International Advance Journal of Engineering Research (IAJER)
Volume 4, Issue 01 (January- 2021), PP 08-14
ISSN: 2360-819X
www.iajer.com

Research Paper **Open Access**

K-NEAREST NEIGHBOR FOR BEST CLUSTERING STATION IN SELECTION OF ACCESS POINTS ON THE 802.11G MULTI RATE WIRELESS NETWORK

Hendro Suwignyo[1*], Muhammad Aziz Muslim[2], Herman Tolle[2]
[1](Electrical EngineeringDepartement/ Brawijaya University, Indonesia)
[2](InformaticsDepartement/ Brawijaya University, Indonesia)
Corresponding author: Hendro Suwignyo

ABSTRACT : Selection of access points (AP) in a wireless local area networks extended service set (WLAN ESS) generally has problems with multi transmission rates, because stations with low rates will make AP reduce the quality of data transfer to other stations that have large transmission rates and signal. The k-nearest neighbors (KNN) algorithm based on signal to noise ratio and transmission rates is used in this study to make the best cluster stations and then reject stations outside the cluster to be connected to other APs to produce uniform throughput on the AP according to the transmission rate provided. Testing using actual equipment and workin in a building with several spaces and obstacles where the signal was not in the line of sight resulted in a better min-max ratio throughput between Station to APs compared to the SSF, LLF and KNN_distance methods of 11%, 10 % and 8%, with the lowest transmission rate of 12Mbps, while other methods got 9Mbps.

Keywords –*Access Point Selection, K-Nearest Neighbor, 802.11g Multi Rate Wireless Network*

I. INTRODUCTION

In IEEE 802.11 protocol known as Wireless Local Area Network (WLAN), Radio Signal Strength Indicator (RSSI) or Signal to Noise Ratio (SNR) as the main criteria for selecting an access point (AP) based on Strong Signal First (SSF), where AP with the largest RSSI / SNR captured by wireless network devices (Station / Sta) is the first choice to connect regardless of signal interference, traffic jams and network loads on the AP. UsingSNR willonly select AP with the best SNR regardless of load or number of users [1]. This results in an unbalanced traffic load on the APs, decreasing the balance throughput at each Sta and reducing the efficiency of equipment in the WLAN network [2].

Several studies were conducted to overcome this problem by ensuring the reliability of APs in the ESS (Extended Set Service) network, namely by selecting AP for Sta to make the throughput balance on all APs run well and fairly evenly. Intelligent algorithms are used to determine whether a station can last longer connected to an AP or rejected to find another AP, then the algorithm performs computations to balance all APs.In addition, as several IEEE 802.11 physical layers variant ssupport multi transmission rates, the currently implemented selection mechanism (SSF) does not consider the impact of association station on the effective throughput of other (already associated) stations [3].

This research is using the real hardware equipment and conducted not in an open area, but in a building with several spaces and obstacles where the signal was not in the line of sight. Proposed using the K-Nearest Neighbor method to get the best cluster of Sta based on Signal To Noise Ratio (SNR) and transmission rate to get balance between AP (min-max ratio) throughput in the appropriate transmission rate.

The rest of this paper is organized as follows: Section II describes related works; Section III outlines the design in the proposed scheme; Section IV elaborates the scenario and interprets the result of experiment; and Section V concludes the paper.

II. RELATED WORK

2.1. Relatedwork

Before entering the subject of selection of access points, we first discuss some of the problems with multirate WLANs. Usually, antenna theory assumes that the media through which the signal passes is when the network in each room in each building will have a unique order depending on the materials used in the construction, namely physical shape, room content, and the position of the access point. It is difficult to determine the exact make up of the buildings incebricks containing a high metal (iron) content, the plaster board, foil wall paper, lead paint wiremesh safety glass canall have a significant effect on the absorption properties of the building [4]. So that the RSSI is depend by the location of the AP and the location of the Sta which can vary at a certain distance and can be blocked by doors, windows, cupboards and so on. And morefarther the distance between AP and Sta, the transmission rate will also be small, this can be caused by interruptions from other radio signals (AP / Sta). Where this transmission rate in a WLAN is required to keep the bit / frame error to the correct value, and the appropriate transmission rate is chosen depending on the bit error status [5]. Thus, in a multi-rate WLAN environment, however, the total throughput decreases if the AP is shared by the STA where the transmission rates are high and low simultaneously. This issue is reported as a performance anomaly problem in IEEE 802.11 [6].

A lot of research has been developed over the last few years regarding centralized control of AP for user management of STA, several methods are used and developed, several research models are carried out, including simulations on programs, using test tools (testbed) and so on. Research conducted generally uses the Signal Strong First (SSF) method as a comparison of results.

Gong et.al [7] used the NS2 simulator for AP selection by calculating RSSI weights to make Sta freely select the AP with the lowest transmission rate. Comparison with better results from the SSF method and the Fukuda method on the average throughput of Sta (in Kbps).

Larasati et.al [8] researched access points selection using the Software Defined Network Mininet-WiFi with the Extended-LeastLoad First (LLF) algorithm which added the RSSI value to the algorithm calculation. Comparing SSF and LLF algorithms get better values for jitter, packet loss, and load balancing throughput. In the measurement of throughput, it was obtained better results than SSF and LLF at 0.04% and 0.05%.

Ahmed et.al [9] using the Unity 3D Engine software, conducted Load Balancing research on the WIFi / LiFi hybrid net work using the K Nearest Neighbor optimization algorithm based on RSSI and the distance between AP and Sta. He comparison of results with Wang, Wu and Ma methods, through put results are better than the three at the number of stations below 10, for results above 10 stations show the same throughput value.

2.2. ProposedMethod

This research testing with the real equipment where the Aps and Stations are not in the line of sight, but in a house building with several spaces and its obstacles so that the signal was attenuated. With the different obstacle, of course the real distance between the station and the AP is no longer relevant in the changes of signal level and transmission rate. To measure this new distance the K-Nearest Neighbor algoritm is used, which uses the euclidean distance to determine which station closest to the center of the cluster. With K=n-1 will get1 worst Station that is not be member of the cluster of best Station, its mean Sta_worst ∉ Sta_best. This station out side cluster is in the worst condition because it has low SNR and a low transmission rate then it will be a candidate to be associated with other APs.

III. WORKING DESIGN

3.1. Design of Hardware and Software

System design in this research includes software and hardware design. The network design is in a Wireless LAN Extended Service Set (ESS) which is formed from 3 Aps that are distributed by LAN cable. AP uses the HAP lite Access Point RB 941-2nD-T and 12 stations uses an Android-based smartphone in the same brand and same type. A file server computer is used as a download facility for the station to make and generate traffict between Station and AP. One computer as controller for processing data and executing. The network hardwares describe as below on Fig.1.

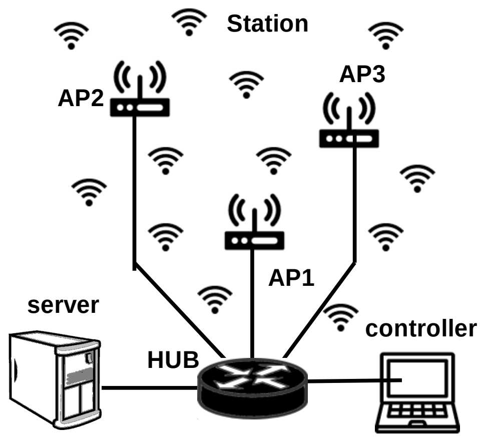

Fig.1. Network Wireless LAN ESS

The ESS WLAN position is set with a distance of AP1 - AP2 = 6 meters, AP2 - AP3 = 7 meters, and AP3 - AP1 = 7meters in such a way that 12 Sta can get SNR data and various transmission rates. Every AP sets in non-overlapping channel for avoiding interferrence between Aps, AP1 sets on channel 1, AP2 in channel 6, and AP3 sets on channel 11. By placing the AP in a position on the table or cupboard within a distance of each AP about 6 to 8 meters. The placement of the station varies according to the condition of the building, so that many Sta do not get a signal directly because they are blocked by walls, wooden doors, tables and chairs, window trellises and several other obstacles. Fig 2 shows the AP and station installation area.

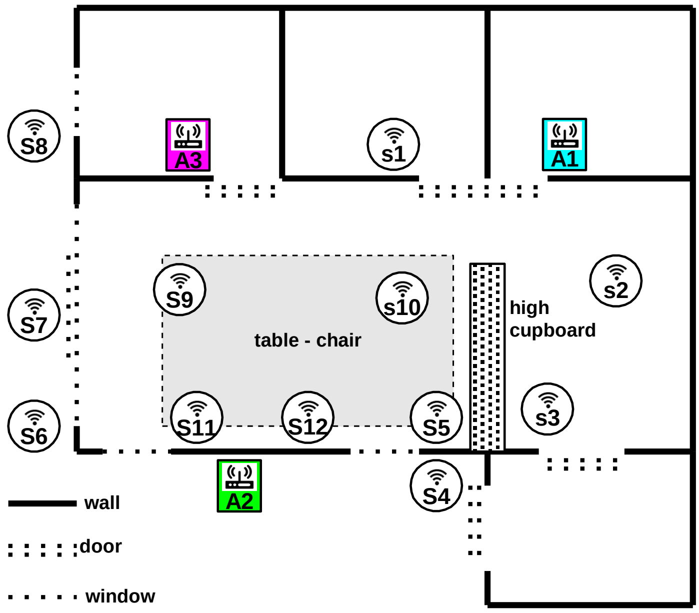

Fig 2. APs and Stations installation.

The software design builted from Delphi 7 desktop programming language using Mikrotik API to read and retrieve data from the RouterOS AP and running under OS Windows 7.

3.2. Algorithm

The KNN algorithm is one of many classification algorithm that is often used as supervised learning and does not require prior training. It is a classification method for data sets based on previously classified data learning, where the results of new query instances will be classified based on the majority of the proximity of the categories in the cluster. Classification is calculated using the euclidian distance,

$$d = \sqrt{(x2 - x1)^2 + (y2 - y1)^2} \qquad (1)$$

where d is euclidian distance, x1 is SNR of Sta_best as cluster center, x2 is SNR of neighboring Sta_best, y1 is transmission rate of Sta_best as cluster center, y2 is transmission rate of neighboring Sta_best.
Value range of SNR between 30 until 80, range of transmission rate are in value 6, 9, 12, 18, 24, 36, 48 and 54.

By determining K = n-1 for each AP, 1 Sta_best is chosen as the cluster center, determined based on the highest SNR value. The determination of the cluster center based on SNR is because SNR is not easy to change when compared to the value of transmission rate. This best cluster members is all stations that are allowed to associate with AP, while members outside the cluster is 1 Sta as a candidate to be associated to other AP.

3.3. Work

Flow of execute program is depicted in Fig. 3. Each Sta randomly associates with the AP, then Sta performs process of downloading the data file from the PC server using its browser. In this initial condition, all of the respective stations get different traffic because there are no download traffic restrictions. So the throughput of each Sta in each AP can be not in the same throughput.

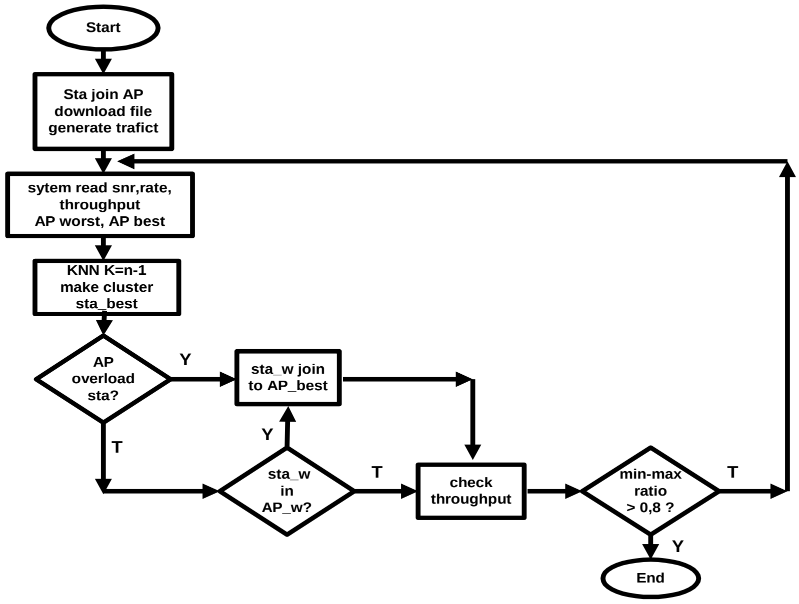

Fig.3. Flow chart of execute

From the calculation of the KNN classification process, one Sta will be a Sta_worst and out side the best cluster which the choosen candidate to be proccesed. Following the pseudo code below the steps to move the Sta_worst association: Once Sta is connected to the AP, it immediately performs a data download process then traffic occurs in the WLAN and produces data Sta throughput (t_Sta) and AP throughput (t_AP).

1. System reads the SNR and transmission rate of all Sta on each AP.
2. System checks the throughput on each AP (t_AP) to get the best AP (AP_best) and the worst (AP_worst).
3. Clustering process by KNN algorithm with K=n-1 to get Sta_worst $\notin$ Sta_best.
4. Looking for AP with the highest number of Sta (AP overload) and selecting Sta_worst to move association to AP_best.
5. If step no.5 does not exist an AP overload, continue to find Sta_worst in AP_worst as a candidate to move association to AP_best.
6. System checks the results of changes in the throughput of each AP to get the balance of each throughput.
7. In few many seconds repeat start step no.2 to get the balance for each throughput AP until the balance of throughput between Aps is met with the min-max ratio reaching 0.8, then iteration procces stoped.

The throughput balance between Aps is the comparison between AP_best and AP_worst throughput, we use min-max ratio and Jain's Fairness index [10].

Min-max ratio Sta calculated by:

$$\text{min-max ratio} = \frac{t\_Sta\_w}{t\_Sta\_b} \quad (2)$$

where *t_Sta_w* is minimum ongoing Sta throughput, *t_Sta_b* is maximum ongoing Sta throughput.

Min-max ratio AP calculated by:

$$\text{min-max ratio} = \frac{t\_AP\_w}{t\_AP\_b} \quad (3)$$

where *t_AP_w* is minimum ongoing AP throughput, *t_AP_b* is maximum ongoing AP throughput. When the min-max ratio maximum value is 1, that is means that all Stas/Aps have the same amount of throughput. In this study, a good min-max ratio was determined by a value above 0.8.

Throughput balance is calculated by Jain's Fairness index:

$$\text{Fairness index} = \frac{(\sum_{i=1}^{n} t\_AP)^2}{n * (\sum_{i=1}^{n} t\_AP^2)} \quad (4)$$

Where n is the number of AP.

IV. RESULTS

The final cluster of Sta from the KNN process is shown in Fig. 4. It shows that the actual distance can not be used as a reference for the Sta to associate with the closest AP because their position is blocked by walls, doors or windows and other object obstructions. Many Sta connected to AP farther than to AP closes to it, as example Sta_11 and Sta_12 preferred to association with AP3 not to AP2 wich very closes but behind the wall, the Sta_9 which put on the chair preferred connectto AP2 than to AP3.

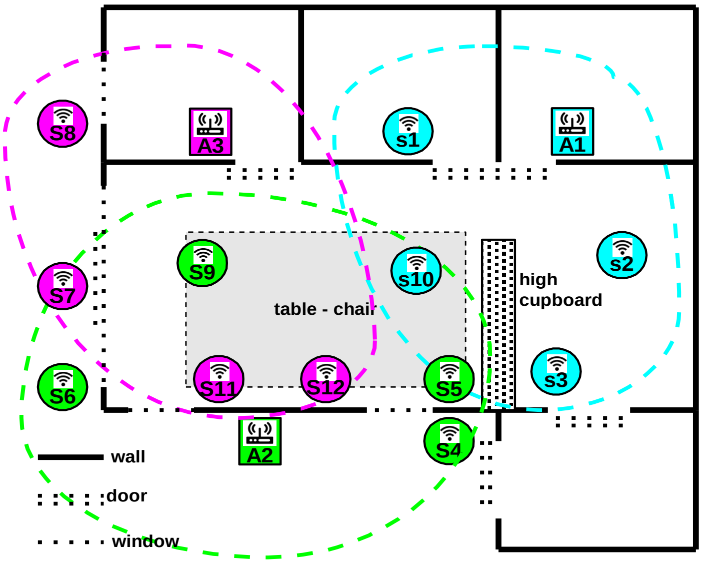

Fig.4. Final result association of Stas and APs clustering with KNN_snr-transmision_rate.

In this paper, the testing process is carried out until the min-max ratio of the AP throughput is above 0.8 after experiencing 6 step iterations. Table.1 below shows the moving of Sta in Aps on each step iteration proccess.

Table.1. Stas Moving Association In Aps.

iteration	**1**	**2**	**3**	**4**	**5**	**6**
	Sta 1	Sta 1	Sta 1	Sta 1	Sta 1	Sta 1
	Sta 2	Sta 2	Sta 2	Sta 2	Sta 2	Sta 2
	Sta 3	Sta 3	Sta 3	Sta 3	Sta 3	Sta 3
		Sta 10	Sta 10	Sta 10	Sta 10	Sta 10
AP1						
	Sta 4	Sta 4	Sta 4	Sta 4	Sta 4	Sta 4
	Sta 5	Sta 5	Sta 5	Sta 5	Sta 5	Sta 5
	Sta 6	Sta 6	Sta 6	Sta 6	Sta 6	Sta 6
	Sta 7	Sta 7	Sta 7	Sta 11	Sta 11	Sta 9
AP2			Sta 11		Sta 9	
	Sta 8	Sta 8	Sta 8	Sta 8	Sta 8	Sta 8
	Sta 9	Sta 9	Sta 9	Sta 9	Sta 12	Sta 12
	Sta 10	Sta 11	Sta 12	Sta 12	Sta 7	Sta 7
	Sta 11	Sta 12		Sta 7		Sta 11
AP3	Sta 12					

Table.1 shows the moving of Sta in 6 iterations process, the steps obtained in each iteration expalined below:

1. AP3 overload, Sta10 as Sta_worst outside best cluster will moved to AP1 which least number of Sta.
2. Sta10 associated to AP1, AP3 is AP_worst, Sta11 is Sta_worst will moved to AP2 as AP_best. Min-max ratio below 0,8.
3. Sta11 associated to AP2, AP2 overload, Sta7 as Sta_worst outside best cluster will moved to AP3 as AP_best.
4. Sta7 associated to AP3, AP3 is AP_worst, Sta9 is Sta_worst will moved to AP2 as AP_best. Min-max ratio below 0,8.
5. Sta9 associated to AP2, AP2 overload, Sta11 as Sta_worst outside best cluster will moved to AP3 as AP_best.
6. Sta11 back again to associated to AP3, but min-max ratio reached 0,8, iteration stopped.

In order to evaluate this proposed methode, comparison with other methods are carried out using a test instrument similar to this method for evaluating the throughput and the min-max ratio. First test on SSF scheme, then LLF by connecting station one by one to AP with a lowest load/trafict and high RSSI until each AP loaded with 4 Sta. Last test with scheme KKN_distance by connecting any Sta according to RSSI dan real distance based on KNN clustering.

Table.2. Lo_Min-Max RatioT_Sta, Hi_Min-Max RatioSta, Min-Max Ratio AP and Fairness Index.

no	Method	lo_min-max ratio t_Sta in AP	hi_min-max ratio t_Sta in AP	min-max ratio Access Point	Jain's fairness index
1	SSF	0,84	0,93	0,8	0,99178816
2	LLF	0,76	0,79	0,81	0,992456014
3	KNN_distance	0,7	0,92	0,83	0,994560369
4	KNN_proposed	0,89	0,94	0,91	0,998523494

Table.2 shows all fairness index of 4 scheme in good fairness, all reach above 0,99. All min-max ratio throughput AP reach above 0,8.

Table.3. T_Sta And T_AP, the mumber Of Sta Associated In AP And Lowest Transmission Rate

Method	SSF		LLF		KNN_dist		KNN_prop	
	n_Sta connect	Throughput (Mbps)	n_Sta connect	Throughput (Mbps)	n_Sta connect	Throughput (Mbps)	n_Sta connect	Throughput (Mbps)
AP1	3	5,5	4	4	4	4,2	4	5,2
AP2	4	6	4	5	4	4,7	4	5,5
AP3	5	4,8	4	4,7	4	5,1	4	5
total throughput (Mbps)		16,3		13,7		14		15,7
Trans rate min(Mbps)	9		9		9		12	

Table.3 shows the numbers of Sta distributed in each AP, the throughput each AP and its lowest transmission rate can be 1 or 2 of 3 AP in every tested scheme.

Table. 4. ResultOfCalibrationTest For MaximumThroughputAP In Multi Transmission Rate.

rate transmission	6 Mbps	9 Mbps	12 Mbps	18 Mbps	24 Mbps	36 Mbps	48 Mbps	54 Mbps
max throughput (Mbps)	3,7	5,3	6,2	9,2	11	17	23	25

A calibration AP has been tested before to know the maksimum throughput AP in multi transmission rate. It shown in Table. 4, explain in every transmission rate the real maximum throughput is about a half less, not in full throughput according mentioned to its transmission rate. The analysis about 4 scheme can be explain by data showing in Table. 2 and Table.3

From the Table.1 in SSF scheme, the lo_min-max ratio and hi_min-max ratio Sta between Sta reaches 0.8 at all. This means that the performance of each AP in distributing traffic is quite good. But it noted from Table.2 that on AP1 only 3 stations associated, this caused by the station chooses the largest signal so AP1 can be associated with 3 stations. And in AP3 gets 5 stations but with the lowest transmission rate of 9 Mbps, that is an 802.11 anomaly occurs which causes other Sta also get a low throughput. Overall, SSF's performance is not good due to overloading of one AP, that is AP3.

In LLF scheme shows the lo_min-max ratio and hi_min-max throughput ratio between Sta below 0.8. This caused by the LLF method each station associated one by one to takes turns filling the AP which low throughput first in the hope for getting more available traffic and eventhough still chooses the best RSSI between 3 APs, but after all the APs are full filled with 4 station there are some stations get a low transmission rate and attract the others to get a low transmission rate too. In this LLF the lowest station transmission rate is 9 Mbps.

In the KKN-distance method, the lo_min-max ratio and the hi_min-max throughput ratio between Sta below 0.7 on AP3. This caused by the KNN_distance method the stations are connected based on the actual distance regardless they are blocked by the shape of the buildings, doors and windows, then some stations get a low transmission rate and make other stations get a low transmission rate too. In this method the lowest station transmission rate is 9 Mbps.

In the KKN-proposed the lo_min-max ratio and hi_min-max throughput ratio between Sta have reached 0.9 on all APs. Although some stations are obstructed by the shape of the buildings, door sand windows, after experiencing 6 iterations of AP connection switching the stations selected to be associate with the appropriate AP and the other result is transmission rate greater than the 3 previous methods. In this method the lowest station transmission rate is 12 Mbps. It shown in Table.1, for all the min-max ratios in the KNN_proposed exceeds all other the method.

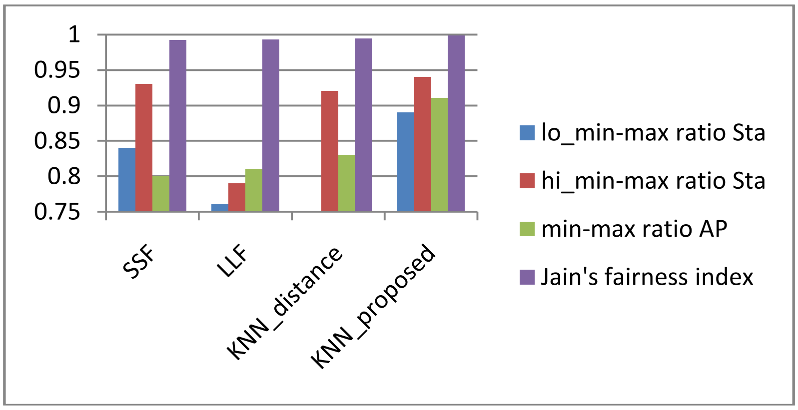

Fig.5. Comparison chart.

Fig.5 shows the comparison chart about min-max ratio Sta, min-max ratio AP and fairness index between 4 scheme tested in this paper. Last, by Table.2, by KNN proposed here the results improvement in min-max ratio throughput AP compares with others, there are with SSF: (0,91-0,8)*100% = 11%, LLF: (0,91-0,81)*100%=10%, KNN_distance: (0,91-0,83) *100% = 8%.

V. CONCLUSION

1. The K Nearest Neighbor method works well to determine the best cluster then to select the worst station.
2. The euclidian distance based on snr and transmission rate is can be a real distance between Sta and AP.
3. By K Nearest Neighbor for the selection of AP based on SNR and the transmission rate shows good improvement results compared to the SSF, LLF and KNN_distance methods by 11%, 10% and 8%., with minimum transmission rate 12 Mbps, while other 9Mbps.

REFERENCES

[1] Glenn Judd and Peter Steenkiste. Fixing 802.11 Access Point Selection. *ACM SIGCOMM Computer Communications Review Volume 32, Number 3: July 2002.*

[2] Yutaka Fukuda Takamitsu Abe Yuji Oie. Decentralized Access Point Selection Architecture For WirelesLANs. *0-7803-8246-3/04/$17.00 @ 2004 IEEE.*

[3] Murad Abusubaih, James Gross, SvenWiethoelter, Adam Wolisz. On Access Point Selection in IEEE 802.11Wireless Local Area Networks. *1-4244-0419-3/06/$20.00 ©2006 IEEE.*

[4] John N. Davies, Rich Picking, VicGrout. Predictionof Wireless Network Signal Strength within a Building. *Glyndŵr University Research Online. Article· January 2008.*

[5] Saeko Iwaki, Throughput Analysis and Measurement on Real Terminal in Multi-rate Wireless LAN.*ICUIMC'12, Feb 20–22, 2012, Kuala Lumpur, Malaysia.*

[6] Fumie Miki, Daiki Nobayashi, Yutaka Fukuda, Takeshi Ikenaga. Performance Evaluationof Multi-Rate Communication in Wireless LANs. *978-1-4244-5176-0/10/$26.00 ©2010 IEEE.*

[7] Huazhi Gong, KitaeNahm, Jong Wn Ki. Distributed Fair AP Selection For Multi Rate Ieee 802.11 WLANs.*IEICE TRANS. INF.&SYST., Vol.Ed 91-D, No.4 April 200.*

[8] Harashta Tatimma Larasati, Rifqy Hakimi, Tutun Juhana.Extended - LLF: A Least Loaded First (LLF) – Based Handover Association Control for Software - Defined Wireless Network. *VOL. 9, NO. 9, September 2017, 203–210 Available online at: www.ijceit.org. E-ISSN 2412-8856 (Online).*

[9] Mohammed Ahmed, Ali Alkahrsan, Muhammad Ilyas. Load Balancing In Hybrid Wifi/ Lifi Networks Based On The Rssi Of The Load Using Optimized Knn Clustering. *978-1-7281-9352-6/20/$31.00 ©2020 IEEE.*

[10] Raj Jain. A Quantitative Measure Of Fairness And Discrimination for resource allocation in shared computer system. *Unlimited distribution. Submitted for publication to ACM Transaction on Computer System. September 1984.*