Optimizing Multi Gateway Wireless Mesh Networks for Throughput Improvement

Soma Pandey[1(✉)], Govind Kadambi[1], and Vijay Pande[2]

[1] Faculty of Engineering and Technology, MSRUAS, Bangalore, India
soma.p@ieee.org
[2] School of Engineering, Amrita Vishwa Vidyapeetham, Bangalore, India

Abstract. This paper applies the concept of subnet virtualization to the edge network comprising of the multi-gateway Wi-Fi mesh. A necessary and sufficient condition for improving the throughput of Wi-Fi mesh network (WMN) is proposed in the paper. A holistic approach of optimizing the mesh topology by fair distribution of gateways (GW) is developed. Subnets (partitions) are created within the mesh such that each partition has one GW and approximately equal amount of Mesh Routers. Thereafter an overload estimation process is defined which indicates the instance when the WMN is overloaded and a Load Management Scheme (LMS) has to be applied. A Steady State Load equation is derived based on the current processing load of each GW. Thereafter a stability condition is defined which can avoid triggering chain of load transitions from one neighbor GW to another. Simulation studies presented in the paper show that after providing a conventional WMN with the features of the proposed LMS, the throughput became more than double, there was a decrease of 22% in the average packet delay and a decrease of 90% in the number of packets dropped.

Keywords: Wi-Fi mesh · IEEE802.11s · Optimizing mesh networks · Multiple gateway load balancing · Virtualization

1 Introduction

The IEEE 802.11s [1, 2] Wi-Fi mesh standard has still not been able to gain the kind of popularity that is enjoyed by its low data rate counterpart; the sensor mesh networks. The main focus of mesh networks is to extend the coverage of Wi-Fi through use of routers. IEEE 802.11s standard gave lot of hope to increase coverage of Wi-Fi broadband but could not gain much traction. Some of the reasons for the standard not gaining popularity are following

- Throughput drops with increase in number of hops [3]
- There is no guarantee of minimal Quality of Service (QoS) support in order to present it as a commercial network
- The standard proposes multiple Internet Access Points (Gateways). Increasing the number of GWs need not increase the throughput necessarily. In fact, in some cases multiple GWs might reduce throughput at some of the MRs due to GW contention issues as proved in [4]

N. Kumar and R. Venkatesha Prasad (Eds.): UBICNET 2019, LNICST 276, pp. 33–52, 2019.
https://doi.org/10.1007/978-3-030-20615-4_3

Devising a suitable MR-GW association process which also incorporates load management amongst multiple GWs is important to achieve better throughput and thereby better QoS in WMNs. This process should be efficient enough to be able to fairly allocate MRs to GWs with low time complexity. Various methods for load balancing and allocation of GWs are available in literature. Some of the papers on load balancing include [5–8]. The main disadvantage of these schemes is their requirement to calculate and save all possible alternative paths to the available GWs. Implementing such schemes results in creating large overheads in both time and space.

In this paper the concept of load sharing in WMNs is proposed. The major advantage of load sharing over load balancing is the elimination of repeated path computations between the MRs and GWs. Conventional load management in WMN is performed continuously thereby consuming the resources actually meant to be used for Internet traffic. Such schemes have high computational time complexity. Therefore there is need for a load aware GW scheduling mechanism. The major challenges in load aware GW scheduling are

- High computational time complexity of scheduling algorithm
- Large packet processing and queuing delays

This paper proposes to overcome these challenges by ‘fixed partitioning’ and ‘load sharing’. Instead of the conventional method of combining the GW scheduling with load balancing, this paper proposes to perform both of these processes independently. Initially partitioning is performed to define a GW and its associated MRs. If load demand of a partition exceeds maximum capacity of its GW, then some MRs are shifted/transited to a less loaded neighbouring partition. This process is called load sharing amongst the GWs.

Another novelty of this paper lies in the fact that none of the literature presents such a detailed study on effect on the throughput of network when the number of Mesh Routers (MRs) is fixed and the number of GWs is increased systematically. Similarly the study also investigates the change in throughput of WMN when number of GWs is kept fixed and number of MRs is increased. This kind of simulation study is very valuable for network planning and designing and correct assessment of change in throughput values with change in number of load generating MRs and load processing GWs.

Section 2 of the paper presents a brief overview of the present IEEE 802.11s mesh architecture. Thereafter it is explained how this architecture can be optimised by the work proposed in this paper. A comparison of the present and the optimised mesh architecture is provided in Figs. 1 and 2. The next Sect. 3 builds the framework to optimise the WMN through Partitioning and load management procedures. Section 3.1 describes partitioning procedure and Sect. 3.2 describes the load management scheme for load optimisation and load sharing within the WMN. Lost nodes and redundant nodes are a possibility when nodes are transited from one partition to another during load management. Section 3.3 extends the proposed partitioning and LMS to map it onto a matrix representation. This matrix representation helps in validating maintaining integrity of mesh topology by providing a check on lost nodes or redundant nodes. Section 4 provides a comparative performance analysis of the mesh throughput obtained in the existing IEEE 802.11s WMN as compared to the throughput obtained in a WMN which is optimised after applying the proposed LMS. In the end, Sect. 5 and Sect. 6 provide conclusions on this work and future extensions of the proposed LMS respectively.

2 Proposed Architecture

This paper proposes city wide Wi-Fi mesh based on IEEE 802.11s standard for WMNs. This standard is compatible to any Wi-Fi (IEEE 802.11x) based end node. The only changes needed will be on the routers and the Network Operation Centre (NOC). The solution addresses QoS provisioning within the WMN through load management. The novelty lies in the proposed architecture as it serves the purpose of deploying single WMN for entire city.

This is achieved in this paper as explained next. First a topology is defined for the WMN which addresses the GW parenting issue. The topology is based on well-defined partitions/service clusters around the GWs. A partitioning algorithm is developed by modifying the Ciarlet and Lamour's graph partitioning algorithm [9] for WMNs with multiple GWs. This was published in [4] by the authors. This partitioning algorithm leverages the matrix representation of graphs for marking and collecting MRs for each partition. To achieve this a node marking algorithm is developed [10]. The partitioning algorithm defines partitions with the GW and a set of MRs around it. The algorithm ensures that each partition has exactly one GW and a set of nodes (MRs) which will be serviced by the GW. The algorithm also computes and assigns nearly same number of MR to each GW to ensure fairness in GW scheduling. This is explained in Sect. 3.1 of this paper.

Once the partitions are in place a mechanism has to be drafted to prevent overloading of the partitions. For this each partition has to be monitored for load. A load monitoring and overloading condition called the Steady State Load (SSL) condition is defined based on the load on each GW. If the SSL condition is violated then load management has to be performed. Overload is managed by moving some MRs from one partition to another. This is achieved by updating the routing table entries managed at each GW. This is explained in Sect. 3.2 of the paper.

The partitions generated by the partitioning algorithm are further mapped onto the matrix model of the WMN graph. The purpose of mapping partitions onto a single WMN matrix is to provide a unified model to represent the partitions of the WMN. A validation equation is formulated which is required to ensure that there are no lost nodes/partitions [18]. The purpose of maintaining a matrix model, is to monitor a subnet GW connectivity and link through the adjacency matrix of the subgraph. This solution also proposes using of validation equation defined on the matrix model of the partitioned WMN. Violation of the validation equation means there are lost nodes or redundant nodes in the system. The validation equation also keeps check on the security aspect of the network wherein no additional malicious MRs can be added to the network. This is explained in Sect. 3.3 of the paper.

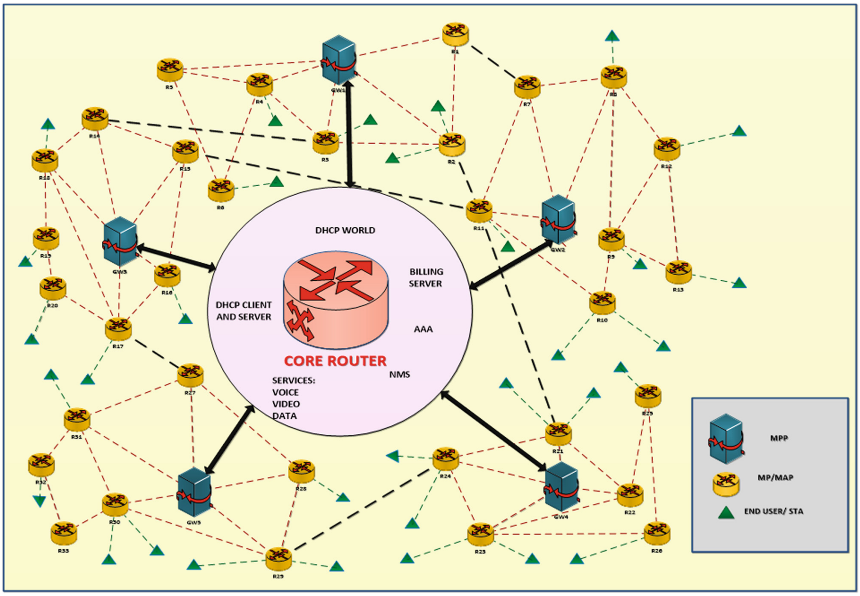

Fig. 1. A generic 802.11s network. (Color figure online)

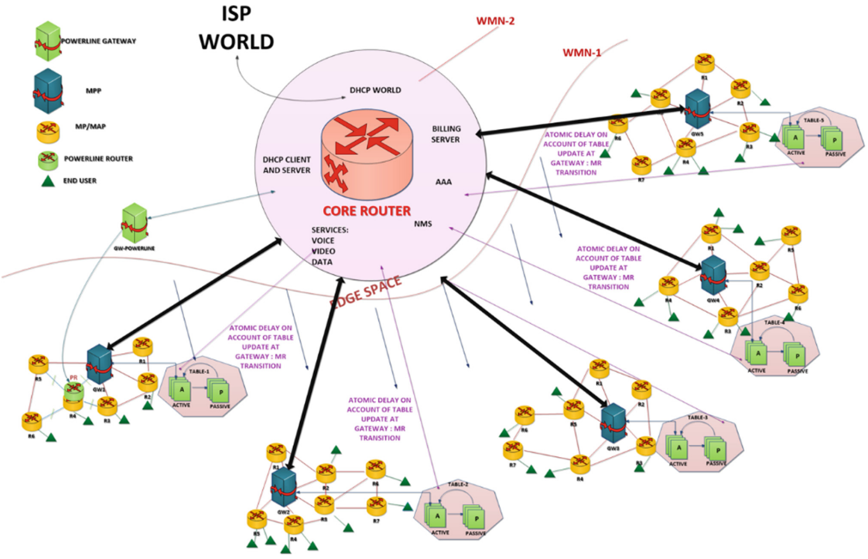

Fig. 2. Optimized 802.11s network with QoS guarantees. (Color figure online)

Figures 1 and 2 give an idea of the topology of Conventional WMN and the WMN obtained after applying the proposed LMS. Figure 1 presents a generic 802.11s network as defined by the IEEE. One may observe that a generic 802.11s network has

Mesh Portal Points (this is IEEE 802.11s term for gateway) shown as boxes in Fig. 1, connect to the Internet through a high bandwidth wired connection. On the other side these MPP are connected to the Mesh Points (this is IEEE 802.11s term for MRs) and Mesh access Points (this is IEEE 802.11s term for those MR which also cater to Wi-Fi clients or end devices) through wireless links. The green color triangles denote the end user devices which can be laptop, smart phone or any other Wi-Fi enabled service. The concept here is to show that the generic IEEE 802.11s can connect distant devices to Internet through multiple wireless hops thereby increasing the coverage area of the Wireless Local Area Network. This property allows for installation of city wide networks.

Figure 2 presents the proposed architecture which attempts to optimize the existing 802.11s mesh network so as to accommodate diverse networks with QoS guarantees. Unlike Fig. 1; Fig. 2, has well defined groups of Mesh Points/Mesh Access Points. In this paper Mesh Points/Mesh Access Points will be collectively called the MRs. Each group has been provided a dedicated Mesh Portal Point (GW) to serve. Therefore these are called the partitions. In a similar extension these partitions can also be thought of composed of other networks also. This network is generically denoted as WMN.

3 Optimizing the IEEE 802.11s WMN

Partitioning of WMN such that each GW has well defined service set, results in better throughput [4]. But since these multiple GWs have their own service set therefore they cannot serve other MRs outside their service set even when they are idle and other GWs are overloaded. This results in unfair scheduling of multiple GWs [20]. In order to have efficient use of all resources and to prevent overloading and drop in QoS it is suggested to manage load amongst the GWs. Load management will be dealt with in later sections. The next section describes the partitioning of WMN.

3.1 Partitioning the WMN

The initial arrangement of partitioned WMN, obtained after applying the modified Ciarlet and Lamour algorithm [9], is taken as reference point. This partitioning of network can be done in planning and deployment phase. Partitioning is done over Graph model of WMN wherein the MRs represent the vertices (nodes) of the graph and wireless links between the vertices represent the graph edges [12]. Network is partitioned by performing graph walk using Breadth First Search [12]. This procedure is called ‘node marking’. Node marking has been explained in detail in [10]. An innovative method of node marking especially defined for wireless networks is available in [10].

Every time a partition is to be created the graph walk begins from an unmarked GW node and ends when fair number of nodes required to create a partition are marked and collected. These MRs along with the GW constitute one partition. This procedure is repeated for every unmarked GW till all the GWs have formed a partition with required number of MRs. This greedy procedure for partition formation is summarized in next paragraph.

Let total number of partitions required be k, number of nodes per partition will be $n_i = \lfloor \frac{n}{k} \rfloor$, where n_i is number of nodes in partition i then, $n = \sum_{i=1}^{k} n_i$. In order to differentiate GW nodes from ordinary nodes, they are pre-marked and are kept in a list of GWs (GW_list). Let V_i be the set of nodes in i^{th} partition (analogous to the ordered set S in NMA). The greedy procedure for partition formation is summarized as follows

1. Select a GW node (already marked in the adjacency matrix) from the GW_list
2. Create this node as start node for the partition in question
3. Accumulate descendants of the GW for the partition in question using the NMA
4. Stop if total number of accumulated nodes = $\lfloor \frac{n}{k} \rfloor$

Detailed algorithm along with pseudocode and comparison study is available in [4]. After partitioning the network is booted up and the GWs start processing network data packets. The normal functioning of WMN continues until one of the GWs gets congested and overloaded. In such a condition load sharing has to be performed. In coming sections we first define congestion of GWs and derive a mathematical equation to define overloading of GWs. Thereafter we present a load management scheme based on load sharing among the GWs.

3.2 Load Management Scheme (LMS) Among the Partitions of WMN for QoS

In this section a LMS is devised through which the GWs can be utilized efficiently and overloading among the GWs can be prevented. The LMS is aimed at maintaining nominal load for GW. Each GW has processing load of the packets being directed to it through the MRs which are assigned to its partition. Initially a steady state load (SSL) condition is derived. If this condition is violated then it implies overloading of one or more of the WMN GWs. Load management is done by identifying and reducing the load of the overloaded GWs by shifting its MRs to a neighboring partition having GW with lesser load. The next subsection defines process for computing load of each partition.

Load Monitoring and Overload Condition

Following assumptions on load and traffic will apply for the rest of the paper

- **Full Coverage:** The term full coverage means that all the MRs must be served by a GW. This requires each MR to be assigned only one GW. Further no MR should be left isolated so that no GW is assigned to it. The assigned GW is treated as the default GW by the MR and it routes all its traffic through it until the assigned GW gets overloaded. On overloading of the assigned GW, the MR is assigned another less loaded GW by the proposed LMS. But at no point, is there a situation where an MR is not covered by a GW.
- **GW Throughput:** All GWs can process data as per their maximum throughput W_{gw}, defined by the Eq. (2) in next paragraph. When the MRs and their traffic demand increases, the corresponding GWs become overloaded. In such circumstances, the proposed LMS moves some MRs from the overloaded GW partition to another partition which has GW with a load lesser than its capacity.

- **MR Throughput:** Similar to the GWs, even the MRs have a maximum throughput. As indicated in Fig. 3, all the MRs direct their local traffic as well as the relay traffic towards their assigned GWs. Since the MR has multiple paths to reach the GW, it can always select the best alternative with the help of Air Time Link Metric (ATLM) computed by the path selection process of the routing protocol. In case a particular MR gets congested and overloaded, it need not accept the relay traffic. In such a situation, the relay traffic can be rerouted through other MRs having better ATLM. The process of reducing congestion at the MRs is far simpler as it is an inbuilt mechanism by which packets are not routed through a neighbor if its ATLM is poor. MR congestion being a localized problem can be solved by the underlying path selection mechanism easily and does not significantly disturb the mesh traffic and its topology.

It is assumed that the NICs associated with each of the *m* edges of the network graph, denoting wireless links can have separate transmit and receive frequency bands of operation. The physical location of a vertex (GW or MR) $v_i \in \mathcal{V}$ is static after deployment and its co-ordinates are denoted by (X_i, Y_i). Additionally each vertex $v_i \in \mathcal{V}$ is connected to a power supply which is not subjected to power constraint. Representation of number of non-interfering channels through a notation in set theory can be written as Eq. (1)

$$CH = \{1, 2, \ldots, c\} \tag{1}$$

Where *c* is the number of non-interfering channels in the wireless system which varies from one wireless standard to another (for IEEE 802.11b value of $c = 3$). Two vertices (MRs or GWs) are connected by an edge if and only if they are within the transmission range of each other and they can communicate on the same channel.

Capacity of a GW: *In a WMN, the maximum capacity of a GW is*

$$W_{gw} = \sum\nolimits_{i=1}^{|\rho(v_{gw})|} w_i \tag{2}$$

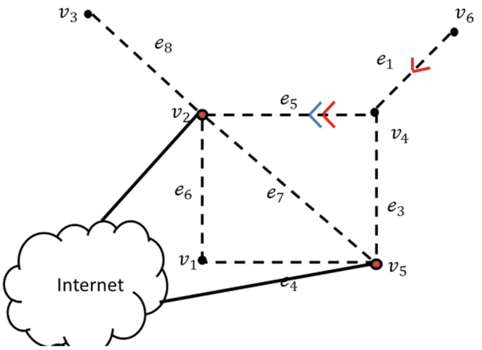

Fig. 3. A WMN graph and its traffic. (Color figure online)

Where, w_i bits/s is the data rate of channel $i \in CH$ and $\left|\rho(v_{gw})\right|$ is number of wireless interfaces configured on a GW denoted by $v_{gw}\,\forall\, gw = 1 \cdots k$. This is because at a given time slot, there are at most $\left|\rho(v_{gw})\right|$ (where, $\left|\rho(v_{gw})\right| \leq c$) interfaces in a GW that can simultaneously transmit and receive data packets to/from its neighbouring MRs.

Given an MR $v_i \in \mathcal{V}$, its traffic may include two parts as shown in Fig. 3.

- Local Internet traffic: This traffic is generated by various mobile devices such as laptops and smartphones which use the MRs as AP.
- Relayed Internet traffic: This is the traffic that is generated by other MRs which are further away from the GW as compared to MR v_1. Such MRs route their traffic through v_1 thereby adding to its local load.

In Fig. 3, all traffic is directed towards the GWs v_2 and v_5. For example MR v_6 can send its traffic to Internet through GW v_2 via MR v_4 considering it as an optimal path. In such a case traffic at MR v_4 will be its local Internet traffic which is depicted with a blue color in addition to relay traffic of v_6 depicted in red color.

Therefore the *bandwidth demand* d_i in terms of the local and relay traffic [13–16] is given by

$$d_i = local(v_i) + relay(v_i) \tag{3}$$

And *load on a GW* denoted by R_{gw} is defined as the current processing requirement of the GW and can be computed by Eq. (4).

$$R_{gw} = \sum\nolimits_{i=1}^{p} d_i \tag{4}$$

Where p is number of MRs assigned to the GW (number of MRs within the partition to which the GW belongs) and d_i is the bandwidth demand at MR v_i.

Derivation of Supergraph for Load Monitoring in a WMN

- A threshold load equation has to be derived which if violated results in invocation of the load sharing process. This section defines a structure called Supergraph from the graph model of WMN to monitor the load state of WMN. The term Supergraph is derived from the fact that it is a graph derived from subgraphs of partitioned WMN.
- A *Supergraph* of $\mathcal{G}$ is denoted by $\mathcal{G}^2$ because it can be perceived as a second order graph of $\mathcal{G}$. Formally a super graph $\mathcal{G}^2\left(\mathcal{V}^2, \mathcal{E}^2\right)$ is defined as a graph with set of vertices $\mathcal{V}^2$ which represent each partition (sub graph) and a set of edges $\mathcal{E}^2$ such that an edge uv $\in \mathcal{E}^2$ if and only if partitions u and v are connected to each other by at least one edge.

The process to derive a Supergraph is explained by using the partitioned WMN of Fig. 4. The WMN of Fig. 4 has five GWs and therefore five partitions are created around these five GWs. This implies that the Supergraph of Fig. 5 will have five vertices corresponding to the subgraphs $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_5$. Therefore $\mathcal{V}^2 = \{\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_5\}$ and there will be 6 edges corresponding to $\mathcal{G}_1\mathcal{G}_2, \mathcal{G}_1\mathcal{G}_3, \mathcal{G}_2\mathcal{G}_3, \mathcal{G}_2\mathcal{G}_4$, $\mathcal{G}_3\mathcal{G}_5$ and $\mathcal{G}_4\mathcal{G}_5$ since these partitions are connected to each other by one or more edges.

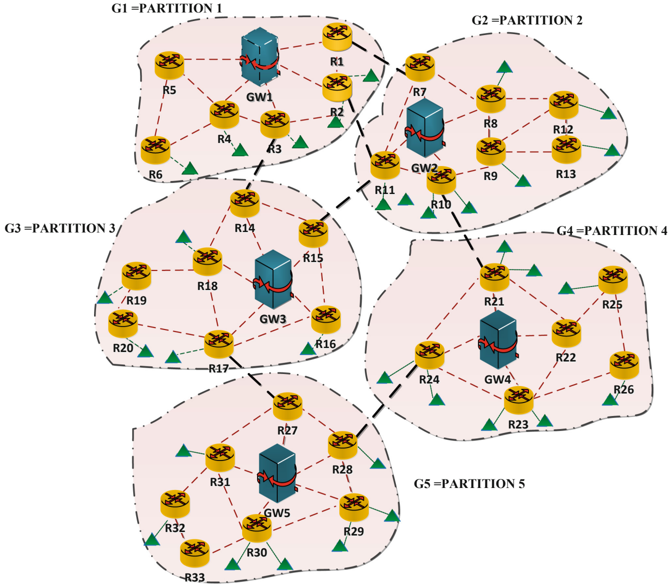

Fig. 4. A partitioned WMN. (Color figure online)

Whereas, there will be no edges connecting $\mathcal{G}_1\mathcal{G}_5$ and $\mathcal{G}_2\mathcal{G}_5$ because they do not have any edges between them. Therefore $\left|\mathcal{V}^2\right| = 5$ and $\left|\mathcal{E}^2\right| = 6$ and the resultant Supergraph is shown in Fig. 5.

Only the MRs which are in the communication range (denoted by a connecting edge) of a neighbor partition can be transited. In this case, if partition 5 gets overloaded than it can either transit R_{27} to partition 3 or R_{28} to partition 4. Since graph model for load management involves transition of traffic bearing MRs, these MRs are presumed to be active. But before performing load sharing, an overload condition has to be defined. Also to avoid chain transitions for load sharing a stability condition has to be defined. If these conditions are satisfied then only the MRs can be transited to neighborhood. Next sections present these conditions.

It can be noted in Fig. 4 that the actual partitions represent their connectivity using red color edges which represent various wireless links whereas the black color edges represent the connectivity between the partitions. These are virtual links which are further used to create the Supergraph of Fig. 5. It may be noted that each vertex V_i^2 of the supergraph is formed by contraction of a subgraph $\mathcal{G}i$ which has n_i number of nodes. For the WMN and its Supergraph corresponding to Figs. 4 and 5, it may be seen that $\mathcal{G}_1$ has $n_1 = 7$ nodes (1 GW and 6 MRs), similarly $\mathcal{G}_2$ has $n_2 = 8$ nodes (1 GW and 7 MRs) and so on.

Properties of Supergraph

The following properties can be derived for the Supergraph of a partitioned WMN

> **Property I**: A WMN with k GWs will have k partitions and therefore number of vertices in Supergraph $\mathcal{G}^2$ of partitioned WMN will be k.
>
> **Property II**: If each partition is represented as a node and this node is of degree $k-1$, then $\mathcal{G}^2$ graph will be a complete graph[1].
>
> **Property III**: $\mathcal{G}$ is a planar graph and $\mathcal{G}^2$ is its Supergraph, then if $\mathcal{G}$ is planar then $\mathcal{G}^2$ will also be a planar graph
>
> **About property III**: By contradiction let us assume that $\mathcal{G}^2$ is non planar. Then $\mathcal{G}^2$ will have intersecting edges. Since $\mathcal{G}$ is contracted to form $\mathcal{G}^2$, therefore this implies that $\mathcal{G}$ also has intersecting edges. Hence $\mathcal{G}$ is *non-planar*. Since $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is planar, therefore $\mathcal{G}^2$ is also planar.

Note: As a consequence of property II one may decipher a complete graph like K_5 graph (non-planar) as a resultant Supergraph. Property III provides justification on this.

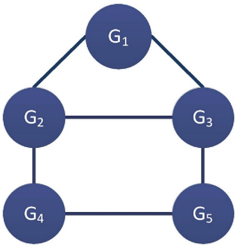

Fig. 5. Supergraph of the WMN of Fig. 4.

SSL and Stability Conditions on Supergraph for Load Sharing

In this section a condition is derived to monitor load imbalance in WMN. This condition is called the SSL condition. Later in this section a stability condition is derived to check if the load sharing can be performed without making the WMN unstable. Suppose $Q(\mathcal{G}_i)$ be the generic service limits[2] defined on subgraph $\mathcal{G}_i$. The generic service limit can be any parameter which might be requiring monitoring and control. The word generic is suggestive of using any threshold value based on formula which can be derived for a combination of various QoS parameters. The floor and ceiling do not imply the strict mathematical operation but these operators indicate those cases wherein the QoS parameters are specified within a range. In such case, the higher range is indicated by the ceiling operator and lower range is indicated by the floor operator. Although load on GWs is one such parameter which is mainly considered in this paper,

---

[1] A complete graph is defined as a simple graph which has connecting edge between all possible pair of vertices.

[2] Generic Service Limit can be QoS with respect to the network under consideration.

but network designers may like to derive some other complex parameter. This is the reason why the discussion on parameters are kept as generic as possible. In this paper the service limit is assumed as the processing capacity of a GW (Eq. 1).

To derive the SSL condition it is needed to define the binary limits on $Q(\mathcal{G}_i)$. The upper and lower limits are also suggestive of defining upper and lower service value which can be derived for a combination of various QoS parameters. Let U_i be the upper service limit where, $U_i = \lceil Q(\mathcal{G}_i) \rceil$. In this case the **maximum capacity** of GW as per Eq. (1) is assumed to be the upper service value. Let L_i be the lower service limit where $L_i = \lfloor Q(\mathcal{G}_i) \rfloor$. **Lower processing limit** can be defined by network planner on the basis of the minimal processing load within the network. This load could comprise of the minimum network management traffic, back-end traffic or protocol related traffic. Mainly it is that processing load of the network which is not generated by end-user. This means that if any vertex of the Supergraph begins to operate at U_i then the GW in partition $\mathcal{G}_i$ must transit some of its MRs to GW of a neighbouring partition $\mathcal{G}_j$ which has operating load of L_j (lower limit of load). But this transition of MRs cannot be done continuously or else it will be an overhead on the system. Therefore a threshold load condition has to be established on the WMN. This is called the *SSL condition* for load monitoring. If this condition is violated then the load sharing process has to be invoked.

SSL Condition for Load Monitoring
Let the existing (present) demand of a partition be denoted by R_i. The inequality $\|R\|^2 \leq \sum_{i=1}^{k} L_i U_i$ must hold true for a WMN to work at the nominal load

Derivation: As it is evident from earlier explanations that $L_i \leq R_i \leq U_i \, \forall \, \mathrm{i} = 1 \cdots k$ must hold true for all the GWs to work at a nominal load. This implies that for nominal load condition the average[3] demand for all k partitions should satisfy the following inequality

$$R_1^2 \leq L_1 U_1 \tag{5}$$

$$\begin{matrix} R_2^2 \leq L_2 U_2 \\ \vdots \end{matrix} \tag{6}$$

$$R_k^2 \leq L_k U_k \tag{7}$$

Summing up the load of individual partitions will give nominal working load for the whole WMN

$$R_1^2 + R_2^2 + \cdots + R_k^2 \leq L_1 U_1 + L_2 U_2 + \cdots + L_k U_k \tag{8}$$

$$\Rightarrow \|R\|^2 \leq L_1 U_1 + L_2 U_2 + \cdots + L_k U_k \tag{9}$$

---

[3] The geometric mean is more appropriate than the arithmetic mean for describing proportional growth like increasing bandwidth demand of Internet [17].

Where, R is the average working load of the whole mesh (WMN). Therefore for SSL operation, nominal load of mesh must satisfy the following inequality

$$\|R\|^2 \leq \sum\nolimits_{i=1}^{k} L_i U_i \tag{10}$$

Therefore the core router which is sending and receiving traffic to the WMN keeps a check on R and the moment value of R violates the Eq. (10), it invokes the load sharing process. In next section, a detailed simulation study is performed to ascertain how far the WMN can continue to remain in steady state with varying load demands. Once the SSL condition to monitor the WMN overloading is determined, a formulation needs to be derived to monitor the WMN systemic stability during MR transitions from the perspective of load sharing. By instability it is meant that an MR transition should not trigger a chain reaction of transitions (ping pong effect). The stability condition established in the next section is used to avoid such a situation.

Bipartite Graph of the Supergraph of WMN

Before presenting the stability condition an overview of bipartite graphs is presented. Formally, a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is said to be bipartite, or 2-partite, if its vertex set can be partitioned into two different sets $\mathcal{V}_1$ and $\mathcal{V}_2$ such that every edge of the graph connects one vertex in $\mathcal{V}_1$ to a vertex in $\mathcal{V}_2$. The two sets $\mathcal{V}_1$ and $\mathcal{V}_2$ are called partite sets.

In simpler words a graph $\mathcal{G}$ is called bipartite if its vertex set $\mathcal{V}$ can be decomposed into two disjoint subsets $\mathcal{V}_1$ and $\mathcal{V}_2$ such that every edge in $\mathcal{G}$ joins a vertex in $\mathcal{V}_1$ with a vertex in $\mathcal{V}_2$ and none of the edges in the graph connect vertices of the same set.

Theorem for Stability of Load Sharing in WMN

This theorem is called as the **stability theorem** in this paper. It is stated as:

Stability Theorem: *A partitioned WMN can share load by transiting MRs from one GW to another if its Supergraph $\mathcal{G}^2$ is bipartite between S and T where S is the set of nodes in $\mathcal{G}^2$ operating at nominal load (load which is greater than or equal to L_i but less than U_i) and T is the set of nodes in $\mathcal{G}^2$ operating at load which is greater than or equal to the upper load limit U_i.*

Proof: Let, $\mathcal{G}^2$ be a Supergraph whose nodes represent WMN partitions. Let the overloaded nodes (partitions/GW) belong to set T and all other nodes belong to set S. Load sharing by MR transition from one partition to another partition can happen if and only if the nodes of Supergraph $\mathcal{G}^2$ form a bipartite graph with S and T.

This theorem is proved as follows:-

Consider that $\mathcal{G}^2$ is non-bipartite for S and T. This implies $\{\mathcal{G}_1 \cdots \mathcal{G}_k\} \in \mathcal{G}^2$ is non-bipartite for S and T. This means that at least one pair of nodes $\mathcal{G}_i$ and $\mathcal{G}_j$ is adjacent to each other and both $\mathcal{G}_i$ and $\mathcal{G}_j$ belong to either S or T.

Case 1. $\{\mathcal{G}_i, \mathcal{G}_j\} \subset T$.

Node transition from one overloaded node to another overloaded node cannot happen, thus proving the theorem.

Case 2. $\{\mathcal{G}_i, \mathcal{G}_j) \subset S$.

Since both the adjacent partitions are less loaded, they will not resort to the process of transition of MRs as a nominally loaded partition does not have a need for load sharing (MR transition).

This implies that edges connecting vertices belonging to the same set are trivial for load sharing.

3.3 Matrix Model of Partitioned WMN and Its Implementation to Validate Integrity of WMN

The authors have presented a matrix model of WMN in their paper [18]. In this model they represent the partitions in form of neighborhood matrix. The matrix not only provides a mathematical validation equation to perform check on lost MRs and hanging MRs of the network, but also provides a basis to represent MRs which have been moved to other networks for temporary offloading. The validation equation is provided on basis of the graph equation as explained in [18].

A brief explanation in to the matrix equation for the partitioned graph is provided as follows. For any connected self-loop free Graph $\mathcal{G}$ the property $\mathcal{B}\mathcal{C}^T = \mathcal{C}\mathcal{B}^T = 0\,(mod\,2)$ must be true [12]. Where $\mathcal{B}$ is the incidence matrix of graph $\mathcal{G}$ of WMN and $\mathcal{C}$ is the cycle matrix of the graph. Therefore for all the connected subgraphs the same property must be true. Extending this property further we can state the following.

For a self-loop free planar graph $\mathcal{G}$ *with partitions* $\mathcal{G}_1, \mathcal{G}_2 \ldots \mathcal{G}_k$ *the partitioned graph is consistent with the original graph if and only if*

$$[\mathcal{B}_i] * [\mathcal{C}_i]^T = 0\,(mod\,2)\,\forall\, i \in [1 \ldots k] \tag{11}$$

Where $\mathcal{G}_i$ *represents the* i^{th} *partition* $\forall\, i = 1, .., k$.

Using Eq. (11) the partitioned graph can be validated for integrity (absence of lost MRs or redundant MRs). This equation is applied every time load sharing happens to ensure that there are no lost MRs or redundant MRs. This is important requirement because every time a MR is moved out of one partition to another there can be chances of lost MRs especially if hand off and hand over from one GW to another does not happen in a seamless manner.

4 Performance Results and Discussion

This section is aimed to analyse the performance of the proposed Load Management Scheme (LMS). In the previous chapter a simulation model of the proposed LMS was developed using MATLAB, and Simulink blocks. This model is used in this chapter to analyse the performance of the proposed LMS. Various test cases for performance analysis are created to compare the performance of the WMN with the proposed LMS and the conventional WMNs with no load management feature. Throughout the simulation process the simulation parameters are kept as per Table 1.

4.1 Analysis of the IMW

This section highlights the importance of the Intelligent Middle Ware (IMW) to invoke the load sharing and internetworking processes for meeting the increased bandwidth demand on a WMN. IMW is the software executing at the core router or the Network Control Centre. This software is supposed to implement the proposed LMS by monitoring the network for overloading. This is done by checking the SSL condition and then proceeding for load sharing if the SSL condition is not satisfied. For the purpose of simulation, the bandwidth demand is increased at randomly chosen MRs. The parameters assumed for simulation are provided in Table 1. As per the MATLAB Simulink model all the packets are routed through the core router to the GWs of the WMN.

Table 1. Simulation parameters.

Parameter	Value
Number of GWs	1–15
Number of MR	Varying from 100–300
Maximum number of mesh clients	250
Mean packet arrival rate	0.01 s (100 packet/s)
Mean hop delay	0.01 s
Flow rate	Markov distribution
Packet size	64 bytes
Core router capacity	100 Mbps
GW Capacity	2 Mbps
Transmission range of MR & GW	250 m
Carrier sensing range	550 m

As shown in Table 1, the increase in bandwidth demand is shown in steps ranging from 0 to 10 Mbps resulting in 10 ranges of bandwidth demand. Each of these was divided to 10 more sub ranges. In all 100 simulations have been performed. In each of these simulations, different MRs are chosen to generate bandwidth demand. Then a count is recorded for number of time the SSL condition is violated for a particular level of bandwidth demand.

The parameters considered for SSL computation are,

U_i = 2 Mbps (GW Capacity)

L_i = 1 Mbps (minimal load)

Then as per SSL condition of Eq. 10,

$$R^2 \leq (k \times U_i \times L_i) = 10 \quad \text{(where k is number of GWs)} \tag{12}$$

Therefore nominal load should satisfy the following inequality

$$R \leq 3.2 \tag{13}$$

It can be seen in Fig. 6 that after the bandwidth demand crosses the 3 Mbps range the WMN load becomes unstable and there is need for load sharing.

Violation of the SSL condition implies overloading of the WMN. Similarly, bipartite reducibility of a graph results in load sharing. Therefore it can be assumed that the numbers of times the Supergraph could be reduced to its bipartite form indicates the number of times load sharing is invoked. In Fig. 6, the X-axis denotes load in Mbps and Y-axis denotes count of number of times an event is invoked for 10 simulations.

Next Fig. 7 compares the throughput obtained in a conventional WMN with the throughput obtained in a WMN with the proposed LMS feature added to it. The results of Fig. 7 are obtained when the throughput of a WMN is compared for the following cases.

- Conventional WMN (no partitioning and load sharing)
- WMN with partitioning
- WMN with partitioning and load sharing.

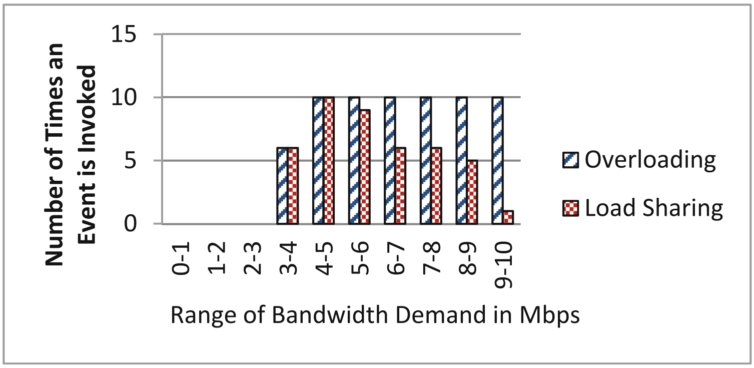

Fig. 6. Frequency of overloading and load sharing with dynamically changing load.

The simulations have been performed keeping the total number of MRs fixed in each of the WMN scenarios but the number of GWs has been changed. This helps to study the effect of increasing the number of GWs, keeping the number of MRs constant.

Figure 8 depicts throughput improvement in WMN with different number of MR and GWs. The first two bar graph of Fig. 8 reveal that a WMN with the same number of GWs but with a relatively larger total number of MRs, exhibits better throughput performance [19].

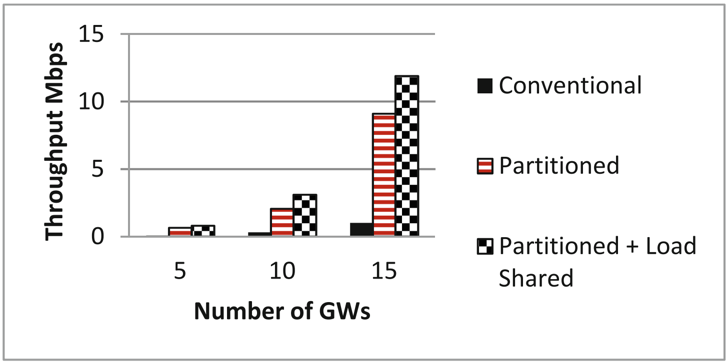

Fig. 7. WMN maximum throughput improvement at each stage of LMS for a 100 MR WMN scenario with increasing number of GW.

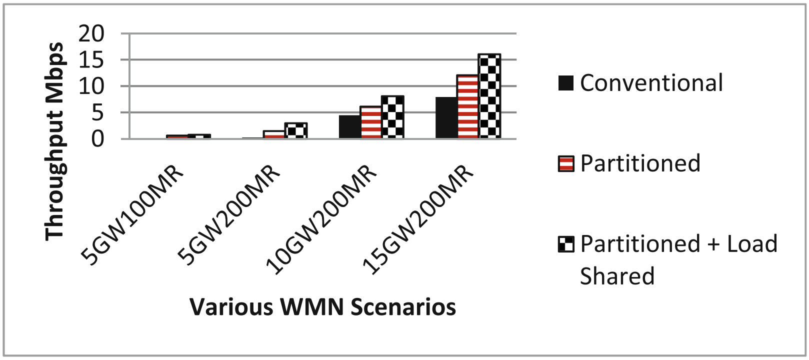

Fig. 8. Throughput performances for each stage of LMS for different WMN scenarios.

4.2 Analysis of Average Packet Transfer Delay in WMN

Since the LMS proposed in this paper works towards reducing the congestion of GWs also, a study of the average packet transfer delay is very important to assess the performance of the LMS.

For the simulation, a WMN with a total number of 100 MRs is considered with the number of GWs being changed from 5 to 10. Figure 9 depicts the average packet delay for the scenario associated with the above stated features. The results are compared with a WMN with no partitioning. The next section studies the effect of the number of packets dropped during each phase of LMS.

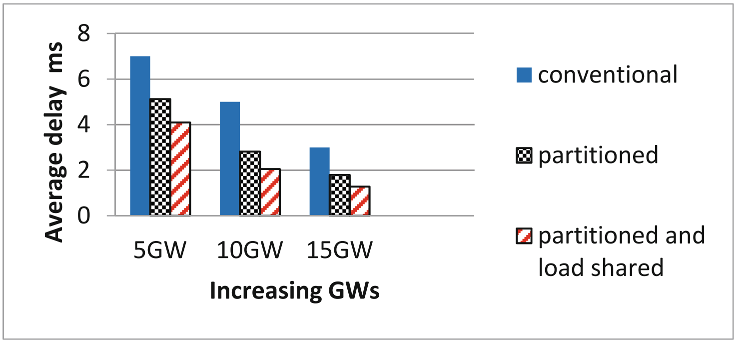

Fig. 9. Average packet delay at each stage of the proposed LMS.

4.3 Analysis of Packet Loss

Packet loss is one of the parameters affecting the QoS. One of the major causes of packet loss in wireless networks is channel congestion. Since the proposed LMS works on relieving the congestion of GWs, it results in a reduced packet loss thereby leading to an overall improved performance of a WMN. For the simulation, a WMN with GWs and MRs is considered. The results of Fig. 10 depict the variation in the packet loss as a function of the elapsed time in 5GW 100 MR WMN. In the simulation, the number of packets lost after a time window of 300 ms. Figure 10 depicts the results obtained for a conventional WMN. The number of packets dropped after 300 ms was found to be 78. Next Fig. 11 captures the packet loss attributed individually to the two basic processes of LMS. For the simulation results showed in Fig. 11, the total number of MRs remain constant at 100 while the number of GWs is varied from 5 to 15.

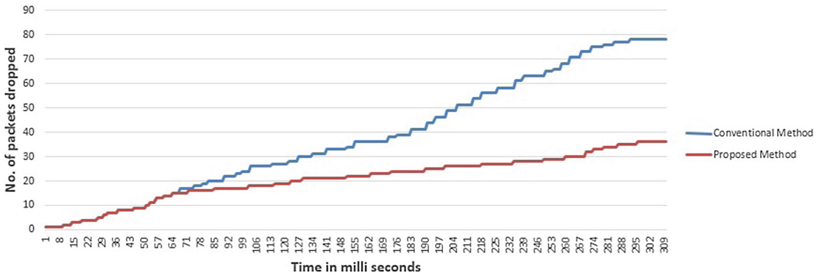

Fig. 10. Packets dropped in 5 GW 100 MR WMN.

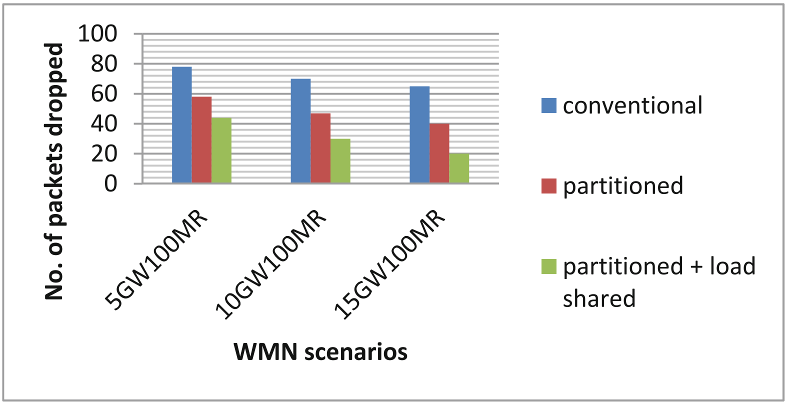

Fig. 11. Number of packets dropped at each stage of the proposed LMS.

5 Conclusions

This paper presented a load management mechanism within a partitioned WMN. Simulation studies indicate that when load on a WMN increases beyond 60% of its capacity, the frequency of occurrence of load sharing increases by 60% and frequency of internetworking increases by 40%. On comparing throughput of a non-partitioned WMN with throughput of a partitioned WMN, it was found that the partitioned WMN showed an average increase of 100% in throughput. Inclusion of the load sharing feature to the partitioned WMN resulted in a throughput improvement of 30%. It was also observed that after providing a conventional WMN with the features of the proposed LMS, on an average, there was a decrease of 22% in the average packet delay. Also there was a decrease of 90% in the number of packets dropped.

The performance analysis depicts that a WMN can provide good throughput if

- it is well defined with respect to the serving GW
- if the serving GWs are monitored for load.

The load sharing further improves the throughput. Such a WMN can be effectively utilized to create city wide Wi-Fi access as well as it can be leveraged to use as an intermediate access networks to couple with the 4G mobile networks to achieve the M2M communication. Besides performing the load management the paper also proposed a method to preserve the integrity of partitioned WMN such that it can be used further for connecting to the Internet of Things as access network. The matrix based platform which is proposed in this paper allows representing disjoint partitions. This representation is leveraged to accommodate other subnets which might be connecting to the WMN access network as part of M2M connectivity infrastructure.

6 Future Work

The paper suggests using the WMN and its subnet representation in form of matrix to address the diversity of Internet of Things (IoT) networks. In this architecture the backhaul for MPPs to the core can be any WAN. Middle mile is achieved through the WMNs which are provisioned with an additional feature of load sharing. Finally the last mile comprise of short range IP and non-IP access networks. In future work non-IP networks should be integrated to the WMN through the matrix based mathematical model. The advantage of such a structure is most appropriate for use with the 4G mobile networks where it gets really difficult to connect sensor based massive IoT networks to the Internet. The architecture should be modified to form a virtualized middle mile network provisioned with load management technique like the one proposed in this paper.

References

1. Institute of Electrical and Electronics Engineers: IEEE Standard for Information Technology–Telecommunications and information exchange between systems–Local and metropolitan area networks–Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 10: Mesh Networking, IEEE Std. 802.11s-2011, 10 September 2011
2. IEEE 802.11-2012: IEEE Standard for Information technology–Telecommunications and information exchange between systems Local and metropolitan area networks–Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications (2012). https://standards.ieee.org/standard/802_11-2012.html
3. Nandiraju, D., Santhanam, L., Nandiraju, N., Agrawal, D.P.: Achieving load balancing in WMN through multiple gateways. In: Mobile Adhoc and Sensor Systems (MASS), pp. 807–812 (2006)
4. Pandey, S., Kadambi, G., Bates, S., Pande, V.: A load sharing and partitioning system for multihop wireless mesh network with multiple gateways. In: IEEE Conference on Open Systems (ICOS), 25–28 September 2011, pp. 369–374 (2011)
5. Sriram, L., Raghupathy, S., Karthikeyan, S.: Multi-gateway association in wireless mesh networks. J. Ad Hoc Netw. **7–3**, 622–637 (2009)
6. Tokito, H., Sasabe, M., Hasegawa, G., Nakano, H.: Routing method for gateway load balancing in wireless mesh networks. In: IEEE Eighth International Conference on Networks, ICN 2009, 1–6 March 2009, pp. 127–132 (2009)
7. Le, A., Kum, D., Cho, Y., Toh, C.: Routing with load-balancing in multi-radio wireless mesh networks. IEICE Trans. Commun. **92**(3), 700–708 (2009)
8. Bruno, R., Conti, M., Pinizzotto, A.: A queuing modelling approach for load-aware route selection in heterogeneous mesh networks. In: Proceedings of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks & Workshops, 15–19 June 2009, pp. 1–9 (2009)
9. Ciarlet Jr., P., Lamour, F.: On the validity of a front oriented approach to partitioning large sparse graphs with a connectivity constraint. Technical report 94-37. Computer Science Department, UCLA, Los Angeles, CA (1994)
10. Pandey, S., Pande, V.: A node marking algorithm for partitioning wireless mesh networks. In: IEEE Conference on Open Systems (ICOS), 25–28 September 2011, pp. 363–368 (2011)

11. He, B., Xie, B., Agrawal, D.P.: Optimizing the internet gateway deployment in a wireless mesh network. In: IEEE International Conference on Mobile Adhoc and Sensor Systems, MASS 2007, October 2007, pp. 8–11 (2007)
12. Deo, N.: Graph Theory with Applications to Engineering and Computer Science. Prentice Hall India, New Delhi (1974). Latest Edition 2000
13. Wolf, J., Hevkmuller, S., Wolfinger, B.E.: Dynamic resource reservation and QoS management in IEEE 802.11e networks. In: Proceedings of the International Symposium on Performance Evaluation of Computer and Telecommunication Systems, SPECTS (2005)
14. Wu, H.-T., Yang, M.-H., Ke, K.-W., Yan, L.: Enhanced QoS mechanisms for IEEE 802.11e wireless networks. In: World Academy of Science, Engineering and Technology, vol. 34, pp. 912–916 (2009)
15. Lin, Y., Wong, V.W.S.: An admission control algorithm for multi-hop 802.11e based WLAN. J. Comput. Commun. **31**, 3510–3520 (2008)
16. Lin, H.-T., Lin, Y.-Y., Chang, W.-R., Cheng, R.-S.: An integrated WiMAX/Wi-Fi architecture with QoS consistency over broadband wireless networks. In: Proceeding of the 6th IEEE Consumer Communications and Networking Conference, CCNC (2009)
17. Fleming, P.J., Wallace, J.J.: How not to lie with statistics: the correct way to summarize benchmark results. Commun. ACM **29**(3), 218–221 (1986)
18. Pandey, S., Pande, V., Kadambi, G., Bate, S.: Partitioning and internetworking wireless mesh network with wired network for delivery maximization and QoS provisioning. Procedia Technol. **3**, 18–29 (2012). ISSN 2212-0173
19. Pandey, S., Tambakad, V., Kadambi, G., Vershinin, Y.: An analytic model for route optimization in load shared wireless mesh network. In: Proceedings of the IEEE European Modelling Symposium (EMS), pp. 543–548, 20–22 November 2013
20. Bejerano, Y., Han, S.J., Kumar, A.: Efficient load-balancing routing for wireless mesh networks. IEEE J. Comput. Netw. **51**(10), 2450–2466 (2007)