Title	A Study on Prediction-based Autonomous Wi-Fi Channel Management for Interference Mitigation in Urban Environment
Author(s)	梶田, 宗吾
Citation	大阪大学, 2019, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/72590
rights	
Note	

A Study on Prediction-based Autonomous Wi-Fi Channel Management for Interference Mitigation in Urban Environment

Submitted to
Graduate School of Information Science and Technology
Osaka University

January 2019

Shugo KAJITA

List of Publications

Related Journal Articles

(1) Shugo Kajita, Hirozumi Yamaguchi, Teruo Higashino, Shigeki Umehara, Fumiya Saitou, Hirofumi Urayama, Masaya Yamada, Taka Maeno, Shigeru Kaneda, Mineo Takai, "An Algorithm for Estimating the Performance of WLAN Systems with Severe Interference" *IPSJ Journal*, vol. 57, no. 2, pp. 745–755, February 2016. (in Japanese)

(2) Shugo Kajita, Tatsuya Amano, Hirozumi Yamaguchi, Teruo Higashino, Mineo Takai, "Prediction-based Channel Selection Algorithm for Autonomous Wi-Fi Access Point Management" *IPSJ Journal*. (in Japanese, conditional accepted)

Related Conference Papers

(1) Shugo Kajita, Tatsuya Amano, Hirozumi Yamaguchi, Teruo Higashino, Mineo Takai, "Wi-Fi Channel Selection Based on Urban Interference Measurement" in *Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous 2016)*, pp. 143-150, November-December 2016.

(2) Shugo Kajita, Hirozumi Yamaguchi, Teruo Higashino, Hirofumi Urayama, Masaya Yamada, Mineo Takai, "Throughput and Delay Estimator for 2.4GHz WiFi APs: A Machine Learning-Based Approach" in *Proceedings of the 8th IFIP Wireless and Mobile Networking Conference (WMNC 2015)*, pp. 223-226, October 2015.

(3) Shugo Kajita, Hirozumi Yamaguchi, Teruo Higashino, Shigeki Umehara, Fumiya Saitou, Hirofumi Urayama, Masaya Yamada, Taka Maeno, Shigeru Kaneda, Mineo Takai, "A Channel Selection Strategy for WLAN in Urban Areas by Regression Analysis" in *Proceedings of the 10th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2014)*, pp. 646-651, October 2014.

(4) Tatsuya Amano, Shugo Kajita, Hirozumi Yamaguchi, Teruo Higashino, Mineo Takai, "A Study on Wi-Fi RSSI Map Construction Based on Crowdsourcing and Simulations" *IPSJ Journal*, vol. 59, no. 2, pp. 450–461, February 2018. (in Japanese)

(5) Tatsuya Amano, Shugo Kajita, Hirozumi Yamaguchi, Teruo Higashino, Mineo Takai, "A Crowdsourcing and Simulation based Approach for Fast and Accurate Wi-Fi Radio Map Construction in Urban Environment" in *Proceedings of the IFIP Networking 2017 (Networking 2017)*, pp.1-9, June 2017.

Other Conference Papers

(1) Tatsuya Amano, Shugo Kajita, Hirozumi Yamaguchi, Teruo Higashino, Mineo Takai, "Smartphone Applications Testbed Using Virtual Reality", in *Proceedings of the 15th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous 2018)*, pp. 422-431, November, 2018.

(2) Tatsuya Amano, Shugo Kajita, Hirozumi Yamaguchi, Teruo Higashino, Mineo Takai, "VRing: Bring Your Mobile App to Virtual Space", in *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers (UbiComp 2018)*, pp. 327-330, October, 2018.

(3) Katsuya Ogura, Yuma Yamada, Shugo Kajita, Hirozumi Yamaguchi, Teruo Higashino, Mineo Takai, "Ground Object Recognition from Aerial Image-based 3D Point Cloud" in *Proceedings of the 11th International Conference on Mobile Computing and Ubiquitous Networking (ICMU 2018)*, October, 2018.

(4) Nathavuth Kitbutrawat, Shugo Kajita, Hirozumi Yamaguchi, Teruo Higashino, "Location Identification of BLE-Embedded HVACs for Smart Building Management" in *Proceedings of the 14th International Conference on Intelligent Environments (IE 2018)*, pp. 1-4, June, 2018.

(5) Ryosuke Narimoto, Shugo Kajita, Hirozumi Yamaguchi, Teruo Higashino, "Wayfinding Behavior Detection by Smartphone" in *Proceedings of the 32nd IEEE International Conference on Advanced Information Networking and Applications (AINA 2018)*, pp. 488-495, May, 2018.

Abstract

The worldwide movement toward increasing availability and usability enhancement of Wi-Fi in public areas has become more active. Several companies and governments have deployed outdoor public Wi-Fi access points (APs) to promote Wi-Fi communication environment improvement. Also, Wi-Fi is expected to be utilized for cellular traffic offloading, smart city platform, wireless sensor network in the Internet of Things (IoT) context etc. Therefore, Wi-Fi becomes one of the most important infrastructures and is expected to be penetrated to support our social life.

On the other hand, since the number of APs is steadily increasing and they compete for the limited bandwidth, 2.4GHz unlicensed band becomes more congested particularly in urban areas. This leads to chaotic and disorderly channel competition, and results in critical performance degradation. As many APs with existing architectures such as IEEE802.11a/g/n have already been installed, interference mitigation techniques that operate on them with few modifications are required for efficient frequency reuse in the current urban environment.

To assure a certain level of communication quality even in such circumstances, Wi-Fi channel selection is a simple but promising technology. However, considering Wi-Fi-specific features, the best quality channel is not easily estimated only from the monitored traffic in each channel. This is mainly because there are several factors that cause interference and noise. Wi-Fi traffic in adjacent channels may become noise signals like non-Wi-Fi devices since Wi-Fi channels are not completely separated in terms of the spectrum they use, particularly in 2.4GHz band (IEEE802.11b/g). Therefore, IEEE802.11 frames in one channel may become noise for another channel. We call it *inter-channel interference problem*. This is very significant in such a situation like urban areas where many Wi-Fi systems use different and uncoordinated channels. Additionally, traffic and RSSI diversity within each channel makes the inter-channel interference problem more complex.

Our research goal is to identify the best quality channel with less interference based on simple passive frame monitoring. For this, we build a numerous number of simulation scenarios to understand the effect of inter-channel interference from adjacent channels, traffic volume and RSSI, and the combination inter-channel interference problem as well as the diversity of traffic and RSSI. Our proposed approach falls into the category of interference prediction. Compared with the previous approaches that pursue the similar goals, we take an approach of leveraging simulation-based big data in modeling and analyzing the performance of Wi-Fi under interference from the traffic in both the same and

different channels with various RSSIs. As far as we know, this is the first approach to assessing Wi-Fi channel quality based on such simple measurement, using simulation-based big data analysis.

Firstly, we present an algorithm to estimate each channel performance by multiple regression functions. As it is often hard to identify the channel with less interference in the urban situation, we present a channel scoring function that estimates the performance level of each channel based on the concept of interference environment sensing. We apply the IEEE802.11 MAC frame monitoring in each channel, which can be obtained by the off-the-shelf devices with low-cost. In order to build the scoring function based on the observations, we conduct exhaustive simulations with a large number of scenarios, and multiple regression analysis is applied where channel occupancy patterns, traffic volumes and RSSI in those channels are used as explanatory variables. To evaluate our method, this scoring function is examined in two kinds of general and realistic scenario (typical and dense scenario) where several APs interfere with the AP of interest in a 150m $\times$ 150m region. We confirm that the scores and the actual performance are well-matched where the Spearman's rank correlation coefficient was over 0.8 and can identify the top-ranked channel as well.

Secondly, we present an improved approach to predict each channel performance for channel selection at the target AP. In order to let APs not select erroneous channels, it is quite essential to provide an estimation function for APs to correctly estimate the channel status without actually moving into it. Therefore, we prepare more than 10,000 scenarios and conducted simulations which are assumed that the own traffic of the target AP moves to the new channel to simulate "channel state change". We apply a machine learning based classification algorithm to estimate the channel saturation due to the traffic movement in channel migration and multiple regression analysis to build a prediction function of channel performance under saturation. We confirm that our function can classify the channel state accurately and estimate the frame delivery ratio with less than 10% error in average with additional 2,000 simulations.

Thirdly, we design the realistic urban scenario based on the actual measurement in Osaka city to evaluate our proposed approach in more realistic environment. We monitored Wi-Fi traffic at ten locations including shopping malls, cafes, commercial buildings and stations around Osaka station on both weekdays and holidays to understand the current traffic situation in such typical urban environment. Also, we obtain the actual AP locations and their corresponding channels from Wi-Fi Radio Map of Osaka City which has been built by our group. We match this information with OpenStreetMap in the network simulator, and then we constructed three real geographical scenarios for this evaluation experiment. In these evaluation scenarios, we demonstrate that it is possible to predict and select the best channel with the highest communication quality and to predict the trend over all channels by our proposed method in all the scenarios. In the subsequent validity evaluation experiment, the experimental result shows that the throughput of the target AP becomes about 1.73 times higher than the value by a random channel selection strategy.

Through these contributions, it will be shown that our channel management mechanism offers efficient frequency reuse by passive frame monitoring at the point of interest. In particular, we focus on the inter-channel interference issue in urban environment where many APs in different systems are densely deployed. This dissertation has established the foundation of prediction-based autonomous channel management with few modification of the current Wi-Fi AP mechanism for reducing the unnecessary interference in urban environment.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Japan is the host country for the Tokyo Olympic and Paralympic in 2020. Since foreign tourists tend to use Wi-Fi to avoid the monetary cost of using LTE, the Ministry of Internal Affairs and Communications issued an action plan, called "SAQ2 JAPAN Project", which promotes Wi-Fi communication environment improvement and realizes the world's highest level ICT environment in June 2014 [4]. A cellular company, Softbank, has also provided nationwide 400,000 Wi-Fi access points (APs) for foreigners [5]. As a result, it is reported that the percentage of users explicitly turning on their Wi-Fi interface of their smartphone during the day increases from 50% to 60% to offload 3G or 4G traffic [6]. Also, outdoor public APs have been deployed by AT&T [7], Time Warner Cable etc. in large cities of US. The worldwide movement toward increasing availability and usability enhancement of Wi-Fi in public areas has become more active.

In addition, Wi-Fi has also been important as an alternative infrastructure of mobile networks in case of disasters as well as low-cost smart city platform. In Barcelona, the urban environmental information such as street lights, human flows, parking, temperature, air quality and noise levels are aggregated through a Wi-Fi-based platform [8, 9]. Such Wi-Fi-based wireless sensor network has been investigated in the Internet of Things (IoT) context [10, 11] and several applications have been proposed [12–14]. Also, Audi has developed a vehicular Wi-Fi system called Audi Connect [15] and the other companies are now focusing on on-board Wi-Fi devices for intra-vehicle (V2V) communication [16]. In the field of Intelligent Transportation System (ITS), applying Wi-Fi infrastructure to Roadside-to-Roadside (R2R), Roadside-to-Vehicle (R2V) and Vehicle-to-X (V2X) communication has been investigated [17–22]. This tread will promote Wi-Fi-based communication in the future ITS systems due to its low cost and high-penetration features. Therefore Wi-Fi becomes one of the most important infrastructures and is expected to be penetrated to support our social life.

On the other hand, congestion in an unlicensed band (especially 2.4GHz band) becomes severer in urban areas as Wi-Fi devices are more and more popular in a variety of infrastructures and services [23]. Cisco Meraki reported the average number of interfering APs was about 55 in 2.4GHz band and has doubled from July 2014 to January 2015 [24]. This is because Wi-Fi APs for private/home use are densely installed and more users have become to hold mobile APs called mobile routers. Therefore the number of Wi-Fi devices is steadily increasing and they compete for the limited bandwidth. In 5G of cellular mobile communications which faces a similar problem, integrated multi radio access technology (RAT) solutions are proposed [25–28]. In an unlicensed band, IEEE802.11ax task group (TGax) addresses such a dense Wi-Fi problem with the goal of enhancing throughput-per-area [29]. For example, TGax reports that Wi-Fi throughput can nearly be doubled using Dynamic Sensitivity Control (DSC) and Transmit Power Control (TPC) in dense networks [30]. However, since many APs with existing architectures such as IEEE802.11a/g/n have been already installed, autonomous interference mitigation techniques that operate on them with few modifications are required for efficient frequency reuse in the current urban environment.

As we believe that this frequency is still the significant frequency that can maximize the number of potential participants to the services, we focus on 2.4GHz band interference problems in this dissertation. In order to assure a certain level of communication quality even in such circumstances, intelligent Wi-Fi channel selection is a simple but promising technology due to low modification cost. To select the best quality channel, a naive but straightforward approach is to directly examine the performance by active probing of each channel [31]. However, it needs the target system to be tuned into each channel and run probing, which is a time-consuming task. In addition, such probe packets may cause severe degradation in performance of Wi-Fi networks [32]. Another possibility is (passive) traffic monitoring of channels to seek interference-free channels [33]. We may estimate the quality of channels based on the traffic information monitored in each channel.

However, considering Wi-Fi-specific features, real performance is not easily estimated only from the monitored traffic in each channel. This is mainly because there are several factors that cause interference and noise. Wi-Fi traffic in adjacent channels may become noise signals like non-Wi-Fi devices since Wi-Fi channels are not completely separated in terms of the spectrum they use, particularly in 2.4GHz band (IEEE802.11b/g). Therefore, IEEE802.11 frames in one channel may be noise for another channel. We call it *inter-channel interference problem.* This is very significant in such a situation like urban areas where many Wi-Fi systems use different and uncoordinated channels. Additionally, traffic and RSSI diversity within each channel makes the inter-channel interference problem more complex. More concretely, it is not easy to assess the impact of both signal strength and traffic volume in the same channel or in a different channel on the performance of the target system. It seems that Wi-Fi channel selection issues have been well investigated, but it has not been discussed how the inter-channel interference affects the performance, how it is closely related with RSSI and traffic volume, and how we should choose a channel in an open, uncoordinated situation.

This dissertation presents a novel approach to select interference-free channels with simple passive frame monitoring in the current 2.4GHz Wi-Fi systems. The goal of this dissertation is to identify the best quality channel with less interference based on the knowledge about the effect of inter-channel interference from adjacent channels, the effect of traffic volume and RSSI, and the combination inter-channel interference problem and the diversity of traffic and RSSI. We address the issue by introducing a simple monitoring scheme to the target AP and predicting each channel performance for channel selection based on exhaustive simulation dataset. In this dissertation, we make the following three primary contributions to embody this idea.

Firstly, we present an algorithm to estimate each channel performance by multiple regression functions. As it is often hard to identify the channel with less interference in the urban situation, we present a channel scoring function that estimates the performance level of each channel based on the concept of interference environment sensing. To cope with the problem, our approach for ranking function derives a relative indicator of channel quality based on realistic, observable parameters like inter-channel distance, RSSI and traffic volume. We apply the IEEE802.11 MAC frame monitoring in each channel, which can be obtained by the off-the-shelf devices with low-cost. In order to build the scoring function based on the observations, we conduct exhaustive simulations with a large number of scenarios, and multiple regression analysis is applied where channel occupancy patterns, traffic volumes and RSSI in those channels are used as explanatory variables. Relying on exhaustive simulations but with a reduced number of simulation cases, our model built by regression analysis achieves sufficient accuracy to estimate better Wi-Fi channels. To evaluate our method, this scoring function is examined in two kinds of general and realistic scenario (typical and dense scenario) where several APs interfere with the AP of interest in a 150m $\times$ 150m region. We confirm that the scores and the actual performance are well-matched where the Spearman's rank correlation coefficient was over 0.8 and can identify the top-ranked channel as well.

Secondly, we present an improved approach to predict each channel performance for channel selection at the target AP. In order to let APs not select erroneous channels, it is quite essential to provide an estimation function for APs to correctly estimate the channel status without actually moving into it. Therefore, we prepare more than 10,000 scenarios and conducted simulations which are assumed that the own traffic of the target AP moves to the new channel to simulate "channel state change". We analyze the interference dataset for understanding the relationship between the observed parameters and interference effect to build our proposed function. We apply machine learning based classification algorithm to estimate the channel saturation due to the traffic movement in channel migration and multiple regression analysis to build a prediction function of channel performance under saturation. We confirm that our function can classify the channel state accurately and estimate the frame delivery ratio with less than 10% error in average with additional 2,000 simulations. Moreover, we demonstrate that our estimator can capture the tendency of overall channel performance in a more general scenario for channel selection. The experimental result shows the correlation coefficient between our estimator output and the groundtruth is above 0.85.

Thirdly, we design realistic urban scenarios based on the actual measurement in Osaka city to evaluate our proposed approach in these scenarios. We conduct the traffic monitoring at ten locations including shopping malls, cafes, commercial buildings and stations around Osaka station on both weekdays and holidays to understand the current traffic situation in such typical urban environments. Then we analyze this actual measurement results and designed the traffic parameters from the future traffic prediction based on the distribution on the number of APs in each channel. Also, we obtain the actual AP locations and their corresponding channels from Wi-Fi Radio Map of Osaka City which has been built by our group. We match this information with OpenStreetMap in the network simulator, and then we constructed three real geographical scenarios for this evaluation experiment. In these evaluation scenarios, we demonstrate that it is possible to predict and select the best channel with the highest communication quality and also predict the trend over all channels by our proposed method in all scenarios. In the subsequent validity evaluation experiment, the experimental result shows that the throughput of the target AP becomes about 1.83 times as compared with the expected value when the target AP select a channel randomly.

Through these contributions, it will be shown that our channel management mechanism offers efficient frequency reuse by passive frame monitoring at the point of interest. In particular, we focus on the inter-channel interference issue in an urban environment where many APs in different systems are densely deployed. This dissertation has established the foundation of prediction-based autonomous channel management with few modifications of the current Wi-Fi AP mechanism for reducing unnecessary interference in the urban environment.

The rest of this dissertation is organized as follows. Chapter 2 reviews related work on Wi-Fi channel management techniques. Chapter 3 describes the design and performance of the simulation-based channel performance estimator by simple MAC frame monitoring. Chapter 4 proposes the improved approach for taking into account "channel state change" due to own channel migration without actual moving. Chapter 5 shows the evaluation result of our channel management approach in realistic scenarios based on the actual observations. Finally, Chapter 6 summarizes and concludes this dissertation.

Chapter 2

Related Work

2.1 Interference Mitigation Algorithm

To avoid interference in cellular networks, inter-cell interference avoidance has been well-investigated [34–42]. For example, Fractional Frequency Reuse (FFR), which allocates different frequencies to users around cell boundaries and the same frequency to those near base stations, has been developed so far [43]. Cooperative network control mechanisms such as cell clustering are also effective for such (closed) cellular networks [44, 45], but cannot be applied to uncontrolled Wi-Fi systems directly. Ref. [46] presents the various analysis of resources and performance with time, frequency and spatial-division multiplexing. Ref. [47] presents an autonomous power and resource control mechanism for efficient spatial reuse.

Also, in the fifth generation of cellular mobile communications (5G), integrated multi-radio access technology (RAT) solutions are proposed [25–28]. 5G focuses on the typical ultra dense wireless networks as shown in Ref. [23]. Ref. [25] shows the relationship between access point density and bandwidth partitioning in such environment. Ref. [27] presents the downlink interference statics in a dense deployed scenario simulating Tokyo. Ref. [28] mentions that the offloading from cellular network to Wi-Fi network is shifting towards a true integration of both technology families. Therefore, Wi-Fi has also been important as an alternative infrastructure of mobile networks.

As smartphone users increase, the interference mitigation approaches from client side are proposed. Since the traffic of video streaming application which smartphone users often use is significant, several approaches focus on to improving the video bitrate by adopting to network dynamics. Ref. [48] presents a channel-aware video streaming mechanism. This approach is based on a lightweight channel characterization method that can provide an accurate airtime estimation by the observation of Wi-Fi management packets. Ref. [49] provide the efficient AP channel scanning mechanism for real-time streaming and they focus on the handover environment. In the ITS field, smartphone users in vehicular also try to offload the cellular traffic to Wi-Fi networks. Ref. [17] proposed auction game-based offloading mechanism to avoid unnecessary interference. We focus on the interference mitigation approach of the AP at the location of interest.

In an unlicensed band, IEEE802.11ax task group (TGax) addresses such a dense Wi-Fi problem with the goal of enhancing throughput-per-area [29]. For example, TGax reports that Wi-Fi throughput can nearly be doubled using Dynamic Sensitivity Control (DSC) and Transmit Power Control (TPC) in dense networks [30]. On the other hand, Ref. [6] shows that the percentage of users explicitly turning on their Wi-Fi interface of their smartphone during the day increases from 50% to 60% to offload 3G or 4G traffic. However, the nearly 40% of all clients remain 2.4GHz only and the most of APs are running in 2.4GHz. Moreover, the overall delivery ratio of 2.4GHz links have degraded over the past six months.

Therefore, since many APs with existing architectures such as IEEE802.11a/g/n have been already installed, autonomous interference mitigation techniques that operate on them with few modifications are required for efficient frequency reuse in the current urban environment. Cisco Meraki reported the average number of interfering APs was about 55 in 2.4GHz band and has doubled from July 2014 to January 2015 [24]. In the existing IEEE802.11ac, Ref. [50] proposed the measurement-based practical approach to improve the performance. They applied the dynamic channel assignment algorithm and the sophisticated ACK mechanism to improve the IEEE802.11ac performance. Ref. [51] proposed a framework to measure and characterize Wi-Fi latency at large scale and investigated the result in Tsinghua campus. Ref. [52] investigated the relationships between 802.11n physical layer transmission features. They focused on PHY parameters such as rate and channel width adaptation, frame aggregation, and MIMO settings.

Frequency Hopping (FH) is another approach to mitigate interference effect in Wi-Fi network. Many FH systems like Bluetooth use static subchannel hopping sequences and they continue hopping along the sequences. However, FH systems move over subchannels regardless of subchannel status, and this often causes serious performance degradation in a new subchannel. Moreover, it causes a certain overhead in hopping, and in particular, Wi-Fi system is not designed for frequent hopping among channels due to its association overhead between APs and clients. Although some work like Ref. [53] considers dynamic channel sequences, it needs channel status estimation by monitoring or some other techniques.

Besides channel selection techniques, adaptive carrier sense threshold control and transmission rate control has been considered for densely-deployed Wi-Fi APs. Interestingly, from the research results in [54], Ref. [55] addresses the fact that the transmission power of most Wi-Fi APs is configured to maximum in the factory settings, which often induces unnecessary interference, but self-control of transmission power by APs may cause unidirectional links. Therefore, cross-layer control is recommended where the carrier-sense threshold is coordinately controlled with transmission power and the transmission power of APs with heavy traffic load should be larger. Ref. [56] presents a distributed channel selection algorithm and an AP selection strategy for clients. However, the goal of this approach is fairness among clients while ours is to identify such channels with the least interference effect. It is worth noting that both [55, 56] predict performance by the Gibbs sampling method and this principle can be used for online learning and building of our function.

2.2 Channel Management Technique

Cisco Meraki reported the importance of channel planning using a utilization measure [24]. To select the best quality channel, a naive but straightforward approach is to directly examine the performance by active probing of each channel. However, it needs the target system to be tuned into each channel and run probing, which is a time-consuming task. Ref. [31] found that up to 90% of probe responses carry redundant information. Moreover, the probe traffic can be as high as 60% of the management traffic. They proposed an algorithm to control the probe traffic for reducing unnecessary active scans. Also, Ref. [32] shows that such probe packets may cause severe degradation in performance of Wi-Fi networks. Then, we decided to apply only passive monitoring for our proposed channel management. Ref. [33] proposed CSpy to find the best quality channel without probing in 5GHz band. They reported CSpy improved the performance by up to 100% in comparison to channel agnostic schemes.

Some researches have focused on channel control of Wi-Fi APs in an autonomous environment where cooperation can not be expected. Ref. [57] provides an survey on some of proposed approaches. The research direction can be divided into three categories (i) centralized [58], (ii) decentralized approach [56], (iii) channel hopping [53]. However, in these works, the methods are mainly designed to utilize only non-overlapping channels. Interference can be reliably avoided if all the APs are under control and use only non-overlapping channels, but it is not realistic considering urban environments. On the other hand, the availability of partially overlapping channels has been attracted attention in recent works [59–62]. In other words, it is promising to increase network capacity and network throughput by allowing inter-channel interference. [63] mentions the interference relationship between the inter-channel distance (the distance of the center frequency) and the physical distance of the AP in the 2.4GHz band. Also, based on the interference relationship in [63], [60, 61, 64] reports that the channel assignment converges to the Nash equilibrium when each access point repeats dynamic channel selection in an autonomously decentralized manner. In this method, assuming that each node acts to reduce the number of adjacent interference nodes, convergence is shown by grasping the behavior of the entire network as a potential game. In [60, 61, 64], it is decided based on the interference graph mentioned above, but more advanced operation becomes possible by incorporating the interference quantification part of our proposed method.

It is difficult to balance communication and monitoring with only one interface, considering the monitoring cost. Ref. [49] provide the efficient AP channel scanning mechanism for real-time streaming which is combined passive scan with active scan. It is possible to aim for reduction of switching overhead by using multiple interfaces and optimizing configuration parameters while switching monitoring and communication interfaces [65]. In this research as well, we will implement monitoring module on the premise of this system.

2.3 Passive Channel Monitoring Method

As stated earlier, active scanning may cause severe communication performance degradation [32]. Therefore, passive channel monitoring methods have been proposed for various purposes. Ref. [33] proposed CSpy to find the best quality channel without probing in 5GHz band. CSpy predicts the strongest channel by probing only a single channel based on channel impulse response (CIR) which can be obtained from some Wi-Fi chipsets. For Wi-Fi performance prediction, Ref. [66] presents an approach of estimating frame collision and loss rates in IEEE 802.11 MAC. It employs probabilistic models to infer backoff occurrence due to carrier-sense operations.

RSSI monitoring is often utilized for wireless devices detection and its merits and demerits have been discussed in the past. For example, [67] has pointed out SNR and RSSI do not provide sufficient information to estimate L2 performance, but recent work [68] presents a novel method to accurately identify the presence of non-Wi-Fi machines using information obtained through off-the-shelf Wi-Fi cards. This is done by machine learning where RSSI variation is modeled as pulse waves.

RSSI information is also used for localization of wireless devices. Some recent work deals with positioning Wi-Fi access points with mobile phones [69,70]. On the contrary, Wi-Fi fingerprinting [71–73] is popularly used for positioning mobile devices. Ref. [74] investigates the AP position estimation error that comes from the difference of the Wi-Fi devices used for Wi-Fi scanning. The method proposed in Ref. [75] gets radio wave incoming direction using the directional antennas and estimates AP location. Ref. [76] localize roadside AP from moving vehicles using beam directional antennas. Ref. [77,78] estimate the direction of arrival of radio wave from the change of the receive signal strength with the movement of the observer, and Ref. [79] estimates the direction of arrival using a smartphone by rotating the observer at the observation points. Ref. [80] uses Channel State Information (CSI) which is information including the phase of Wi-Fi radio waves, which is difficult for ordinary smartphones to obtain.

In summary, RSSI measurement has been used in a variety of methods since it does not require extra hardware dedicated for the measurement. In these works, RSSI measurements is the primary means to estimate the distance between devices. With these developments, the creation of Wi-Fi maps based on smartphone users' crowdsensing is thriving. Participatory sensing and crowdsourcing are considered to be effective methods for constructing a spatial database aggregating Wi-Fi information. Ref. [81] investigate the load of APs in the campus wireless network using Wi-Fi channel scan dataset collected by cooperative smartphone users, and shows that channel scans by ordinary smartphones are useful in monitoring enterprise Wi-Fi network. There are some methods to survey large scale Wi-Fi radio status by crowdsourcing, war-walking and war-driving [82]. Radio maps generated by these methods are mainly used for smartphone localization in an indoor environments [83]. Place Lab [84] collects Wi-Fi fingerprints for client localization by war-driving. Ref. [85] uses crowdsourcing and users' motions to construct the indoor radio map of a floor plan. The difference between scan data by war-driving and war-walking is addressed in Ref. [86]. A localization method as well as an indoor and radio map construction indoor is proposed in Ref. [87], which estimates the relative positions of APs using multidimensional scaling. MCNet [85] demonstrates Wi-Fi performance measurement system using crowdsourcing. They aggregate performance data directly measured by mobile Wi-Fi client and delect problems in Wi-Fi networks. There are some crowdsensing system for collecting Wi-Fi beacon data such as Wigle [1]. However, it is basically difficult to rely solely on observational data from war-walking or wa-driving for constructing radio maps covering a wide range of cities, because the number of collaborators and behavior patterns have a large influence on observation density. Although interpolation methods such as Kriging mainly used in Geographic Information System (GIS) have been applied to the radio map construction [88–90], it is difficult to consider reflection of radio waves by buildings with those methods in urban areas.

2.4 Cross Technology Interference in 2.4GHz band

In the unlicensed band, effective avoidance of "cross technology interference" should be considered for better performance of wireless communication [91–93]. This is because different wireless technologies (*i.e.* Wi-Fi, Bluetooth, game controllers and ZigBee devices) follow different protocols and most of them are not designed with the coexistence of multiple technologies. For example, Ref. [10] presents the network interface between ZigBee and Wi-Fi in the IoT context. Interference between Wi-Fi and Bluetooth has been discussed so far [94, 95], and recent work [96] presents a new Wi-Fi MAC design for the coexistence of Wi-Fi and Bluetooth.

For more general coexistence problems, Ref. [97] presents TIIM, a Technology-Independent Interference Mitigation solution that detects and reacts to cross technology interference in realtime in IEEE 802.15.4. To detect and identify the type of cross technology interference, TIIM applies the link quality indicator value, used in ZigBee networks. If TIIM detects the interference, countermeasures are automatically determined by the classifier made using a decision tree algorithm. To cope with the issue of coexistence in ISM band, IEEE 802.19 Wireless Coexistence Working Group [98] is working for standardization.

Chapter 3

An Algorithm for Estimating Channel Performance by Multiple Regression Analysis

3.1 Introduction

Unlicensed band becomes more congested particularly in an urban environment as Wi-Fi devices are more and more popular in a variety of infrastructures and services. For example, as illustrated in Figure 3.1, public Wi-Fi services at shops, cafeteria and convenience stores as well as private Wi-Fi utilization in offices and homes compete for the limited bandwidth. Although another band (such as 5GHz band) is available in many countries for Wi-Fi, 2.4GHz is still significant frequency that can maximize the number of potential participants to the services.

Therefore, in order to assure a certain level of communication quality even in such circumstances, it is necessary for APs to fully utilize the available channels although neighboring 4 channels are overlapped in 2.4GHz Wi-Fi. In order to observe the status of 13 channels, a naive but straightforward approach is to directly examine the performance by active probing of each channel. However, it needs the target AP to be tuned into each channel, to ask a client to participate in the probing procedure and finally to run probing, which is a time-consuming task. Another possibility is (passive) traffic monitoring of channels. We may estimate the quality of channels based on the traffic information monitored in each channel. However, considering Wi-Fi-specific features, real performance is not easily estimated only from the monitored traffic in each channel. This is mainly because there are several factors that cause interference and noise. Not only noise from non-Wi-Fi systems, but Wi-Fi traffic in adjacent channels may also become noise signals like non-Wi-Fi devices since Wi-Fi channels are not completely separated in terms of the spectrum they use, particularly in 2.4GHz band (IEEE802.11b/g). Therefore, IEEE802.11 frames in one channel may be noise for another channel. We call it *inter-channel interference problem*, which is referred to as *ICI* problem hereafter. This is very significant in such a

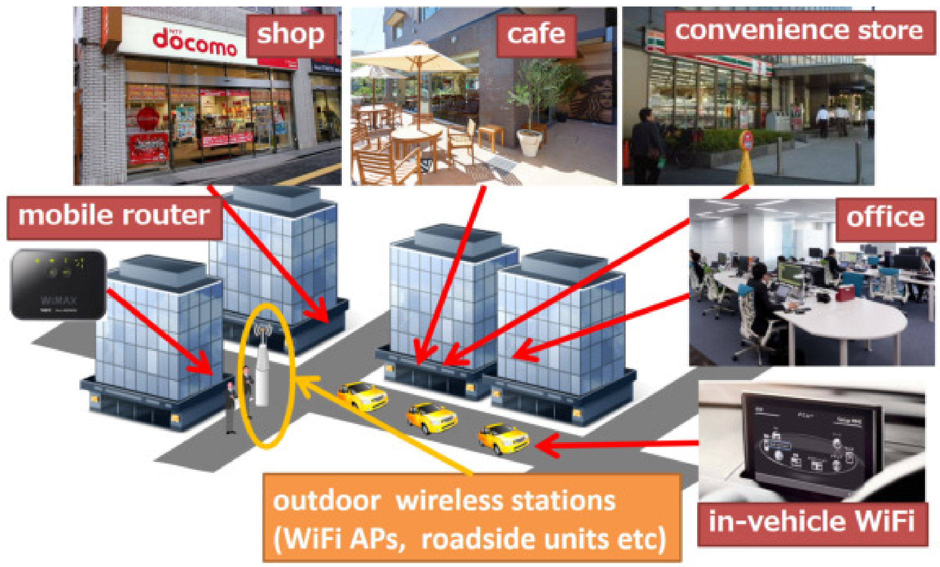


Figure 3.1: Wi-Fi concentration in urban environment

situation like an urban environment where many Wi-Fi systems use different channels. Additionally, traffic and received signal strength (RSSI) diversity within each channel makes the ICI problem more complex. More concretely, it is not easy to assess the impact of both signal strength and traffic volume in the same channel or in a different channel on the performance of the target system.

In this chapter, we present a novel strategy to choose Wi-Fi channels in an urban environment. As stated earlier, adjacent channels interfere with each other in the current Wi-Fi systems. This often makes it very complex to estimate the performance of the target system in presence with other Wi-Fi systems at different locations, which use different channels with different traffic volume. In our method, we build a function to predict how much the target system is affected by interference from the other systems, by taking inter-channel distance, RSSI levels and traffic volumes into account. This function is built based on a number of data with different parameter values, and the dataset has been obtained by exhaustive simulations using realistic network simulator called *Scenargie* 1.7 [99], which can simulate the complete protocol stack from the PHY level (OFDM subchannel spectrum spread) to the application layer, as well as the IEEE802.11 family. Finally, multiple regression analysis is employed to represent the performance metrics (delay and frame delivery ratio) by the observed values.

To evaluate our method, this scoring function was examined in two kinds of general scenarios where several APs interfere with the AP of interest in a 150m $\times$ 150m region. In the first scenario, we assumed the target AP in a typical ITS scenario where the several interference nodes affect the target AP to confirm the basic performance of our designed function. The second scenario was designed like an urban dense environment. We deployed 50 AP-client pairs as the interference sources in a

150m × 150m region randomly. In both scenarios, we have confirmed that the scores and the actual performance are well-matched where the Spearman's rank correlation coefficient was sufficiently high and can identify the top-ranked channel as well.

3.2 Approach Design

This section presents our approach of proposed method to cope with the urban interference problem. We firstly describe the overview of our approach and the problem which we focus on. Secondly, we define our target scenario in urban environment and assume the situation. Finally, we describe the design of explanatory parameters obtained from simple MAC frame monitoring. We note that this is a unique approach to the urban Wi-Fi problem, and as far as we know, this is the first approach to ranking channel quality based on simple measurement of inter-channel distance, RSSI and traffic.

3.2.1 Overview

As we have discussed earlier, we focus on (i) inter-channel interference where adjacent channels interfere with each other in Wi-Fi systems and (ii) urban situations where many APs in different systems are deployed in an uncoordinated way. To cope with the problem, our approach for ranking function derives relative indicator of channel quality based on realistic, observable parameters like inter-channel distance, RSSI and traffic volume. We predict how much channel is affected by interference compared with the other channels for channel selection. We monitor the IEEE802.11 data frames in each channel, which can be obtained by the off-the-shelf devices with low-cost. However, it is difficult to gather such information in real urban environment to quantify the effect of the inter-channel interference and the diversity of RSSI and traffic volume. Then we decide to use the network simulator *Scenargie* [99] to model such interference effect and build the exhaustive dataset. Scenargie is a commercial network simulator that supports the IEEE802.11 specification [100] such as IEEE802.11a/g/n/ac and can simulate the complete protocol stack from the PHY level (OFDM subchannel spectrum spread) to the application layer. Since the implementation is accurate and reliable, the simulation results are sufficiently dependable [101]. As the relative indicator, we have built a function that anticipates the performance level in each channel from a given set of traffic and RSSI information in the channels. In order to build the function, we have applied multiple regression analysis to the exhaustive simulation dataset. We selected the linear function to estimate the channel performance level by considering the computational cost.

3.2.2 Preliminaries

We consider a target system is an IEEE802.11g AP with clients (the AP is referred to as *target AP* in this chapter) and propose a method to rate its performance (L2 delay simply called *delay* and L2 frame delivery ratio (we use the term *FDR* to refer to the ratio without confusion) for each channel as a channel selection algorithm. Basically we assume urban situation where a number of APs and their clients of different systems or mobile routers exist among the target AP. In indoor environment, building administrators constitute operational policy of utilization of Wi-Fi channels and in such a coordinated environment, static allocation is much better. Instead, in urban outdoor, a number of public APs are deployed each of which is not under control, *i.e.* highly uncoordinated situation needs to be considered. Our collaborator Sumitomo Electric Industries LTD. is now investigating possibilities to exploit 2.4GHz ISM band devices for ITS roadside units for V2R and P2R communication. A typical scenario is that roadside CCTVs and IR-beacon transmitters detect vehicles driving through the streets, and these information is aggregated to the AP at the intersection.

We assume that target APs have an IEEE802.11MAC monitoring function as well as RSSI detection. It is worth noting that IEEE MAC level information is easy to be captured using off-the-shelf devices and tuned drivers. For example, using Atheros WLAN chips, such information is available in promiscuous mode. The target AP monitors the traffic of other systems (*interference source*) that use the same or other channels.

As we have discussed earlier, it is not easy to estimate interference effect in a “chaos” environment like urban environment. As an example, we will show in the experiment section the following scenarios where channels 1, 7 and 11 have been used by IEEE802.11g APs of different systems. In this case, it is not easy to assess which channels are better than others due to the ICI problem and due to different traffic volume and RSSI from those APs. The detailed results will be presented later in Section 3.4, but channel 1 is the best in the case. We design the rating function that determines the levels of channel status for given information about traffic and RSSI in each channel.

3.2.3 Channel Monitoring and Explanatory Parameters Definition

We denote the set of all APs and their clients (except the target AP and its clients) that use a channel k as $I(k)$. Each AP or client in $I(k)$ is called *interference source*. For each channel k, we obtain the following information about interference sources by IEEE802.11MAC frame monitoring and some additional information.

(a) *Normalized received interference signal strength*

This is called *RSSI indicator* of channel k and denoted as $s(k)$. We define $s(k)$ as the normalized averaged RSSI ($SS_{ave}(k)$) of data-frames transmitted by interference sources in $I(k)$ as follows.

$$s(k) = \begin{cases} \frac{SS_{ave}(k)-\theta_{min}}{\theta_{max}-\theta_{min}} & \theta_{max} \geq SS_{ave}(k) \\ 1 & \theta_{max} < SS_{ave}(k) \end{cases} \tag{3.1}$$

where θ_{min} and θ_{max} represent the minimum RSSI threshold of data frame reception (-90dBm in IEEE802.11g) and expected maximum RSSI (usually -50dBm or around), respectively. From my preliminary experiments of channel monitoring in urban outdoor with the wireless traffic packet capture *AirPcap* [102] and the network tool *iperf* [103], the expected $SS_{ave}(k)$ value is lower than -50 dBm. But, in the case of $SS_{ave}(k) > -50$ dBm, we define $s(k)$ value dose not exceed 1.

(b) *Normalized traffic volume*

This is called *traffic indicator* of channel k and denoted as $t(k)$. We define $t(k)$ as the normalized data bytes transmitted by interference sources in $I(k)$ as follows.

$$t(k) = \frac{8 \cdot d(k) + b \cdot q(k) \cdot T_{preamble}}{b} \tag{3.2}$$

where $d(k)$ is the byte amount of data frames per second and b is the transmission rate of IEEE 802.11b/g. Control information of the PHY layer is transmitted, but in MAC frame monitoring this cannot be observed. In equation3.2, the occupancy duration of PHY control information is added to $t(k)$. For this purpose, we define $q(k)$ and $T_{preamble}$. $q(k)$ denotes the total received count of data frame per second from all interference sources in $I(k)$ and $T_{preamble}$ denotes the length of time that the control information of the PHY layer is transmitted (*i.e.* the length of $T_{preamble}$ in IEEE 802.11g is 20μs).

(c) *Inter-channel distance*

This is called *inter-channel distance indicator* and denoted as $c(c_{\mathrm{t}}, k)$. We define $c(c_{\mathrm{t}}, k)$ as the normalized inter-channel distance between the channel c_{t} of the target AP and channel k used by at least one interference source;

$$c(c_{\mathrm{t}}, k) = \frac{|c_{\mathrm{t}} - k|}{c_{\max}} \tag{3.3}$$

where $c_{\max}$ is the maximum channel distance within which two nodes interfere. From our preliminary experiments, two nodes with inter-channel distance larger than 3 do not significantly interfere with each other with any RSSI and traffic. Therefore, we set $c_{\max} = 3$ and interference source with $|c_{\mathrm{t}} - k| > c_{\max}$ is ignored.

As briefly stated earlier, $SS_{ave}(k)$ and b are parameters from the PHY layer, but off-the-shelf WLAN devices (*e.g.* those using Atheros chips) can obtain this information through normally-provided drivers. For example, the above information can be displayed by *iwconfig* command.

3.3 Design of a Rating Function by Multiple Regression Analysis

This section presents how to build our rating function to quantify the inter-channel interference effect. In order to investigate the interference effect with different traffic volume, RSSI and channel distance, we have built the exhaustive simulation scenario. In following section, we describe the design of the simulation scenario and how to model the function.

3.3.1 Basic Strategy

In order to estimate the performance of the target AP in each channel, it is required to understand how RSSI, traffic volume and inter-channel distance indicators affect the performance. Furthermore, usually multiple channels are occupied by interference sources. For example, let us consider a scenario; the target AP scanned all the channels and observed that (a) some APs use channels 2 and 8 with heavy traffic and weak RSSI and (b) some others use channels 5 and 11 with marginal traffic and strong RSSI. The question is which channel is the best for the target AP, and a naive answer is to choose one without any APs (such as channel 3, 4, 6, $7\cdots$). However, traffic from channels 2 and 8 may become noise signal in those channels and it is not easy to estimate their effect. Therefore, we conducted exhaustive simulations to obtain the model to estimate the interference effect.

Before addressing exhaustive simulations and multiple regression analysis, we investigate the number of cases that we need for the exhaustive simulations. We let n_s and n_t denote the number of "levels" that are contained in $s(k)$ and $t(k)$, respectively. We also let K denote the set of all the possible channel occupation patterns. In order to completely explore all the possible cases, we need

$$\sum_{h\in 1..2c_{\max}+1} \left\{ {}_{(2c_{\max}+1)}\mathrm{C}_h \cdot (n_s \cdot n_t)^h \right\} \tag{3.4}$$

cases where C denotes a combination. In the above, ${}_{(2c_{\max}+1)}\mathrm{C}_h$ denotes the number of occupancy patterns of h channels, $(n_s \cdot n_t)^h$ denotes the number of RSSI and traffic patterns for each occupancy pattern of h channels. For example, in case that $n_{rs} = n_{tr} = 30$ and $c_{\max} = 3$ (the settings used later), totally we need 4.8×10^{20} cases. As this number is not realistic even in an offline process, we try to reduce the number of combinations by taking the symmetry of occupancy pattern of channels into account. For instance, if the target AP uses channel 6 and interference sources use channel 3, 5, 7, occupancy pattern of channels is (-3, -1, 1). I consider the result of pattern(-1, 1, 3) is the same of pattern(-3, -1, 1). Therefore, the number of occupancy patterns of h channels is as follows.

$$\frac{{}_{(2c_{\max}+1)}\mathrm{C}_h + {}_{c_{\max}}\mathrm{C}_{\lfloor \frac{h}{2} \rfloor}}{2} \tag{3.5}$$

In addition, I also try to reduce the cases by taking the following strategy.

(1) For each $c(c_\mathrm{t}, k)$, we conduct simulations for all the combinations of $s(k)$ and $t(k)$.

(2) We apply linear regression analysis to obtain the regression model of the performance for given $s(k)$ and $t(k)$. Linear multiple regression analysis is one of the analytical methods which is used for multivariate analysis. This prepares a polynomial formula consisting of a linear sum of independent variable groups that explain observation values. In this analysis, we calculate and determine a linear coefficient that best explains the observation value. In our method, this is called *single ICI model* (ICI denotes inter-channel interference) and denoted as $f_{c(c_\mathrm{t},k)}(s(k), t(k))$. This represents how RSSI and traffic affect the performance if only channel k is occupied by interference sources.

(3) For each set $\{k_1, k_2, ..., k_L\} \in K$ of channels occupied by interference sources, we conduct simulations for limited combinations of $s(k_1)$, $t(k_1)$, $s(k_2)$, $t(k_2)$, ... , $s(k_L)$ and $t(k_L)$.

(4) We apply linear regression analysis to obtain the regression model of the performance for given $f_{c(c_t,k_1)}(s(k_1), t(k_1))$, $f_{c(c_t,k_2)}(s(k), t(k))$, ... and $f_{c(c_t,k_L)}(s(k_2), t(k_2))$ as well as $c(c_t, k_1)$, $c(c_t, k_2)$, ... and $c(c_t, k_L)$ to obtain the final function for given RSSI, traffic and inter-channel distance indicator values. This is called *aggregated ICI model* and denoted as f_{multi}. Our channel selection algorithm uses f_{multi} to choose a channel.

3.3.2 Single ICI Model

We have built the single-ICI model by analysis of simulation results. Simulation settings are defined by generalizing the situation of Section 3.2. The client uploads obtained information to the AP periodically, and the physical distance between them is set to 100m. In the scenario, the nodes are static and deployed the fixed location. The interference sources are a pair of AP and its client, and traffic between them is created by iperf implemented on the Scenargie simulator [99] with changing the parameter *iperf-udp-rate-bps*, in order to arrange different $t(k)$ values. For different $s(k)$ values, the distance between the target AP and interference source AP is changed from 10m to 300m with step 10m. We have measured (i) the frame delivery ratio (which we call *FDR*) from the client to the target AP and (ii) the MAC layer transmission duration (which we call *delay*) which is obtained as time duration from the moment that a frame is queued at the client till the moment that the frame is queued at the target AP. The simulation scenario is illustrated in Figure 3.2 and the setting is shown in Table 3.1. The APs and clients follow IEEE802.11g standards and work in 2.4GHz. The clients is associated the determined APs and works at the same channel number as the AP's initial channel setting. In this scenario, every nodes use the fixed data rate BPSK 3/4 in IEEE802.11g.

In this scenario, we have 30 $s(k)$ values and 30 $t(k)$ values ($n_s = n_t = 30$). In addition, we have 4 $c(c_t, k)$ ($= c_{\max} + 1$) values for an interference source pair of AP and client. Therefore, we have 3600 simulation cases as a total. Each case was simulated for 30 seconds.

We have used the following linear function for regression analysis.

$$f_{c(c_t,k)}(s(k), t(k)) = c_1 + c_2 \cdot s(k) + c_3 \cdot t(k) + c_4 \cdot s(k) \cdot t(k) \tag{3.6}$$

We designed the above function which contains $t(k)$,$s(k)$ and the interaction term as the descriptive variables. We have built this function with different 4 $c(c_t, k)$ from 0 to 1.

3.3.3 Determining Parameters of Single ICI Model

The result of the analysis is summarized in Tables 3.2 and 3.3.

With $c(c_t, k) = 0$, the target AP and interference sources reside in the same channel. In this case, the target AP can hear the preamble of the other APs' frame transmission, which may explicitly

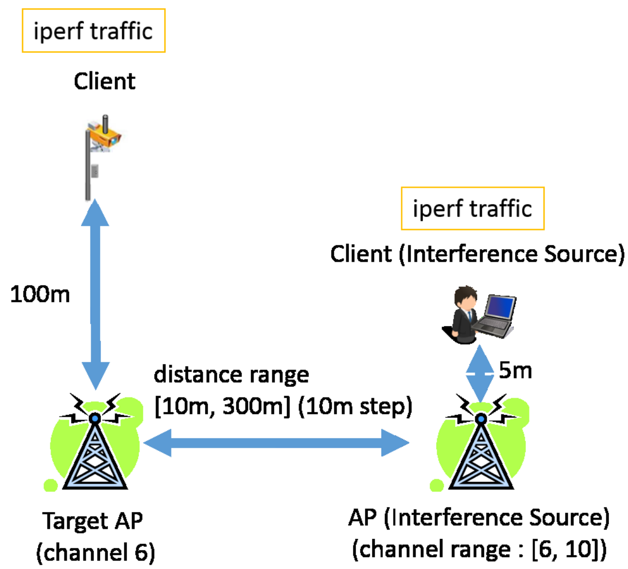


Figure 3.2: Simulation scenario for single ICI model

prevent the target AP's frame transmission. Therefore, $t(k)$ is more significant than $s(k)$. Meanwhile, with $c(c_{\mathrm{t}}, k) = 1/3$ or larger, $s(k)$ as well as $t(k)$ also affects the performance since the traffic from the interference sources becomes "noise" for the target AP, which affects the AP's carrier-sense behavior and SNR (this causes frame error, *i.e.*, FCS is likely to be false). It seems that $s(k)$ is more tightly related with the performance than $t(k)$, which supports our hypothesis. As a total, the adjusted R^2 is mostly above 0.8 in all the cases except $c(c_{\mathrm{t}}, k) = 3/3$, and around 0.75 even in case of $c(c_{\mathrm{t}}, k) = 3/3$. The adjusted R^2 value is an index which shows the explanation accuracy of the linear expression constructed by the regression analysis. As this value is closer to 1, it is judged that the tendency of the objective variable is better grasped. This shows that our linear regression successfully represents the effects of $s(k)$ and $t(k)$ for each inter-channel distance (0, 1, 2 and 3). Based on this result, we move to the step (3) to analyze the effect of multiple k's.

Here, we describe the preliminary experiment about $cmax$. In the same scenario in Figure 3.2 and 3.1, we have conducted 900 scenario where the interference AP's channel number is configured as 10, that is, the inter-channel distance is 4. In this scenario, we have 30 $s(k)$ values and 30 $t(k)$ values ($n_s = n_t = 30$). The simulation result are shown in Figure 3.3. From these figures, we confirmed that there was no interference effect for the 98% results in the sight of the simulated value of delay and the frame delivery ratio. We also confirmed that the remaining 2% results are very limited cases where the target AP and the interference source are the nearest among the set values of 10m. Since the proposed method assumes an outdoor environment such as an intersection in the major cities, we ignore these cases where the inter-channel distance is 4.

Table 3.1: Simulation settings for single ICI model

Common Parameter	Values
Area Size	400m×400m
Target AP - its STA Distance	100m
Interference AP - its STA Distance	5m
Wireless Standards	IEEE802.11g
Wireless Band	2.4GHz
Target AP's Channel	6
Transmission Power	20dBm
IEEE802.11g Data Rate	9Mbps (BPSK 3/4)
Antenna Height	1.5m
$T_{preample}$	20μs
Propagation Model	Free Space
L7 application	*iperf-udp-data-rate*
L7 traffic (Target AP)	9Mbps
Payload Size	1470byte
for Single ICI Model	Values
Target Pair - Interference Pair Distance	[10m, 300m] step=10m
Interfernece AP's Channel	{6,7,8,9,10}
L7 traffic (Interference AP)	Iperf [0.3Mbps, 9Mbps] step=0.3Mbps

Table 3.2: Regression analysis result with $f_{c(c_t,k)}(s(k), t(k))$ (delay)

$c(c_t, k)$	Coefficient				adjusted R^2 (delay)
	c_1	c_2	c_3	c_4	
0	-0.38498	-0.86602	5.89684	1.27298	0.905
1/3	1.3917	-3.7342	-12.7026	35.1980	0.9029
2/3	1.5988	-3.8891	-16.6614	40.9565	0.8895
3/3	0.4015	-0.9238	-11.2069	25.4772	0.7379

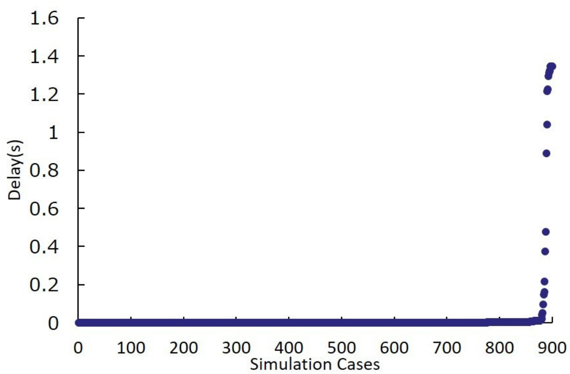
3.3.4 Aggregated ICI Model

Based on the strategy in the Basic Strategy section, we prepare the following regression function f_{multi}. For simplicity, we have dealt with the case that the number of occupied channels is 2 in this chapter.

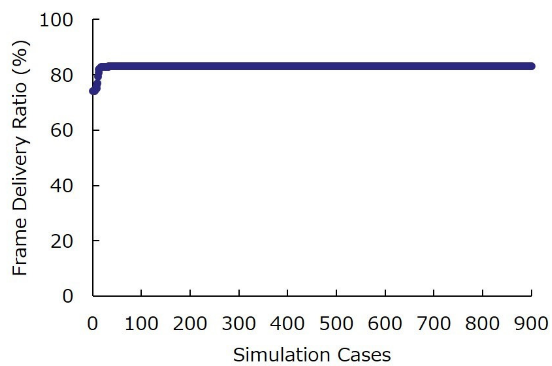
Then using the scenario of Figure 3.4, we have conducted 1,134 simulations where we have 7 $c(c_t, k_1)$ values, 54 $f_{c(c_t,k_1)}(s(k_1), t(k_1))$, 7 $c(c_t, k_2)$ values and 54 $f_{c(c_t,k_2)}(s(k_2), t(k_2))$. In case of multiple channel occupancy, two sides should be taken into account, *i.e.* $7 = 1 + 2 * c_{\max}$. The simulation setting is shown in Table 3.4 and the common parameter is also shown in Table 3.1. As the same as the simulation for single ICI model, the target client uploads obtained information to the AP periodically, and the physical distance between them is set to 100m. The nodes are also static and deployed the fixed location. In order to reduce the simulation cases, we pick up some $f_{c(c_t,k)}(s(k), t(k))$ values. For different $s(k_1)$ and $s(k_2)$ values, the distance between the target AP and interference source

Table 3.3: Regression analysis result with $f_{c(c_t,k)}(s(k),t(k))$ (Frame Delivery Ratio)

$c(c_t,k)$	Coefficient				adjusted R^2 (FDR)
	c_1	c_2	c_3	c_4	
0	0.86200	0.17056	-0.51439	-0.48568	0.8413
1/3	0.80081	0.14843	0.94830	-2.86823	0.8223
2/3	0.81915	0.07458	1.12194	-2.93707	0.8339
3/3	0.81033	0.05432	1.06558	-2.44258	0.7635



(a) Delay



(b) Frame Delivery Ratio

Figure 3.3: Simulation result with inter-channel distance 4

Table 3.4: Simulation settings for aggregated ICI model

Parameters for Aggregated ICI Model	Values
Interference AP1's Channel	{3,4,5,6,7,8,9}
Interference AP2's Channel	{3,4,5,6,7,8,9} which is not selected by AP1
Target Pair - Interference Pair 1 Distance	[10m, 100m, 300m]
Target Pair - Interference Pair 1	[10m, 100m, 300m] which is not selected by AP1
L7 traffic (Interference AP1)	[1Mbps, 3Mbps, 9Mbps]
L7 traffic (Interference AP2)	[1Mbps, 3Mbps, 9Mbps]

AP is changed from 10m, 100m and 300m. When the interference pair 1 select 10m, the interference pair 2 can not select 10m at the same time. As well as $s(k_1)$, the interference pair 1 and 2 select the different setting about the channel number.

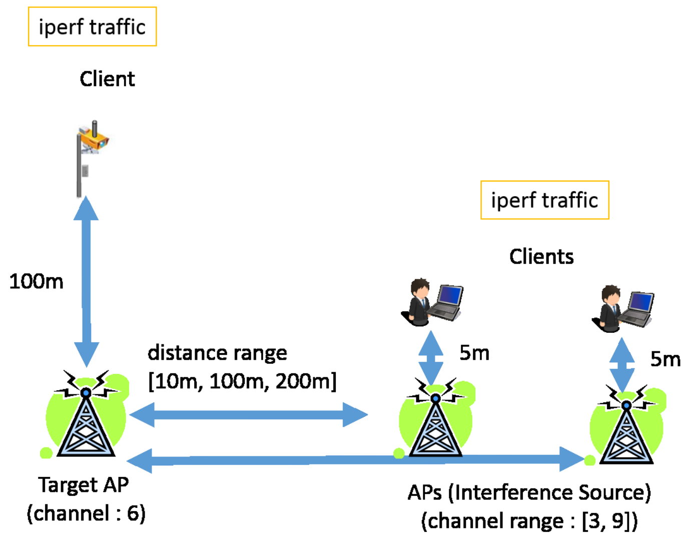


Figure 3.4: Simulation scenario for aggregated ICI model

We designed the aggregated ICI model as following equation 3.7.

$$
\begin{aligned}
& f_{\mathrm{multi}} \\
= \quad & d_1 + d_2 \cdot c(c_{\mathrm{t}}, k_1) + d_3 \cdot f_{c(c_{\mathrm{t}},k_1)}(s(k_1), t(k_1)) \\
& + d_4 \cdot c(c_{\mathrm{t}}, k_2) + d_5 \cdot f_{c(c_{\mathrm{t}},k_2)}(s(k_2), t(k_2)) \\
& + d_6 \cdot c(c_{\mathrm{t}}, k_1) \cdot f_{c(c_{\mathrm{t}},k_1)}(s(k_1), t(k_1)) \\
& + d_7 \cdot c(c_{\mathrm{t}}, k_2) \cdot f_{c(c_{\mathrm{t}},k_2)}(s(k_2), t(k_2))
\end{aligned} \tag{3.7}
$$

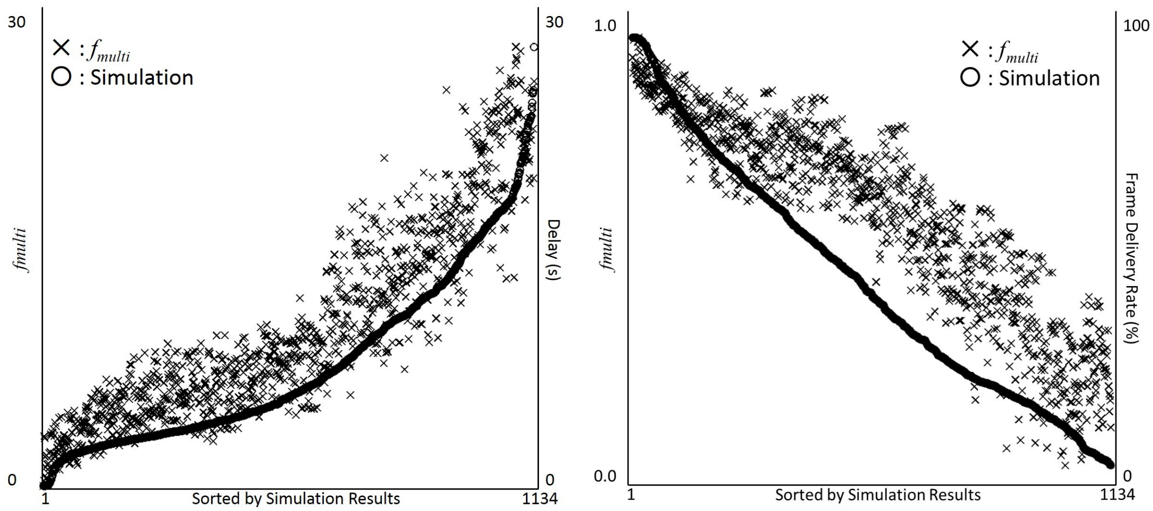
We designed the above function which contains both channel distances $c(c_{\mathrm{t}}, k_1)$, $c(c_{\mathrm{t}}, k_2)$, and the function outputs $f_{c(c_{\mathrm{t}},k_1)}(s(k_1), t(k_1))$,$f_{c(c_{\mathrm{t}},k_2)}(s(k_2), t(k_2))$ as the descriptive variables. In addition, the interaction terms between $c(c_{\mathrm{t}}, k_1)$ and $f_{c(c_{\mathrm{t}},k_1)}(s(k_1), t(k_1))$ and between $c(c_{\mathrm{t}}, k_2)$ and $f_{c(c_{\mathrm{t}},k_2)}(s(k_2), t(k_2))$ are introduced as the descriptive variables. This decision is based on the indicator AIC (Akaike's Information Criterion) which is used for model selection on multiple regression analysis. As a result, we ignore the interaction terms between $c(c_{\mathrm{t}}, k_1)$ and $f_{c(c_{\mathrm{t}},k_2)}(s(k_2), t(k_2))$, between $c(c_{\mathrm{t}}, k_2)$ and $f_{c(c_{\mathrm{t}},k_1)}(s(k_1), t(k_1))$ and others. This result is very reasonable because the inter-channel distance between the target pair and the interference pair 1 and $f_{c(c_{\mathrm{t}},k_2)}(s(k_2), t(k_2))$ value influence independently on the target AP.

3.3.5 Determining Parameters of Aggregated ICI Model

The multiple regression results of our proposed linear function 3.7 are shown in Tables 3.5. Similarly with the single ICI cases, the both values of adjusted R^2 is above 0.8, and for the case of delay, it is

Table 3.5: Regression analysis result for aggregated ICI model

	Coefficient							adjusted R^2
	d_1	d_2	d_3	d_4	d_5	d_6	d_7	
delay	2.24359	-1.10688	0.70291	-2.66502	0.05354	0.54053	1.63336	0.8685
FDR	-0.09109	-0.29729	0.44932	-0.34902	0.34982	0.48801	0.56516	0.8064



(a) Delay

(b) Frame Delivery Ratio

Figure 3.5: Indicator values by f_{multi} and simulations

close to 0.85. Therefore, we can say that the linear function f_{multi} successfully models the delay and frame delivery ratio performance with derived coefficients.

We note that for reference, we have plotted the values of delay and frame delivery ratio indicators obtained by the models and simulations in Figure 3.5. The 1,134 cases are sorted along X-axis by the simulation values. We can also confirm that our multiple regression-based models for delay and frame delivery ratio well-represent the treads of the values by simulations.

In the next section, we conducted experiments to confirm that our aggregated ICI model can be used for general scenarios.

3.4 Performance Evaluation in Typical Scenario

We have examined using more general scenarios the ability of f_{multi} to find out the "best" channel in terms of expected delay and frame delivery ratio. In particular, when the currently-chosen channel does not provide expected quality due to traffic situation changes, our rating scheme provides useful information for the target AP to move to another channel (the target AP may examine one by one from the top-ranked channel to the bottom, which extensively reduce the overhead of channel selection

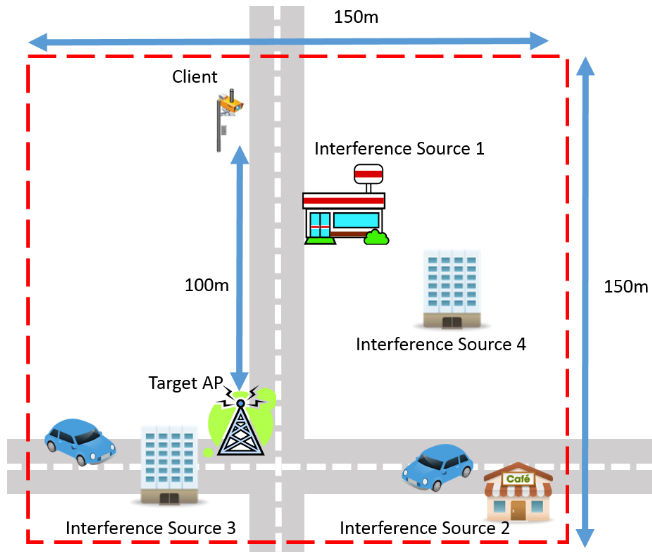


Figure 3.6: Typical simulation scenario

in dynamic situations).

3.4.1 Typical Scenario Setting

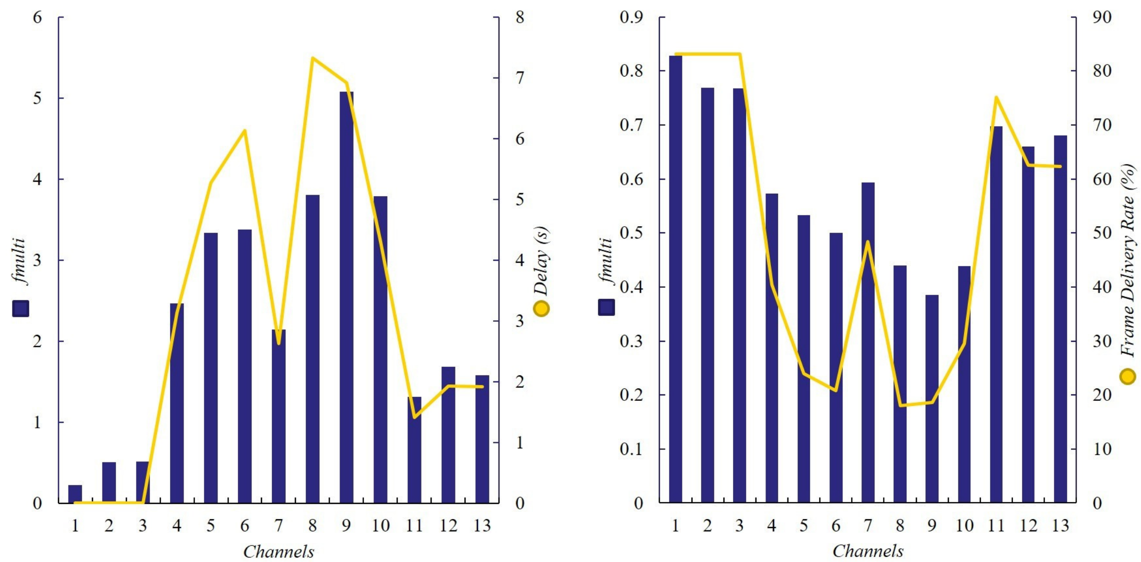
For the target AP and its client, we have deployed four AP-client pairs as interference sources in a 150m × 150 area with crossed roads (Figure 3.6). Table 3.6 shows the coordinates of those interference sources. We assume that the target AP-client pairs is in an ITS roadside unit system that generates *iperf-udp-rate-bps* with 5Mbps. Interference sources are those in a convenience store (AP1), public Wi-Fi AP (AP2) and APs in office buildings (AP3 and AP4). Each client is at 5m north from the location of its associated AP, and their rates are 1.5Mbps, 3Mbps 2Mbps and 3Mbps, respectively. They employ BPSK 3/4 (thus $bitrate = 9Mbps$). Other parameters are configured as the same as the simulation for model building in Table 3.1. The target AP monitors the 13 channels for 30 seconds each. We have compared the ranking of f_{multi} values by the proposed models and that of the real performance metrics by the simulations to examine the accuracy of rating. In order to see the performance in each channel, we have run simulations changing the target AP's channels from 1 to 13.

3.4.2 Evaluation Result

We have summarized the results in Figure 3.4.2 and in Table 3.7. From the figures, the treads of f_{multi} values over 13 channels well-match the actually simulated performance. The tables show the ranking results. In both cases, the Spearman's rank correlations are 0.965035 (delay) and 0.9352028 (frame delivery ratio), which mean very high correlation between the models and the real performance. Therefore, we confirmed that our model could estimate the top-ranked channel and the whole ranking

Table 3.6: Node coordinates and channels in typical scenario

Node	Coords.	Channel
Target AP	(75.000, 25.000)	to be determined
Target Client	(75.000,125.000)	to be determined
Interference AP 1	(87.220,105.632)	1
Interference AP 2	(148.151, 14.946)	7
Interference AP 3	(18.433, 20.508)	7
Interference AP 4	(139.297, 85.083)	11



(a) Delay (b) FDR

Figure 3.7: f_{multi} values (Y1-axis with boxes) and simulated value (Y2-axis with Lines) over 13 channels in typical scenario

with reasonable accuracy.

The top-ranked channel number is 1 in this result. However, the interference pairs uses channel 1, 7 and 11. As we have discussed earlier, we can confirmed that the target pair is affected by inter-channel interference clearly. Especially, we can confirm that channel number 9 is affected from channel 7 and channel 11. Basically, commodity Wi-Fi AP adjusts its own channel based on the existence of other APs by beacon frame scanning. From this result, contrary to our expectations, it was better to use the same channel as the interference APs. Therefore such Wi-Fi APs will be wrong and will be affected by such interference. Our rating function can determines the levels of channel status for given information about traffic and RSSI in each channel by taking inter-channel interference effect into account.

Table 3.7: Experimental results of ranking channel performance in typical scenario

(a) Delay

Channel ID	Model		Simulation	
	Indicator	Ranking	Delay (s)	Ranking
1	0.221914	1	0.002356	1
2	0.501899	2	0.002854	2
3	0.512307	3	0.003902	3
11	1.314429	4	1.414476	4
13	1.574846	5	1.916814	6
12	1.683511	6	1.922286	5
7	2.141591	7	2.62581	7
4	2.463841	8	3.129002	8
5	3.337122	9	5.267372	10
6	3.380002	10	6.132739	11
10	3.789207	11	4.306259	9
8	3.808662	12	7.328154	13
9	5.076176	13	6.919489	12

(b) Frame Delivery Ratio

Channel ID	Model		Simulation	
	Indicator	Ranking	FDR (%)	Ranking
1	0.827794	1	83.08427	1
2	0.769155	2	83.08427	1
3	0.767343	3	83.06915	3
11	0.69729	4	75.08957	4
13	0.680698	5	62.3103	6
12	0.659767	6	62.61121	5
7	0.593243	7	48.34523	7
4	0.57273	8	40.40192	8
5	0.532401	9	23.92966	10
6	0.500817	10	20.76474	11
8	0.439509	11	18.02075	13
10	0.438814	12	29.57189	9
9	0.385057	13	18.61199	12

3.5 Performance Evaluation in Dense Scenario

In previous section, we have confirmed that our proposed function can capture the trend of overall channel quality. However, the investigation about the diversity of traffic volume and RSSI of interference pairs is not enough because the traffic and distance settings of interference pairs are fixed. In addition, the investigation about the density of interference nodes is not enough because we assumed that our proposed method was used in urban environment. Therefore, we have examined using more general 4 urban scenarios the ability of f_{multi} to find out the "best" ranked channel in terms of expected delay and frame delivery ratio.

3.5.1 Dense Scenario Setting

We have designed more 4 scenario which is assumed the urban environment. Comparing the typical scenario, the node density and the diversity of traffic and RSSI are different. In any scenarios, we have deployed 50 AP-client pairs as interference sources in a 150m $\times$ 150m area for the target AP and target client. Interference sources are deployed randomly in this area. Figure 3.8 show the coordinates of those interference sources. The target AP and client are represented as red node (bottom) and red node (top) respectively. Interference sources are represented as blue node and the AP-client pairs are located at the same point.

The transmission is one-way from a client to an AP (uplink) which employs IEEE 802.11g. The payload size for one frame transmission is set to 1470 byte. We assume that the target AP-client pair is in urban Wi-Fi system that employs UDP protocol and generates *iperf-udp-rate-bps* with 5 Mbps. All AP-client pairs employ BPSK 3/4 (thus $b = 9Mbps$). Other simulation settings are the same as the common parameters in the simulation for model building in Table 3.1.

The channel of interference sources is selected among 1, 6 and 11 randomly. These channels are often used as usual if there is a network manager. He/she adjust the AP's channel to 1, 6 or 11 manually when he/she has to manage multiple APs because these channels are not overlapped completely. In these cases, 13, 17 and 20 pairs of interference sources are running in channel 1, 6 and 11 respectively. The target AP monitors the 13 channels for 30 seconds each. To measure the performance of each channel, we have run simulations changing the target AP's channels from 1 to 13.

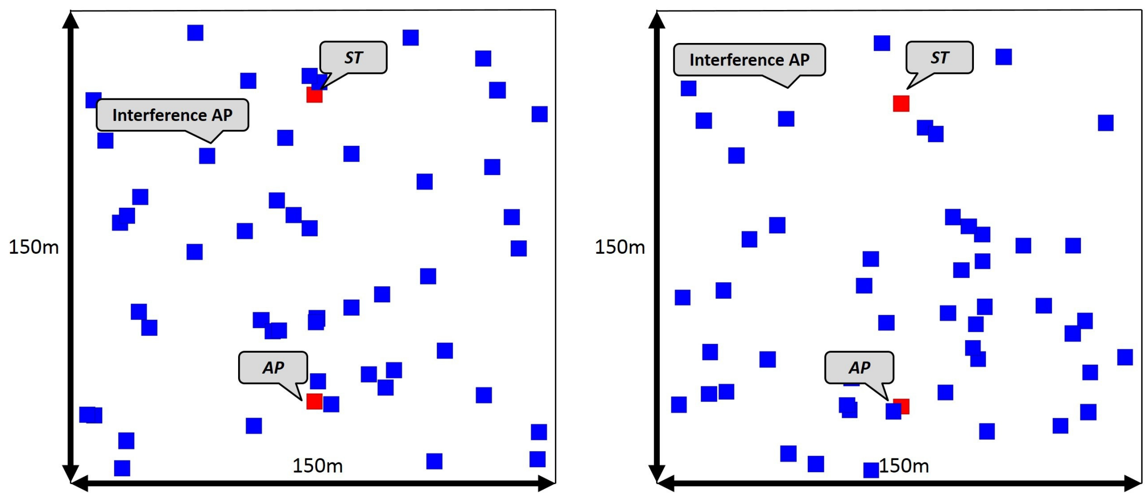
In the following, we explain the four scenarios ($sim1$-$sim4$). The transmission power, the transmission rate and the location of interference sources are different in these scenarios. As a result, the average RSSI and traffic volume which are monitored at the target AP are different, which means that the observed $s(k)$ and $t(k)$ are different in these scenario. The settings of interference sources are determines at random according to Table 3.8. $sim1$ is the normal scenario and $sim2$ is different from the normal scenario in TxPower, which means that the expected $s(k)$ values in $sim2$ is lower than in $sim1$. $sim3$ is designed for checking whether our proposed method can work in the higher $t(k)$ environment. Transmission power and data rate are represented by changing the *dot11-tx-power-dbm* and *iperf-udp-rate-bps* in Scenargie simulation. In $sim1$ and $sim4$, the setting of the transmission power and rate is the same, but the location of interference sources is different. Thus, regardless of the interference sources' position, we confirm that f_{multi} can rank all channels exactly. We confirmed that the target pair can sense all packets of the interference nodes because we used the Free Space propagation model in these scenario. Therefore The signal of all nodes reaches from end to end and we did not consider about hidden node problem in our proposed model.

3.5.2 Evaluation Result

We have summarized the results in Figure 3.9 and 3.10. From the figures, the trends of f_{multi} values over 13 channels well-match the actually simulated performance. The Table 3.10 and Table 3.11 show

Table 3.8: Tx Power settings of interference sources

Simulation	TxPower (dBm)	Transmission Rate (Mbps)
*sim*1, *sim*4	20	[0.1,0.5] (uniform distribution)
*sim*2	[10,20] (uniform distribution)	[0.1,0.5] (uniform distribution)
*sim*3	20	[0.1,0.9] (uniform distribution)



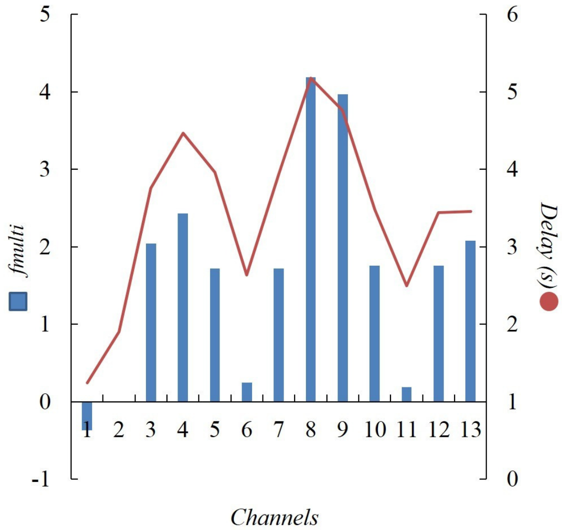
(a) *sim*1-*sim*3 (b) *sim*4

Figure 3.8: Evaluation environment in dense scenario

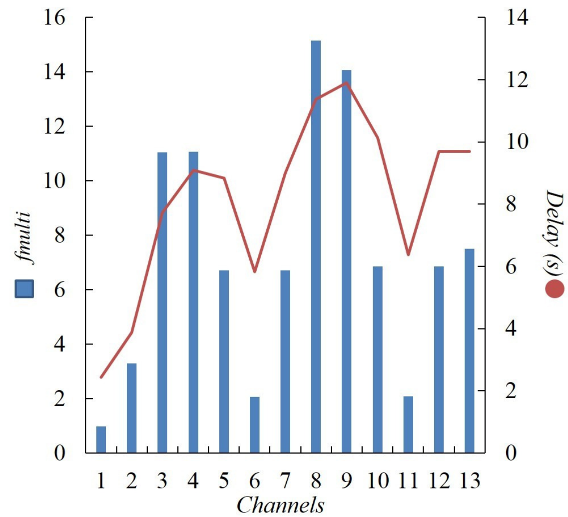
Table 3.9: Summary of Spearman's rank correlations in dense scenario

Simulation	Delay	Frame Delivery Ratio
*sim*1	0.845	0.840
*sim*2	0.931	0.820
*sim*3	0.826	0.853
*sim*4	0.787	0.801

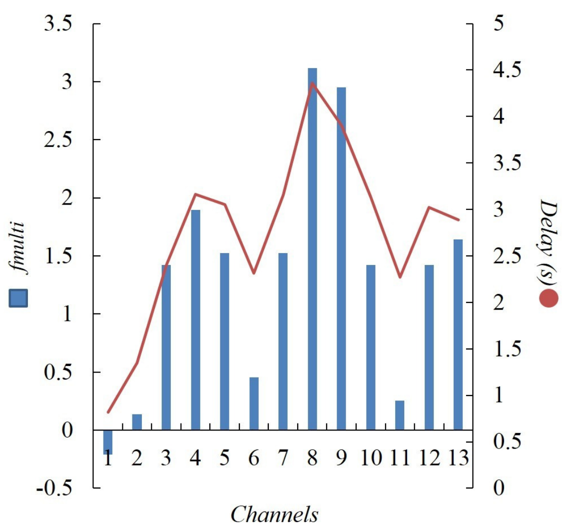
the ranking results of *sim*1 - *sim*4. We can confirm that, in all dense scenario, this function can select the top-ranked channel. Table 3.9 shows the Spearman's rank correlations of *sim*1 are 0.845 (delay) and 0.840 (frame delivery ratio), which means very high correlation between the models and the real performance. Moreover, the Spearman's rank correlations of *sim*2 and *sim*3 are above 0.8 and the correlations of *sim*4 are about 0.8. Especially, in the scenario *sim*3, the error of the predicted channel rank is 3 or less. Therefore, we confirmed that this ranking function could estimate the top-ranked channel and the whole ranking with reasonable accuracy.



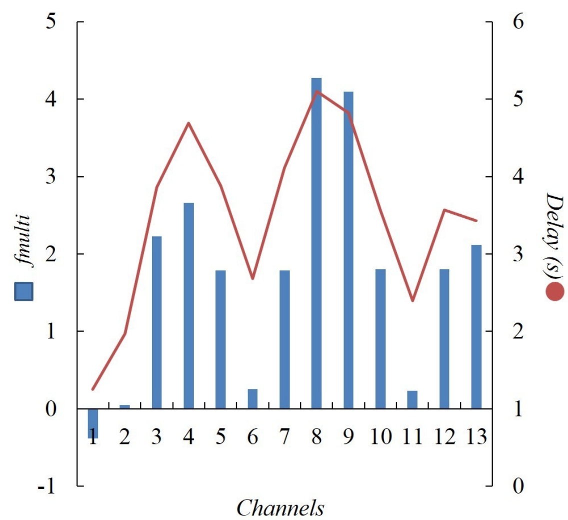
(a) $sim1$



(b) $sim2$



(c) $sim3$



(d) $sim4$

Figure 3.9: f_{multi} values (Y1-axis with boxes) and simulated delay (Y2-axis with lines) over 13 channels in dense scenario

Table 3.10: Experimental result of ranking channel performance in dense scenario (Delay)

(a) *sim*1

Channel ID	*fmulti*		Simulation	
	Indicator	Ranking	Delay (s)	Ranking
1	-0.36543	1	1.244631	1
2	0.005587	2	1.906194	2
3	2.046134	9	3.756066	8
4	2.432525	11	4.466259	11
5	1.723872	5	3.963021	10
6	0.250092	4	2.634693	4
7	1.723872	5	3.945966	9
8	4.189695	13	5.182272	13
9	3.971627	12	4.761439	12
10	1.757880	7	3.478857	7
11	0.192882	3	2.496695	3
12	1.757880	7	3.442613	5
13	2.077132	10	3.454188	6

(b) *sim*2

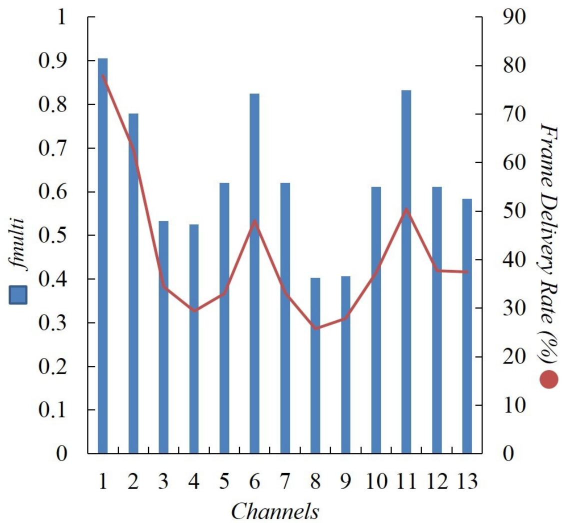
Channel ID	*fmulti*		Simulation	
	Indicator	Ranking	Delay (s)	Ranking
1	0.982828	1	2.430742	1
2	3.301078	4	3.87934	2
3	11.03612	10	7.724116	5
4	11.0678	11	9.086685	8
5	6.716086	5	8.836211	6
6	2.056418	2	5.833651	3
7	6.716086	5	9.006583	7
8	15.15444	13	11.37322	12
9	14.06839	12	11.89855	13
10	6.852393	7	10.12858	11
11	2.075302	3	6.373814	4
12	6.852393	7	9.695095	10
13	7.500844	9	9.687548	9

(c) *sim*3

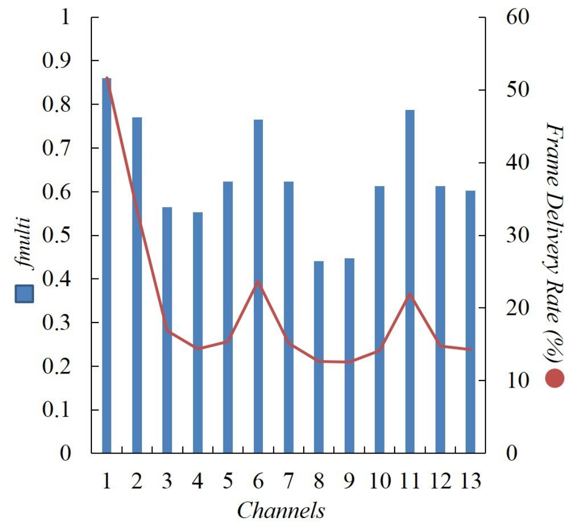
Channel ID	*fmulti*		Simulation	
	Indicator	Ranking	Delay (s)	Ranking
1	-0.20997	1	0.814904	1
2	0.13705	2	1.34891	2
3	1.42199	5	2.404265	5
4	1.895942	11	3.165355	11
5	1.524706	8	3.049665	8
6	0.456469	4	2.3148	4
7	1.524706	8	3.158945	10
8	3.116167	13	4.3608	13
9	2.948739	12	3.902907	12
10	1.422928	6	3.135849	9
11	0.25539	3	2.272857	3
12	1.422928	6	3.0195	7
13	1.639897	10	2.889659	6

(d) *sim*4

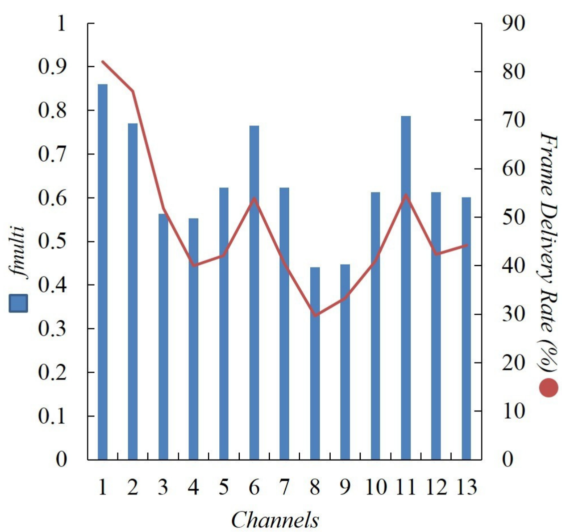
Channel ID	*fmulti*		Simulation	
	Indicator	Ranking	Delay (s)	Ranking
1	-0.37948	1	1.250785	1
2	0.051725	2	1.973766	2
3	2.227958	10	3.863806	8
4	2.661229	11	4.68657	11
5	1.791085	5	3.87793	9
6	0.255992	4	2.684064	4
7	1.791085	5	4.113609	10
8	4.269861	13	5.099492	13
9	4.094081	12	4.818102	12
10	1.803741	7	3.567336	6
11	0.235959	3	2.393839	3
12	1.803741	7	3.568948	7
13	2.114708	9	3.431979	5



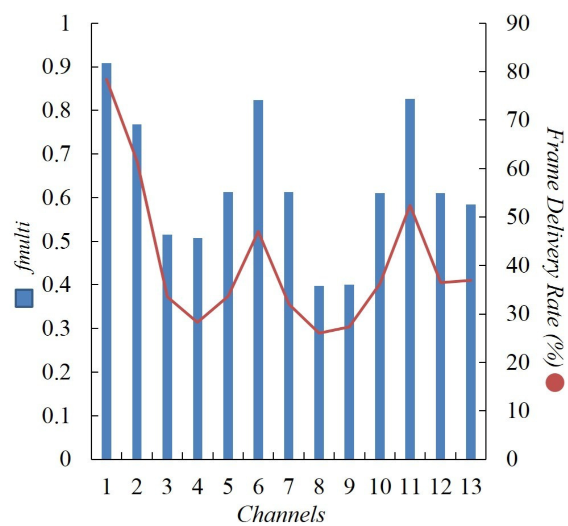
(a) $sim1$



(b) $sim2$



(c) $sim3$



(d) $sim4$

Figure 3.10: f_{multi} values (Y1-axis with boxes) and simulated Frame Delivery Ratio (Y2-axis with lines) over 13 channels in dense scenario

Table 3.11: Experimental result of ranking channel performance in dense scenario (Frame Delivery Ratio)

(a) *sim*1

Channel ID	*fmulti*		Simulation	
	Indicator	Ranking	FDR	Ranking
1	0.905575	1	77.97019	1
2	0.779163	4	62.6963	2
3	0.532546	10	34.41564	8
4	0.524977	11	29.43684	11
5	0.619962	5	32.99166	10
6	0.824607	3	48.00905	4
7	0.619962	5	33.20000	9
8	0.402955	13	25.69994	13
9	0.407218	12	27.86288	12
10	0.611563	7	37.37542	7
11	0.832017	2	50.51788	3
12	0.611563	7	37.72920	5
13	0.583828	9	37.51464	6

(b) *sim*2

Channel ID	*fmulti*		Simulation	
	Indicator	Ranking	FDR	Ranking
1	0.860446	1	51.65802	1
2	0.769958	3	33.41523	2
3	0.563978	10	16.8546	5
4	0.553015	11	14.41043	9
5	0.623696	5	15.42628	6
6	0.765651	4	23.64116	3
7	0.623696	5	15.17963	7
8	0.440856	13	12.69411	12
9	0.446858	12	12.59375	13
10	0.613258	7	14.16637	11
11	0.7867	2	21.9735	4
12	0.613258	7	14.7931	8
13	0.601604	9	14.32164	10

(c) *sim*3

Channel ID	*fmulti*		Simulation	
	Indicator	Ranking	FDR	Ranking
1	0.860446	1	82.1121	1
2	0.769958	3	75.90467	2
3	0.563978	10	51.79418	5
4	0.553015	11	40.03245	11
5	0.623696	5	42.05052	8
6	0.765651	4	53.90625	4
7	0.623696	5	40.43032	10
8	0.440856	13	29.61142	13
9	0.446858	12	33.34396	12
10	0.613258	7	40.94609	9
11	0.7867	2	54.62949	3
12	0.613258	7	42.35375	7
13	0.601604	9	44.14194	6

(d) *sim*4

Channel ID	*fmulti*		Simulation	
	Indicator	Ranking	FDR	Ranking
1	0.9085	1	78.42636	1
2	0.76787	4	61.4238	2
3	0.515168	10	33.56088	9
4	0.507686	11	28.29132	11
5	0.61268	5	33.68212	8
6	0.823724	3	46.99765	4
7	0.61268	5	31.96655	10
8	0.397411	13	26.07525	13
9	0.401091	12	27.37579	12
10	0.609705	7	36.27089	7
11	0.826305	2	52.41948	3
12	0.609705	7	36.50668	6
13	0.583793	9	36.96734	5

3.6 Conclusion

This chapter presents a strategy to choose Wi-Fi channels in urban environment and we have studied the effect of interference in 2.4GHz Wi-Fi. In particular, we consider (i) inter-channel interference where adjacent channels interfere with each other in Wi-Fi systems and (ii) urban situations where many APs in different systems are deployed in an uncoordinated way. It seems that Wi-Fi channel selection issues have been well-investigated, but it has not been discussed how the inter-channel interference affects the performance, how it is closely related with RSSI and traffic volume, and how we should choose a channel in an open, uncoordinated situation. As it is often hard to identify the channel with less interference in such a situation, we present a channel scoring function that estimates the performance level of each channel.

To build the scoring function, we have conducted exhaustive simulations with a large number of scenarios, and multiple regression analysis has been applied where channel occupancy patterns, traffic volumes and RSS in those channels are used as explanatory variables. Relying on exhaustive simulations but with a reduced number of simulation cases, our model built by regression analysis achieves sufficient accuracy to estimate better Wi-Fi channels.

To evaluate our method, this scoring function was examined in two kinds of general and realistic scenario (typical and dense scenario) where several APs interfere with the AP of interest in a 150m $\times$ 150m region. In the first scenario, we assumed the target AP in a typical ITS scenario where the several interference nodes affect the target AP to confirm the basic performance of our designed function. The second scenario was designed like an urban dense environment. We deployed 50 AP-client pairs as the interference sources in a 150m $\times$ 150m region randomly. In both scenario, we have confirmed that the scores and the actual performance are well-matched where the Spearman's rank correlation coefficient was over 0.8 and can identify the top-ranked channel as well.

We note that some contents in this chapter refer our previous publications [104, 105].

Chapter 4

Improvement of the Performance Estimator for Channel Selection by Machine Learning Approach

4.1 Introduction

In Chapter 3, we have presented a basic ranking function to estimate relative quality levels of Wi-Fi channels in urban areas. Taking into account the inter-channel interference, we have built a function to predict how much the target system is affected by such interference from the other systems. As it is often hard to identify the channel with less interference in such a situation, we present a channel scoring function that estimates the performance level of each channel. To build the scoring function, we have conducted exhaustive simulations with a large number of scenarios, and multiple regression analysis has been applied where channel occupancy patterns, traffic volumes and RSSI in those channels are used as explanatory variables. Relying on exhaustive simulations but with a reduced number of simulation cases, our model built by regression analysis achieves sufficient accuracy to estimate better Wi-Fi channels.

However, the following significant issue has not been considered yet, that is, estimation of "channel state change" due to new clients' participation. Since the function of our previous work is to "rank" the current channel status (*i.e.* it is a "diagnosis" function), it is difficult to estimate the channel status *after* an AP actually moves from the current channel into that channel to expect quality improvement. In the worst case, an AP with high volume traffic may move into a channel, which causes serious congestion and saturation. In order to let APs (some of them may be selfish) behave in more intelligent ways and to pursue stability in such an autonomous and uncoordinated Wi-Fi environment, it is quite essential to provide such an estimation function for APs to correctly estimate the status without actually moving into it.

In this chapter, we propose a channel migration technology that can control congestion among Wi-

Fi channels based on the concept of interference environment sensing. Given the information about the other APs' traffic (volume, signal strength and the channel) and its own traffic (volumes) that may be moved to a target channel c, the function predicts capacity saturation in channel c caused by this movement, and then estimates the expected performance under the capacity saturation. Based on the same policy in Chapter 3, we decided to use the network simulator and build the interference dataset in order to quantify the interference effect in channel migration. We have prepared more than 10,000 scenarios and conducted simulations using them. In order to build such estimator, we applied Support Vector Machine (SVM) based machine learning and multiple regression analysis based on the knowledge of investigating the interference dataset. It is possible to indicate a channel that is expected to provide the highest quality.

To validate the model accuracy, we have conducted simulation experiments with different realistic scenarios. For this purpose, we have conducted additional 2,000 simulations. As a result, the function can estimate the frame delivery ratio with less than 10% error in average. Finally, we demonstrate that the proposed function can be used in evaluation scenarios for channel selection at each AP. We show that this function can specify the best channels and APs can increase communication performance compared with other channels. In addition, we confirmed the correlation coefficient between our estimator output and the groundtruth is above 0.85 and our estimator can capture the tendency of overall channel performance.

We note that our proposed method falls into the category of interference prediction based on the passive monitoring of L2 information and RSSI, both of which can be obtained by the off-the-shelf Wi-Fi devices. Compared with the previous approaches that pursue the similar goals, we take an approach of leveraging simulation-based big data in modeling and analyzing the performance of Wi-Fi under interference from the traffic in both the same and different channels with different RSSIs. As far as we know, this is the first approach to assessing Wi-Fi channel quality based on such simple measurement, using simulation-based big data analysis.

4.2 Approach of Improved Method for Channel Performance Estimation

In this section, we present how to improve our basic function for channel migration. We firstly describe the improvement strategy to consider the channel migration effect in our proposed method. Secondly, we mention about the problem formulation and the redesign of monitoring parameters. Finally, we design and build the simulation dataset for investigating the interference effect in the channel migration.

4.2.1 Consideration of Channel Migration

Our goal is autonomous and efficient frequency reuse at each AP which adopts the existing architecture like IEEE802.11a/g/n in urban environment. Assuming the urban environment, it is often hard to identify the channel with less interference. Therefore, we have proposed a basic strategy of building ranking function to estimate relative quality levels of Wi-Fi channels in such environment based on reasonable MAC frame monitoring at the target AP in Chapter 3. In this function, we consider the effect of the diversity of traffic and RSSI from the other Wi-Fi system, and the inter-channel interference. To this end, we have conducted exhaustive simulations with a large number of scenarios because it was not realistic to gather measurement dataset in real environment. In this step, we devised to reduce the number of simulation cases in the step of building the simulation dataset for building the function. We took the observed parameters as explanatory variables which are *normalized received interference signal strength*, *normalized traffic volume* and *inter-channel distance*. Then we applied multiple regression analysis to build the function. We have confirmed that this function can capture the overall channel status trend.

However, "channel state change" has not been considered in this function yet, which means that this function is not suitable for autonomous and dynamic channel migration at the target AP. Since the function is to "rank" the current channel status (*i.e.* it is a "diagnosis" function), it is difficult to estimate the channel status *after* an AP actually moves from the current channel into that channel to expect quality improvement. In the worst case, for example, an AP with high volume traffic may move into a channel, which may cause serious congestion and saturation. As a result, the AP would make an erroneous channel selection. In order to let APs (some of them may be selfish) behave in more intelligent ways and to pursue stability in such an autonomous and uncoordinated Wi-Fi environment, it is quite essential to provide such an estimation function for APs to correctly estimate the status without actually moving into it. Since we introduced the MAC frame monitoring to grasp the current channel status in our approach, we could estimate the channel status after moving with the assumption that the network condition is stable.

For this purpose, we have redesigned the function based on the same strategy as the previous research. We have conducted exhaustive simulations (more than 10,000 cases) which are assumed that the own traffic of the target AP moves to the new channel to simulate "channel state change". We have investigated the simulation dataset to capture not only the interference effect but also the channel state change due to the channel migration of the target AP. As a result, we decided to apply machine learning based classification algorithm to estimate the channel saturation due to the traffic movement in channel migration. In addition, we applied multiple regression analysis to build a prediction function of channel performance (latency and data delivery ratio).

4.2.2 Preliminaries

We let AP denote an IEEE802.11g AP of interest (called *target* AP) and ST denote each Wi-Fi STA which is associated with AP. We assume that AP has commodity Wi-Fi chipsets for traffic monitoring, which are not used for the communication. Suppose AP and its corresponding ST currently uses channel c_{cur} in the Wi-Fi channel set (denoted as C) and monitors the traffic of other AP(s) (interference AP(s)) that use the same or other channels. In particular, AP calculates the following two values for each channel k; (i) temporal channel utilization ratio (or simply channel utilization ratio) denoted as $t(k)$, and (ii) their received signal strength denoted as $s(k)$, (iii) inter-channel distance denoted as $c(k,l)$, respectively. Obtaining these values does not require driver modification, which is a preferable feature for easy installation and operation. For instance, these values such as RSSI, frame length, and transmission rate can be acquired from Atheros chipset in "monitor mode". It is difficult to balance communication and monitoring with only one interface, considering the monitoring cost. It is possible to aim for reduction of switching overhead by using multiple interfaces and optimizing configuration parameters while switching monitoring and communication interfaces [65]. In this research as well, we assumed that the target AP has multiple Wi-Fi interfaces and monitoring module is implemented in the target AP. We note that in order to obtain other information like the number of transmitters, AP needs to maintain a unique address group, which incurs high calculation cost. Instead, our method just needs to update a vector that consists of captured time, frame length, Tx rate and RSSI.

4.2.3 Explanatory Parameters Definition

(i) Indicator of Channel Utilization Ratio: $t(k)$

$t(k)$ is defined as follows.

$$t(k) = min(\sum_{f \in F(k)} \frac{8 \cdot frame\_size(f)}{data\_rate(f)} + |F(k)| \cdot T_{preamble}, 1.0) \tag{4.1}$$

Let $F(k)$ denote the set of all MAC frames observed on channel k, and $frame\_size(f)$ denote the byte size of each observed MAC frame. The data rate $data\_rate(f)$ is determined by the destination client of each MAC frame f, and the data rate of IEEE802.11g is 6, 9, 12, 18, 24, 36, 48 or 54Mbps (in case of OFDM PHY). In the simulation of this thesis, dynamic control of the data rate by the fallback function is not assumed, but it can also be applied to the case where the data rate is different for each client by the above definition. $t(k)$ is normalized to become the value from 0 to 1. In addition, $t(k)$ is corrected based on the total number of received frames and the duration of the preamble. Specifically, $T_{preamble}$ denotes the length of time that the control information of the PHY layer is transmitted, and the length in IEEE 802.11g is 20μs. Due to CSMA/CA features and inter-frame spacing, $t(k)$ cannot be 1.0, but a larger value means higher utilization.

(ii) Indicator of Received Signal Strength: $s(k)$

$s(k)$ is defined as follows.

$$s(k) = \begin{cases} \frac{ave\_rss(k) - \theta_{\min}}{\theta_{\max} - \theta_{\min}} & \theta_{max} \geq ave\_rss(k) \\ 1 & \theta_{max} < ave\_rss(k) \end{cases} \tag{4.2}$$

where $\theta_{\min}$ and $\theta_{\max}$ represent the minimum RSS threshold of data frame reception (-90dBm in IEEE802.11g) and expected maximum RSS (usually -40dBm or around), respectively. $s(k)$ is

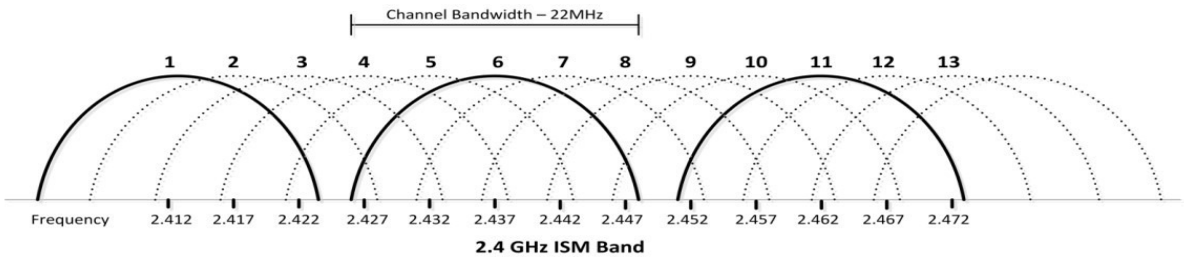


Figure 4.1: Channel allocation in 2.4GHz band

also normalized to become the value from 0 to 1. *ave_rss(k)* is the averaged RSSI of all frames from all APs and STAs in the observation channel k. From my preliminary experiments of channel monitoring in urban outdoor with the wireless traffic packet capture *AirPcap* [102] and the network tool *iperf* [103], the expected $ave\_rss(k)$ value is lower than -40 dBm. But, in the case of $ave\_rss(k) > -40$ dBm, we define $s(k)$ value dose not exceed 1. As well as $t(k)$, a lager $s(k)$ values also means higher interference power.

(iii) Inter-channel Distance: c(k,l)

To cope with the inter-channel interference problem, we define the absolute inter-channel distance (simply called *channel distance* hereafter) between channel k and channel l, denoted as $c(k, l)$, by Equation (4.3).

$$c(k, l) = |k - l| \tag{4.3}$$

The channel distance is expressed as the absolute value of the difference between channels. In 2.4GHz band, two APs are recommended to be operated with channel distance of 5 or larger to avoid the inter-channel interference. For example, channels 1, 6 11 are popularly used in many real situations like Figure 4.1. However, from our preliminary site survey at the downtown in Osaka, APs are operated in an uncoordinated way in such urban outdoor environment because there is no network manager to manage all APs for avoiding the chaotic frequency usage. In this research, we focus on the inter-channel interference under the channel distance $c(k, l)$ 3. In our preliminary simulation experiment, we confirmed that interference from those with channel distance 4 or larger has little affect on the performance. Therefore, in our method, we regard that those channels with 3 or smaller channel distance from channel c interfere with c and we call them *adjacent channels* of c. $AC(k)$ denotes the set of adjacent channels and is defined in Equation (4.4).

$$AC(k) = \{l | c(k, l) \leq 3\} \tag{4.4}$$

In the following section, we describe the design of the dataset to build our proposed estimator.

4.2.4 Design Interference Dataset for Considering Channel Migration

We assume the situation when the target AP is planning to conduct channel migration as follows. AP likes to move the channel from c_{cur} to c_{new}. When the channel migration is conducted, the traffic related with AP in channel c_{cur} is brought into c_{new}. If some surrounding nodes works in the channel set $AC(c_{\mathrm{new}})$, the inter-channel interference between AP and them may occur. Furthermore, usually multiple channels are occupied by interference sources in urban environment. In order to estimate the performance of the target AP in channel c_{new} without actual channel migration, it is required to understand the interference effect caused by the channel migration. Assuming the channel status is stable in the long term, the channel performance could be estimated with the current MAC frame observation result.

Then, our channel selection strategy is designed based on the same as our previous work. Given c_{new} to which AP likes to move from c_{cur}, given observations $t(c_{\mathrm{inf}})$ and $s(c_{\mathrm{inf}})$ where c_{inf} is each channel that may affect the performance of c_{new}, and $t(c_{\mathrm{cur}})$ which is the own traffic volume of the target AP, we provide two functions f_{D} and f_{T}, which return the expected L2 delay and (normalized) L2 frame delivery ratio (*i.e.* channel utilization ratio) after channel migration from c_{cur} to c_{new}, respectively. Having these two estimators, AP can predict the performance when it moves from c_{cur} to c_{new}, just by observing IEEE802.11 MAC frames and their RSSI in channel c_{inf} at AP. Also, taking into consideration the ease of installation and lightweight operation, we only implement the pre-built function on AP and do not conduct resource-consuming operations like online learning. As a result, we decided to apply machine learning based classification algorithm and multiple regression analysis to build a prediction function as described in following section.

In order to design accurate f_{D} and f_{T}, we have to understand relation between the observed $t(c_{\mathrm{inf}})$, $s(c_{\mathrm{inf}})$ and $t(c_{\mathrm{cur}})$ and the corresponding delay/frame delivery ratio at c_{new}. Our basic policy is to use a large dataset, each of which shows the relation to reveal the performance-observation relations and trends.

However, such a dataset is generally hard to obtain in the real world because there are too many combinations of $t(c_{\mathrm{inf}})$, $s(c_{\mathrm{inf}})$ and $t(c_{\mathrm{cur}})$ and for each combination, real equipment has to be configured. This is definitely unrealistic, and we therefore rely on the highly-accurate commercial simulator (*Scenargie* 1.8 [99]) to obtain the dataset. Since it has an accurate OFDM sub-channel spectrum spread model and complete and reliable implementation of the IEEE802.11 family, the simulation results are sufficiently dependable.

4.2.5 Building Interference Dataset for Estimator

According to the above, we have prepared simulation scenario where AP move from channel c_{cur} to c_{new} like in Figure 4.1. The detailed simulation settings are shown in Table 4.1. We put two APs and their corresponding STAs in 400m×400m. One AP is the target AP and another is an interference AP. The distance between each AP and its corresponding STA is fixed 10m and every node is static.

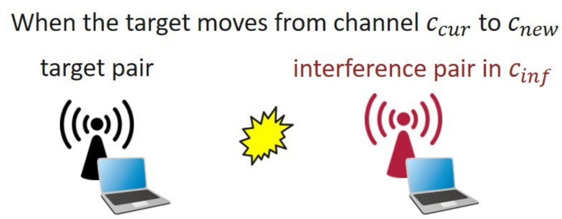


Figure 4.2: Simulation scenario for building interference dataset

Table 4.1: Simulation settings for building interference dataset

Parameter	Values
Area Size	400m×400m
AP - STA Distance	10m
Target Pair - Interference Pair Distance	[20m, 400m] step=20m
Wireless Standards	IEEE802.11g
Wireless Band	2.4GHz
Channels	$c_{\text{new}} = 6$, $c_{\text{inf}} \in \{6, 7, 8, 9\}$
Transmission Power	20dBm
IEEE802.11g Data Rate	9Mbps (BPSK 3/4)
Antenna Height	1.5m
Propagation Model	Free Space
$T_{preample}$	20μs
L7 application	*iperf-udp-data-rate*
L7 traffic (Interference AP)	[0.5Mbps, 9Mbps] step=0.5Mbps
L7 traffic (Target AP)	[1Mbps. 9Mbps] step=1Mbps
Payload Size	1470byte

The distance between the target pair and interference pair is configured for each scenario. Every node follows IEEE 802.11g standards in 2.4 GHz band. STAs follows the channel setting of the associated APs, and the channel settings of APs is set for each scenario. All nodes send packets with the transmission power 20dBm with the modulation BPSK 3/4 in which the maximum data rate is 9Mbps. In these scenario, the traffic is assumed as uplink communication and the traffic demands is configured for each scenario.

We designed these simulation scenarios that gradually change their traffic demand and distance. By changing their traffic demand of application IPERF with the parameter *iperf-udp-data-rate*, channel utilization ratios ($t(c_{\text{cur}})$ and $t(c_{\text{inf}})$) are varied accordingly. We set the traffic demand parameter of the target AP from 1Mbps to 9Mbps by step 1Mbps, which is 9 cases as a total. In addition, we set the traffic demand parameter of the interference AP from 0.5Mbps to 9Mbps by step 0.5Mbps, which is 18 cases as a total. Similarly, by changing their distance, the received signal strength ($s(c_{\text{inf}})$) is varied. We set the distance parameter from 20m to 400m by step 20m, which is 20 cases as a total. Moreover, we prepared the 4 kinds of inter-channel interference situation. After AP's channel migration, their channel distance ($c(c_{\text{new}}, c_{\text{inf}})$) becomes 3 or smaller in these simulation scenarios, which causes the inter-channel interference problem. Consequently, We have totally prepared 12,960 scenarios, and from the obtained simulation-based dataset, we can grasp the impact of traffic/RSSI diversity and inter-channel interference on the performance.

4.3 Design and Build the Performance Estimator

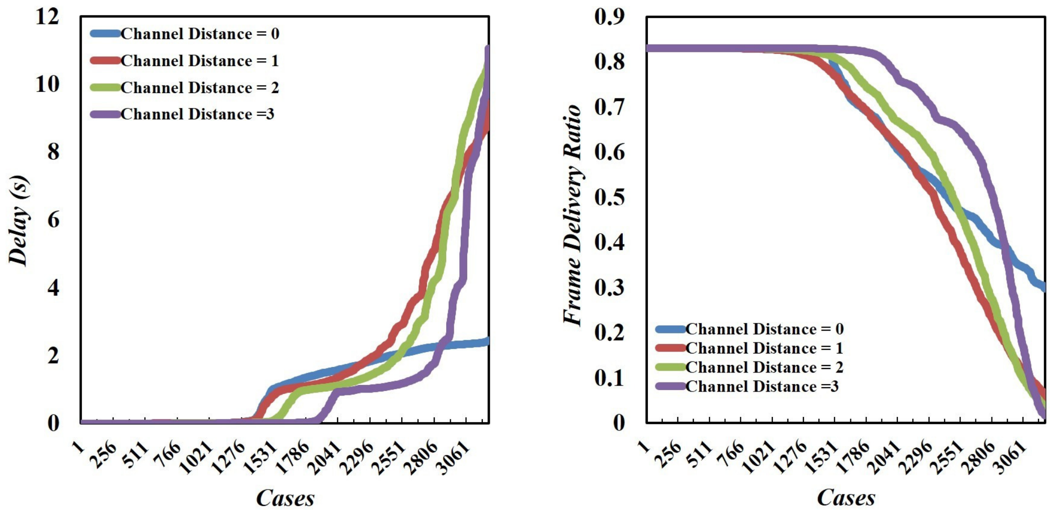
In this section, we present how to build the performance estimator based on the simulation dataset as stated earlier. We firstly describe the investigation result of the interference dataset and how to model the interference effect. Secondly, we design the performance estimator by applying machine learning and multiple regression analysis. Finally, we mention about the way to merge the effect from multiple interference nodes in different channels.

4.3.1 Modeling Estimator Based on Interference Dataset

At first, we show the simulated delay and frame delivery ratio in Figure 4.3. In the Figure 4.3, delay and frame delivery ratio are respectively plotted in the ascending and descending order for each channel distance. We can confirm that the flat trend changes suddenly around the middle of X-axis from the result shown in Figure 4.3. Clearly, this occurs due to channel saturation by the target AP's traffic and interference traffic. From this findings, in order to improve the accuracy of our performance estimator, we should model the saturated situations and unsaturated situations independently because it seems difficult for a single function to capture the joint behavior of flat and increasing/decreasing trends. Then, we provide a two-state prediction function with a binary state classifier (denoted as *sat*) that determines the channel c_{new}'s state as "unsaturated" or "saturated". By surveying these preliminary experiment results shown in Figure 4.3, we empirically define that c_{new}'s state is "saturated" when the observed delay is 100ms or larger. We will use this saturation threshold (delay = 100ms) for determining both of the delay estimator and the frame delivery ratio estimator. We confirmed that the classification results almost match when it is determined by the threshold (frame delivery ratio =0.789).

Our state classifier *sat* is trained by Support Vector Machine (SVM). SVM is a typical clustering method for supervised learning. If *sat* determines that the state of c_{new} is "unsaturated", we can regard that c_{new} can achieve desirable performance. On the other hand, in case of "saturated", AP then predicts how severe the current saturation is.

However, assuming a realistic urban environment, there may be several interference sources for all channels, and there may be no satisfactory channels for AP. Even in that case, we will predict how severe the interference effect is so as to select the channel with the least effect. For this prediction, we provide log-linear-mixed regression function for delay and frame delivery ratio and apply multiple regression analysis using the "saturated" state data. AP only needs to have the function with the determined coefficients, which contributes to lightweight operation.



(a) Delay

(b) Frame Delivery Ratio

Figure 4.3: Definition of saturation (delay/frame delivery ratio)

Our regression functions are given in Equation 4.5 and Equation 4.6 for delay and frame delivery ratio respectively.

$$
\begin{aligned}
& f_{\mathrm{D}}(c_{\mathrm{new}}, c_{\mathrm{inf}}, t(c_{\mathrm{inf}}), s(c_{\mathrm{inf}}), t(c_{\mathrm{cur}})) \\
&= \begin{cases}
0 \quad (|c_{\mathrm{new}} - c_{\mathrm{inf}}| > 3 \textbf{ or } sat = \text{“\textit{unsaturated}”}) \\[2ex]
u_0 + u_1 \log(t(c_{\mathrm{inf}}) + t(c_{\mathrm{cur}})) \\
\ +u_2 \cdot t(c_{\mathrm{inf}}) + u_3 \cdot s(c_{\mathrm{inf}}) + u_4 \cdot t(c_{\mathrm{cur}}) \\
\qquad (|c_{\mathrm{new}} - c_{\mathrm{inf}}| = 0 \textbf{ and } sat = \text{“\textit{saturated}”}) \\[2ex]
v_0 + v_1 \cdot t(c_{\mathrm{inf}}) + v_2 \cdot s(c_{\mathrm{inf}}) + v_3 \cdot t(c_{\mathrm{cur}}) \\
\ +v_4 \cdot t(c_{\mathrm{inf}}) \cdot s(c_{\mathrm{inf}}) + v_5 \cdot s(c_{\mathrm{inf}}) \cdot t(c_{\mathrm{cur}}) \\
\ +v_6 \cdot t(c_{\mathrm{inf}}) \cdot t(c_{\mathrm{cur}}) + v_7 \cdot t(c_{\mathrm{inf}}) \cdot s(c_{\mathrm{inf}}) \cdot t(c_{\mathrm{cur}}) \\
\qquad (0 < |c_{\mathrm{new}} - c_{\mathrm{inf}}| \leq 3 \textbf{ and } sat = \text{“\textit{saturated}”})
\end{cases}
\end{aligned}
\tag{4.5}
$$

$$
\begin{aligned}
& f_{\mathrm{T}}(c_{\mathrm{new}}, c_{\mathrm{inf}}, t(c_{\mathrm{inf}}), s(c_{\mathrm{inf}}), t(c_{\mathrm{cur}})) \\
&= \begin{cases}
1.0 \quad (|c_{\mathrm{new}} - c_{\mathrm{inf}}| > 3 \textbf{ or } sat = \text{“\textit{unsaturated}”}) \\[2ex]
u_0 + u_1 \log(t(c_{\mathrm{inf}}) + t(c_{\mathrm{cur}})) \\
\ +u_2 \cdot t(c_{\mathrm{inf}}) + u_3 \cdot s(c_{\mathrm{inf}}) + u_4 \cdot t(c_{\mathrm{cur}}) \\
\qquad (|c_{\mathrm{new}} - c_{\mathrm{inf}}| = 0 \textbf{ and } sat = \text{“\textit{saturated}”}) \\[2ex]
v_0 + v_1 \cdot t(c_{\mathrm{inf}}) + v_2 \cdot s(c_{\mathrm{inf}}) + v_3 \cdot t(c_{\mathrm{cur}}) \\
\ +v_4 \cdot t(c_{\mathrm{inf}}) \cdot s(c_{\mathrm{inf}}) + v_5 \cdot s(c_{\mathrm{inf}}) \cdot t(c_{\mathrm{cur}}) \\
\ +v_6 \cdot t(c_{\mathrm{inf}}) \cdot t(c_{\mathrm{cur}}) + v_7 \cdot t(c_{\mathrm{inf}}) \cdot s(c_{\mathrm{inf}}) \cdot t(c_{\mathrm{cur}}) \\
\qquad (0 < |c_{\mathrm{new}} - c_{\mathrm{inf}}| \leq 3 \textbf{ and } sat = \text{“\textit{saturated}”})
\end{cases}
\end{aligned}
\tag{4.6}
$$

These equations are configured with unknown parameters (u_i, v_j, u_i and v_j). These unknown parameters are determined by multiple regression analysis. Also predicate $sat_{c(c_{\mathrm{new}},c_{\mathrm{inf}})}(t(c_{\mathrm{inf}}), s(c_{\mathrm{inf}}), t(c_{\mathrm{cur}}))$ is decided by applying machine learning approach.

As stated earlier, predicate *sat* is a binary classifier to predict whether or not the traffic of the target AP in the current channel c_{cur} causes saturation if it is moved to the new channel c_{new}. The first case of both functions represents the unsaturated situation where the delay and frame delivery ratio are assumed to be 0.0 (sec.) and 1.0, respectively.

In the the case of saturated situation where $c(c_{\mathrm{new}}, c_{\mathrm{inf}}) == 0$, we utilize a logarithmically curved function, which models the performance of the CSMA/CA-based systems. The target AP can avoid the collision because AP in $c_{\mathrm{new}} = c_{\mathrm{inf}}$ hears the frames from the interference node directly. It is considered to be different from other cases and the reason for using such a logarithmically curved function can be found in Figure 4.3. Clearly, the trend of interference effect is different from other channel distance results. In this figure, we can see that the increase of delay in zero channel distance case is clearly slower than the other cases. As a result, we designed this function which contains a logarithmically term, $\log(t(c_{\mathrm{inf}}) + t(c_{\mathrm{cur}}))$. Based on this term, this function also has $t(c_{\mathrm{inf}})$, $s(c_{\mathrm{inf}})$ and $t(c_{\mathrm{cur}})$ as the descriptive variables.

Finally, for the rest cases where the channel distance is between 1 to 3, we employ a linear function to model the interference from adjacent channels. We designed the function which contains $t(c_{\mathrm{inf}})$, $s(c_{\mathrm{inf}})$ and $t(c_{\mathrm{cur}})$ as the descriptive variables. In addition, the interaction terms among $t(c_{\mathrm{inf}})$, $s(c_{\mathrm{inf}})$ and $t(c_{\mathrm{cur}})$ are introduced as the descriptive variables.

4.3.2 Determining Classifier and Model Parameters

Firstly, in order to obtain predicate *sat*, we have applied Support Vector Machine (SVM) based learning. We have labeled "*saturated*" or "*unsaturated*" to each data in the dataset. This labeling is simply done by the delay values where 100ms delay is considered as the saturation point. Then using the set of vectors ($t(c_{\mathrm{inf}})$, $s(c_{\mathrm{inf}})$, $t(c_{\mathrm{cur}})$) with labeled delay or frame delivery ratio as a training dataset, we finally obtained a state classifier with different channel distance from 0 to 3, which is directly used as *sat*.

Secondly, we have applied multiple regression analysis to determine all the unknown parameters (u_i and v_j, $0 \leq i \leq 4$ and $0 \leq j \leq 7$) to model the performance in the case of saturation. In order to obtain the values of these parameters, we extracted the vectors labeled with "*saturated*" from the original dataset. We applied multiple regression analysis for this subset where delay and frame delivery ratio are groundtruth and the unknown parameters are explanatory variables. Parameters u_i and v_j obtained by multiple regression analysis are summarized in Table 4.2. The value 0 in the table indicates that there is no interference effect by that term. This decision is based on the indicator AIC (Akaike's Information Criterion) which is used for model selection on multiple regression analysis.

Table 4.2: Regression analysis results for f_{D} and f_{T}

(a) Coefficients of f_{D} and f_{T} (Saturated, $c(c_{\mathrm{new}}, c_{\mathrm{inf}}) == 0$)

Model	coefficients				
	u_0	u_1	u_2	u_3	u_4
u^d	10.08839	11.33052	-6.43820	-0.200706	-9.13417
u^t	-0.091064	-1.581287	0.489509	0.109054	0.712960

(b) Coefficients of f_{D} (Saturated, $c(c_{\mathrm{new}}, c_{\mathrm{inf}}) > 0$)

Channel Distance	coefficients							
	v_0	v_1	v_2	v_3	v_4	v_5	v_6	v_7
1	5.1669	-12.7752	-9.9034	-2.9089	-33.8512	6.3304	1.8806	0
2	3.809	-23.179	-5.935	-1.185	48.670	2.096	10.822	-13.644
3	-5.232	-23.425	11.473	7.862	38.979	-14.945	14.505	-14.738

(c) Coefficients of f_{T} (Saturated, $c(c_{\mathrm{new}}, c_{\mathrm{inf}}) > 0$)

Channel Distance	coefficients							
	v_0	v_1	v_2	v_3	v_4	v_5	v_6	v_7
1	0.98471	0.37795	0.13484	-0.17870	-1.88417	-0.49107	0	0
2	1.42418	0.56237	-0.70279	-0.66546	-2.00872	0.41862	0	0
3	2.35717	0.85577	-2.19927	-1.69353	-1.95952	2.12111	-0.17065	0

4.3.3 Modeling the Effect from Multiple Adjacent Channels

Our prediction function proposed so far can grasp the interference state from a single channel and predict the communication quality. However, considering the application of the proposed method in the scenario imitating the real world, it is necessary to comprehensively capture the influence from adjacent channels occurring due to frequency overlap in the allocation of 2.4 GHz band. On the other hand, according to the design policy of the proposed method, if we attempt to build an interference dataset in which setting parameters are changing step by step in order to consider the influence from multiple channels, the number of simulation scenarios increases. The total number of combinations explosively increases, which is not realistic. Therefore, in the proposed method, we attempt to quantify the influence from multiple channels by a combination of predicted values for each single channel.

In order to quantify the multi-channel effect, we prepared simulation scenarios in which each two interference pair operates on two channels one by one like Figure 4.4. Each detailed simulation setting is shown in the Table 4.3. We put three APs and their corresponding STAs in 400m×400m. The distance between each AP and its corresponding STA is fixed 10m and every node is static. The distance between the target pair and interference pairs are configured fixed 100m like Figure 4.4. Every node follows IEEE 802.11g standards in 2.4 GHz band. STAs follows the channel setting of the associated APs, and the channel settings of APs is set for each scenario. All nodes send packets with the transmission power 20dBm with the modulation BPSK 3/4 in which the maximum data rate is 9Mbps. In these scenario, the traffic is assumed as uplink communication and the traffic demands is configured the fixed value. The traffic demands of each AP's application are set by *iperf-udp-data-rate*.

To quantify the effect from multiple adjacent channels, interference pair 1 uses the channel number

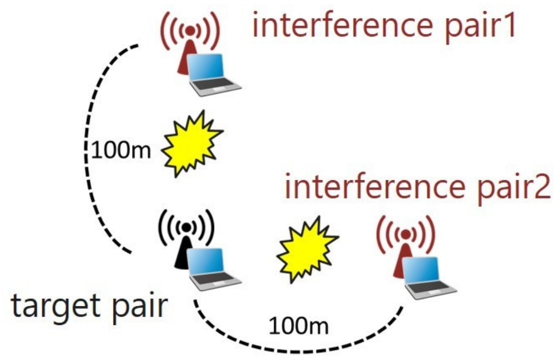


Figure 4.4: Multi-channel scenario about f_{D} and f_{T}

Table 4.3: Simulation settings for multi-channel scenario about f_{D} and f_{T}

Parameter	Values
Area Size	400m×400m
Interference APs - Target AP Distance	100m
AP - STA Distance	10m
Wireless Standards	IEEE802.11g
Wireless Band	2.4GHz
Channel of target AP	$c_{\mathrm{new}} = 6$
Channel of Interference 1	$c_{\mathrm{inf}} = 6$
Channel of Interference 2	$c_{\mathrm{inf}} \in \{7, 8, 9\}$
Transmission Power	20dBm
IEEE802.11g Data Rate	9Mbps (BPSK 3/4)
Antenna Height	1.5m
$T_{preample}$	20μs
Propagation Model	Free Space
L7 application	*iperf-udp-data-rate*
L7 traffic (Interference APs)	3Mbps
L7 traffic (Target AP)	9Mbps
Payload Size	1470byte

6 which is the same as the c_{new}. On the other hand, the channel of interference pair 2 is configured for each scenario. Its channel is selected from the number 7 to 9, which means that the channel distance is from 1 to 3.

4.3.4 Determining Multiple Effect Model

The delay time observed at the target AP is shown the blue boxes in Figure 4.5 as a simulation result. It can be confirmed that the effect of the closer channel distance is strongly received in the case of the influence from multiple channels. This means that the simulated delay value is the highest in the most nearest channel distance case ($c(c_{\mathrm{new}}, c_{\mathrm{inf1}}) = 0$ and $c(c_{\mathrm{new}}, c_{\mathrm{inf2}}) = 1$). Moreover, we can confirm that the influence of the father channel distance is somewhat, but weak influence occurs as the inter-

Table 4.4: Correlation result in multi-channel scenario with function f_{D} output and simulated delay value

Channel Distance	Function output	Groundtruth
(0,1)	1.417	0.968
(0,2)	1.125	0.800
(0,3)	0.879	0.584
(0,none)	0.671	0.554
Correlation		0.979

channel distance increase. This means that the simulated delay value in the case of $c(c_{\mathrm{new}}, c_{\mathrm{inf1}}) = 0$ and $c(c_{\mathrm{new}}, c_{\mathrm{inf2}}) = 3$ is lower than in the case of $c(c_{\mathrm{new}}, c_{\mathrm{inf1}}) = 0$ and $c(c_{\mathrm{new}}, c_{\mathrm{inf2}}) = 1$, but the simulated delay value in the case of $c(c_{\mathrm{new}}, c_{\mathrm{inf1}}) = 0$ and $c(c_{\mathrm{new}}, c_{\mathrm{inf2}}) = 3$ is slightly higher than in the case of $c(c_{\mathrm{new}}, c_{\mathrm{inf1}}) = 0$ and $c(c_{\mathrm{new}}, c_{\mathrm{inf2}}) = none$.

From these findings, we designed the way to aggregate the multiple interference effect indicators f_{D} or f_{T}. When c_{new} is assumed to be affected by multiple channels, the target AP is strongly influenced by the nearer one. Then we quantify by the weighted sum inversely proportional with the weight determined based on the channel distance between channel c_{new} and the interference channel c_{inf}, and the function output for each single channel as following equation 4.7 and 4.8.

$$score_{f_{\mathrm{D}}}(\mathrm{c}_{\mathrm{new}}) = \sum_{i \in AC(c_{\mathrm{new}})} w(c_{\mathrm{new}}, i) \cdot f_{\mathrm{D}} \tag{4.7}$$

$$score_{f_{\mathrm{T}}}(\mathrm{c}_{\mathrm{new}}) = \sum_{i \in AC(c_{\mathrm{new}})} w(c_{\mathrm{new}}, i) \cdot f_{\mathrm{T}} \tag{4.8}$$

We designed the two kinds of the weight function based on the channel distance as following equation 4.9 and 4.10. The first equation 4.9 is the simple reciprocal of the channel distance. The second equation 4.10 is the square of the reciprocal of the channel distance. In these equation, 1 is added so as to be equal magnitude influence when the channel distance is equal to 0. By comparing the error between the function output and the simulated channel performance, we decided to use the equation 4.10

$$w(c_{\mathrm{new}}, c_{\mathrm{inf}}) = \frac{1}{c(c_{\mathrm{new}}, c_{\mathrm{inf}}) + 1} \tag{4.9}$$

$$w(c_{\mathrm{new}}, c_{\mathrm{inf}}) = \frac{1}{(c(c_{\mathrm{new}}, c_{\mathrm{inf}}) + 1)^2} \tag{4.10}$$

In this simulation experiment, the correlation coefficient between $score_{f_{\mathrm{D}}}$ and the simulated delay value is 0.97 as shown in Table 4.4. This result shows that our method can sufficiently follow the effect from multiple channels.

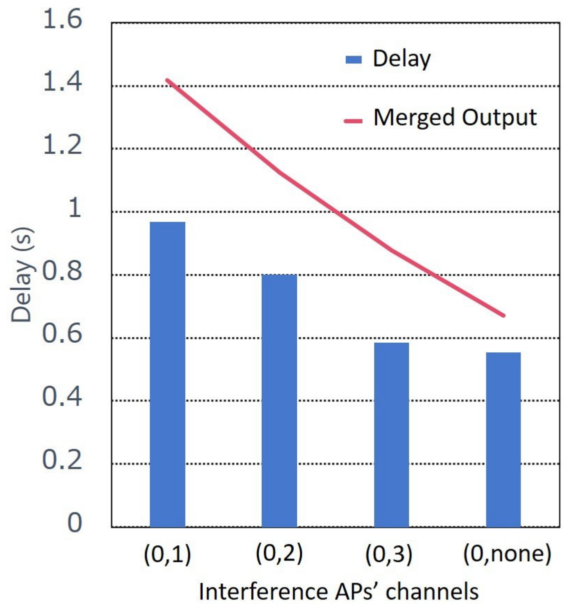


Figure 4.5: Quantify the multiple channel effects about f_{D}

4.4 Estimator Validation

In this section, we show the capabilities of our functions f_{D} and f_{T} that estimate the delay and frame delivery ratio.

4.4.1 Validation Dataset

In order to validate our proposed performance estimator, we divided the interference dataset into 2 groups. One group is used for the training and the other group is used for this validation. We picked up 2,592 scenarios for testing from the dataset randomly which are 20% of all simulation cases. The remaining 80% cases is used for training the classifier *sat* and determining the coefficients by applying multiple regression analysis.

4.4.2 Validation Result

Classifier *sat*

Firstly, we show the classification capability of *sat* classifier by applying *sat* to the test dataset. The classification result is shown in Table 4.5. We show four subtables (confusion matrices) categorized by channel distance $c(c_{\mathrm{new}}, c_{\mathrm{inf}})$ =0, 1, 2 or 3. We note that the rows show the groundtruth and the columns show *sat* outputs. From these results, the accuracy is quite high. Even in the worst case ($c(c_{\mathrm{new}}, c_{\mathrm{inf}}) = 1$), the accuracy is 97.685%. We found that the classification is rather false positive. The average accuracy was 98.34%, and their false positive and false negative are negligible values.

Table 4.5: *sat* classifier output with different channel distance $|c_{\text{new}} - c_{\text{inf}}|$

(a) $|c_{\text{new}} - c_{\text{inf}}| = 0$

channel	distance 0	classifier output *saturated*	classifier output *unsaturated*
groundtruth	*saturated*	371	0
	unsaturated	5	272

(b) $|c_{\text{new}} - c_{\text{inf}}| = 1$

channel	distance 1	classifier output *saturated*	classifier output *unsaturated*
groundtruth	*saturated*	346	0
	unsaturated	15	287

(c) $|c_{\text{new}} - c_{\text{inf}}| = 2$

channel	distance 2	classifier output *saturated*	classifier output *unsaturated*
groundtruth	*saturated*	314	1
	unsaturated	12	321

(d) $|c_{\text{new}} - c_{\text{inf}}| = 3$

channel	distance 3	classifier output *saturated*	classifier output *unsaturated*
groundtruth	*saturated*	265	0
	unsaturated	10	373

Multiple Regression Function

Secondly, we have examined the accuracy of regression functions by using the test dataset. However, we remove some scenario manually to test the regression function because the test dataset contains "unsaturated" condition and the function does not work in this situation. As we have used regression analysis, we may directly refer to the coefficient of determination (R^2) to validate the fitting to the values. The adjusted R^2 value is an index which shows the explanation accuracy of the linear expression constructed by the regression analysis. As this value is closer to 1, it is judged that the tendency of the objective variable is better grasped. We show the results in Table 4.6. As seen in the table, in all the cases, the model well captures the delay and frame delivery ratio behavior as they are almost close to 0.8 or larger. In particular, the frame delivery ratio with $c(c_{\text{new}}, c_{\text{inf}}) = 0$ is the best case where 0.98 is achieved. On the other hand, the model does not capture the trend of delay and frame delivery ratio with $c(c_{\text{new}}, c_{\text{inf}}) = 1$. It is difficult to grasp the inter-channel interference effect by only MAC frame monitoring. Even so, R^2 value of f_{T} with $c(c_{\text{new}}, c_{\text{inf}}) = 1$ is over 0.75.

Finally, we have shown the accuracy of the delay and frame delivery ratio estimation results. For visualization purpose, we have also shown the graphs in Figs. 4.6 and 4.7 that show both groundtruth and the corresponding estimation results where the plots are sorted by the groundtruth values. The visualized scenario is limited by the channel state which is labeled "saturated". We can confirmed that

Table 4.6: Coefficients of determination of f_{D} and f_{T} with different channel distance $|c_{\mathrm{new}} - c_{\mathrm{inf}}|$

Channel Distance	Co. of Det. (R^2)	
	f_{D}	f_{T}
0	0.8215	0.9815
1	0.7906	0.7520
2	0.8463	0.8161
3	0.8306	0.8796

Table 4.7: Average mean square errors of f_{D} and f_{T} with different channel distance $|c_{\mathrm{new}} - c_{\mathrm{inf}}|$

Channel Distance	Ave. Mean Square Errors	
	f_{D} (sec.)	f_{T} (ratio)
0	0.1759	0.0178
1	1.1545	0.1055
2	0.9894	0.0897
3	1.1358	0.0879

the most accurate case is the channel distance 0 and the estimator captures the groundtruth well. We have also summarized the average mean square errors in Table 4.7. For other cases, the error of delay estimation is about 1 second and the error of frame delivery ratio estimation is 10% or less. These values are quite reasonable considering the fact that we only use MAC frame passive observation and this is a lightweight estimation function that can easily be implemented on any APs as a value-added function.

4.5 Evaluation of Estimator

In this section, we show the evaluation result of our proposed estimator in more general scenario.

4.5.1 Evaluation Scenario Setting

In order to evaluate our proposed estimator in more general scenario, we have prepared the evaluation environment in the simulator. Each detailed simulation setting is summarized in the Table 4.8. From our simple site survey result, we confirmed that there are several APs in any channels in urban environment. Therefore we have prepared a representative interference AP-STA pair in each channel, and these 13 pairs are deployed in a position away about 100m from the target AP-STA pair in the simulation area. Like Figure 4.8, the topology of representative interference pairs is like a circle in which the target AP and its STA are the center of the circle. Each STA is deployed 5m away from its AP and all nodes are static. Every node follows IEEE 802.11g standards in 2.4 GHz band. All nodes send packets with the transmission power 20dBm with the modulation BPSK 3/4 in which the maximum data rate is 9Mbps. In these scenario, the traffic is assumed as uplink communication and the traffic demands is configured the fixed value. The traffic demands of each AP's application are set

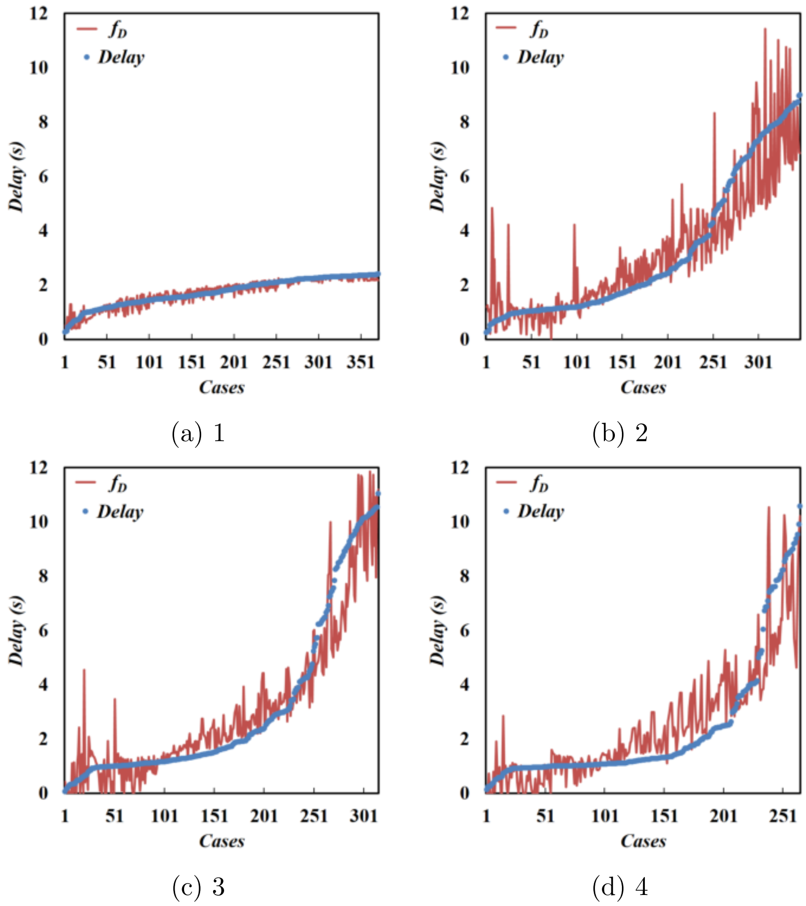


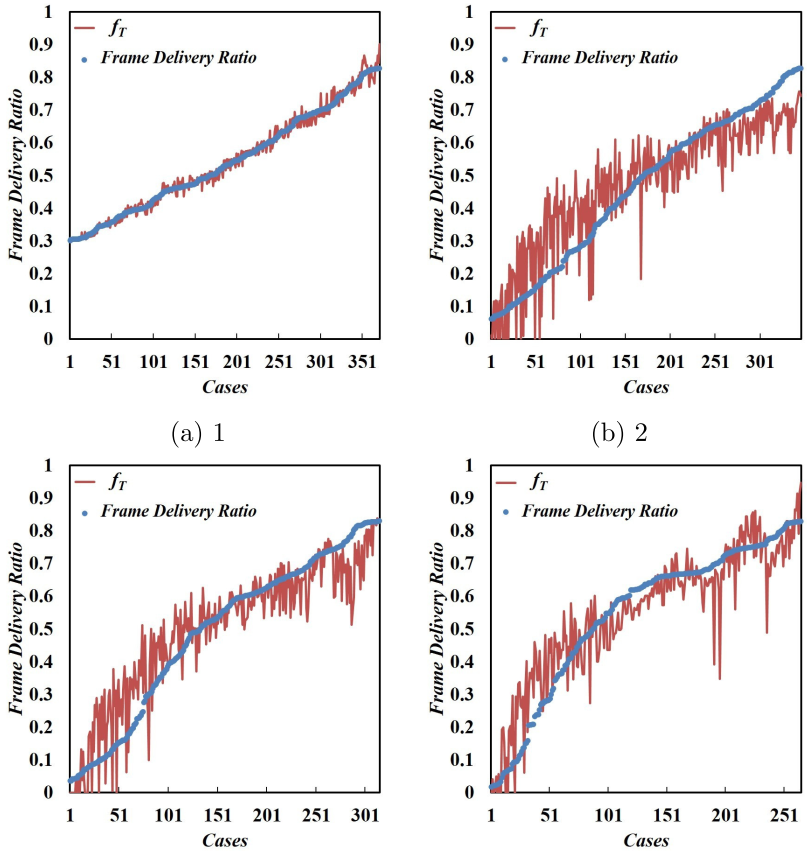
Figure 4.6: f_{D} (delay) output and groundtruth with different channel distance $|c_{\mathrm{new}} - c_{\mathrm{inf}}|$

by *iperf-udp-data-rate*. Their traffic configurations are adjusted to IPERF 5 Mbps (channel 1, 6, 11), 1 Mbps (channel 3, 4, 13) and 3 Mbps (others). This traffic volume setting is designed based on the AP channel distribution. In general, channel 1, 6 and 11 are often used for avoiding the inter-channel interference by the network manager. Then, higher traffic volume is set in channel 1, 6, 11 and, on the contrary, the setting of channel 3, 4, 11 is lower than others. The target AP-STA pair communicates from channel 1 to 13 with IPERF 4.5Mbps. In this evaluation, we compare the estimator output and groundtruth and confirm that the estimator can capture the best channel.

4.5.2 Evaluation Result

At first, we plotted the function output and the simulated delay values in Figure 4.9. The green boxes shows the simulated delay values and the blue line shows the function output. We can confirmed that our estimator can capture the overall tendency of the simulated delay values from this figure.

Secondly, table 4.9 shows the simulated delay value (groundtruth) of target pair and the estimator



(a) 1 (b) 2

(c) 3 (d) 4

Figure 4.7: f_{T} (frame delivery ratio) output and groundtruth with different channel distance $c(c_{\mathrm{new}}, c_{\mathrm{inf}})$

output in each channel. From this table, we can confirmed that the target pair should select the channel number 13 based the simulated delay value, which is the lowest value in all 13 channels. That is, we want to confirm whether our proposed estimator can choose channel number 13. The outputs of our proposed estimator are also shown in Table 4.9. Comparing the function output of each channel, the lowest value is the one of channel number 13. Therefore we have confirmed that our estimator imply the best channel is 13. We also confirmed the correlation coefficient is above 0.85 and our estimator can capture the tendency of groundtruth. From this simulation result, the delay values of the worst channel 9 was about 3.7 times the value of the best one. It is worth noting that our channel selection based on the interference prediction shows significant improvement of communication quality.

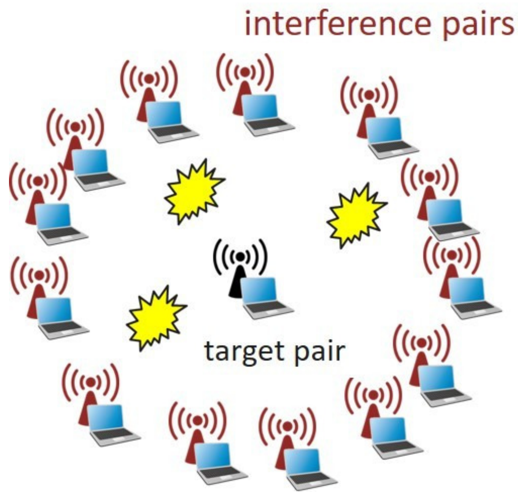


Figure 4.8: Evaluation Scenario

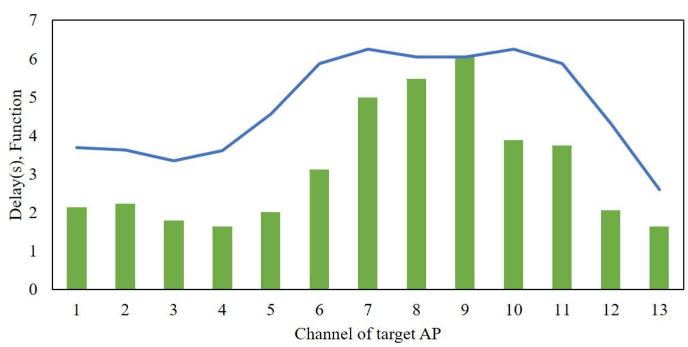


Figure 4.9: Evaluation result

4.6 Conclusion

This chapter presents a channel migration technology that can control congestion among Wi-Fi channels based on the concept of interference environment sensing. Since our previous function is to "rank" the current channel status (*i.e.* it is a "diagnosis" function), it is difficult to estimate the channel status *after* an AP actually moves from the current channel into that channel to expect quality improvement. As a result, the AP would make an erroneous channel selection. In order to let APs (some of them may be selfish) behave in more intelligent ways and to pursue stability in such an autonomous and uncoordinated Wi-Fi environment, it is quite essential to provide such an estimation function for APs to correctly estimate the status without actually moving into it.

Given the information about the other APs' traffic (volume, signal strength and the channel) and

Table 4.8: Simulation settings for evaluation scenario

Parameter	Values
Area Size	400m×400m
Interference APs - Target AP Distance	[50m, 150m] (uniform distribution)
AP - STA Distance	5m
Wireless Standards	IEEE802.11g
Wireless Band	2.4GHz
Channel of target AP	to be determined
Channel of Interference	selected one by one from 1 to 13
Transmission Power	20dBm
IEEE802.11g Data Rate	9Mbps (BPSK 3/4)
Antenna Height	1.5m
$T_{preample}$	20μs
Propagation Model	Free Space
L7 application	*iperf-udp-data-rate*
L7 traffic (Interference APs)	determined based on the AP density
L7 traffic (Target AP)	4.5Mbps
Payload Size	1470byte

its own traffic (volumes) that may be moved to a target channel c, the function predicts capacity saturation in channel c caused by this movement without actual channel migration. In the case of channel saturation, our proposed function estimates the expected performance under the capacity saturation. In order to build our proposed function, we decided to use the network simulator to obtain the interference dataset for understanding the relationship between the observed parameters and interference effect. We have prepared more than 10,000 scenarios and conducted simulations which are assumed that the own traffic of the target AP moves to the new channel to simulate "channel state change". We have investigated the simulation dataset to capture not only the interference effect but also the channel state change due to the channel migration of the target AP. As a result, we decided to apply machine learning based classification algorithm to estimate the channel saturation due to the traffic movement in channel migration. In addition, we applied multiple regression analysis to build a prediction function of channel performance (latency and data delivery ratio).

To validate the model accuracy, we have conducted simulation experiments with different realistic scenarios. For this purpose, we have conducted additional 2,000 simulations. As a result, we confirmed that the classifier *sat* can determine the accurate channel state without actual channel migration. Also, our proposed function can estimate the frame delivery ratio with less than 10% error in average.

Finally, we demonstrate that the proposed function can be used in more general scenario for channel selection at the target AP. To evaluate our function, we prepared the evaluation environment where the representative interference AP in each channel interferes with the target AP. We designed the traffic parameter based on the density of APs in each channel. In this scenario, we found that our channel selection function can predict the best channel and the performance trend of overall 13 channels. In addition, we confirmed the correlation coefficient between our estimator output and the groundtruth

Table 4.9: Correlation result in evaluation scenario with function output and simulated delay value (groundtruth)

Channel	Function output	Groundtruth
1	3.684	2.133
2	3.632	2.224
3	3.352	1.788
4	3.606	1.634
5	4.569	2.009
6	5.869	3.125
7	6.245	4.993
8	6.043	5.484
9	6.043	6.082
10	6.246	3.889
11	5.868	3.743
12	4.315	2.063
13	2.592	1.635
Correlation		0.853

is above 0.85 and our estimator can capture the tendency of overall channel performance.

We note that some contents in this chapter refer our previous publications [106, 107].

Chapter 5

Performance Analysis of Prediction-based Channel Management in Realistic Urban Scenario

5.1 Introduction

We have initially proposed a method to quantify the interference effect by multiple regression analysis with simple and reasonable MAC frame monitoring in chapter 3. In that literature, in order to consider the inter-channel interference problem and the diversity of traffic and RSSI, we have designed an indicator that quantifies the channel interference using a dataset generated by precise simulations. Then, in the following chapter 4, we have improved our initial method to consider the channel migration of the target AP. This algorithm can estimate the real delay and frame delivery ratio values, which can be used to support decision making by access points in migrating to other channels for better performance.

However, the performance analysis of our proposed method was not examined enough. In this chapter, in order to evaluate our method in more realistic urban situations, we modeled the urban environment in the simulator based on the real site survey results in Osaka. As a result, we have built three realistic urban scenarios which was simulating Osaka downtown.

For this purpose, firstly, we should understand the current situations of urban Wi-Fi environment. We focused on the traffic conditions and Wi-Fi AP deployment. To understand the traffic condition in such environment, we conducted traffic monitoring around Osaka station. Figure 5.1 shows the congestion of Wi-Fi APs around Osaka station. As you can see, the congestion is quite high and we regarded this environment as a typical urban situation in our research. To build the Wi-Fi traffic model in an urban area, we conducted the monitoring at 10 locations including shopping malls, cafes,

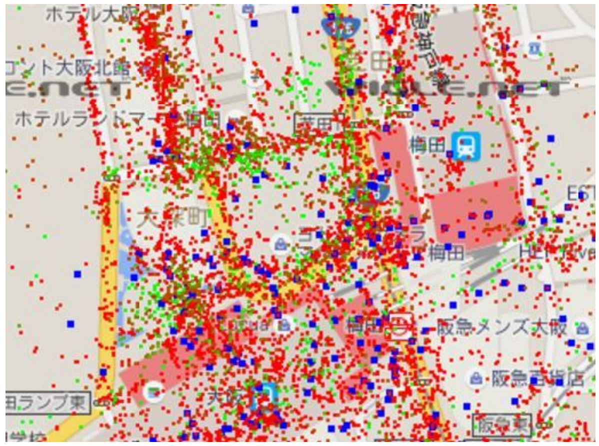


Figure 5.1: Congestoin of Wi-Fi APs around Osaka station [1]

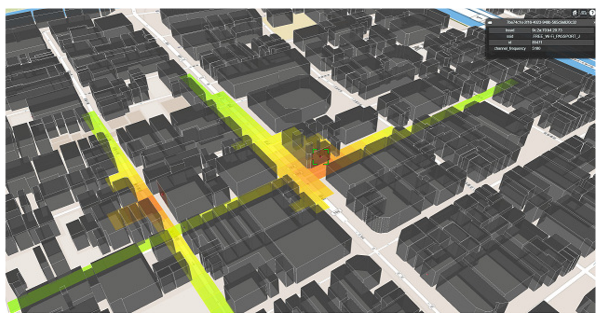


Figure 5.2: Wi-Fi Radio Map of Osaka City

commercial buildings and stations on both weekdays and holidays. This monitoring was carried out in each channel for five minutes, and we got the number of APs and STAs from the source and destination MAC addresses of Beacon frames and Probe Response frames. Then we analyzed this actual measurement and added more traffic considering the future growth of Wi-Fi traffic. This future traffic prediction is determined based on the distribution of the number of APs in each channel.

In addition, in order to understand the current Wi-Fi AP deployment, we used *Wi-Fi Radio Map of Osaka City* which was provided by our Wi-Fi bigdata project [2, 3], as shown in Figure 5.2. This dataset provided the estimated locations of the actual APs, the actual channel usage information and the radio map. We used this dataset to obtain the actual APs' location and their channel usage.

Finally, we demonstrate that our proposed function can be used in this urban model for channel selection at each AP. Through the simulation experiments with such real data in real geography model, we show that our function can identify the best channels and APs can migrate to them accordingly. Moreover, we compared our proposed method with naive channel selection approaches and confirmed the effectiveness of channel management in such a realistic environment. As a result, the throughput of the target AP can be 1.83 times higher than that of the AP which randomly selects channels.

5.2 Wi-Fi Channel Selection Strategy

In this section, we present how to conduct the channel selection based on our proposed channel performance estimator. We firstly describe the overview of our proposed AP's architecture with our method. Then, we introduce prediction-based channel selection strategy.

5.2.1 Overview

We assumes the situation in which the IEEE82.11g AP of interest is deployed at the major intersection in urban environment as shown in Figure 5.3. The target AP is affected by interference from surrounding APs and their STAs. In addition, the target AP is exposed to inter-channel interference because the channel frequency in 2.4 GHz is partially overlapped. In order to let the target AP behave in more intelligent and autonomous ways and to pursue stability in an uncoordinated Wi-Fi environment, the target AP has the two kinds of functions to estimate the best channel for autonomous channel management. The first function is MAC frame monitoring function to understand the interference environment around the target AP. In this step, the target AP monitors MAC frames in each channel by using multiple Wi-Fi interfaces for monitoring. Basically the target AP gathers the received signal strength (RSSI), the length and the transmitted data rate of observed MAC frames in each channel. RSSI is used for understanding the interference signal strength, and the frame information is for understanding the interference traffic volume. These information are summarized for each channel. Then, the summarized parameters are input of the second performance prediction function. The performance prediction function is constructed in two steps. In the first step, the target AP judges whether channel saturation occurs when the target AP conducts channel migration. When the channel migration is conducted, the traffic related with the target AP moves together into the new channel. Since the brought traffic may causes huge traffic saturation, we introduced the binary classifier to our channel selection strategy. When the classifier labels the channel condition as "unsaturated", the performance prediction function judges that the channel migration may not cause the traffic saturation. In the other case, the function applies regression analysis-based performance prediction to understand how severe the saturation is. This two-step prediction function is trained by using a large simulation dataset beforehand. The simulation dataset is designed to understand the relationship between the interference effect and observed parameters. Finally, the target AP estimates the best channel by comparing the function output which represents the interference effect.

5.2.2 Prediction-based Channel Selection

We let AP denote an IEEE802.11g AP of interest (called *target* AP) and ST denote each Wi-Fi device associated with AP. We assume that AP has commodity Wi-Fi chipsets for traffic monitoring, which

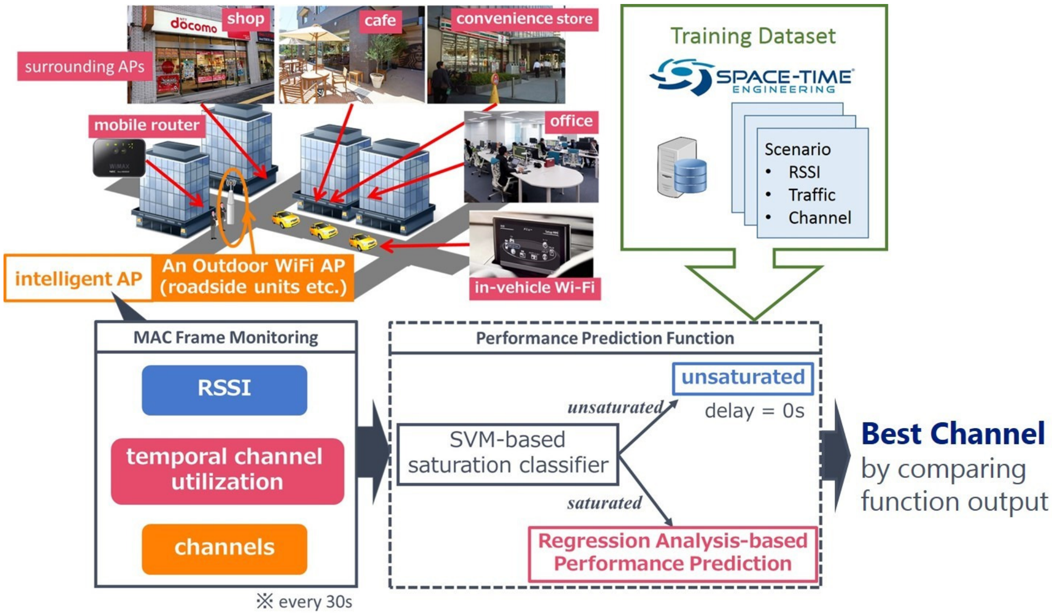


Figure 5.3: Overview of prediction-based channel selection

are not used for the communication. Suppose AP and its corresponding ST currently use channel c_{cur} in the Wi-Fi channel set (denoted as C) and monitors the traffic of other AP(s) (interference AP(s)) that use the same or other channels. In particular, AP calculates the following two values for each channel k; (i) temporal channel utilization ratio (or simply channel utilization ratio) denoted as $t(k)$, and (ii) the received signal strength denoted as $s(k)$. Obtaining these values does not require driver modification, which is a preferable feature for easy installation and operation. For instance, these values such as RSSI, frame length, and transmission rate can be acquired from Atheros chipset in "monitor mode". We note that in order to obtain other information like the number of transmitters, AP needs to maintain a unique address group, which incurs high calculation cost. Instead, our method just needs to update a vector that consists of captured time, frame length, Tx rate and RSSI.

We let $F(k)$ denote the set of observed frames in a certain observation time window W for channel k. We also let $len(f)$ and $dr(f)$ denote the length (bit) of an observed frame f and the transmission rate of IEEE802.11g frames, which is one of 6, 9, 12, 18, 24, 36, 48 and 54 Mbps (in case of OFDM PHY), respectively. We regard that $t(k)$ is as the collection of occupation time by frame f and preamble (denoted as $T_{preamble}$) and is defined in Equation (5.1).

$$t(k) = \sum_{f \in F(k)} \left\{ \frac{len(f)}{dr(f)} + T_{preamble} \right\} \tag{5.1}$$

We adopt $T_{preamble} = 20\mu$s based on the preamble duration in the 802.11g ERP-OFDM.

We let $rss(f)$ denote the RSSI value of frame f. $s(k)$ is defined in Equation (5.2).

$$s(k) = \frac{\frac{\sum_{f \in F(k)} rss(f)}{|F(k)|} - \theta_{\text{min}}}{\theta_{\text{max}} - \theta_{\text{min}}} \tag{5.2}$$

$\sum_{f \in F(k)} rss(f)/|F(k)|$ represents the average RSSI value within W. θ_{min} and θ_{max} represent the minimum RSSI threshold of data frame reception (-90dBm in IEEE802.11g) and the expected maximum RSSI (usually -40dBm or around), respectively. θ_{min} is determined based on the preamble threshold value on frame reception.

We define inter-channel distance (simply called *channel distance* hereafter) between channel k and channel l, denoted as $c(k, l)$, by Equation (5.3).

$$c(k, l) = |k - l| \tag{5.3}$$

The channel distance is expressed as the absolute value of the difference between channels.

In 2.4GHz band, two APs are recommended to operate with channel distance of 5 or larger to avoid interference. For example, channels 1, 6 11 are popularly used in many real situations. However, in our preliminary simulation experiment, we confirmed that interference from those with channel distance 4 or larger has little affect on the performance. Therefore, in our method, we regard that those channels with 3 or smaller channel distance from channel c interfere with c and we call them *adjacent channels* of c. $AC(k)$ denotes the set of adjacent channels and is defined in Equation (5.4).

$$AC(k) = \{l | c(k, l) \leq 3\} \tag{5.4}$$

Then our channel selection strategy is given as follows. Given c_{new} to which AP likes to move from c_{cur} and given observations $t(c_{\text{inf}})$, $s(c_{\text{inf}})$ and $t(c_{\text{cur}})$ where c_{inf} is each channel that may affect the performance of c_{new}, we provide two functions f_{D} and f_{T}, which return the expected L2 delay and (normalized) L2 frame delivery ratio (*i.e.* channel utilization ratio), respectively. Having these two estimators, AP can predict the performance when it moves from c_{cur} to c_{new}, just by observing IEEE802.11 MAC frames and their RSSI in channel c_{inf} at AP. Also, taking into consideration the ease of installation and lightweight operation, we only implement the pre-built function on AP and do not conduct resource-consuming operations like online learning.

Table 5.1: Model parameters of channel performance function

(a) Coefficients of f_{D} (Saturated, $c(c_{\mathrm{new}}, c_{\mathrm{inf}}) == 0$)

Model	coefficients				
	u_0	u_1	u_2	u_3	u_4
u^d	10.08839	11.33052	-6.43820	-0.200706	-9.13417

(b) Coefficients of f_{D} (Saturated, $c(c_{\mathrm{new}}, c_{\mathrm{inf}}) > 0$)

Channel Distance	coefficients							
	v_0	v_1	v_2	v_3	v_4	v_5	v_6	v_7
1	5.1669	-12.7752	-9.9034	-2.9089	-33.8512	6.3304	1.8806	0
2	3.809	-23.179	-5.935	-1.185	48.670	2.096	10.822	-13.644
3	-5.232	-23.425	11.473	7.862	38.979	-14.945	14.505	-14.738

Our functions are given in Equation (5.5) and Table 5.1.

$$
\begin{aligned}
&f_{\mathrm{D}}(c_{\mathrm{new}}, c_{\mathrm{inf}}, t(c_{\mathrm{inf}}), s(c_{\mathrm{inf}}), t(c_{\mathrm{cur}})) \\
&= \begin{cases}
0 \quad (|c_{\mathrm{new}} - c_{\mathrm{inf}}| > 3 \textbf{ or } sat = \textit{“unsaturated”}) \\
\\
u_0 + u_1 \log(t(c_{\mathrm{inf}}) + t(c_{\mathrm{cur}})) \\
\ +u_2 \cdot t(c_{\mathrm{inf}}) + u_3 \cdot s(c_{\mathrm{inf}}) + u_4 \cdot t(c_{\mathrm{cur}}) \\
\quad (|c_{\mathrm{new}} - c_{\mathrm{inf}}| = 0 \textbf{ and } sat = \textit{“saturated”}) \\
\\
v_0 + v_1 \cdot t(c_{\mathrm{inf}}) + v_2 \cdot s(c_{\mathrm{inf}}) + v_3 \cdot t(c_{\mathrm{cur}}) \\
\ +v_4 \cdot t(c_{\mathrm{inf}}) \cdot s(c_{\mathrm{inf}}) + v_5 \cdot s(c_{\mathrm{inf}}) \cdot t(c_{\mathrm{cur}}) \\
\ +v_6 \cdot t(c_{\mathrm{inf}}) \cdot t(c_{\mathrm{cur}}) + v_7 \cdot t(c_{\mathrm{inf}}) \cdot s(c_{\mathrm{inf}}) \cdot t(c_{\mathrm{cur}}) \\
\quad (0 < |c_{\mathrm{new}} - c_{\mathrm{inf}}| \leq 3 \textbf{ and } sat = \textit{“saturated”})
\end{cases}
\end{aligned}
\tag{5.5}
$$

When c_{new} is assumed to be affected by multiple channels, we quantify by the weighted sum inversely proportional to $(c(c_{\mathrm{new}}, c_{\mathrm{inf}})+1)^2$ of the function output for each single channel using equation (5.6) and (5.7).

$$
w(c_{\mathrm{new}}, c_{\mathrm{inf}}) = \frac{1}{(c(c_{\mathrm{new}}, c_{\mathrm{inf}}) + 1)^2} \tag{5.6}
$$

$$
score_{f_{\mathrm{D}}}(\mathrm{c_{new}}) = \sum_{i \in AC(c_{\mathrm{new}})} w(c_{\mathrm{new}}, i) \cdot f_{\mathrm{D}} \tag{5.7}
$$

Finally, the target AP estimates and determines the best channel by comparing the function output $score_{f_{\mathrm{D}}}(\mathrm{c_{new}})$ which represents the interference effect.

5.3 Traffic Model Based on Actual Traffic Measurement in Osaka Downtown

In this section, we describe the Wi-Fi traffic measurement result around Osaka station. As mentioned above, this measurement was conducted for understanding the current traffic condition in typical urban environment. We focused on the distribution of APs over 13 channels, the number of APs and STAs, the RSSI, and the traffic volume. Then we analyzed these actual measurement and added more traffic

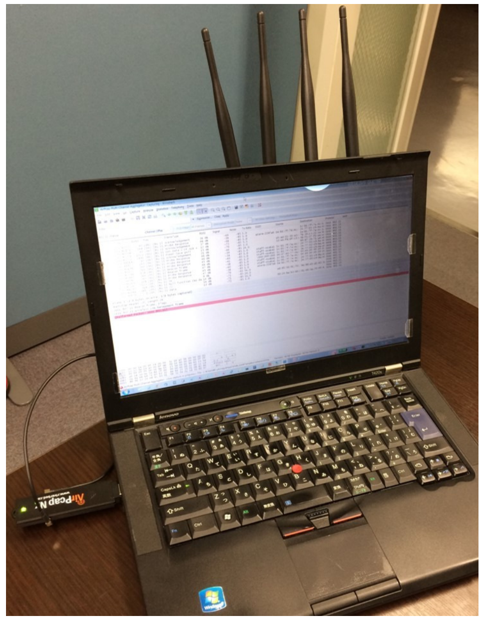


Figure 5.4: *AirPcap Nx* and *Wireshark*

considering the future growth of Wi-Fi traffic. This future traffic prediction is determined based on the distribution of the number of APs in each channel.

5.3.1 MAC Frame Monitoring in Osaka Downtown

For the reproduction of Wi-Fi overcrowded environment in simulation, we conducted frame monitoring in the real environment. As mentioned above, we used a USB wireless LAN protocol analyzer, *AirPcap Nx*, and the analysis software, *Wireshark*, as shown in Figure 5.4. *AirPcap Nx* can capture 802.11 MAC frames including frames that Frame Check Sequence (FCS) is not correct. *Wireshark* can analyze the measurement log files (.pcap files).

Figure 5.5 shows the ten locations where we performed the MAC frame monitoring around Osaka station. As Osaka station is reported the busiest train station in Osaka, these locations can be regarded as typical urban environment. Also, we confirmed the congestion of Wi-Fi APs around Osaka station in Figure 5.1. In our research, we concluded that this location is suitable for our assuming environment. We selected the monitoring locations including stations, outdoor environment, shopping malls, cafes and commercial buildings to find the difference with the location and situation. These locations are shown in Figure 5.5. In Figure 5.5, the green, orange, blue, and black boxes represent the station, shopping malls and commercial buildings, cafes, and outdoor environment, respectively. We surveyed every location on both weekdays and holidays to find the difference between them because the visitors may be different on weekdays and holidays.

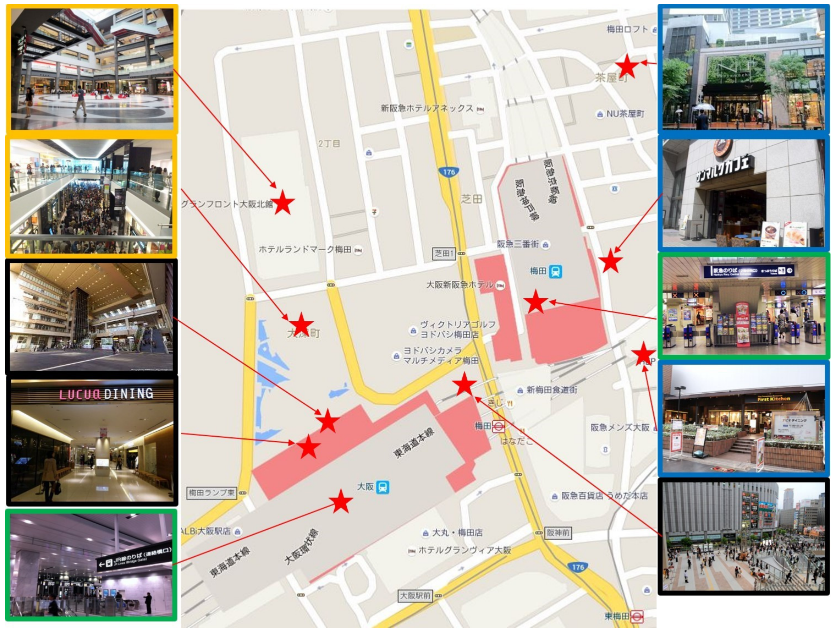


Figure 5.5: Traffic monitoring locations around Osaka station

We have captured the MAC frames that are sent and received at APs and STAs in each channel for five minutes in one location with *AirPcap Nx*. Due to the limited number of *AirPcap Nx*, we could monitor only six channels at the same time. Since 2.4 GHz band has 13 channels, we divided the channels into three groups and we measured the channels in each group simultaneously.

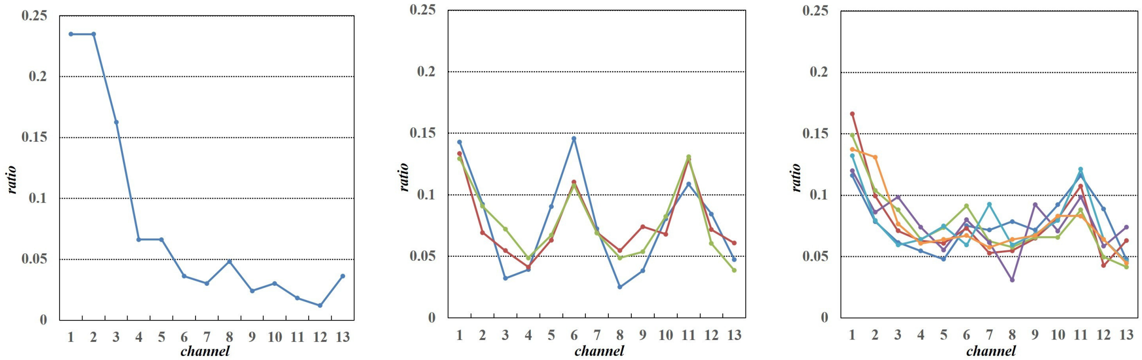
5.3.2 Traffic Modeling by Analyzing Measurement Result

From this observation, we obtain some statical values. As stated earlier, we focused on the distribution of APs over 13 channels, the number of APs and STAs, the RSSI, and the traffic volume of all sensed frames.

AP distribution patterns over 13 channels

At first, we focused on the APs distribution over 13 channels since the concentration of APs to a specific channel directly affects the characteristics of Wi-Fi utilization. In order to obtain the number of APs, we counted the source MAC addresses of Beacon frames and the destination MAC addresses of Probe Response frames. In the same way, to obtain the number of STAs, we counted the destination MAC addresses of Probe Response frames. We summarized the number of APs in each channel to survey the AP distribution.

As stated earlier, it is easily imagined that most APs are configured at either channel 1, 6 or 11 due to the partial channel overlapping in 2.4GHz band. At first, to find such various distribution patterns,



(a) concentrated on 1, 2 (b) concentrated on 1, 6, 11 (c) scattered

Figure 5.6: Normalized AP distribution in each channel and each location

we applied k-means clustering to this dataset. k-means algorithm is a typical clustering method for unsupervised learning. We apply k-means algorithm to the dataset captured on holidays by different k values. As a result, when we use the parameter $k = 3$, we can classify three kinds of distribution patterns clearly. These three kinds of distribution patterns which we found are (a) concentrated on channel 1 and 2, (b) concentrated on channel 1, 6 and 11, (c) scattered as shown in Figure 5.6.

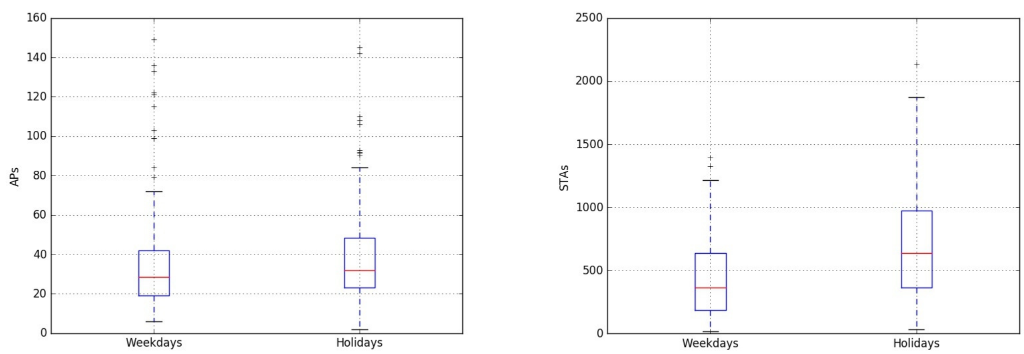
The first pattern, (a) concentrated on channel 1 and 2 in Figure 5.6, can be seen at a shopping mall and shows the ratio of channels 1 and 2 is extremely high. We can consider that this case is the result of easy installation by default settings. The second pattern, (b) concentrated on channel 1, 6 and 11 in Figure 5.6 can be seen at office buildings and shopping malls and shows some peaks at channels 1, 6 and 11. We can consider that the channels of APs are managed so as not to interfere with each other since such a building has a network administrator. The last pattern, (c) scattered in Figure 5.6, can often be seen outdoors and at busy areas. This shows the number of APs in channels 1 and 11 is slightly higher than the others. We consider that this pattern is the most chaotic situation.

Relationship of AP distribution between holidays and weekdays

Secondly, we compared the AP distributions of holidays with the one of weekdays at the same location. We assume that two distributions of holidays and weekdays at the same location have strong correlation since the most APs are considered to be permanently installed. This means that it becomes possible to reduce the frequency of traffic monitoring when the trends are the same. To confirm the relationship, we calculated the correlation coefficients in each location as shown in Table 5.2. We can confirmed that the coefficients are almost above 0.70, except the cafe 3 case. This means the two distributions have high correlation. In addition, we plotted and compared the total sum of APs between in holidays and weekdays in Figure 5.7 (a). We found that the mean values do not really change but the variance change slightly. From this result and for simplicity, we can assume that the distribution of APs in holidays and weekdays are the same.

Table 5.2: Correlation of AP distributions (Holidays and Weekdays)

Location	Correlation
outdoor 1	0.82
outdoor 2	0.78
shopping mall 1	0.75
shopping mall 2	0.91
shopping mall 3	0.73
station 1	0.86
station 2	0.78
cafe 1	0.80
cafe 2	0.77
cafe 3	0.56



(a) APs

(b) STAs

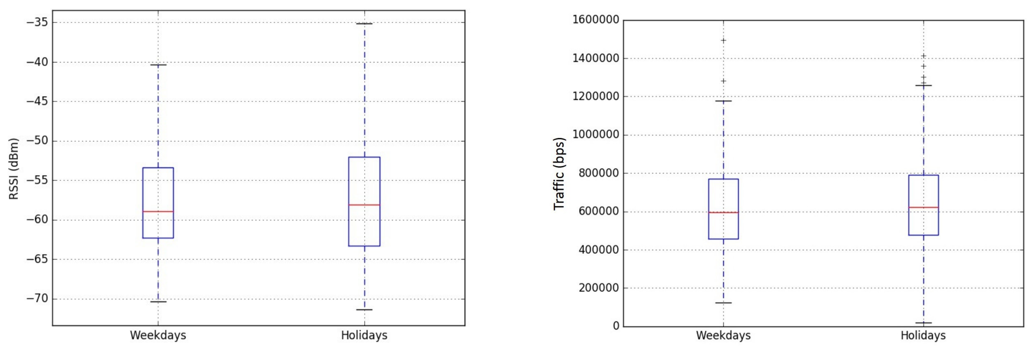
Figure 5.7: Difference of these total volume between weekdays and holidays

Relationship between the number of APs and STAs

Thirdly, we compared the distribution of APs and the distribution of STAs at the same time in the same location. In the same way, we assume that two distributions of APs and STAs at the same time and location have strong correlation since STAs need to associated with an AP to access the internet and the number of STAs is considered to be roughly proportional to the number of APs. To clarify this correlation, we calculated the correlation coefficients at the same time in each location as shown in Table 5.3. As shown in Table 5.3, except the station 1 and cafe 3, coefficients are above 0.7. In addition, we plotted and compared the total sum of STAs between in holidays and weekdays in Figure 5.7 (b). As shown in Figure 5.7 (b), the mean of the number of STAs increases 60% on holidays. This is because more people use their laptops in public areas like cafes on holidays. We need to consider that the traffic gain on holidays this is part of our future work. For simplicity, we also consider that the distribution of STAs is the same as APs.

Table 5.3: Correlation of AP and STA distributions

Location	Correlation
outdoor 1	0.95
outdoor 2	0.92
shopping mall 1	0.91
shopping mall 2	0.85
shopping mall 3	0.93
station 1	0.63
station 2	0.92
cafe 1	0.74
cafe 2	0.72
cafe 3	0.53



(a) RSSI

(b) Traffic

Figure 5.8: Difference of these total volume between weekdays and holidays

Averaged RSSI and the traffic volume of all sensed frames

Finally, we compared the average of RSSI and the traffic volume on holidays and weekdays. The volume of traffic is calculated by averaging the sum of the frame length transmitted and received in each channel, and RSSI is similarly calculated by averageing RSSI of all frames in each channel. This result is shown in Figure 5.8. We found that the mean values do not really change but the variances change slightly. We concluded that we can ignore the effect of day of the week and this knowledge can be used for channel monitoring policy.

5.3.3 Problems of Urban Channel Allocation

By analyzing our measurement results in typical urban environment, we found problems of urban channel allocation.

As seen in Figure 5.6 (c), especially in urban outdoor environments, we can confirm that the uncoordinated AP installation is underway. Therefore, we have investigated the interference problems

Table 5.4: Simulation settings with urban channel allocation

Parameter	Values
Area Size	200m×200m
Coordinates of APs	randomly deployed
AP - STA Distance	5m
Wireless Standards	IEEE802.11g
Wireless Band	2.4GHz
Channel	selected according to channel distribution
Transmission Power	20dBm
IEEE802.11g Data Rate	9Mbps (BPSK 3/4)
Antenna Height	1.5m
$T_{preample}$	20μs
Propagation Model	Free Space
L7 application	*iperf-udp-data-rate*
L7 traffic volume	configured based on actual measurement result
Payload Size	1470byte

with such channel allocation in urban situation. In order to investigate the uncoordinated channel allocation problem in such dense Wi-Fi environment, we prepared the simulation environment which imitated Osaka station area.

We have prepared a 200m×200m field in the simulator, called *Scenargie*, and one hundred APs are deployed randomly in this simulation environment. Each AP has one corresponding STA, and the distance distance between the AP and its STA is 5m. We ignored the mobility of the nodes in this scenario. All node follows IEEE 802.11g standards in 2.4 GHz band and transmits packets with the transmission power 20dBm and the modulation BPSK 3/4 in which the maximum data rate is 9Mbps. We constructed three types of scenario with the AP distributions which are (a), (b) and (c) in Figure 5.6. In these scenario, the channel setting of APs are determined according to the three AP distribution respectively. In addition, the traffic volume of APs are set based on the actual measurement result in each channel. The summary of simulation parameters are shown in Table 5.4.

We show these three simulation results in Figure 5.9. In this figure, we show the averaged delay values of all AP-STA pairs in each scenario, (a) concentrated on channel 1/2, (b) concentrated on channel 1/6/11, and (c) scattered. From this result, the worst delay value is observed in the scenario (a). In (c), since the channel allocation of APs is dispersed throughout, the delay is smaller than (a). This means that the concentration of APs to one channel causes very huge interference and let the overall communication quality low because the interference influences each other. Also, we confirmed that the best channel allocation is (b). In this coordinated scenario, the APs are distributed in channel 1, 6 and 11 to mitigate the interference. These facts indicate the importance to select interference-free channels as the scenario (b) models administrated environments. Our proposed method is designed to avoid interference and the same facts have been confirmed in the following experiment.

We prepared the following scenario in the simulator. We have prepared a representative interference

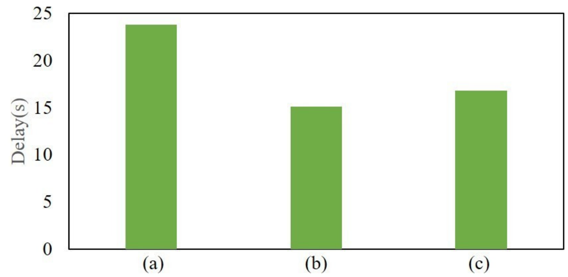
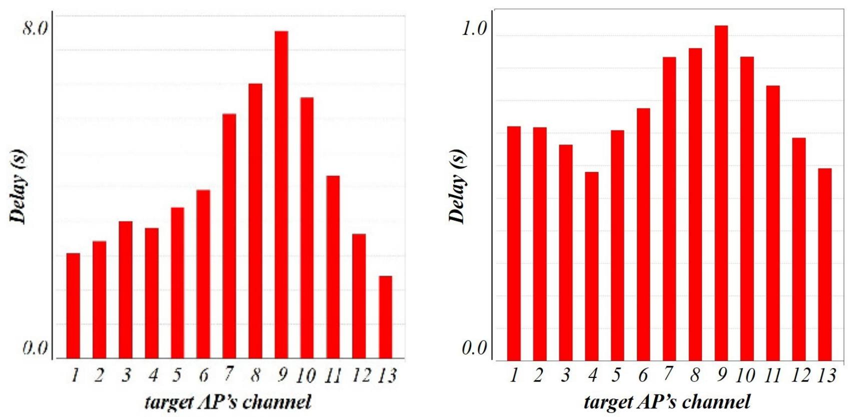


Figure 5.9: Averaged delay values in three scenario ((a) concentrated on channel 1/2, (b) concentrated on channel 1/6/11, (c) scattered)



(a) target AP

(b) all APs

Figure 5.10: Result of delay values in evaluation scenario

AP-STA pair in each channel, and these 13 pairs are deployed in a position away from about 100m from the target AP-STA pair. The topology of representative interference pairs is like a circle in which the target AP and its STA are the center of the circle. Each STA is deployed 5m away from its corresponding AP and we ignored the mobility of all nodes. The traffic demands of each interference AP's iperf application are determined based on the real traffic measurement according to the channel. The traffic setting of the target AP is set to 4.5Mbps. The target AP-STA pair communicates from channel 1 to 13 and we observed the delay value of the target AP and the averaged delay value of all APs. The other simulation settings are the same in the Table 5.4.

Figure 5.10 shows the simulated delay value. The delay value of the target AP is shown in this figure (a), and the averaged delay value of all APs is shown in (b). From these figures, it is found that the target pair should select channel 13. In addition, we can confirm that the overall network quality is the best when the target pair select the channel number 13 as shown in Figure 5.10 (b). We have confirmed that our proposed method can choose the channel 13.

In a realistic scenario constructed in the subsequent sections, it is possible to avoid such problems confirmed in this section by selecting and using a channel with less interference influence by our proposed method.

5.4 Urban Scenario Construction

In this section, we describe how to construct the urban scenario with the traffic model in section 5.3. As stated earlier, we use the actual information about the current positions and channels of existing APs in Osaka downtown. These information is obtained from *Wi-Fi Radio Map of Osaka City* [2] which our research group have constructed [3]. In following section, firstly, we introduce Wi-Fi Radio Map of Osaka City and then we describe the design of the evaluation scenario.

5.4.1 Wi-Fi Radio Map of Osaka City

Our Wi-Fi bigdata project provides Wi-Fi Radio Map of Osaka City, called Wi-Fi radio map hereafter. This Wi-Fi radio map is constructed by using crowdsourcing and radio propagation simulation as shown in Figure 5.11.

In this system, cooperative smartphone users install the observation application which performs Wi-Fi channel scan at regular intervals. The observation application performs this scanning procedure when the user has the smartphone in his/her hands and walks/stops in urban areas. This application collects ESSIDs (SSID texts), BSSIDs (MAC addresses), RSSI, channels and bandwidth (20MHz/40MHz etc.) by beacon advertisement from APs and reports them to the cloud with the timestamp and scanned location from GPS module in the smartphone. By using the collected observations and conducting radio propagation simulation at the cloud server, the "virtual positions" of APs are estimated by the simple range-free 3D localization algorithm using the scanned position and RSSI values. It is difficult to estimate the accuraete location of APs which are indoor based on the observations. Therefore we estimate the virtual AP location, called *"tx tile"*, that represents the virtual radio wave transmission point on the surface of the buildings. In addition, this system provides the 3D radio map as shown in Figure 5.2. Given the virtual AP location and the RSSI observations, the system conducts the radio propagation simulation to build the RSSI radio map of each AP. The radio map contains simulated RSSI values from each AP at any points of outdoor spaces such as major and minor streets, public spaces and parks in urban ares.

We have already conducted large-scale crowdsensing and constructed the radio map in Osaka city [2]. The target area of observation is about $5km^2$ as shown in the following Figure 5.12. The observers holding the smartphone in their hands walked to cover almost all the roads in this area. We got observations covering all areas on three different days. The number of total observation points is 42,202, and the number of observed APs is 78,170. From those observation data, we have constructed the radio map. By excluding common APs on all days, mobile AP are exclued from targets.

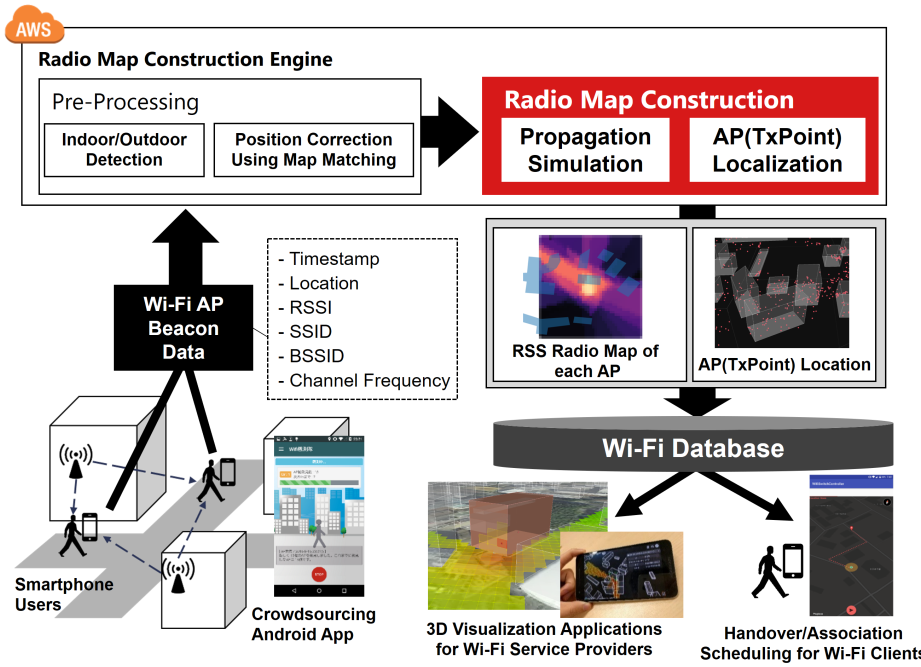


Figure 5.11: Wi-Fi Radio Map of Osaka City [2, 3]

We used the name of APs which is represented as ESSID or BSSID, the channel which the AP uses and the virtual position which are represented 3D coordinates for realistic urban scenario construction. The name and channel data is actual observation, but the position of APs is estimated value. However, the estimated results are quite reasonable because we have predicted the virtual location that reduce contradiction among multiple observed values.

5.4.2 Scenario Setup

In order to confirm that our function can estimate the delay and frame delivery ratio in more realistic scenarios, we have designed several scenarios. We show the overview of this scenario construction in Figure 5.12.

At first, we selected three locations as shown in Figure 5.13. These three targeted locations are about 300m square area. The reason for this selection is that they are crowded urban environments including a large intersection. The network simulator *Scenargie*, which we use in the step of modeling our function, can handle OpenStreetMap (.osm file) and geographical information such as buildings and streets. Then, we can use the tx-tile location from the Wi-Fi radio map directly. From the Wi-Fi radio map, we obtained the 3D coordinates of tx-tiles which represent the virtual AP locations. As a result, we constructed the geographical scenarios as shown in Figure 5.14. The target AP which will conduct channel selection by using our proposed method is deployed at an intersection of streets

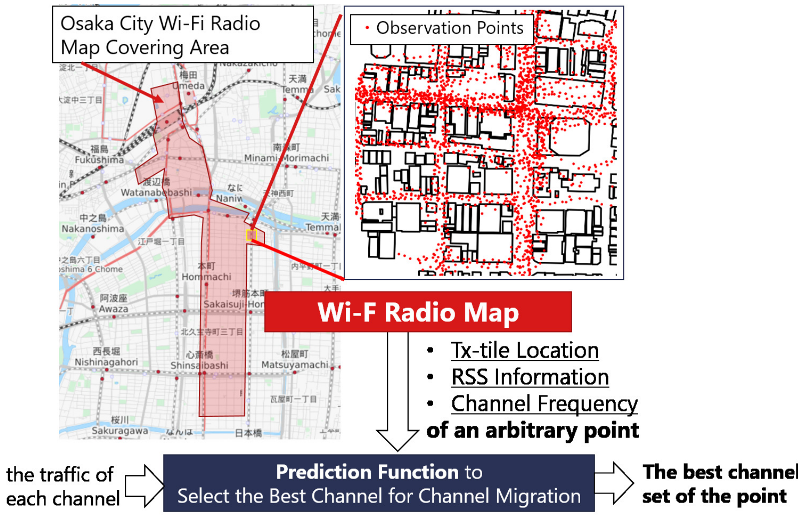


Figure 5.12: Scenario construction overview using Wi-Fi Radio Map of Osaka City

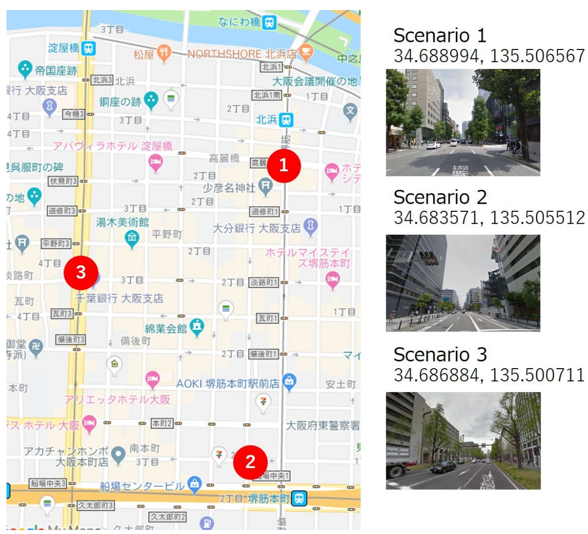
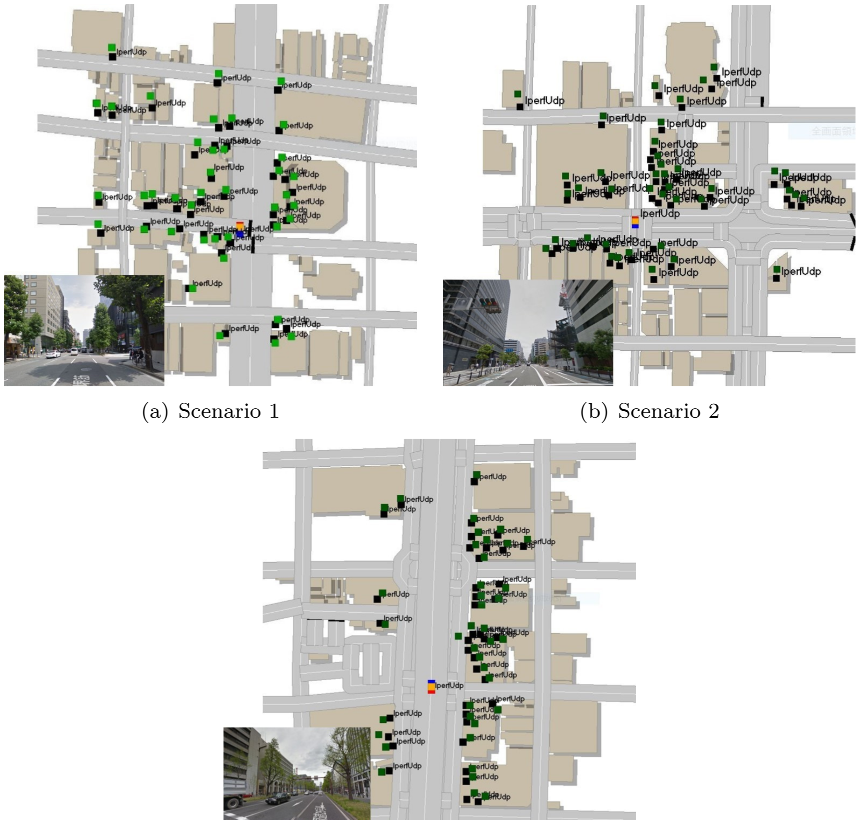


Figure 5.13: Location of evaluation scenario in Osaka City (Scenario 1 - Scenario 3)

assuming a road side unit for intelligent transportation system. The target AP is located at the approximate center in each scenario. In these figures, the red blue and orange nodes represent target APs, their clients and AP's Wi-Fi monitoring devices, respectively. Also, as surrounding interference sources, such APs are represented by black nodes and their client are represented by green nodes. One AP has one corresponding client in our scenario and the client is located 5m away from the AP. In these scenarios, the target AP conducts MAC frame monitoring and applies our proposed function to select the best channel in each environment.

Then, we obtained the channel information of each AP from the Wi-Fi radio map. The number



(a) Scenario 1 (b) Scenario 2

(c) Scenario 3

Figure 5.14: Evaluation scenario based on real urban measurement

of APs in each scenario and in each channel are shown in Table 5.5. We confirmed that APs were concentrated particularly on channel 1 through all the scenarios. Also, channel 6 and 11 are used for a relatively large number of APs. In most cases, at least one AP is running in all channels. However, the target AP may not be exposed to interference from all channels since we take into account attenuation by the walls of buildings in this scenario. Since the Wi-Fi radio map does not have any traffic information, we have determined the traffic parameter of Wi-Fi APs based on the real traffic measurement in urban areas as we described in section 5.3. By analyzing traffic measurement results, the average traffic volume of each channel is set as the application demands parameter *iperf-udp-data-rate*.

The other simulation parameters are shown in Table 5.6. All node follows IEEE 802.11g standards in 2.4 GHz band and transmits packets with the transmission power 20dBm and the modulation BPSK 3/4 in which the maximum data rate is 9Mbps. The target AP send packets with fixed data rate 4.5Mbps in IPERF application. We used the *WallCount* model as the radio wave propagation model. This model is based on free space propagation model attenuating simply according to distance. In this

Table 5.5: The number of APs in each scenario

Scenario	Channel Number												
	1	2	3	4	5	6	7	8	9	10	11	12	13
1	14	1	0	2	0	9	1	1	0	1	8	2	2
2	8	3	2	3	2	5	3	2	1	1	7	1	2
3	8	1	2	3	1	6	1	2	3	1	6	1	5

Table 5.6: Simulation settings for realistic urban scenario

Parameter	Values
Area Size	300m×300m
Coordinates of target AP	center of the area
Coordinates of interference APs	configured based on Wi-Fi Radio Map of Osaka city
AP - STA Distance	5m
Wireless Standards	IEEE802.11g
Wireless Band	2.4GHz
Channels of target AP	to be determined
Channels of interference APs	configured based on Wi-Fi Radio Map of Osaka city
Transmission Power	20dBm
IEEE802.11g Data Rate	9Mbps (BPSK 3/4)
Propagation Model	*WallCount*
Attenuation by wall	5dBm
Antenna Height	1.5m
$T_{preample}$	20μs
L7 application	*iperf-udp-data-rate*
L7 traffic (target AP)	4.5Mbps
L7 traffic (interference AP)	configured based on actual measurement result
Payload Size	1470byte

WallCount model, basically the pathloss value is calculated by the distance between the transmitter and the receiver as the same as free space propagation model. In addition to this, if there is a wall between the transmitter and the receiver, the simulator calculates to attenuate by 5dBm with a fixed value for each wall. We regard each side of the building as one wall in this scenario.

In the following section, we confirmed whether our proposed function can estimate the best channel in these urban scenarios.

5.5 Performance Evaluation in Realistic Urban Scenarios

In this section, we describe the performance evaluation result of our proposed channel selection method in the realistic urban scenarios built in section 5.4. In addition, we show the effectiveness of channel selection with our proposed method by comparing two kinds of different channel selection techniques which are randomly selection and selection based on the number of APs.

5.5.1 Evaluation Result

We conducted performance evaluation experiments of our proposed method with prepared three urban environment scenarios as shown in Figure 5.14. In these scenarios, the target AP and its client communicate while changing from channel 1 to channel 13 under the same condition. We confirm whether or not the channel estimated to be the highest communication quality by relatively comparing the function output matches the best channel determined based on the actual measured value in these simulations.

The evaluation results in three urban realistic scenarios are shown in Figure 5.15. In this figure, the simulated delay values are represented by a bar graph, and the estimator outputs of our proposed method are also represented by a line graph. In each scenario, the best channel in which the delay value of the target system achieves the lowest one is represented in red. By comparing the simulated channel performance values in each channel, the best channel is judged to be channel 9, channel 12, channel 7 respectively. Therefore, we check whether the predicted channel by applying our proposed method matches these channels. From our proposed method, channel 9, channel 12, channel 7 were also judged to be best in each scenario, respectively. It was confirmed that the best channel can be predicted by using the proposed method even in realistic scenarios simulating the urban environment. It is worth noting that our proposed method can estimate the optimal channel even though the best channel is different in each scenario. This shows the adaptability of our proposed method to various environments. In addition, it can be confirmed that the output value of our prediction function follows the trend of the delay values over all 13 channels in each scenario. This information can be used to select a relatively good channel subset. This channel subset is useful in the case of the system which uses multiple channels like MIMO (multiple-input and multiple-output). When the system operates in multiple channels, it can select channels from the channel subset which do not interfere with each other. Even if the channels are partially overlapped, our estimator can predict such inter-channel interference.

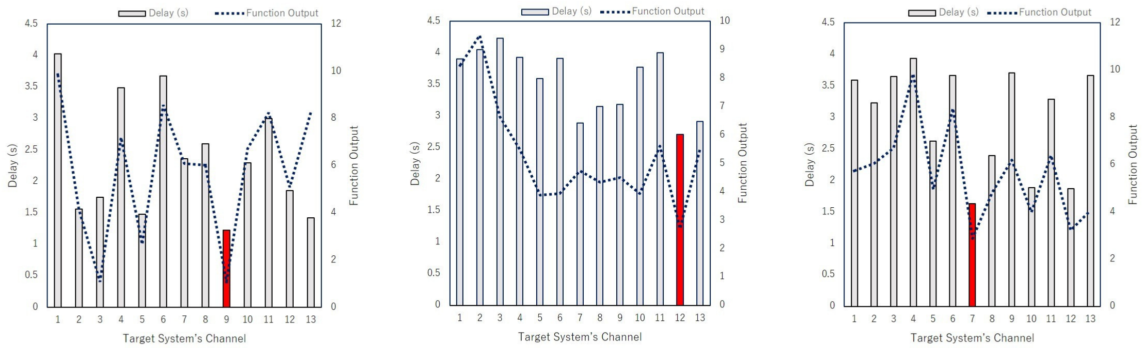
5.5.2 Effectiveness of Channel Selection

Moreover, we have evaluated whether the target AP can achieve the higher throughput by comparing two channel selection methods to confirm the effectiveness of our channel selection approach. This evaluation environment was designed based on scenarios in Figure 5.14 and the simulation setting also follows the parameters as shown in Table 5.6.

We selected the following naive selection approaches.

(1) *Randomly Selection, RS*

In this approach, the target AP selects the channel randomly. We considered this approach as a baseline. In reality, people who are not familiar with IT technology are likely to do this kind of operation. This throughput is calculated as the expected value, which is the averaged throughput



(a) Scenario 1 (b) Scenario 2 (c) Scenario 3

Figure 5.15: Evaluation result in urban scenarios

value of all channels.

(2) *Least Congested Channel Scan, LCCS* [108]
In this approach, the target AP selects the channel based on the number of APs in the same channel of the target one. This LCCS algorithm selects one channel which is the smallest number of competitive Wi-Fi APs to avoid the conflict in the same channel. Therefore, this approach considers a part of interference problems, not including the inter-channel interference problem. The target AP senses beacon frames and counts the unique number of interference APs. Such operation is usually conducted in commodity products. When multiple channels are selected, the average throughput among selected channels is calculated.

(3) *Least Traffic Channel (Single Channel), LTC (SC)*
In this approach, the target AP selects the channel based on the traffic measurement. This LTC (SC) algorithm selects one channel which has the least traffic volume. This approach does not consider the interference from adjacent channels and the signal strength from surrounding nodes.

(4) *Least Traffic Channel (Adjacent Channels), LTC (AC)*
As well as LTC (SC), the target AP selects the channel based on the traffic measurement in this approach. This LTC (AC) algorithm selects one channel which has the least sum of traffic volume in the same and one adjacent channel to consider the inter-channel interference problem.

Table 5.7 shows the throughput result of comparison approaches and our proposed method in each scenario. Except for scenario 1, we can confirm that only our proposed method can select the channel that achieves the best throughput, which shows bold in the table. As shown in Table 5.5, in Scenario 1, the surrounding APs are concentrated on channel 1, 6, 11 which are separated completely. In such scenario, LTC (AC) can select the best quality channel as well as our proposed method. The throughput of the best channel is 1.83 times higher that RS in Scenario 1. Also, the throughput can

Table 5.7: Evaluation result about effectiveness of our proposed approach (throughput (Mbps))

	RS	LCCS [108]	LTC (SC)	LTC (AC)	Proposed
Scenario1	2.35	3.41	3.41	**4.30**	**4.30**
Scenario2	0.85	1.27	0.73	1.32	**1.76**
Scenario3	1.27	1.96	1.11	2.72	**3.21**

achieve 1.26 times higher than LCCS and LTC (SC) by considering the inter-channel interference.

Compared with Scenario 1, in Scenario 2 and 3, the channels which the surrounding APs use are scattered. In these scenarios, we confirmed that comparison methods cannot achieve the best channel because these methods did not consider the relationship among inter-channel distance, RSSI, and traffic volume. The throughput of our proposed approach can achieve 1.33 and 1.18 times higher than LTC (AC) in Scenario 2 and 3 respectively. Therefore, we confirmed the importance to consider inter-channel interference problem for using the channel resources effectively and concluded that it is significant to select interference-free channels with our proposed method.

5.6 Discussion

Channel Usage Trend

As a result of actual traffic measurement using AirPcap shown in Figure 5.6 and in Table 5.5, we confirmed that APs tend to concentrate on a specific channel. Using these normalized AP distributions, we can design probabilistic models of the potential number of transmitters. The number of transmitters in each channel is very useful information to predict channel performance. This is because the channel capacity is affected by this number in CSMA/CA system. On the other hand, it is not desirable to observe it considering the small memory area of low-cost Linux equipment. We believe that it is possible to obtain such information from some Web sites or Osaka Wi-Fi Scan Map as prior knowledge. Especially in the Osaka Wi-Fi radio map, the information about mobiles APs such as mobile routers is omitted, so it is suitable for this use case.

In addition, around some shops like a coffee shop, continuous traffic like video streaming is well detected. In the scenario where fixed APs are installed at intersections, it can be considered that it is possible to acquire surrounding geographical information. For example, AP can use probabilistic parameters about the occurrence of continuous traffic by using this geographical information and the opening hours of the surrounding shops.

Furthermore, we are planning to utilize the location and the temporal characteristics of mobile routers. Since business people and visitors have multiple Wi-Fi devices recently, they tend to hold mobile routers. Based on the geographical information, it is possible to utilize the statistics about human congestion at the roadside. Then, in the rush hour, the traffic from mobile routers will be increased in the vicinity of people gathering such as the station.

Figure 5.16: Measurement experiment in laboratory environment

By applying such a spatiotemporal trend as a probabilistic model to the channel prediction, we aim at the sophistication of our proposed method.

Implementation of prototype

We are implementing the prototype of an autonomous/intelligent AP with frame monitoring module, channel quality prediction module, and quality measurement module. We are planning to show feasibility at low cost by using small Linux devices with existing Wi-Fi module. Now we are developing in Scenargie Comm Node which has two Wi-Fi interfaces for transmitting and monitoring. Measurement of Wi-Fi usage status and throughput in the laboratory environment is also currently underway.

We are also designing an experiment to predict channel quality by applying a prediction module based on traffic, RSSI and channel information which is input parameter of this method observed by the monitoring module. The channel quality will be evaluated by the throughput using the IPERF application. Through this experiment, we would like to ascertain the effectiveness of our proposed method in the actual environment rather than simulation.

As a preliminary experiment in Figure 5.16, we conducted channel monitoring and quality measurement in the laboratory environment of midnight on holiday. We deployed AP and ST in the laboratory. Then AP measures the throughput for 10 seconds in each channel by the IPERF application. AP and ST changes the channel from 1 to 13 sequentially. Since MCS index was set to 0, the maximum transmission rate is up to 6.5 Mbps.

As a result, even though AP observed the same level of traffic ($t(k)$) and RSSI ($s(k)$) at the channel 1 and 2, the throughput of channel 1 was much greater than that of channel 2 without interference from any other channels. Therefore, in that case, our prediction function will not be worked well. One possible reason for this is that the interference effect may differ depending on the type of observed frames. It was confirmed that the frame error rate of channel 2 was higher than that of channel 1. In channel 2, many Data frames were observed. On the other hand, in channel 1, most of the observed frames were Beacon frames. From these findings, we can guess that frames whose priority is low like Beacon frames are likely to be suppressed by CSMA/CA. In our current proposed method, since we do not sufficiently consider the various frame types and protocols, we have to propose the function design which can estimate the channel performance in the real environment.

5.7 Conclusion

In this chapter, we have designed and modeled the urban environment based on the real measurement in Osaka city to evaluate our proposed method in more realistic environment. To this end, we conducted the traffic monitoring at ten locations including shopping malls, cafes, commercial buildings and stations around Osaka station on both weekdays and holidays to understand the current traffic situation in such typical urban environment. Then we analyzed this actual measurement results and designed the traffic parameters from the future traffic prediction based on the distribution on the number of APs in each channel. Also, we obtained the actual AP locations and their corresponding channels from Wi-Fi Radio Map of Osaka City which our research group have built. We matched this information with OpenStreetMap in the network simulator, and then we constructed three real geographical scenarios for this evaluation experiment.

By using these evaluation scenario, we demonstrated whether our proposed function can be used in these urban scenarios for channel selection at the target AP. We confirmed that it was possible to predict and select the best channel with the highest communication quality and also predict the trend over all channels by our proposed method in all scenarios. In the subsequent validity evaluation experiment, it was confirmed that the throughput of the target AP becomes about 1.83 times as compared with the expected value when the target AP select a channel randomly. As a result, our proposed approach can select interference-free channels considering the inter-channel interference problem for using the channel resources effectively.

We note that some contents in this chapter refer our previous publications [109].

Chapter 6

Conclusion

This dissertation has presented a novel approach to select interference-free channels in 2.4GHz Wi-Fi. The goal of this dissertation is to cope with interference problems; how the inter-channel interference affects the performance, how it is closely related with RSSI and traffic volume, and how we should choose a channel in an open, uncoordinated situation. We have addressed the issue by introducing a simple monitoring scheme to the target AP and predicting each channel performance for channel selection based on exhaustive simulation dataset. In particular, we consider (i) inter-channel interference where adjacent channels interfere with each other in Wi-Fi systems and (ii) urban situations where many APs in different systems are deployed in an uncoordinated way. In this dissertation, we have made the following three primary contributions to embody this idea.

Firstly, we have presented an algorithm to estimate each channel performance by multiple regression functions. As it is often hard to identify the channel with less interference in the urban situation, we present a channel scoring function that estimates the performance level of each channel based on the concept of interference environment sensing. To cope with the problem, our approach for ranking function derives a relative indicator of channel quality based on realistic, observable parameters like inter-channel distance, RSSI and traffic volume. We apply the IEEE802.11 MAC frame monitoring in each channel, which can be obtained by the off-the-shelf devices with low-cost. In order to build the scoring function based on the observations, we have conducted exhaustive simulations with a large number of scenarios, and multiple regression analysis has been applied where channel occupancy patterns, traffic volumes and RSSI in those channels are used as explanatory variables. Relying on exhaustive simulations but with a reduced number of simulation cases, our model built by regression analysis achieves sufficient accuracy to estimate better Wi-Fi channels. To evaluate our method, this scoring function was examined in two kinds of general and realistic scenario (typical and dense scenario) where several APs interfere with the AP of interest in a 150m $\times$ 150m region. We have confirmed that the scores and the actual performance are well-matched where the Spearman's rank correlation coefficient was over 0.8 and can identify the top-ranked channel as well.

Secondly, we have presented an improved approach to predict each channel performance for channel selection at the target AP. In order to let APs not select erroneous channels, it is quite essential to provide an estimation function for APs to correctly estimate the channel status without actually moving into it. Therefore, we have prepared more than 10,000 scenarios and conducted simulations which are assumed that the own traffic of the target AP moves to the new channel to simulate "channel state change". We analyzed the interference dataset for understanding the relationship between the observed parameters and interference effect to build our proposed function. We applied a machine learning based classification algorithm to estimate channel saturation due to the traffic movement in channel migration and multiple regression analysis to build a prediction function of channel performance under saturation. We confirmed that our function can classify the channel state accurately and estimate the frame delivery ratio with less than 10% error in average with additional 2,000 simulations. Moreover, we demonstrate that our estimator can capture the tendency of overall channel performance in the more general scenario for channel selection. The experimental result shows the correlation coefficient between our estimator output and the groundtruth is above 0.85.

Thirdly, we have designed realistic urban scenarios based on the actual measurement in Osaka city to evaluate our proposed approach in more realistic environments. We conducted the traffic monitoring at ten locations including shopping malls, cafes, commercial buildings and stations around Osaka station on both weekdays and holidays to understand the current traffic situation in such typical urban environments. Then we analyzed this actual measurement results and designed the traffic parameters from the future traffic prediction based on the distribution on the number of APs in each channel. Also, we obtained the actual AP locations and their corresponding channels from Wi-Fi Radio Map of Osaka City which has been built by our group. We matched this information with OpenStreetMap in the network simulator, and then we constructed three real geographical scenarios for this evaluation experiment. In these evaluation scenarios, we demonstrated that it was possible to predict and select the best channel with the highest communication quality and also predict the trend over all channels by our proposed method in all scenarios. In the subsequent validity evaluation experiment, the experimental result shows that the throughput of the target AP becomes about 1.73 times as compared with the expected value when the target AP select a channel randomly.

Through these contributions, it has been shown that our channel management mechanism offers efficient frequency reuse by passive frame monitoring at the point of interest. Our study leaves potentials of further studies for cooperative channel management of multiple APs at the target area. For instance, our proposed estimator predicts the interference effect by MAC frame monitoring when the target AP uses some channel. Since the interference effects are mutual, we can also estimate the negative impact for other surrounding APs. Such information will be helpful for channel management of the surroundings APs as the target AP sends it with a management frame such as a beacon frame. Also, cooperation between our proposed approach and Wi-Fi Radio Map of Osaka City can be considered. In this dissertation, we use only the estimated location and operating channel of each observed AP based on the smartphone user's crowdsensing from this radio map. This radio map also has the transmission range for each observed AP based on the estimated RSSI by the network simulator. Therefore, by assuming traffics from actual observations in a typical environment, we can estimate the best channel at arbitrary points by using the information on radio wave range and operating a channel of each AP provided by this radio map. We believe that it would be worth continuing to seek such further possibilities toward the more efficient frequency management solution.

Acknowledgement

I would like to express my sincerely appreciation to continued support of my supervisor Professor Teruo Higashino of Osaka University through trials and tribulations of this Ph.D thesis. Again I express my heartiest gratitude to him for his encouragement and invaluable comments in preparing this thesis.

I am very grateful to Professor Masayuki Murata, Professor Takashi Watanabe, Professor Toru Hasegawa and Professor Morito Matsuoka of Osaka University for their invaluable comments and helpful suggestions concerning this thesis.

I am heartily grateful to Associate Professor Hirozumi Yamaguchi for the precious advices and technical discussions provided through out the research.

I would like to thank Associate Professor Akihito Hiromori and Assistant Professor Akira Uchiyama of Osaka University, and Associate Professor Takaaki Umedu of Shiga University for their valuable comments.

I also thank Dr. Mineo Takai at University of California, Los Angeles for his friendly cooperation.

Thanks go to everyone of Higashino laboratory for their feedback, encouragement and support.

Finally, I would like to thank my family and my friends for their help and understanding.

Bibliography

[1] Arkasha, Bobzilla, and Uhtu, “Wigle: Wireless network mapping,” https://wigle.net/.

[2] T. Amano, “Wi-Fi Radio Map of Osaka City,” https://www.wifibigdata.org.

[3] T. Amano, S. Kajita, H. Yamaguchi, T. Higashino, and M. Takai, “A Crowdsourcing and Simulation based Approach for Fast and Accurate Wi-Fi Radio Map Construction in Urban Environment,” in *Proceedings of the IFIP Networking 2017 (Networking ’17)*, June 2017, pp. 1–9.

[4] Ministry of Internal Affairs and Communications, JAPAN, “SAQ2 JAPAN Project,” http://www.soumu.go.jp/main_sosiki/joho_tsusin/eng/Releases/Telecommunications/140612_02.html.

[5] SoftBank Corp., “FREE Wi-Fi PASSPORT,” https://www.softbank.jp/en/mobile/special/freewifi/en/.

[6] K. Fukuda, H. Asai, and K. Nagami, “Tracking the Evolution and Diversity in Network Usage of Smartphones,” in *Proceedings of the 2015 Internet Measurement Conference (IMC ’15)*, USA, 2015, pp. 253–266.

[7] AT&T Inc., “Wi-Fi from AT&T,” https://www.att.com/wi-fi/.

[8] Barcelona, “Smart City 22@,” http://www.22barcelona.com/index.php.

[9] S. Zygiaris, “Smart City Reference Model: Assisting Planners to Conceptualize the Building of Smart City Innovation Ecosystems,” *Journal of the Knowledge Economy*, vol. 4, no. 2, pp. 217–231, June 2013.

[10] Firdaus, E. Nugroho, and A. Sahroni, “ZigBee and wifi network interface on Wireless Sensor Networks,” in *Proceedings of the 2014 Makassar International Conference on Electrical Engineering and Informatics (MICEEI ’14)*, November 2014, pp. 54–58.

[11] Y. Yu, F. Han, Y. Bao, and J. Ou, “A Study on Data Loss Compensation of WiFi-Based Wireless Sensor Networks for Structural Health Monitoring,” *IEEE Sensors Journal*, vol. 16, no. 10, pp. 3811–3818, May 2016.

[12] L. Li, H. Xiaoguang, C. Ke, and H. Ketai, "The Applications of WiFi-based Wireless Sensor Network in Internet of Things and Smart Grid," in *Proceedings of the 6th IEEE Conference on Industrial Electronics and Applications (ICIEA '11)*, June 2011, pp. 789–793.

[13] D. Saeed and R. Iqbal, "NFC and Wi-Fi enabled smart clock for smart communication," in *Proceedings of the 2015 International Conference on Open Source Systems Technologies (ICOSST '15)*, December 2015, pp. 30–35.

[14] M. E. E. Alahi, L. Xie, S. Mukhopadhyay, and L. Burkitt, "A Temperature Compensated Smart Nitrate-Sensor for Agricultural Industry," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 9, pp. 7333–7341, September 2017.

[15] Audi of America, "Audi connect," https://www.audiusa.com/technology/intelligence/audi-connect.

[16] J.-R. Lin, T. Talty, and Ozan K. Tonguz, "An Empirical Performance Study of Intra-vehicular Wireless Sensor Networks under WiFi and Bluetooth Interference," in *Proceedings of the 2013 IEEE Global Communications Conference (GLOBECOM '13)*, December 2013, pp. 581–586.

[17] N. Cheng, N. Lu, N. Zhang, X. Zhang, X. Shen, and J. W. Mark, "Opportunistic WiFi Offloading in Vehicular Environment: A Game-Theory Approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, pp. 1944–1955, 2016.

[18] C. Chou, C. Li, W. Chien, and K. Lan, "A Feasibility Study on Vehicle-to-Infrastructure Communication: WiFi vs. WiMAX," in *Proceedings of the 10th International Conference on Mobile Data Management: Systems, Services and Middleware (MDM '09)*, May 2009, pp. 397–398.

[19] K. Dhondge, S. Song, B. Choi, and H. Park, "WiFiHonk: Smartphone-Based Beacon Stuffed WiFi Car2X-Communication System for Vulnerable Road User Safety," in *Proceedings of the IEEE 79th Vehicular Technology Conference (VTC Spring 2014)*, May 2014, pp. 1–5.

[20] D. Zhang and C. K. Yeo, "Enabling Efficient WiFi-Based Vehicular Content Distribution," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 3, pp. 479–492, 2013.

[21] Z. Zheng, P. Sinha, and S. Kumar, "Sparse WiFi Deployment for Vehicular Internet Access With Bounded Interconnection Gap," *IEEE/ACM Transactions on Networking*, vol. 20, no. 3, pp. 956–969, 2012.

[22] P. Lv, X. Wang, X. Xue, and M. Xu, "SWIMMING: Seamless and Efficient WiFi-Based Internet Access from Moving Vehicles," *IEEE Transactions on Mobile Computing*, vol. 14, no. 5, pp. 1085–1097, 2015.

[23] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-Dense Networks: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2522–2545, Fourthquarter 2016.

[24] S. Biswas, J. Bicket, E. Wong, R. Musaloiu-E, A. Bhartia, and D. Aguayo, "Large-scale Measurements of Wireless Network Behavior," in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication (SIGCOMM '15)*, August 2015, pp. 153–165.

[25] M. Bennis, M. Simsek, A. Czylwik, W. Saad, S. Valentin, and M. Debbah, "When Cellular Meets WiFi in Wireless Small Cell Networks," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 44–50, June 2013.

[26] S. Stefanatos and A. Alexiou, "Access Point Density and Bandwidth Partitioning in Ultra Dense Wireless Networks," *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3376–3384, September 2014.

[27] B. Soret, K. I. Pedersen, N. T. K. Jørgensen, and V. Fernández-López, "Interference Coordination for Dense Wireless Networks," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 102–109, January 2015.

[28] O. Galinina, A. Pyattaev, S. Andreev, M. Dohler, and Y. Koucheryavy, "5G Multi-RAT LTE-WiFi Ultra-Dense Small Cells: Performance Dynamics, Architecture, and Trends," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 6, pp. 1224–1240, June 2015.

[29] E. Khorov, A. Kiryanov, A. Lyakhov, and G. Bianchi, "A Tutorial on IEEE 802.11ax High Efficiency WLANs," *IEEE Communications Surveys & Tutorials*, pp. 1–1, September 2018.

[30] K. Shin, I. Park, J. Hong, D. Har, and D.-H. Cho, "Per-node Throughput Enhancement in Wi-Fi Densenets," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 118–125, 2015.

[31] D. Jaisinghani, V. S. Naik, S. K. Kaul, R. Balan, and S. Roy, "Improving the Performance of WLANs by Reducing Unnecessary Active Scans," *arXiv e-prints*, 2018, arXiv:1807.05523 [cs.NI].

[32] D. Jaisinghani, V. Naik, S. K. Kaul, and S. Roy, "Sniffer-based Inference of the Causes of Active Scanning in WiFi Networks," in *Proceedings of the 2017 Twenty-third National Conference on Communications (NCC '17)*, March 2017, pp. 1–6.

[33] S. Sen, B. Radunovic, J. Lee, and K.-H. Kim, "CSpy: Finding the Best Quality Channel Without Probing," in *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking (MobiCom '13)*, September 2013, pp. 267–278.

[34] C. M. Barnhart, J. E. Wieselthier, and A. Ephremides, "A Neural Network Approach to Solving the Link Activation Problem in Multihop Radio Networks," *IEEE Transactions on Communications*, vol. 43, no. 234, pp. 1277–1283, February 1995.

[35] J. E. Wieselthier, C. M. Barnhart, and A. Ephremides, "A Neural Network Approach to Routing Without Interference in Multihop Radio Networks," *IEEE Transactions on Communications*, vol. 42, no. 1, pp. 166–177, January 1994.

[36] R. Ogier and D. Beyer, "Neural Network Solution to the Link Scheduling Problem Using Convex Relaxation," in *Proceedings of the 1990 IEEE Global Communications Conference (GLOBECOM '90)*, Dec 1990, pp. 1371–1376 vol.2.

[37] N. Funabiki and S. Nishikawa, "A Binary Neural Network Approach for Link Activation Problems in Multihop Radio Networks," *IEICE Transanctions on Networking*, vol. E79-B, no. 8, pp. 1086–1093, August 1996.

[38] A. A. Bertossi and M. A. Bonuccelli, "Code Assignment for Hidden Terminal Interference Avoidance in Multihop Packet Radio Networks," *IEEE Transactions on Networking*, vol. 3, no. 4, pp. 441–449, August 1995.

[39] N. Funabiki and J. Kitamichi, "A Gradual Neural-network Algorithm for Jointly Time-slot/code Assignment Problems in Packet Radio Networks," *IEEE Transactions on Networking*, vol. 9, no. 6, pp. 1523–1528, November 1998.

[40] B. Hajek and G. Sasaki, "Link Scheduling in Polynomial Time," *IEEE Transactions on Information Theory*, vol. 34, no. 5, pp. 910–917, September 1988.

[41] U. Mukherji, "A Periodic Scheduling Problem in Flow Control for Data Communication Networks," *IEEE Transactions on Information Theory*, vol. 35, no. 2, pp. 436–443, March 1989.

[42] V. S. A. Kumar, M. V. Marathe, S. Parthasarathy, and A. Srinivasan, "Algorithmic aspects of capacity in wireless networks," in *Proceedings of the 2005 ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS '05)*, June 2005, pp. 133–144.

[43] R. Giuliano, C. Monti, and P. Loreti, "WiMAX Fractional Frequency Reuse for Rural Environments," *IEEE Wireless Communications*, vol. 15, no. 3, pp. 60–65, 2008.

[44] J. Sheng, T. David NC, S. Joseph B, H. Jilei, S. John E, P. Roberto, *et al.*, "Multicell Downlink Capacity with Coordinated Processing," *EURASIP Journal on Wireless Communications and Networking*, vol. 2008, pp. 18:1–18:19, 2008.

[45] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network Coordination for Spectrally Efficient Communications in Cellular Systems," *IEEE Wireless Communications*, vol. 13, no. 4, pp. 56–61, 2006.

[46] N. Jindal, J. G. Andrews, and S. Weber, "Bandwidth Partitioning in Decentralized Wireless Networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 5408–5419, 2008.

[47] R. Veronesi, V. Tralli, J. Zander, and M. Zorzi, "Distributed Dynamic Resource Allocation with Power Shaping for Multicell SDMA Packet Access Networks," in *Proceedings of the IEEE*

Wireless Communications and Networking Conference (WCNC '04), vol. 4, march 2004, pp. 2515–2520.

[48] L. Song and A. Striegel, "SEWS: A Channel-Aware Stall-Free WiFi Video Streaming Mechanism," in *Proceedings of the 28th ACM SIGMM Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV '18)*, 2018, pp. 67–72.

[49] H. C. Lee, K. T. Kim, H. Y. Youn, and O. Song, "An Efficient AP Channel Scanning Scheme for Real-time Streaming over WLAN," in *Proceedings of the 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, March 2010, pp. 558–563.

[50] A. Bhartia, B. Chen, F. Wang, D. Pallas, R. Musaloiu-E, T. T.-T. Lai, and H. Ma, "Measurement-based, Practical Techniques to Improve 802.11Ac Performance," in *Proceedings of the 2017 Internet Measurement Conference (IMC '17)*, 2017, pp. 205–219.

[51] K. Sui, M. Zhou, D. Liu, M. Ma, D. Pei, Y. Zhao, Z. Li, and T. Moscibroda, "Characterizing and Improving WiFi Latency in Large-Scale Operational Networks," in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '16)*, 2016, pp. 347–360.

[52] A. Abedi and T. Brecht, "Examining Relationships Between 802.11N Physical Layer Transmission Feature Combinations," in *Proceedings of the 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM '16)*, 2016, pp. 229–238.

[53] A. Mishra, V. Shrivastava, D. Agrawal, S. Banerjee, and S. Ganguly, "Distributed Channel Management in Uncoordinated Wireless Environments," in *Proceedings of the ACM 12th Annual International Conference on Mobile Computing and Networking (MobiCom '06)*, 2006, pp. 170–181.

[54] A. Akella, G. Judd, S. Seshan, and P. Steenkiste, "Self-management in Chaotic Wireless Deployments," in *Proceedings of the ACM 11th Annual International Conference on Mobile Computing and Networking (MobiCom '05)*, 2005, pp. 185–199.

[55] V. Mhatre, K. Papagiannaki, and F. Baccelli, "Interference Mitigation Through Power Control in High Density 802.11 WLANs," in *Proceedings of the 26th IEEE International Conference on Computer Communications (INFOCOM '07)*, May 2007, pp. 535–543.

[56] B. Kauffmann, F. Baccelli, A. Chaintreau, V. Mhatre, K. Papagiannaki, and C. Diot, "Measurement-Based Self Organization of Interfering 802.11 Wireless Access Networks," in *Proceedings of the 26th IEEE International Conference on Computer Communications (INFOCOM '07)*, May 2007, pp. 1451–1459.

[57] S. Chieochan, E. Hossain, and J. Diamond, "Channel Assignment Schemes for Infrastructure-based 802.11 WLANs: A Survey," *IEEE Communications Surveys Tutorials*, vol. 12, no. 1, pp. 124–136, First 2010.

[58] K. N. Ramachandran, E. M. Belding, K. C. Almeroth, and M. M. Buddhikot, "Interference-Aware Channel Assignment in Multi-Radio Wireless Mesh Networks," in *Proceedings of the 25th IEEE International Conference on Computer Communications (INFOCOM '06)*, April 2006, pp. 1–12.

[59] A. Şahin, E. Bala, İ. Güvenç, R. Yang, and H. Arslan, "Partially Overlapping Tones for Uncoordinated Networks," *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3363–3375, September 2014.

[60] J. Zheng, Y. Cai, W. Yang, Y. Xu, and A. Anpalagan, "A Game-Theoretic Approach to Exploit Partially Overlapping Channels in Dynamic and Distributed Networks," *IEEE Communications Letters*, vol. 18, no. 12, pp. 2201–2204, December 2014.

[61] Y. Xu, Q. Wu, J. Wang, L. Shen, and A. Anpalagan, "Opportunistic Spectrum Access Using Partially Overlapping Channels: Graphical Game and Uncoupled Learning," *IEEE Transactions on Communications*, vol. 61, no. 9, pp. 3906–3918, September 2013.

[62] P. B. F. Duarte, Z. M. Fadlullah, A. V. Vasilakos, and N. Kato, "On the Partially Overlapped Channel Assignment on Wireless Mesh Network Backbone: A Game Theoretic Approach," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 119–127, January 2012.

[63] Y. Ding, Y. Huang, G. Zeng, and L. Xiao, "Using Partially Overlapping Channels to Improve Throughput in Wireless Mesh Networks," *IEEE Transactions on Mobile Computing*, vol. 11, no. 11, pp. 1720–1733, November 2012.

[64] J. Zheng, Y. Cai, Y. Xu, and A. Anpalagan, "Distributed Channel Selection for Interference Mitigation in Dynamic Environment: A Game-Theoretic Stochastic Learning Solution," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 9, pp. 4757–4762, November 2014.

[65] S. Sen, R. Roy Choudhury, and S. Nelakuditi, "No Time to Countdown: Migrating Backoff to the Frequency Domain," in *Proceedings of the 17th Annual International Conference on Mobile Computing and Networking (MobiCom '11)*, 2011, pp. 241–252.

[66] D. Malone, P. Clifford, and D. Leith, "MAC Layer Channel Quality Measurement in 802.11," *IEEE Communications Letters*, vol. 11, no. 2, pp. 143–145, February 2007.

[67] D. Aguayo, J. Bicket, S. Biswas, G. Judd, and R. Morris, "Link-level Measurements from an 802.11b Mesh Network," in *Proceedings of the ACM 2004 Conference on Applications, Technolo-*

gies, Architectures, and Protocols for Computer Communications (SIGCOMM '04), 2004, pp. 121–132.

[68] S. Rayanchu, A. Patro, and S. Banerjee, "Airshark: Detecting non-WiFi RF Devices Using Commodity WiFi Hardware," in *Proceedings of the 2011 Internet Measurement Conference (IMC '11)*, 2011, pp. 137–154.

[69] S.-M. Chun, S.-M. Lee, J.-W. Nah, J.-H. Choi, and J.-T. Park, "Localization of Wi-Fi Access Point using smartphone's GPS information," in *Proceedings of the 2001 International Conference on Selected Topics in Mobile and Wireless Networking (iCOST '11)*, October 2011, pp. 121–126.

[70] A. Achtzehn, L. Simic, M. Petrova, and P. Mahonen, "IEEE 802.11 Wi-Fi Access Point Density Estimation with Capture-recapture Models," in *Proceedings of the 2015 International Conference on Computing, Networking and Communications (ICNC '15)*, February 2015, pp. 153–159.

[71] F. Li, C. Zhao, G. Ding, J. Gong, C. Liu, and F. Zhao, "A Reliable and Accurate Indoor Localization Method Using Phone Inertial Sensors," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (Ubicomp '12)*, 2012, pp. 421–430.

[72] J. Yin, Q. Yang, and L. M. Ni, "Learning Adaptive Temporal Radio Maps for Signal-Strength-Based Location stimation," *IEEE Transactions on Mobile Computing*, vol. 7, no. 7, pp. 869–883, 2008.

[73] P. Bahl and V. N. Padmanabhan, "RADAR: An In-building RF-based User Location and Tracking System," in *Proceedings of the IEEE 19th International Conference on Computer Communications (INFOCOM '00)*, vol. 2, March 2000, pp. 775–784.

[74] Y. Cho, M. Ji, Y. Lee, and S. Park, "WiFi AP Position Estimation Using Contribution from Heterogeneous Mobile Devices," in *Proceedings of the 2012 IEEE/ION Position, Location and Navigation Symposium (PLANS '12)*, 2012, pp. 562–567.

[75] H. Satoh, S. Ito, and N. Kawaguchi, "Position Estimation of Wireless Access Point Using Directional Antennas," in *Proceedings of the 1st International Workshop on Location-and Context-Awareness (LoCA '05)*, 2005, pp. 144–156.

[76] A. P. Subramanian, P. Deshpande, J. Gao, and S. R. Das, "Drive-by Localization of Roadside WiFi Networks," in *Proceedings of the 27th IEEE International Conference on Computer Communications (INFOCOM '08)*, 2008, pp. 1391–1399.

[77] D. Han, D. G. Andersen, M. Kaminsky, K. Papagiannaki, and S. Seshan, "Access Point Localization Using Local Signal Strength Gradient," in *Proceedings of the 10th International Conference on Passive and Active Network Measurement (PAM '09)*, 2009, pp. 99–108.

[78] Z. Fang, L. U. O. Haiyong, G. Hao, and S. U. N. Qijin, "An RSSI Gradient-based AP Localization Algorithm," *China Communications*, vol. 11, no. 2, pp. 100–108, 2014.

[79] Z. Zhang, X. Zhou, W. Zhang, Y. Zhang, G. Wang, B. Y. Zho, and H. Zheng, "I Am the Antenna-Accurate Outdoor AP Location using Smartphones," in *Proceedings of the 17th ACM International Conference on Mobile Computing and Networking. (MobiCom '11)*, 2011, pp. 109–120.

[80] X. Zheng, C. Wang, Y. Chen, and J. Yang, "Accurate Rogue Access Point Localization Leveraging Fine-grained Channel Information," in *Proceedings of the IEEE 2014 Conference on Communications and Network Security (CNS '14)*, 2014, pp. 211–219.

[81] J. Shi, L. Meng, A. Striegel, C. Qiao, D. Koutsonikolas, and G. Challen, "A Walk on the Client Side: Monitoring Enterprise Wifi Networks Using Smartphone Channel Scans," in *Proceedings of the 35th IEEE International Conference on Computer Communications (INFOCOM '16)*, 2016, pp. 1–9.

[82] S. Yang, P. Dessai, M. Verma, and M. Gerla, "FreeLoc: Calibration-free Crowdsourced Indoor Localization," in *Proceedings of the 32nd IEEE International Conference on Computer Communications (INFOCOM '13)*, 2013, pp. 2481–2489.

[83] Q. Jiang, Y. Ma, K. Liu, and Z. Dou, "A Probabilistic Radio Map Construction Scheme for Crowdsourcing-Based Fingerprinting Localization," *IEEE Sensors Journal*, vol. 16, no. 10, pp. 3764–3774, 2016.

[84] A. LaMarca, Y. Chawathe, S. Consolvo, J. Hightower, I. Smith, J. Scott, T. Sohn, J. Howard, J. Hughes, F. Potter, J. Tabert, P. Powledge, G. Borriello, and B. Schilit, "Place Lab: Device Positioning Using Radio Beacons in the Wild," in *Proceedings of the 3rd IEEE International Conference on Pervasive Computing and Communications (PerCom '05)*, 2005, pp. 116–133.

[85] S. Rosen, S. J. Lee, J. Lee, P. Congdon, Z. Mao, and K. Burden, "MCNet: Crowdsourcing Wireless Performance Measurements through the Eyes of Mobile Devices," *IEEE Communications Magazine*, vol. 52, no. 10, pp. 86–91, 2014.

[86] A. W. T. Tsui, W. C. Lin, W. J. Chen, P. Huang, and H. H. Chu, "Accuracy Performance Analysis between War Driving and War Walking in Metropolitan Wi-Fi Localization," *IEEE Transactions on Mobile Computing*, vol. 9, no. 11, pp. 1551–1562, 2010.

[87] J. Koo and H. Cha, "Unsupervised Locating of WiFi Access Points Using Smartphones," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 42, no. 6, pp. 1341–1353, 2012.

[88] M. L Stein, *Interpolation of Spatial Data: Some Theory for Kriging*. Springer, 1999.

[89] H. Zhao, B. Huang, and B. Jia, "Applying Kriging Interpolation for WiFi Fingerprinting based Indoor Positioning Systems," in *Proceedings of the 2016 IEEE Wireless Communications and Networking Conference (WCNC '16)*, 2016, pp. 1–6.

[90] C. Liu, A. Kiring, N. Salman, L. Mihaylova, and I. Esnaola, "A Kriging Algorithm for Location Fingerprinting based on Received Signal Strength," in *Proceedings of the 10th Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF '15)*, 2015, pp. 1–6.

[91] S. Gollakota, F. Adib, D. Katabi, and S. Seshan, "Clearing the RF Smog: Making 802.11n Robust to Cross-Technology Interference," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 4, pp. 170–181, 2011.

[92] C.-J. M. Liang, N. B. Priyantha, J. Liu, and A. Terzis, "Surviving Wi-Fi Interference in Low Power Zigbee Networks," in *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems (SenSys '10)*, 2010, pp. 309–322.

[93] F. Hermans, O. Rensfelt, T. Voigt, E. Ngai, L.-Å. Norden, and P. Gunningberg, "SoNIC: Classifying Interference in 802.15.4 Sensor Networks," in *Proceedings of the 12th ACM International Conference on Information Processing in Sensor Networks (IPSN '13)*, 2013, pp. 55–66.

[94] J. Lansford, A. Stephens, and R. Nevo, "Wi-Fi (802.11b) and Bluetooth: Enabling Coexistence," *IEEE Network*, vol. 15, no. 5, pp. 20–27, September 2001.

[95] R. Garroppo, L. Gazzarrini, S. Giordano, and L. Tavanti, "Experimental assessment of the coexistence of wi-fi, zigbee, and bluetooth devices," in *Proceedings of the 2011 IEEE International Symposium on World of Wireless, Mobile and Multimedia Networks (WoWMoM '11)*, June 2011, pp. 1–9.

[96] A. C.-C. Hsu, D. S. Wei, and C.-C. J. Kuo, "Coexistence Wi-Fi MAC Design for Mitigating Interference Caused by Collocated Bluetooth," *IEEE Transactions on Computers*, vol. 64, no. 2, pp. 342–352, 2015.

[97] A. Hithnawi, H. Shafagh, and S. Duquennoy, "TIIM: Technology-Independent Interference Mitigation for Low-power Wireless Networks," in *Proceedings of the ACM 14th Annual International Conference on Information Processing in Sensor Networks (IPSN '15)*, April 2015, pp. 1–12.

[98] IEEE 802.19 Wireless Coexistence Working Group (WG), "IEEE 802.19," http://www.ieee802.org/19/.

[99] Space-Time Engineering., LLC., "Scenargie High Quality System Simulation Framework," http://www.spacetime-eng.com.

[100] *IEEE Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, 2007.

[101] M. Takai, Y. Owada, and K. Seki, "A Comparative Study on Network Simulators for ITS Simulation IEEE802.11 Medium Access Control (MAC) Models," *16th ITS World Congress and Exhibition on Intelligent Transport Systems and Services*, September 2009.

[102] Riverbed Technology, "AirPcap Nx," http://www.riverbed.com/.

[103] NLANR/DAST and ESnet/Lawrence Berkeley National Laboratory, "iPerf-The ultimate speed test tool for TCP, UDP and SCTP," https://iperf.fr/.

[104] S. Kajita, H. Yamaguchi, T. Higashino, S. Umehara, F. Saitou, H. Urayama, M. Yamada, T. Maeno, S. Kaneda, and M. Takai, "A Channel Selection Strategy for WLAN in Urban Areas by Regression Analysis," in *Proceedings of the IEEE 10th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob '14)*, October 2014, pp. 642–647.

[105] ——, "An Algorithm for Estimating the Performance of WLAN Systems with Severe Interference," *IPSJ Journal*, vol. 57, no. 2, pp. 745–755, 2016.

[106] S. Kajita, H. Yamaguchi, T. Higashino, H. Urayama, M. Yamada, and M. Takai, "Throughput and Delay Estimator for 2.4GHz WiFi APs: A Machine Learning-Based Approach," in *Proceedings of the 8th IFIP Wireless and Mobile Networking Conference (WMNC '15)*, 2015, pp. 223–226.

[107] S. Kajita, T. Amano, H. Yamaguchi, T. Higashino, and M. Takai, "Prediction-based Channel Selection Algorithm for Autonomous Wi-Fi Access Point Management," *IPSJ Journal*, (submitted).

[108] M. Achanta, "Method and Apparatus for Least Congested Channel Scan for Wireless Access Points," *US Patent No. 20060072602*, April 2006.

[109] S. Kajita, T. Amano, H. Yamaguchi, T. Higashino, and M. Takai, "Wi-Fi Channel Selection Based on Urban Interference Measurement," in *Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous '16)*, November 2016, pp. 143–150.