A cross-layer adaptive algorithm for multimedia QoS fairness in WLAN environments using neural networks

C. Wang, T. Lin and J-L. Chen

Abstract: The authors address the problem of providing fair multimedia quality-of-service (QoS) in IEEE 802.11 distributed co-ordination function-based wireless local area networks in the infrastructure mode where mobile hosts experience heterogeneous channel conditions due to mobility and fading effects. It was observed that unequal link qualities can pose significant unfairness of channel sharing, which may thereby lead to the degradation of multimedia QoS performed in adverse conditions. A cross-layer adaptation scheme that provides fair QoS by online adjusting the multidimensional medium access control layer backoff parameters in accordance with the application-layer QoS requirements as well as the physical-layer channel conditions was proposed. The solution is based on an optimisation approach, which utilises neural networks to learn the cross-layer function. Simulation results demonstrate that the proposed adaptation scheme can tackle heterogeneous channel conditions and random joining (or leaving) of hosts to achieve fair QoS in terms of throughput and packet delay.

1 Introduction

With the popularity of IEEE 802.11 based wireless local area networks (WLAN) capable of providing high data rates ranging from 1 Mbps up to 54 Mbps, the demands of multimedia services for mobile users are growing. Various kinds of multimedia applications such as streaming video, Internet phone and net meeting, require differentiated quality-of-service (QoS) guarantees because of their differentiated traffic types. Thus, the provision of QoS in 802.11 medium access control (MAC) becomes increasingly important. Most of current 802.11 MAC employ distributed co-ordination function (DCF) [1], a random access protocol based on a carrier sense multiple access with collision avoidance (CSMA/CA), on account of its distributed nature for the simplicity of implementation [2]. To provide multimedia services in such contention-based networks, fairness is of particular concern because the QoS performed essentially depends on the sharing of transmission mediums among users.

The fairness of IEEE 802.11 DCF has been largely investigated in previous works [2–14]. The study reported by Koksal *et al.* [3] shows that 802.11 DCF presents short-term unfairness of channel sharing due to the backoff protocol in CSMA/CA, which can therefore make a significant impact on delay-sensitive applications like real-time audio and streaming video [3]. However, Berger-Sabbatel *et al.* [4]

provided a contrary perception that DCF indeed presents pretty fine fairness. They argued that the confusion of fairness problems in the previous work [3] is as a result of analysing the behaviour of CSMA/CA protocol specific to Wavelan system [15] instead of that characterised in 802.11 standards. Actually, the two access methods present a significant difference between them: the Wavelan CSMA/CA protocol executes exponential backoff when the channel is sensed busy, whereas 802.11 protocol does that only when a collision is experienced. Although the analysis of Berger-Sabbatel *et al.* [4] is rather consistent with the behaviour of present 802.11 protocol; however, the conclusion is valid only under the assumption of homogeneous link qualities among hosts, which may be impractical. In fact, hosts in WLANs commonly experience different signal qualities, especially in an indoor environment, that hosts close to the access point (AP) have a line-of-sight signal whereas those obstructed by physical objects experience much degraded signal quality. To cope with dynamic channel conditions to provide a good transmission quality, for example an acceptable level of bit error rate (BER), 802.11 standards on physical layer (PHY) [1] support a link adaptation mechanism, which dynamically selects one modulation and coding scheme (MCS) such that BER of the selected MCS with the highest data rate is within the prescribed performance bound. If all hosts have the same BERs, their throughputs will be equal regardless of their transmission rates [6]. This phenomenon is so called 'performance anomaly' [5]. However, even with a link adaptation mechanism, the link qualities among hosts will not likely be the same at most of the time since the MCS available are limited, which may therefore cause the unfairness of throughput sharing among hosts.

To demonstrate the significant impact of heterogeneous link qualities on fairness, we have conducted a simulation scenario of error-prone channels (Ecs) as shown in Section 2. The numerical results show that unequal link

Table 1: System parameters

Payload = 1023 bytes	MAC header = 28 bytes	Initial window size = 32	Maximum window size = 1024
Slot time = 20 us	PHY header = 24 bytes	DIFS = 50 us	Window increasing factor = 2
Propagation delay =1 us	ACK = 38 bytes	SIFS = 10 us	Maximum retry limit = 5

qualities among hosts can pose severe unfairness of channel sharing. For example, the average throughput variation between hosts in an error-free condition and those with a BER of 2×10^{-5} can be as large as 48%. Such unfairness behaviour can degrade the QoS performed in adverse channel conditions.

The objective of this work is to provide fair multimedia QoS for hosts under heterogeneous channel conditions in 802.11 DCF-based WLANs. In this paper which substantially extends our previous work [16], we present a cross-layer adaptive scheme that determines the multidimensional MAC-layer parameters based on PHY-layer channel conditions as well as the application-layer QoS requirements, for example the desired level of throughput or the tolerable bound of delay. Neural networks are used to learn the cross-layer correlation between the parameters adopted and QoS performed under time-varying channel conditions. We explore a cost−reward function to quantify the overall QoS performance, and then use the nonlinear correlation learned with neural networks to evaluate the gradient of this cost−reward function with respect to each parameter by means of a back-propagation manner. The multidimensional parameters can therefore be jointly determined for achieving fair QoS in terms of minimising the cost−reward function. We have applied this neural network-based learning technique for providing weighted QoS in 802.11e enhanced distributed co-ordination function (EDCF) [17] by adjusting the initial window size and arbitrary inter frame space (AIFS) [18].

We focus on the transmission scenario that the participating hosts are with similar real-time services of equal access priorities in an infrastructure mode. Three simulation scenarios are conducted to evaluate the effectiveness of our adaptation scheme. The results demonstrate that our scheme can tackle a variety of channel conditions and the dynamic participation or departure of hosts to provide fair multimedia QoS in terms of both throughput and packet delay. The remainder of this paper is organised as follows. Section 2 presents numerical results to show the performance variation due to heterogeneous channel conditions. Section 3 formalises the addressed problem and presents our adaptive algorithm as a solution. In Section 4, we construct simulation scenarios to evaluate the effectiveness of the proposed scheme. Section 5 draws the conclusions.

2 Unequal channel sharing in heterogeneous channel conditions

In this section, we conduct simulations to explore the unequal channel sharing due to heterogeneous channel conditions and its impact on multimedia QoS. We provide numerical results in both cases of hosts transmitting at an equal data rate and at different rates with a link adaptation mechanism. Consider K hosts in an infrastructure 802.11b WLAN environment. Assume each host transmits a saturated data traffic (i.e. always has a packet to send) to AP. Assume that half of the hosts, named ideal-channel (IC) hosts, are always in an IC condition; the others, named EC hosts, are initially in an ideal condition and later suffer from the channel degradation with an average BER

level of BER_{EC} when moving away from AP. The system parameters are shown in Table 1.

In the first experiment, all hosts use the same MCS and transmit at the data rate of 1 Mbps. The saturated throughputs of an IC or EC host are shown in Table 2. It is shown that when all the hosts are initially in an ideal condition ($BER_{EC} = 0$), their throughputs are equal. When BER_{EC} later deteriorates to 2×10^{-5} and 4×10^{-5} sequentially, the performance variation between an IC and EC host is gradually enlarged. Assume there are ten hosts with each one demanding the throughput of 64 kbps for meeting its own QoS. If all the hosts are in an IC condition, the throughput of each host is 80 kbps as shown in Table 2 and thus QoS is assured. However, if the link qualities of EC hosts deteriorate with BER of 2×10^{-5}, their throughputs degrade to 56 kbps and can no more meet the required QoS. In the meanwhile, the throughputs of IC hosts increase to 94 kbps, much exceeding the demand. The performance difference between these two groups of hosts becomes as large as 47.5% (38 kbps/80 kbps = 47.5%).

In the second experiment, IC hosts transmit at the data rate of 11 Mbps while EC hosts transmit at 1 Mbps. Table 3 shows the result of saturated throughputs. Initially, when IC and EC hosts are in an IC condition, they have equal throughput lower than 1 Mbps. This phenomenon is so called 'performance anomaly' [5] meaning that if some hosts transmit at a lower data rate, the throughput of others at higher rates will be degraded below the level of the lower rate. However, when BER_{EC} deteriorates later, it is shown that the throughput-based fairness (i.e. performance anomaly) gradually fades away. For example, when there are an IC and EC hosts with BER_{EC} equal to 4×10^{-5} in the network, the throughput of the EC host is degraded to 0.507 Mbps while that of the IC host is increased to 1.295 Mbps beyond the range of 1 Mbps.

The performance variation arises as a result of the following facts. Due to its higher BER, an EC host averagely experiences more retries to succeed a transmission than an

Table 2: In the first experiment, the saturated throughput of an IC and EC hosts, respectively, against the number of hosts varying with the BER level of EC hosts

	The number of hosts: K				
	2	4	6	8	10
The saturated throughput of an IC host (kbps)					
BER = 0	436	211	137	100	80
$BER_{EC} = 2 \times 10^{-5}$	494	244	160	118	94
$BER_{EC} = 4 \times 10^{-5}$	565	280	184	135	107
The saturated throughput of an EC host (kbps)					
BER = 0	436	211	137	100	80
$BER_{EC} = 2 \times 10^{-5}$	319	152	97	71	56
$BER_{EC} = 4 \times 10^{-5}$	219	107	37	49	38

Table 3: In the second experiment, the saturated throughput of an IC and EC hosts, respectively, against the number of hosts varying with the BER level of EC hosts

	The number of hosts: K				
	2	6	8	9	10
The saturated throughput of an IC host with 11 Mbps (kbps)					
$\text{BER} = 0$	779	379	246	181	142
$\text{BER}_{EC} = 2 \times 10^{-5}$	978	491	325	241	190
$\text{BER}_{EC} = 4 \times 10^{-5}$	1295	656	433	319	250
The saturated throughput of an EC host with 1 Mbps (kbps)					
$\text{BER} = 0$	779	379	246	181	142
$\text{BER}_{EC} = 2 \times 10^{-5}$	637	309	200	147	115
$\text{BER}_{EC} = 4 \times 10^{-5}$	507	246	159	116	91

IC host does. When a retransmission is performed, according to CSMA/CA protocol, the backoff window will be increased exponentially until the retries reach a certain limit. Thus, an EC host would averagely adopt a larger backoff timer and then has less chance to access the channel. From our numerical results, it is also demonstrated that 802.11 CSMA/CA can only present fairness of channel usage on the condition of homogeneous link qualities; the presence of heterogeneous channel conditions can pose significant unfairness to hosts at either an equal data rate or different rates with a link adaptation mechanism due to the reason of limited MCSs available.

3 Proposed cross-layer adaptive algorithm using neural networks

Based on the above observation, it is shown that fair QoS cannot be achieved by the approach of fixed parameters in a varying heterogeneous wireless environment. In this section, we formalise this problem as an optimisation problem and present our solution of a cross-layer adaptation scheme. For illustration, we use throughput for the QoS indicator and the backoff parameters for the adaptable arguments, which consist of initial window size, window increasing factor and maximum retry limit.

3.1 Problem description

Consider K hosts with each one transmitting a real-time traffic flow. To host i, the throughput for satisfying QoS is given as T_THR_i; the initial window size, window increasing factor and maximum retry limit adopted is $CW_{i,\min}$, σ_i, and $L_{i,\text{retry}}$, respectively; the throughput is T_i. Since the achievable throughput of a given host is affected by the parameters adopted of its own as well as other hosts, the overall throughputs performed of the K hosts are then modelled as a correlation function $f(\cdot)$ associated with the joint setting of parameters. That is

$$(T_1, T_2, \ldots, T_K) = f(CW_{1,\min}, \sigma_1, L_{1,\text{retry}}$$
$$\ldots, CW_{K,\min}, \sigma_K, L_{K,\text{retry}}) \tag{1}$$

The optimal parameters can be chosen by minimising a cost–reward function such that the overall throughputs are close to the prescribed QoS levels. The cost–reward function, C_{QoS} is defined as

$$C_{\text{QoS}} = \sum_{i=1}^{K} \frac{(T_i - T\_THR_i)^2}{T\_THR_i} \tag{2}$$

where the denominator T_THR_i is utilised to normalise the difference between the achievable throughput and the corresponding QoS requirement. C_{QoS} will be minimised if the channel is fairly shared in proportion to T_THR_i.

To minimise C_{QoS}, each backoff parameter is iteratively updated based on the gradient of C_{QoS} with respect to itself. To calculate the gradient, the knowledge of $f(\cdot)$ is needed. However, $f(\cdot)$ strongly depends on the characteristics of communication environments such as the collision probability among hosts and the link quality of each host. For example, if some hosts experience link quality degradation, the channel sharing of them as well as that of other hosts will be influenced, leading to a skewed sharing of overall throughputs. Hence $f(\cdot)$ is a nonlinear, complicated and time-variant function, which is rather difficult to be depicted with analytical formulas [19]. We are thus motivated to exploit neural networks (NN) to model the function $f(\cdot)$. Then the learned $\hat{f}(\cdot)$ is utilised to evaluate the gradient of C_{QoS} with respect to each parameter by the technique of the back-propagation manner. The function modelling stage and parameter determining stage of our adaptation scheme will be discussed in the following subsections.

3.2 Learning the nonlinear function

We exploit the multilayer perceptron (MLP) to model the correlation function between the backoff parameters and corresponding throughputs, $f(\cdot)$. The nonlinear and feed-forward MLP is capable of learning $f(\cdot)$ adaptively in a supervised manner [20], and can approximate it to an arbitrary degree of the accuracy [21]. Fig. 1 shows the architecture of the exploited MLP (indicated as 'the original NN for modelling the nonlinear function' inside the dotted line). It consists of 3K-x-K sensory units at the input layer, hidden layer and output layer, respectively, to model the nonlinear function from 3K backoff parameters to the corresponding K throughputs. The output of the ith neuron at the lth layer can be described as

$$u_i(l) = \sum_{j=1}^{N_{l-1}} \omega_{ij}(l) a_j(l-1) + \theta_i(l) \tag{3}$$

$$a_i(l) = h(u_i(l)) \quad 1 \le i \le N_l; \quad l = 1, 2 \tag{4}$$
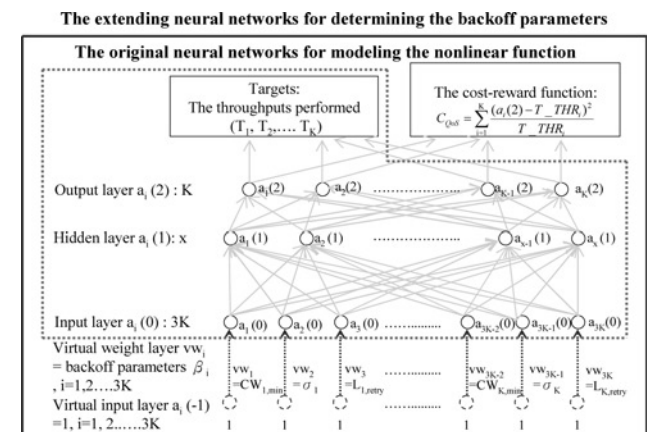
Fig. 1 *Exploited neural network architecture*

where N_l is the number of neurons at the lth layer; $u_i(l)$ and $a_i(l)$ are correspondingly the activation and output values of the ith neuron at the lth layer. The input units are represented by $a_i(0)$ and the output units by $a_i(2)$. $\omega_{ij}(l)$ refers to the weight connecting the output of the jth neuron at the $(l-1)$th layer to the activation of the ith neuron at the lth layer. $\theta_i(l)$ refers to the bias associated with the ith neuron at the lth layer. The transfer function $h(\cdot)$ is sigmoid at the hidden layer and linear at the output layer.

The nonlinear function $f(\cdot)$ can be modelled with MLP by recursively adjusting $\omega_{ij}(l)$ and $\theta_i(l)$ to minimise mean squares error (MSE) between the performed throughput, T_i and actual outputs, $a_i(2)$. That is

$$E = \frac{1}{2}\sum_{m=1}^{M}\sum_{i=1}^{K}\left(T_i^{(m)} - a_i^{(m)}(2)\right)^2 \qquad (5)$$

where M is the number of teacher patterns. The learning procedure of MLP in details can be found in [20].

3.3 Adaptively determining the backoff parameters

Here we denote parameter i of the nth adaptation as $\beta_i^{(n)}$. The gradient learning formula to minimise C_{QoS} with respect to $\beta_i^{(n)}$ is

$$\beta_i^{(n+1)} = \beta_i^{(n)} + \Delta\beta_i^{(n)} \qquad (6)$$

$$\Delta\beta_i^{(n)} = -\mu\partial C_{\text{QoS}}/\partial\beta_i^{(n)}, \quad 1 \le i \le 3K \qquad (7)$$

where μ is the adjusting rate. To evaluate the gradient of C_{QoS} with respect to β_i, we add a virtual input layer $a_i(-1)$ and a virtual weight layer vw_i under $a_i(0)$. The extending architecture is shown in Fig. 1 (indicated as 'the extending NN for determining the backoff parameters'). The virtual weight vw_i is set with the backoff parameter, β_i. By adding the virtual input layer and virtual weights, we can separately model $f(\cdot)$ in the original NN and thus utilise the learned $\hat{f}(\cdot)$ to evaluate the gradient of C_{QoS} in the extending networks.

All the units in the virtual input layer $a_i(-1)$ are set to 1, and the transferring function at the input layer $a_i(0)$ is linear. That is $a_i^{(n)}(0) = \beta_i^{(n)}$. Thus, the problem of evaluating the minus gradient of C_{QoS} for minimising C_{QoS} with respect to $\beta_i^{(n)}$ is equivalent to that with respect to $a_i^{(n)}(0)$. To compute the minus gradient of C_{QoS} with respect to $a_i^{(n)}(0)$, $-\partial C_{\text{QoS}}/\partial a_i^{(n)}(0)$, a chain-rule technique is utilised

$$-\partial C_{\text{QoS}}/\partial a_i^{(n)}(0) = \lambda_i^{(n)}(1)\left[\partial a_i^{(n)}(1)/\partial a_i^{(n)}(0)\right] \qquad (8)$$

where $\lambda_i^{(n)}(l)$ is defined as the minus gradient of C_{QoS} with respect to $a_i^{(n)}(l)$: $\lambda_i^{(n)}(l) \equiv -\partial C_{\text{QoS}}/\partial a_i^{(n)}(l)$. By means of the back-propagation and chain-rule methods, $\lambda_i^{(n)}(l)$ can be derived with the knowledge of $\lambda_i^{(n)}(l+1)$ from the upper layer

$$\lambda_i^{(n)}(l) = \sum_{j=1}^{N_{l+1}} \lambda_j^{(n)}(l+1)h'\left(u_j(l+1)\right)w_{ji}(l+1) \qquad (9)$$

In particular, the minus gradient of C_{QoS} with respect to the output layer, $\lambda_i^{(n)}(2)$ can be derived from the cost–reward function as shown in (2). Therefore $\lambda_i^{(n)}(l)$ can be successively derived in the order of $l = 2$, 1 and 0. Since the value of $a_i^{(n)}(0)$ is equal to $\beta_i^{(n)}$, (6) becomes

$$\beta_i^{(n+1)} = \beta_i^{(n)} + \mu\lambda_i^{(n)}(0) \qquad (10)$$

From (10), the multidimensional backoff parameters $\beta_1^{(n+1)} \ldots \beta_{3K}^{(n+1)}$ for minimising C_{QoS} can be determined concurrently.

3.4 Alternatively modelling the function and determining parameters

New backoff parameters are therefore applied into the system. In order to learn $f(\cdot)$ in current wireless environments, the training data will be updated with the recent used backoff parameters and the corresponding throughputs. In a recursive manner, this algorithm models the nonlinear function and adjusts the backoff parameters for optimising the overall QoS in terms of minimising C_{QoS}. It is noticeable that in this paper, we particularly focus on the WLAN with limited (or just adequate) resources available to satisfy overall QoS requirements and then perform our adaptation scheme to achieve QoS fairness. In case that the network resources available are sufficient (e.g. by using an admission control scheme to limit the amount of hosts) and all the participating nodes can meet their requirements, our adaptive mechanism need not perform any action. Accordingly, our adaptation mechanism will act only on the premise that some hosts cannot meet their requirements due to the imbalance resource usage among hosts, and the resources available are adequate to satisfy overall QoS requirements.

This adaptation mechanism works in a centralised sense, and thus is designed for an infrastructure 802.11 WLAN and performed at AP. At the run time, AP will request the associated hosts to send packets indicating their QoS requirements, for example the requirements for the lowest QoS and high-fidelity QoS. According to the network capability, AP will select adequate levels for the cost–reward function (i.e. determines T_THR$_i$ described in (2). By logging the information about payload size and packet arrival timestamp from successfully received packets over a sample time periodically, AP can evaluate QoS achievable by factors such as throughput and packet delay of each associated host. With the information of achievable QoS needed for the training data and cost–reward function, AP can therefore centrally model the nonlinear function and adjust backoff parameters alternatively. Whenever the parameters are adjusted, AP will send packets to notify the associated hosts of their new parameters.

3.5 Dynamic NN architecture for hosts departing from or joining in the network

In Section 3.4 we illustrated this adaptation mechanism using the example of a fixed NN architecture, that is a constant amount of hosts. However, hosts in wireless environments can randomly join in or depart from a network, which introduces the need of a flexible NN architecture to accommodate the situation. For this purpose, we devise a dynamic NN scheme similar to the optimal brain damage (OBD) [22]. OBD can remove unimportant weights from NN to optimise both network complexity and training set error [22]. In a similar manner, we disable or enable some
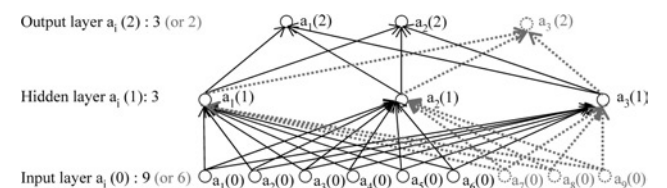

Fig. 2 *Dynamic NN architecture*

NN components in accordance with the current network topology. When some hosts leave the network, their associated weights and input/output neurons will therefore be set to 0 fixedly. If hosts join in, some NN components will be activated corresponding to the current network topology. Fig. 2 shows the example of a dynamic NN architecture for two or three participating hosts. In case of three hosts, all NN components are active to model the function and adjust parameters. If one host departs away from the network later, its weights and input/output neurons (indicated as red dotted lines) will therefore be set to 0 fixedly, resulting in an equivalent architecture for only two hosts. With such a dynamic NN architecture, our adaptation framework works well in time-varying wireless environments.

3.6 Properties of this adaptation framework applied in 802.11 WLAN environments

As mentioned in Section 3.4, the achievable QoS needed for the training data is evaluated by means of a measurement-based approach. Thus, in order to learn the nonlinear function $f(\cdot)$ that accurately reflects the rapidly changing channel conditions, the sample time of measurements should be less than the channel coherence time, that is the time period within which the wireless channel characteristic can be regarded unchanged. At a mobile speed of 1 m/s (nearly the walking speed) in IEEE 802.11b WLANs with the centre frequency of 2.4 GHz, the channel coherence time is about 120 ms [23]. As for the measuring time period, it can be regarded as multiple numbers of packet transmissions which depend on many factors including transmission data rate, payload size, packet collision rate, and so on. Consider an 802.11b WLAN with the capacity of 1 Mbps in which each host transmits a saturated traffic flow of about 1000-bytes payload size. If there are no more than five contending hosts (2–5 hosts), the typical value of the MAC frame transmission time ranges from 10 to 50 ms [24, 25]. That is the coherence time of 802.11 wireless channels is generally on the order of multiple packet transmission times [23]. This fact indicates that using the measurement-based approach for obtaining training data has the potential to learn the nonlinear function accurately reflecting the varying 802.11 channels.

The computing speed of NN is also of particular concern for applying this adaptation framework in rapidly varying wireless environments. Today, NN can be implemented with field-programmable gate array (FPGA) chips for combining the flexibility of software solutions (e.g. the dynamic NN scheme aforementioned) and the efficiency of hardware computing [26]. Such a hardware approach can offer solutions quite quickly, may be in a few microseconds [27]. Thus the hardware-based implementation for the proposed NN-based adaptation framework could be a promising approach to varying 802.11 WLAN environments.

4 Experimental results

In this section, we show the simulation results to demonstrate that our approach can effectively provide fair multimedia QoS in varying heterogeneous WLAN environments. We examine three representative scenarios as follows in our simulations.

Scenario I: hosts with heterogeneous channel conditions that transmit at an equal data rate (using the same MCS).

Scenario II: hosts with heterogeneous channel conditions that transmit at unequal data rates with a link adaptation mechanism (different MCS used).

Scenario III: hosts depart from or join in the network.

For each scenario, we assume an IEEE 802.11b infrastructure WLAN in which each participating host transmits one saturated video streaming with a fixed payload size of 1023 bytes through AP to the corresponding receiver. Assume the throughput requirements of this video application for the high fidelity and the lowest quality are at least 320 and 160 kbps, respectively. We compare the proposed adaptive scheme with IEEE 802.11 DCF protocol, which adopts fixed backoff parameters with the initial window size of 32, window increasing factor of 2 and maximum retry limit of 5 [1]. The other system parameters used are shown in Table 1. For our adaptive algorithm, the adaptation space of initial window size is from 8 to 64, window increasing factor from 1.1 to 4 and maximum retry limit from 1 to 10. The parameter, T_THR_i, for the cost−reward function is set to the requirement level for the lowest QoS (160 kbps) or high-fidelity QoS (320 Kbps) depending on the network capacity.

We use the well-known Gilbert-Elliott two-state discrete Markov model [28] to model time-varying channels. In this model, the channel condition could be either 'bad' or 'good'. The time of the channel staying in the bad and good states are exponentially distributed with rates μ_b and μ_g, respectively. Thus the state transition probability from 'bad' state to 'good' state, $p_{b,g}$ is $\mu_b/(\mu_g + \mu_b)$ and that from 'good' state to 'bad' state, $p_{g,b}$ is $\mu_g/(\mu_g + \mu_b)$. The values of $p_{b,g}$ and $p_{g,b}$ indicate the statistic of wireless channel conditions. For instance, the values of $p_{b,g}$ near 1 (0) implies that the channel state is good (bad) in most cases. This model is commonly used for modelling fading channels such as the Rayleigh channel [29]. For the simulation we assume that BER in the 'good' state is 1×10^{-7} and 1×10^{-4} in 'bad' state. The hosts in the network are assumed composed of IC and EC hosts. IC hosts are with the parameter $p_{b,g}$ of 1, that is always in the 'good' state with a BER of 1×10^{-7}, while EC hosts are with $p_{b,g}$ with lower values depending on simulation scenarios.

4.1 Scenario I: hosts with the same MCS in heterogeneous channel conditions

The simulation set-up for this scenario assumes two IC hosts and two EC hosts in the network. EC hosts are assumed with $p_{b,g}$ of 0.8 (the corresponding average BER is about 2×10^{-5}) initially, and later suffer from channel degradation with $p_{b,g}$ of 0.6 (the corresponding average BER is about 4×10^{-5}) when they move away from AP. This experimental scenario is set up such that the network capacity is sufficient for all users with the lowest requirements, 160 kbps. The parameter T_THR_i is accordingly set to 160 kbps. To clearly show the result of throughputs with respect to IC and EC hosts, we use the following simplified NN architecture consistent with our design goal. The input space consists of six components which are three backoff parameters for IC and EC hosts, respectively. The output space consists of two components, which are the throughputs achievable of IC and EC hosts separately. six hidden nodes are used. Therefore the network topology consists of 6-6-2 nodes from the input to the output layer. In the function modelling stage, we use an online learning strategy, that is the NN is updated when a new training data is available. The NN weights and biases are adjusted till the MSE falls below 1×10^{-6} or the training epochs are over 1000 times. In the parameter determining stage, the parameters are updated with the adjusting rate μ of 0.1.
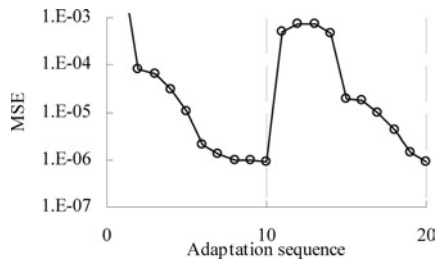
Fig. 3 *MSE varying with the adaptation sequence*

Fig. 3 presents the level of MSE varying with the adaptation sequences, that is the numbers of the parameters are adjusted. It is shown that MSE continues to decrease mostly when NN can learn the function gradually. Although the communication environment changes (as the parameter $p_{b,g}$ of EC hosts becomes 0.6) during the sequence of 11, MSE rises immediately due to the changes of the nonlinear function. With the learning capability of NN and the up-to-date teacher data, MSE decreases rapidly after the adaptation sequence of 15. The results demonstrate that our adaptive algorithm can model the real-time communication system under varying heterogeneous channel conditions.

Fig. 4 presents the saturated throughputs for IC and EC hosts with 802.11 DCF and the proposed adaptation scheme, respectively. It is shown that our approach achieves not only much better fairness but also a significant improvement of QoS. With 802.11 DCF, the throughput of an EC host is 151.7 kbps initially and down to 104 kbps as the communication environment deteriorates; both are below the minimum QoS requirement (160 kbps). The QoS degradation of EC hosts is posed by a skewed channel sharing because of using fixed parameters in heterogeneous channel conditions.

With the proposed adaptation scheme, the fairness of channel sharing is improved significantly such that the achievable throughput of EC hosts can progressively meet the QoS requirement. It is shown in Fig. 4 that the throughput of EC hosts drops before the adaptation sequence of five because the gradient of the cost–reward function is not well-evaluated yet due to limited teacher patterns. While the gradient is evaluated rather accurately with more teacher patterns latter, the throughputs of IC and EC hosts are almost equal (193 kbps) after the sequence of seven, which are better than the QoS requirement level. When BER of EC hosts deteriorates later, the backoff parameters adjusted earlier cease to be effective in the present condition. Consequently, the throughput of EC host degrades
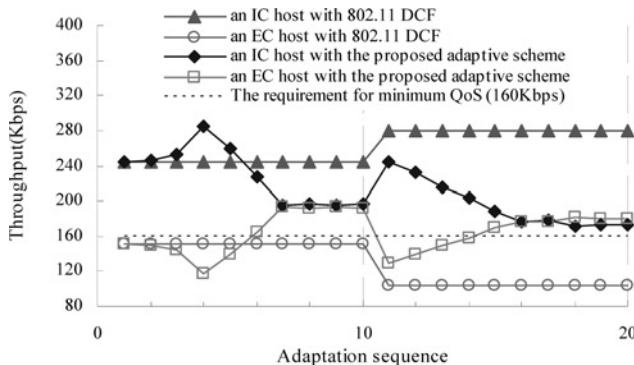
to 131.5 kbps and can no more satisfy QoS requirement. With the proposed algorithm online learning the current function and adjusting parameters, the throughputs of IC and EC hosts are almost equal (176 kbps) after the sequence of 16 which again meets the required QoS. The results demonstrate that our adaptation scheme can effectively tackle a variety of channel conditions, providing fair multimedia QoS for hosts with the same data rates.

4.2 Scenario II: hosts with different data rates in heterogeneous channel conditions

Now we examine the proposed adaptation mechanism using the scenario that hosts transmit at unequal data rates with a link adaptation mechanism. The simulation set-up of this scenario is similar to that of Scenario I. In particular, here IC hosts are assumed to transmit at the data rate of 1 Mbps while EC hosts at 11 Mbps. Because of the higher network capacity, the parameter T_THR_i of the cost–reward function is set as the high-fidelity level of QoS, 320 kbps.

Fig. 5 presents the experimental results of saturated throughput. With 802.11 DCF, it is shown the achievable throughput of EC host with 11 Mbps is even lower than that of an IC host with 1 Mbps. The reason is as follows: when the participating hosts' link qualities and packet lengths are similar, CSMA/CA protocol provides throughput-based fairness regardless of their transmission rates [6] (referred to the phenomenon of 'performance anomaly' [5]). Thus, if the hosts at higher data rates experience worse channel conditions (such like EC hosts), the throughputs can even be less than that of the hosts at lower rates with better channels (such like IC hosts). We also find in Fig. 5 that EC hosts cannot meet the QoS requirement of high fidelity, while IC hosts exceed the requirement. The results demonstrate that the unbalanced throughput is caused by heterogeneous link qualities rather than unequal data rates.

With the proposed adaptation scheme, the throughput of an EC host can be progressively improved and then meet the QoS requirement. Eventually, the throughputs of IC and EC hosts are close to each other, both satisfying high-fidelity QoS. In addition, the aggregated throughput is increased from 1.2 to 1.46 Mbps on average because our adaptation scheme provides more transmission opportunities for EC hosts with higher data rates. From the results shown in Sections 4.1 and 4.2, it is demonstrated that under a variety of channel conditions, our adaptation scheme can effectively provide fair multimedia QoS for hosts at either

Fig. 4 *In Scenario I, the saturated throughput of an IC and EC hosts with 802.11 DCF and the proposed adaptive scheme, respectively*
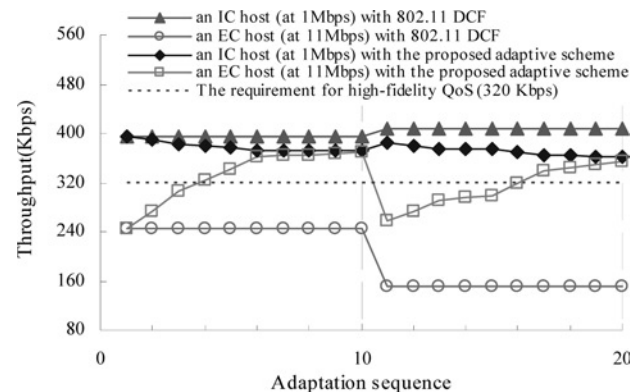
Fig. 5 *In Scenario II, the saturated throughput of an IC and EC hosts with 802.11 DCF and the proposed adaptive scheme, respectively*
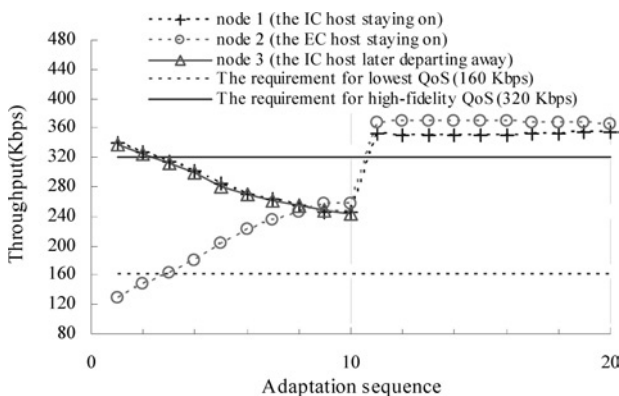
equal data rates or different rates with a link adaptation mechanism.

4.3 Scenario III: hosts departing from or joining in the network

Now we examine the effectiveness of the proposed adaptive scheme using the scenarios of hosts departing from and joining in the network. In both scenarios, EC hosts are assumed with the parameter $p_{b,g}$ of 0.6. In the departing scenario, there are two IC hosts and one EC host in the network at first, and one IC host departs later. In the joining scenario, there are an IC host and an EC host in the network and later one more EC host joins in. For the simplicity of illustration in both scenarios we denote node 1 as the IC host staying on, node 2 as the EC host staying on, and node 3 as the IC (or EC) host departing away from (joining in) the network. The parameter T_THR_i of the cost−reward function is set to be 160 kbps (the require-ment of the lowest QoS) in case of three hosts existing and 320 kbps (the requirement of high-fidelity QoS) in case of two hosts.

4.3.1 Hosts departing from the network: The simu-lation results of the host departing scenario are presented in Fig. 6. While the number of hosts is 3, the throughputs of node 1, node 2 and node 3 are 246 kbps (33%), 257 kbps (34%) and 244 kbps (33%) around the adaptation sequence of ten individually. When node 3 departs away at the sequence of 11, the throughputs achievable for node 1 and node 2 simultaneously increase to 352 kbps (49%) and 368 kbps (51%), respectively, which still present equal share of channel usage. The fairness is kept at this moment since the parameters achieving fair QoS for three hosts is still effective for the two staying-on hosts. Afterward, this algorithm performs adaptation correspond-ing to only node 1 and node 2 by using the dynamic NN architecture. It is shown that the throughputs of the two hosts are almost unchanged with time. The results demon-strate that when some hosts depart away from wireless environments, the proposed adaptive scheme can automati-cally provide fair QoS among staying-on hosts.

4.3.2 Hosts joining in the network: The simulation results of host joining scenario are shown in Fig. 7. When the number of hosts is two, the throughputs of node 1 and node 2 are 365 kbps (50%) and 359 kbps (50%) around the adaptation sequence of ten individually. When node 3 joins in the network around the sequence of 11, the
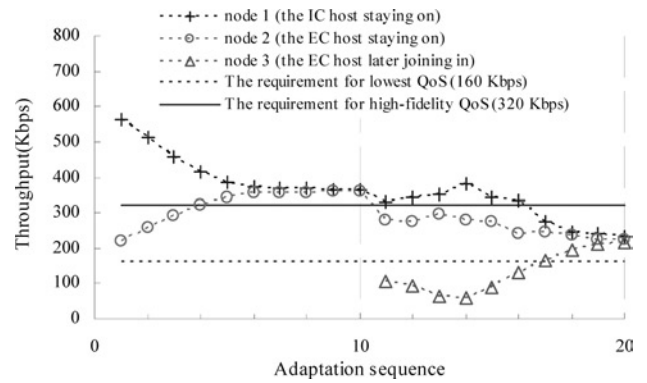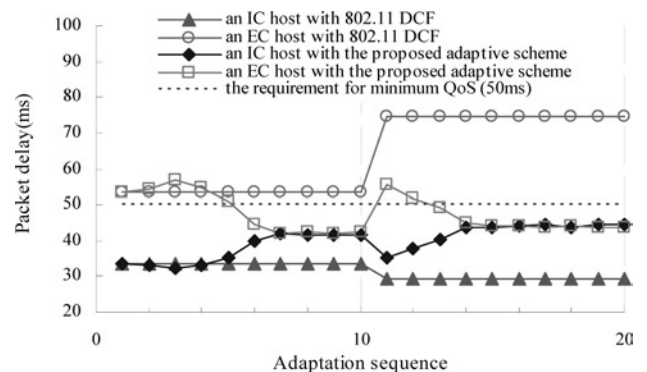

Fig. 7 *In Scenario III that the number of participating hosts increases from 2 to 3, the saturated throughput of each host using the proposed adaptive scheme*

throughputs for node 1, node 2 and node 3 simultaneously become 328 kbps (46%), 278 kbps (39%) and 106 kbps (15%) correspondingly. The fairness deteriorates at this moment such that node 1 meets high-fidelity QoS, whereas node 2 only meets the lowest QoS and node 3 even cannot meet the lowest QoS. The parameters achieving fair QoS for node 1 and node 2 cease to be effective at this moment since node 3 uses 802.11 DCF default parameters. After NN gradually learns the new knowledge and adjust parameters, it is shown that the throughputs of node 1, node 2 and node 3 become 232 kbps (35%), 224 kbps (33%) and 215 kbps (32%), respectively, around the sequence of 20, presenting the fairness of QoS among hosts again. Therefore all the three hosts can meet the requirement of the lowest QoS. The results demonstrate that when some hosts join in wireless environments, the pro-posed adaptive scheme can automatically provide the fair-ness of QoS among hosts.

4.4 Fair QoS in terms of packet delay

Finally, we use Scenario I to illustrate that fair QoS in terms of packet delay can also be achieved with the same tech-nique. Each video streaming is assumed with the delay bound of 50 ms for a smooth playback. Here this proposed adaptation framework use packet delay as QoS indicator. The simulation result is presented in Fig. 8, similar to what shown in Fig. 4, that with 802.11 DCF, an IC host has better channel usage beyond the requirement (its packet delay is further lower than 50 ms) whereas an EC host cannot meet the requirement. With the proposed scheme, the performance variation is gradually narrowed

Fig. 6 *In Scenario III that the number of participating hosts decreases from 3 to 2, the saturated throughput of each host using the proposed adaptive scheme*

Fig. 8 *In Scenario I, the saturated packet delay of an IC and EC hosts with 802.11 DCF and the proposed adaptive scheme, respectively*

864

such that both IC and EC hosts can satisfy QoS requirement. It is shown that our proposed adaptation scheme can effectively provide the fairness of QoS in terms of both throughput and packet delay under a variety of channel conditions.

5 Conclusion

In this paper, we investigate the fairness of multimedia QoS in heterogeneous WLAN environments. We have observed that unequal link qualities among hosts can cause severe unfairness of channel sharing, which thereby leads to the degradation of QoS in adverse conditions. To provide fair channel shares, this paper proposes a cross-layer adaptive algorithm which dynamically adjusts MAC-layer backoff parameters based on the application-layer QoS requirements as well as PHY-layer channel conditions. Our solution is provided with an optimisation approach, which utilises NN to learn the cross-layer function. The simulation results illustrate that under a variety of channel conditions and host amounts, our adaptive scheme can provide fair QoS in terms of both throughput and packet delay.

6 Acknowledgment

7 References

1 IEEE 802.11a/b: 'Wireless LAN medium access control (MAC) and physical layer (PHY) specifications, Standard, IEEE', August 1999
2 Nandiraju, N., Gossain, H., Cavalcanti, D., Chowdhury, K., and Agrawal, D.P.: 'Achieving fairness in wireless LANs by enhanced IEEE 802.11 DCF'. Proc. IEEE WiMob, 2006, pp. 132–139
3 Koksal, C.E., Kassab, H., and Balakrishnan, H.: 'An analysis of short-term fairness in wireless media access protocols.' Proc. ACM SIGMETRICS, 2000, pp. 118–119
4 Berger-Sabbatel, G., Duda, A., Heusse, M., and Rousseau, F.: 'Short-term fairness of 802.11 networks with several hosts'. Proc. IEEE MWCN, 2004, pp. 263–274
5 Heusse, M., Rousseau, F., Berger-Sabbatel, G., and Duda, A.: 'Performance anomaly of 802.11b'. Proc. IEEE INFOCOM, 2003, vol. 2, pp. 836–843
6 Tan, G., and Guttag, J.: 'Time-based fairness improves performance in multi-rate wireless LANs'. USENIX Annual Technical Conference, Boston, MA, June 2004
7 Huang, X.L., and Bensaou, B.: 'On max–min fairness and scheduling in wireless ad-hoc networks: analytical framework and implementation'. Proc. ACM MobiHoc, 2001, pp. 221–231
8 Luo, H., Cheng, J., and Lu, S.: 'Self-coordinating localized fair queueing in wireless ad hoc networks', IEEE Trans. Mobile Comp., 2004, 3, (1), p. 86 – 98
9 Jiang, L.B., and Liew, S.C.: 'Proportional fairness in wireless LANs and ad hoc networks'. Proc. of IEEE WCNC, 2005, vol. 3, pp. 1551–1556
10 Pong, D., and Moors, T.: 'Fairness and capacity trade-off in IEEE 802.11 WLANs'. Proc. IEEE LCN, 2004, pp. 310 – 317
11 Nandagopal, T., Kim, T.E., Gao, X., and Bharghavan, V.: 'Achieving MAC layer fairness in wireless packet networks'. Proc. ACM MOBICOM, 2000, pp. 87–98
12 He, J., and Pung, H.K.: 'A fairer multiple access protocol for multi-hop wireless networks: hybrid asynchronous time division multiple access protocol (HATDMA)'. Proc. IEEE LCN, 2003, p. 356–365
13 Ozugur, T., Naghshineh, M., Kermani, P., Olsen, C.M., Rezvani, B., and Copeland, J.A.: 'Balanced media access methods for wireless networks'. Proc. MOBICOM, 1998, pp. 21–32
14 Vaidya, N.H., Bahl, P., and Gupta, S.: 'Distributed fair scheduling in a wireless LAN'. Proc. ACM MOBICOM, 2000, pp. 167–178
15 The Lucent Home Page, http://www.lucent.com/
16 Wang, C., and Lin, T.: 'A neural network based adaptive algorithm for multimedia quality fairness in WLAN environments. Proc. IEEE ICME, 2006, pp. 1233–1236
17 IEEE 802.11e/D5.0: 'Draft supplement to Part 11: wireless medium access control (MAC) and physical layer (PHY) specifications: medium access control (MAC) enhancements for quality of service (QoS)', June 2003
18 Wang, C., Lin, P., and Lin, T.: 'A cross-layer adaptation scheme for improving IEEE 802.11e QoS by learning', IEEE Trans. Neural Netw., 2006, 17, (6), pp. 1661–1665
19 van der Schaar, M., and Shankar, N.S.: 'Cross-layer wireless multimedia transmission: challenges, principles, and new paradigms', IEEE Wirel. Commun., 2005, 12, (4), pp. 50–58
20 Haykin, S.: 'Neural networks: a comprehensive foundation' (Prentice-Hall, 1999, 2nd edn.), pp. 161–171
21 Hornik, K., Stinchcombe, M., and White, H.: 'Multilayer feedforward networks are universal approximators', Neural Netw., 1989, 2, (5), pp. 359–366
22 Cun, Y.L., Denker, J.S., and Solla, S.A.: 'Optimal brain damage' in Touretzky, D.S. (Ed.): 'Advances in neural information processing systems, 1990, vol. 2, pp. 598–605
23 Sadeghi, B., Kanodia, V., Sabharwal, A., and Knightly, E.: 'Opportunistic media access for multirate ad hoc networks'. Proc. ACM MOBICOM, 2002, pp. 24–35
24 Chatzimisios, P., Boucouvalas, A.C., and Vitsas, V.: 'IEEE 802.11 packet delay – a finite retry limit analysis'. Proc. IEEE GLOBECOM, 2003, pp. 950–954
25 Yin, J., Wang, X., and Agrawal, D.P.: 'Optimal packet size in error-prone channel for IEEE 802.11 distributed coordination function'. Proc. IEEE WCNC, 2004, vol. 3, pp. 1654–1659
26 Schurmann, F., Hohmann, S., Schemmel, J., and Meier, K.: 'Towards an artificial neural network framework'. Proc. NASA/DOD Conference on Evolvable Hardware, 2002, pp. 266–273
27 Ahn, C.W., and Ramakrishna, R.S.: 'QoS provisioning dynamic connection-admission control for multimedia wireless networks using a Hopfield neural network', IEEE Trans. Veh. Technol., 2004, 53, (1), pp. 106 – 117
28 Modiano, E.: 'An adaptive algorithm for optimizing the packetsize used in wireless ARQ protocols', Wirel. Netw., 1999, 5, (4), pp. 279–286
29 Qiao, D., Choi, S., and Shin, K.G.: 'Goodput analysis and link adaptation for IEEE 802.11a wireless LAN', IEEE Trans. Mobile Comp., 2002, 1, (4), pp. 278–292