Load Balancing in Wireless Mesh Network: a Survey

Maryam Asgari¹, Mohammad Shahverdy², Mahmood Fathy³, Zeinab Movahedi⁴

¹Department of Computer Engineering, Islamic Azad University-Prof. Hessabi Branch, Tafresh, Iran
Asgari@iautb.ac.ir

²Department of Computer Engineering, Islamic Azad University- Prof. Hessabi Branch, Tafresh and PHD student of University of Science and Technology, Tehran Iran
Shahverdy@iautb.ac.ir

³, ⁴Computer Engineering Faculty, University of Science and Technology Tehran, Iran
zmovahedi@iust.ac.ir, mahfathy@iust.ac.ir

Abstract
Wireless mesh network (WMN) is a state of the art networking standard for next generation of wireless network. The construction of these networks is basis of a network of wireless routers witch forwarding each other’s packets in a multi-hop manner. All users in the network can access the internet via Gateways nodes. Because of the high traffic load towards gateway node, it will become congested. A load balancing mechanism is required to balance the traffic among the gateways and prevent the overloading of any gateway. In this paper, we investigated different load balancing techniques in wireless mesh networks to avoid congestion in gateways, as well as we survey the effective parameters that is used in these techniques.

Keywords: clustering, Gateway, Load Balancing, Wireless Mesh Network (WMN).

1. Introduction
Wireless mesh networking is a new paradigm for next generation wireless networks. Wireless mesh networks (WMNs) consist of mesh clients and mesh routers, where the mesh routers form a wireless infrastructure/backbone and interwork with the wired networks to provide multi-hop wireless Internet connectivity to the mesh clients. Wireless mesh networking has generated as a self-organizing and auto-configurable wireless networking to supply adaptive and flexible wireless Internet connectivity to mobile users. This idea can be used for different wireless access technologies such as IEEE 802.11, 802.15, 802.16-based wireless local area network (WLAN), wireless personal area network (WPAN), and wireless metropolitan area network (WMAN) technologies. WMNs Potential application can be used in home networks, enterprise networks, community networks, and intelligent transport system networks such as vehicular ad-hoc networks. Wireless local area networks (WLANs) are used to serve mobile clients access to the fixed network within broadband network connectivity with the network coverage [1]. The clients in WLAN use of wireless access points that are interconnected by a wired backbone network to connect to the external networks. Thus, the wireless network has only a single hop of the path and the Clients need to be within a single hop to make connectivity with wireless access point. Therefore to set up such networks need access points and suitable backbone. As result a Deployment of large-scale WLANs are too much cost and time consuming. However, The WMNs can provide wireless network coverage of large areas without depending on a wired backbone or dedicated access points [1, 2]. WMNs are the next generation of the wireless networks that to provide best services without any infrastructure. WMNs can diminish the limitations and to improve the performance of modern wireless networks such as ad hoc networks, wireless metropolitan area networks (WMANs), and vehicular ad hoc networks [2, 3, 4 and 5]. WMNs are multi-hop wireless network which provide internet everywhere to a large number of users. The WMNs are dynamically self-configured and all the nodes in the network are automatically established and maintain mesh connectivity among themselves in an ad hoc style. These networks are typically implemented at the network layer through the use of ad hoc routing protocols when routing path is changed. This character brings many advantages to WMNs such as low cost, easy network maintenance, more reliable service coverage. Wireless mesh network has different members such as access points, desktops with wireless network interface cards (NICs), laptops, Pocket PCs, cell phones, etc. These members can be connected to each other via multiple hops. In the full mesh topology this feature brings many advantages to WMNs such as low cost, easy network maintenance and more reliable service coverage. In the mesh topology, one or multiple mesh routers can be connected to the Internet. These routers can serve as GWs and provide Internet connectivity for the entire mesh network. One of the most important challenges in these networks happens on GW, when number of nodes which connected to the internet via GW, suddenly increased. It means that GWs will be a bottleneck of network and
performance of the network strongly decreases [4, 5, and 6].

2. Related Work
The problem of bottleneck in wireless mesh networks is an ongoing research problem although much of the literature [7, 8, 9, 10] available, addresses the problem without an introducing method for removing bottleneck and/or a well-defined way to prevent congestion. In [11], the authors proposed the Mesh Cache system for exploiting the locality in client request patterns in a wireless mesh network. The Mesh Cache system alleviates the congestion bottleneck that commonly exists at the GW node in WMNs while providing better client throughput by enabling content downloads from closer high throughput mesh routers. There is some papers related to optimization problems on dynamic and static load balancing across meshes [11]. Optimal load balancing across meshes is known to be a hard problem. Akyildiz et al. [12] exhaustively survey the research issues associated with wireless mesh networks and discusses the requirement to explore multipath routing for load balancing in these networks. However, maximum throughput scheduling and load balancing in wireless mesh networks is an unexplored problem. In this paper we survey different load balancing schemes in wireless mesh networks and briefly introduce some parameters which they used in their approaches.

3. Load Balancing Techniques
Increasing Load in a wireless mesh network causes Congestion and it is lead to different problems like packet drop, high end to end delay, throughput decline etc. various techniques have been suggested that considers load balancing are discussed below.

3.1 Hop-Count Based Congestion-Aware routing [13]

In this routing protocol, each mesh router rapidly finds out multiple paths based upon hop count metric to the Internet gateways by routing protocol designing. Each mesh router is equipped to a bandwidth estimation technique to allow it to forecast congestion risk, then router select high available bandwidth link for forwarding packets. Multipath routing protocol consists two phases: Route discovery phase and path selection phase.

In the route discovery phase, whenever a mesh router tries to find a route to an internet gateway, it initiates a route discovery process by sending a route request (RREQ) to all its neighbors. The generator of the RREQ marks the packet with its sequence number to avoid transmitting the duplicate RREQ.

A mesh router hearing the route request uses the information in the RREQ to establish a route back to the RREQ generator.

During the path selection phase a source should decide which path is the best one among the multiple path configured out in the first phase. The path selection can be prioritized in following order:

(a) If there exist multiple paths to a source’s primary gateway then, take the path with minimum hop count and if there is still a tie, we can randomly opt a path.

(b) If there is no path to source’s primary gateway but a several paths to secondary gateways then take the path with minimum hop count and if there is still a tie opt a path randomly.

As it’s clear, congestion control is based on bandwidth estimation technique, therefore available bandwidth on a link should be identified. Here the consumed bandwidth information can be piggy packed on to the “Hello” message which is used to maintain local connectivity among nodes. Each host in the network determines its devoted bandwidth by monitoring the packets it sends on to the network. The mesh router can detect the congestion risk happening on its each link by the bandwidth estimation techniques. A link is in risk of congestion whenever the available bandwidth of that link is less than a threshold value of bandwidth. If a link cannot handle more traffic, it will not accept more requests over that link. The primary benefit of this protocol is that it simplifies routing algorithm but it needs precise knowledge about the bandwidth of each link.

3.2 Distributed Load Balancing Protocol [14]

In this protocol the gateways coordinates to reroute flows from congested gateways to other underutilized gateways. This technique also considers interference which can be appropriate for practical scenarios, achieving good results.
and improving on shortest path routing. Here the mesh network is divided into domains. A domain \( d_i \) can be defined as set of routers which receive internet traffic and a gateway which serve them. For each domain a specific capacity is assigned and is compared against the load in the domain. The domain is considered as overloaded if the load exceeds the sustainable capacity. To avoid congestion in a domain we can reroute the traffic. This technique does not impose any routing overhead in the network.

3.3 Gateway–Aware Routing [15]

In [15] a gateway mesh aware routing solution is proposed that selects gateways for each mesh router based on multihop route in the mesh as well as the potentiality of the gateway. A composite routing metric is designed that picks high throughput routes in the presence of multiple gateways. The metric designed is able to identify congested part of each path, and select a suitable gateway. The gateway capacity metric can be defined as the time needed to transmit a packet of size \( S \) on the uplink and is expressed by

\[
gwETT = \text{ETXgw}S/Bgw
\]  

Where \( \text{ETXgw} \) is the expected transmission count for the uplink and \( Bgw \) is the capacity of the gateway. For forwarding packets a GARM (Gateway Aware Routing Metric) is defined which is follows:

\[
\text{GARM} = \beta M_i + (1-\beta)(mETT+gwETT)
\]  

This Gateway-aware Routing Metric has two parts. The first part of the metric is for bottleneck capacity and the second part accounts the delay of the path. The \( \beta \) is used for balancing between these two factors. The gateway with minimum GARM value can be chosen as the default gateway for balancing the load. This scheme overcomes the disadvantage of accurate bandwidth estimation suggested in [6] and also improves network throughput.

3.4 DGWLBA: Distributed Gateway Load balancing Algorithm [16]

In [16] gateways execute DGWLBA to attain load balancing. DGWLBA starts by assigning all routers to their nearest gateway that is called the NGW solution. Next steps consist in trying to reroute flows from an overloaded domain \( d_1 \) to an uncongested domain \( d_2 \) such that the overload of both domains is reduced.

If domain is overloaded, its sinks are checked in descending order of distance to their serving gateway. This is done to give preference to border sinks. The farther a sink is from its serving gateway the less it will harm other flows of its domain if it is rerouted. And its path to other domains will be shorter, thus improving performance. For the same reason, when a sink is chosen, domains are checked in ascending order of distance to the sink. Next, to perform the switching of domains, the overload after the switch must be less than the overload before the switch (lines 9-11).

Lastly, the cost of switching is checked. ngWsis the gateway nearest to \( \Delta s \). Only if the cost is less than the switching threshold \( \Delta s \) will be performed (line 12). This rule takes into account the existence of contention, because it prevents the establishment of long paths, which suffer from intra-flow interference and increase inter-flow interference in the network, and gives preference to border sinks. Hence this approach successfully balances load in overloaded domains considering congestion and interference.

\begin{verbatim}
for each gateway GWi do di = {}; for each sink s do if ( distance(s,GWi) = minimum) Add sink s to di; For domain d_i in D do if load(d_i) > C_{d_i} then For sink s in d_i do
\end{verbatim}
For domain $d_1$ in $D$ do
  If $d_1 = d_2$ then
    Continue
  Ovld_{before} = ovld(d_1) + ovld(d_2)
  Ovld_{after} = ovld(d_1 - \{s\}) + ovld(d_2 \cup \{s\})
  If ovld_{after} < ovld_{before} then
    If dist(s,GW_2) / (dist(s,nGW_n) < \Delta_s then
      d_1 = d_1 - \{s\}
      d_2 = d_2 \cup \{s\}
      break;
  If load(d_1) \leq C_d then
    Break;

3.5 Load Balancing in WMNs by Clustering[17]

In [17] authors proposed a load balancing schemes for WMNs by clustering. In first step all nodes are clustered to control the workload of them. If workload on a GW is increased up to maximum capacity of the GW then the cluster is broken. With the respect to the gateways capacity, the gateways overload can be predictable. Because selecting a new GW and establish a route table is time consuming, thus third scheme is proposed which GW selection and creating route table is done before breaking the cluster. Also they considered some parameters for selecting the new GW in new cluster which is offered in following formula:

$$G_{value} = \frac{\text{Power_{supply}} \times \text{Power_{CPU}} \times \text{Constancy} \times \text{Distance}}{\text{Velocity}}$$

Where $\text{Power_{supply}}$ is the power of a node, $\text{Power_{CPU}}$ is the processing power of each node, Constancy is the time which a node actively exists in cluster, Velocity is the speed of each node and Distance is the distance of the node to central of the cluster. In the above formula, they calculate $G_{value}$ for each node in a cluster and then each node that has larger $G_{value}$ is more suitable for being a GW.

Although the paper considers most of the design aspects of the proposed infrastructure, it leaves some open issues and questions. For instance, surveying load balancing of multi-channel GWs in clustering wireless mesh networks, finding maximum throughput of nodes in cluster based wireless mesh networks. Another open issue is using fuzzy logic for breaking the clusters.

4. Conclusion

Load balancing is one of the most important problems in wireless mesh networks that needs to be addressed. The nodes in a wireless mesh network trend to communicate with gateways to access the internet thus gateways have the potential to be a bottleneck point. Load balancing is essential to utilize the entire available paths to the destination and prevent overloading the gateway nodes. In this paper we surveyed different load balancing scheme with various routing metrics that can be employed to tackle load overhead in the network. Table1 summarizes the load balancing techniques which we surveyed in this paper.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Metric</th>
<th>Advantages</th>
<th>Issues that not Addressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hop Count based Congestion Aware routing</td>
<td>Hop Count</td>
<td>No routing overhead.</td>
<td>Computational overhead. accurate bandwidth information required.</td>
</tr>
<tr>
<td>DISTRIBUTED GATEWAY LOADBALANCING</td>
<td>Queue Length</td>
<td>Low end to end delay</td>
<td>Routing and Computational Overhead.</td>
</tr>
<tr>
<td>Load Balancing in WMNs by Clustering</td>
<td>Queue Length</td>
<td>Low end to end delay</td>
<td>Cluster initial formation parameter</td>
</tr>
</tbody>
</table>
References


