A Mobile Wi-Fi Access Point Considering Types and Volumes of Traffic

Yuto Kobayashi, Junji Takahashi, and Yoshito Tobe

***Abstract*— In recent years, due to the spread of mobile terminals equipped with wireless LAN functions, various studies to improve the communication environment of a wireless LAN Access Point (AP) have been conducted. Though many studies have assumed that the AP is fixed, we have already proposed a system that designs a movable AP and dynamically optimizes it according to the communication environment [1]. However, this system has a problem that the type of traffic generated by each user is not considered. Therefore, we propose a system CHASA which optimizes the communication environment by considering the type of traffic. In this paper we describe the design and implementation of CHASA and show the results of experiments.**

Keywords—access point; traffic;

I. Introduction

Wi-Fi is becoming inevitable communication infrastructure for our daily lives. In addition to the popularization of mobile terminals in recent years, the spread of streaming services such as moving pictures and music, the stability of communication over wireless LAN is required. IEEE802.11 formulated by IEEE as a wireless LAN standard exists, and it is used for many wireless LAN products. At present, the most popular standard is IEEE802.11n developed in 2009, and it became possible to speed up by channel bonding and MIMO.

Many techniques have been studied for speeding up communication by wireless LAN, and the communication speed has been dramatically improved. However, in order to make comfortable communication with wireless LAN, it is also necessary to consider connection status such as user's location and traffic. However, in the present situation, since the AP is fixed at a fixed position, it cannot cope with the change of the position of the user.

We have proposed a mobile AP designing system which dynamically optimizes according to the communication environment. However, this system has a problem that the type of traffic generated by the user is not considered. Introducing a mechanism that takes into consideration the type of traffic makes it possible to optimize considering the characteristics of traffic types. Therefore, in this research, we propose a system CHASA that optimizes the communication environment by considering the type of user's traffic.

Y. Kobayashi is with Graduate School of Science and Engineering, Aoyama Gakuin University. Sagamihara, Japan. (email: yuto@rcl-aoyama.jp).

J. Takahashi and Y. Tobe are with Department of Integrated Information Technology, Aoyama Gakuin University. Sagamihara, Japan. (email: takahashi@it.aoyama.jp, yoshito-tobe@rcl-aoyama.jp)

The rest of the paper is organized as follows. Section 2 describes related work. Sections 3 and 4 explain the design and implementation of CHASA, respectively. Section 5 describes the results of experiment.

II. Related Works

There have been many trials on investigating the type of Internet traffic. The first application in the early age focused on intrusion detection[2]. It is crucial to detect anomaly in the network traffic. However, classification at the beginning utilized only simple methods with TCP/UDP port numbers. As the machine learning (ML) techniques matured, ML was applied to classify the type of traffic based on time-series signals. Dewes et al.[3] detected a chat session by using the characteristics of the traffic in terms of flow duration. Witten et al. [4] applied ML to clustering of traffic.

More recent work utilizes deep neural networks (DNN). The DNN is becoming popular in many fields such as computer vision[5], speech recognition[6], and language processing[7]. In ImageNet Challenge 2012, Hinton et al.[8] exhibited error rate as low as 15%. Thus, DNN can be expected to be used in network-traffic classification. Some work uses reinforcement[9] in learning network traffic.

Our work utilizes a classification technique, but our objective is the determination of the suitable point of a WiFi AP.

In the research field of WiFi networks, Hamaguchi et al.[10] considers a deployment of WiFi APs. Unlike our work, they do not assume mobility of APs.

III. Design

This section describes the assumption of the system and explains how we design the traffic-type-aware movable AP.

A. Definition

MA：Movable AP

Ci：the i-th client

N　: the number of clients

Wi：Weight by traffic type

Tu：Optimum position update interval [s]

B. Overview

CHASA is a system that dynamically updates the optimum location considering the traffic of all clients connected to the MA. The MA has the following three functions.

1) Movement：

The MA always moves to the optimum position and stops when it reaches the optimum position. The optimum position is updated every *Tu* time. When the optimum position changes, the MA starts moving again. In determining whether or not the MA has reached the optimum position, the MA prevents itself from fluctuation around the optimum position using a certain threshold value.

2) Update of optimum position：

The optimum position is calculated by finding the optimal solution of the objective function with the weights by type of throughput and traffic as variables. CHASA uses the hill climbing method to solve the optimal solution. The optimum position is updated at the frequency of 1/*Tu*.

3) Classification of Traffic：

The traffic type of each client is classified by using the feature quantity of the traffic flow: traffic generated during *Tu*. We use decision tree for classification. The result of classifying traffic when obtaining the optimum position is used. As a result, the MA approaches clients that emit traffic with high importance.

C. Equations

The optimum position of MA is the position where the sum of the throughput of each client becomes the largest. In addition, the throughput is multiplied by the weight of each type of traffic in order to make importance different depending on the type of traffic. The weight is a variable parameter and it is necessary to adjust it as the importance of the traffic. Considering the above, the objective function for finding the optimal position of CHASA is given by the equation (1).

$$\max \sum_{i=1}^{N} W_i Th_i \qquad (1)$$

Let Th_i be the communication throughput of Ci. It is known that Th_i fluctuates depending on the position of MA and attenuates in proportion to the square of distance in free space. Therefore, Th_i is a function of the distance between MA and Ci. Let the distance between MA and Ci be *di*.

In actual environments, it can be considered that the attenuation rate changes as compared with that in free space. Therefore, a model formula expressing the relationship between *di* and Th_i was created based on experiments.

$$Th_i = k_i d_i^{\,-\alpha} + e \qquad (2)$$

where, α is a variable parameter that varies depending on the receiving terminal.

D. Update of optimum position

In CHASA, Eq. (1) is used to find the optimum position, but the equation is not straightforwardly solves since it is nonlinear. Therefore, we use an algorithm for obtaining an approximate value. There are various kinds of algorithms for finding approximate values such as hill climbing methods, Newton method, steepest descent method, etc. CHASA adopts a hill climbing method which can be processed at high speed.

The hill climbing method initially evaluates the neighborhood point of the initial value and sets that point as the optimum position if it is higher than the initial value evaluation. Next, neighborhood points of the optimum position are evaluated, and when there are neighboring points with high evaluation value, the optimum position is updated. The process of update continues until there are no points with high evaluation value at the neighboring point of the optimum position, and the optimum position at the time of updating is set as the final optimum position. A pseudo code of the method is shown in Fig. 1. Here, the neighbor function returns the position around the current position and is the objective function of (1).

```
bestEval ← −∞
  for x in neighbor(currentPosition)
   if(eval(x) > bestEval)
      bestEval ← eval(x)
        bestPosition ← x
```

Fig.1 pseudo code

The approximate solution obtained by using the hill climbing method is the optimum position. By updating this optimum position at intervals of Tu, MA can always move to the optimum position of the current environment.

E. Classification of traffic type

Since CHASA weights according to the importance of each type of traffic, the type of traffic via the MA needs to be classified. Many methods for classifying traffic types have already been proposed, and many of them use information in IP packets as feature quantities. A representative method using IP packet information is shown below.

1) Port number: A well-known port is used to classify traffic types. For example, port 80 is allocated to HTTP. Although it is a very simple classification method, since many of the new protocols are not assigned port numbers and HTTP is used for various services, this method makes it difficult to correctly classify traffic characteristics precisely.

2) Signature-based traffic identification: This is a method of using the payload of a packet as a feature amount. The payload has a property in such a way that the pattern of the bit string of the payload except the header is similar if the same protocol is used. Therefore, traffic can be classified by pattern matching of the payload.

3) Statistical features and machine learning: Statistical methods and machine learning can classify traffic by using *various features of traffic flow. A Traffic flow is a series of packets from a specific source to one destination. Features often used at this time are shown below.*

- Total number of packets
- Average packet size
- Standard deviation of packet size
- Average packet arrival interval
- Standard deviation of packet arrival interval
- TCP packets rate
- UDP packets rate

In the traffic classification of CHASA, a traffic classification method by machine learning which has high classification accuracy and can correspond to a new protocol is used. Traffic classification is done by using the above-mentioned feature quantity of traffic flow during Tu. In my study, Tu is set to 10 [s].

Traffic types to be classified are categorized into four: video, VoIP, Web, Others. Video is a data transfer of movies and traffic of streaming service. VoIP is traffic of voice transmission and reception using VoIP, and Web is traffic generated at web browsers and servers. Table 1 shows weights for each traffic type used in CHASA. Since streaming service and VoIP traffic require much bandwidth or severe delay constraints for a high quality service is, weight of Video and VoIP is set high.

Table.1

symbol	Traffic type	Weight
W0	Video	2.5
W1	VoIP	2.2
W2	Web	1.0
W3	Others	0.2

There are many classification methods using machine learning such as decision trees, naive Bayes classifier, neural network, etc. In CHASA, a decision tree is used. The advantage of the decision tree is that the classification rule is easy to be visually understood. Furthermore, since classification can be performed at high speed, a large amount of packets can be classified without delay. We use a machine learning library called Scikit-learn to generate decision trees. Scikit-learn uses CART optimized for machine learning. In the CART algorithm, the Gini-coefficient is used as a criterion of classification. It takes a value from 0 to 1 and is a measure representing balance / imbalance. When the Gini coefficient is 0, it cannot be classified any more.

Learning is performed by using labeled learning data already having the correct answer, and a classification model is generated. The learning data is packet data with original label collected by performing packet capture from actual traffic. The traffic generated above is traffic of classification class such as streaming movies, web browsing, IP calling.

F. Classification model

The classification model actually obtained by the CART algorithm is shown in Fig.2.

```
class = null
if MeanLength <= 796.481
   if UDPrate <= 0.7818
      if MeanInterval <= 0.659
         class = “Web”
      else
         class = “Others”
   else
      class = “VoIP”
else
   class = “Video”
```

Fig.2. Classification model by decision tree

The output result of the decision tree consists of nodes and leaves. The node represents a branch condition for classification, and the leaf represents the final classification result.

The feature quantities used in the node in Fig.2 are shown below.

Mean Length：Average packet length

UDP rate：Percentage of UDP of all packets

Mean Interval：Average packet arrival interval

IV. IMPLEMENTATION

MA is built with RaspberryPi. This OS uses Raspbian ver.4.1. In order to operate RaspberryPi as AP, WLI-UC-GNM2 which is a wireless LAN adapter is used, hostAPD and bridge-utils are installed and set to operate as AP. The hostAPD is an application for running the wireless LAN adapter as an AP. The specification of MA is shown in Table.2.

Table.2. The specification of MA

Model	RaspberryPi model B
CPU	ARM1176JZF-S 700MHz Single core
OS	Raspbian ver.4.1
Memory	512 MB
Network adapter	WLI-UC-GNM2
Wireless LAN standard	IEEE802.11n

In RaspberryPi, General Purpose Input / Output (GPIO), motor, motor driver and power supply are connected to form a circuit. For the motor, use TAKARATOMY 's Plarail S - 23 E

257 Series AZSHA and connect the motor driver TA7291P. CHASA can use three commands: moving forward, moving backward, and stopping. The created MA is shown in Fig.3.

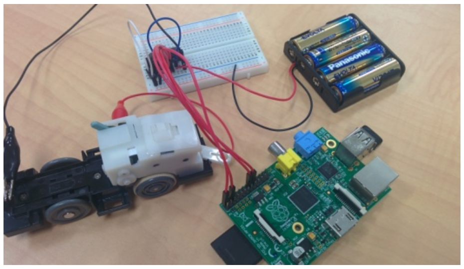

Fig.3 A structure of the MA

V. Experiment

A. Environment

In order to evaluate the performance of CHASA, four terminals are prepared to generate different types of traffic, and the total throughput in the case of using CHASA at this time is compared with the overall throughput. In order to make a difference in traffic generated at each terminal, the terminal 1 performs Web browsing, the terminal 2 does not operate, the terminal 3 performs IP call, and the terminal 4 plays the streaming video. Table 3 shows the terminals used.

Table.3. Specification of terminal used for experiment

Terminal	Model	OS
Terminal1	Nexus5	Android6.0
Terminal2	Nexus5	Android6.0
Terminal3	iPad mini 4	iOS10.0
Terminal4	Let's note	Windows7

a) Web browsing: This corresponds to daily operations such as search and browsing using a Web browser. Traffic occurs when searching or clicking a link.

b) No operation: Since no operation is performed, basically no traffic occurs.

c) IP telephony: This corresponds to an IP telephony such as Skype. The call is a one-to-one voice only. Due to the nature of the call, constant traffic always occurs, but the amount of traffic is relatively small.

d) Streaming video: This corresponds to watching YouTube live streaming delivery videos on a web browser. The amount of traffic is large, and traffic is continuously generated.

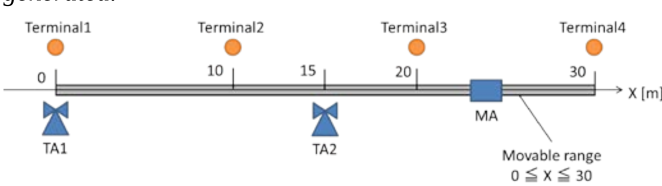

Fig.4. Placement of AP and terminal

The value of α, which is a variable parameter varying depending on the receiving terminal, is 0.15 for terminals 1 and 2, and 0.05 for terminals 3 and 4.

In the experiment method, each terminal is first connected to the AP to be evaluated, and traffic is generated. On the AP side, traffic from each terminal is monitored and communication throughput is measured. This experiment is conducted with three types of AP (MA, TA1, TA2) and the results are compared. In addition, TA1 and TA2 are conventional fixed APs and will not move from the initial position. The performance of all three APs is the same. The positions of each AP and each terminal are shown in Fig 4. The initial position of MA is 15 m and the movable range is 30 m.

B. Result

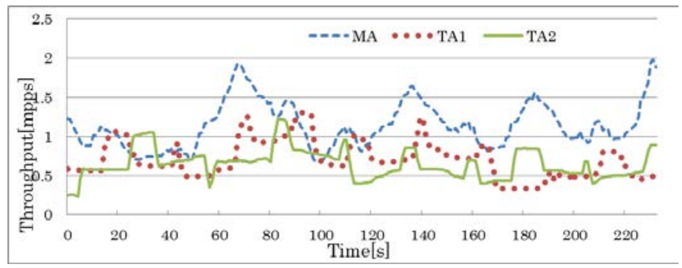

Fig.5. Comparison of throughput of each AP [mbps]

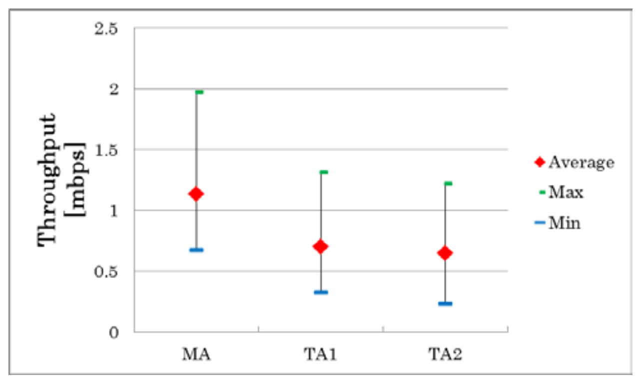

Fig.6. Average, maximum, minimum of throughput of each AP

The transition of the throughput measured by each AP is shown in Fig.5. However, the throughput at this time is moving average in the range of 10 seconds before and after, and it is smoothed. Also, the average value, the maximum value, and the minimum value of the communication throughput of each AP are shown in Fig.6. As observed in the graph, we see that the throughput of MA is higher than TA1, TA2. The reason for this result is that as the MA approaches a terminal that generates traffic requiring a lot of bandwidth, such as IP calls and streaming movies, packet loss due to the distance of wireless communication is reduced and efficient communication It was thought that it was possible to do. In addition, when comparing TA 1 and TA 2, it is understood that

TA 2, which is close to the terminal performing the IP call or streaming moving image, has lower average throughput than TA 1. The throughput should be higher for TA 2, which is closer to a terminal that needs more bandwidth if originally, but in this case, TA 1 is close to terminal 1 as much as possible, and because it is not a MIMO environment but an antenna It is thought that this is because it is susceptible to interference.

VI. CONCLUSION

In this study, we have proposed a system CHASA which optimizes the communication environment by considering the type of traffic. By introducing a mechanism that takes into consideration the type of traffic, priority can be set for each communication of the client terminal, and optimization can be performed according to traffic characteristics. In addition, it was found that by actually comparing the case of using CHASA with the case of not using CHASA, the throughput is higher when CHASA is used and the communication efficiency is improved.

The MA proposed in this research does not correspond to the MIMO environment and is not a directional antenna. Therefore, beamforming is not used. Since the expression of communication throughput used in the optimization formula of CHASA assumes that the environment with one antenna without directivity, when there are multiple antennas or when there is a directional antenna, the throughput can change. Currently, even small terminals such as smartphones are equipped with multiple antennas, and thus MIMO environments are expected to increase in the future. Based on the fact, it is considered that the communication environment can be further improved by applying MIMO and beam forming to CHASA.

REFERENCES

[1] Y. Kobayashi, J. Takahashi and Y. Tobe, “Measurement of WiFi Traffic for a Movable Access Point”, SICE Annual Conference, 2016.

[2] V. Paxson, “Bro: A system for detecting network intruders in real-time,” Computer Networks, no. 31(23-24), pp. 2435?2463, 1999.

[3] C. Dewes, A. Wichmann, and A. Feldmann, “An analysis of Internet chat systems,” in ACM/SIGCOMM Internet Measurement Conference 2003, Miami, Florida, USA, October 2003.

[4] I. Witten and E. Frank, Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations (Second Edition). Morgan Kaufmann Publishers, 2005.

[5] Le Q V. "Building high - level features using large scale unsupervised learning," Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on. IEEE, 2013:8595 - 8598.

[6] Zeiler M D, Ranzato M, Monga R, et al. "On rectified linear units for speech processing," Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on. IEEE, 2013: 3517 - 3521.

[7] Socher R, Pennington J, Huang E H, et al. "Semi - supervised recursive autoencoders for predicting sentiment distributions," Proceedings of the Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2011: 151 - 161.

[8] Krizhevsky A, Sutskever I, Hinton G E. "Imagenet classification with deep convolutional neural networks," Advances in neural information processing systems. 2012: 1097 - 1105.

[9] Li Li,Yisheng Lv,Fei-Yue Wang. "Traffic Signal Timing via Deep Reinforcement Learning," IEEE/CAA Journal of Automatica Sinica, 2016, 3(3): 247-254.

[10] T. Hamaguchi, T. Komata, T. Nagai, and H. Shigeno, "A Framework of Better Deployment for WLAN Access Point Using Virtualization Technique," IEEE 24th Int. Conf. on Advanced Information Networking and Applications Workshops (WAINA), 2010.