A New Load Balancing Procedure in IEEE 802.11 WLANs

Ahmet BARAN

Associate Professor, Computer Engineering Department,Erzincan University, Erzincan, Turkey

---***---

Abstract - *In this paper, a new load balancing procedure has been proposed in IEEE 802.11 wireless local area networks (WLANs). This procedure balances load of access points (APs) of an extended service set (ESS) domain without using any network entity. Proposed procedure balances load of APs belong to different ESS, too. In procedure, a network-assisted station-controlled handoff schema has been implemented and a fuzzy expert system in stations (STAs) selects optimal AP for each STA by using information of STAs and APs in vicinity. Simulation results clearly show that proposed procedure has balanced all STA throughputs and also has increased total wireless network throughput.*

Key Words: IEEE 802.11, Load balancing, handoff algorithm, fuzzy expert systems.

1. INTRODUCTION

IEEE 802.11 wireless networks are becoming increasingly popular and beginning to be deployed nearly everywhere from the office to the home to public hotspots in cafes and restaurants. With the spread of wireless local area network (WLAN), the number of wireless stations (STAs) connected with wireless LAN also has increased. Therefore, multiple access points (APs) are required to serve many STAs and to improve transmission capacity in the WLAN. Although IEEE 802.11 wireless LAN can serve many STAs through the multiple APs, the following significant issue can arise in those cases: how to achieve network load balancing.

There is no standardization for load balancing in IEEE 802.11 architecture. In the existing architecture, the received signal strength (RSS) is usually used to select an AP [1]. This procedure assumes that the quality of service (QoS) of the selected AP would be the best on the basis of its RSS alone [2]. However, this strategy may cause the concentration of STAs at specific APs: many STAs may associate with only a few APs because their signal strengths measured by STAs are strong; while only a few of STAs may associate with the remaining APs. This results in imbalanced traffic load on APs in WLAN, thereby degrading the fairness in STA throughput and harming an efficient use of WLAN resource [3]. Indeed, studies of deployments of public-area wireless networks have shown the traffic characteristics and user behavior in WLAN [4-7], and they report that traffic load is often distributed quite unevenly among APs.

In order to solve this issue, various association control schemes have been considered by both the research community and the industry. However, up to now, load balancing has been studied for only one extended service set (ESS). Recently, a lot of ESSs' access regions which are belong to different organizations or Internet Service Providers (ISP), overlap other ESSs' access regions and overlapped areas occur. In this case, stations in these overlapped areas have encountered a new problem: which AP in these ESSs provides the best QoS.

In this paper, we present a novel load balancing procedure for wireless LANs. There are two key issues in our procedure. Firstly, we have designed a novel communication procedure among APs for load balancing in an ESS without using any wired backbone entity. Secondly, proposed procedure balances load of APs of different ESS. To select the most suitable AP for load balancing, we have proposed a fuzzy expert system.

The rest of the paper is organized as follows. The related works are reviewed in Section 2. The architecture of proposed load balancing procedure is presented in Section 3 and optimal AP selection procedure is presented in Section 4. Section 5 contains our simulation results. Finally, our conclusions are given in Section 6.

2. RELATED WORKS

Load balancing in WLANs has been studied intensely and all of the suggested load balancing techniques distribute traffic load by managing AP-STA associations. Depending on which part of the network is in charge of such managements, these approaches could be classified into two types: STA-based and network-based [8].

In STA-based approaches, each STA tries to associate with an AP providing maximum throughput. In selection process, STA obtains and uses APs' load levels instead of using APs' RSS. STAs can obtain load levels of APs in several ways. In [9], STAs estimate potential upstream and downstream bandwidth of APs by using delays of beacon signals of APs which are in coverage area. This method does not modify APs and does not use any network entity.

On the other hand, there are some methods, where load levels of APs are reported to STAs, using beacon frames or probe response frames. In [3,10,11], APs assist to STAs with broadcasting their traffic loads or number of associated STAs. Also in [12], suggested approach proposed that a dedicated server assists to STAs for AP selection. According to this approach firstly each STA associates a AP quickly in scanning phase. Then, STA learns load of APs in coverage area from dedicated server via associated AP connection. With additional estimate on its bandwidth consumption, STA decides whether a handoff should be conducted to distribute to load. In [13], STA associates to each AP quickly in scanning phase and tests AP using reference servers in order to estimate bandwidth of each AP. Then, STA associates with the AP that provides the highest performance service for a long stay.

As for network-based approaches, STAs behave passively and AP or a dedicated server or a switch controls the distribution of APs load. Two basic techniques are used in these approaches: Coverage Adjustment and Association Management. In coverage adjustment technique, concentrated APs reduce their transmission powers of beacon signals and lightly-loaded APs raise their transmission powers of beacon signals [14]. Moreover, continuous coverage area is ensured with cooperation of APs [15,16]. In association management technique, overloaded APs do not accept new STAs and only under-loaded APs accept new association requests [17]. Furthermore, overloaded APs send unsolicited disassociation frame to STAs and hope that STAs will associate with a less loaded AP. In dynamic state, APs must consider of other APs' load levels for balancing of network in a rapidly-changing networking environment. In mentioned studies, APs are aware of other APs' load levels via a wired backbone. Some of studies use Inter Access Point Protocol (IAPP) with slight modifications [16,17] and some of them use a dedicated server connected to wired backbone [18,19] or a switch [20] for cooperation of APs. However, if BSSs belong to different organizations or ESS, none of mentioned studies is suitable and load of APs of these BSSs does not balance properly with these approaches.

3. PROPOSED LOAD BALANCING PROCEDURE

Load balancing schemes may be used in wireless LANs if at least two access points serve the same physical location. These APs can belong to the same or different ESS. Such a topology has been given in Fig.-1. As it can be seen from Fig.-1, although STA1 and STA2 can only connect to AP1, STA6 to AP2, STA9 to AP3 and STA11 to AP4, overlapped area stations (STA3, STA4, STA5, STA7, STA8, STA10) that are in the access regions of more than one AP can connect to one of these APs. For example, STA7 can associate with one of the AP1, AP2 belong to ESS-A or AP3 belongs to ESS-B. The aim of load balancing is to balance the traffic loads of each AP. In this respect only overlapped area stations can balance traffic load because STA1, STA2, STA6, STA9 and STA11 cannot associate with other APs.

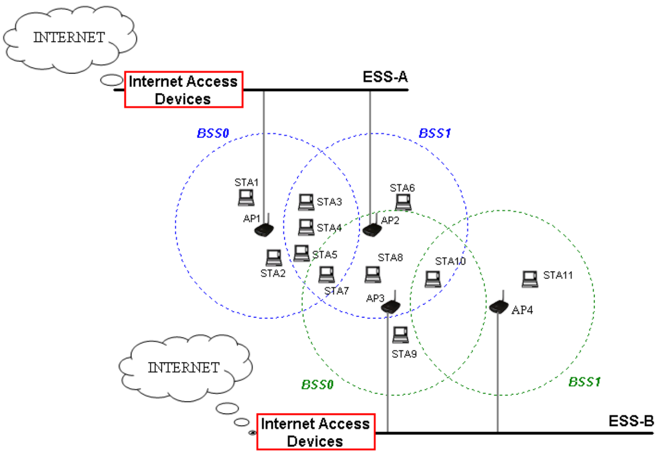

Fig -1:Multi-ESS IEEE 802.11 wireless network structure

In existing classical infrastructure network architecture [1], each STA must associate with an AP within a specific basic service set (BSS). Accordingly, STA must find the AP to communicate with others in an infrastructure network firstly, and the process of finding APs is called scanning. Scanning may be either passive or active. Passive scanning involves only listening to IEEE 802.11 traffic. In the passive scanning procedure, STA sweeps from channel to channel and records information from all beacons it receives. Active scanning requires the scanning STA to transmit and elicit responses from APs. In an active scan, STA will move to a channel and transmit a probe request frame. If there is a BSS on the channel that matchs the Service Set Identifier (SSID), the AP in that BSS will respond by sending a probe response frame to the scanning STA. Once the scanning STA has processed any response or has decided there is no response, it may change to another channel and repeat the process. At the conclusion of the passive or active scan, STA has accumulated information about BSSs and ESSs in its vicinity. After compiling these scan results, STA can select one of BSSs, and establish the association with the selected BSS [1].

In this paper, instead of classic handoff procedure, a new procedure has been proposed and phases of this procedure have been given in Fig.-2. According to our procedure the handoff procedure consists of four phases.

1. *Scanning Phase:* As existing architecture, STA scans all available channels from the first channel to the last available channel. When STA switches to each channel, it sends a probe request frame and waits a probe response frame. If there is an

AP in the channel, it responses this request. Distinct from existing probe response frame [1], this frame includes packet loss (PL) of AP and internet access rate (IAR) of ESS. STA scans all channels in that way and saves collected information to internal database. This database includes identifier and IAR of ESS and also identifier, channel, encryption state, RSS and PL of AP.

2. *Inform Phase:* The same STA switches to all channels except idle channels in the same order and sends collected data to that channel AP via a new modified probe request (MPR) frame. When AP gets MPR frame it updates internal database with this information packet. This frame includes collected information of STA and this additional information located in frame body. Rest of MPR frame is same as probe request frame. AP does not response to MPR frame and STA do not wait any answer from APs in this phase. In this way, APs, that have overlapped areas, learn load of other APs in the same or different ESS via overlapped area stations.
3. *Decision Phase:* STA selects the optimal AP among the candidate APs for load balancing by using a fuzzy expert system. The selection process is detailed in following section.
4. *Association or Re-Association Phase:* STA associates with the selected AP and starts data communication.

Dynamic state, originated with changing network conditions, is managed by APs. Network conditions change because new stations are joined to network or traffic characteristics of stations or locations of stations. In dynamic state, APs balance load of inner ESS or inter-ESS. In our design, dynamic load balancing is maintained by checking and balancing procedure (CBP), which has been given in Fig.-3. CBP may start by 3 triggers: *Time Limit (TL), Access Point Packet Loss Limit (APPLL) and Station Complain (STAC).*

First trigger, TL is used to start to CBP periodically. To take into account changes of network conditions, each AP makes CBP in initially configured periods (e.g. 60 seconds). With this process, overlapped area STAs, which have reported two or more APs in the inform phase, are forced to handoff procedure one by one. In this way, APs are informed about other APs by overlapped area STAs.

Second trigger, APPLL is triggered if an AP reaches to the packet loss limit initially configured. When AP reaches to the packet loss limit, AP hopes that a STA will associate an under loaded AP and AP selects a station in overlapped area and Selected STA (SSTA) is forced to handoff procedure.

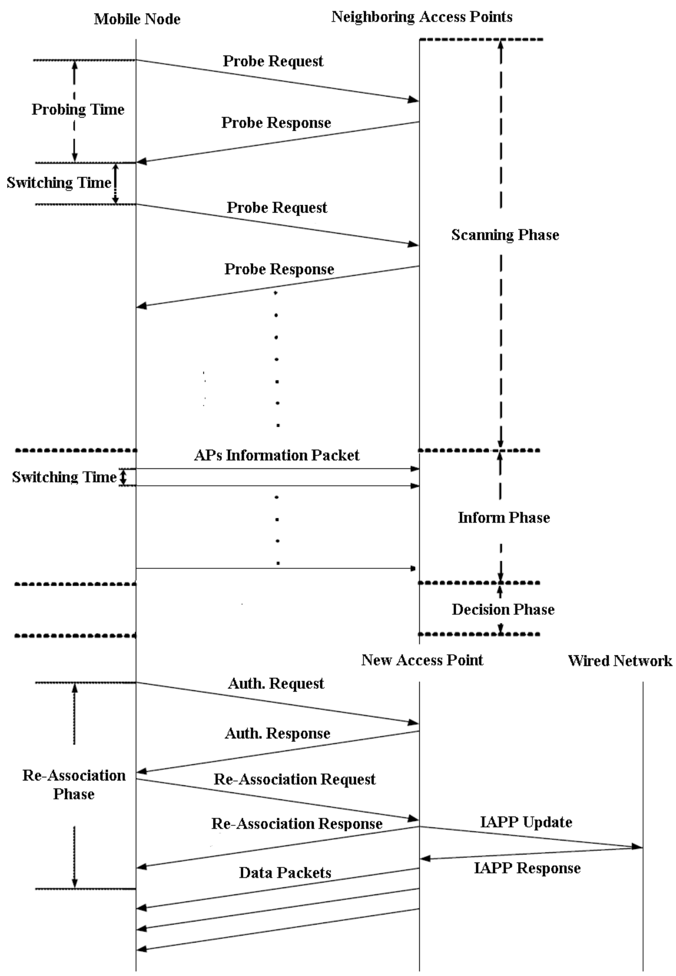

Fig -2:Proposed handoff procedure and its phases

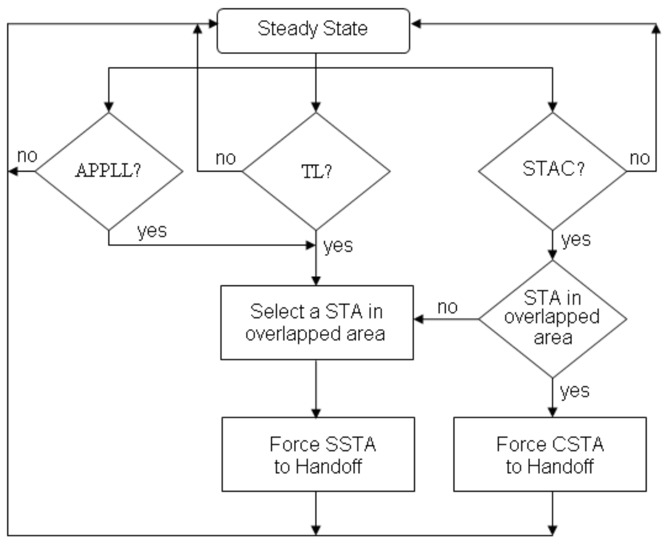

Fig -3: Dynamic state

Third trigger, STAC is triggered if there is a Complainant STA (CSTA), which complains about its high packet loss. When STAC is triggered, firstly, AP checks whether CSTA is in an overlapped area or not. If CSTA is in overlapped area, AP forces CSTA to handoff procedure, else AP selects a station in overlapped area and SSTA is forced to handoff procedure. If CBP has been triggered with APPLL or STAC, and load cannot be balanced with SSTA, then AP selects another STA in overlapped area and new SSTA is forced to handoff procedure.

In this point, SSTA enters a new scanning phase. After finishing all handoff phases, STA may associate with a new AP or the same AP. This loop continues until network is balanced by SSTA in overlapped areas. If there is no change in network, it is concluded that network is balanced or system resources are insufficient and network waits one of the next triggers.

4. AP SELECTION PROCEDURE FOR LOAD BALANCING

Another novelty in proposed procedure is AP selection method. In our procedure, STA makes decision according to some metrics using a Fuzzy Expert System (FES). Fuzzy expert systems utilize human knowledge by giving the fuzzy or linguistic descriptions a definite structure. Some of the popular configurations include Takagi and Sugeno's fuzzy system, and Mamdani's fuzzy system [21]. This paper utilizes Mamdani's inference engine [22], which consists of fuzzifier, knowledge base, fuzzy inference engine, and defuzzifier.

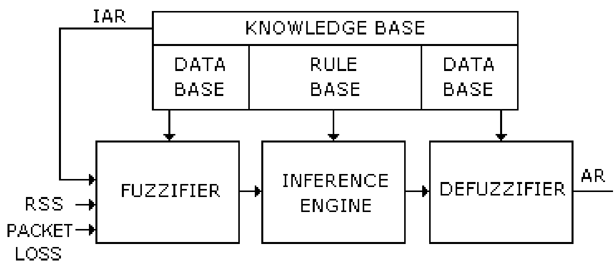

Fig -4:Fuzzy expert system

Defining and computing of load is one of the key elements in a load balancing procedure. Today, there is no standardized definition of load. In [3], AP load is estimated by using packet error rate (PER) and number of associated stations (N). However, because of wireless network conditions change rapidly, PER and N are not adequate to predict AP load. IEEE 802.11e [23] proposed to use channel utilization, the percentage of time an AP is busy transmitting or receiving data during some interval, as a gauge of channel loading. Channel utilization is not an appropriate AP selection metric, however, as it does not capture transmission capabilities of respective APs: an 80% utilized IEEE 802.11g AP can offer ever more bandwidth than a 40% utilized IEEE 802.11b AP [8]. Also frame drop rate [15] and the time between beacon frames [9] are proposed load metrics for load balancing. Expected throughput [17] is another proposed metric to compute load metric in load balancing system. However throughput is affected by time varying channel conditions. In [19], the authors have implemented a scheme in which access points aim to fairly distribute bandwidth among all users. The mobile terminals will choose to associate with the AP that can the best provide its bandwidth request. However, the predicted load coincides with observed workload only if STA consumes all the bandwidth allocated to it [8].

In this study we have used packet loss (PL) for load metric proposed by Bianchi and Tinnirello [11]. They are presented to several traffic-based metrics such as the packet loss and the gross load including the number of stations and the retransmission probability because each station in a wireless network generates a variable amount of traffic.

The inputs to the proposed FES, depicted in Fig.-4, are RSS, Packet Loss (PL) and Internet Access Rate (IAR) and the output of the FES is Appropriateness Rank (AR). These inputs are the best possible load balancing metrics to select the target AP as the one, which minimizes the expected packet loss percentage after the addition of a new STA [11]. PL of an AP is calculated in AP by following equation:

$$PL = 1 - \frac{N}{TGL}$$

Where N is the number of the associated STAs in actual BSS and TGL is Total Gross Load. TGL is calculated in AP side by this equation:

$$TGL = N + \sum_{i=1}^{N} \frac{p_i}{1 - p_i}$$

Where p_i is transmission error probability suffered by the i-th STA.

In FES, firstly, a fuzzifier fuzzifies inputs. Membership functions of inputs and output are given in Fig.-5.

Inference engine calculates an AR for each AP by using 125 rules in the rule base. Some sample rules in rule base are given in Table 1. Rule base is constructed to calculate the highest AR when the best input values are applied. In addition to this, adding a new rule or removing or modifying an existing rule can be realized by interface software in STAs practically. According to rule base, AP, which has the highest RSS, minimum PL and the best IAR, is ranked as the highest rank. After ranking all the APs in the region, these ranks are put in order, and then the highest scored AP is selected. If STA cannot associate with this AP, STA attempts to associate with other APs subsequently.

If the reason of associate failure is authentication, then this AP is saved to black AP list and then STA does not scan and attempt associate with this AP.

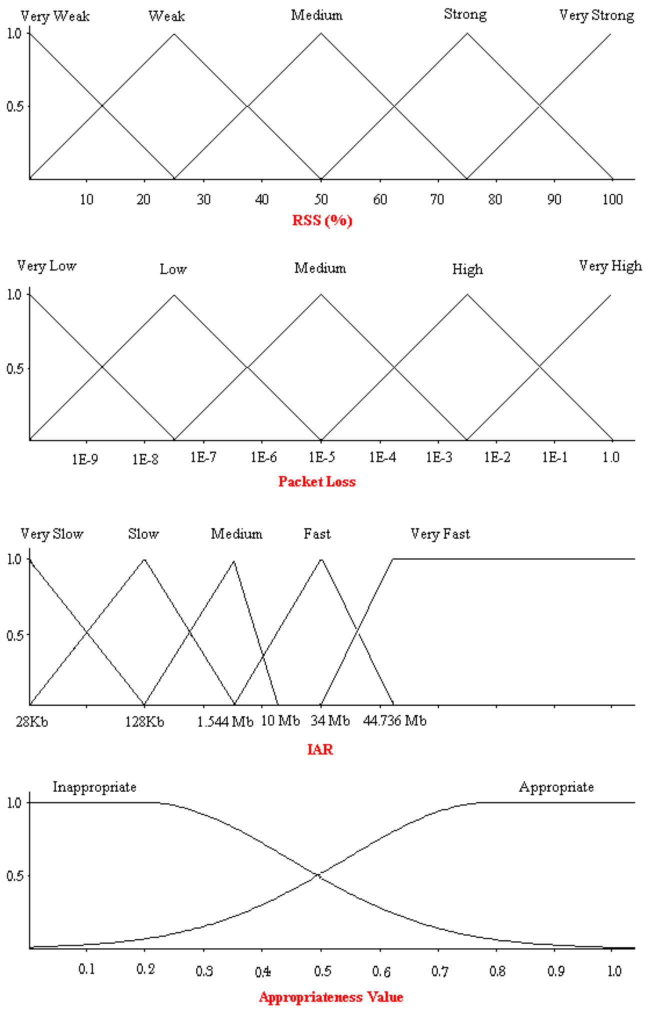

Fig -5:Membership functions of RSS, Packet loss, IAR and Appropriateness value

Table -1: Some rules in the rule base

RSS	PL	IAR	Appropriateness
Very Weak	Very High	Very Slow	Inappropriate
Very Weak	Medium	Fast	Inappropriate
Medium	High	Medium	Appropriate
Medium	Low	Fast	Appropriate
Strong	Very Low	Very Fast	Appropriate

5. SIMULATION RESULTS

In this section, performance and load balancing ability of the proposed method has been compared to RSS based association procedure [1], maximizing local throughput (MLT) approach [3], potential bandwidth (PBW) approach [9] and Virgil [13] load balancing procedures. We have implemented simulations in ns-2 [24] after modifying it to 802.11b. Orinoco 11b NIC's specifications for STA and APs are used and we have used shadowing model over wireless link. The size of the simulation area is 80m x 50m and 4 APs in different channels and 11 STAs located on simulation area. AP1 and AP2 have been connected to ESS-A and AP3 and AP4 have been connected to ESS-B and each AP has been connected to wired backbone via 10 Mbps connection link. For comparability with other load balancing algorithms, IAR parameter values of ESSs' have been adjusted to same value and two wired stations run as a FTP server have been connected to ESS-A and ESS-B. All APs and STAs have been powered up at the begining of simulation. Simulation has been analyzed for 100 seconds between 100th and 200th seconds. The value of TL parameter is selected as 25 seconds in three load balancing procedures. In our simulations, it has been observed that, with using 13 available channels, average scanning phase delay and average inform phase delay of proposed procedure are about 353ms and 40ms, respectively. In addition, suggested fuzzy decision system spends about 2ms for selection of optimal AP in the decision phase.

Fig.-6 illustrates the simulation scenario and it can be seen that, if RSS is used for association algorithm, most of the users will select AP1 and AP3 for communication. However, this scheme overwhelms the APs. AP1 and AP3 can only offer data rates between 880 Kb/s and 1277 Kb/s to their associated users because of the congestion, while AP2 and AP4 offer data rates around 4000 Kb/s, as shown in Table 2.

As it can be seen in Table 2, proposed load balancing method is relieved the congestion of AP1 and AP3 by association of some STAs to AP2 and AP4. Using proposed load balancing procedure, performance of all users except for STA6 and STA11 have been increased between 49% and 144% compared to RSS association algorithm.

To evaluate the performance of our method, we have also compared our method with three methods: MLT [3], PBW [9] and Virgil [13] with the same scenario and same parameters. Simulations show that, all of the three load balancing procedures are balanced to network resources of the given scenario. However, proposed procedure is the fairest load balancing procedure among these three procedures. Fig.-7 draws the average throughputs of RSS algorithm, proposed procedure and other three procedures.

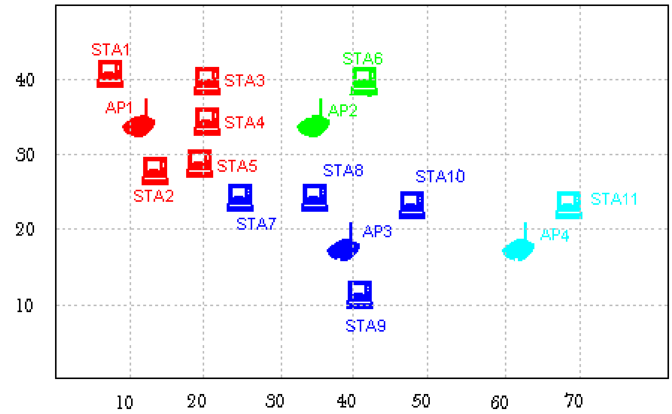

Fig -6:Illustration of simulation set-up

Table -2:Average throughputs of stations (Kb/s)

Node ID	RSS	MLT	PBW	Virgil	Our Method
STA1	972	1670	1544	1210	1891
STA2	1019	1804	1643	1312	1903
STA3	951	1645	1699	1548	1822
STA4	1053	1249	1503	1344	1601
STA5	976	1497	1577	1456	1670
STA6	4099	2044	1933	2445	1876
STA7	880	1465	1522	1332	1590
STA8	974	1453	1566	1401	1676
STA9	1277	1850	1866	1656	1901
STA10	915	2157	2192	1769	2230
STA11	3977	2450	2350	2666	2403

Total throughputs and throughput increment rates compared with RSS of simulated procedures have been given in Table 3. It can be seen from Table 3 that proposed procedure has increased the total throughput in the maximum rate compared to the other methods.

In Fig.-8, offered AP loads of simulated procedures are given. According to Fig.-8, all of load balancing procedures have balanced the load of 4 APs clearly in various rates. However, proposed procedure has decreased total AP load to 19% compared to RSS procedure, while MLT, PBW and Virgil has decreased total AP load 5, 10 and 6%, respectively. On the other hand, we have analyzed PL among the simulation procedures.

According to simulation results, total ESS PL is 7.50E-05 in the proposed procedure while total ESS PL is 6.00E-2, 8.2E-4, 9.12E-4 and 2.01E-3 in the RSS, MLT, PBW and Virgil association procedures, respectively, shown in Fig.-9. Observed low PL value has been expected because of proposed procedure uses PL based load balancing approach.

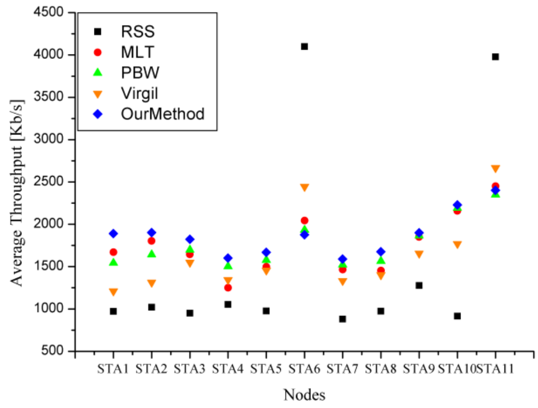

Fig -7: Comparison of average throughputs

Table -3:Total throughputs and throughput increment according to RSS

	RSS	MLT	PBW	Virgil	Our Method
Total throughput (Kb/s)	34186	38568	38790	36278	41126
Throughput increment according to RSS (%)	--	12.82	13.47	6.12	20.30

6. CONCLUSIONS

In this paper, a new load balancing and optimal AP selection procedure is presented in order to use future wireless local area networks. Two novelties are presented in proposed procedure. The first one is that a novel communication procedure among APs has been developed for load balancing without using any wired backbone entity. The second novelty is that, with the proposed procedure, the load can also be balanced among APs of different ESSs. Furthermore, a new fuzzy expert system based AP selection procedure has been introduced. The proposed structure can be realized by adding some slight software additions to existing structure without a hardware addition.

To evaluate the performance, proposed procedure has been compared with MLT, PBW and Virgil load balancing procedures and RSS association procedure. Simulation results show that proposed procedure has balanced throughputs of all STA connected to two ESSs. Total wireless network throughput of two ESSs has also been increased clearly with proposed procedure. We have also observed the load of each AP and PL of every BSS. Consequently, we have shown that proposed procedure have provided the best results compared to other procedures.

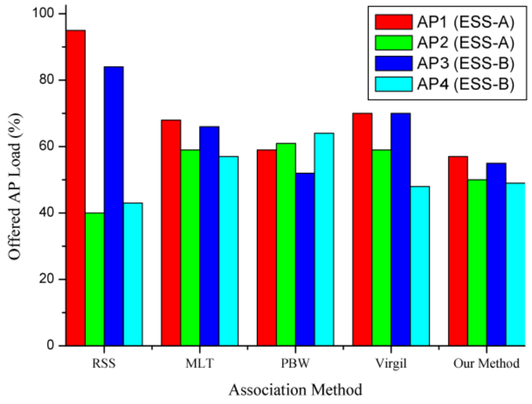

Fig -8:Comparison of offered AP loads

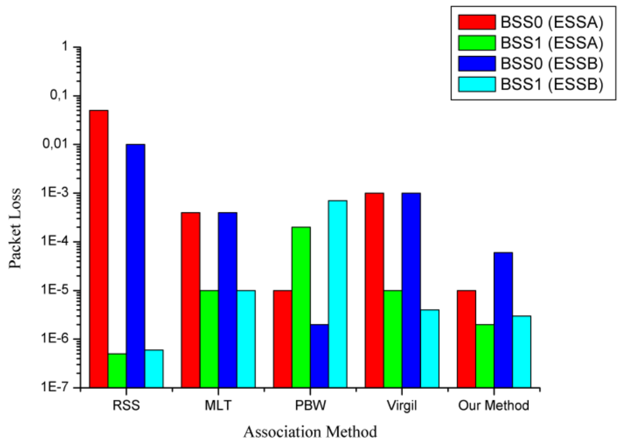

Fig -9:Comparison of packet losses

REFERENCES

[1] IEEE Std 802.11. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Technical report, IEEE, 1999.

[2] E.G. Villegas, R.V. Ferre, J.P. Aspas, “Load Balancing in WLANs through IEEE 802.11k Mechanisms,” in. Proceedings of the 11th IEEE Symposium on Computers and Communicatios, pp. 844-850,2006.

[3] Y. Fukuda, Y. Oie, “Decentralized access point selection architecture for wireless LANs”, IEICE Transaction on Communications, Vol. E90-B, pp. 2513-2523, 2007.

[4] D. Tang, M. Baker, “Analysis of a local-area wireless network”, in the Proceedins of ACM Mobicom 2000, pp. 1–20, 2000.

[5] D. Tang, M. Baker, “Analysis of a Metropolitan-Area Wireless Network”, in the Proceedins of ACM Mobicom 1999, pp. 12–23, 1999.

[6] D. Kotz, K. Essien, “Analysis of a campus-wide wireless network”, in Proceedings of the Eighth Annual International Conference on Mobile Computing and Networking, pp. 107–118, 2002.

[7] A. Balachandran, G.M. Voelker, P. Bahl, V. Rangan, “Characterizing user behavior and network performance in a public wireless LAN”, in the Proceedings of ACM SIGMETRICS’02, pp. 195-205, 2002.

[8] L.H. Yen, T.T. Yeh, K.H. Chi, “Load Balancing in IEEE 802.11 Networks”, IEEE Internet Computing, vol. 13, pp. 56-64, 2009.

[9] S. Vasudevan, K. Papagiannaki, C. Diot, J. Kurose, D. Towsley, “Facilitating access point selection in IEEE 802.11 wireless networks”, in Proc. ACM SIGCOMM/USENIX Internet Measurement Conf., 2005.

[10] I. Papanikos, M. Logothetis, “A study on dynamic load balance for IEEE802.11b wireless LAN”, in Proc. 8. Intl. Conf. Advances in Communication and Control: Telecommunications/Signal Processing, 2001.

[11] G. Bianchi, I. Tinnirello, "Improving load balancing mechanisms in wireless packet networks," in Proc. IEEE Int. Conf. Commun., pp. 891-895, 2002.

[12] L.H. Yen, T.T. Yeh, “SNMP-based approach to load distribution in IEEE 802.11 networks”, in Proc. IEEE VTC2006-Spring, 2006.

[13] A.J. Nicholson, Y. Chawathe, M.Y. Chen, “Improved access point selection”, in Proc.ACM MibiSys 06, 2006.

[14] Y. Bejerano, S.J. Han, “Cell Breathing Techniques for Load Balancing in Wireless LANs”, IEEE Transactions on Mobile Computing, pp. 735-749, 2009.

[15] O. Brickley, S. Rea, D. Pesch, “Load balancing for QoS enhancement in IEEE802.11e WLANs using cell breathing techniques”, in IFIP Mobile and Wireless Communication Networks Conf., 2005.

[16] Y.Wang, L. Cuthbert, J. Bigham, “Intelligent radio resource management for IEEE 802.11WLAN”, in Proc. IEEE WCNC, 2004.

[17] H Velayos, V. Aleo, G. Karlsson, “Load balancing in overlapping wireless LAN cells", in Proc. IEEE ICC, 2004.

[18] A. Balachandran, G.M. Voelker, P. Bahl, “Hot-spot congestion relief in public-area wireless networks”, in proceedings of WMCSA’02, pp. 70–82, 2002.

[19] Y. Bejerano, S.-J. Han, L. Li, “Fairness and load balancing in wireless LANs using association control”, in Proc. ACM MobiCom'04, pp. 315-329, 2004.

[20] P. J. Iyer, P. Narasimhan, M. Andrade, J. Taylor, “System and method for centralized station management”, U.S. Patent 0213579, Sept. 2005.

[21] L.-X. Wang, Adaptive Fuzzy Systems and Control. Englewood Cliffs, New Jersey: PTR Prentice Hall, 1994.

[22] E. H. Mamdani, “Applications of fuzzy algorithms for simple dynamic plant”, in Proc. IEEE, vol. 121, pp. 1585-1588, 1974.

[23] IEEE 802.11e/D4.0, Draft Supplement to Part 11:Wireless Medium Access Control (MAC) and physical layer (PHY) specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS), November 2002.

[24] The VINT Project. (2010, Jan.). Network Simulator NS. [Online]. Available : http://www.isi.edu/nsnam/ns/

BIOGRAPHIES

Ahmet Baran graduated from Erciyes University, Turkey, in 1998. He took M.Sc. degree from Erciyes University, Turkey, in 2001, all in Electronics Engineering. He took Ph.D. degree from Computer Sciences, Atatürk University, Turkey, in 2001. His interest areas include computer programming, wireless computer networks, nonlinear system identification,electronic circuits and management information systems.

