Novel Wi-Fi Throughput Estimation Method Considering CSMA/CA Behavior

Yuichi Imagaki, Kanshiro Kashiki, Kosuke Yamazaki and Akira Yamaguchi
Cognitive Radio Laboratory
KDDI R&D Laboratories Inc.
2-1-15 Ohara, Fujimino, Saitama 356-8502, JAPAN

Abstract – Wi-Fi is one of the solutions to overcome the explosion of mobile data traffic and improvement of Wi-Fi performance intensifies the cellular data offload to Wi-Fi. Wi-Fi AP (Access Point) selection at the STA (STAtion) is a key function to improve Wi-Fi throughput. Our proposed method is practical for the AP selection. If an STA can know the throughputs between the STA and surrounding APs before exchanging Wi-Fi packets with an AP, the STA can select the optimum AP that promises high throughput. In this paper, we focus on a throughput estimation method that is applicable to AP selection, and discuss the accuracy of the estimated throughput. We propose a novel Wi-Fi throughput estimation method with passive measurement and without modification for the Wi-Fi standard. The proposed method consists of three stages: measurement of the time ratio used by each Wi-Fi device, selection of dominant Wi-Fi devices and calculation of estimated throughout. Our proposed method considers the influence of CSMA/CA that affects Wi-Fi throughput that changes based on the number of Wi-Fi devices sharing the same frequency band and the traffic load of each device. We evaluate the effectiveness of the proposed method by experiments (estimated before communication) and compare the estimated throughput and the actual throughput (measured during communication) in certain environments. The proposed method is effective for estimating Wi-Fi throughput. The RMSE (Root Mean Square Error) of the proposed method decreases to 29~33% that of conventional methods.

Estimated throughput with our proposed method is always higher than measured throughput. Therefore, modification of the proposed method to decrease the difference between the estimated and measured value is a subject for future study. As some systems influence Wi-Fi, considering these influences in a throughput estimation process is also recommended.

Index Terms – Wi-Fi, WLAN, Wireless LAN, CSMA/CA, time occupancy ratio, IEEE 802.11

I. Introduction

Wi-Fi is a popular system used in various locations and one of the solutions to control mobile data traffic explosion. Wi-Fi is a potent scheme for cellular data offloading. Our proposed method increases the Wi-Fi throughput by selecting the best AP and the effect of data offload using Wi-Fi improves. Wi-Fi devices share a Wi-Fi channel based on CSMA/CA. Here, Wi-Fi devices involve APs and STAs. If several Wi-Fi devices send a large amount of Wi-Fi traffic simultaneously, the resources available to them drop and the throughput of each device decreases dramatically. In contrast, an STA that will start sending packets is expected to use almost all of the available resources and achieve the high throughput communication when there are few transmitted Wi-Fi packets. Wi-Fi throughputs are affected by the number of Wi-Fi devices using the same frequency band and traffic load based on the CSMA/CA behavior.

One of the approaches to these Wi-Fi throughput contingencies is to develop a method to estimate Wi-Fi throughput. Thus, a method to estimate Wi-Fi throughput per AP is necessary to use a Wi-Fi system comfortably. In the environment where several APs are available, an STA can select the AP to maximize the estimated throughput if Wi-Fi throughput estimation is possible.

Several related works on estimated throughput have been conducted. In [1], Wi-Fi throughput is discussed analytically but some parameters used in the analytical approach are difficult to measure in a real Wi-Fi system. In [2], load balancing between APs is discussed. This is based on the simple assumption that Wi-Fi resources are allocated to Wi-Fi devices evenly. However, allocated resources change based on the number of devices and the traffic load of each device, and our experiments confirm that this method is valid under high Wi-Fi traffic load. Under low Wi-Fi traffic load, there is a tendency for estimated throughput to be below actual throughput. The classical approach is active measurement sending probe packets to measure throughput. But, the throughputs of surrounding Wi-Fi devices decrease while performing active measurement.

We proposed a novel Wi-Fi throughput estimation method considering CSMA/CA and the traffic difference of each Wi-Fi device, and the method does not include active measurement in the estimation process. Our proposed method has three stages. First, a Wi-Fi STA that estimates throughput captures Wi-Fi packets and collects information on the traffic amount of each Wi-Fi device around the STA. Second, the device selects the dominant devices that significantly influence the behavior of CSMA/CA. Third, the Wi-Fi device calculates the estimated throughput based on the assumption that CSMA/CA works among those dominant Wi-Fi devices.

We evaluated our proposed method in experiments comparing our proposed method and two conventional methods. The RMSE of our proposed method is about 24~39% that of conventional methods and can estimate Wi-Fi throughput even if Wi-Fi devices transmit many packets or not.

In Section II, our proposed method is outlined. Section III addresses experimental environments and Section IV discusses experimental results. Finally, the conclusion and further studies are summarized and addressed in Section V.

II. Proposed Method

In this section, we introduce our proposed method to estimate Wi-Fi throughput considering CSMA/CA. The main point of our proposed method is to find dominant

978-1-4673-0990-5/12/$31.00 ©2012 IEEE

Wi-Fi devices that send many packets compared to other devices. Dominant devices are defined through our proposed method which has three stages. The first stage is capturing Wi-Fi packets and calculating the occupancy ratio per Wi-Fi device. Wi-Fi devices are operated in descending order of the occupancy ratio. The second stage is to find dominant Wi-Fi devices for CSMA/CA using the sorted device list. The third stage calculates the estimated throughput based on the second stage result and the data rate between an AP and an STA that is estimated by beacon RSSI.

The process of estimating Wi-Fi throughput uses the "occupancy ratio", details of which are explained in [3]. This is our previous research whose objective was to develop a method to measure Wi-Fi channel congestion. The occupancy ratio is the time ratio used to communicate and this occupancy ratio includes waiting time, for example, DIFS or Backoff.

This process assumes the following:

a) Wi-Fi devices whose occupancy ratios are not high have no influence on CSMA/CA behavior.

b) Occupancy ratios of each dominant Wi-Fi device are the same value.

Fig. 1 shows the flow diagram of our proposed method.

A. Capturing Packets and Calculating Occupancy Ratio (First Stage)

In this stage, an STA that will start communication captures Wi-Fi packets around the STA and makes a list of surrounding Wi-Fi devices.

1. An STA that will estimate throughput measures the occupancy ratio per Wi-Fi device by capturing Wi-Fi packets.
2. A list is created containing Wi-Fi devices' MAC address and occupancy ratio per Wi-Fi device.
3. The Wi-Fi devices list is sorted in descending order of the occupancy ratio.

In this stage, a measuring tool measures the occupancy ratio per Wi-Fi device. We developed a tool that can measure the channel occupancy ratio [3] which is the sum of the occupancy ratio of each Wi-Fi device using the same channel. Although our developed tool does not have the function of measuring the occupancy ratio of each Wi-Fi device that is required to realize our proposed method in this paper, it is easy to modify the tool to enable measurement of the occupancy ratio per Wi-Fi device. The reason is that the tool in our previous work reads a MAC header of Wi-Fi packets. Wi-Fi packets include information about a sender and a receiver. Modification is simply adding information to be used in the occupancy ratio measuring process. As a result, the processes in this stage are realizable in a real Wi-Fi system.

B. Finding Dominant Devices and Obtaining Available Occupancy Ratio (Second Stage)

In this stage, our proposed method finds dominant Wi-Fi devices sending many packets. The process of the second stage is recursive. The basic principle is to define the number of dominant devices. The process evaluates the assumption about the number of dominant devices. If the assumption is invalid, the process repeats after incrementing the number of dominant devices.

"Occ_Ratio(i)" means the channel occupancy ratio of Wi-Fi device i. "Sum(k)" means "a summation of the occupancy ratios of first k device". "RR" means "1-Sum(k)". "(Sum(k) + RR)/(k+1)" means the calculation of the available resources of each Wi-Fi device considering CSMA/CA behavior between devices sending background traffic and a new STA that will start sending packets. Comparison between this value and Occ_Ratio(i) is important to find dominant Wi-Fi devices for CSMA/CA.

1. Calculate the remaining channel resources (RR) by calculating "1-Sum(k)".
2. Compare "(Sum(k) + RR)/(k+1)" and "Occ_Ratio(k)". If "(Sum(k) + RR)/(k+1)", is larger than Occ_Ratio(k), and the estimation flow is finished. And the maximum estimated throughput is represented by the No.1 formula at the third stage of the proposed method.
3. If "Occ_Ratio(k)" is larger than or equal to "(Sum(k) + RR)/(k+1)", compare "(Sum(k) + RR)/(k+1)" and "Occ_Ratio(k+1)".
4. If "Occ_Ratio(k+1)" is larger than or equal to "(Sum(k) + RR)/(k+1)", increment the value of "k" and return to comparing "(Sum(k) + RR)/(k+1)" and "Occ_Ratio(k)".
5. If "(Sum(k) + RR)/(k+1)" is larger than "(Sum(k) + RR)/(k+1)", the estimated throughput is represented by the No.2 formula at the third stage of the proposed method. The estimation process is then concluded.

C. Calculating Estimated Throughput (Third Stage)

Two formulas (Eq. 1 and 2) are available to calculate the estimated throughput.

"Throughput_MAX" is the maximum throughput based on the data rate that is determined by the distance between an AP and an STA when there are no other devices that can cause interference. α means the total throughput decrease caused by the increase of an average waiting time of CSMA/CA or packets collision. These situations increase based on the number of Wi-Fi devices that share the same Wi-Fi channel. The range of α is $0 < \alpha <= 1$. In this paper, we determined $\alpha=1$. This means that we do not consider time ratios that are consumed by waiting time, packet collision etc.

"Estimated Throughput" is the output of the flow and these values are calculated per AP.

- Estimated throughput is

$$\text{“}\alpha \times \text{Throughput\_MAX} \times RR\text{”.} \tag{1}$$

- Estimated throughput is

$$\text{“}\alpha \times \text{Throughput\_MAX} \times (\text{Sum}(k) + RR) / (k+1)\text{”.} \tag{2}$$

If there are many available resources of the Wi-Fi channel, an STA that will start communication is expected to be able to use most resources. This assumption is represented by Eq. 1. In contrast, dominant Wi-Fi devices share resources equally when there are moderate available resources. This assumption is represented by Eq. 2.

As mentioned above, there is no tool that can measure the occupancy ratio of each Wi-Fi device. Hence, we approximated the occupancy ratio per Wi-Fi device with the following formula.

$$(\text{Measured throughput})/(\text{Throughput\_MAX}). \tag{3}$$

If modulations of each device are the same, throughput is determined by only the occupancy ratio, and we can say (Occupancy ratio = 100%) is nearly equal to (throughput = "Throughput_MAX").

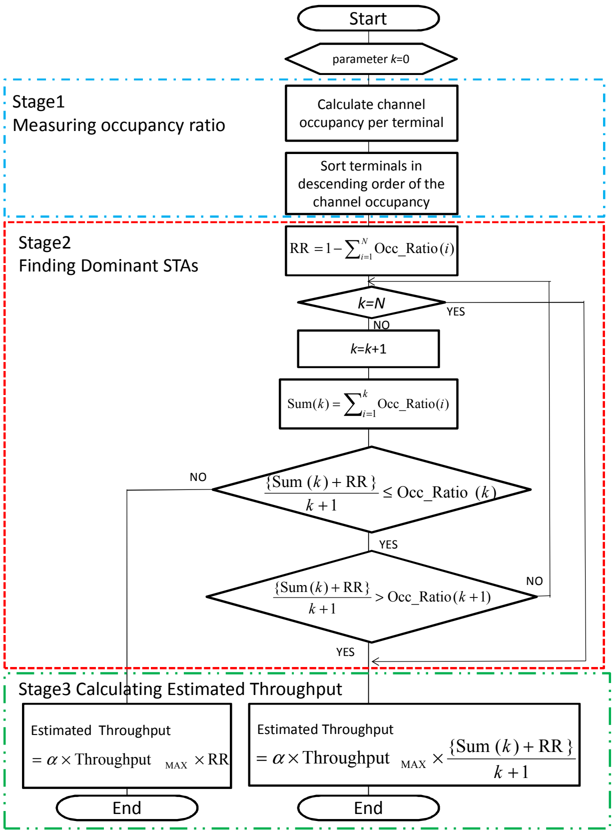

Figure 1. Flow diagram of the proposed method.

III. EXPERIMENTAL ENVIRONMENT

In this section, we describe the experiments to evaluate the proposed method. We compared the three methods and measured throughput, and then evaluated the RMSE of those methods.

A. Equal Method and Rest Method

The first comparative method is the "Equal" method. This method assumes that all Wi-Fi devices influence CSMA/CA behavior and the time occupancy ratio of every Wi-Fi device is exactly the same. The estimated throughput of "Equal" is calculated as follows:

(Throughput_MAX) / (the number of devices). (4)

The number of devices means the number of STAs and APs sharing the same frequency band.

The second comparative method is the "Rest" method. This method assumes that available resources are the remnants of background traffic usage. This method calculates the estimated throughput using the following formula:

(Throughput_MAX) – (Sum of throughput). (5)

(Sum of throughputs) means the summation of throughputs of STAs and APs that are in communication before the estimating process. We extract them from the value "Throughput_MAX" and "sum of throughputs" in experiments using iperf in this paper. There are variations regarding the output of iperf (UDP). For this reason, the estimated throughput with "Rest" sometimes becomes negative. In most cases, the value is not a negative value. This method is based on the assumption that an STA that will start sending packets can then use the entire unused time ratio.

B. Experiment Environment

Our proposed method is effective in the environment in which there are several APs. In this paper, we used one AP as a basic experiment. Fig. 2 shows the environment image. There is one AP (ICOM AP-56W) and several STAs (ICOM STA-56W). Some Wi-Fi STAs generate background traffic, and the other STA calculates the estimated throughput. Then, the STA starts sending Wi-Fi packets to the maximum extent and measures the actual throughput, and we compare the estimated and measured throughputs. All STAs can communicate with an AP with "64QAM, 3/4". We made the data rates of each Wi-Fi device the same because we used Eq. 3 to approximate the occupancy ratio in the experiments. If there is a tool that can measure the occupancy ratio of each Wi-Fi device, the data rates of each device do not need to be the same. To measure throughputs, we use iperf (UDP). The throughputs are the average of the three samples measured in 60 sec. Table I. shows the parameters of the experiments. Tables II.~IV. show the experiment scenarios regarding background traffic load. Capital letters, such as A, B, C…, mean scenarios and numbers such as 1, 2, 3… mean STAs for background traffic. The value in the chart means the occupancy ratio that corresponds to the background traffic of each STA. If STA1's value is 50%, it means STA1's measured background throughput is 0.5 times Throughput_MAX in the experiment environment. The scenarios of experiments have two categories. One is a "HIGH" background occupancy ratio where the background ratio is equal to or greater than 99%. The other is a "MEDIUM" background occupancy ratio where the background ratio is from 60% to 80%.

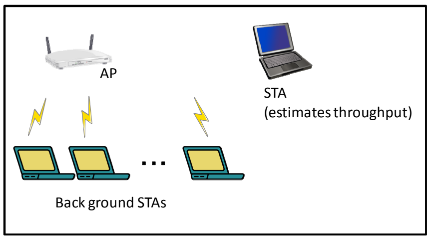

Figure 2. Environment image of the experiment.

TABLE I. PARAMETERS OF THE EXPERIMENT ENVIRONMENT

Standards	IEEE802.11a
Number of APs	1
Number of STAs	{3, 5, 9, 12}
Data Rate (All STAs and AP)	64QAM, 3/4
Protocol	UDP
UDP Packet Size(Data)	1470 Bytes
Measuring Duration	60 sec
Number of Trials	3
Throughput_MAX	29.7 Mbps

TABLE II. BACKGROUND OCCUPANCY RATIOS (STAS = 3 AND STAS=5)

	1	2
A	50	50
B	90	10
C	70	30
D	40	40
E	60	20
F	70	10
G	30	30
H	40	20
I	50	10

	1	2	3	4
A	25	25	25	25
B	40	30	20	10
C	40	20	20	20
D	40	40	10	10
E	28	21	21	21
F	28	21	14	7
G	17.5	17.5	17.5	17.5
H	28	28	7	7

TABLE III. BACKGROUND OCCUPANCY RATIOS (STAS = 9)

	1	2	3	4	5	6	7	8
A	12.5	12.5	12.5	12.5	12.5	12.5	12.5	12.5
B	0.5	4	7.5	11	14.5	18	21.5	23
C	16	16	16	10.4	10.4	10.4	10.4	10.4
D	26	26	8	8	8	8	8	8
E	12.8	12.8	12.8	8.32	8.32	8.32	8.32	8.32
F	20.8	20.8	6.4	6.4	6.4	6.4	6.4	6.4
G	8.75	8.75	8.75	8.75	8.75	8.75	8.75	8.75
H	0.35	2.8	5.25	7.7	10.15	12.6	15.05	16.1

TABLE IV. BACKGROUND OCCUPANCY RATIOS (STAS = 12)

	1	2	3	4	5	6	7	8	9	10	11
A	12	12	12	8	8	8	8	8	8	8	8
B	12	12	12	12	12	12	12	4	4	4	4
C	9	9	9	9	9	9	9	9	9	9	9
D	2	3.4	4.8	6.2	7.6	9	10.4	11.8	13.2	14.6	16
E	7.2	7.2	7.2	4.8	4.8	4.8	4.8	4.8	4.8	4.8	4.8
F	7.2	7.2	7.2	7.2	7.2	7.2	7.2	2.4	2.4	2.4	2.4
G	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4
H	1.2	2.04	2.88	3.72	4.56	5.4	6.24	7.08	7.92	8.76	9.6

IV. EXPERIMENTAL RESULT AND ANALYSIS

In this section, we explain the results of experiments. Fig. 3~6 show the results of measured throughput and estimated throughputs using the three methods. Capital letters, such as A, B, C…, on the horizontal axis mean scenarios for experiments. The vertical axis is the measured throughput and estimated throughputs. Fig. 7~9 show scatter diagrams. The horizontal axis (X-axis) is the estimated throughput, and the vertical axis (Y-axis) is the measured throughput. In the scatter diagrams, Y=X is drawn. If all dots are on the line, it means the estimated throughputs are equal to real throughputs. In scatter diagrams, cross marks indicate data when the background occupancy ratio is high. A background occupancy ratio whose value is over 99%, is regarded as high. In contrast, rhombus marks indicate data when the background occupancy ratio is not high. We introduce the RMSE to evaluate the accuracy of the estimation methods. Table V.~VII. show the RMSE of our proposed method and conventional methods. The RMSE is calculated by the following formula:

$$\mathrm{RMSE} = \left(\sum_{i=1}^{N} (\mathrm{Error}_i)^2 \Big/ N \right)^{1/2}. \tag{6}$$

Here, "N" is the sample number.

Error_i = (Estimated throughput) – (Measured Throughput).

Fig. 3~6 indicate that estimated throughputs with our proposed method behave with the same trend of measured throughputs and the estimated value is almost equal to the measured throughput. Fig. 7~9 provide a statistical analysis. It is obvious that conventional methods are sometimes able to estimate Wi-Fi throughput, but cannot estimate throughput regardless of background occupancy ratios. For example, the estimated throughput of the "Equal" method is nearly equal to the measured throughput when the background occupancy ratio is high. But the estimated throughput of the "Equal" method is different from measured throughput when the background occupancy ratio is not high. The "Rest" method cannot estimate throughput whether the background occupancy ratio is high or not. Only our proposed method can estimate regardless of the background occupancy ratio. However, estimated throughput with our proposed method is always higher than measured throughput possibly because non-Wi-Fi devices collide or interfere with Wi-Fi packets. Modification of the proposed method to decrease the difference between the estimated and measured value is a subject for future study. As there are systems that influence Wi-Fi, considering these influences in a throughput estimation process is also recommended.

Considering the RMSE under a high background occupancy ratio (Table V), the RMSE of the proposed method and "Equal" are almost the same value. However, comparing the two methods under a medium background occupancy ratio, there is 2.5Mbps difference (Table VI). The RMSE of the proposed method is the minimum value in the situation that considers a high and medium background occupancy ratio (Table VII). In summary, our assumptions and the method for defining dominant devices are outlined as follows:

a) Wi-Fi devices whose occupancy ratios are not high do not influence CSMA/CA behavior.

b) Occupancy ratios of each dominant Wi-Fi device are the same value.

As a result, our proposed Wi-Fi throughput estimation method is effective.

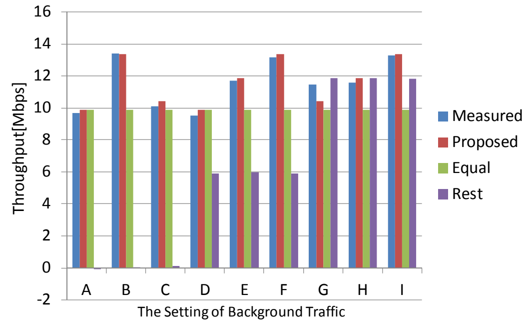

Figure 3. Measured and estimated throughputs. (STAs = 3)

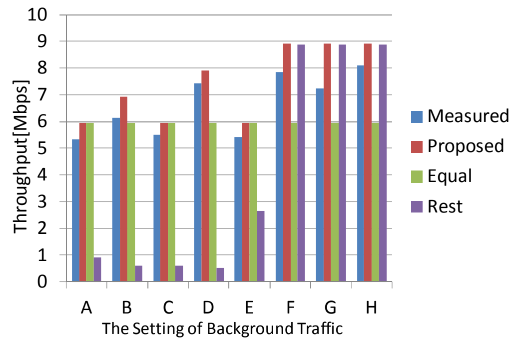

Figure 4. Measured and estimated throughputs. (STAs = 5)

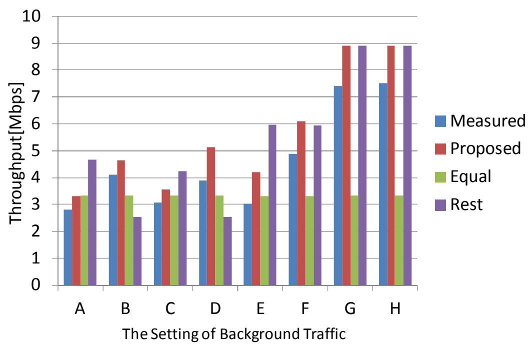

Figure 5. Measured and estimated throughputs. (STAs = 9)

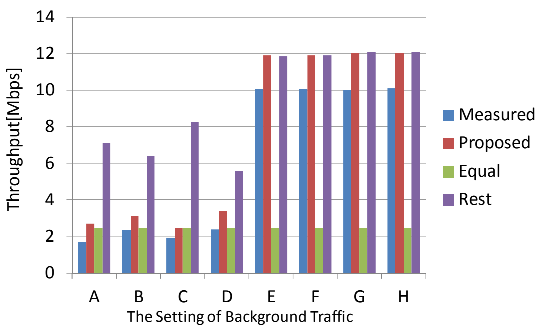

Figure 6. Measured and estimated throughputs. (STAs = 12)

TABLE V. RMSE OF THREE METHODS
(HIGH BACKGROUND OCCUPANCY RATIO)

Method	RMSE
Proposed	0.71
Equal	1.09
Rest	6.36

TABLE VI. RMSE OF THREE METHODS
(MEDIUM BACKGROUND OCCUPANCY RATIO)

Method	RMSE
Proposed	1.43
Equal	3.93
Rest	2.91

TABLE VII. RMSE OF THREE METHODS
(HIGH AND MEDIUM BACKGROUND OCCUPANCY RATIO)

Method	RMSE
Proposed	1.17
Equal	3.02
Rest	4.79

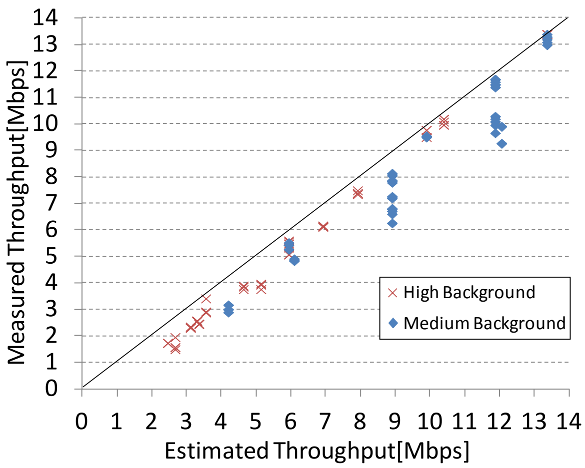

Figure 7. Scatter diagram. (Proposed)

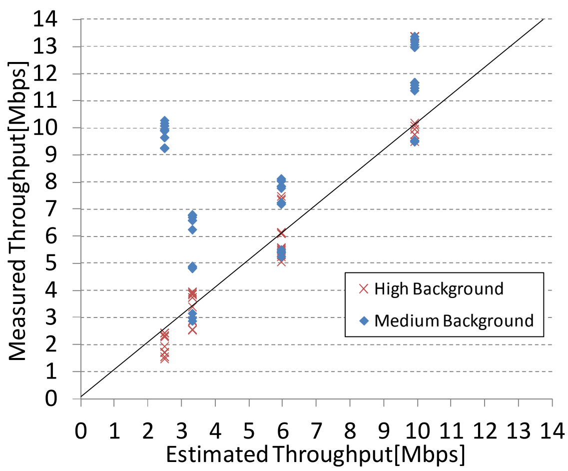

Figure 8. Scatter diagram. (Equal)

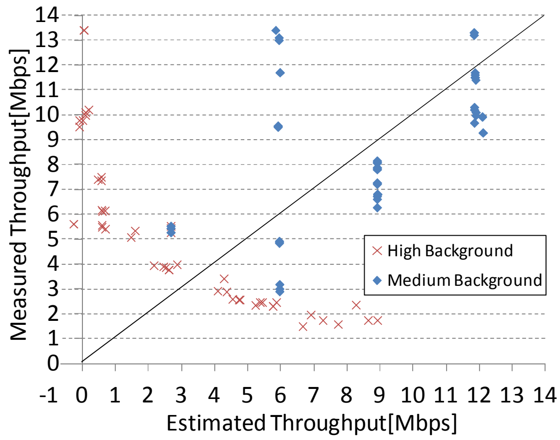

Figure 9. Scatter diagram. (Rest)

V. CONCLUSION

In this paper, we proposed a method to estimate Wi-Fi throughput considering CSMA/CA. The process of estimation is composed of three stages. The first stage measures the occupancy ratios per Wi-Fi device. The second finds dominant devices for CSMA/CA and calculates the time ratio available for an STA that has not sent packets yet. The third stage calculates a throughput considering available resources and data rates. We showed that our proposed method can estimate Wi-Fi throughput regardless of the background occupancy ratio. The RMSE of the Equal method is almost equal to that of our proposed method only under a high background occupancy ratio but the RMSE of the Equal method is large under a medium background occupancy ratio. The RMSE of the proposed method is 24~39% that of conventional methods under a high and medium background occupancy ratio. Our proposed method assumes that only dominant Wi-Fi devices divide resources equally through CSMA/CA. From the experiments, it is obvious that our proposed method based on the assumption is appropriate whether the background occupancy ratio is high or not. Our proposed method can estimate Wi-Fi throughput accurately, and it improves the throughput of the Wi-Fi system and contributes to cellular data offload.

ACKNOWLEDGMENT

This research was conducted under a research contract organized by the Ministry of Internal Affairs and Communications, Japan.

REFERENCES

[1] David Malone; Ken Duffy; Doug Leith; , "Modeling the 802.11 Distributed Coordination Function in Nonsaturated Heterogeneous Conditions," Networking, IEEE/ACM Transactions on , vol.15, no.1, pp.159-172, Feb. 2007

[2] Fukuda, Y.; Oie, Y.; , "Decentralized access point selection architecture for wireless LANs Deployability and robustness," Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th , vol.2, no., pp. 1103- 1107 Vol. 2, 26-29 Sept. 2004

[3] Fukuhara, T.; Yamaguchi, A.; Kashiki, K.; Suzuki, T.; Takeuchi, K.; , "Channel occupancy estimation method for Wireless LAN," Personal, Indoor and Mobile Radio Communications, 2009.