MARMARA UNIVERSITY

INSTITUTE FOR GRADUATE STUDIES
IN PURE AND APPLIED SCIENCES

GAP FREE LOAD BALANCING IN WIRELESS LAN NETWORKS USING CELL BREATHING TECHNIQUE

İLHAN DEMİRCİ

MASTER THESIS
Department of Computer Engineering

ADVISOR
Asst. Prof. Ömer KORÇAK

ISTANBUL, 2014

MARMARA UNIVERSITY

INSTITUTE FOR GRADUATE STUDIES
IN PURE AND APPLIED SCIENCES

GAP FREE LOAD BALANCING IN WIRELESS LAN NETWORKS USING CELL BREATHING TECHNIQUE

İLHAN DEMİRCİ
(524111005)

MASTER THESIS
Department of Computer Engineering

ADVISOR
Asst. Prof. Ömer KORÇAK

ISTANBUL, 2014

MARMARA UNIVERSITY
INSTITUTE FOR GRADUATE STUDIES
IN PURE AND APPLIED SCIENCES

İlhan DEMİRCİ, a Master of Science student of Marmara University Institute for Graduate Studies in Pure and Applied Sciences, defended his thesis entitled "**Gap free Load Balancing in Wireless LAN Networks using Cell Breathing Technique**", on 11./06/2014 and has been found to be satisfactory by the jury members.

Jury Members

Asst. Prof. Dr. Ömer KORÇAK (Advisor) ....
Marmara University

Asst. Prof. Dr. Müjdat SOYTÜRK (Jury Member) ....
Marmara University

Asst. Prof. Dr. Didem GÖZÜPEK (Jury Member) D.Gözüpek ....
Gebze Institute of Technology

APPROVAL

Marmara University Institute for Graduate Studies in Pure and Applied Sciences Executive Committee approves that İlhan DEMİRCİ be granted the degree of Master of Science in Department of Computer Engineering, Computer Engineering Program on 16.06./2014. (Resolution no: 2014/17-37)

Director of Institute

Prof. Dr. Abdülkerim KAR

ACKNOWLEDGMENTS

I would like to thank all those who supported me and made this research project possible. First and foremost I would like to express my deepest gratitude and respect to my advisor Asst. Prof. Dr. Ömer KORÇAK for his guidance, helpful comments, constructive feedbacks, and suggestions throughout this study. I would like to thank Asst. Prof. Dr. Müjdat SOYTÜRK and Asst. Prof. Dr. Didem GÖZÜPEK for their valuable comments on improving this study. I extend my sincere appreciation to Dr. Ferhat UÇAN for his suggestions, guidance, and psychological support. I would like to thank to Dr. Etem DENİZ in TÜBİTAK, who has supported me during my thesis, not only for his valuable guidance, suggestions and criticism but also for his everlasting encouragement, trust, and his lasting positive attitude towards me.

The last but never the least, my deepest appreciation goes to my family. I am unable to express my thankfulness to my mother, my father, and my sister for their love and help in all my life.

And Selin ÖZEN, the kindest and loveliest person I’ve ever met, please accept my sincere gratitude for your support and extra help in writing of this thesis.

The research presented was supported by Commission of Scientific Research Project (BAPKO) under D-type project and TUBITAK – BIDEB (2224) Support for Participating Scientific Activities Abroad Program. I would like to thank Scientific Research Fund of Marmara University for their financial support in realizing this study.

May, 2014 **İlhan DEMİRCİ**

TABLE OF CONTENTS

ÖZET

Hücre Soluma Yöntemiyle Kablosuz Ağlarda Boşluksuz Yük Dengelemesi

Yük dengesinin eşit dağıldığı bir kablosuz alan ağı genellikle erişim noktalarında adaletli kullanıcı dağılımına sahiptir ve bu durum toplam ağ kazancının artırılmasına vesile olur. Fakat, kalabalık bölgelerdeki bazı kullanıcılar, erişim noktalarında adaletsiz yük dağılımına neden olabilir. Bu nedenle kablosuz alan ağı işletmecileri, popüler bölgelerde düzensiz olarak görülen kullanıcı yoğunluğunu dağıtmak için yöntem arayışı içerisindedirler. Mevcut yük dengeleme yöntemleri çoğunlukla kullanıcı ve erişim noktası tarafında, yazılım ve donanım desteğine ihtiyaç duyarlar. Hücre soluma, erişim noktası tarafında uygulanan ve kullanıcı cihazında değişikliğe ihtiyaç duymayan bir sinyal gücü ayarlama yöntemidir. Bu yöntem, bir yük dengeleme mekanizması olarak hücresel ağlarda tanınmış bir kavramdır ve biz bu yöntemi kablosuz yerel alan ağlarına uyguladık. Bu yaklaşım kullanıcılarda ya da haberleşme protokolünde değişikliğe gerek duymamaktadır. Tek değişiklik, erişim noktalarının işaret sinyallerinde gerçekleştirilmiştir ve kullanıcıların servis kalitesini düşürmemek için veri taşıma sinyalinin seviyesinde herhangi bir değişiklik yapılmamıştır. Yoğun erişim noktaları yük yoğunluklarını azaltmak için işaret sinyal seviyelerini düşürebilirler. Fakat, kapsama alanlarının küçültülmesi bazı bölgelere hizmet verilememesine neden olabilir.

Sinyal gücü ayarlama yöntemine dayanan Dinamik En Küçük-En Büyük Yük Dengeleme algoritması, kapsama alanlarında oluşabilecek boşlukları devingen şekilde kontrol ederek servis garantisi sunar. Biz, kullanıcı dağılımının heterojen yapısına uygun olarak bu algoritmayı iyileştirdik. Yük seviyelerini baz alarak erişim noktalarının sınıflandırılmasını sağlayan Boşluksuz İstatistiksel En Küçük-En Büyük Yük Dengeleme algoritması, erişim noktalarının kapsama alanlarını etkin şekilde ayarlayabilmek için geliştirilmiştir. Boşluksuz Çevirim içi En Küçük-En Büyük Yük Dengeleme algoritması, en küçük güç seviyesini devingen şekilde atayarak boşluksuz bir servis bölgesi sağlayabilmek için geliştirilmiştir. Kapsamlı bir dizi simülasyon deneyi sonucu, algoritmaların etkili ve adil olduğu gösterilmiştir.

May, 2014 **İlhan DEMİRCİ**

ABSTRACT

Gap free Load Balancing in Wireless LAN Networks using Cell Breathing Technique

A balanced Wireless Local Area Network (WLAN) generally has a fair user distribution among the Access Points (APs) and enables a maximized overall network throughput. However, there are some hotspot areas resulting in uneven load distribution among APs. Hence, WLAN administrators generally deal with the sporadic client congestion in hotspot regions such as popular locations within the network. Existing load balancing techniques mostly require hardware or software contribution at both AP and user side to control the association of users with APs. Cell breathing is a power adjustment technique, which is implemented at AP side and requires no modification at user devices. This technique is a well-known concept in cellular networks as a load balancing mechanism and we applied this technique to the WLAN network. This approach does not require modification to clients or the protocol. The only change is performed at the beacon signal level and the data transmission signal level is not changed in order to avoid degrading client's performance. Congested APs can reduce their beacon signal levels in order to decrease their load level. However, coverage reduction can create unserviced areas.

We provide a dynamic min-max load balancing algorithm based on power adjustment technique, which provides service availability guarantee by dynamically checking the coverage holes. We enhanced the algorithm by increasing awareness to heterogeneous nature of user distribution. We propose Gap-free Statistical Min-Max Priority Load Balancing algorithm, which provides classification for APs by means of their load level, in order to adjust the coverage areas effectively. In addition we propose Gap-free Online Min-Max Priority Load Balancing algorithm in order to provide a gap-free environment by defining the minimum power level assignments dynamically. Our extensive set of simulations showed that our proposed algorithms are effective and fair.

May, 2014 **İlhan DEMİRCİ**

SYMBOLS

$\boldsymbol{\mu_i}$: Minimum power level of AP_i

$\mathbf{A}$: List of access points

$\mathbf{AP_i}$: i^{th} access points

APLoc : Access point locations

$\mathbf{AP_{num}}$: Number of access points

$\mathbf{D}$: Maximum loaded access point index

$\tilde{\mathbf{d}}$: Recorded maximum loaded access point index

$\mathbf{F}$: Fixed access point list

$\mathbf{K}$: Maximum power level number

$\mathbf{MR_i}$: Minimum power range of AP_i

$\mathbf{p_i}$: Beacon power level of AP_i

$\mathbf{P_{max}}$: Maximum power level

$\mathbf{P_{min}}$: Minimum power level

$\mathbf{R}$: Rectangular region covered by access points

$\tilde{S}$: Maximum load value of non-fixed AP list

$\mathbf{U}$: Users

$\boldsymbol{\Lambda}$: Congestion load value

$\widetilde{\boldsymbol{\Lambda}}$: Congestion load of recorded state

$\boldsymbol{\lambda_i}$: Load of AP_i

ABBREVIATIONS

AC : Access Controller
AP : Access Point
BCC : Binary Convolutional Coding
BPSK : Binary Phase Shift Keying
BR : Beacon Range
CAPWAP : Control And Provisioning of Wireless Access Points
CCK : Complementary Code Keying
DPSK : Differential Phase Shift Keying
DSSS–OFDM : Direct Sequence Spread Spectrum – Orthogonal Frequency Division Multiplexing
ERP–DSSS : Extended Rate Physical layer – Direct Sequence Spread Spectrum
ERP–OFDM : Extended Rate Physical layer – Orthogonal Frequency Division Multiplexing
ERP–PBCC : Extended Rate Physical layer – Packet Binary Convolution Coding
FEC : Forward Error Correction
GF-MMPLB : Gap-free Min-Max Priority Load Balancing
GF-OMMPLB : Gap-free Online Min-Max Priority Load Balancing
GF-SMMPLB : Gap-free Statistical Min-Max Priority Load Balancing
GUI : Graphical User Interface
HR–DS : High Rate – Direct Sequence
HR–DSSS : High Rate – Direct Sequence Spread Spectrum
HT–OFDM : High Throughput – Orthogonal Frequency Division Multiplexing
IAPP : Inter Access Point Protocol
IEEE : Institute of Electrical and Electronics Engineers
LLF : Least Loaded First
LDPC : Low – Density Parity – Check Coding
LTE : Long-term Evolution
MAC : Media Access Control
MIMO : Multiple – Input/Multiple – Output

NP-hard	: Non-deterministic Polynomial-time hard
OFDM	: Orthogonal Frequency Division Multiplexing
QAM	: Quadrature Amplitude Modulation
QoS	: Quality of Service
QPSK	: Quadrature Phase Shift Keying
RSN	: Robust Secure Networking
RSSI	: Received Signal Strength Indicator
SC	: Single Carrier
SNMP	: Simple Network Management Protocol
SNR	: Signal-to-Noise Ratio
SQPSK	: Staggered Quadrature Phase-Shift Keying
SSF	: Strongest Signal First
WEP	: Wired Equivalent Protocol
WLAN	: Wireless Local Area Network
WPA	: Wi-Fi Protected Access
WS	: Wireless Station

LIST OF FIGURES

PAGE

LIST OF TABLES

PAGE

1. INTRODUCTION

Recently, Wireless Local Area Network (WLAN) networks have been deployed enormously and in many public places such as university campuses, businesses, and public places. Users having mobile devices can connect to the Internet using the Institute of Electrical and Electronics Engineers (IEEE) 802.11 protocol. Since each user selects independently the Access Point (AP) to associate with, the user distribution may compromise congestion load on APs. Uneven user distribution is the common characteristics of these types of networks, which results in overloading problem among the APs. Although most of the APs may be lightly loaded, some APs can get seriously loaded. An overloaded AP is defined as congested that could lead to slow performance or dropped connections. Hence, it is necessary to provide a load balancing solution to the standard WLAN protocol. An effective load balancing among APs can increase overall system throughputs.

IEEE 802.11 is a commonly used as a standard due to its advantages different from the similar standards e.g. 3G. Both technologies provide wireless access for wireless stations; however, 802.11 provides greater bandwidth, much faster connection, and cheaper deployment cost. In addition, since wired network is widely used and IEEE 802.11 is fully interoperable with Ethernet, the protocol is reasonable to access Internet.

An AP to a wired infrastructure is used by a user having mobile device with an 802.11 interface to send and receive frames. APs play the linkage role between wireless and wired worlds and serve as link-layer attachment points to the Internet. Since each user select the closest AP to associate required by the standard, traffic load between the APs might be unevenly distributed, resulting in congested regions and overloading problems. This network state results in low data throughput for users at the congested regions. Traffic load problems are most likely to occur in public access areas such as stations, airports, and convention or exhibition sites. Although several solutions have been proposed by the researchers, existing methods continue to exhibit insufficiencies.

In next sections, the history of the 802.11 standards is mentioned by giving detailed information about 802.11 a/b/g/n/ac/ad specifications. Then the current load balancing methods is explained in Section 1.2.

1.1. History of 802.11

The IEEE 802.11 standard was initially approved in 1997, and then the subsequent amendments are being proposed and added to the standard. Later specifications enabled a robust standard and comprehend variety of frequencies and connection speeds with worldwide regulations. When the specification was first approved, the maximum transmission rate was about 1 and 2 Mbps and the frequencies from 2.4 to 2.4835 GHz in the US and Europe [1]. In 1997 the 802.11a is defined providing transmission rates of 6, 9, 12, 18, 24, 36, 48, and 54 Mbps in frequency range of 5 GHz band available in the US. The 802.11b is the amendment including changes to the media access control (MAC) layer only to the standard, which is approved in 1999. This amendment provides maximum transmission rates of 5.5 and 11 Mbps for different modulations. In 2003, 802.11g amendment was approved, which aims to improve the frequency range of 2.4 GHz band. The maximum transmission rate of the 802.11g is 54 Mbps and this amendment is backward compatible with the 802.11b [1]. In 2009, 802.11n amendment was proposed that provides enhancements for higher throughput up to 600 Mbps. The amendment includes features like Multiple – Input/Multiple – Output (MIMO) with multiple antenna, frame aggregation, and security inclusion [2]. In order to increase the transmission rates to Gigabit speeds the 802.11ac and 802.11ad are being developed upon existing approved protocols. The IEEE 802.11ac provides maximum transmission rate of 7 Gbps with frequency range of 5 GHz band and 802.11ad provides about 6.7 Gbps within 60 GHz frequency range. 802.11ac is introduced in 2013 and 802.11ad is introduced in 2012.

Wi-Fi Alliance was formed in 1999 in order to certify and ensure the WLAN devices interoperable with same type of devices based on IEEE 802.11. Before the association is formed, the 802.11 devices had many issues in terms of communication and association. This Alliance certified the standard so that the devices are designed to cover the specification requirements resulting in reliable communication. Currently, more than 600 companied are joined to the alliance including Cisco, Conexant, Apple, and Samsung [3]. If a device is certified with Wi-Fi, the MAC and physical layer interoperability must be ensured for either 802.11 a/b/g, or n specifications. In addition the device must be included with WPA2 and EAP wireless security [3]. Devices with these certifications and enhancements provide a reliable usage in home, work, or public place.

Beside the Wi-Fi Alliance, wireless security is an important issue for the IEEE 802.11. Shared medium is utilized to transmit wireless signals. Hence, an intruder has the advantage to snoop easily all the unencrypted traffic data generated by the wireless devices in the same area. In order to provide a secure connection, Wired Equivalent Protocol (WEP) is introduced in 1999. The aim of the WEP is to provide security like a wired point to point connection. WEP was considered insecure in 2000 by the revelation of the weaknesses [4]. MAC filtering is used for security purposes; however an authorized user can easily sniff the packets by altering the MAC address of their own wireless card. The IEEE 802.11i proposed Robust Secure Networking (RSN) in 2001 in order to address the issues related to wireless security. However, it took some time to develop and accept the RSN standard [5]. Hence, need for finding a better security solution evoked the Wi-Fi Alliance to take an early draft of 802.11i that was Wi-Fi Protected Access (WPA) released in 2003. When the 802.11i RSN specification is completed, the full specification was implemented until the WPA2 standard is declared. Although these specifications have some weaknesses, WPA and WPA2 are the mostly used security standards.

1.1.1. 802.11a

The 802.11a standard proposed Orthogonal Frequency Division Multiplexing (OFDM) to the physical layer specifications in 1999. The standard specification is presented in hardware to the users around two years later. With this specification the 5 GHz radio frequency spectrum was utilized without license. Input frequency is encoded into multiple sub-channels transmissions with OFDM [5]. The divided sub-channels are multiplexed in order to create a faster uniformed channel. Like in traditional frequency division multiplexing, OFDM split up the spectrum into parallel channels. However, the difference is that the orthogonally split channels do not require space between sub-channels in order to provide protection against interference from neighboring transmitting channels in OFDM [5]. The range of 802.11a channels is 20 MHz and contains 52 sub-channels. The minimum bandwidth is 6 Mbps and maximum bandwidth can be 54 Mbps. Other than the original standard specification for the U.S.A. spectrum, other countries created additional specifications related to 802.11a for use.

1.1.2. 802.11b

The 802.11b standard specification proposed the High Rate – Direct Sequence (HR–DS) or High Rate – Direct Sequence Spread Spectrum (HR–DSSS) to the WLAN specification in 1999. The specification resembles the 802.11 DSSS, however high data rate version provides 11 Mbps bandwidth and uses 11 channels above 2.4 GHz range. 802.11 DS uses the original chipping method, but 802.11b uses Complementary Code Keying (CCK). Chipping stream is divided into 8-bit code symbols where code transformation enables more information to be stored in that particular symbol in CCK technique. Operating method of CCK is shown in Figure 1.1.

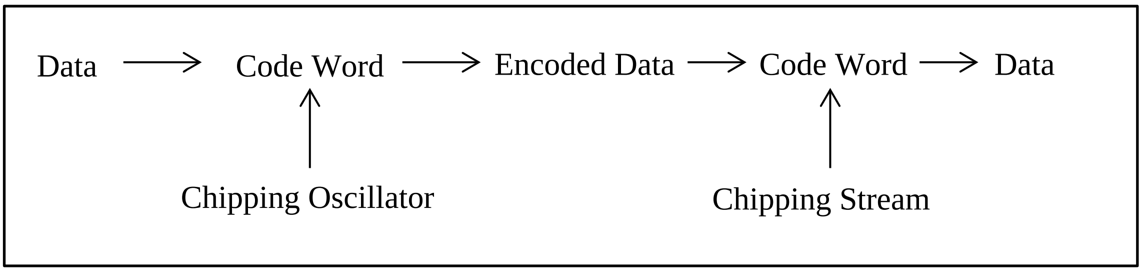


Figure 1.1. Code symbols in CCK

Differential Phase Shift Keying (DPSK) also is used by the IEEE 802.11b specification in order to encode data that is based on transmitting a bit code as a particular phase shift of a waveform. DPSK encoding can be illustrated as follows. DPSK uses following bit patterns: 00, 01, 11, and 10. These symbols correspond to phase shifts of 0, 90, 180, and 270 degrees respectively. During a transmission in order to form a bit pattern of 10, one could shift the waveform by 270 degrees. However, this technique is used with additional techniques to select phase angles in the 802.11b that is based on transmitting even or odd blocks. Additionally 802.11b has 5.5 and 11 Mbps speed and different DPSK techniques accordingly. This standard becomes popular and used in many places such as home and businesses, because it had sufficient bandwidth for many purposes [5]. However, later standards brought even better throughput and transmission speeds.

1.1.3. 802.11g

The 802.11g standard specification was introduced to provide higher bit rates in 2003. Additionally this standard is backward compatible with 802.11b and contains additional physical layer specifications. This standard includes the following physical layer modes:

- Extended Rate Physical layer – Packet Binary Convolution Coding (ERP–PBCC),
- Direct Sequence Spread Spectrum – Orthogonal Frequency Division Multiplexing (DSSS–OFDM),
- Extended Rate Physical layer – Direct Sequence Spread Spectrum (ERP–DSSS), and
- Extended Rate Physical layer – Orthogonal Frequency Division Multiplexing (ERP–OFDM).

ERP–PBCC specification is an optional extension that provides 22 and 33 Mbps speeds. However, this specification is not widely used. DSSS–OFDM is a hybrid specification. Headers of packets are encoded with DSSS and payloads of frames are encoded with OFDM in this mode. This is another optional extension, which is not widely used. ERP–DSSS physical layer specification provides 1, 2, 5.5, and 11 Mbps speeds and backwards compatible with the 802.11b. ERP–OFDM specification provides speeds of 6, 9, 12, 18, 24, 36, 48, and 54 Mbps. This mode is defined as 802.11a running in the 2.4 GHz range essentially.

The IEEE 802.11g specification generally involves the assuring that devices supporting 802.11b can work with devices supporting 802.11g without experiencing any problem. The 802.11g standard has the same number of channels and frequency ranges as in 802.11b because of the backwards compatibility [5].

1.1.4. 802.11n

802.11n standard specification proposed High Throughput – Orthogonal Frequency Division Multiplexing (HT–OFDM). This standard proposes up to 600 Mbps and uses

20 MHz channels operating in both 2.4 and 5 GHz bands. 40 MHz channels can be used instead of 20 MHz channels in order to allow higher capacity in the 5 GHz band.

Techniques like DPSK are utilized in order to modulate the signal; however the amplitude manipulation can also be performed in addition to phase shifting to encode additional information. Phase and amplitude modulation can be performed using Quadrature Amplitude Modulation (QAM) method to encode the signal. 16-QAM and 64-QAM encoding are used in 802.11n specification. The prefixed values indicate the number of symbols that can be encoded into carrier wave. In order to visualize the encoding, phase and amplitude information can be utilized. The phase can be indicated with the angle from the positive x axis and amplitude can be indicated with the distance from the origin. 16-QAM encoding is shown as an example in Figure 1.2. In the figure the constellation points refers to symbols on the grid, the amplitude is indicated by the arrow, and the phase is illustrated by the angle from the x axis to the symbol. Then, the value of this particular point can be decoded into 4-bit value [2].

These can be visualized using constellation points that show symbols on a grid where the angle from the positive x axis indicates phase and distance from origin indicates amplitude. Figure 1.2 shows an example of this using 16-QAM, where an amplitude is indicated by the arrow and the phase by the arch from the x-axis to the dot. Then the value of this particular point in the constellation can be decoded into a 4-bit value [6].

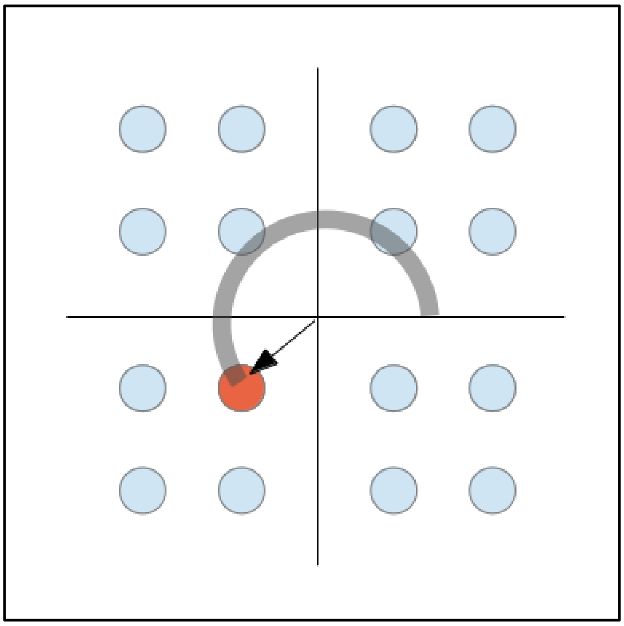


Figure 1.2. 16-QAM encoding

MIMO is another important feature of the IEEE 802.11n. In this method use of multiple antennas is allowed by stations in order to transmit and receive data simultaneously. Hence, by adding more antennas to the system the bandwidth can be greatly increased. For example, a person can buy a 3x3 antenna configuration for home use. The configuration consists of 3 transmitter antennas and 3 receiver antennas. The throughput is improved with this method and multipath interference is diminished [5]. This standard becomes popular and more widely used. Hence, recent mobile devices and wireless routers provide this capability.

The overall information about the standards is shown in Table 1.1. The table shows the frequency spectrum, typical data rate, maximum data rate, and modulation methods of the IEEE 802.11 a/b/g/n/ac and ad standards.

1.1.5. 802.11ac

The IEEE 802.11ac standard operates on 5 GHz unlicensed band, which is more advantageous. This standard provides wider bandwidth availability and lower interference levels as opposed to 2.4 GHz. The data rate is from 6.5 Mbps up to 7 Gbps achieved through combination of modulation schemes [7].

According to 802.11ac, the physical layer data subcarriers are modulated using Binary Phase Shift Keying (BPSK), Quadrature Phase Shift Keying (QPSK), 16-QAM, 64-QAM, and 256-QAM. Note that 256-QAM is not supported by 802.11n. Forward Error Correction (FEC) coding is used with coding rates of 1/2, 2/3, 3/4, and 5/6. Use of Binary Convolutional Coding (BCC) is mandatory, but Low – Density Parity – Check Coding (LDPC) is optional.

IEEE 802.11ac is backward compatible with 802.11n at 5 GHz ensuring the interoperability of 802.11ac and the already deployed 802.11n devices. In the development of 802.11ac, a variety range of usage models based on applications that require gigabit throughput are defined by IEEE. The high data rates of 802.11ac enable to operate more than one concurrent high bandwidth applications. The main usage of this high-rate includes wireless displays, video streaming, high definition televisions, downloading large files, video security, smart phones, schools, business automation etc.

Table 1.1. Evolution of the IEEE 802.11 standards

Standard	Year Introduced	Frequency	Typical Data Rate	Max Data Rate	Modulation
802.11a	1997	5 GHz	6 Mbps	54 Mbps	OFDM
802.11b	1999	2.4 GHz	5.5 Mbps	11 Mbps	HR–DSSS, HR–DS, CCK, DPSK
802.11g	2003	2.4 GHz	18 Mbps	54 Mbps	ERP–PBCC, ERP–DSSS, ERP–OFDM, DSSS–OFDM
802.11n	2009	2.4 / 5 GHz	74 Mbps	600 Mbps	HT–OFDM
802.11ac	2013	5 GHz	6.5 Mbps	7 Gbps	BPSK, QPSK
802.11ad	2012	60 GHz	385 Mbps	6.7 Gbps	SC, OFDM, QPSK

1.1.6. 802.11ad

The IEEE 802.11ad standard operates within 60 GHz band. While the range is short, data rate is from 385 Mbps up to 6.7 Gbps achieved through combination of modulation schemes [7]. Figure 1.3 shows the spectrum allocation for unlicensed worldwide operation of 60 GHz. This standard defines 4 channels, each with 2.16 GHz bandwidth operating at 60 GHz band.

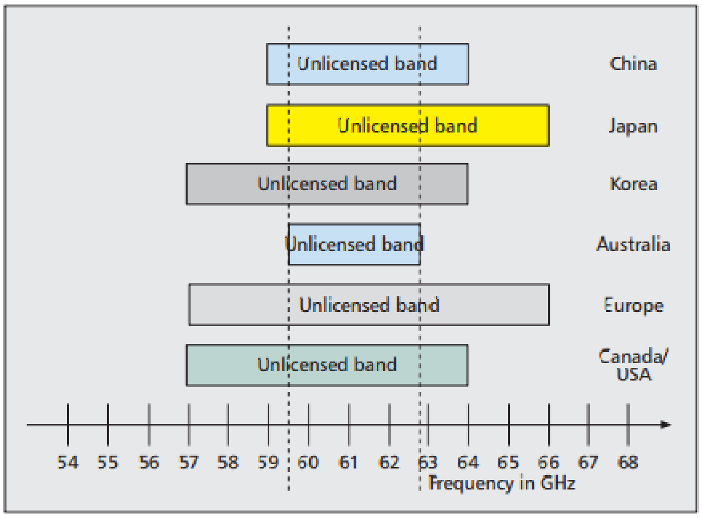


Figure 1.3. Unlicensed worldwide frequency allocation at 60 GHz band [7]

Single Carrier (SC) modulation and OFDM modulation are defined in 802.11ad specification. OFDM enables longer distance communication and greater delay spreads. The advantage is the flexibility in handling obstacles and reflected signals. OFDM allows Staggered Quadrature Phase-Shift Keying (SQPSK), QPSK, 16-QAM, and 64-QAM modulation with the maximum achievable physical layer data rate of 6.756 Gbps. Singe carrier physical layers consume low power and focus on small form factor devices like handsets. The p/2-B/SK, p/2-QPSK, and p/2-16-QAM modulations are used by SC with the maximum achievable physical layer data rate of 4.620 Gbps. In this standard, the data is encoded by an LDPC encoder with 1/2, 5/8, 3/4, and 13/16 code rates.

Table 1.2. Coding schemes and modulation of 802.11n, 802.11ac, and 802.11ad [7]

Standard	Code rate	Physical Layer (Mbps)					Spatial streams	Modulation
		20 MHz channel	**40 MHz channel**	**80 MHz channel**	**160 MHz channel**	**20 MHz channel**		
802.11n[2]	1/2	6.5	13.5	–	–	–	1	BPSK
802.11n	3/4	19.5	40.5	–	–	–	1	QPSK
802.11n	3/4	26	81	–	–	–	1	16-QAM
802.11n	5/6	65	135	–	–	–	1	64-QAM
802.11n	5/6	260	540	–	–	–	4	64-QAM
802.11ac[2]	1/2	6.5	13.5	29.3	58.5	–	1	BPSK
802.11ac	3/4	19.5	40.5	87.8	175.5	–	1	QPSK
802.11ac	3/4	39	81	175.5	351	–	1	16-QAM
802.11ac	5/6	65	135	292.5	585	–	1	64-QAM
802.11ac	5/6[1]	78	180	390	780	–	1	256-QAM
802.11ac	5/6[1]	312	720	1560	3120	–	4	256-QAM
802.11ac	5/6[1]	624	1440	3120	6240	–	8	256-QAM
802.11ad	1/2	–	–	–	–	385	1	p/2-BPSK
802.11ad	3/4	–	–	–	–	1155	1	p/2-BPSK
802.11ad	3/4	–	–	–	–	2310	1	p/2-QPSK
802.11ad	3/4	–	–	–	–	4620	1	p/2 16-QAM
802.11ad	13/16	–	–	–	–	6756.75	1	64-QAM

[1] Code rate of 3/4 for 20 MHz channel width.
[2] Guard interval = 800 ns.

Table 1.2 shows modulation and coding schemes used in 802.11n, 802.11ac, and 802.11ad to achieve multi-megabit and multi-gigabit physical layer data rates. Note that both 802.11n and 802.11ac support a long guard interval of 800 ns and optionally a short guard interval of 400 ns between transmissions of two symbols. The guard interval is 48.4 ns in 802.11ad. We assumed the long guard interval for 802.11n and 802.11ac in Table 1.2. With short guard interval, the data rates increase accordingly. For example, maximum data rate of 802.11n increases from 540 Mbps to 600 Mbps, and maximum data rate of 802.11ac increases from 6.240 Gbps to 6.933 Gbps. The 802.11n physical layer data rate is from 6.5 Mbps up to 600 Mbps, which is achieved through various combinations of modulation scheme, code rate, channel bandwidth, guard interval, and number of spatial streams. Physical layer data rate of 802.11ac is from 6.5 Mbps up to 6.933 Gbps. Physical layer data rates achieved by 802.11ac with 256-QAM modulation scheme, 800 ns guard interval, 40 MHz channel bandwidth, and 4 spatial streams is 720 Mbps. Physical layer data rates achieved by 802.11n with 64-QAM modulation scheme, 800 ns guard interval, 40 MHz channel bandwidth, and 4 spatial streams is 540 Mbps. By the advantage of 256-QAM modulation scheme, there is 33 percent increase in the physical layer data rate.

In previous sections, we give detailed information about the IEEE 802.11 a/b/g/n/ac and ad specifications. However in following sections, the network environment is assumed to support the IEEE 802.11b standard due to being popular and used in many places and it has sufficient bandwidth for many purposes.

1.2. Load Balancing Protocols

Recent load balancing protocols distribute the traffic load by controlling the association between users and APs. These protocols can be compared in terms of server requirement, load distribution control, modification required at AP side, and modification required at client side. The Table 1.3 shows that the load balancing protocols can be classified according to the proposed metrics.

Table 1.3. Load balancing methods in IEEE 802.11 networks [8]

Proposal	Needs a Server	Load-Distribution Control	Changes required at AP	Changes required at Client
Dynamic load balance	No	WS	Yes	Yes
Maximizing local throughput	No	WS	Yes	Yes
Vasudevan and colleagues' scheme	No	WS	No	Yes
Application-Layer Load Distribution Protocol	Yes	WS	No	Yes
Virgil	Yes	WS	No	Yes
US patent 2004/0039817	No	WS	Yes	Yes
Load balancing agent	No	AP	Yes	No
Cell breathing	No	AP	Yes	No
Admission control server	Yes	Server	Yes	Yes
Bejerano and colleagues' scheme	Yes	Server	No	Yes
US patent 2005/0213579	No	Switch	Yes	No

1.2.1. Client based load distributions

In client based approaches, users choose among the available APs that maximizes their potential benefits e.g. bandwidth, throughput etc. The users play an active role during association process. Since each user selects the APs for connection according to their own interest, this approach does not generally provide a system wide load balanced state. Given APs utilization is based on maximal available bandwidth, the implementation of load balancing technique becomes Least Loaded First (LLF) AP selection.

1.2.1.1. Load condition estimation

The AP load conditions can be received by the users in several ways. One of the techniques is the measuring channel utilization. Another technique used by the clients, calculating the delay between the actual and scheduled beacon-frame transmission times [9]. The proposed approaches require no network based modification. If a network with APs supporting IEEE 802.11e standard, the traffic population of APs can be broadcasted in a probe response or beacon frame preferably in a Quality of Service (QoS) Basic Service Set Element. These approaches maximize the local throughput and dynamic load balance.

A dedicated server can help measuring the load values in a WLAN network. Initially a user is associated with an AP in the application layer load distribution protocol. Then, the user accesses to the load metrics by accessing the dedicated server [10]. The server periodically polls the management information base object, which contains the throughput information, in order to maintain the load states of all APs to the users using Simple Network Management Protocol (SNMP). User decides to change the association to other available APs, if the new available bandwidth is superior to current value.

In 2006, a project proposed an association technique based on client based approach, called Virgin project [11]. In this approach, before the association, the user connects to a dedicated internet server to generate test traffic for each AP. This method provides a performance metric for each AP to the user. After the test, the user associates with the AP that provides the highest performance based on bandwidth and round trip time.

1.2.1.2. Association management

The AP association management of users can be classified as static or dynamic. In static approach, the user first selects the AP before the association operation and does not seek for a new association until the current connection lost. The most obvious drawback of the static AP selection is that the users cannot adapt to network changes. However, in dynamic AP selection, the user seek for an AP association, which increases the potential benefit, and can re-associate with another AP even if the current connection still holds. Since users are adaptive to network changes, this approach is better for highly dynamic networks. However, if the network environment changes too frequently, the user associations may get unstable, which means the user changes its association from one AP to another repeatedly, called Ping-Pong effect.

The main reason of the Ping-Pong effect is that the users change their association from one AP to another without any coordination. In order to illustrate the situation, assume that an AP is congested and another neighboring AP is powered up suddenly. At this moment, all the users decide to associate with new available AP simultaneously. The resulting state will be no different from the previous state, because the new AP get immediately overloaded due to burst migration. However, at this moment the first AP becomes lightly loaded than the second AP. Hence, the re-association repeats continuously.

The Ping-Pong effect can be avoided by using static AP selection or time variant association is required. One proposed method is users periodically search for the best AP providing the least load value [10]. When a least loaded AP is found, the user generates a random integer value, called *d*. The user switches from one AP to another only when the user identifies the candidate AP as the best for *d* successive trial. This scheme avoids burst association problem of users.

1.2.1.3. Pros and cons of client based approaches

The client based approach enables the developers to use commercial APs, because the developers do not need any change at the AP side. However, letting users to choose the APs results in association problems proposed in 1.2.1.2. An alternative solution can be managing the load distribution through the network side. Since users are unable to decide which AP to choose using the network based approach, the problems seen in client based schemes do not occur.

1.2.2. Network based load balancing techniques

In network based load distribution, a network side controller, such as an AP, a switch, or a server, manages the load distribution among the APs. The users play a passive role in the association process. There are three available techniques that enables APs to control their own load level; association management, admission control, and coverage adjustment.

1.2.2.1. Association management

In association management technique, a congested AP can transmit a disassociation frame to selected and connected users. The frame possibly contains a status code, which indicates the APs load state in order to show the inability to handle all currently connected users. Supposedly, the main objective of the technique is to re-associate the users with other lightly loaded APs. In theory, the best user, who received disassociation frame, re-associates in a way that balances the network load among related APs. However, this approach does not fit to frequently changing network environment, because the optimum solution holds only for the current state, not the changed state. One approach that finds decent candidate APs is that the AP might need to know the load level of neighboring APs and set of APs that are accessible by each of the associated users. This approach might lead to Ping-Pong effect as the re-associated users are disconnected again later on.

1.2.2.2. Admission control

In admission control technique, when an AP is congested, the new association requests are refused by that AP. If an AP is not congested, the users gain grant for their association request based on the workload status of the AP. Each AP is given a load level threshold. Hence, an AP accept the association request if the future load level after the association does not exceed the predefined threshold level. Load balancing agent approach illustrates the situation in [12].

1.2.2.3. Coverage adjustment

In coverage adjustment technique, congested APs reduce their beacon signal’s transmission power in order to drop the association of users that stay outside the signal coverage. Additionally new users becomes less likely to discover that AP because of reduction in beacon signal’s transmission power [13]. APs also change their transmission power in order to enable the lightly loaded APs to cover more area and eliminate the coverage holes [14]. This technique is applied to generally telephony systems referred as cell breathing.

1.2.2.4. Estimating load distribution

When the global load distribution is considered, the load status information of APs can be exchanged through a wired backbone. In order to provide the information, an existing protocol, such as Inter Access Point Protocol (IAPP), can be utilized with effective modification. The overloaded APs can use this information in order to relieve their load levels [12, 15]. As an alternative, a wired infrastructure server can be used in order to collect and distribute the load information of APs. Admission control server approach is illustrated in [16]. In this approach, the dedicated server is informed about the load distribution and recommend and direct the users to change the AP associations. This can be achieved by creating communication with the peers running on user devices. The ability to estimate the bandwidth consumption of each user can be used to decide on changing the association from one AP to another. This information makes it easy to determine the set of available APs that satisfies the implicit bandwidth demand of a particular user and estimate the load shift between APs for each handoff and determine the best AP for association.

Estimating the load levels of APs is currently difficult because the current protocols support multiple transmission rates. For example, the IEEE 802.11a that has the range from 6 to 54 Mbps, supports eight transmission rates. The time-variant channel conditions determine the actual transmission rate between AP and associated users. In addition, when users with high transmission rates are connected to low-rate wireless stations, the throughput of users suffer. Hence, the handoff decision becomes difficult because it is hard to predict the load shift due to association migrations.

1.2.2.5. Switch based load balancing

Switches can be used to manage the user-AP association. In this mechanism, a centralized switch is connected to a set of APs that provides load balancing functionality in WLAN network [17]. In this technique, the APs receive information from the switch to decide on accepting new users. In order to actualize the technique APs and switch must share some protocol. Hence, APs of different vendors may encounter interoperability issues.

1.3. Motivation

There are several approaches proposed for congestion relieving in WLAN networks. One promising approach is the cell breathing technique, which doesn't necessitate change in users' equipment. Cell breathing looks like the human breathing, which the APs changes their coverage area by modifying the beacon signal strength. By decreasing the signal strength the coverage area shrinks that looks like exhaling. Inversely, by increasing the signal strength the coverage area enlarge, which resembles inhaling. The mechanism behind the cell breathing is to drop some users from connecting to the AP by decreasing beacon signal strength when the cell becomes heavily loaded. The dropped users then redirected to the neighboring cells that are more lightly loaded in order to satisfy connectivity.

1.4. Our Contributions

Our solution to the load balancing is based on cell breathing technique. Our first concern is to provide load balancing and system utilization with sparse wireless AP deployment where coverage area shrinking can result in uncovered areas among them. For this purpose, we extend the min-max load balancing algorithm stated in [18]. Secondly, we propose a method to define minimum possible beacon power levels by using statistical network usage information. Lastly, we present an online deterministic min-max load balancing solution that determines the minimum beacon levels by considering the current load values.

1.5. Organization

A brief overview of this thesis follows. First, in the next chapter, we start with giving relevant information and point out the related works done in Load Balancing. Additionally, in the same chapter, we give the materials used during the development. We describe the problem definition by giving the methodologies, network environment and the service availability guarantee. Next in this chapter, the proposed algorithms are defined that are gap free min-max priority load balancing algorithm, gap free statistical min-max priority load balancing algorithm, and gap free online min-max priority load balancing algorithm. Then, in Chapter 3, the construction of simulation environment is explained and the load measurement method is described in detail. Then in the same chapter, the comparison of the algorithms in terms of load balancing capabilities is shown. We conclude the thesis with further discussion and future work in Chapter 4.

2. MATERIAL AND METHOD

2.1. Load Balancing in WLAN Networks

Recent studies show that, in most of the WLAN networks, majority of connections are generated among small number of APs [19]. 802.11b protocol does not provide any mechanism to balance the load in a congested network. Hence, researchers proposed techniques to manage the network load. These techniques focus on client-supported [16, 20-22] or network-controlled load balancing mechanisms [18, 23].

In [16], Balachandran focuses on client-supported load balancing algorithms. In order to relieve congestion, two approaches proposed; explicit channel switching and network directed roaming. Main purposes of the algorithms are; increasing bandwidth usage by dynamically directing users where capacity available, improving overall network utilization and guaranteeing minimum amount of bandwidth to users. It is expected that both users and APs cooperatively adapt themselves to the changing network load conditions. However, negotiation of such users with APs can only be achieved by an admission control protocol, which is not feasible for all available APs and mobile user devices. Additionally, directing users to available APs requires implementing software module in both user and AP sides, which may not be applicable to all devices. In [20], D. Gong and Y. Yang use an algorithm to maximize the MAC efficiency at the client side to achieve proportional fairness among users. In [22], A. Kumar focuses on utility maximization by associating the users to APs optimally using proportional fairness scheduling.

Client supported schemes require software or hardware modifications at the client side. On the other hand, network controlled schemes can achieve load balancing without any change in user equipment. Bahl et al. [6], proposed a cell breathing algorithm to maximize the overall network throughput. They formulate cell breathing problem as mixed integer linear programming, assuming continuous beacon power levels. A later work by Bejerano et al. also considers cell breathing approach and aim to achieve min-max fair load balancing [18]. Different than Bahl et al., they assume discrete levels of beacon power levels, which can be considered as more practical assumption. The network is min-max load balanced if there is no AP whose load can be reduced without increasing the load of other APs. The algorithms do not require user intervention; hence the algorithms in [18] can be classified as network-controlled schemes. Another approach proposed by S. Wang [23] that is also based on cell-breathing technique. They proposed Full BR algorithm to find the optimal beacon power level of APs, which covers a given region. Additionally, a polyhedron genetic algorithm is proposed, which provides a near optimal Beacon Range (BR) for each AP in the network environment.

In [18], a global optimal solution using a cell breathing method is proposed that uses discrete set of power levels, which is identical with our power level assignment, but APs must be located without causing gaps even if all the APs transmit at minimum power level. By assigning the minimum beacon power levels using radio coverage survey method, they provided a static solution to coverage hole issue. However, in our algorithms we provide a gap-free environment by defining the minimum power level assignments dynamically. In addition to this, [23] provides an algorithm for assigning beacon power levels with a service availability guarantee in a continuous manner. In this work, we use more realistic assumption stated in [18], such that only discrete power levels are feasible.

2.2. Materials

MatLab R2013a Development Environment is used for analysis and algorithm development. In this study, a notebook computer is used as research material. Each sample network environment with user and AP locations are prepared in Cartesian coordinate system. The simulation results are printed in a Graphical User Interface (GUI) prepared in MatLab. The GUI consists of four buttons as shown in Figure 2.1. This output helped to visualize the user distribution among the network and to show the association between users and APs. The information on each AP is utilized that shows the number of users connected, load, and beacon power level. Each O symbol represents an AP and X symbol represents users. The circles around APs represent the beacon coverage areas.

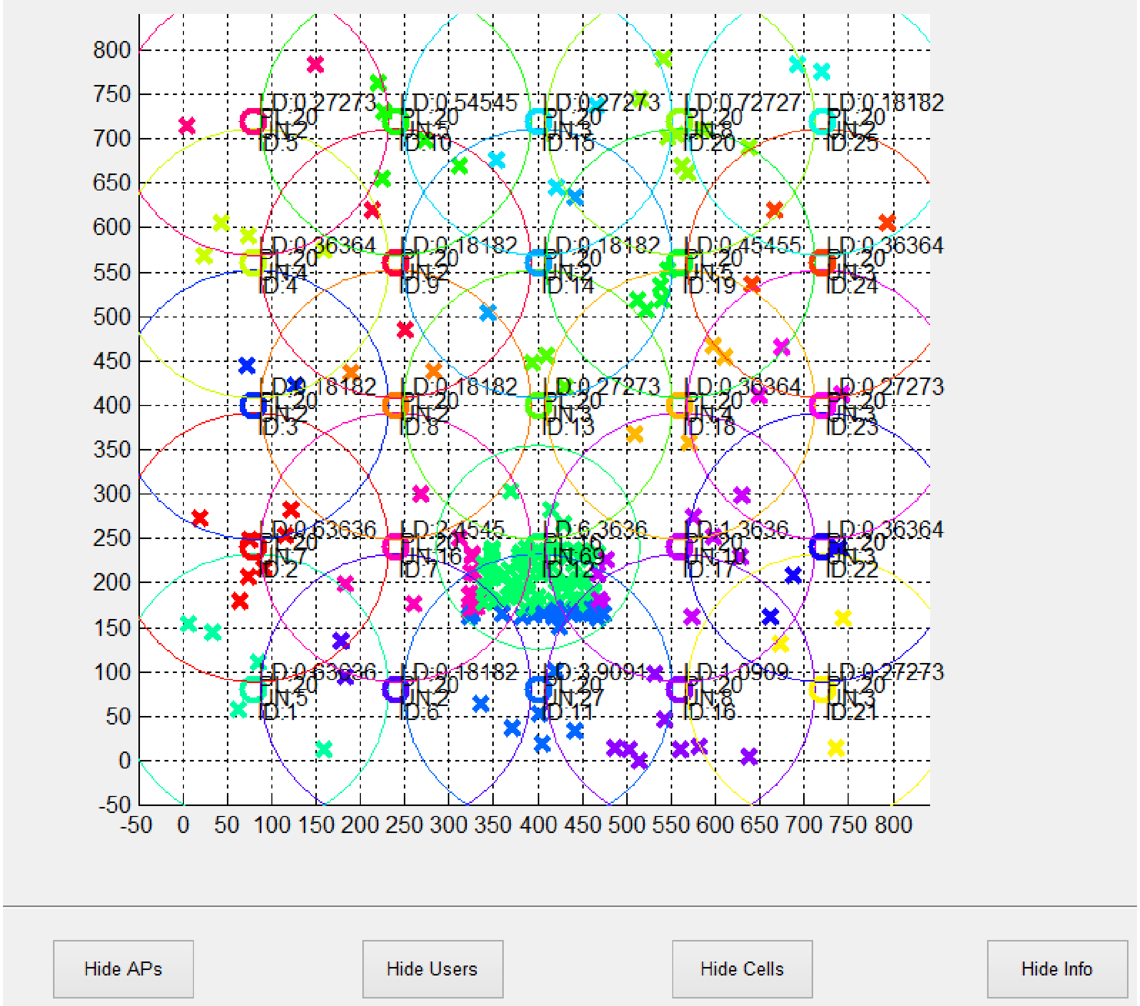


Figure 2.1. The GUI showing the user distribution among the network

In addition, the algorithms provide service availability guarantee stated in section 2.3.3. This constraint satisfied with defining minimum values for beacon power levels. Minimum power levels of each AP are determined with *Minimum Range Coverage* algorithm (Algorithm 1). The output of the algorithm is shown in Figure 2.2. Each AP is symbolized with O. The minimum power level values are provided on each AP.

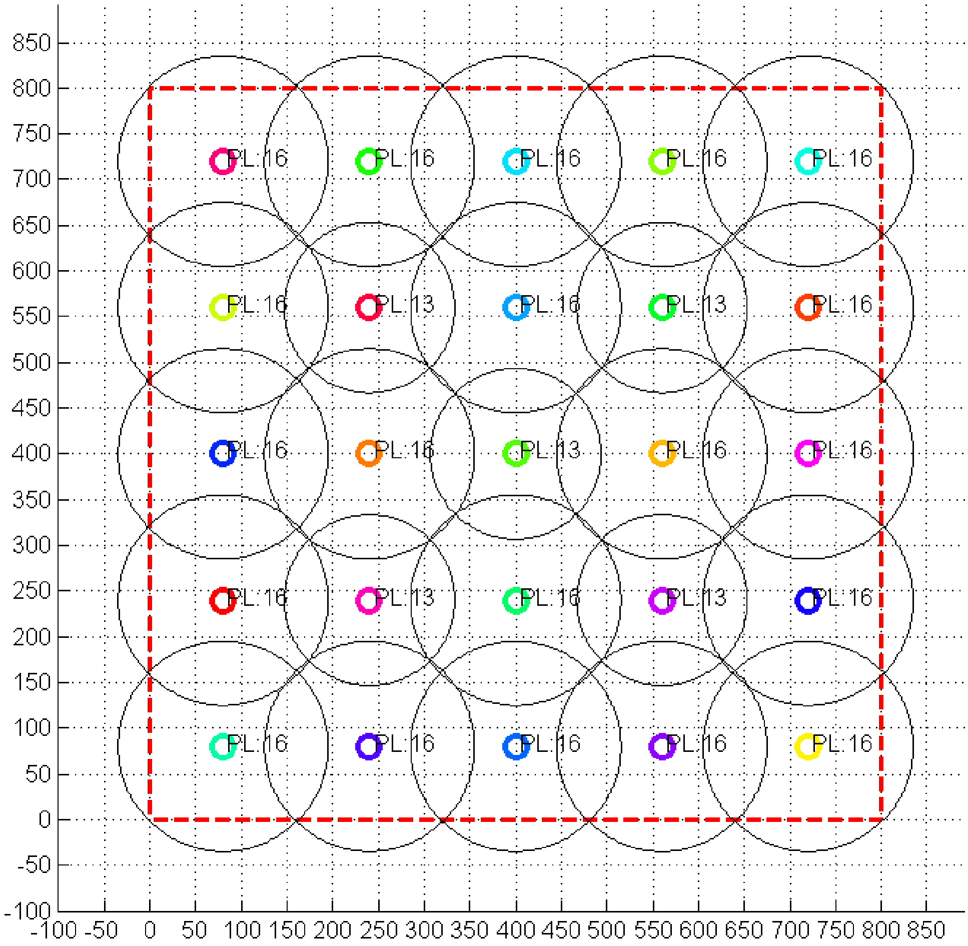
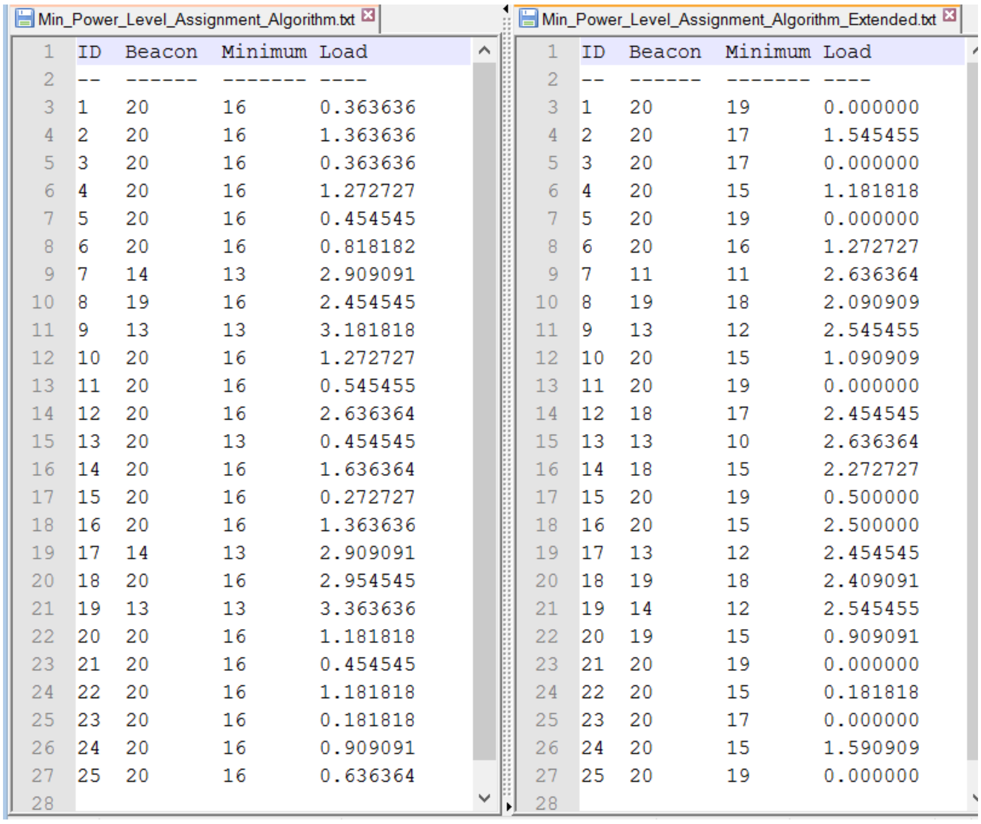


Figure 2.2. Minimum values of APs beacon power levels

Finally the comparison of two APs is performed using the outputs provided by each algorithm as shown in Figure 2.3. The output information is beacon power levels, minimum power levels, and the load values of each AP after the algorithms executed.



Min_Power_Level_Assignment_Algorithm.txt

```
ID  Beacon  Minimum Load
--  ------  ------- ----
1   20      16      0.363636
2   20      16      1.363636
3   20      16      0.363636
4   20      16      1.272727
5   20      16      0.454545
6   20      16      0.818182
7   14      13      2.909091
8   19      16      2.454545
9   13      13      3.181818
10  20      16      1.272727
11  20      16      0.545455
12  20      16      2.636364
13  20      13      0.454545
14  20      16      1.636364
15  20      16      0.272727
16  20      16      1.363636
17  14      13      2.909091
18  20      16      2.954545
19  13      13      3.363636
20  20      16      1.181818
21  20      16      0.454545
22  20      16      1.181818
23  20      16      0.181818
24  20      16      0.909091
25  20      16      0.636364
```

Min_Power_Level_Assignment_Algorithm_Extended.txt

```
ID  Beacon  Minimum Load
--  ------  ------- ----
1   20      19      0.000000
2   20      17      1.545455
3   20      17      0.000000
4   20      15      1.181818
5   20      19      0.000000
6   20      16      1.272727
7   11      11      2.636364
8   19      18      2.090909
9   13      12      2.545455
10  20      15      1.090909
11  20      19      0.000000
12  18      17      2.454545
13  13      10      2.636364
14  18      15      2.272727
15  20      19      0.500000
16  20      15      2.500000
17  13      12      2.454545
18  19      18      2.409091
19  14      12      2.545455
20  19      15      0.909091
21  20      19      0.000000
22  20      15      0.181818
23  20      17      0.000000
24  20      15      1.590909
25  20      19      0.000000
```

Figure 2.3. Output information of algorithms to compare each AP

2.3. Problem Definition

2.3.1. Methodology

Cell breathing technique is one of the approaches as a solution to load balancing. In traditional WLANs, users scan all available channels and associate with AP having strongest Received Signal Strength Indicator (RSSI). However, RSSI information is insufficient alone to design a strong algorithm enabling users to be distributed among the network. Our contribution is to enhance the methodologies suggested in [18]. We provide an algorithm to determine minimum beacon power levels depending on locations of APs. We also provide two extensions to this algorithm for better adjusting minimum power levels to improve load balancing considering the loads of APs. The first concern is based on load balancing and system utilization with sparse wireless AP deployment where coverage area shrinking can result in uncovered areas among them.

In popular places having wireless connection to the network, it is required to have neighboring APs around to satisfy the bandwidth need and distribute the internet connection among people. Since users are connected via wireless devices, they are generally in mobile state. Hence, the probability of having some APs heavily loaded is inevitable in crowded places. The congestion at an AP results generally in minimization of throughput and unfair network distribution among the users. In Example 1, the necessity of a routing algorithm is depicted.

Example1: As an illustration, let's assume there are 2 APs according to the IEEE 802.11b standard and 7 users with tablet, laptop or smart phone etc. Bandwidth rate of APs is defined as 40 Mbps and users' download rate is defined as 8 Mbps. Assuming first five users are associated with the first AP where the bandwidth usage is at full capacity. Last two users associated with the second AP. In this scenario, if an additional 8th user enters the coverage area of both APs, there is no other option but new user can be associated with either one of the APs according to the standard. Because, by default, a user scans all available channels and associates itself to the AP with the strongest RSSI value. Hence, association of the new user with the first AP can decrease the download rate of all first five users.

The example shows that, the standard association algorithm is not well designed to distribute load among APs. Hence, additional information is required to determine the load of APs and distribute the users evenly.

2.3.2. Network environment

In my thesis, we consider WLAN network with multiple APs in region R. AP regions are assumed to be circular. There are two signal levels for each AP; data transmission signal level and beacon signal level. Beacon signal level can be equal to or less than the data transmission signal level. We assume that with highest beacon signal level, the whole area R can be covered by APs. We provide a full data coverage area in region R where each point in the area is served by at least one AP. APs transmit the beacon signals in discrete manner. Each AP can use one of $K+1$ beacon power levels, denoted by $\{P_k \mid k \in [0...K]\}$, where the minimal and maximal levels are denoted by $P_{min} = P_0$ and $P_{max} = P_K$.

In next sections a notation called minimum power level, μ_i, is mentioned, which defines the minimum value of beacon power level that an AP can set without violating service availability. So, an AP i is able to set its beacon power level, p_i, to the power levels that are higher than μ_i. On the other hand, APs transmit the data traffic with maximal power, i.e. data transmission signal level of an AP is always equal to P_{max}. In this environment, beacon signal adjustment of APs is handled by an Access Controller (AC), and communication between AP and AC can be based on a protocol like The Control and Provisioning of Wireless Access Points (CAPWAP) [24].

2.3.3. Service availability guarantee

In sparse AP deployments, reducing beacon signal level may expose some uncovered areas (gaps), which violate service availability. However, in our algorithm we propose a minimum power level assignment of AP to prevent users being out of network service even if each AP set their minimum power level, μ_i, to the assigned minimum level. Hence, it is ensured that whole area R is covered. Let BR_i is the circular range that is served by AP_i and p is a point defined in R. Enclosure of R is satisfied if and only if the following conditions stated in [23] hold;

1) Boundary of R intersects with at least one BR.

2) If $p \in R \cap BR_i$, then $\exists\, j$ such that $p \in BR_j \wedge p \notin$ boundary of BR_j.

3) If $p \in R \wedge p \in$ boundary of $BR_i \cap BR_j$, where $i \neq j$, then $\exists k$ such that $p \in BR_k \wedge$ p $\neq$ boundary of BR_k.

Our algorithms check these conditions for detecting possible gaps in the service area. The algorithm is as follows:

```
Algorithm 1 Minimum Range Coverage
Input: $A, R, p_i = P_{max}$ provides full coverage in $R$
Output: State if full coverage is satisfied in $R$
 1: $fullcover \leftarrow false$
 2: $U \leftarrow NULL$
 3: if $MR_i \cap \partial R = NULL$ for $\forall i \in \{1 \ldots AP_{num}\}$ then
 4:     $fullcover \leftarrow false$
 5:     $U \leftarrow U \cup \{i | i$ is the closest AP to $\partial R\}$
 6: end if
 7: for each $p \in (\partial MR_i \cap \partial R)$ do
 8:     if $p \notin MR_j$ for $\forall i \neq j$ then
 9:         $fullcover \leftarrow false$
10:         $U \leftarrow U \cup \{i | i$ is the closest AP to $p\}$
11:     end if
12: end for
13: for each $p \in R \wedge p \in \partial MR_i \cap \partial MR_j, i \neq j$ do
14:     if $p \notin MR_k$ for $\forall k \neq i, j$ then
15:         $fullcover \leftarrow false$
16:         $U \leftarrow U \cup \{i | i$ is the closest AP to $p\}$
17:     end if
18: end for
19: if $U = NULL$ then
20:     $fullcover \leftarrow true$
21: end if
```

The objective of *Minimum Range Coverage* algorithm is to check whether the region R is covered by the minimum power levels, μ_i, of APs. The algorithm returns TRUE if the region is covered, otherwise the algorithm returns FALSE. Initially, minimum power levels are assigned to P_{max}. This configuration must provide full coverage in region R. Initially AP list A and predefined region R are given as input. The output of the algorithm is the check, which indicates if the full coverage is satisfied or not. U is the set of APs whose minimum power range, MR_i, need to be adjusted to provide full minimum power coverage. If U is empty, then full MR_i coverage is satisfied. From line 3 to 6, for all AP_i, if the intersection of MR_i, with the boundary of region R, ∂R, is empty, then we set the *fullcover* to false and closest AP to the ∂R is put in U. From line 7 to 12, for each crossing point of MR_i and ∂R, if there does not exist an MR_j, which covers the crossing point p, then we set the *fullcover* to false. Additionally closest AP to the p is put in U. From line 13 to 18, for each crossing point of MR_i and MR_j, if there does not exist an MR_k, which covers the crossing point p, then we set the *fullcover* to false. Additionally closest AP to the *p* is put in *U*. Finally, the algorithm returns true if the set *U* is empty since there does not exist any AP whose BR needed to be adjusted.

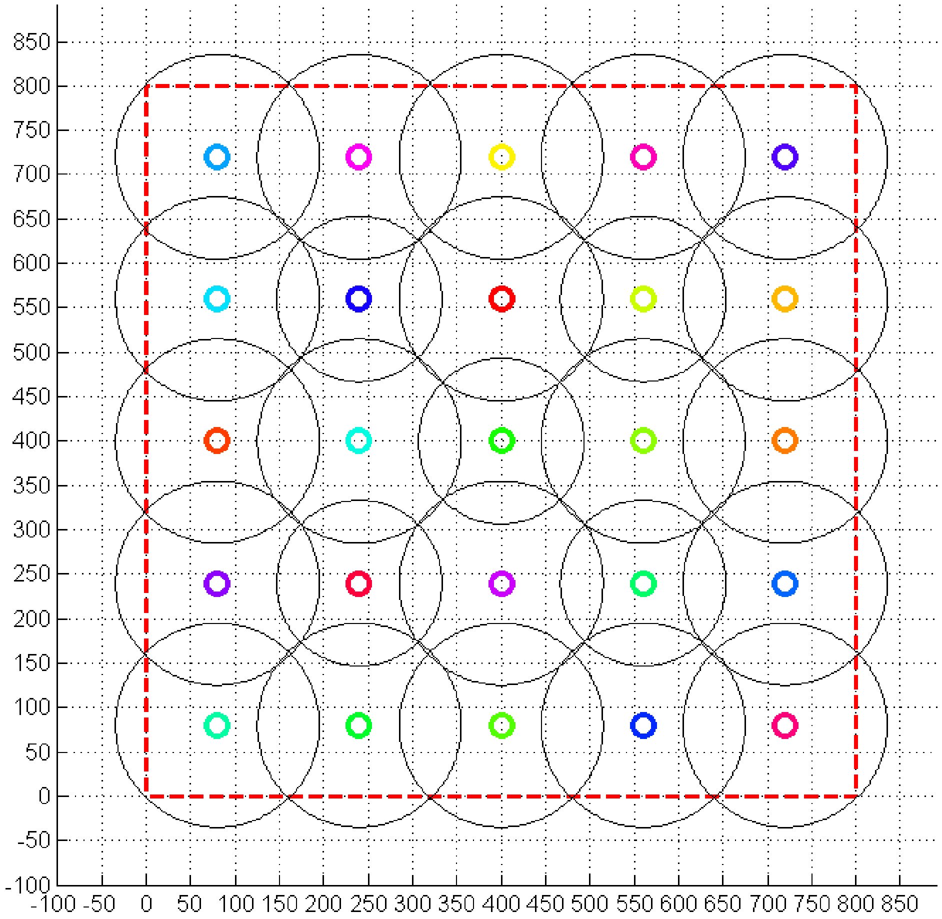


Figure 2.4. Four hotspot point using set minimum power levels algorithm in 800x800 m^2 region

The *Set Minimum Power Levels* algorithm in 2.4.1 is used to set the possible minimum μ_i values of APs that covers the *R*. Initially the μ_i values are set to P_{max}. Each iteration Minimum Range Coverage algorithm is called to check if the coverage is satisfied. For each AP *i*, if the algorithm returns false, then the μ_i value of AP_i is decreased by one. This process repeats until the algorithm returns true. Finally, all μ_i values are set to the possible minimum power levels equally. Load Balancing Algorithms uses the above algorithm to check the service availability guarantee. The μ_i levels of APs after Set Minimum Power Levels algorithm is depicted in Figure 2.4.

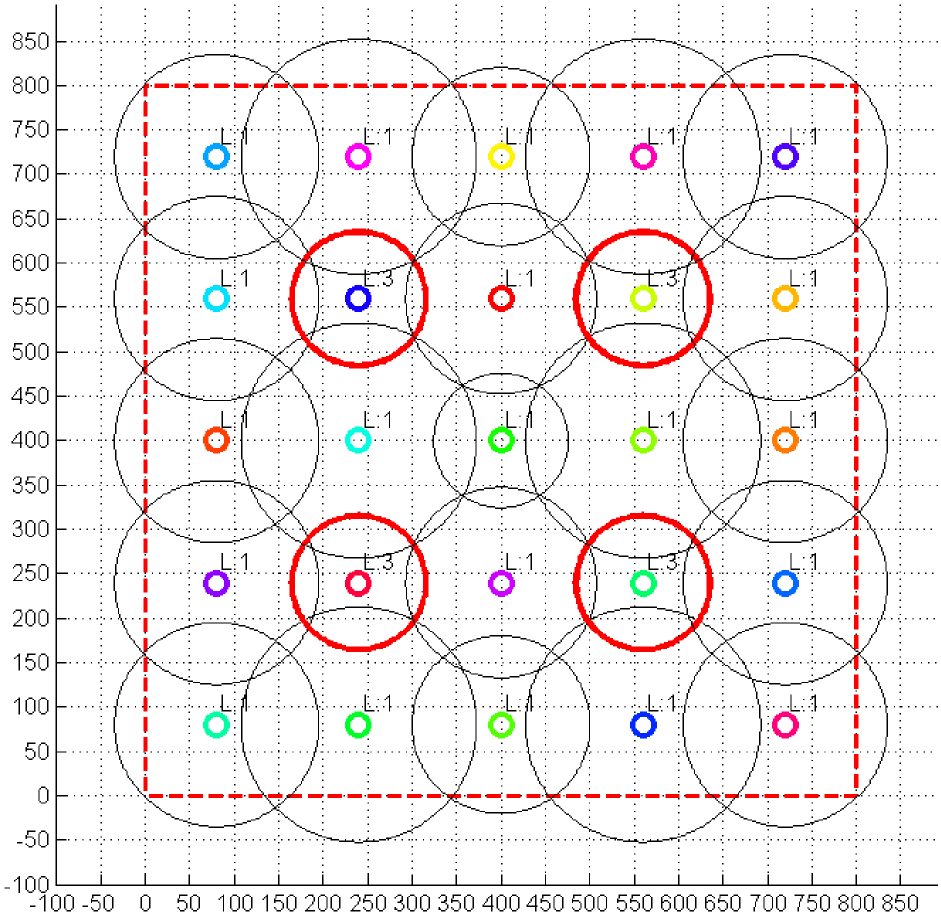


Figure 2.5. Four hotspot point using set statistical minimum power levels algorithm in 800x800 m^2 region

The *Set Statistical Minimum Power Levels* algorithm in 2.4.2 works as similar with algorithm stated above, which is used by the Gap-free Statistical Min-Max Priority Load Balancing (GF-SMMPLB) algorithm. The only difference is each AP is given a level from the set [1…L]. The APs having the highest level, *L*, are the most loaded ones and the APs having the lowest level are the least loaded among all APs. Starting from the level L, the μ_i values of APs is decreased until any further reduction violates the service availability in *R*. All the APs of each level are traversed accordingly. Since the reduction process started from the most loaded APs, their minimum coverage areas become smaller, which enables them to decrease beacon power level, p_i, more than other APs in the environment. The μ_i levels of APs after Set Statistical Minimum Power Levels algorithm is depicted in Figure 2.5. It is clear from the figure that the hotspot points have smaller μ_i coverage areas.

2.3.4. Problem statement

A network is in Min-Max Load Balanced state, if the following condition holds. Assume Y is the set of load values of APs, $Y = \{y_1, \dots, y_{AP_{num}}\}$ where AP_{num} is the number of APs in the network. If there does not exist a set of load values $Y' = \{y'_1, \dots, y'_{AP_{num}}\}$, which is same or lexicographically lower load value, then we say the network is in Min-Max Load state. Min-Max Load Balancing is Non-deterministic Polynomial-time hard (NP-hard) problem, which means it is very hard to find an optimal solution [18]. We can illustrate the problem of finding min-max load balancing is NP-hard with an example:

Example 2: Given a WLAN network having some of the APs congested, set a lower bound for congestion load value, which can be achieved by the cell breathing approach. The problem of finding a min-max load balanced state by minimizing the number of congested APs is NP-hard as stated in [18]; therefore, as the number of congested APs increases, the probability of finding an optimal solution decreases dramatically.

Hence, a variant of min-max load balancing problem is presented. For each AP i there exist a unique priority, called $w_i \in [1 \dots AP_{num}]$ and aggregated load called l_i. The definition of priority load is described in 2.4.1. Using priority load definition, at any network state, the congested APs set contains only one AP. Using this property, the min-max load balancing solution is simplified, because at each iteration, only unique congested AP is defined.

In addition to min-max load balancing problem, the users must be covered by at least one AP in the network. There are two critical parameters involved in this process; the beacon range and the data transmission range. The BR of an AP i is the range within which the beacon power of AP i can be sensed. The data transmission rage of an AP i is the range within which an associated client can transmit and receive data through AP i.

Using these definitions, a user can associate with an AP only when that user is located within the BR of that AP. A user can communicate with an AP only when that user is located within the data transmission range of that AP. When a client actively searches the existence of APs, the AP having strongest RSSI is selected. We assume that the data transmission power is set to maximum available power level in order to avoid degrading client's performance.

2.4. Proposed Algorithms

In next sections, gap free min-max priority load balancing, gap free statistical min-max priority load balancing, and gap free online min-max priority load balancing algorithms are presented and explained in detail, respectively.

2.4.1. Gap-free min-max priority load balancing algorithm

We considered min-max priority load balancing problem that is defined in [18]. Each AP a is given a distinct priority w_a. Priority load of an AP a, denoted by λ_a is defined as the ordered pair $\lambda_a = (l_a, w_a)$ where l_a is the aggregated load of AP a. AP a has higher load than AP b if $\lambda_a > \lambda_b$, i.e. if i) $l_a > l_b$ or ii) $l_a = l_b$ and $w_a > w_b$. A network state is said to be min-max priority load balanced if there exist no way to reduce the load of any AP without increasing the load of another AP with higher load. Min-max Priority Load Balancing algorithm stated in [18] iteratively finds min-max priority load balanced state. Our first algorithm is based on this algorithm; however we also check the gaps and provide service availability. In [18], it is assumed that APs are deployed in such a way that there is no gap even if their beacon powers are set to the lowest level. Our limitation is that there should not be any gap even if the beacon powers are set to the highest level.

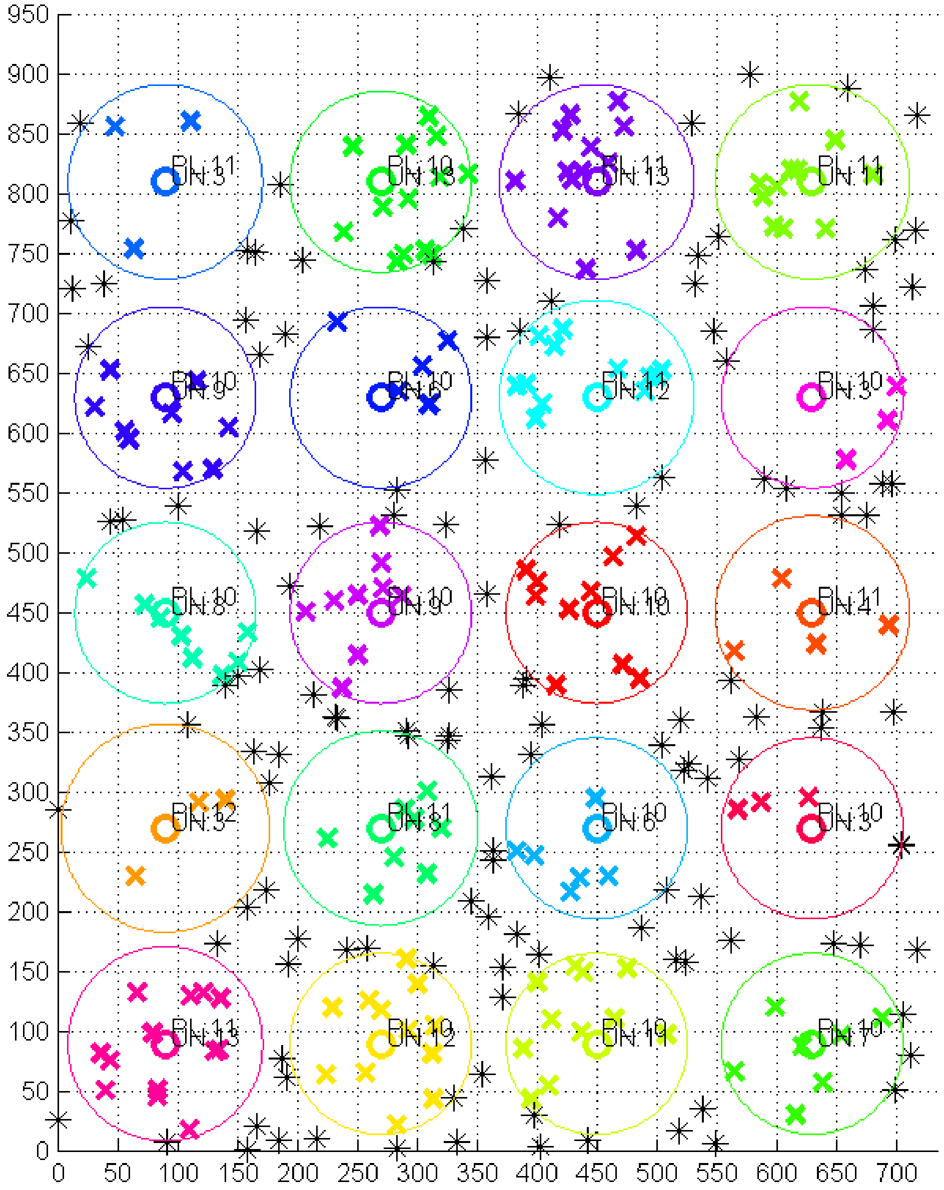


Figure 2.6. Sparse AP deployment resulting in unserviced areas using min-max priority load balancing algorithm

Figure 2.6 represents the network setting, which is not recommended by [18]. The PL values represent the beacon power levels assigned by the min-max priority load balancing algorithm. UN represents the number of users that are associated with AP. The service area is 720x900 m^2 and the distance between the APs is 180 meters. However minimum coverage area of an AP is 75 m. Hence, when neighboring APs transmit the beacon power level at P_{min}, then coverage holes might exist. The circles represent the beacon power coverage of APs after the min-max priority load balancing algorithm is executed. The X's represents the users associated with the AP and * (star) represent the unassociated users with any APs. Each AP that is the O symbol in the graph has its own beacon coverage area and X users inside the beacon coverage area is connected to that AP.

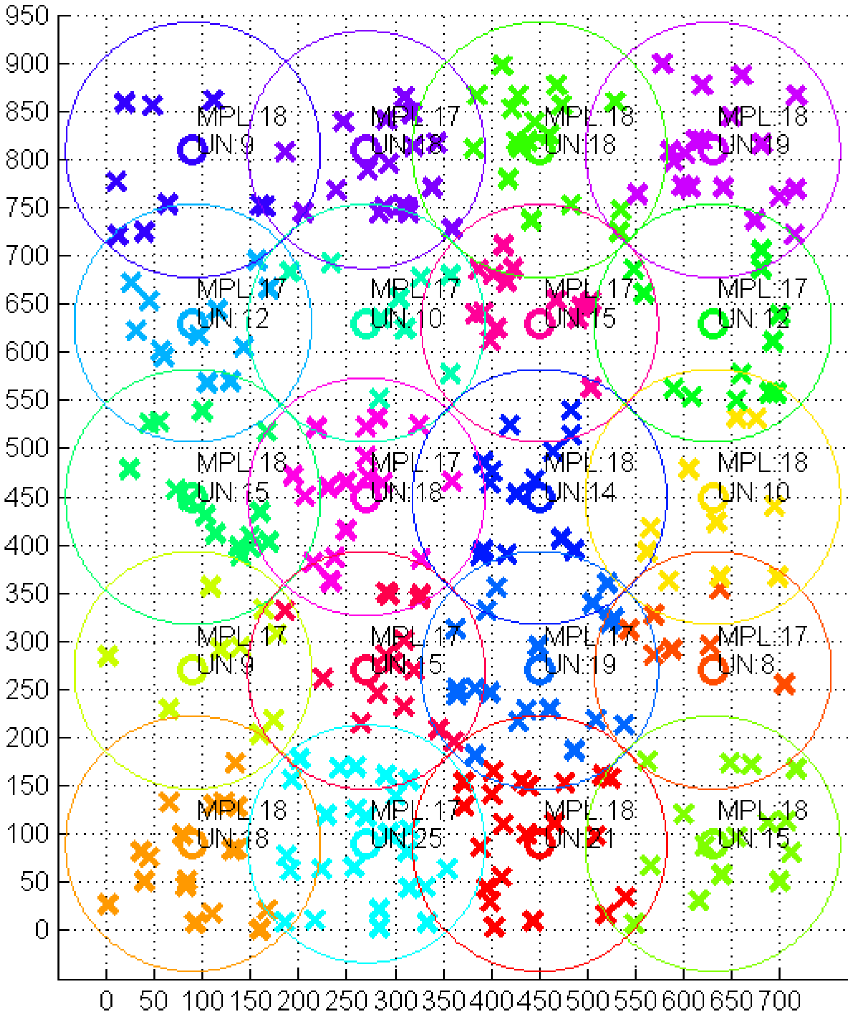


Figure 2.7. Minimum power level demonstration that covers all the area providing service availability

Figure 2.7 shows the same network environment as in Figure 2.6. However, the only difference is the minimum power level, μ_i, of the APs are set without causing gaps. The MPL values represent the minimum power levels assigned by *Set Minimum Power Levels* algorithm. UN represents the number of users that are associated with AP. The beacon power levels cannot be lower than minimum power levels. By providing this constraint, service availability is guaranteed.

Gap-free Min-Max Priority Load Balancing (GF-MMPLB) algorithm is described in Algorithm 2. Algorithm uses three inputs: list of users, U, list of APs, A, and a rectangular closed region, R. In order to ensure the service availability in a closed region, we need to provide a gap-free environment such that the region R should be covered by data transmission signals of available APs. Network state S is defined as beacon power levels, p_i, and minimum power levels, μ_i of all APs in the network. Initially, for all AP i, we set the minimum power level, μ_i and beacon power level, p_i to P_{max}. The *Set Minimum Power Levels* algorithm works as follows. We decrement μ_i values of all APs one by one, until any further reduction violates the service availability in R. This provides a gap-free environment and ensures the service availability. We have a fixed AP list called F, which is initially empty. If an AP is determined as fixed, $\tilde{d}$, from Routine called *Minimize-m-coordinate*, then the AP is stored in F. There exist a variable called S, which stores the minimized mth coordinate of the load vector, without modifying the load of APs stored in F. The Algorithm 2 terminates when all APs are stored in F. At the end, we fairly set the minimum power levels of each AP.

Minimize-m-Coordinate routine, which is described in [18], minimizes mth coordinate of the load vector at iteration m without modifying the load of APs stored in F. This routine assumes “limited knowledge” model, in which user-AP association and AP load may not be known a priori for all possible beacon power settings, since such information may not be readily available. Instead, only information on the user-AP association and AP load for the current beacon power setting is available. Input variables used in the algorithm are listed accordingly.

Algorithm 2 GF_MMPLB

```
Input: $U$, $A$, $R$
Output: List of APs where users in R associated with
$S \leftarrow \{(a, p_a = P_{max}, \mu_a = P_{max}) \mid \forall a \in A\}$
$F \leftarrow NULL$
while $F \neq A$ do
    for $i$=1...$AP_{num}$ do
        $\mu_i = \mu_i - 1$; if gap occurs $\mu_i = \mu_i + 1 \wedge F \leftarrow F \bigcup i$
    end for
end while
$F \leftarrow NULL$
while $F \neq A$ do
    S,d $\leftarrow$ Minimize_m_Coordinate($S$, $F$)
    $F \leftarrow F \bigcup d$
end while
```

Routine 1 Minimize_m_Coordinate

```
Input: $S_{init}$, $F$
Output: $S$, $\tilde{d}$
$\tilde{S} \leftarrow S_{init}$
$\tilde{\Lambda} \leftarrow \Lambda = \max_{a \in A - F} \lambda_a$
$\tilde{d} \leftarrow d = a$ s.t. $a \in A - F \wedge \lambda_a = \Lambda$
$end\_flag \leftarrow FALSE$

while $end\_flag == FALSE$ do
    if $p_d == \mu_d$ then
        $end\_flag \leftarrow TRUE$
    else
        $p_d \leftarrow p_d - 1$
        $\Lambda \leftarrow \max_{a \in A - F} \lambda_a$
        $d \leftarrow a$ s.t. $a \in A - F \wedge \lambda_a = \Lambda$
        if exist $a \in F$ s.t. $\tilde{y}_a < \lambda_a$ then
            $end\_flag \leftarrow TRUE$
        else if $\Lambda < \Lambda$ then
            $\tilde{S} \leftarrow \{(a, p_a, \mu_a)\}$
            $\tilde{\Lambda} \leftarrow \Lambda$
            $\tilde{d} \leftarrow d$
        end if
    end if
end while
```

S_{init} contains the initial load condition of APs and μ_i, $i \in [1 \ldots AP_{num}]$ and fixed AP list, F. The routine halts when AP$_i$ with maximum load that is not listed in F reaches to μ_i. Additionally, the routine terminates if one of the fixed APs' load value is increased. At each iteration in routine, we store the state of APs, $\tilde{S}$, maximum load value of non-fixed AP list, $\tilde{\Lambda}$, and maximum loaded AP index, $\tilde{d}$. After that, the algorithm reduces the beacon power level of congested AP, p_d, by one and user association is performed to store new values of $\tilde{\Lambda}$ and $\tilde{d}$. If the load level of any fixed AP is increased the routine terminates. If the congestion load of new state is less than the previous iteration, then we store new state to $\tilde{S}$, congestion load to $\tilde{\Lambda}$ and maximum load index to $\tilde{d}$.

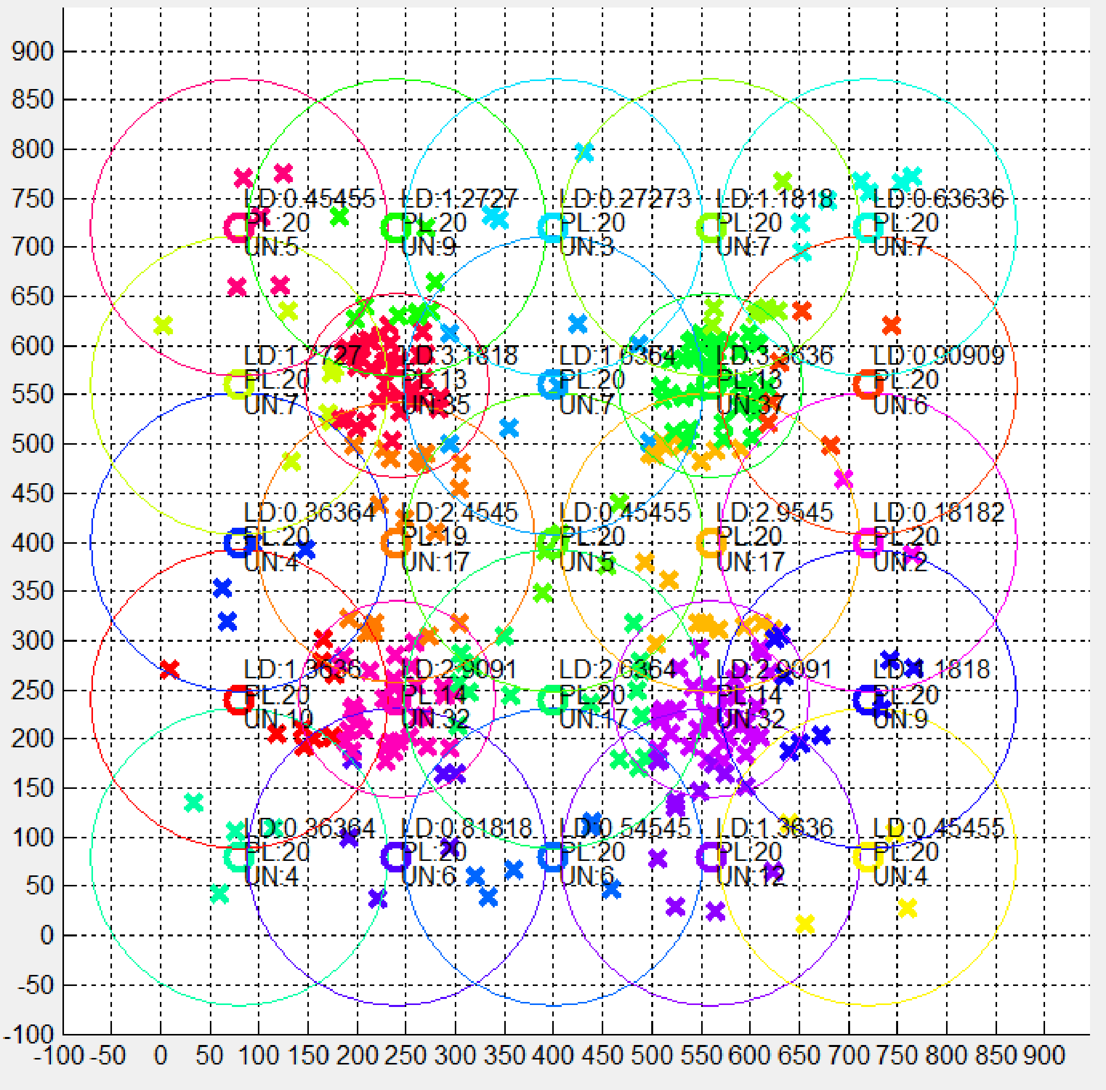


Figure 2.8. Single run of GF-MMPLB with 300 random users where 4x50 users are concentrated on four APs

In Figure 2.8, GF-MMPLB algorithm is executed one time with 300 users where 4x50 users are concentrated on four APs. In this figure the maximum loaded AP has the load level of approximately 3.18. The beacon power level of the AP is set to 13 and the minimum power level is set to 13 at the end of the algorithm. In next section, we see that the load of the congested AP is decreased in same simulation settings via GF-SMMPLB algorithm.

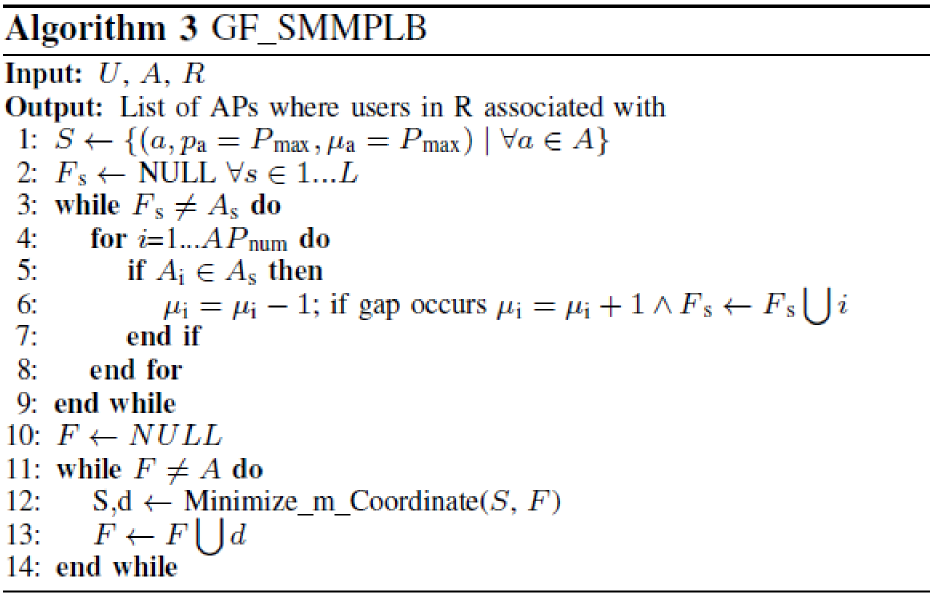
2.4.2. Gap-free statistical min-max priority load balancing algorithm

This algorithm uses the same routine that is *Minimize_m_Coordinate*. GF-MMPLB sets the minimum power level of APs to P_{max} and reduces them one by one until any further reduction result in a gap. On the other hand, GF-SMMPLB algorithm categorizes the APs into L levels considering estimated loads using the past user distribution statistics. In order to set the minimum power levels, *Set Statistical Minimum Power Levels* algorithm is used. The algorithm works as follows. Each AP is given a level from the set [1…L]. The APs having the highest level, *L*, are the most loaded ones and the APs having the lowest level are the least loaded among all APs. Starting from the level L, the μ_i values of APs is decreased until any further reduction violates the service availability in *R*.

An AP *a* is in level *s* if the following situation holds.

$$\tilde{\lambda}_{\mathrm{min}} + \mathrm{s}.\Delta \leq \tilde{\lambda}_{\mathrm{a}} \leq \tilde{\lambda}_{\mathrm{min}} + (\mathrm{s}+1).\Delta \qquad (2.1)$$

Where $\Delta = \frac{\tilde{\lambda}_{\mathrm{max}} - \tilde{\lambda}_{\mathrm{min}}}{\mathrm{L}}$, $\tilde{\lambda}_{\mathrm{a}}$ is the estimated load of AP a, $\tilde{\lambda}_{\mathrm{max}}$ is the estimated load of maximum loaded AP, and $\tilde{\lambda}_{\mathrm{min}}$ is the estimated load of minimum loaded AP. We set a level for each AP according to Equation 2.1. We define A_s such that, an AP $a \in A_s$ if *a* is in level *s*.



Algorithm 3 GF_SMMPLB

```
Input: $U$, $A$, $R$
Output: List of APs where users in R associated with
$S \leftarrow \{(a, p_a = P_{max}, \mu_a = P_{max}) \mid \forall a \in A\}$
$F_s \leftarrow$ NULL $\forall s \in 1...L$
while $F_s \neq A_s$ do
    for $i$=1...$AP_{num}$ do
        if $A_i \in A_s$ then
            $\mu_i = \mu_i - 1$; if gap occurs $\mu_i = \mu_i + 1 \wedge F_s \leftarrow F_s \bigcup i$
        end if
    end for
end while
$F \leftarrow NULL$
while $F \neq A$ do
    S,d $\leftarrow$ Minimize_m_Coordinate($S$, $F$)
    $F \leftarrow F \bigcup d$
end while
```

After determining the boundaries of each level, we decrease minimum power level of the APs of highest level L as much as possible. This lets the most loaded APs to decrease their beacon power levels more than the least loaded APs. Then the APs in level $L-1$ are admitted to decrease minimum power level if possible. We decrease the minimum power level of APs in subsequent levels consecutively. Finally, minimum power level of APs in level 1 is set. This procedure determines the minimum level of beacon coverage areas. After the assignment, GF-SMMPLB algorithm continues to execution as if the GF-MMPLB algorithm. Pseudo code is given in Algorithm 3.

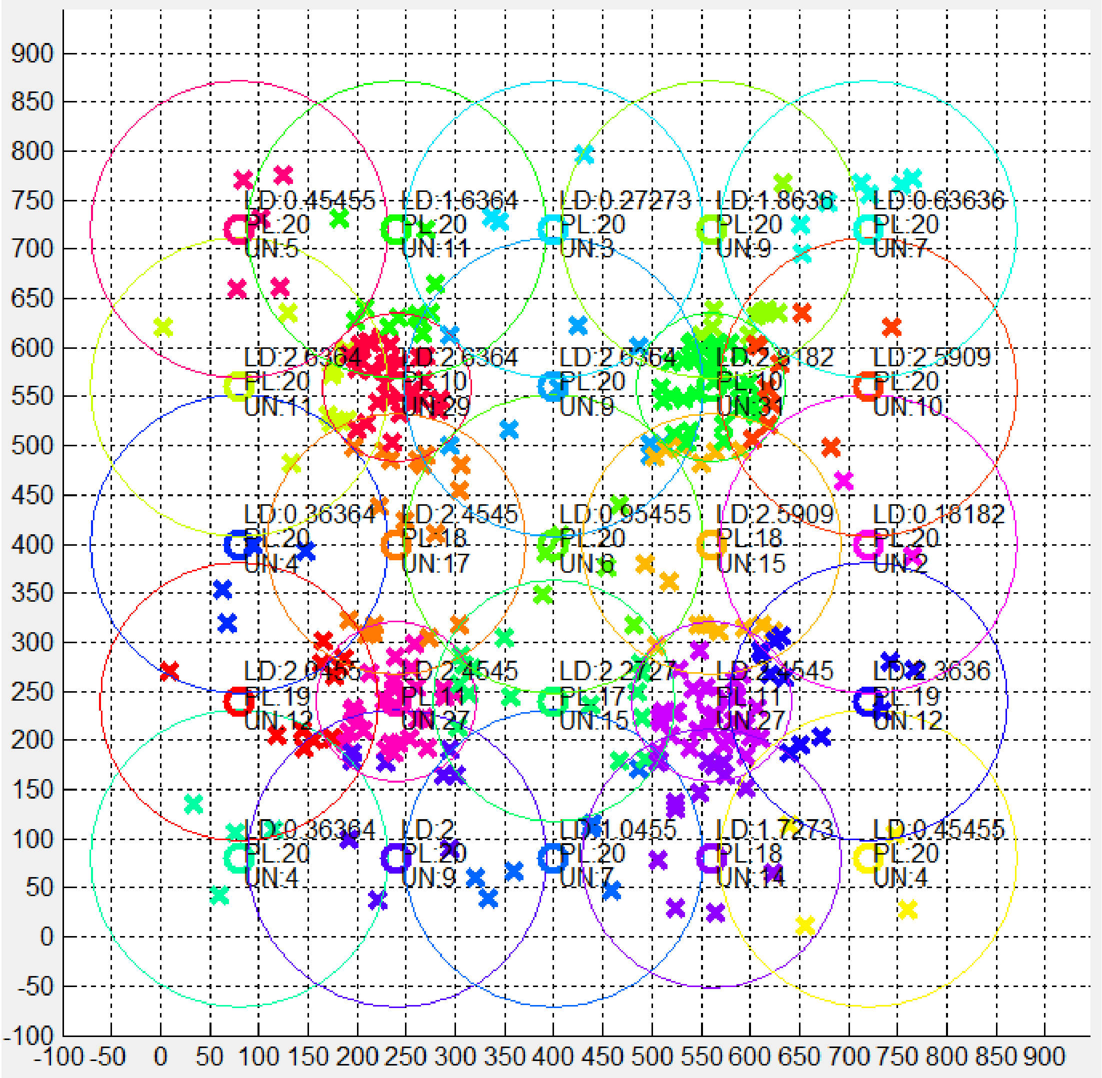


Figure 2.9. Single run of GF-SMMPLB with 300 random users where 4x50 users are concentrated on four APs

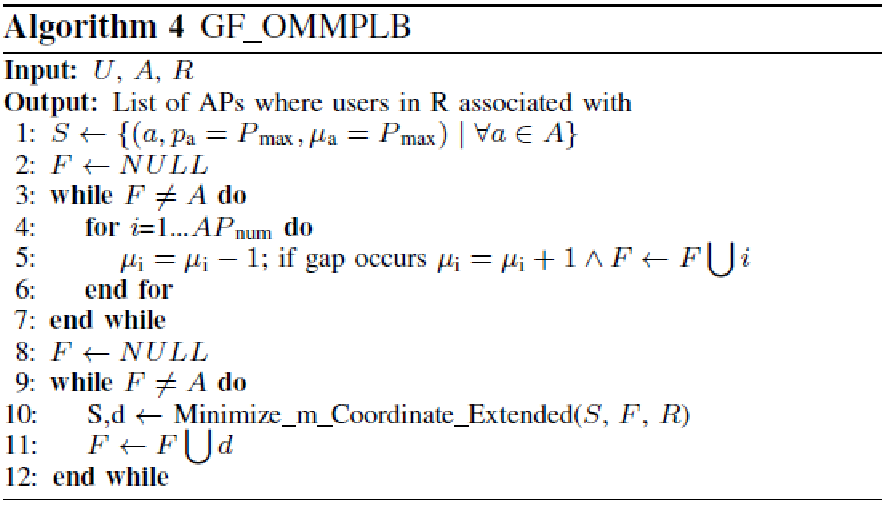
In Figure 2.9, GF-SMMPLB algorithm is executed one time with 300 users where 4x50 users are concentrated on four APs. In this figure the maximum loaded AP has the load level of approximately 2.63, which is better as compared to simulation run of GF-MMPLB algorithm depicted in Figure 2.8. The beacon power level of the AP is set to 10 and the minimum power level is set to 10 at the end of the algorithm. In next section, we see that the load of the congested AP is decreased in same simulation settings via GF-OMMPLB algorithm (Algorithm 4).

2.4.3. Gap-free online min-max priority load balancing algorithm

Our final algorithm concerns about min-max load balancing by adjusting minimum beacon power levels according to current load values. Minimum power levels, μ_i, of APs are updated according to congestion load index at each iteration. This approach gives the flexibility to congestion loaded APs to decrease the beacon power level, which increases the chance of congestion load reduction. This algorithm also provides service availability guarantee to users dynamically in a pre-determined region.

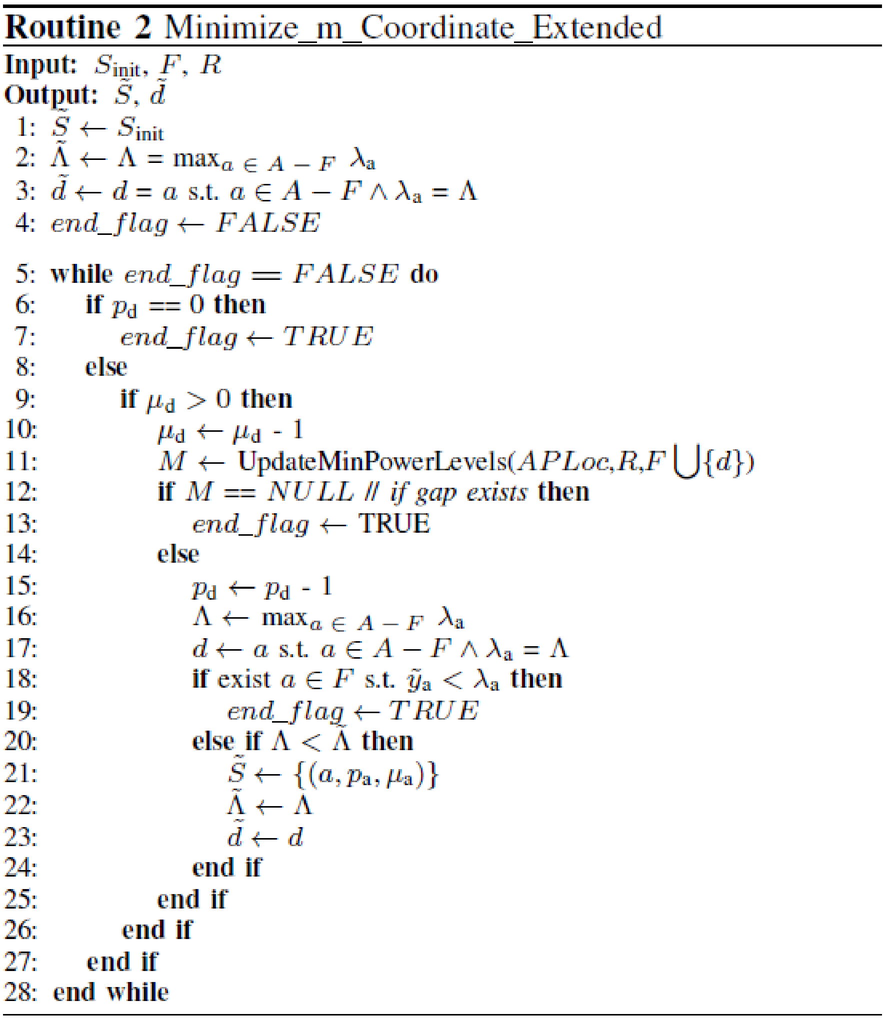
The algorithm uses an extended version of the *Minimize-m-Coordinate* routine where the input variables are S_{init}, F, *and* R. S_{init} contains the initial load condition of APs and minimum power levels. The algorithm halts when an AP with maximum load, which is not listed in F, reaches to minimum possible power level. In addition, when a gap occurs due to minimum power level reduction, the routine halts to store maximum loaded AP.

At each iteration in routine, we store the state of APs, $\tilde{S}$, maximum load value of non-fixed AP list, $\tilde{\Lambda}$, and maximum loaded AP index, $\tilde{d}$. After that, the algorithm reduces the minimum power level of congested AP, μ_d, to admit the beacon power level reduction. After μ_d power reduction, there may be gap between APs for signal strength at minimum power level. Hence, we update the minimum power levels of non-fixed APs where the new fixed AP set is $F \cup \{d\}$. Updating the μ_d of non-fixed APs may not cover the region R. If gap occurs in the region, then we terminate the routine. Otherwise, the beacon power level, p_d, is reduced by one and user association is performed to store new values of $\tilde{\Lambda}$ and $\tilde{d}$. If the load level of any fixed AP is increased, then the routine terminates. Otherwise, if the congestion load of new state is less than the previous iteration, then we store new state to $\tilde{S}$, congestion load to $\tilde{\Lambda}$ and maximum load index to $\tilde{d}$.



Algorithm 4 GF_OMMPLB

```
Input: $U$, $A$, $R$
Output: List of APs where users in R associated with
$S \leftarrow \{(a, p_a = P_{max}, \mu_a = P_{max}) \mid \forall a \in A\}$
$F \leftarrow NULL$
while $F \neq A$ do
    for $i$=1...$AP_{num}$ do
        $\mu_i = \mu_i - 1$; if gap occurs $\mu_i = \mu_i + 1 \wedge F \leftarrow F \bigcup i$
    end for
end while
$F \leftarrow NULL$
while $F \neq A$ do
    S,d $\leftarrow$ Minimize_m_Coordinate_Extended($S$, $F$, $R$)
    $F \leftarrow F \bigcup d$
end while
```



Routine 2 Minimize_m_Coordinate_Extended

```
Input: $S_{init}$, $F$, $R$
Output: $\tilde{S}$, $\tilde{d}$
$\tilde{S} \leftarrow S_{init}$
$\tilde{\Lambda} \leftarrow \Lambda = \max_{a \in A - F} \lambda_a$
$\tilde{d} \leftarrow d = a$ s.t. $a \in A - F \wedge \lambda_a = \Lambda$
$end\_flag \leftarrow FALSE$

while $end\_flag == FALSE$ do
    if $p_d == 0$ then
        $end\_flag \leftarrow TRUE$
    else
        if $\mu_d > 0$ then
            $\mu_d \leftarrow \mu_d - 1$
            $M \leftarrow$ UpdateMinPowerLevels($APLoc$,$R$,$F \bigcup \{d\}$)
            if $M == NULL$ // if gap exists then
                $end\_flag \leftarrow$ TRUE
            else
                $p_d \leftarrow p_d - 1$
                $\Lambda \leftarrow \max_{a \in A - F} \lambda_a$
                $d \leftarrow a$ s.t. $a \in A - F \wedge \lambda_a = \Lambda$
                if exist $a \in F$ s.t. $\tilde{y}_a < \lambda_a$ then
                    $end\_flag \leftarrow TRUE$
                else if $\Lambda < \tilde{\Lambda}$ then
                    $\tilde{S} \leftarrow \{(a, p_a, \mu_a)\}$
                    $\tilde{\Lambda} \leftarrow \Lambda$
                    $\tilde{d} \leftarrow d$
                end if
            end if
        end if
    end if
end while
```

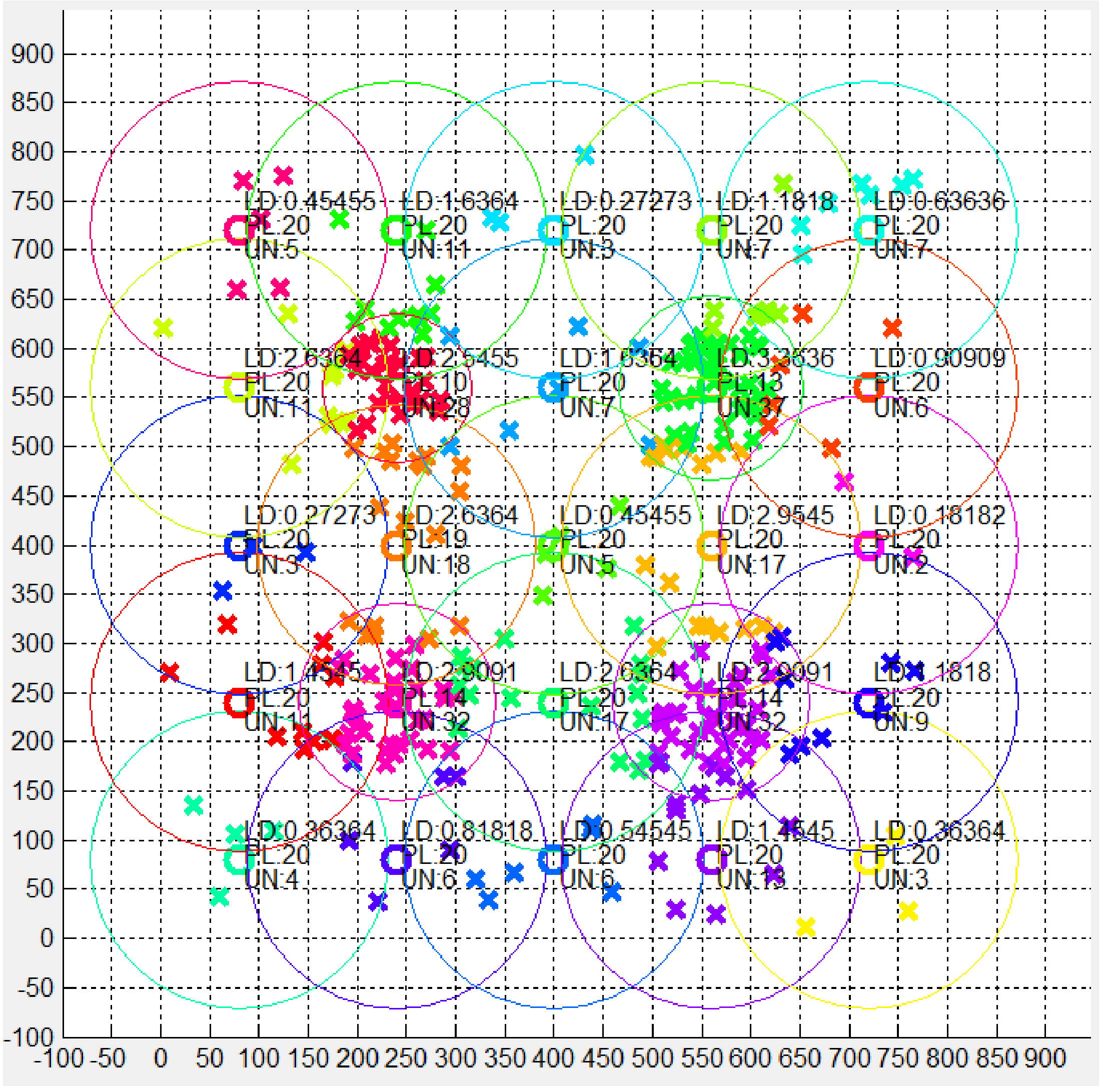


Figure 2.10. Single run of GF-OMMPLB with 300 random users where 4x50 users are concentrated on four APs

In Figure 2.10, GF-OMMPLB algorithm is executed one time with 300 users where 4x50 users are concentrated on four APs. In this figure the maximum loaded AP has the load level of approximately 2.54, which is better as compared to simulation run of GF-SMMPLB algorithm depicted in Figure 2.9. The beacon power level of the AP is set to 10 and the minimum power level is set to 10 at the end of the algorithm. We have shown that the load of the congested AP is decreased better via GF-OMMPLB algorithm.

3. RESULTS AND DISCUSSION

3.1. Simulation Results

We performed extensive set of simulations to provide comparison of the algorithms stated above in terms of data load of APs and bandwidth rate of users. There are two additional association algorithms depicted in the simulation graphs that are Strongest Signal First (SSF) and LLF. The LLF association technique enables users to be connected with the AP that has least load among all reachable APs. The SSF algorithm is an association technique in which users are connected with the AP that has the strongest signal transmission as stated in IEEE 802.11b Standard.

In our simulation, a WLAN with 25 APs is deployed in an 800x800 m^2 region R where the APs positions are equidistant to neighbors. The distance between APs is set to 160 meters and AP bandwidth is set to 10 Mbps. Furthermore, power levels are assumed to be circular shape and intersecting regions may exist between neighboring APs. In region R, it is assumed that APs provide full network coverage by transmitting with highest power level. User-AP association is determined by the Signal-to-Noise ratio (SNR). SNR determines the load that a user contributes on the associated AP. We assume discrete level of load determination from SNR. We used the SNR-to-Bitrate calculation stated in [18] to determine the bit rate between user and AP. The relationship between bit rate and SNR is indicated as follows. For SNR ≥ 1dBW → 1 Mbps, SNR ≥ 3dBW → 2 Mbps, SNR ≥ 5dBW → 5.5 Mbps, SNR ≥ 9dBW → 11 Mbps is used. Additionally, P_{max} and P_{min} are given as 20 dBm and 10 dBm respectively. By defining d as the distance between user and AP, the path loss formulation is given as follows, $P_{loss} = 40 - 10 \times 3.3 \times log_{10} d$ [25]. This formulation is used with background noise -93 dBm, to calculate the coverage range of a cell [25]. Using this model, the minimal coverage range is found as 75 m and maximal coverage range is found as 150 m as shown in the following steps.

$$SNR[dBW] = 10 \times log_{10}\left(\frac{P_{rx}}{N}\right) = 10 \times log_{10} P_{rx} - 10 \times log_{10} P_{rx} N \tag{3.1}$$

$$SNR[dBW] = 10 \times log_{10} P_{rx} - 10 \times log_{10} P_{rx} N \tag{3.2}$$

$$SNR[dBW] = P_{rx}[dBW] - N[dBW] \tag{3.3}$$

$$P_{rx}[dBW] = SNR[dBW] + N[dBW] \tag{3.4}$$

$$P_{rx}[dBW] = 1.19\ dBW + (-93 - 30)dBW = \ -121.81\ dBW \tag{3.5}$$

$$P_{rx}[dBW] = P_{tx}[dBW] - P_{loss} = P_{tx}[dBW] - (40 - 10 \times 3.3 \times log_{10}d) \tag{3.6}$$

$For\ P_{tx}[dBm] = \ 20\ dBm;$

$$-121.81\ dBW = (20 - 30) - 40 + 10 \times 3.3 \times log_{10}d \tag{3.7}$$

$$-71.81\ dBW = -33 \times log_{10}d \tag{3.8}$$

$$d = 10^{71.81/33} \cong \ 150\ m \tag{3.9}$$

In order to determine the maximum coverage distance of a cell, we used the calculations above. SNR is by subtracting noise power, N, from received signal power, P_{rx}, in step 3.3. The background noise power level is set to -93 dBW. 1 Watt is equal to 1000mW = 30dBm. Hence, we used the conversion from dBW to dBm as, $1\ dBW = 1\ dBm - 30$. Minimum SNR value is set to 1.19 dBW that is required for successful signal transmission. If the signal to noise ratio increases the coverage area shrinks. In order to simulate an indoor environment, the path loss exponent is chosen as 3.3 [25]. Hence, the path loss calculation is performed as $P_{loss} = 40 - 10 \times 3.3 \times log_{10}d$ in step 3.6. The variable d represents the distance between a user and an AP. When the maximal transmission power, P_{tx}, of beacon signals is set to 20 dBm, the coverage diameter of a cell is 150 m approximately in step 3.9.

$For\ P_{tx}[dBm] = \ 10\ dBm;$

$$-121.81\ dBW = (10 - 30) - 40 + 10 \times 3.3 \times log_{10}d \tag{3.10}$$

$$-61.81\ dBW = -33 \times log_{10}d \tag{3.11}$$

$$d = 10^{61.81/33} \cong \ 75\ m \tag{3.12}$$

When the maximal transmission power of beacon signals is set to 10 dBm, the coverage diameter of a cell is 75 m approximately in step 3.12.

AP load is defined as an ordered pair of the aggregated load contributions of its associated users and a unique AP priority. A user load on an AP can be calculated using the bitrate information. The load of an AP_i can be defined as follows, $\boldsymbol{\lambda_i} = \sum_{u \in U_i} \frac{1}{r_{u,i}}$, where $r_{u,i}$ is the bitrate with which the user u communicates with AP_i and U_i is the set of users that are associated with AP_i.

The network setup can cause coverage holes between APs, however the algorithms prevent occurrence of this problem. While simulating GF-SMMPLB algorithm, we assumed that distribution of the users are time-invariant, hence loads can be perfectly estimated. Moreover, we choose L=3.

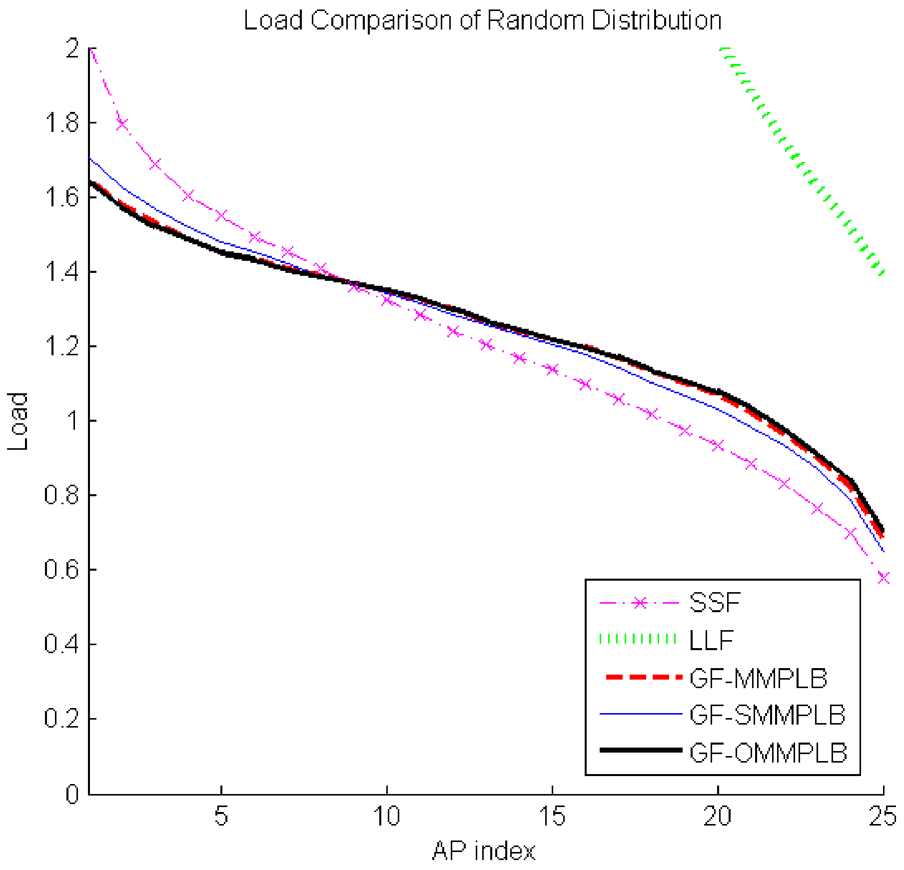


Figure 3.1. Load comparison of algorithms with 300 randomly distributed users

Figure 3.1 depicts the load of APs in a network environment with 300 randomly distributed users in the square region *R*. The x-axis represents the AP's index and the y-axis represents average load value of an AP. At each simulation run, the APs are sorted by the load values in decreasing order. After 50 simulation runs, the λ_i of each AP n is computed by finding the average value of the nth highest loaded AP. The load values at each AP index are joined to make the graph sightly. In Figure 3.1, Gap-free Online Min-Max Priority Load Balancing (GF-OMMPLB) algorithm is shown with a thick solid line, GF-SMMPLB is plotted with a thin solid line and GF-MMPLB algorithm is represented with a thick dashed line that is the min-max algorithm stated in [18].

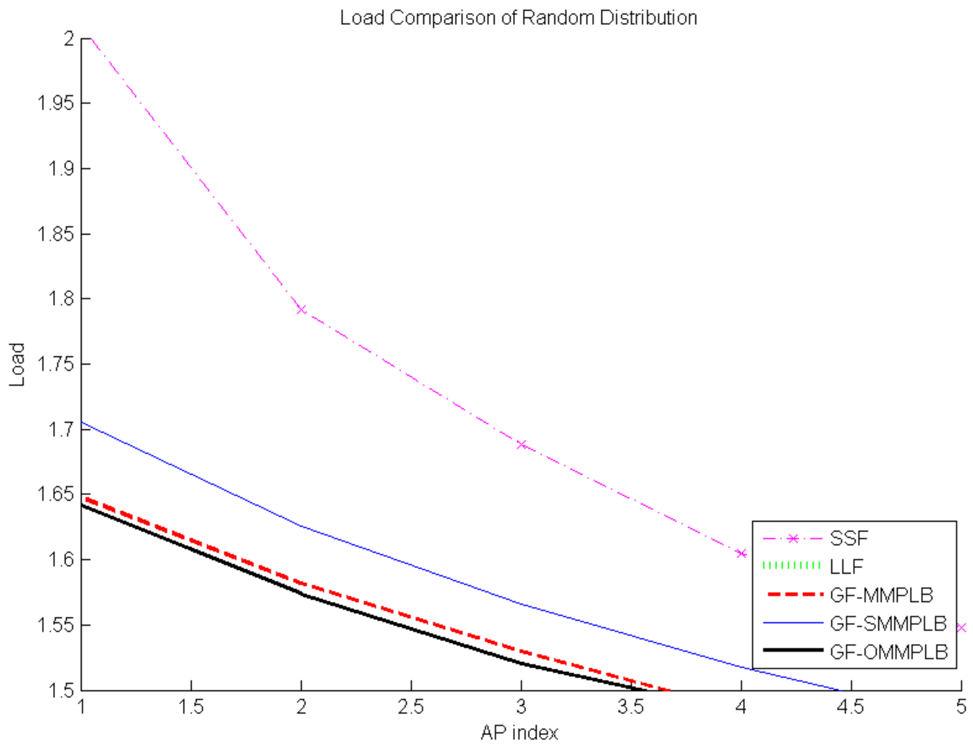


Figure 3.2. Load comparison of algorithms with 300 randomly distributed users with focused view

The Figure 3.2 illustrates the focused view of the previous figure to provide more information about congestion relieving of the maximum loaded APs. The LLF algorithm, which is not visible in the focused area, performs the worst with approximate load values of 15.5 for the first AP index and 13.5 for the second AP index. Although GF-MMPLB algorithm generates a load vector with fairly lower lexicographical order, the GF-OMMPLB algorithm decreases the maximum load value of APs better than other algorithms. All proposed algorithms clearly outperform the SSF and LLF algorithms. Difference between three proposed algorithms is not so evident in this scenario. The reason is that users are uniformly distributed to the region, while GF-SMMPLB and GF-OMMPLB are designed for networks with hot-spot areas.

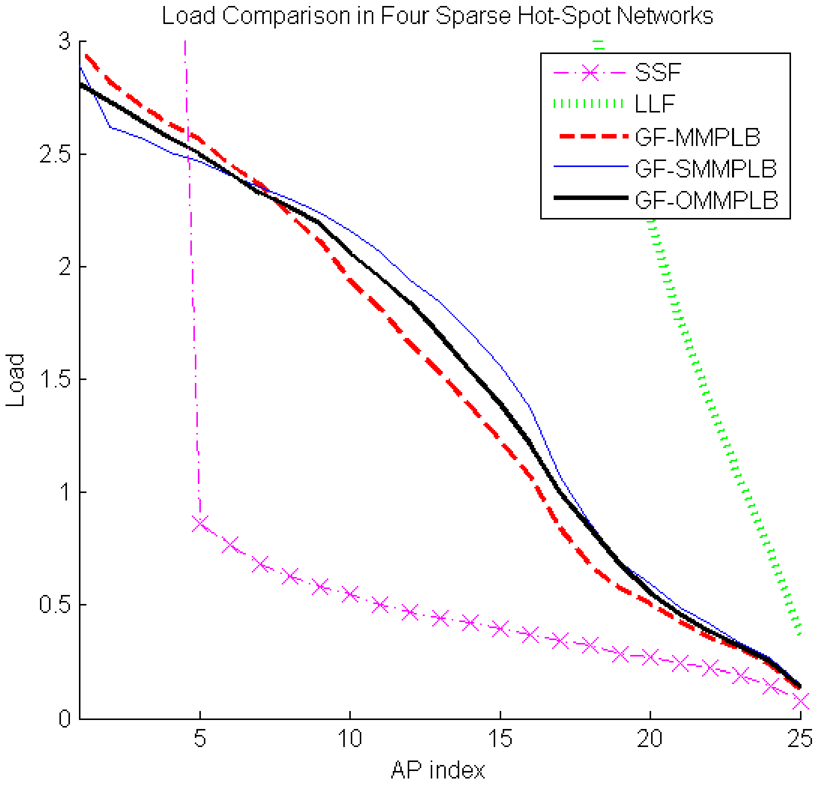


Figure 3.3. Load comparison of algorithms with 300 random users where 4x50 users are concentrated on four APs

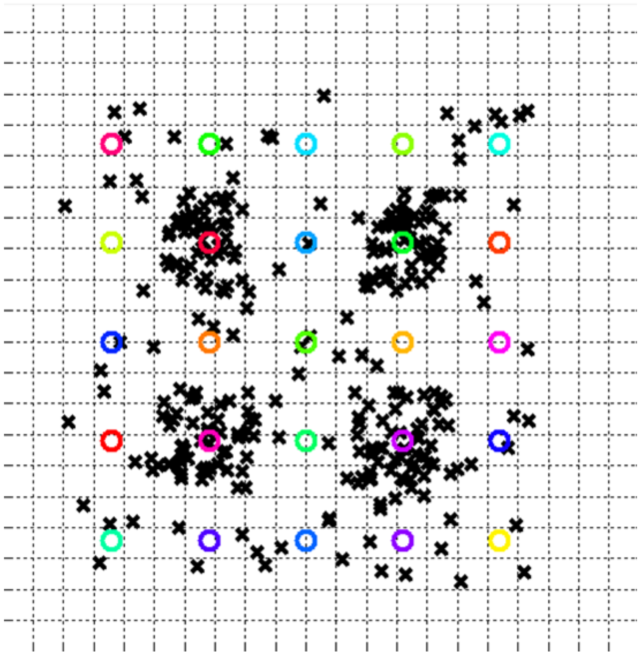


Figure 3.4. Distribution of 300 random users where 4x50 users are concentrated on four APs

Figure 3.3 presents a network environment with unbalanced distribution of 300 users. There are four hot-spot regions apart with 50 users. Last 100 users are distributed randomly in *R*. Each hotspot that is square shaped of 160x160 m^2 consists of random user distribution as shown on Figure 3.4. The crosses represent the users and circles represent the APs.

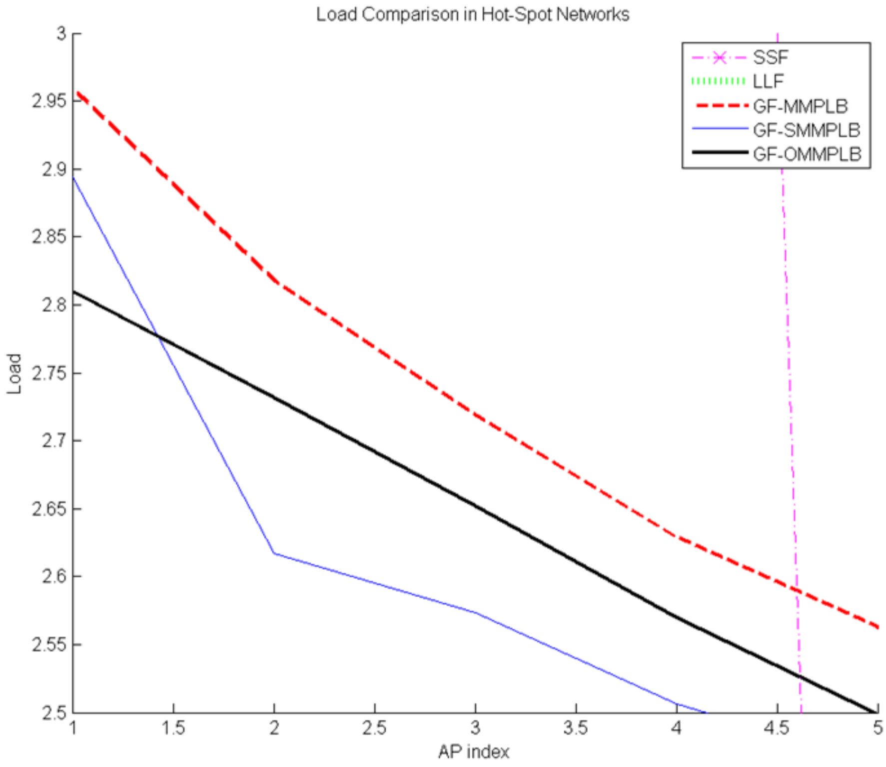


Figure 3.5. Load comparison of algorithms with 300 random users where 4x50 users are concentrated on four APs with focused view

The Figure 3.5 illustrates the focused view of the Figure 3.3 to provide more information about congestion relieving of the maximum loaded APs. The LLF algorithm, which is not visible in the focused area, performs the worst with approximate load values of 25.5 for the first AP index and 19 for the second AP index. The SSF algorithm creates the approximate load values of 5.7, 5.5, 5.3, and 5.1 for the first four indices. The hotspots creates leveling for APs such that statistical algorithm uses the level information to determine the μ_i value of APs. Hence, for this setting, the GF-SMMPLB performs better than GF-MMPLB algorithm. However, GF-OMMPLB algorithm decreases the maximum load value better at highest AP index. Also, the SSF algorithm performs ineffective for the first four APs, because the hotspots create immense load at the four APs.

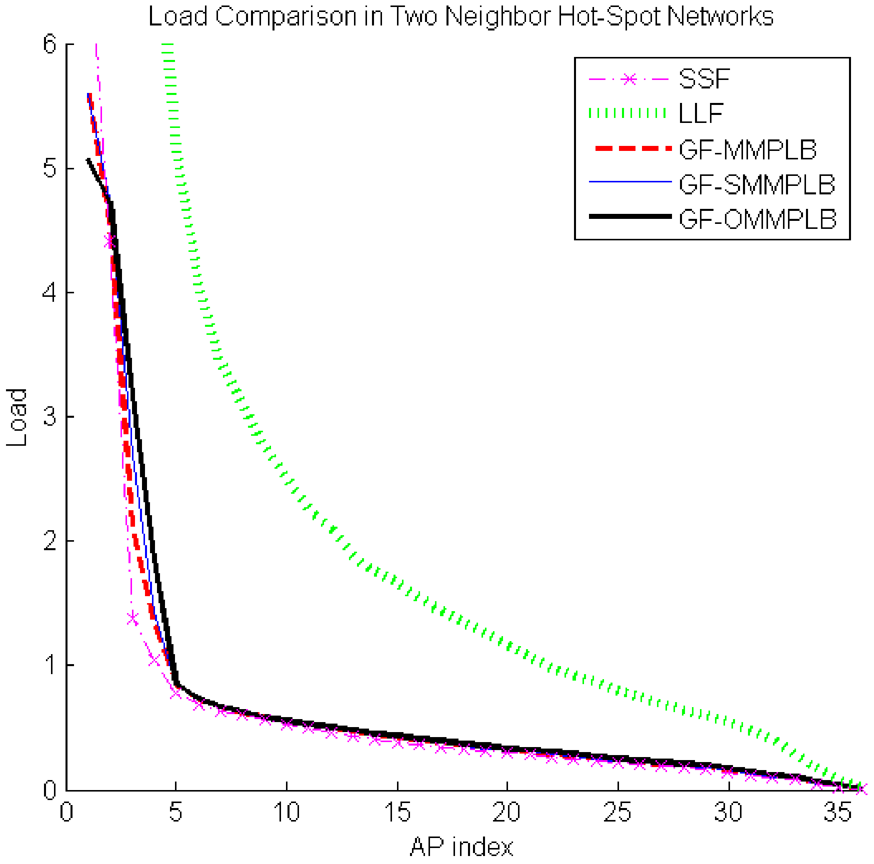


Figure 3.6. Load comparison of algorithms with 200 random users where 2x50 users are concentrated on two neighbor APs

Figure 3.6 shows a network environment with random distribution of 200 users. There are two neighbor hot-spot regions. Each hotspot consists of 50 random user distributions. Last 100 users are distributed randomly in *R*.

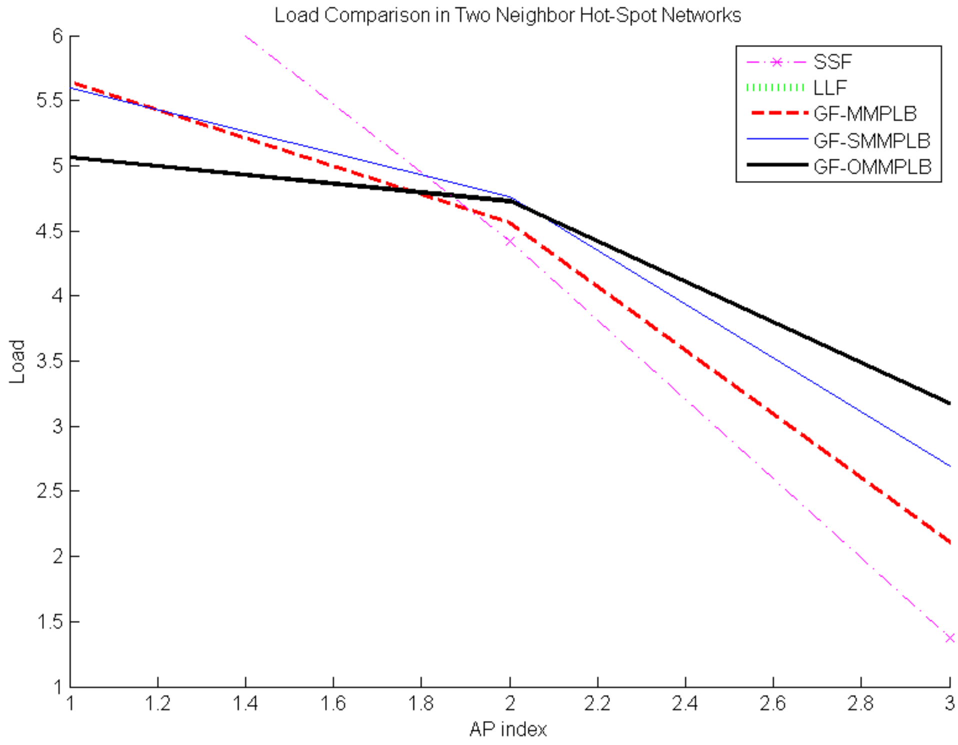


Figure 3.7. Load comparison of algorithms with 200 random users where 2x50 users are concentrated on two neighbor APs with focused view

The Figure 3.7 illustrates the focused view of the Figure 3.6 to provide more information about congestion relieving of the maximum loaded APs. The LLF algorithm, which is not visible in the focused area, performs the worst with approximate load values of 21.5 for the first AP index and 16 for the second AP index. The SSF algorithm provides approximate load value of 17 for the first index. GF-OMMPLB algorithm decreases the maximum load value better at first and second AP index. The load values of first two APs of LLF algorithm are 21.5 and 15.9 respectively. The LLF algorithm performs ineffective for the first two APs, because the hotspots create immense load at that APs.

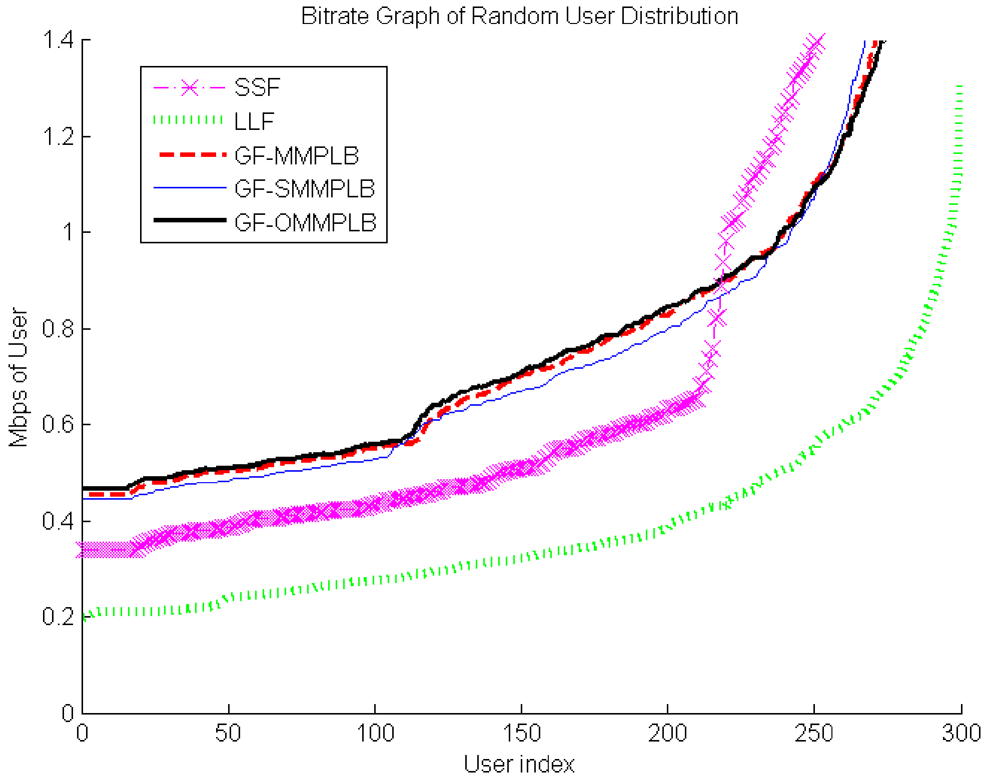


Figure 3.8. Per-user bandwidth rate for 300 random user distribution

Figure 3.8 illustrates the bandwidth distribution among the 300 random users as in Figure 3.1. The SSF reaches to maximum rate of 5.5 Mbps, the GF-SMMPLB reaches to maximum rate of 4.8 Mbps, the GF-MMPLB reaches to maximum rate of 4.5 Mbps, the GF-OMMPLB reaches to maximum rate of 4.2 Mbps, and the LLF reaches to maximum rate of 1.3 Mbps. In this setting, GF-OMMPLB algorithm performs the best in terms of bandwidth allocation and fairness, because the online solution performs the load balancing among the APs more effectively. GF-MMPLB performs better than the GF-SMMPLB algorithm because of random user distribution. Since users are not creating a hot-spot region the GF-SMMPLB algorithm cannot provide a better bandwidth to overall users. The focused view for the first 50 users is presented in Figure 3.9.

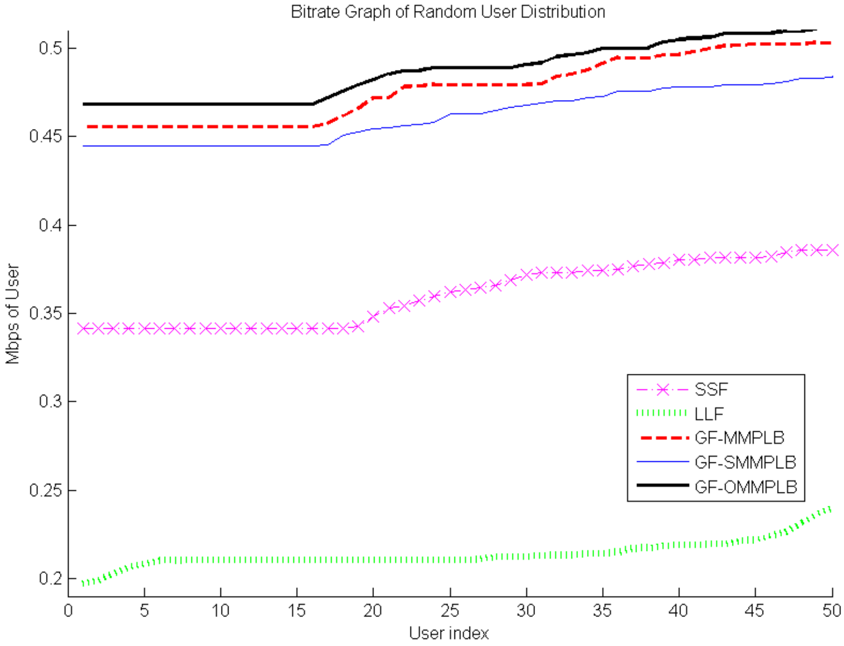


Figure 3.9. Per-user bandwidth rate for 300 random user distribution with focused view

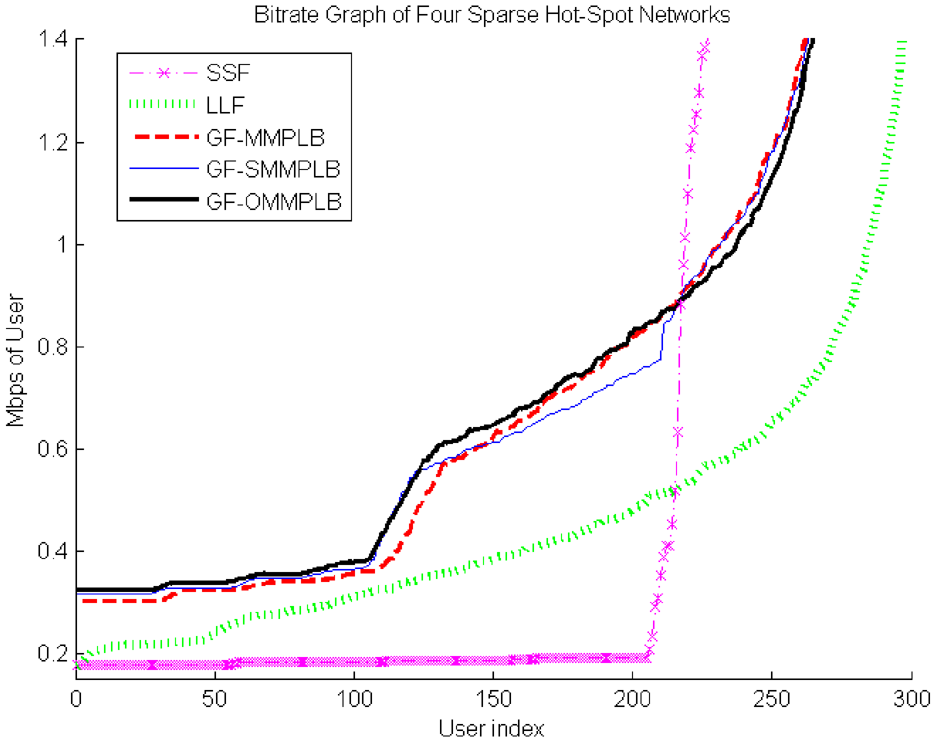


Figure 3.10. Per-user bandwidth rate for 300 random users in four sparse hot-spot scenario

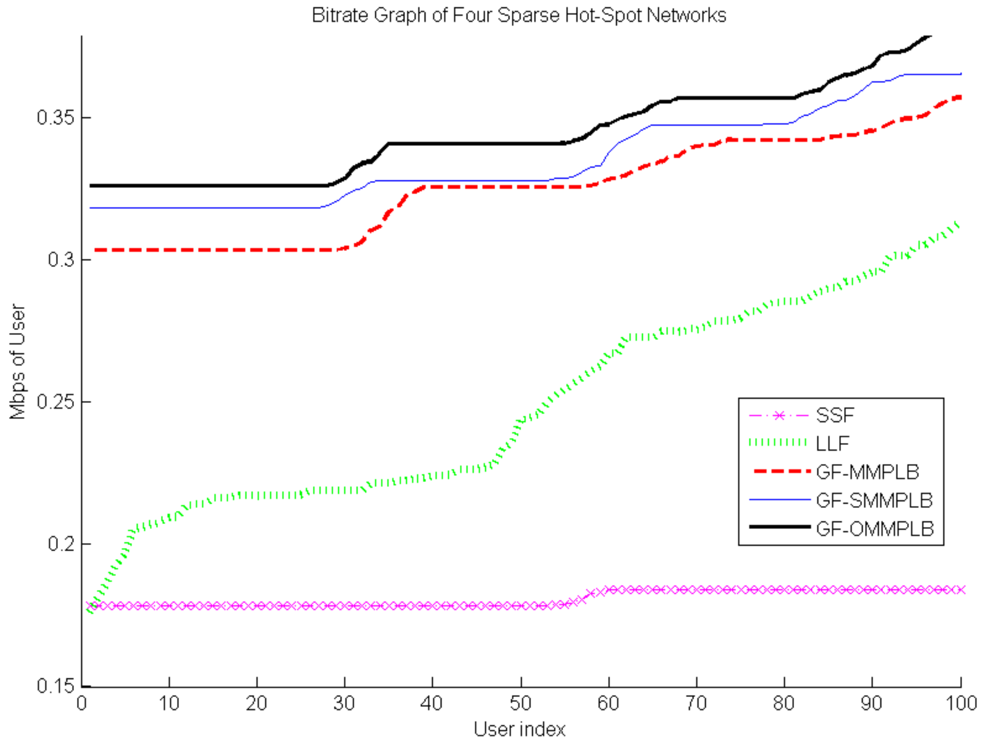


Figure 3.11. Per-user bandwidth rate for 300 random users in four sparse hot-spot scenario with focused view

Bandwidth distribution among the users is depicted in Figure 3.10, where the simulation settings are set to same configuration with the experiment in Figure 3.3. The SSF reaches to maximum rate of 8.9 Mbps, the GF-MMPLB reaches to maximum rate of 7.1 Mbps, the GF-OMMPLB reaches to maximum rate of 6.9 Mbps, the GF-SMMPLB reaches to maximum rate of 6.6 Mbps, and the LLF reaches to maximum rate of 1.9 Mbps. It is observed that the bandwidth allocation is performed best with the GF-OMMPLB algorithm in terms of fairness, because the load balancing among the APs is done more effectively. GF-SMMPLB also increases per user bandwidth in hot-spot areas and it is fairer than GF-MMPLB. The focused view for the first 100 users is presented in Figure 3.11.

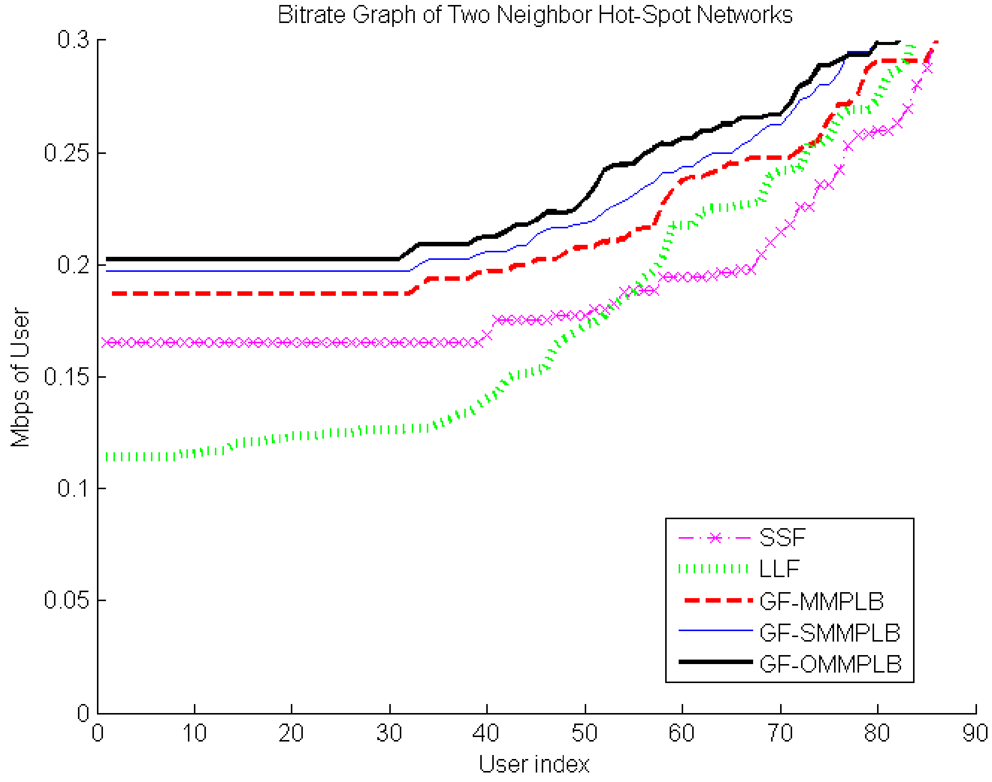


Figure 3.12. Per-user bandwidth rate for 200 random users where 2x50 users are concentrated on two neighbor APs

Figure 3.12 illustrates the bandwidth distribution among the 200 random users where 100 users are concentrated on two neighbor APs randomly as in Figure 3.6. In this setting, GF-OMMPLB algorithm performs the best in terms of bandwidth allocation and fairness, because the online solution performs the load balancing among the APs more effectively. GF-SMMPLB also enhances per user bandwidth in hot-spot areas and it is fairer than GF-MMPLB. SSF algorithm performs better than the LLF. When a user associates with least loaded available AP, that user most probably connected to the internet from a long distance resulting in heavy load condition. Hence, the throughput of the system and users decreases.

4. CONCLUSIONS

Cell-Breathing technique is a promising approach for load balancing in the IEEE 802.11 WLAN networks, since it does not require any modifications in the user equipment. Nevertheless, it can suffer from AP service availability problem because coverage reduction can create some unserviced areas. In this thesis, a min-max priority load balancing algorithm is provided that sets a minimum level for the beacon power of each AP such that no gap occurs. Additionally, we proposed two novel enhancements to this algorithm. Firstly, we described an algorithm that utilizes the statistical past user distribution information while deciding on the minimum beacon power levels. Secondly, we provided an online algorithm that adjusts the minimum beacon power levels according to current load values. Our simulations showed the effectiveness and fairness of our load balancing algorithms.

4.1. Future Work

The effectiveness and fairness of cell breathing technique is demonstrated in this work and the technique significantly outperforms popular fixed power adjustment techniques. In the simulation we assumed a fixed network traffic for each iteration. As future work, we will assume a dynamic network traffic for the simulation environment. We are planning to extend our work for the case of mobile users. Secondly, we will consider QoS of connectivity for different network configurations. We used the IEEE 802.11b standard as wireless standard. In future study, we can use other IEEE 802.11 specifications such as 802.11a/g/n/ac or ad. Additionally we can adapt the these studies to Long-Term Evolution (LTE) or heterogeneous networks.

REFERENCES

[1] Ferro, E., Potorti, F. (2005). Bluetooth and Wi-Fi wireless protocols: a survey and a comparison. Wireless Communications, IEEE, 12(1), 12-26.

[2] Ieee standard for information technology (2009). Local and metropolitan area networks- specific requirements- part 11: Wireless LAN medium access control (mac) and physical layer (phy) specifications amendment 5: Enhancements for higher throughput. IEEE Std 802.11n-2009, pages 1 -565.

[3] Wi-Fi Alliance. (1999). Wi-Fi Alliance, http://www.wi-fi.org/ (24 May 2014)

[4] Lehembre, G. (2005). Wi-Fi security–WEP, WPA and WPA2. Hackin9 (January 2006).

[5] Gast, M. (2002). Wireless LAN security: A short history. Retrieved July 25, 2005.

[6] Bahl, P., Hajiaghayi, M. T., Jain, K., Mirrokni, S. V., Qiu, L., Saberi, A. (2007). Cell breathing in wireless LANs: Algorithms and evaluation. Mobile Computing, IEEE Transactions on, 6(2), 164-178.

[7] Verma, L., Fakharzadeh, M., Choi, S. (2013). Wi-Fi on steroids: 802.11ac and 802.11ad. Wireless Communications, IEEE, 20(6), 30-35.

[8] Yen, L. H., Yeh, T. T., Chi, K. H. (2009). Load balancing in IEEE 802.11 networks. Internet Computing, IEEE, 13(1), 56-64.

[9] S., Papagiannaki, K., Diot, C., Kurose, J., Towsley, D. (2005, October). Facilitating access point selection in IEEE 802.11 wireless networks. In Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement (pp. 293–298). USENIX Association.

[10] Yen, L. H., Yeh, T. T. (2006, May). SNMP-based approach to load distribution in IEEE 802.11 networks. In Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd (Vol. 3, pp. 1196-1200). IEEE.

[11] Nicholson, A. J., Chawathe, Y., Chen, M. Y., Noble, B. D., Wetherall, D. (2006, June). Improved access point selection. In Proceedings of the 4th international conference on Mobile systems, applications and services (pp. 233-245). ACM.

[12] Velayos, H., Aleo, V., Karlsson, G. (2004, June). Load balancing in overlapping wireless LAN cells. In Communications, 2004 IEEE International Conference on (Vol. 7, pp. 3833-3836). IEEE.

[13] Aleo, V. (2003). Load distribution in IEEE 802.11 cells. Master of Science Thesis, KTH School of Electrical Eng., Royal Institute of Technology, Stockholm, Sweden.

[14] Brickley, O., Rea, S., Pesch, D. (2005, September). Load balancing for QoS enhancement in IEEE 802.11e WLANs using cell breathing techniques. In 7th IFIP International Conference on Mobile and Wireless Communications Networks, Maroc.

[15] Wang, Y., Cuthbert, L. G., Bigham, J. (2004, March). Intelligent radio resource management for IEEE 802.11 WLAN. In Wireless Communications and Networking Conference (Vol. 3, pp. 1365-1370). IEEE.

[16] Balachandran, A., Bahl, P., Voelker, G. M. (2002). Hot-spot congestion relief in public-area wireless networks. In Mobile Computing Systems and Applications, 2002. Proceedings Fourth IEEE Workshop on (pp. 70-80). Callicoon, NY. IEEE.

[17] Iyer, P. J. et al. (2005). System and method for centralized station management. U.S. Patent No. 0,213,579. Washington, DC: U.S. Patent and Trademark Office.

[18] Bejerano, Y., Han, S. J. (2009). Cell breathing techniques for load balancing in wireless LANs. Mobile Computing, IEEE Transactions on, 8(6), 735-749.

[19] Schwab, D., Bunt, R. (2004, March). Characterising the use of a campus wireless network. In INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies (Vol. 2, pp. 862-870). IEEE.

[20] Gong, D., Yang, Y. (2012, March). Ap association in 802.11 n WLANs with heterogeneous clients. In INFOCOM, 2012 Proceedings IEEE (pp. 1440-1448). IEEE.

[21] Kauffmann, B., Baccelli, F., Chaintreau, A., Papagiannaki, K., Diot, C. (2005). Self organization of interfering 802.11 wireless access networks. INRIA, Rapport de recherche RR-5649.

[22] Kumar, A., Kumar, V. (2005, January). Optimal association of stations and APs in an IEEE 802.11 WLAN. In Proceedings of the National Conference on Communications (pp. 1-5).

[23] Wang, S., Huang, J., Cheng, X., Chen, B. (2014). Coverage adjustment for load balancing with an AP service availability guarantee in WLANs. Wireless Networks, 20(3), 475–491.

[24] Calhoun, P., Montemurro, M., Stanley, D. (2009). Control and provisioning of wireless access points (CAPWAP) protocol binding for IEEE 802.11. IETF RFC5416.

[25] Rappaport, T. S. (2001). Wireless communications: principles and practice (Vol. 2). New Jersey, USA: Prentice Hall PTR.

AUTOBIOGRAPHY

Name Surname: İlhan DEMİRCİ

Place and Date of Birth: Düzce / May 22, 1988

Foreign Language: English

E-Mail : demirci.ilhan@gmail.com

Educational Status

Grade	Department/Programme	University/High School	Graduation Year
High School	Physical Sciences	Gebze Anadolu High School	2006
University	Computer Science&Engineering	Sabancı University	2011

Work Experience

Year	Company/Corporation	Position
2011-2012	Tubitak-BTE, Gebze/KOCAELİ	Part-time worker
2012-…	Tubitak-BTE, Gebze/KOCAELİ	Researcher

Academic Publications

1. İ. Demirci and Ö. Korçak, “Gap-free load balancing in wireless LAN networks using cell breathing technique,” in *The Eleventh International Symposium on Wireless Communication Systems 2014: Track 2: Networking, protocols, cognitive radio, wireless sensor networks, services and applications* (ISWCS’2014 - Track 2-), Barcelona, Spain, Aug. 2014. [In submission]