depth-first search algorithm for solving this problem.

V. OPTIMAL LATENCY BALANCING ALGORITHM

Most of the cases, WMNs are not used independently. WMNs are designed to access to other types of wired networks. For such purpose, the function of MPP that acts as a portal is very essential. In this paper, we focus on proposing a feasible path routing metric for the HWMP and a traffic balancing algorithm for the multiple portal issue. Since the ALM that specified in IEEE802.11s reflects the amount of channel resources consumed by transmitting the packet over a particular link. When the transmitting packet is failed due to the interference effect or the packet collision, the ALM captures this measurement as a function of the frame loss probability. In other words, the ALM metric is indirectly taken the physical interference into account. In this paper, we take the ALM metric and transform it in a generic path routing metric, i.e., a latency. The latency for a wireless network is defined as the time of the start of packet transmission at the source node to the time of the end of packet reception at the destination node. This latency does not comprise of the time to process the packet at both source and destination nodes. It is important to distinguish between *network latency* and *latency* to avoid confusion. Since the network latency can be measured either one-way (the time from the source sending a packet to the receiving destination) or round-trip (the sum of the one-way latency from source to destination and the one-way latency from destination to source), we interpret *latency* metric is similar to the one-way *network latency*.

Algorithm 1 Optimal Latency Balancing Algorithm

```
01: Definition: $S$ is source, $D$ is destination.
02: Input: all the pairs of $S$ and $D$, the path routing weight $w_{SD}$.
03: Output: an optimum value of $\mathcal{C}$ with the switch flows.
04: Begin
05:     $C$ = 0; // $C$ is a variable of optimization value.
06:     Set g = $\mathcal{G}$ parameter // $g$ is the number of serving MPPs in a WMN.
07:     Compute $\mathcal{W}_g$ and $\mathcal{C}$;
08:     $C$ = $\mathcal{C}$;
09:     Set $W_{max} = max\{\mathcal{W}_g\}$;
10:     Set $W_{min} = min\{\mathcal{W}_g\}$;
11:     Set $w_{SD} = argmax\{W_{max}\}$; // Set the highest path routing weight.
12:     Choose $S$ from $w_{SD}$;
13:     $S$ is marked as visited; disconnected_flag = 0;
14:     while All the MAPs of $\mathcal{D}_{max}$ is not visited do {
15:         $A$ = $S$; // $A$ is the next MAP.
16:         while $D$ is not reached && disconnected_flag = 0 do {
17:             if $A$ = $S$ // $S$ has no unvisited neighbors
18:                 disconnected_flag = 1; // $S$ and $D$ are disconnected
19:             if $A \neq S$ && $A$ has no unvisited neighbors
20:                 return to itself; $A$ = sender to $A$;
21:             if $A$ has unvisited neighbors
22:                 $A$ sorts all unvisited neighbors using ALM criterion;
23:                 $A$ sends the packet to the first neighbor $B$ in the list;
24:                 $B$ memorizes $A$ as the sender;
25:                 $B$ is marked as visited;
26:                 $A$ = $B$;
27:         }
28:     }
29:     if disconnected_flag = 1
30:         Go to the line 11 with the next highest path routing weight;
31:     else
32:         Compute $\mathcal{W}_g^{'}$ and $\mathcal{C}^{'}$;
33:         if $\mathcal{W}_g^{'} \leq \mathcal{W}_g$ && $\mathcal{C}^{'} \leq C$
34:             Switch the flow of $S$ to the $\mathcal{D}_{min}$ domain;
35:             $C$ = $\mathcal{C}^{'}$;
36:             Go to the line 9 for the next domain weight;
37:         else
38:             Break;
39: End
```

A. *Proposed Traffic Balancing Algorithm*

We propose an optimal latency balancing (OLB) algorithm, dynamically computes the domain weight of MPPs and balances the traffic that are directed to or from a wired network through the serving MPPs. Because MSs are randomly accessed to the MAP and can generate connections at any time, network conditions can constantly change. It is therefore important that the algorithm adapts to the current conditions, and the solution must converge, while avoiding route flapping. The proposed OLB is an adaptive algorithm, which continually monitors the network conditions of WMNs. Based on the current conditions and the knowledge of the network topology, the proposed OLB calculates an optimum solution $\mathcal{P}$. The OLB algorithm is self-corrective and executes periodically to adapt to current conditions. The basic idea of the OLB algorithm is to balance the traffic load among the serving MPPs in order to increase inter-domain flow fairness and minimize the latency of packet sending.

Through the HWMP protocol with the ALM criterion, the traffic of the MAPs can be transfered to or from the serving MPPs with the shortest cost path. The OLB algorithm starts with the path routing weight computation for each active flows of the WMNs. Then, the OLB algorithm will first initialize a solution of $\mathcal{P}$ is equal to $\mathcal{C}_{initial}$ based on the ALM criterion. To guarantee an optimal solution of $\mathcal{P}$, it is impossible to evaluate every possible solution due to time and memory limitations. Therefore, a DFS method is used to find the feasible solution quickly so that the OLB algorithm can compare the feasible solution with initial solution. During the iteration process, the DFS method looks for possible feasible solution and increase the balance of the domain weights by finding the optimal solution of $\mathcal{P} = \mathcal{C}_{optimal}$ with the condition that the total domain weight of the optimal solution $= \mathcal{C}_{optimal}$ must be smaller than the total domain weight of the previous solution $= \mathcal{C}_{previous}$.

The general idea of DFS method is as follow. Each MAP memorizes if it has already been *visited* by the DFS traversal, and the *sender* from where the traffic flow was received for the first time. The MAP that is currently holding the routing table will sort all its *unvisited* neighbors by the ALM criterion. The first MAP in the list of routing table is selected to transport the traffic flow. If a MAP has no *unvisited* neighbors to proceed, it returns to itself as the *sender*, which will transport the traffic flow to the next *unvisited* neighbor in the list of routing table. Eventually, the traffic flow either reaches the destination, or returns back to the source, which has no *unvisited* neighbors. The OLB algorithm with DFS that integrates latency metric based on the ALM criterion operates as presented in **A**lgorithm 1. The OLB algorithm executes every fixed period of time, which should be greater than a few times of PREQ broadcast interval time. To improve the inter-domain flow fairness, the OLB algorithm will attempt to switch MPPs from domains $\mathcal{D}_1$ with an above average number of flows to domains $\mathcal{D}_2$ with a below average number of flows. If switching from $\mathcal{D}_1$ to $\mathcal{D}_2$ is not suspended, the OLB algorithm progressively finds other flows to be switchable until the optimal solution of $\mathcal{P}$ is achieved. To avoid the same flow switch back to the original domain, the OLB algorithm can use the 'scavenge stale resource record' to prevent oscillations.

VI. PERFORMANCE EVALUATION

We evaluate and study the performance of the proposed OLB algorithm over the HWMP protocol in the WMN environment with multiple portals and a varying number of flows. We use C++ console application to construct our simulator

TABLE I. SIMULATION PARAMETERS AND SETTINGS

Hardware specification	IEEE 802.11a
MAC protocol	CSMA/CA
Maximum contention window	1023
Minimum contention window	15
PLCP preamble length	20 μs
PLCP header length	4 μs
Slot time	9 μs
SIFS	16 μs
DIFS	34 μs
Network protocol	HWMP
RANN size	28 bytes
PREP size	24 bytes
RANN interval	15 s
Number of MPPs	2
Network coverage	500 m × 500 m
Energy model	Two-ray
Propagation loss exponent	3.5
Transmit power	100 mW
Pattern of MAP	static
Transmission bit rate	54 Mbps
Queue size	100 packets
Traffic type	UDP
Average traffic interval	100 ms
Data payload size	1000 bytes
Data header size	24 bytes

program. This program is event-driven application. All the events are defined in the configuration file. In the program, we simulate each MAP as an independent object with its own properties. In the startup, the application first reads the configuration file to initialize the parameters and generate all the events. When an event meets its time, the generated packets will be forwarded to the specified MAP to process event. We compare our proposed OLB algorithm with the NMPP solution.

A. *Simulation Environment and Setup*

The IEEE802.11a hardware specification with CSMA/CA protocol is used to investigate the performance of proposed OLB algorithm and the NMPP solution over the HWMP protocol in the WMNs environment. The parameter types and values are shown in Table I. In our simulation, two serving MPPs are considered as a preliminary stage of performance evaluation studies. One MPP is located at (250 m, 0) and another one is located at (0, 250 m) whereas other MAPs are randomly distributed over 500 m × 500 m field. We assume that all the MAPs is identical. This means MAPs have same transmit power and transmission bit rate. We assume that all the MAPs is static throughout the entire simulation. We model our traffic based on user datagram protocol (UDP). MAPs connect to the serving MPP, which generates downlink and uplink flows. For each MAPs, the time between start of connections follow an exponential distribution with ten seconds. The UDP traffic consists of 1000-byte frame size, which sends at the exponential distribution with traffic interval of 100 milliseconds. We ran our simulation for ten minutes and ten topologies with different seeds for pseudo-random number generator are averaged. In our simulations, we focus the influence of the number of flows on the performance of the OLB algorithm and the NMPP solution over the HWMP protocol. To observe the effect of flow increment, we vary number of flows from 10 to 50 flows in increment of 10 flows for both uplink and downlink flows.

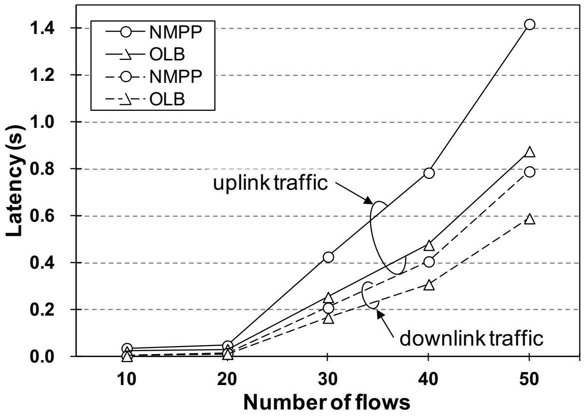

Fig. 2. Latency as a function of number of flows.

TABLE II. HOP COUNT AS A FUNCTION OF NUMBER OF FLOWS

		Number of flows				
		10	20	30	40	50
Uplink	NMPP	4.65	4.31	3.42	3.19	2.98
	OLB	4.65	4.31	3.34	3.12	2.95
Downlink	NMPP	4.53	4.21	3.34	3.11	2.91
	OLB	4.53	4.21	3.25	3.05	2.88

B. *Performance Metrics*

The simulation results are observed with four sets of performance metrics; latency, hop count, network throughput, and packet delivery ratio. Latency is the average end-to-end delay of all delivered data packets of all the flows. Network throughput is the total data bytes delivered by the network divided by the time elapsed since the first packet was sent and the last packet was received. Hop count is the average end-to-end hop count of all the source and destination pairs. Packet delivery ratio is the ratio of received data packets to those transmitted by the source.

C. *Simulation Results and Discussion*

It is essential to note that each value shown in the graph is the average value taken from the number of different topologies. Fig. 2 shows the performance for the latency versus the number of flows. We can obtain that OLB outperforms NMPP significantly as the number of flows increases. The reason for this behavior is the possibility of switching flow is getting increase as the number of flows increases. OLB can decrease the latency by average of 37.3% and 19.8% for the uplink traffic and downlink traffic, respectively. This means that the path of each flow becomes shorter, leading to decreased contention and interference in the whole network. We can observe this phenomena with the hop count metric as depicted in Table II.

Fig. 3 shows the performance for the network throughput versus the number of flows. With a small number of flows there is no contention and so the improvement achieved by OLB is small. As the number of flows goes up, the flow throughput decreases for NMPP but by OLB is able to improve flow throughput and therefore the network throughput is increased. OLB is noticeably achieved network throughput improvement for the uplink traffic and the downlink traffic are about 10.0%

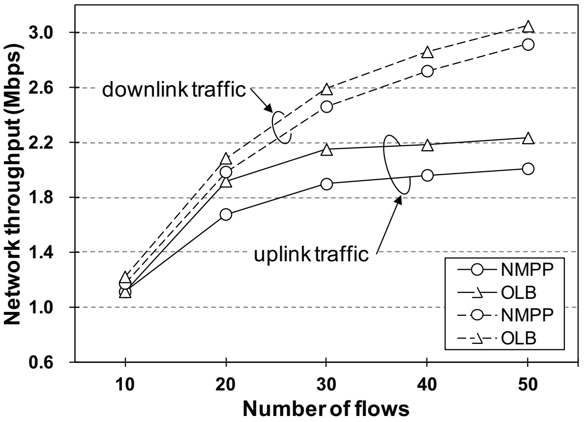

Fig. 3. Network throughput as a function of number of flows.

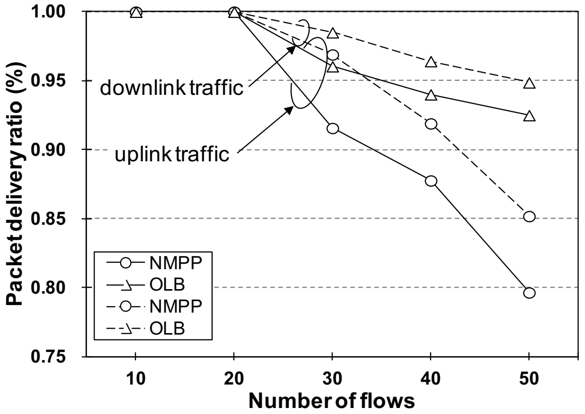

Fig. 4. Packet delivery ratio as a function of number of flows.

and 4.8%, respectively although the entire network throughput for the downlink traffic is higher than the uplink traffic. This is because OLB performs the same amount the switching flows to the domains with lower average number of flows for both uplink traffic and downlink traffic, leading to the improvement ratio is higher for uplink traffic when the network throughput of the uplink traffic is low.

As the number of flows increases, contention in the network increases, which leads to packet drops. The performance for packet delivery ratio is shown in Fig. 4. In general, NMPP drops faster. OLB performs slightly better than NMPP. The packet delivery ratio drops slightly with OLB, which indicates that it can effectively balance the entire inter-domain traffic fairness. From different viewpoint, we can see that the traffic becomes imbalance when the number of flows increases. This leads to the packet delivery ratio decreases. This reason is that increasing traffic imbalance progressively produces high contention and more interference in one domains. However, OLB is less affected by this issue, because it can self-corrective and dynamic monitoring the traffic always in the balance mode.

VII. CONCLUDING REMARKS

In this paper, we addressed the advantages of multiple portal for attaining the high traffic volume that goes through single mesh gateway. We therefore proposed the OLB algorithm for the HWMP protocol to be able to handle the issue of multiple portal in the WMNs environment. The key features of the OLB algorithm are as follow. (i) it can use to balance the traffic load among the MPPs; (ii) it can use to improve the latency of packet sending; and (iii) it effectively can find the potential traffic flows to be switchable until the inter-domain traffic volume is balanced.

We also examined the performance of the OLB algorithm compared to the NMPP solution over the HWMP protocol under the multiple portal WMNs environment. Simulation results reveal that the OLB algorithm can attain a high reduction of latency by up to about 37.3% for the uplink traffic and also can accomplish an improvement in the network throughput by up to about 10.0% for the uplink traffic. In addition to that, the OLB algorithm can achieve a merely improvement in terms of hop count and packet delivery ratio. For this evaluation it can be concluded that using the OLB algorithm can always provide advantages and beneficial to the routing protocol for the WMN environment. Our future work also will focus on examining the performance effect of the proposed OLB algorithm when the MPP placement is considered and investigating the performance effect of the proposed OLB algorithm in the WMNs environment with the issues of security, mobility, reliability and quality of service provision.

REFERENCES

[1] Amendment 10: Mesh Networking, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications*, IEEE802.11s Standard, 10 September 2011.

[2] I. F. Akyildiz and X. Wang, “A survey on wireless mesh networks,” *IEEE Commun. Mag.*, vol.43, no.9, pp.S23–S30, 2005.

[3] D. De Couto, D. Aguayo, J. Bicket, and R. Morris, “A high-throughput path metric for multi-hop wireless routing,” *Proc. of the Int. Conf. on Mobile Comput. and Netw. (MobiCom)*, pp.134–146, 2003.

[4] R. Draves, J. Padhye, and B. Zill, “Routing in multi-radio, multi-hop wireless mesh network,” *Proc. of the Int. Conf. on Mobile Comput. and Netw. (MobiCom)*, pp.114–128, 2004.

[5] Y. Yang, J. Wang, and R. Kravets, “Load-balanced routing for mesh networks,” *ACM SIGMOBILE Mobile Comput. and Commun. Review*, vol.10, no.4, pp.3–5, 2006.

[6] H. Aiache, V. Conan, L. Lebrun, and S. Rousseau, “A load dependent metric for balancing Internet traffic in wireless mesh networks,” *Proc. of the Int. Conf. on Mobile Ad Hoc and Sensor Syst. (MASS)*, pp.629–634, 2008.

[7] X. Tao, T. Kunz, and D. Falconer, “Traffic balancing in wireless mesh networks,” *Proc. of the Int. Conf. on Wireless Netw., Commun. and Mobile Comput. (WirelessCOM)*, vol.1, pp.169–174, 2005.

[8] S. Maurina, R. Riggio, T. Rasheed, and F. Granelli, “On tree-based routing in multi-gateway association based wireless mesh networks,” *Proc. of the Int. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC)*, pp.1542–1546, 2009.

[9] J.J. Galvez, P.M. Ruiz, and A.F.G. Skarmeta, “A feedback-based adaptive online algorithm for multi-gateway load-balancing in wireless mesh networks,” *Proc. of the Int. Symp. on World of Wireless Mobile and Multimedia Netw. (WoWMoM)*, pp.1–9, 2010.

[10] U. Ashraf, S. Abdellatif, and G. Juanole, “Route selection in IEEE 802.11 wireless mesh networks,” (Published Online), *Int. J. Telecommun. Syst.*, 17 June 2011.

[11] X. Wang and A.O. Lim, “IEEE 802.11s wireless mesh networks: Framework and challenges,” *Elsevier Ad Hoc Netw.*, vol.6, no.6, pp.970–984, 2008.