CHAPTER 8: DYNAMIC LOAD BALANCING FOR IEEE 802.11E WLAN

8.1 NEED FOR LOAD BALANCING

WLANs are the predominant option for accessing the wireless broadband Internet. The wireless stations associated to a WLAN in Infrastructure mode as in figure 8.1 can normally move from one access point to another. These APs are interconnected by means of a shared Ethernet cable. Each access point covers an area, and its coverage area depends on the data rate and the location.

Figure 8.1 WLAN in Infrastructure Mode of Communication

The IEEE standard describes the channel access methods, the characteristics, bandwidth of physical layers and the operation modes. There is no specific standard to measure and to support the load balancing in wireless local area network [3]. Consequently some of the access points may provide service to number of wireless stations, while some other neighboring AP may provide service to only few wireless terminals.
In this section the load balancing problem of WLAN in infrastructure mode is analyzed and dynamic load balancing mechanism is proposed [3]. A decentralized approach used in [40], is mainly based on the concept of AP selection technique deployed by the wireless stations. It considers the number of wireless stations connecting to the AP and supports only best-effort traffic. This approach of balancing the load will consider more number of wireless terminals with lower data rate than the terminals requires transmitting high data rate signals. As a result of this, the channel resource is not efficiently utilized and also degrades the WLAN performance.

With a suitable load balancing technique, the AP can associate with more number of wireless stations for providing channel access. This enables to shares its total bandwidth among all other wireless stations. The access point can accommodate an additional load, as long as the average load does not surpass its transmitting power. If the load on an AP attains to its maximum capacity, the performance does not increase further. This is known as overloading of a QAP. A simple approach to overcome these overloading issues is, to share the load of currently heavily loaded APs with the neighboring APs.

The proposed approach balances the load among all APs and also efficiently uses the radio resources. As in figure 8.2, there are seven wireless stations, associated to AP1, while only three and five are associated with AP2 and AP3. The reason for this asymmetry is the stations very close to the AP will have strongest SNR. This allows more number of stations to associate with these access points. Hence a load balancing mechanisms in a WLAN can balance the load among these APs with equal number of wireless stations associated with each of the APs as in figure 8.3. It results an improved performance of the WLAN from this symmetry.
8.2 LOAD BALANCING

Two main approaches which are to be considered to address load balancing problem are;

- The method to define and measure the load and load related metrics for an AP.
- The process to distribute the traffic load from the heavily loaded access point to an access point loaded minimum and balance the load among all access points.

The algorithm proposed here distributes the wireless stations among all APs connected to a common Ethernet backbone. For this, the wireless stations scan the available channels of each AP, to discover the APs with sufficient bandwidth. The algorithm first verifies number of wireless stations connected to an access point and their data transmission rate for balancing the load. A wireless station performs probing process as in figure 8.4, by an active scan or a passive scan. During active scan it transmits a probe request frame with channel access parameters. If the AP is able to fulfil the channel access of the respective wireless station, it responds with a probe response message to it; with accept or reject information. Also the admission
control mechanism implemented in the AP prioritizes the real-time traffic to access the channel.

![Diagram of Message Communications between Wireless Station and the AP](image)

**Figure 8.4 Message Communications between Wireless Station and the AP**

In case of passive scan, a wireless terminal receives the beacon frames, which is broadcasted periodically by an AP. Depending on the information gathered from the probe response; the wireless terminal chooses an access point to connect to. The channel scanning provides the information about corresponding received signal strength indicator (RSSI) value, for each AP-to-wireless station links. Comparing these RSSI values, the wireless stations then selects the best AP, which has the greater RSSI value.

If the wireless station selects an AP with large RSSI value, using admission control procedure, it can select the traffic load with high data rate requirement to access the wireless channel. This process can optimally utilize the wireless channel. Therefore the access point can support larger traffic, and provide better throughput.
8.2.1 Load Aware AP Selection Metrics

A fundamental technique in load aware AP selection process is to measure the workload of APs. Let us consider that there are X wireless stations and Y APs. A wireless station can only select an access point to connect to, and each access point accomplishes support to at least X wireless stations. Hence a perfect load metric of an AP is the number of wireless terminals connected with it. Generally a wireless terminal selects access point to associate, based on the RSSI value. It is used to measure the distance between the wireless stations from the AP. In free space the relationship between RSSI and distance is described as,

\[
P_i(d) = P_o - 20 \log_{10}\left(\frac{4\pi d}{\lambda}\right)\ dBm
\]

Where \( P_o \) is an empirical constant and for this set, it is considered as 31.0 dBm.

\[
\lambda = \frac{c}{f}\quad c = 3 \times 10^8 m/\sec, \quad f = 2.4\ GHz
\]

The association of an AP with a wireless station is defined by the RSSI value. A greater value of RSSI indicates a better transmission condition. Let,

- \( SN_i \) - is the number of wireless terminals connected with an AP, \( AP_i \)
- \( SR_i(j) \) - is the RSSI value between the access point \( AP_i \) and the wireless terminals \( WS_j \) when it is contending to access the access point.
- \( AR_{av} \) - is the average RSSI value in the set \( S_i \)

Let \( S_i \) denote a set of stations connected to an access point \( AP_i \) and let \( SR_i \) denote its RSSI value. When \( AP_i \) accepts the traffic sent by a new wireless terminal \( WS_j \) the average RSSI value, \( AR_{av} \) in set \( S_i \) is,
\[ AR_{av} = \frac{\sum_{j \in S_{i}} SR_{i}(j)}{SN_{i}} \]

In the above equation \( SN_{i} \) is the total number of STAs connected in the set \( S_{i} \). Let the variations in average RSSI is \( AR_{av} \) and the number of wireless terminals in the set \( S_{i} \) is; \( VAR_{av} \) and \( VSN_{i} \) respectively.

8.3. DYNAMIC LOAD BALANCING ALGORITHM (DLBA)

In DLBA for a given WLAN having \( X \) wireless stations and \( Y \) number of APs, each wireless station will choose an appropriate AP to associate. Using the proposed algorithm the average RSSI value \( AR_{av} \) in a WLAN is maximized, and the variations in the average RSSI value \( VAR_{av} \), and the number of STAs \( VSN_{i} \), in a set \( S_{i} \) is minimized. To access an AP, a newly joining wireless station \( (WS_{j}) \) will perform an active scan by sending a probe request to each channel. This process evaluates all the links between \( WS_{j} \) and all the APs. The AP now records \( SR_{i}(j) \), as in figure 8.4, the RSSI value between the access point \( AP_{i} \) and the wireless station \( WS_{j} \). Once \( WS_{j} \) probes all the channels, it will select the access point with highest RSSI value to join.

Unfortunately this process of selecting an AP may cause serious unbalancing of load, because the wireless stations may be connected to a few APs while other access points continue to be idle. Thus the load is not equally distributed among each access point.

The wireless station \( WS_{j} \) can select the AP that maximizes the average RSSI value based on the following weighted parameters. These parameters of the wireless station
are defined as,

\[ W_i(j) = D_i(j) \times P_i(j) \]  (8.1)

Where \( D_i(j) \) is the difference between \( SR_i(j) \) and \( AR_{av} \)

\[ D_i(j) = SR_i(j) - AR_{av} \]

\( AR_{av} \) - is the average RSSI value of set \( S_i \) calculated by considering the effect of the contending wireless station \( W_{S_j} \) into the set \( S_i \).

If \( SR_i(j) > AR_{av} \); \( D_i(j) > 0 \)

This implies that the wireless station \( W_{S_j} \) has a positive contribution to the set \( S_i \).

On the contrary, \( SR_i(j) \leq AR_{av} \), indicates that, admitting a new wireless station \( W_{S_j} \) to the set \( S_i \) will degrade its average RSSI value. In equation (8.1) \( P_i(j) \) is the proportional weighted parameter, where,

\[
P_i(j) = \begin{cases} 
1 + \frac{AR_{av}}{SR_i(j)} & \text{if } D_i \geq 0 \\
1 - \frac{AR_{av}}{SR_i(j)} & \text{if } D_i \leq 0
\end{cases}
\]

However this mechanism still does not guarantee that all the access points share the load uniformly, i.e., some of the APs are associated with more number of wireless stations and others with less numbers of stations. This issue can be solved by carefully considering the RSSI value between the AP-to-wireless station links.

If the RSSI value of those channels is less than new \( AR_{av} \) the wireless stations are forced to change to another set \( S_i \) by hand-off process. More number of such hand-off processes will degrade the performance of the WLAN. To alleviate this, each
wireless station maintains a hand-off counter (HC) [15]. When a new wireless station joins, the existing wireless stations in the set will also get the probe request and compares its RSSI value with $AR_{av}$ value. If this value is less than the $AR_{av}$, its HC is incremented by one. Once the HC value reaches to a threshold value say $HC_{max}$, it will leave the current set $S_i$ to join the other set.

![Dynamic Load Balancing State Machine](image)

This information is updated in the load balancing module of each AP and broadcast to the common Ethernet backbone, and the wireless stations now can perform the joining process to the new AP. Meanwhile the HC value is reset to zero. This hand-off process occurs only when the load in an access point reaches the threshold value. As a result all the wireless stations are re-arranged into a relatively better condition and
balance the load in the WLAN. The above operation can be defined in a state machine with three different states as shown in figure

8.4. PERFORMANCE EVALUATION

The proposed DLBA is implemented in NS-2 and the performance is simulated. The wireless station and access points considered for simulation operate with IEEE standards 802.11b and 802.11e. The algorithm and the pseudo code is as shown in figure 8.6 and 8.7. A bandwidth of 2 Mbps and the simulation traffic considered are VoIP (voice), video and data. Different data rates are chosen based on the space between the AP and the terminals viz, 11Mbps, 5.4Mbps and 2 Mbps, for a distance of 25mts, 40mts and 60mts respectively. The other simulation parameters for the different ACs are in table 8.1.

<table>
<thead>
<tr>
<th>AC</th>
<th>AIFS</th>
<th>CWmax</th>
<th>CWmin</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>7</td>
<td>1023</td>
<td>31</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1023</td>
<td>31</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>31</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>15</td>
<td>7</td>
</tr>
</tbody>
</table>

The simulation is conducted by distributing the wireless stations uniformly and non-uniformly around the access points. A VoIP traffic (AC-VO) at 84Kbps, is produced by a fixed interval of 20 milliseconds and has a traffic load of length 0.208Kbyte is used as voice traffic. The video traffic (AC-VI) is at770Kbps, generated by a constant interval of 13.33 milliseconds and a fixed length of 1.28Kbyte. The traffic in AC-BK is 800Kbps, is produced at a fixed interval. The simulation output is obtained every
1200 milliseconds. From the simulation result, the throughput and delay variation for different traffic streams like VoIP, video and data is evaluated [53].

Figure 8.6 Flow Chart of the Dynamic Load Balancing Algorithm
Input: $\text{RSSI, AR}_w$, throughput, Probe request  
Output: True/false  
1. Channel = 1  
2. If received probe response from AP within timeout then  
3. Average Load in AP, $\text{AR}_w = \left( \sum \text{AR}_w \right) / n$  
4. If $\text{Rx}(y) \leq \text{AR}_w$ then  
5. select $\text{LessLoadedAP} \leftarrow \text{true}$  
6. $\text{Dx}(y) \leftarrow \text{AR}_w - \text{Rx}(y)$  
7. else  
8. If $\text{Dx}(y) > \text{Dx}(i)$ then  
9. reject $\leftarrow$ false  
10. else  
11. reject $\leftarrow$ true  
12. endif  
13. endif  
14. else  
15. If $t =$ timeout then  
16. reject $=$ true  
17. else  
18. repeat  
19. endif  
20. else  
21. If $\text{AR}_w \leq \text{AR}_w$ then  
22. $\text{HC} = \text{HC} + 1$  
23. If $\text{HC} \geq \text{HC}_{\text{max}}$ then  
24. Hand – Off $\leftarrow$ true  
25. else  
26. Hand – Off $\leftarrow$ false  
27. endif  
28. endif  
29. endif  
30. endif  
31. Return  

Figure 8.7 Dynamic Load Balancing Algorithm  

The simulation output obtained for non-uniform load conditions and for a fixed number of VoIP, video and data traffic is considered for performance evaluation. Throughput and delay performance is compared at different traffic load. The
performance of DLBA is verified with the nSTA method, which used number of active wireless stations sharing technique with each AP [1].

8.4.1 Simulation Result

8.4.1.1 Uniformly Distributed Wireless Stations

In this case, the wireless stations are uniformly distributed over the APs. Therefore there is a marginal difference in the performance among the two load balancing mechanisms. The throughput performance and delay variations for VoIP, video and data load are described as below. Figure 8.8 8.9 8.10 8.11, and 8.12 shows the average throughput performance and delay for VoIP, video and data load. Using the proposed dynamic load balancing and admission control mechanism, the real-time traffic has the better channel access. The higher performance is due to the priority based reservation system used in admission control mechanism for IEEE 802.11e. This retained the throughput optimally, for both real-time and BE services.

Figure 8.8 Average Throughput of VoIP Applications
In nSTA method, more channel access is provided for data traffic than real-time traffic. As a result the data transmission rate for high priority application is reduced, because the high priority traffic is allowed to access the channel only by those APs, which has higher data rate. This degrades the performance of real-time traffic. Therefore it results an inefficient utilization of the radio resources.
Figure 8.1 shows average delay for uniformly distributed wireless stations under each AP. Using the proposed algorithm the delay is maintained optimally, compared to nSTA method. In nSTA method the delay increased as the load increases.
8.4.1.2 Non-Uniform Distribution of Wireless Stations

The wireless stations in this case are non-uniformly distributed. When the traffic load is low there are no much changes in throughput and delay, both in proposed algorithm and nSTA methods. The proposed algorithm supports more real-time traffic. Figure 8.13, 8.14, 8.15, 8.16 and 8.17, shows that, the throughput and delay of the wireless channel is maintained, even if the traffic load is increased. This performance stability of the proposed algorithm is because, it supports the traffic with high transmission rate and hand-off those wireless stations to other APs, only after its load reaches to the threshold value. This provides a better resource utilization and load balancing.

In case of nSTA, if the traffic loads increase the throughput performance of VoIP and video signals decreases, and accordingly the delay increases, because nSTA method of load balancing does not support traffic prioritization. The high priority traffic may be blocked in a particular access point. Thus for high priority traffic there is a decrease in throughput performance. Also the stations with low transmission rate traffic, located nearer to the AP get better transmission rate than which is far from them. As a result, the average throughput of WLAN system using nSTA is degraded when the traffic load increases.

![Figure 8.13 Average Throughput of VoIP Applications](image-url)
Figure 8.14 Average Throughput of Video Applications

Figure 8.15 Average Throughput of Data Applications
The proposed algorithm delivers the load balancing and performs efficient use of wireless channel resources. Simulation result shows the performance improvement from 20% to 54% under different load conditions, compared to nSTA method in [1]. Considering this it is clear that the proposed algorithm outperforms nSTA; it can achieve better performance for the IEEE 802.11e WLAN.
8.5. SUMMARY

In the proposed dynamic load balancing algorithm which is running in each AP; exchanges the load information with other APs. It considers the current state in an access point to balance the load and its data rate at which a station can access the access point. This enables a wireless station to participate in the BSS efficiently to share and utilize channel resources. Balanced access points only accept new wireless stations entering to its coverage area. Over-loaded access points do not admit more wireless stations and run the hand-off procedure for the connected wireless stations until the load is balanced.