Interference-free AP identification and shared information reduction for tabular Q-learning-based WLAN coordinated spatial reuse

Yuto Kihira¹, Koji Yamamoto¹a), Akihito Taya², Takayuki Nishio³, Yusuke Koda¹, and Kazuto Yano⁵

¹ Graduate School of Informatics, Kyoto University
Yoshida-honmachì, Sakyo-ku, Kyoto, 606-8501 Japan
² College of Science and Engineering, Aoyama Gakuin University
5-10-1 Fuchinobe, Chuo-ku, Sagamihara-shi, Kanagawa, 252-5258, Japan
³ School of Engineering, Tokyo Institute of Technology
Ookayama, Meguro-ku, Tokyo, 152-8550, Japan
⁴ Centre for Wireless Communications, University of Oulu
90014 Oulu, Finland
⁵ Wave Engineering Laboratories, Advanced Telecommunications Research Institute International
2-2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan

Abstract: Access point (AP) coordinated spatial reuse with Q-learning enables efficient spectrum utilization [1]. Although sharing of transmission schedules among APs is necessary for coordination, there is no mechanism to identify the APs with which the schedules are to be shared, resulting in excess information being shared among APs. In this study, we propose a scheme to identify the interference-free APs that are not required for sharing of information by comparing Q-values. A simple simulation demonstrates that this scheme successfully reduces shared information without throughput degradation.

Keywords: WLAN, coordinated spatial reuse, reinforcement learning, dimensionality reduction

Classification: Wireless communication technologies

References

[1] Y. Kihira, Y. Koda, K. Yamamoto, T. Nishio, and M. Morikura, “Adversarial reinforcement learning-based robust access point coordination against uncoordinated interference,” Proc. IEEE VTC2020-Fall, Victoria, BC, Canada, pp. 1–5, Dec. 2020. DOI:10.1109/VTC2020-Fall49728.2020.9348462

[2] D. López-Pérez, A. García-Rodríguez, L. Galati-Giordano, M. Kasslin, and K. Doppler, “IEEE 802.11 be extremely high throughput: The next generation of Wi-Fi technology beyond 802.11 ax,” IEEE Commun. Mag., vol. 57, no. 9, pp. 113–119, Dec. 2019.
1 Introduction

With the widespread prevalence of wireless local area networks (WLANs), there is increasing competition for the limited transmission opportunities. To identify the factors responsible for frame loss under the situation, in our previous study, we focused on the application of access point (AP) coordination for spatial reuse [1]. In the proposed framework, APs share their transmission schedules beforehand (e.g., via backhaul links). By utilizing the transmission schedules and transmission results, the AP can identify the factors responsible for frame losses, specifically, the interference from other transmitters or low received-signal power. Depending on these factors, the AP determines its own transmission schedules and transmission data rate using Q-learning. As a result, high spectrum efficiency is achieved.

However, since the APs do not know which APs the transmission schedules should be shared with, the APs may unnecessarily share the schedules with all surrounding APs. If the shared schedules contain information that is redundant for AP coordination, the traffic for information sharing is needlessly large, which should be avoided for efficient operation of coordinated spatial reuse. This problem was not discussed in [1].

To solve this problem, we propose a method to reduce the amount of shared information. In this method, we focus on the fact that the transmissions from interference-free APs have almost no impact on the Q-values. This method identifies the interference-free APs by examining the Q-table and subsequently stops sharing the transmission schedules with those APs. As a result, the traffic for information sharing is reduced.

2 System model

The system model is almost the same as in our previous study [1], i.e., there are coordinated $N + 1$ basic service sets (BSSs) and each of them consists of an AP and a station (STA) with only downlink traffic. Note that this assumption can be easily extended to multi-STA scenarios. The BSSs perform coordinated time-division resource assignment in the same frequency band in periods for multi-AP coordination, which is under discussion for the IEEE 802.11be standard [2]. In this study, the unit of resource assignment is called a slot. Each BSS is numbered from $n = 0, 1, \ldots, N$, and the AP and STA in the BSS $n$ are named AP $n$ and STA $n$, respectively. The other APs randomly decide in advance whether or not to transmit for each slot and share the transmission schedules with AP 0. Note that we only consider slots for multi-AP coordination and the details have been discussed in [1].

AP 0 determines whether to transmit a frame as well as the transmission data rate (i.e., a modulation and coding scheme (MCS) index), according to the information on the transmission schedules of the other APs.

3 Determination of transmission timing by reinforcement learning

This section describes how AP 0 determines its transmission schedule and data rate. In this scheme, the agent uses Q-learning to decide its action as
At each learning step, AP $0$ observes a state that is determined by whether each surrounding AP is going to transmit or not. The state space is given as follows:

$$S := S_1 \times S_2 \times \cdots \times S_N, \quad S_i := \{0, 1\}, \ i \in N_1 := \{1, \ldots, N\}. \ (1)$$

Here, $S_i$ represents the set of each AP's condition. The state is set to be zero if the AP is going to transmit; otherwise, the state is set to be one. It is assumed that the conditions of APs are shared via backhaul link.

The agent decides whether or not to transmit; if so, it selects an MCS index. The action space of AP $0$ is given as follows:

$$\mathcal{A} := \{0\} \cup M, \ (2)$$

where $M := \{1, \ldots, M\}$ denotes a set of MCS indices and 0 denotes an action of not transmitting a frame. The respective number denotes which MCS index AP $0$ selects when transmitting a frame, and $M$ represents the number of the MCS indices.

If AP $0$ successfully transmits a frame at $x \text{Mbit/slot}$, the agent receives a reward of $x$. Otherwise, the agent is given a negative reward of $-1$ for a frame loss. Owing to this penalty, the agent attempts to avoid collisions with other transmissions. The agent selects actions by the $\epsilon$-greedy method.

### 4 Shared information reduction

Because the observation of state (1) requires exhaustive information sharing, this section describes how to reduce the volume of transmission schedules to be shared with other APs. This reduction is achieved by avoiding redundant information sharing with some of the APs.

The agent identifies the APs that can be excluded from the information sharing by examining the values of the Q-table during the learning process as follows. After some learning, the agent extracts the parts of the Q-table for a state with only one 1 in the state vector, for example, $s = (0, \ldots, 0, 1), (0, \ldots, 0, 1, 0), \ldots, (1, 0, \ldots, 0)$. Here, we define $s_i$ as the vector whose $i$th element is 1 and the others are 0 ($i \in N_1$). Among them, the agent focuses on the states whose Q-values are close to the values of $s_0 := (0, 0, \ldots, 0)$. The fact that the Q-values are close for $s_0$ and $s_i$ implies that the transmissions of AP $i$ have almost no impact on the communication of AP $0$. Because the information from these APs is not necessary for efficient management, the agent can ask these APs to stop sharing information and reduce the number of dimensions of the state space.

In this study, AP $0$ stops sharing information with AP $i$ upon satisfying the following inequality:

$$\left| \frac{Q(s_i, a) - Q(s_0, a)}{Q(s_0, a)} \right| \leq \beta, \ \forall a \in \mathcal{A}, \ (3)$$

where $Q(s, a), s \in S$ and $a \in \mathcal{A}$, denotes the Q-value, and $\beta \geq 0$ denotes a parameter. Then, we redefine $S$ by removing $S_i$ from (1). As a result,
this method can reduce the shared information that AP 0 has to receive. In addition, the size of Q-table can be reduced by decreasing the dimensionality of the state space. This results in faster learning and less data storage.

To comply with reduced state space, the Q-table is downsized by repeatedly averaging over the axis. First, if \( i \) satisfies (3), the Q-table is reduced by the following operation:

\[
Q(s_{\text{after}}, a) := \frac{1}{2} (Q(s_{\text{before}0}, a) + Q(s_{\text{before}1}, a)) \quad \forall a \in \mathcal{A},
\]

where \( s_{\text{before}0} \) and \( s_{\text{before}1} \) denote the states before the reduction. They are defined as follows:

\[
s_{\text{before}0} := (e_1, e_2, \ldots, e_{i-1}, 0, e_{i+1}, \ldots, e_{N-1})
\]

\[
s_{\text{before}1} := (e_1, e_2, \ldots, e_{i-1}, 1, e_{i+1}, \ldots, e_{N-1}) \quad e_j \in \{0, 1\}, \ j \in \mathcal{N}_1 \setminus \{i\},
\]

where \( s_{\text{after}} \) denotes the state after the reduction and is represented as follows:

\[
s_{\text{after}} := (e_1, e_2, \ldots, e_{i-1}, e_{i+1}, \ldots, e_{N-1}) \quad e_j \in \{0, 1\}, \ j \in \mathcal{N}_1 \setminus \{i\}.
\]

Note that \( s_{\text{after}} \) has one less dimension than \( s_{\text{before}0} \) and \( s_{\text{before}1} \). By repeating the same operation, the size of Q-table is reduced in size. If information sharing with \( n \) APs is stopped, the size of the Q-table is reduced by a factor of \( 2^n \).

5 Evaluation

To facilitate understanding, we simplify the relationship between the transmission data rate and the distance between APs. A set of MCS indices is denoted as \( \mathcal{M} = \{1, 2, 3\} \), where 1, 2, and 3 denote transmissions at 1, 2, and 3 Mbit/slot. The distances between APs are assumed to be of four levels: 1, 2, 3, and “no interference”. Distance \( d \) is the distance at which transmission fails at a transmission rate of \( d \) Mbit/slot or higher when the other AP is transmitting. Additionally, “no interference” means that these communications have no effect on each other. The distance between AP 0 and the surrounding six APs (i.e., \( N = 6 \)) are shown in Table I.

Table I. Setting of distance of APs and the result of interference-free AP identification; values below \( \beta \) are shown in bold, meaning they were identified as interference-free APs.

| Index of APs | 1 | 2 | 3 | 4 | 5 | 6 |
|--------------|---|---|---|---|---|---|
| Distance from AP 0 | 2 | 2 | 3 | (Interfering) | (Interference-free) |
| At 100,000th slot |
| LHS in (3) | 1.5 | 1.3 | 1.3 | 0.0007 | 0.0009 | 0.0002 |
| Decision | Interfering | Interference-free |
We compare the proposed scheme with the following two baselines. *W/o reduction scheme* continues to share transmission schedules from all APs forever. This scheme does not reduce the shared information. *W/o sharing scheme* does not share the transmission schedules with other APs. This scheme corresponds to the case of not sharing transmission schedules. As in the proposed scheme, this scheme uses Q-learning with $\epsilon$-greedy action selection rules.

Fig. 1. Maximum value of LHS in (3) of AP 0 for all actions.

Fig. 1 shows the maximum values of LHS in (3) of AP 0 for all actions. We can see that the values for $s_4$, $s_5$, and $s_6$ converge to zero. In this evaluation, identification of interference-free APs and consequent shared information reduction are conducted at 100,000th slot with $\beta = 1/3$. As summarized in Table I, from six APs, three interference-free APs 4, 5, and 6 are successfully identified. Thus, the amount of information-shared APs and the shared information are reduced by half.

Fig. 2 shows the simulation results of the proposed and comparison schemes. It presents the learning curves of the throughput of AP 0. The vertical axis represents the amount of data successfully transmitted by AP 0 in 50 slots. Despite the shared information reduction, we can see that there is no performance gap between the proposed scheme and *w/o reduction scheme*, thus the proposed scheme does not reduce the throughput. Note that how much the shared information can be reduced depends on radio environment and BSS density. According to Fig. 2, the throughput of *w/o sharing scheme* is lower than that of the other two schemes. This indicates that the information sharing with surrounding APs contributes to the improvement of the throughput.
6 Conclusion

We successfully identified the interference-free APs. By excluding these APs, this method can achieve a reduction in the shared information for coordinated spatial reuse.

Acknowledgments

This research and development work was supported in part by the MIC/SCOPE #JP196000002 and JSPS KAKENHI Grant Number JP18H01442.