Study of Contention Window Adjustment for CSMA/CA by Using Machine Learning

Yen-Wen Chen
Communication Engineering Department
National Central University
Tao-Yuan City, Taiwan
ywchen@ce.ncu.edu.tw

Kuo-Che Kao
Communication Engineering Department
National Central University
Tao-Yuan City, Taiwan
sora811010@gmail.com

***Abstract*— In IEEE 802.11, CSMA/CA protocol applies the exponential backoff scheme to relax the contention problems among different clients wishing to transmit data at the same time. A client shall randomly choose a number of time slots bounded by the contention window. As the size of initial contention window is fixed for each device without considering the congestion status of the network, it may worsen the congestion condition for smaller contention window size, or may waste radio resource for larger window size in traditional scheme. In this paper, we propose the reinforcement learning model rewarded by throughput to dynamically adjust the contention window periodically. The Q-learning model is utilized in the proposed scheme and the reward function is determined by the current throughput measured during a certain interval and which measured in previous interval. The simulation results were compared with the rule based approach used in current 802.11 networks. The results show that the proposed scheme can effectively decrease the collision rate and the system throughput increases significantly accordingly. We also change the number of clients during the simulation time to examine the adjustment capability of the proposed scheme. The results also show that the proposed scheme is still superior to the rule based scheme in such environment.**

***Keywords*— CSMA/CA, Contention Window, Collision, Reinforcement Learning, Throughput.**

I. Introduction

With the rapid development of the Internet of Things (IoT), it is estimated that there will be more than 40 billion devices by 2025. Although the IoT device does not require large bandwidth, the number of network accesses is much higher than that of other services due to the large amount of IoT devices and frequent short data uplink transmission. The uplink access behavior of IoT services is quite different from that of traditional human oriented network services. The intensive uplink short data transmissions greatly increase the collision probability. Therefore, to resolve collisions among huge number of devices is one of the critical issues that affect the transmission performance. Under the IEEE 802.11 (Wireless LAN) standard, the Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) was proposed to deal with the contention problem in IEEE 802.11 wireless local area network. Its main concept is to randomly choose the back-off time by referring to the contention window. In 802.11, the initial value of contention window of each device is fixed to be the smallest and will enlarge exponentially for each collision [1]. The device set its initial contention window regardless of actual transmission performance of the system. The rule based contention window adjustment scheme without transmission performance feedback may result in either high collision rate when the number of transmission devices is large or the low radio resource utilization if the number of transmission devices is small. The data transmission behavior of IoT devices is quite different from the traditional network services, the wireless access network shall provide more flexible and rapid adjustment scheme to satisfy various IoT applications. In this paper, we try to deal with the adjustment of contention window in an alternative way. The reinforcement learning scheme is applied to dynamically adjust the initial value of the contention window by referring to the system throughput. The reinforcement learning process is basically a feedback based approach to gradually improve system performance, which is different from the traditional rule based scheme. The simulation results of the proposed scheme are compared to that of the traditional rule based scheme to investigate the improvement of the collision rate as well as the system throughput. We also examine the adjustment capability by changing the number of clients during simulation time to investigate the efficiency of the proposed algorithm.

II. Background and Related Works

802.11 provide two modes in MAC layer, one is point coordination function (PCF) and the other is distributed coordination function (DCF). PCF is designed for contention-free service, and DCF is designed for contention service which achieve by CSMA/CA mechanism. In order to maintain the transmission quality, the avoidance of collision is the critical issue toward effective transmission. In 802.11, the device randomly chooses the waiting time before its transmission to avoid collision in probability way. And the random choosing interval is determined by the contention window (CW) as [0, CW-1]. The initial CW of each device is fixed and it will be double for each collision till the maximum CW reaches. Therefore, if the contention window is small with excessive transmission devices, then the collision probability increases. However, if the contention window is larger with less number of contending devices, then more idle time will be left and more radio resource will be wasted.

Several studies were proposed to improve the performance of CSMA/CA [2-5] in either throughput or fairness. In [2], the authors proposed the Adaptive Contention Window Backoff (ACWB) to replace the original Binary Exponential Backoff algorithm (BEB) to improve the system throughput while maintaining the original fairness. The Dynamic Contention Window Adjustment (DCWA) scheme was proposed in [3] to allow each node dynamically adjusts the size of its CW according to the observed channel average idle so as to improve throughput. In [4], the idle sense scheme was proposed to enable each node to observes the mean number of idle slots between transmission attempts to dynamically control its contention window. Their results show that the system throughput can be improved and can achieve time-fair channel allocation. In [5], the authors considered the traffic loads of uplink and downlink to dynamically adjust the minimum contention window for the fairness between uplink and downlink flows.

Comparing to the above rule based scheme, the machine learning approach provides a more heuristic way to make its decision by referring to the experienced data from the target application. Several researches were proposed to deal with the improvement of wireless communication systems in changeable condition [6, 7, 8, 9]. In [6], the concept of intelligent edge was proposed to support the artificial intelligent (AI) enable applications. The authors suggested design guidelines for wireless communication in edge learning. Generally, the random access is one of the critical issues that affect the service performance for massive end devices such as IoT. The potential contributions of machine learning in Beyond 5G (B5G), especially for massive end devices, were studied in [7]. In [8], the authors proposed a distributed method to dynamically allocate RA lots to machine type communication (MTC) devices in Non-Orthogonal Multiple Access (NOMA) 5G network. And the Q-Learning based approach was applied to improve network throughput. There is no exact correct policy or rules to get the optimal performance of RA in practical communication networks because of the changeable factors caused by the number of end devices and their accessing behaviors.

III. The Proposed Learning Model and Window Adjustment Algorithms

We assume that there are M devices covered by an access point (AP) without hidden node problem. Each device will start its countdown after a DIFS time if there is no other device that is transmitting in this time period. The contention window CW is determined by AP through broadcast to all device for the periodical time T.

The Q-learning model (QL) is applied to determine the contention window size (CWS). In QL model, an agent will be set at beginning of learning process, and observe environment to select the most suitable action from a finite set. The agent then obtains reward from environment and updates the Q-value in Q-table where each Q-value means the quality of an action in a state. We can express the method of updating Q-value which called Q-function as

$$Q(s,a) \leftarrow (1-\propto)Q(s,a)+\propto [R + \gamma max_{\alpha} Q(s',a') - Q(s,a)]$$

where $Q(s,a)$ means Q-value of action a in current state s, $max_{\alpha}Q(s',a')$ means the maximum Q-value at next state s', R means reward of perform action a in current state, $\propto$ means learning rate, γ means discount factor. In Q learning model, we first set agent in initial state, it will choose an action a to change state from s to s' and update Q value $Q(s,a)$ by above formula. We can overserve that the greater learning rate α, the less effective of previous training will be retained, and the $max_{\alpha}Q(s',a')$ have greater effect if discount factor is greater by the above formula. The $max_{\alpha}Q(s',a')$ means the benefit in memory, it is the best Q value in state s'. If the agent gets benefit at a certain action which in state s' at the past training episode, the above formula will reflect the benefit in current training episode.

The proposed model uses contention window size as the system state and is given a Q value as $Q(CWS)$. Then we can rewrite the above equation to the following equation

$$Q(CWS) \leftarrow (1-\propto) \times Q(CWS)+\propto [R - Q(CWS)]$$

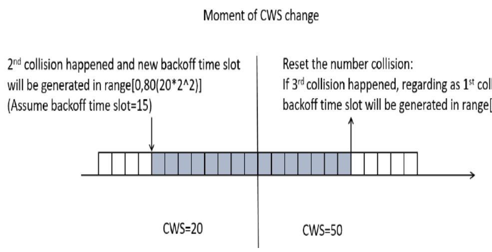
AP will observe and record current throughput for a period of time (T_c), and compare with previous throughput(T_p). If current throughput is better than previous throughput, set the reward to 1, otherwise it is -1. Thus the proposed algorithm adjust the suitable CWS by learning from the change of throughput and the applied policy. Exploration and exploitation mechanism also applied in this model to prevent the local sub-optimal problem. We set the epsilon-greedy parameter as ε. The proposed algorithm is illustrated in the Figure 1.

Algorithm 1: Dynamic contention windows selection algorithm

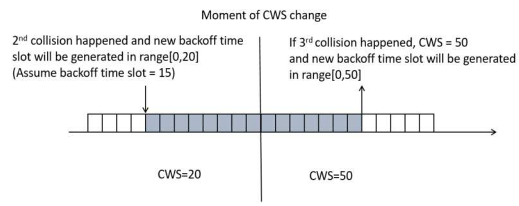
```
Initialize Q(CWS) , α , ε, past throughput Tp
foreach episode do
    if number of client change then
        Initialize α, ε;
    Choose CWS using policy derived from Q (e.g., ε-greedy);
    Take action CWS, observe current throughput Tc;
    if Tc>Tp then
        R=1;
    else
        R=-1;
    end if
    Q(CWS) ← (1 − α)Q(CWS) + α[R−Q(CWS)];
    Tp ← Tc;
    update α, ε;
end foreach
```

Figure 1 The Proposed Algorithm

According to the above algorithm, Then AP will select contention window size (CWS) by the exploring rate and the observed throughput for a period of time to determine reward. Finally, update Q-value by reward. The client determines its backoff time slots by referring to the CWS periodically broadcasted by AP, however, two modes for backoff selection are proposed in the algorithm. The client will randomly choose the backoff time slots within broadcasted for the initial transmission in both modes. And, in model 1, if collision is occurred, the client will double the received CWS as the range for random selection. Thus the backoff time slots will be randomly chosen within [0. 2*CWS]. The mode 2 adopts more aggressive approach to decide the backoff time slots. The client will randomly choose the backoff time slots within [0, CWS] without double the received CWS in mode 2. The AP adjusts the CWS according to the achieved throughput and the associated applied policy. The objective of the proposed two modes, one tends to be conservative (mode 1) and the other tends to be aggressive (mode 2), are designed to compare the effect when collision is occurred. The examples of deciding the backoff time slots for mode 1 and 2 are illustrated in figures 2(a) and 2(b), respectively.



(a) Backoff Time Slots Decision in Mode 1



(b) Backoff Time Slots Decision in Mode 2

Figure 2 Examples of Backoff Selection for Mode 1 and 2

It is noted that, as AP will broadcast CWS periodically, the client will always apply the latest received CWS. As shown in the examples, the applied CWS by the client will be 50 because the CWS has changed from 20 to 50.

IV. EXPERIMENTAL SIMULATIONS

In order to examine the performance of the proposed algorithm, exhaustive simulations were performed. The simulation results of the proposed algorithm (including mode 1 and mode 2) are compared to the traditional rule based CW adjustment and random backoff decision scheme. The minimum and maximum CW of the traditional rule based scheme are assumed to be 7 and 255, respectively. The parameters used during the simulations time are illustrated in the following Table 1. The limit to double CWS of mode 1 is assumed to be 5.

Table 1 Simulation Parameters

Packet payload	8184 bits
MAC header	272 bits
PHY header	128 bits
ACK length	112 bits + PHY header
Channel Bit Rate	1 Mbit/s
SIFS	28 µs
DIFS	128 µs
Slot time	50 µs
Episode	1 sec

There are 128 actions to decide the CWS in the action set in the proposed QL based algorithm. The learning rate is assumed to be 0.8 at the beginning of training process and decreases 0.1 after every 10 episodes till 0.2. Similar to the learning rate, the explore rate ε decreases 0.2 after every 10 episodes till 0.2. The parameters of QL is Table 2.

Table 2 Parameters Adopted in QL model

Action CWS	128 actions (3,7,11,.......,511)
Learning rate ∝	0.8 to 0.2
Explore rate	0.8 to 0.2

Figure 3 shows the change of throughput over time for the number of clients being 50. It demonstrates that the proposed algorithm, either mode 1 or mode 2, has better performance than the rule based scheme. It also illustrates that there are some low throughput points, e.g. at the 33-second of mode 1 and the 40-second of mode 2, of the proposed algorithm. We find that those low throughput points are mainly due to the exploration of the QL model.

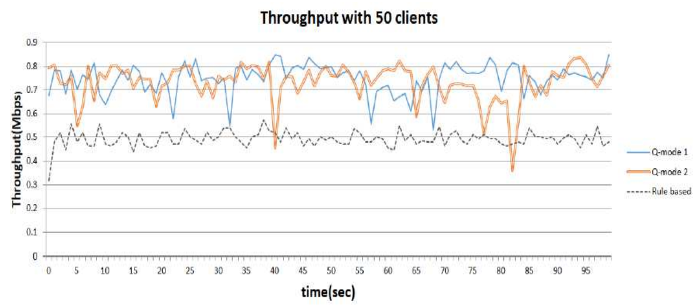


Figure 3 Throughput v.s. Time for 50 Clients

The throughput comparison of the proposed scheme and the rule based scheme is shown in Figure 4. The proposed scheme achieves higher throughput than the traditional rule based scheme. It also indicates the result of mode 1 is superior to the mode 2, especially when the number of clients becomes larger. The main reason is that mode 1 enlargers the chosen range (CWS) of backoff time slots when collision occurred to reduce the collision probability. Figure 5 provides the comparison of collision rate to illustrate this phenomenon.

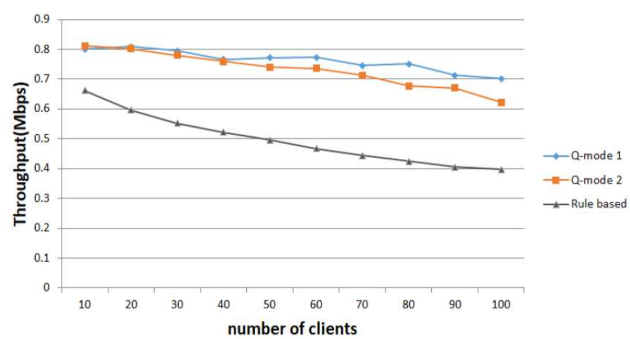


Figure 4 Comparisons of Throughput v.s. Number of Clients

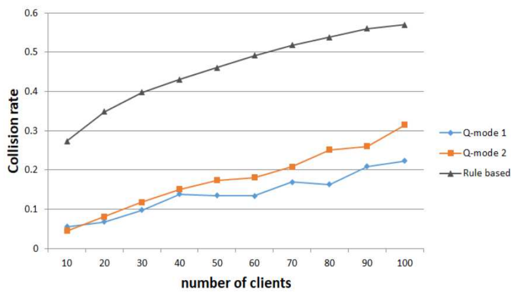


Figure 5 Comparisons of Collision Rate v.s. Number of Clients

Additionally, in order to examine the performance of the environment with changeable number of clients. During the simulations, we randomly change the number of client every 10 seconds. Figures 7 and 8 show the results of throughput and collision rate for the change of different numbers of clients. As shown in the figures, the initial number of clients is 8, and it changes to 48 after 10 seconds, and so on. According to the results, the proposed algorithm has higher and stable throughput than the traditional rule based scheme except when the number of clients is 1. The traditional rule based scheme has higher throughput when there is only one client. The main reason is that the minimum CWS is always applied for the client in the rule based scheme because no collision is occurred. The collision rate is zero when the number of clients is 1.

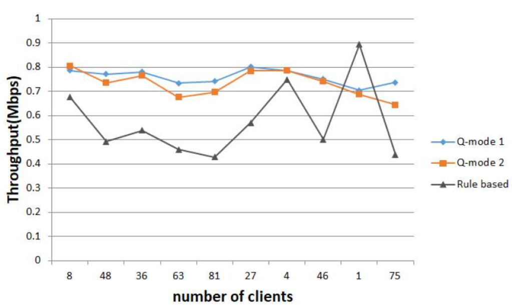


Figure 7 Comparisons of Throughput v.s. Changeable Number of Clients

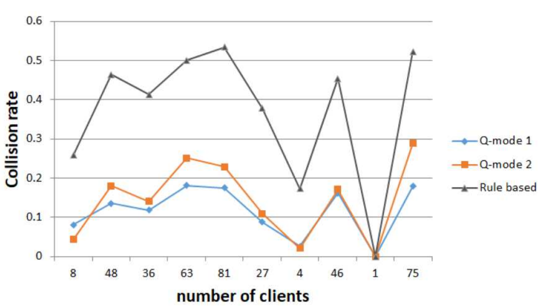


Figure 8 Comparisons of Collision Rate v.s. Changeable Number of Clients

V. Conclusions

The CSMA/CA has been adopted by IEEE 802.11 for a long time. The decision of CW and random backoff mechanism are also deployed by several WiFi based applications. In the traditional scheme, each client decides its own CW based on the collision status and generates its individual backoff time slots, however, the system throughput is not considered. In this paper, we try to deal with this issue from the throughput point of view to investigate the possible improvement when the machine learning approach is applied for the decision of CW. Two backoff slots decision modes were analyzed by referring to the proposed Q-learning based CWS decision algorithm. The simulation results show that the proposed algorithm can effectively reduce the collision rate and the system throughput can then be improved accordingly. As the number of clients may change from time to time in practical environment, we also examine the performance by letting the number of clients randomly changed during the simulation time. The results also illustrate that the proposed algorithm can achieve better performance than the traditional scheme. The CWS is determined by AP to inform the clients through centralized broadcast. Although it is quite different from the traditional scheme, in which the CWS is determined by each client in distributed manner, the proposed algorithm presents a preliminary results to show that the system throughput of WiFi can be improved. We think this study can provide reference for the associated protocol design in the future.

Acknowledgment

This work was supported in part by the Ministry of Science and Technology (MOST) (grant number: 109-2221-E-008 -052 -MY2), Taiwan.

References

[1] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," IEEE Journal on Selected Areas in Communications, vol 18, pp.535 – 547, Mar.2000.

[2] Xin Zhou, Changwen Zheng, Xiaoxin He, "Adaptive contention window tuning for IEEE 802.11," 2015 22nd International Conference on Telecommunications ,2015.

[3] Qin Yu, Yiqun Zhuang, Lixiang Ma, "Dynamic contention window adjustment scheme for improving throughput and fairness in IEEE 802.11 wireless LANs," 2012 IEEE Global Communications Conference, 2013.

[4] M. Heusse, F. Rousseau, R. Guillier, A. Duda, "Idle sense: an optimal access method for high throughput and fairness in rate diverse wireless LANs," SIGCOMM 2005, pp. 121-132, 2005.

[5] B. A. Hirantha Sithira Abeysekera, Takahiro Matsuda, Tetsuya Takine, "Dynamic Contention Window Control Mechanism to Achieve Fairness between Uplink and Downlink Flows in IEEE 802.11 Wireless LANs," IEEE Transactions on Wireless Communications, vol 7, pp. 3517 – 3525, Sep. 2008.

[6] G. Zhu, D. Liu, Y. Du, C. You, J. Zhang and K. Huang, "Toward an Intelligent Edge: Wireless Communication Meets Machine Learning," in *IEEE Communications Magazine*, vol. 58, no. 1, pp. 19-25, January 2020.

[7] M. E. Morocho-Cayamcela, H. Lee and W. Lim, "Machine Learning for 5G/B5G Mobile and Wireless Communications: Potential, Limitations, and Future Directions," in *IEEE Access*, vol. 7, pp. 137184-137206, 2019.

[8] Matheus Valente da Silva, Richard Demo Souza, Hirley Alves, Taufik Abrão"A NOMA-Based Q-Learning Random Access Method for Machine Type Communications," IEEE Wireless Communications Letters, vol 9, pp. 1720 – 1724, Oct. 2020.