Copyright @ 2011 KSII

Load Balancing and Mobility Management in Multi-homed Wireless Mesh Networks

Minh Tri Tran[1], Young Han Kim[2] and Jaehwoon Lee[3]
[1, 2]School of Electronic Engineering Soongsil University
DongJak-Gu, Seoul, 156-743, Korea
[e-mail: {minhtri, yhkim}@dcn.ssu.ac.kr]
[3] Dongguk University
Pil-dong 3-ga Jung-gu, Seoul, 100-715, Korea
[e-mail : jaehwoon@dongguk.edu]
*Corresponding author: Young Han Kim

Received November 15, 2010; revised February 19, 2011; revised May 4, 2011; accepted May 11, 2011; published May 31, 2011

Abstract

Wireless mesh networks enlarge the wireless coverage area by interconnecting relatively stationary wireless routers (mesh routers). As wireless mesh networks are envisioned to provide high-bandwidth broadband Internet service to a large community of users, the Internet gateway, which acts as a central point of Internet attachment for the mesh networks, is likely to suffer heavily from the scramble for shared wireless resources because of aggregated traffic toward the Internet. It causes performance decrement on end-to-end transmissions. We propose a scheme to balance the load in a mesh network based on link quality variation to different Internet gateways. Moreover, under the mesh coverage, mobile nodes can move around and connect to nearby mesh routers while still keeping the connections to the Internet through the best gateway in terms of link quality. In this structure, gateways perform the balancing procedure through wired links. Information about gateways and mobile node's location is distributed appropriately so that every mesh router can quickly recognize the best gateway as well as the positions of mobile nodes. This distributed information assists mobile nodes to perform fast handoff. Significant benefits are shown by the performance analysis.

Keywords: Load balance, mobility management, wireless mesh networks, OLSR protocol

A preliminary version of this paper appeared in ICOIN 2010, January 26-28, Busan, Korea. This version includes a concrete analysis and supporting simulation results on multi-homed Wireless mesh networks. This research is supported by the Ministry of MKE, Korea, under the "program for CITG" support program supervised by the NIPA (NIPA-2010-0004) and the KCC (Korea Communications Commission), Korea, under the R&D program supervised by the KCA(Korea Communications Agency) (KCA-2011-08913-05001).

DOI: 10.3837/tiis.2011.05.005

1. Introduction

Wireless mesh networks (WMNs) [1] are considered as a solution to extend the coverage of wireless access networks in which mesh routers (MRs) are responsible for routing, maintaining network topology, and providing connections for a large number of mobile nodes (MNs). Some of the MRs play the role of gateways. Gateways function as portal devices that provide access to the wired Internet. Among various potential application scenarios provided by WMNs, broadband Internet access using WMNs is one of the most promising directions for applying this technology [2]. There are great advantages when the WMN is connected to the global Internet via multiple Internet gateways (IGWs), referred to as multi-homed, in terms of achieving both reliability and load balancing. When multi-homed, an end-user can still enjoy a network connection if at least one gateway is working. Furthermore, the traffic load inside the mesh network can be kept balanced among different paths to gateways to avoid congestion of certain heavily loaded paths. Furthermore, when an MN changes its associated MR, multiple multi-hop paths are available, linking to different IGWs. Among these paths, the best one will be chosen. The path selection for handoff decisions guarantees transparency to applications during and after the handoff. To ensure transparency, mobility management solutions should be designed and implemented especially for WMNs, considering the differences between WMNs and conventional wireless networks. One prominent difference is that messages, including data and signaling, have to traverse a multi-hop wireless path to reach the destination.

A comprehensive survey of mobility management in hybrid WMNs is given in [3]. In sum, existing work on mobility management in WMNs is classified according to address management for mobility support, multi-hop routing to facilitate handoffs (e.g. iMesh [4], M^3 [5]), or integrated routing and mobility management. The authors mentioned some issues to address handoff management. Among them, the gateway selection issue needs to be solved. Our proposal is concerned with this issue.

For load balancing purposes, the authors of [6] proposed a method to achieve load balancing in WMNs based on the average queue length in IGWs. Furthermore, Thomas et al [7] implemented the cross-layer metric to provide an improved load-balanced wireless mesh network with multi-homing. They showed that balancing the traffic variation to gateways can improve the throughput in wireless mesh networks. However, these papers do not focus on how to keep the ongoing transmissions continuous.

The authors of [8] showed the problems in multi-homed ad hoc networks where gateways may use different technologies, e.g. NAT-based and MIP-FA based gateways, to connect ad hoc networks to the Internet, as well as security problems such as ingress filtering. They proposed applying explicit tunneling, mobile IP with NAT-traversal, and reverse tunneling to solve these problems. As a result, the source node (MN) will add more necessary information in the sending of packets to ensure that the session is continuous while moving and to avoid the packets' ingress filtering. Our proposal is to design an architecture such that MNs are exempt from participating in mobility management. IP-in-IP encapsulation is also used but between MRs and IGWs and between IGWs.

The authors of [9] designed an architecture with multiple MANET border gateways (MBGs), but the number of MBGs working as IGWs was kept minimized, usually at one. In this environment, MNs keep IP addresses intact while moving in the MANET coverage. Data packets of the MANET are tunneled between an MN and its associated access router and between an access router and its gateway. This paper tried to shift the overhead for suppressing route reconstruction from within MANET to within the Internet. The basic architecture of my proposal is similar to that of [9]. However, [9] used a global IP address and tunneling between an MN and its associated access router. In my proposal, local IP address space is used and MRs perform tunneling on behalf of MNs. In addition, IGWs control the load balancing and choose path from MRs to GWs. By using distributed information, our proposal reduces frequent queries to IGWs where the wireless resource is scarce because of aggregated traffic toward the Internet.

In this paper, we consider a wireless mesh network with three sets of nodes: mobile nodes, mesh routers, and multiple Internet gateways. Each MN connects to the nearest MR with the strongest signal. The load of an MR is the load of all MNs connecting to this MR. The metric of a complete path does not consider the link quality from MNs to the associated MR. When an MR performs switching to another IGW, the aggregated load will be routed through the new IGW.

To achieve load balancing in our solution, the path selection from a mesh router to an IGW needs to rely on a specific metric. In this study, Optimized Link State Routing (OLSR) [10] with the minimum loss (ML) metric, i.e. OLSR-ML [11], has been selected as the routing protocol inside the WMN. Traffic is directed via links with the highest probability of successful transmission.

This paper presents an architecture in which IGWs negotiate with each other via wired links to decide which MRs they will serve; the decision is made relying on the link quality from MRs to IGWs. Information about the set of MRs served by a certain IGW along with the MNs' location, i.e. where the MR MNs reside, will be held by some of MRs instead of being stored only in IGWs. These MRs are chosen based on multipoint relay selection in the OLSR protocol so that any given MR can acquire the necessary information from them with an allowable time cost. To send packets toward the Internet, MRs access these selected MRs to determine which IGWs are the most appropriate. When the handoff happens, the time for the serving MRs to obtain the information about the former MR of the handoff MN is small; this leads to faster handoff. In addition, MNs can route packets via the best IGW after the handoff. When an MR has a better path to a new IGW, load balancing occurs. The balancing process makes the network performance increase. The balancing process, as well as the handoff procedure, does not require the participation of MNs in signaling. MRs and IGWs control the balancing and mobility management on behalf of MNs; therefore, the signaling workload in MNs is eliminated.

The remainder of this paper is organized as follows. First, the proposed solution is presented. Then, the performance analysis is carried out in section 3. Finally, concluding remarks are given.

2. Protocol Design

2.1 Architecture

OLSR-ML is an OLSR link quality extension using the minimum loss probability as a metric [11]. Multipoint-relays (MPRs) selection in OLSR-ML is based on link quality information: a neighbor is selected as an MPR if it has the best route to any 2-hop neighbor. Initially, MPR selection in OLSR-ML is the same as the MPR selection of RFC OLSR because of its lack of traffic.

The probability of the successful transmission of a link is calculated using forward and reverse link delivery ratios. The delivery ratio is the probability that a data packet successfully arrives at the next hop. The expected probability that a transmission is successfully received and acknowledged is the product of the forward delivery ratio (d_f) and the reverse delivery ratio (d_r) of a link:

$$P_{link} = d_f \times d_r$$

In a multi-hop path, the probability of successful transmission over the complete path should be the product of the probabilities of each path. The best route from one source to a specific destination is the one with the highest probability of successful transmission, i.e. the one with the minimum loss probability.

The delivery ratios are measured using modified OLSR HELLO packets sent every t seconds. Each node calculates the number of HELLOs received in a w second period and divides it by the number of HELLOs that should have been received in the same period. Each modified HELLO packet contains information about the number of HELLOs received by the neighbor during the last w seconds to allow each neighbor to calculate the reverse delivery ratio.

In the proposed solution, IGWs join a multicast group in which they exchange information with each other via wired backbone links. Initially, IGWs are assigned priorities and advertise the priority information into the multicast group to elect a primary IGW. The IGW with the highest priority will play the role of the primary IGW; the others are secondary. The primary IGW is responsible for providing IP addresses for MNs.

From the point of view of the primary IGW, the set of MPRs of rank 0 acts as the IGW itself and the set of MPRs of rank 1 acts as the MPR itself. Let us define the set of MPRs of rank k+1, for the k integer, as the union of the MPRs set of all MR elements of the MPR set of rank k. The MPRs ranked with even numbers, 2, 4, and so on, are selected as the delegated mesh routers (DRs) [10]. These DRs are kept unchanged even if the MPR set changes later (**Fig. 1**).

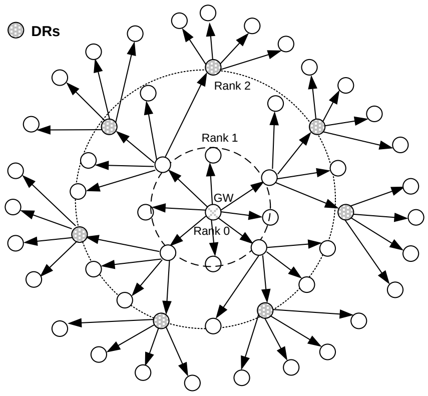

Fig. 1. Delegated router selection example in the proposal

In our multi-homed WMN, IGWs exchange routing information with each other via wired links and decide which MRs they will take care of. The decision is made based on the best routes from MRs to IGWs, i.e. those with the minimum loss probability. **Fig. 2** shows that MR1 and MR2 use IGW1 to access the Internet; while MR3 and MR4 forward traffic toward the Internet via IGW2. This mapping is depicted in the form of IGW: [MR1-MRn]. The information about the mapping between MRs and IGWs is kept in all IGWs and distributed to

DRs by the primary IGW. MRs will query DRs to acquire which IGW they will use to route traffic toward the Internet.

To maintain the locations of MNs inside mesh networks, IGWs store the mappings between MNs and MRs in the form of tuples [MN_MAC, Mesh_ID, MN_IP]. As a result, IGWs and DRs store MN mapping information and IGW/MR combination information. However, only the primary IGW is responsible for allocating IP addresses for MNs. Upon receipt of address allocation requests, the other IGWs forward them to the primary IGW via wired links.

A distribution tree with the primary IGW as the root is created in order for information delivery to the DRs to work efficiently. First, the primary IGW sends messages to all the MPRs at rank 1. Then the MPRs at rank k (k>=1) are responsible for disseminating the messages to the MPRs at rank k+1 and so on. It runs like a multicast technique.

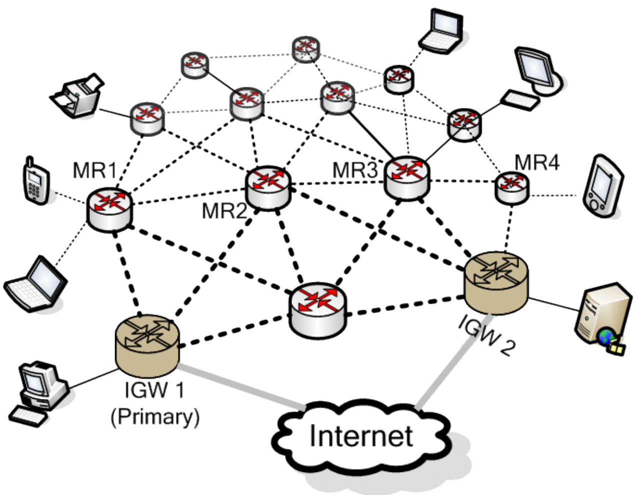

Fig. 2. Model of multi-homed WMNs

2.2 Initial Joining Procedure

Suppose that IGW1 is the primary IGW; the others are secondary. When an MN joins an MR (serving MR), this router queries the DR with the best link quality (**Fig. 3**).

- The DR replies a negative result; i.e. this MN has not joined the network yet and indicates the best IGW (IGWn) for this MR (steps 2 and 3). This IGW information is stored in the serving MR to forward traffic later.
- The serving MR sends a request to the IGWn to require a new IP address for the MN (step 4).
- IGWn forwards the request to IGW1 via a wired link (step 5).
- IGW1 allocates a new IP address, updates its database with the new MN, and sends a reply to IGWn. Then, IGWn sends back the new IP address to the serving MR (step 6).
- Upon receipt of the reply, the serving MR advertises the new IP address to the MN (step 7).
- The serving MR announces the new MN’s mapping to DRs in the vicinity (step 8). The vicinity is defined as a set of DRs such that the distance to the current MR is equal to or fewer than 4 hops.
- Note that, as an MN leaves the network, the serving MR announces the MN’s leaving to the primary IGW, which, in turn, updates its database and distributes this information to

all DRs via the distribution tree and to IGWs in its multicast group. DRs remove the MN's mapping if it exists.

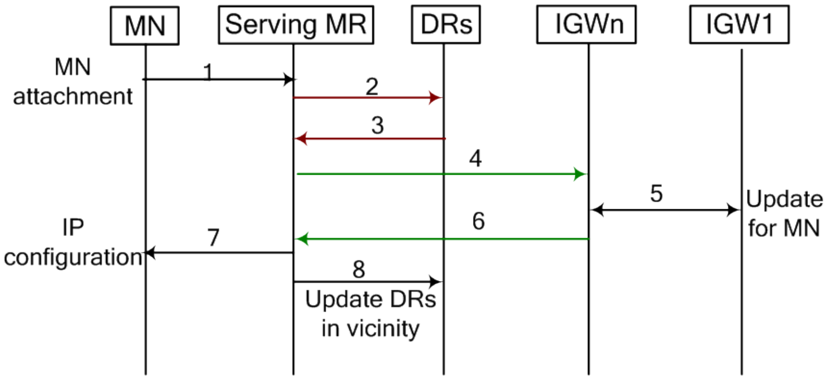

Fig. 3. Initial joining procedure

2.3 Network Address Translation (NAT) Operation and Tunneling Technique

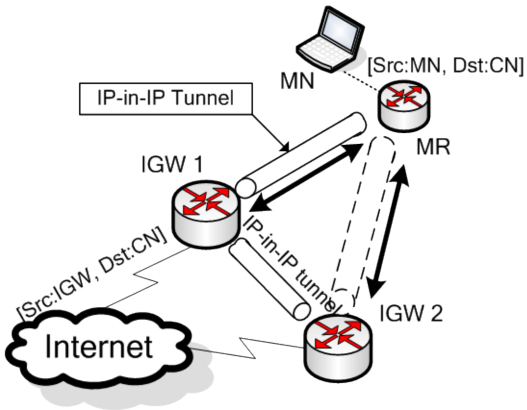

Fig. 4. Illustration of packet transmission

Upon joining the network, an MN begins to send data. Packets from the MN are encapsulated with the MN's private source IP address and a public destination IP address and sent to its associated MR (**Fig. 4**). This MR will perform IP-in-IP packet encapsulation, in which the outer header has the associated MR's source IP address and the respective IGW's destination IP address. When the packet arrives at the IGW specified in the outer header, depending on the initial IGW of the MN, the packet may be forwarded to another IGW or NATed in the current IGW after decapsulation. If the packet belongs to another IGW, i.e. the initial IGW, it will be encapsulated and dispatched to the initial IGW. When the packet arrives at this IGW, its outer header is decapsulated. The IGW then will translate the MN's source IP address to the IGW's public IP address before sending the packet to the Internet. For the reverse direction from the Internet to the MN, the receiving IGW finds out what IGW is currently responsible for the associated MR of the MN. If the current IGW takes care of it, this IGW performs NAT, encapsulation with the respective MR's destination IP address based on the MN mapping, the

IGW mapping, and the NAT rule table, and then it sends the encapsulated packet toward the MN. If the MN's associated MR is served by another IGW, the current IGW encapsulates the packet with the respective IGW's outer destination IP address and transmits it to the specified IGW. Upon receipt of the packet, the destination IGW will NAT, encapsulate, and forward it to the MN as described above.

2.4 Balancing Process

From **Fig. 2**, suppose that, at the beginning, MR1-MR2 use IGW1 to access the Internet; while MR3-MR4 uses IGW2. **Fig. 5** shows that, initially, traffic from the serving routers belonging to IGW1 arrives at IGW1; here, its source IP address is translated to IGW1's public IP address before it is forwarded to the correspondent node (CN). For the reverse direction from CN to MN, NAT also takes place in IGW1; i.e., the destination IP address is translated to the MN's IP address at IGW1. At time t1, IGWs update their routing tables and exchange routing information with each other through wired links and then decide that one active source MR (a source with traffic), say MR2, and the inactive MRs belonging to IGW1, which have link quality from MRs to IGW2 that is better than those to IGW1, will route packets through IGW2. It is likely that more than one active MR will have better routes to IGW2. Among these MRs, only the MR with the highest metric, i.e. the one with the highest probability of successful transmission, is switched to the new IGW. This is because, when an MR changes to another IGW, traffic distribution in the network will change. This probably leads to the metric update and route change. The next negotiation between IGWs will decide whether or not there is more active MR switching. In this way, the ping-pong effect can be avoided.

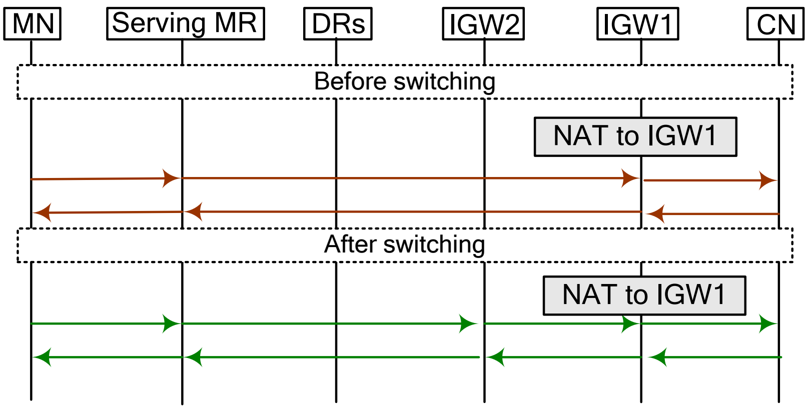

Fig. 5. Traffic flow before and after balancing

At this time, IGW1 requires IGW2 to add MR2 to its supervision set. Upon receiving an acknowledgment from IGW2, IGW1 eliminates MR2 from its supervision and sends a request to instruct MR2 to begin using IGW2. At the same time, it also updates the DRs via the distribution tree. At this moment, the traffic travels between CN and the MN residing in the serving router, as shown in **Fig. 5**:

- Uplink: MN->Serving MR->IGW2->IGW1 (NAT source to IGW1)->CN or MN->Serving MR->IGW1 (NAT source to IGW1)->CN if the serving MR has not received the update yet)
- Downlink: CN->IGW1->IGW2->Serving MR->MN

2.5 Handoff Procedure

When an MN moves around and connects to a nearby MR (**Fig. 6**), this MR is either served by the IGW of the former MR or by another IGW. Upon receiving the MN's attachment trigger, the new MR queries the best DR in terms of link quality and hop count to know where the MN comes from and to get IGW information (step 1). The acquired IGW is always the best one thanks to balancing. Then the new MR requires the former MR to establish a tunnel and forward traffic to the new MR (steps 2-3). At the same time, it sends the MN's mapping update to the specified IGW; i.e., the IGW is serving the MN's associated MR (step 4). This mapping is also disseminated to the DRs in the vicinity for future handoff (step 6). The IGW then updates its database regarding the new location of the MN and dispatches this information to other IGWs in the IGW multicast group (step 5). In the case of different IGWs, upon receipt of an update, the initial IGW will forward incoming traffic to the new IGW. The subsequent steps happen as described in section 2.3.

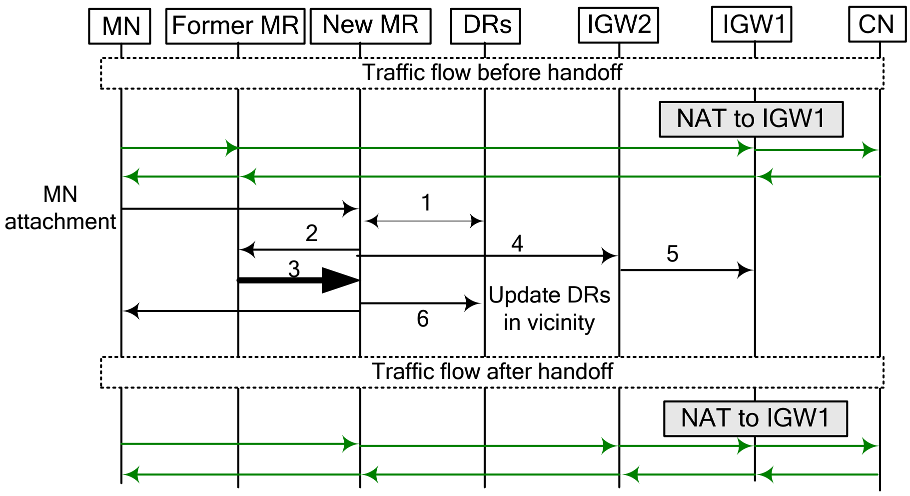

Fig. 6. Handover procedure

The traffic travels between CN and the MN before and after handoff, as shown in **Fig. 6** (in the case of different IGWs):

Before handoff:

- Uplink: MN->former MR->IGW1 (NAT source to IGW1)>CN
- Downlink: CN->IGW1->former MR->MN

After handoff:

- Uplink: MN->new MR->IGW2-> IGW1 (NAT source to IGW1)>CN
- Downlink: CN->IGW1->IGW2->new MR->MN

3. Performance Analysis

3.1 Design Aspect Consideration

First, it is proven that any given MR can get the MN location information at zero, one, or two hops' distance from DRs.

Reference [10] defines $d_F(X,Y)$ as the minimal number of hops, provided that the intermediate relay MRs are forwarders. It was proven that if, for two MRs X and Y, $d_F(X,Y) = k+1$ for the k integer, then Y is at a distance of 1 from the multipoint relay set of rank k of X.

In our proposal, X is the primary Internet gateway and Y is any MR. It is shown that router Y can access a DR at zero, one, or two hops' distance. The number of hops from router Y to the DR falls into the following cases:

- Zero hops: MN moves to Y and Y is the DR itself.
- One hop: in this case, k is an even number. The MPRs at rank k of the GW are the DRs, so Y is at a distance of 1 from the DR.
- Two hops: from the above theorem, there is a path $XM_1M_2\ldots M_{k-1}M_kY$ where M_1 is a multipoint relay of X, M_{i+1} is a multipoint relay for M_i, and Y has a one-hop distance from M_k. We consider k to be an odd number, so k-1 is an even number. Therefore, M_{k-1} is a DR. Thus, to reach M_{k-1}, Y will travel via M_k.

Second, it is prove that, when an MN moves to a new MR, this MR can get updated information from DRs. As defined above, the vicinity of an MR is a set of DRs with a distance to the current MR equal to or fewer than 4-hop. There is a circle with the current MR at the center and a 4-hop radius. Suppose that A1A2A3A4A5A6A7A8A9 is the diameter of the circle with A5 as the current MR, and other A's are any MRs on the path where the MN may move to. The MN may move to A3, A4, A6, or A7 (one- or two-hop movements). According to the proof above, an MR can get information from DRs within a 2-hop distance; therefore, A3, A4, A6, or A7 can get the correct MN information from DRs.

The next analysis is the number of DRs in the mesh network.

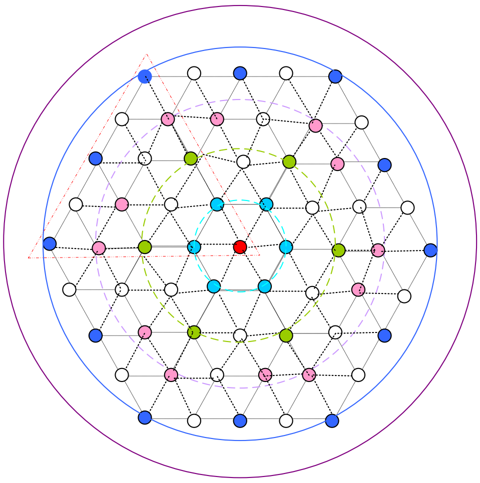

Fig. 7. Illustration of hexagonal topology

We consider the network topology as having the shape of a hexagon in which the primary Internet gateway is allocated in the center. Each router connects to six other routers in the form of hexagon, as shown in **Fig. 7**. Therefore, the total number of routers is:

$$N = 1 + \sum_{i=0}^{K} 6 * i ,$$

where K is the number of ranks.

The MPR heuristic currently used in the OLSR implementation follows a "degree-greedy" strategy that selects neighbors with the largest remaining cover of uncovered two-hop routers.

Upon observation of one of the six parts of the big hexagon (the red triangle in the figure), we can see that each router at the i-th rank has 2 links toward 2 routers at the (i+1)-th rank, and the i-th rank in this part has i+1 routers. From the point of view of the IGW, to cover 3 routers at rank 2, we need 2 routers at rank 1; to cover 4 routers at rank 3, we need 2 routers at rank 2 because each router has only 2 connections to the outside. Therefore, in general, to cover i+1 routers of the i-th rank of one part, we need $round(\frac{i+1}{2})$ routers at the *(i-1)-th* rank.

Thus, to cover all routers at the i-th rank, the number of MPRs at the *(i-1)-th* rank is:

$$N_{i-1} = round(\frac{i+1}{2}) * 6 - 6$$

As a result, the total number of MPRs selected as DRs is:

$$N_{DRs} = \sum_{i=3}^{K} (\frac{i+1}{2}) * 6 - 6, \qquad (1)$$

where i is odd and K is the number of ranks.

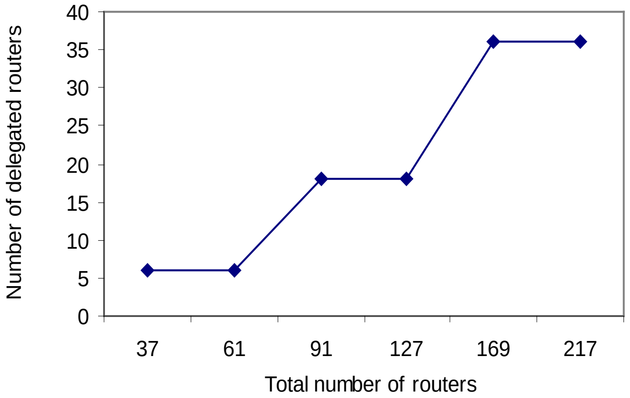

Fig. 8. Number of DRs over the total number of MRs

Based on formula (1), we derive the chart (**Fig. 8**) denoting the number of MRs used to store mapping information over the total number of MRs. In iMesh, every MR has to store the mapping information. On the other hand, the number of DRs in our scheme varies from 10% to 17% of the total number of routers.

Third, we examine an analytical model for calculating handoff delay. The layer 3 handoff delay can be described as the total amount of delay to send a request and receive the information to and from the DR, processing time at the DR (lookup time), and the delay to send a handoff request from the new router to the former router.

$$\begin{aligned} D_{L3\_HO} = {} & 2 * (T_{DIFS\text{-}OR} + T_{BACKOFF\text{-}OR} + T_{TRANSMISSION\text{-}OR} + T_{PROPAGATION\text{-}OR}) \\ & + 2 * (T_{DIFS\text{-}IN} + T_{BACKOFF\text{-}IN} + T_{TRANSMISSION\text{-}IN} + T_{PROPAGATION\text{-}IN}) \\ & + T_{PROCESSING\_TIME} \\ & + (T_{DIFS\text{-}RQ} + T_{BACKOFF\text{-}RQ} + T_{TRANSMISSION\text{-}RQ} + T_{PROPAGATION\text{-}RQ}) \\ & + (T_{DIFS\text{-}RL} + T_{BACKOFF\text{-}RL} + T_{TRANSMISSION\text{-}RL} + T_{PROPAGATION\text{-}RL}) \qquad (2) \end{aligned}$$

where $T_{\_\text{-}OR}$: delay at the originating router (new MR and DR)

$T_{\_\text{-IN}}$: delay at the intermediate router
$T_{\_\text{-RQ}}$: delay at the sending request router (new MR)
$T_{\_\text{RL}}$: delay at the router to relay the request message to the former router

$T_{\_\text{-OR}}$ is zero if new the MR is a DR itself
$T_{\_\text{-IN}}$ is zero with a 0- or 1-hop distance to the DR
$T_{\_\text{RL}}$ is zero with a 1-hop movement

$T_{PROCESSING\_TIME}$ is dominated by the lookup time. With sequence search, it is O(N), where N is the number of mobile nodes.

The propagation delay $T_{PROPAGATION}$ for IEEE 802.11a/b is estimated [12] to be 1μs.

The transmission delay $T_{TRANSMISSION}$ at the router is modeled by an M/G/1 non-pre-emptive priority queue [13]. The routers are represented by an M/G/1 queue to take into account the multiple queues that can have different priorities resulting from packets originating at the router and the packets to be relayed. The packets originating from a router may have a priority higher than that of the packets to be forwarded. For a priority queue at each router with 'c' classes, class 1 as the highest priority, class 2 as the second highest, and up to class c which is the lowest class, we use the following notations:

- c : number of queuing classes at each router
- μ_K : service rate in the k-th class of a router
- λ_K : arrival rate in the k-th class of a router
- ρ_k : utilization factor for class k. $\rho_k = \lambda_k \mu_k$

Considering M/G/1 non-pre-emptive priority queuing at the routers and using results from queuing theory,

$$T_{TRANSMISSION\_k} = \frac{R}{(1-\rho_1-...-\rho_{k-1})(1-\rho_1-...-\rho_k)} + \mu_k$$

$$\text{where } R = \frac{1}{2}\sum_{k=1}^{c} \lambda_k \mu_k^2 \text{ : mean residual service time}$$

Depending on packets originating from a router or packet to be relayed, k has different values:

- Packets to be relayed at a router: because these packets have the least priority, the transmission delay is:

$$T_{TRANSMISSION} = \frac{R}{(1-\rho_1-...-\rho_{c-1})(1-\rho_1-...-\rho_c)} + \mu_c$$

- Packets originating from a router: the transmission delay for these packets depends on the class of the packets in the router.

3.2 Simulation

We use an ns-2 simulator to simulate our proposed scheme. The distance between MRs is set to about 250m in this simulation, the carrier sense is approximately 550m, and the access technology is 802.11b. Two MRs function as IGWs that connect the wireless mesh network to the Internet. Between two IGWs, it is supposed that they are connected by a wired link. OLSR-ML is running as the routing protocol supporting the link quality-based routing. Clients under MRs: MR1, MR3, and MR4, generate traffic flow 1, flow 3, and flow 4, respectively, at the same rate of 250 Kbps, while MR2 has a 200 Kbps flow (flow 2). Traffic from MR1, MR2, and MR3 starts at the same time.

Initially, there is no traffic in the network. Therefore, after IGWs negotiate and disseminate IGW information to the DRs, the MRs will gain the IGWs with the shortest hop count. MR1, MR2, and MR4 route traffic via IGW1, while MR3 uses IGW2 to access the Internet (**Fig. 9**).

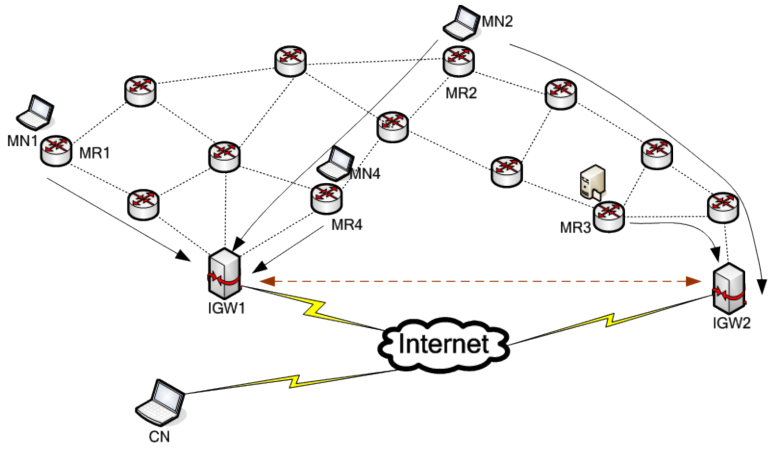

Fig. 9. Illustrating balancing in multi-homed WMNs

To evaluate the handoff latency, we consider the delay in sending a request and receiving the information to and from the DR, plus the delay in sending a handoff request from the new router to the former router. We create two distinct movement scenarios: MN moves with 1 hop and 2 hops away from the former router.

When an NR recognizes an MN's attachment, it will send a query to the nearest DR to get the previous MN's location as well as an appropriate IGW. There are three possible cases: the NR is a DR, so it gets information from itself (1); if the NR is different from the DR, the query message will travel from the NR to a DR at a one-hop (2) or two-hop distance (3).

Fig. 10 shows that the round-trip propagation delay with background traffic is approximately 30ms when the NR needs 2 hops to reach the required DR and less than 20ms for a 1-hop distance. These values combined with the layer 2 handoff delay are appropriate to keep the MNs' connection continuous. In fact, with an even distribution of DRs, the probability of leading to 2-hop path to a DR is small; this only happens to the outmost MRs.

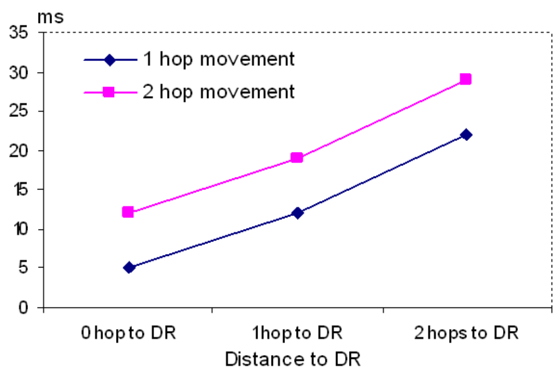

Fig. 10. Round-trip delay with background traffic

In comparison with M^3, the latency in our proposal is equal to the latency in M^3 when the new router is the DR itself with a 1-hop movement of the MN and higher than that in M^3 when the new router is different from the DR. This is because our proposal does not require MNs to participate in any mobility signaling. Therefore, to achieve transparency to MNs, the new router has to find out where MNs come from by querying the nearest DR. This results in the higher latency in our proposal than that in M^3. Nevertheless, in M^3, if MNs move more than 1 hop, it is impossible for intermediate MRs to forward packets to the new router.

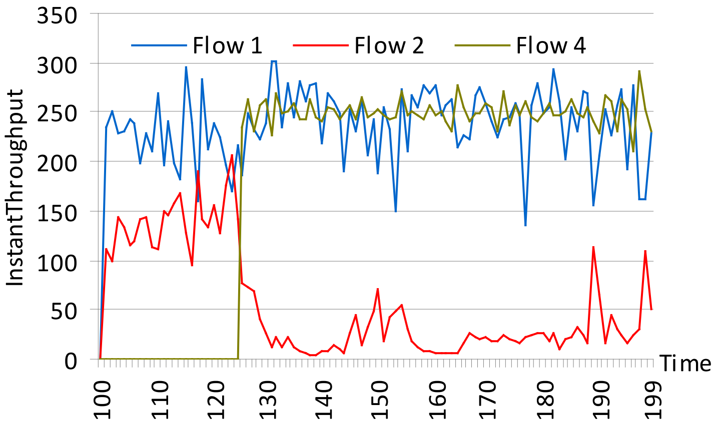

Fig. 11. Instant throughput without balancing

The next evaluation is the effect of load balancing on the throughput and end-to-end delay parameter. After traffic from MR4 starts the throughput as well as the delay from MR1 and MR2 suffers due to a scramble for wireless resources, especially MR2, it is allocated far from IGW1. This results in a link quality change and route update. The route metric from MR2 to IGW2 now becomes better than the one to IGW1. After negotiation, the IGWs decide that MR2 will route packets through IGW2 while ensuring that the connection is kept continuous. **Fig. 11** and **Fig. 12** show the instant throughput without and with balancing, respectively.

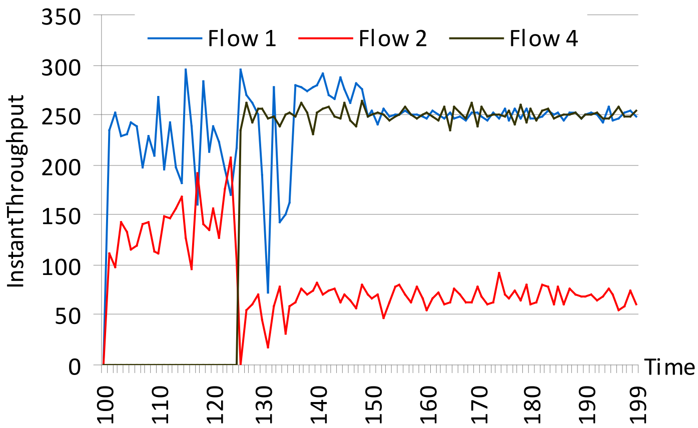

Fig. 12. Instant throughput with balancing

Fig. 11 and **12** show that, as flow 4 starts, the throughput of existing flows 1 and 4 descends immediately. With a 3-hop distance to IGW1, flow 2 suffers considerable decrement. However, with a balancing scheme, this flow can improve its throughput and remain stable until the end of the session.

Fig. 13 shows that, when MR2 finds a better route and switches to IGW2, its throughput increases even if the hop count of the new path is greater than that of the previous one. It is worth noting that the throughput of MR3 is almost intact regardless of the additional flow 2. In addition, traffic flow from MR1 and MR4 gains more shared wireless resources; this leads to higher throughput. Flow 3 is affected slightly, as shown in **Fig. 13**; since MR3 is located near IGW2, it needs only two hops to reach IGW2. Hence, the appearance of flow 2 reduces the wireless resources but does not affect flow 3 too much. Even through the throughput of flow 3 is affected by the appearance of flow 2, the overall throughput of the mesh network increases thanks to the appropriate traffic distribution after balancing.

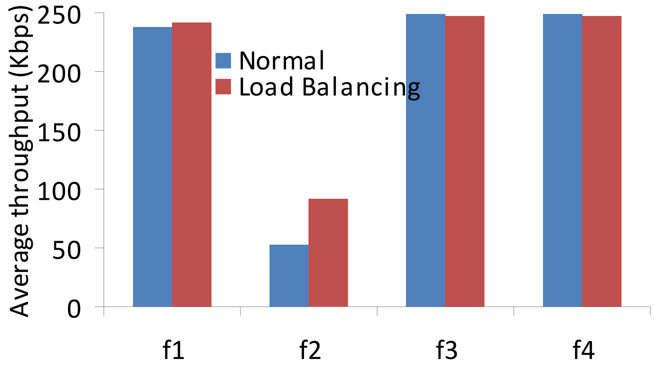

Fig. 13. Average throughput without and with balancing

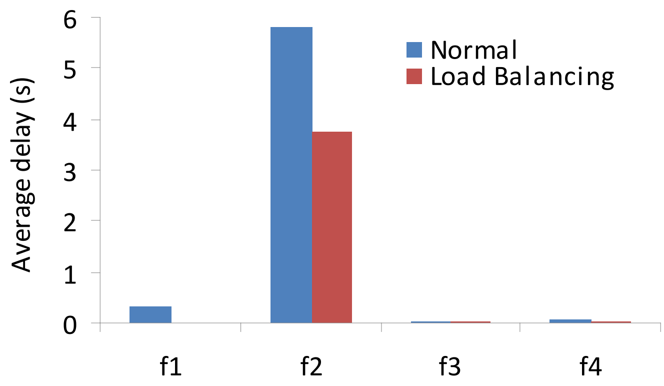

Fig. 14. Average delay without and with balancing

We also observe from **Fig. 14** that the average delay in the network using our proposed scheme is smaller than a scheme in which there is no balancing. As traffic distribution is more appropriate, the waiting time to access the shared wireless resources lessens. Even if the traffic of MR2 has to traverse the IGW1-IGW2 link, the delay caused by the wired link is negligible.

"Packet delivery ratio" is defined as the number of packets received at the sink node over the number of packets sent by the source node. We can recognize that flow 2 with a balancing scheme can get a better packet delivery ratio than a normal scheme (**Fig. 15**). In addition, the delivery ratio of other flows remains almost unchanged.

Signaling overhead also causes performance decrement in the network. When route re-calculation occurs due to link quality variance, some MRs likely have better paths to other IGWs. As a result, IGWs negotiate with each other and decide which MRs they will take care of, and then the primary IGW disseminates information to the DRs via the distribution tree. The negotiation process happens through wired links, and the distribution to DRs is based on multicast routing; this will partially reduce overhead.

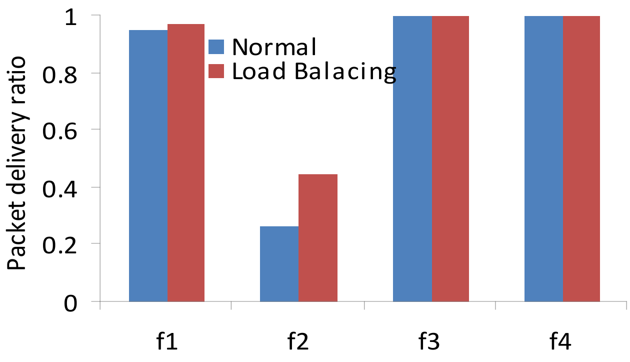

Fig. 15. Packet delivery ratio without and with balancing

Next, we take into account the handoff aspect. This proposal ensures that every MRs can get information at a zero-, one-, or two-hop distance from the DRs. Besides ensuring that the number of hops is less than or equal to two, it also pays attention to link quality; i.e., the path to the DR with the highest quality will be chosen when there are more than one nearest DRs with equal hop counts. This increases the probability of successfully getting information from DRs and leads to faster handoff. In addition, when an MN roams and attaches to an MR, this scheme ensures that the MN can always route packets via the best IGW thanks to load balancing between IGWs.

Besides the handoff latency, the convergence time is also important. The convergence time is a period when the traffic begins traveling on a temporary path after handoff until it can go on its optimal path. In this proposed solution, traffic can go on its optimal path when the original IGW receives and updates the MN's new location. For a topology with one IGW, if t is the time for the farthest MR to send an update message to the IGW, then the necessary time in a topology with two IGWs is approximately t/2 thanks to the wired link transmission between two IGWs.

4. Conclusion

We design a load balancing scheme in which IGWs exchange routing information via a wired link and make a decision about which MRs they will serve. Then, the primary IGW will distribute the information to DRs via the distribution tree. The balancing process ensures that ongoing transmissions are kept continuous. Moreover, thanks to load balancing, MNs can route packets via the best IGW when a handoff occurs. The handoff time cost varies within an acceptable range thanks to the distribution of the MN location information in the vicinity of the current delegated MR. Finally, the convergence time, the period of time when traffic begins traveling on a temporary path after handoff until it can go on its optimal path, is reduced.

References

[1] Y. Zhang, J. Luo and H. Hu, "Wireless Mesh Networking: Architectures, Protocols and Standards," *Auerbach Publications*, 2007. Article (CrossRef Link).

[2] I. F. Akyildiz and X. Wang, "A Survey on Wireless Mesh Networks," *IEEE Commun. Mag.*, vol. 43, no. 9, pp. S23–S30, Sep. 2005. Article (CrossRef Link).

[3] J. Xie and X. Wang, "A Survey of Mobility Management in Hybrid Wireless Mesh Networks," *IEEE Network*, vol. 22, no. 6, pp. 34-40, Dec. 2008. Article (CrossRef Link).

[4] V. Navda, A. Kashap and S. R. Das, "Design and Evaluation of iMesh: An Infrastructure-Mode Wireless Mesh Networks," in *Proc. of 6th IEEE Int'l. Symp. World of Wireless Mobile Multimedia Networks*, pp. 164-70, Jun. 2005. Article (CrossRef Link).

[5] R. Huang, C. Zhang and Fang, "A Mobility Management Scheme for Wireless Mesh Networks," in *Proc. of IEEE Global Telecom. Conf*, pp. 5092-96, Nov. 2007. Article (CrossRef Link).

[6] Nandiraju.D, Santhanam.L, Nandiraju.N et al, "Achieving Load Balancing in Wireless Mesh Networks through Multiple Gateways," in *Proc. of Mobile Ad hoc and Sensor Systems (MASS), 2006 IEEE International Conference on*, pp. 807-812, 2006. Article (CrossRef Link).

[7] Aure. T and Li. F.Y, "An optimized path-selection using airtime metric in OLSR networks: Implementation and testing," in *Proc. of Wireless Communication Systems, ISWCS*, pp. 359–363, 2008. Article (CrossRef Link).

[8] P.E. Engelstad, A. Tonnesen, A. Hafslund and G. Egeland, "Internet connectivity for multi-homed proactive ad hoc networks," in *Proc. of IEEE International Conference on Communications 2004*, vol. 7, pp. 4050-4056, 2004. Article (CrossRef Link).

[9] K. Mase and Y. Owada, "Gateway Aggregation Protocol (GAP) for Mobile Ad hoc Networks," *draft-mase-autoconf-gap-02, Internet Draft*, 2007. Article (CrossRef Link).

[10] P. Jacquet, P. Muhlethaler, T. Clausen, A. Laouiti, A. Qayyum, L. Vienot, "Optimized Link State Routing Protocol for Ad Hoc Networks," in *Proc. of INMIC*, pp. 62-68, 2001. Article (CrossRef Link).

[11] D. Passos, D. Teixeira, D.C. Muchaluat-Saade, L.C. Schara Magalhaes and C. Albuquerque, "Mesh Network Performance Measurements," in *Proc. of 5th International Information and Telecommunication Technologies Symposium, Cuiaba, MT, Brazil*, Dec. 2006. Article (CrossRef Link).

[12] Y. Xiao, J. Rosdahl, "Throughput and Delay Limits of IEEE 802.11," *IEEE Comm. Letters*, vol. 11, no. 8, Aug. 2002. Article (CrossRef Link).

[13] Santosh S.Venkatesh, "Networking- Theory and Fundamentals," *University of Pennsylvania, 1997*. Article (CrossRef Link).

Minh Tri Tran received the B.S. degree in CanTho University, Vietnam. He is currently studying master course in Soongsil University, Data and Communication Network lab. His interests include wired and wireless networking, mobile computing and ubiquitous networking.

Younghan Kim received the B.S. degree in electronic engineering from Seoul National University, Korea, in 1984. He received the M.S. and Ph. D. degrees in electrical engineering from KAIST, Korea, in 1986 and 1990, respectively. From 1987 to 1994, he worked for Digicom Institute of Telemetics as a senior research engineer, where he developed ISDN and packet switching systems. Since 1994, he has been a professor in the School of Electronics at Soongsil University in Korea. His research interests include wired and wireless networking, QoS, mobile computing and ubiquitous networking. He is a member of IEICE and IEEE.

Jae-Hwoon Lee received the B.S degree in electronic engineering from Hanyang University, Seoul Korea, in 1985, and the M.S. and Ph.D. degrees in electrical and electronics engineering from Korea Advanced Institute of Science and Technology (KAIST), Korea, in 1987 and 1995, respectively. From 1987 to 1990, he served as a Researcher in the DACOM. At DACOM, his research area was line of packet switching network protocols. From 1990 to 1999, he was with Samsung Electronics, where he worked at Network Systems Group. Since March 1999, he has been with Dongguk University, where he is a Professor of Information and Communications Engineering. His research interests include Internet protocol, IP mobility and high-speed networks.