國立交通大學

電機學院通訊與網路科技產業研發碩士班

碩 士 論 文

適用於高速下行封包擷取技術之

乏晰 Q-Learning 式混合自動重傳機制

Fuzzy Q-Learning based Hybrid ARQ for HSDPA

研 究 生：蔡世宏

指導教授：張仲儒　博士

中 華 民 國 九 十 六 年 六 月

適用於高速下行封包擷取技術之

乏晰 Q-Learning 式混合自動重傳機制

Fuzzy Q-Learning based Hybrid ARQ for HSDPA

研 究 生：蔡世宏　　　　　Student：Shih-Hung Tsai

指導教授：張仲儒 博士　　　Advisor：Dr. Chung-Ju Chang

國 立 交 通 大 學

電機學院通訊與網路科技產業研發碩士班

碩 士 論 文

A Thesis

Submitted to College of Electrical and Computer Engineering

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Industrial Technology R & D Master Program on
Communication Engineering
June 2007
Hsinchu, Taiwan, Republic of China

中 華 民 國 九 十 六 年 六 月

適用於高速下行封包擷取技術之
乏晰 Q-Learning 式混合自動重傳機制

研究生：蔡世宏　　　　　　　　　　　　指導教授：張仲儒 博士

國立交通大學電機學院產業研發碩士班

摘　　　　要

為了在現有的第三代通訊系統下提供更高速且安全的下鏈路資料封包傳送，第三代合作夥伴計畫（3rd generation partnership project, 3GPP）提出了一種高速下行封包擷取技術（high speed downlink packet access, HSDPA）。HSDPA 採用一種稱為調變編碼（adaptive modulation and coding, AMC）的技術，可以針對不同的通道情況來使用不同的調變方式和不同的編碼速率，並提供更大量的多碼（multi-code）以進行傳輸，也採取了混合自動重傳機制（hybrid automatic retransmission request, HARQ）使得傳輸機制更有效率。為了能夠適應通道的變化，HSDPA 也把每個傳送時間間隔（transmission time interval, TTI）縮短為 2ms，以求達成更有效率的資源分配模式。

本篇論文提出了一個適用於高速下行封包擷取技術之乏晰 Q-learning 式混合自動重傳機制（FQL based HARQ）。針對已被模擬成離散時間馬可夫決策過程（Markov decision process, MDP）的 HARQ 機制，並且期望能夠符合傳輸服務品質（quality of service, QoS），我們使用了名為乏晰 Q-learning（fuzzy Q-learning）的即時加強型學習演算法來不斷地學習並決定每次傳輸的方式，以期滿足 QoS 並維持高速傳輸速率。由模擬結果可以看到，我們所提出的機制可以在滿足 QoS 的同時也維持住高速的傳輸速率，並且也維持在很低的丟棄封包之機率，確實達成了預期的目標。

Fuzzy Q-Learning based Hybrid ARQ for HSDPA

Student：Shih-Hung Tsai   Advisor：Dr. Chung-Ju Chang

Industrial Technology R & D Master Program of
Electrical and Computer Engineering College
National Chiao Tung University

ABSTRACT

In order to provide higher speed and more effective downlink packet data service in 3G, high speed downlink packet access (HSDPA) is proposed by 3rd generation partnership project (3GPP). HSDPA adopts a technique called adaptive modulation and coding (AMC) to use different modulation orders and coding rates corresponding to different channel conditions. It provides more multi-codes for transmissions, and also adopts the hybrid automatic retransmission request (HARQ) scheme to make the transmission mechanism more effective. In order to adapt the changes of channel conditions, HSDPA adopts shorter transmission time interval (TTI) as 2ms to reach more effective resource allocation.

In this thesis, a fuzzy Q-learning based HARQ scheme for HSDPA is proposed. For the HARQ scheme modeled as Markov decision process (MDP), we use fuzzy Q-learning algorithm to learn the modulation and coding rates of initial transmissions. Our object is to satisfy the quality of service (QoS) and keep high data rate. The simulation results show that the proposed scheme can indeed reach the object and is feasible in different channel conditions.

誌謝

　　這篇論文的誕生，得感謝許多人所給予我的力量與幫助。首先我得感謝我的指導教授張仲儒博士，在我撰寫此篇論文的過程中給予我不少的提點與指引；再來則是得感謝早已自學校畢業的芳慶學長，犧牲自己的時間解答我所有的疑惑；感謝立峰、志明、詠漢和文祥學長不定時的提點和建議，讓我能更快地進入學術領域；感謝家源、煖玉、俊帆、琴雅，親切的帶領我熟悉並融入這整個實驗室的大家庭；感謝一起走來的建安、建興、佳泓、佳璇以及正昕，和我一起相互扶持走過這兩年，在我低潮的時候給予我繼續前進的勇氣；感謝尚樺、宗利、英奇、巧瑩、維謙、邱胤、浩翔以及惟娸，是你們讓實驗室變得更熱鬧，讓苦悶的生活增添不少色彩。

　　最後，還得感謝我的父母、家人以及朋友，沒有你們的支持、關心和鼓勵，我也無法順利而無憂的走過這兩年。

<div align="right">

世宏　謹誌

民國　九十六年六月

</div>

Contents

List of Tables

List of Figures

Chapter 1

Introduction

High Speed Downlink Packet Access (HSDPA) is an evolution of the Universal Mobile Telecommunication System (UMTS) standard to provide high speed downlink with peak data rate up to 8~10 Mbps. Based on a new channel – High Speed Downlink Shared Channel (HS-DSCH), it significantly reduces the downlink transmission delay compared to previous releases. It achieves a short round trip delay for transmission by defining each transmission time interval (TTI) to be 2ms instead of 10ms, which is defined in previous releases. Shared among several users, this channel is dedicated to support high data rates in downlink. An advanced technique called adaptive modulation and coding (AMC) technique is adopted for link adaptation in HSDPA. It can adjust the transmission rate dependent on the channel quality.

On the other hand, hybrid automatic repeat request (H-ARQ) scheme is proposed for the purpose of processing the link adaptation errors and improving the performance of retransmissions. H-ARQ combines ARQ technique and forward error correction (FEC) scheme to obtain advantages of high reliability in ARQ and good throughput even in poor channel condition. There are several types of H-ARQ: chase combining (CC), full incremental redundancy (full IR), partial incremental

redundancy (partial IR) [4], and adaptive and asynchronous incremental redundancy (A^2IR) [5]. Chase combining adopts a code combining method to reduce the number of retransmissions, while each retransmission is identical to the original transmission. A decoder combines these multiple copies of the transmitted packet weighted by their received SNR to obtain time diversity gain. Full IR uses a parity retransmission strategy. Instead of sending repeats of the same coded packet, the full IR will add an additional redundancy in the retransmission packet incrementally if an error is detected. The retransmissions are not identical with the first transmission, and the receiver will combine all previous transmitted packets to help correcting errors. Partial IR is a particular case of full IR in which every transmission is self-decodable. Each packet contains all information bits necessary for a correct reception of the data with different parity bits. Due to the self-decidable property, the partial IR can extract the information either from the last received packet or by combining all previous packets, which is generally the case with full IR. An enhanced version of the IR scheme called A^2IR was proposed in [5]. It selects a different modulation and coding scheme for each retransmission.

Maximal ratio combining (MRC) is used to combining (re)transmissions for HARQ, but it is not an optimal method under the condition of imperfect channel estimation. A modified MRC scheme was proposed in [6] to decrease the impact of the channel estimation errors as a weighting coefficient such that it can improve the performance. An enhancement HARQ method for TDD-CDMA supporting HSDPA techniques was proposed in [7]. The idea is to discard the retransmission and change the modulation and coding when the link quality is lower than the threshold which is pre-defined. Thus it can reduce the number of unnecessary retransmissions and improve the performance. The performance comparison between chase combining and IR schemes was investigated in [8] and [9]. It showed that if the scheduler has no

knowledge of the actual channel quality or for poor channel estimation, the IR is better than the CC. Simulation results showed in [9] have presented that full IR is a better scheme since full IR is better than the other schemes for most signal to noise ratio (SNR) values; while partial IR is better than full IR just under some moderate SNR values. In [10], the performance comparison among different schemes (A^2IR, CC and IR) has been discussed, and it also showed that A^2IR can improve system throughput under various conditions.

In the H-ARQ procedure, the object is to find a suitable decision for an initial transmission of packets dependent on the feedback information of channel quality indicator (CQI). Although in H-ARQ procedure AMC can adjust the transmission rate for each TTI according to the received CQI value, another mechanism is still needed to help AMC throttling the decision in each TTI such that the system could reach higher utilization. Since the AMC makes decision according to the instant channel quality information only, combining the past information into the decision procedure seems necessary. Based on this concept, we adopt a learning scheme into the design of the H-ARQ procedure. Observing the progress of the H-ARQ procedure it can be found that the decision making is based on the current channel information, and the current channel information correlates with only the previous channel information and decision. Therefore, it can be concluded that the H-ARQ procedure can be modeled as a discrete-time Markov decision process (MDP).

In the conventional H-ARQ scheme, the relation between the feedback channel quality information and the output decision of modulation and coding scheme (MCS) is already determined. It implies the service provider should have the pre-knowledge of the thorough system. To achieve more system efficiency, a new H-ARQ procedure based on Q-learning algorithm [11] has been proposed in [12]. The Q-learning algorithm is a reinforcement learning method with a so-called Q-value function (or

Q-function). The Q-value function is an evaluation function, which could be taken as the expected total discounted cost counting from the initial transmission. A system with Q-learning can choose proper action according to the updated Q-values and finally reach a convergent solution. In other words, eventually the system in each state will choose the best suitable action correspondingly after learning process. The simulation results of [12] have shown that Q-learning based H-ARQ scheme can improve the system throughput over conventional schemes.

In the Q-learning based H-ARQ scheme [12], the author tries to find a best suitable decision for each input state and the object is to satisfy the system QoS requirement. But sometimes we could take another decision which is not the best suitable one to increase the system throughput furthermore still satisfy the system QoS requirement. In other words, we want the system to give judicious guidance according to circumstances, or say, make a more human-based decision. Therefore, we consider that the fuzzy system would be helpful to achieve the object. Fuzzy system has the advantage of simplifying the complication relation between inputs and outputs of a system based on the pre-knowledge of the system. It turns the absolute relation between inputs and outputs into the comparative relation, such that the system can reach a slightly worse performance but much faster procedure. To combine the advantages of Q-learning algorithm and fuzzy system, a novel learning algorithm called fuzzy Q-learning (FQL) algorithm was proposed in [14]. It can be thought as efficient adaptations of Q-learning with fuzzy rules. With the knowledge representation ability of fuzzy rules, prior knowledge can be incorporated into the initial set of rules and then accelerate the learning stage. It has been shown in [15] that the initialization of Q-values by fuzzy rules can exactly accelerate a system with Q-learning technique.

In this paper, we propose a fuzzy Q-learning (FQL) based H-ARQ scheme for

HSDPA systems. Since for the H-ARQ scheme the system throughput is influenced by the number of the re-transmissions, we try to find the best suitable action for each transmission time interval (TTI). Here the action is defined as the modulation order and coding rate to transmit data. We try to describe the system state through fuzzy theorem, and then combine the Q-learning algorithm to find a suitable action for each state. Then we can decrease the number of retransmissions effectively and get high throughput. To verify the result, we will simulate the system throughput later and compare it with other the conventional H-ARQ scheme.

Chapter 2

System Model

_____

In this paper, the HSDPA system in Release 5 is considered. We try to make suitable decision for H-ARQ to decrease the number of retransmissions. To simplify the problem, we take the channel state as the input variable and the modulation order and coding rate as the output variables in H-ARQ learning scheme. Notice that the channel state includes the measured signal to interference and noise ratio (SINR); candidates of modulation order are QPSK and 16QAM, and candidates of coding

Table 2-1 : The peak data rates of all modulation and coding schemes

Modulation level	Coding rate	Peak data rate
QPSK	$1/3$	2.4 Mbps
	$1/2$	3.6 Mbps
	$2/3$	4.8 Mbps
	$3/4$	5.4 Mbps
	1	7.2 Mbps
16QAM	$1/2$	7.2 Mbps
	$2/3$	9.6 Mbps
	$3/4$	10.8 Mbps
	1	14.4 Mbps

rates are ($\frac{1}{3}$, $\frac{1}{2}$, $\frac{2}{3}$, $\frac{3}{4}$, 1). Table 2-1 shows the total candidates of actions that can be adopted in our proposed scheme. In this thesis, there are total 9 candidates of actions.

2.1 HSDPA System

HSDPA is designed to increase downlink packet data throughput by means of fast physical layer (L1) retransmission and transmission combining, as well as fast link adaptation controlled by the Node B. Notice that the retransmission for HSDPA is processed in the Node B instead of in the radio network control (RNC). The advantage is the faster retransmission and shorter delay when a retransmission is needed. Figure 2.1 shows the difference of retransmission handling in HSDPA between Release 5 and Release'99.

The medium access control (MAC) layer protocol architecture of HSDPA is shown in Figure 2.2. A new MAC functionality called MAC-hs is added in the Node B. The MAC-hs is to handle the automatic repeat request (ARQ) functionality, scheduling, as well as priority handling.
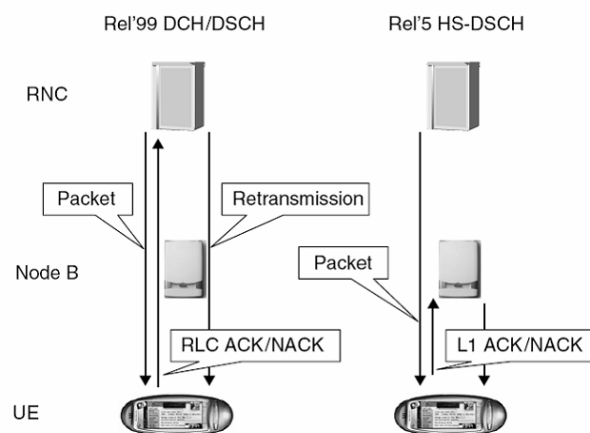

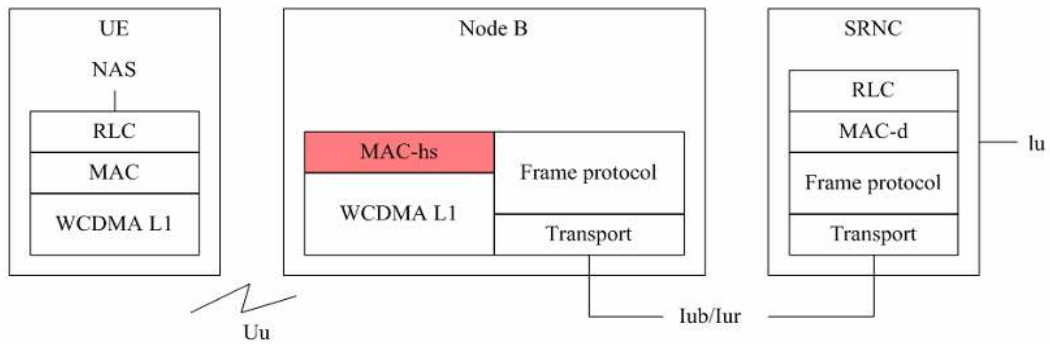Figure 2.1 : Release'99 and Release 5 HSDPA retransmission control

Figure 2.2 : HSDPA protocol architecture

In HSDPA, there are three new channels introduced in the physical layer specification. They are high-speed downlink shared channel (HS-DSCH), high-speed shared control channel (HS-SCCH), and uplink high-speed dedicated physical control channel (HS-DPCCH).

HS-DSCH carries the data to the user in the downlink direction with the peak rate reaching up to 14Mbps with 16QAM. The TTI or interleaving period has been defined to be 2ms (three slots) to achieve a short round rip delay between the Node B and the terminal for retransmission. The spread factor (SF) is always fixed at 16, and multi-code transmission as well as code multiplexing of different users can take place in HS-DSCH. This means that the maximum number of codes which can be allocated is 15 since there is a need to have code space available for common channels. The only coding scheme of HS-DSCH is turbo code. To achieve the multiplexing coding gain, HARQ functionality is added to vary the transport block size, modulation scheme and the number of multicodes.

HS-SCCH carries the necessary physical layer control information to decode the data of HS-DSCH and perform the possible physical layer combining of the data sent on HS-DSCH when a retransmission is needed. If there is no data on the HS-DSCH, there is no need to transmit the HS-SCCH. Each HS-SCCH block has three slots and is divided into two functional parts. The first slot carries the time-critical information

needed to start the demodulation process in due time to avoid over-buffer in chip level. The remaining two slots contain less time-critical parameters, like cyclic redundancy check (CRC), to check the validity of HS-DSCH information and HARQ process information. Figure 2.3 shows the timing relationship between HS-SCCH and HS-DSCH. From this figure, we can see that the terminal has time duration of one slot to determine which codes to de-spread from the HS-DSCH.
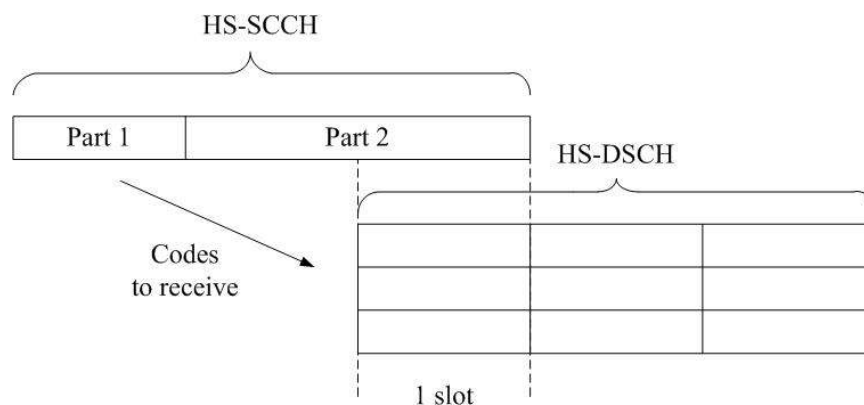
Figure 2.3 : HS-SCCH and HS-DSCH timing relationship

HS-DPCCH carries the necessary control information in the uplink direction and is divided into two parts, which carries ACK/NACK messages and downlink quality feedback information, respectively. The second part is also called channel quality indicator (CQI) feedback. The information on HS-DPCCH can be used by the Node B scheduler to decide which terminal to transmit and at what data rate.

The HSDPA physical layer operation goes through the following steps:

(i) The scheduler in the Node B evaluates several channel quality information for different users. The information includes the channel conditions, how much data is pending in the buffer for each user, for which users retransmissions are pending and how much time has elapsed since a particular user was last served and so forth.

(ii) Once a terminal has been determined to be served in a particular TTI, the

Node B identifies the necessary HS-DSCH parameters. These parameters are, for example, the number of available codes, the kind of modulation order that can be used and the terminal capability limitations. The terminal soft memory capability also defines which kink of HARQ can be used.

(iii) The Node B starts to transmit the HS-SCCH two slots before the corresponding HS-DSCH TTI to inform the terminal of the necessary parameters. The HS-SCCH selection is free if there was no data for the terminal in the previous HS-DSCH frame.

(iv) The terminal monitors the HS-SCCHs given by the network. While the terminal has decoded part 1 (as shown in Figure 2.3) from an HS-SCCH intended for that terminal, it will start to decode the rest of that HS-SCCH and will buffer the necessary codes from the HS-DSCH.

(v) Until the HS-SCCH parameters has been decoded from part 2, the terminal can determine which H-ARQ process the data belongs to and whether it needs to be combined with data received previously in the soft buffer.

(vi) After decoding the combined data, the terminal sends an ACK/NACK indicator in the uplink HS-DPCCH, depending on the outcome of the CRC check conducted on the HS-DSCH data.

(vii) If the network continues to transmit data for the same terminal in consecutive TTIs, the terminal will stay on the same HS-SCCH which was used during the previous TTI.

2.2 HARQ Scheme

HARQ plays an important role in HSDPA. It can combine the previous packet and redundancy from present packet to help decoding if received a NACK signal.

Figure 2.4 shows an example of bit allocation in adjacent TTIs. By increasing the redundancy for previous failure packet, it can improve the probability of successful decoding of previous packet. With an appropriate HARQ functionality, the delay time of retransmission can be reduced. HARQ can determine the transport block size, modulation scheme, and the code rate according to the channel quality information received from the CQI on HS-DPCCH. The CQI feeds the corresponding modulation scheme and coding rate back to the Node B scheduler according to the CQI table. The CQI table stores the information that which modulation order (QPSK or 16QAM) and coding rate are suitable for the channel condition. CQI sends the information to the Node B based on the instant channel condition. The Node B receives the information from CQI and determines the exact (re)transmission modulation order and coding rate since the information is just a suggestion for the Node B. By changing the modulation order according to the instant channel condition, HARQ can lessen the number of retransmission effectively and improve the throughput.
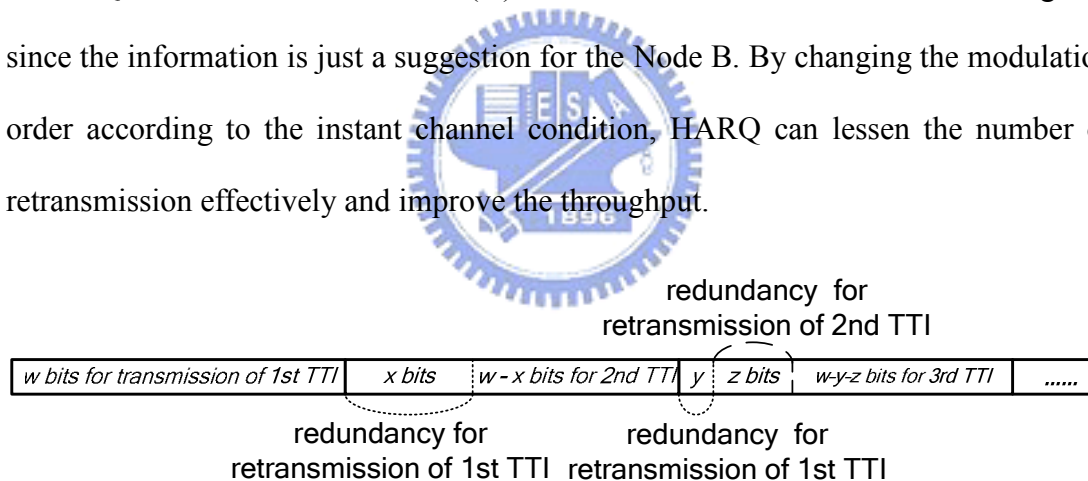

Figure 2.4 : The allocation of bits in adjacent TTIs

2.3 Channel Model

The channel condition in WCDMA at time t, denoted by $F(t)$, is modeled mainly by long-term fading and short-term fading, which is shown as:

$$F(t) = \xi(r) \times 10^{\eta/10} \times \zeta(t), \tag{2-1}$$

where $\xi(r) \times 10^{\eta/10}$ is the long-term fading including path loss and shadowing, r is the distance between base station and user equipment (UE), η is the normal-distributed random variable with zero mean and variance σ_L^2, and $\zeta(t)$ is the short-term fading caused by multi-path effect. The $\zeta(t)$ is assumed to be with Jakes model, which is given by

$$\zeta(t) = 2\sigma\sqrt{\frac{2}{L}}\sum_{m=1}^{M}\cos(2\pi f_D t \cos(2\pi m/L) + \theta_m)e^{j\beta_m}, \tag{2-2}$$

where σ is the radical of the average power signal, L is the number of signal path and equals to $4M+2$, $\beta_m = \pi m/(M+1)$ and

$$\theta_m = \beta_m + 2\pi ms/(M+1), \quad s = 0, 1, 2, \cdots, M-1. \tag{2-3}$$

Since $\zeta(t)$ is mutually independent, it can produce up to M independent short-term fading. Hence we could choose M equal to the number of the total links in all cells of the system.

The shadowing effect for a user will be changed according to the motion of the user. But in practical, the degradation degree of shadowing between two sampling time is small since the duration of two sampling time in HARQ is very short compared to the motion of the user. As mentioned above, the shadowing effect of these sampling points are highly correlated. We assume that the model of the correlated shadow fading is defined as a normalized autocorrelation function $\rho(\Delta x)$, where Δx is the distance between two adjacent TTI. The $\rho(\Delta x)$ can be obtained by

$$\rho(\Delta x) = e^{-\frac{|\Delta x|}{d_{cor}}\ln 2}, \tag{2-4}$$

where d_{cor} is the decorrelation length.

Chapter 3

Fuzzy Q-learning based H-ARQ Scheme

Conventional H-ARQ procedures try to select a proper modulation and coding scheme for transmission according to the CQI value from UE, but they cannot adapt the modulation and coding scheme (MCS) of transmissions to the channel condition precisely. The reason is that the range of CQI value is confined and obviously cannot represent entire possible channel condition precisely. We try to find an advanced scheme so that the adaptation between the channel quality and the MCS can be more precisely. As shown in Figure 3.1, we can consider a learning system interacting with its environment to accommodate the rate adjustment to the entire possible channel condition. We take the practical channel condition as the system state and the MCS adopted by Node B as the action. Then the object is to find an appropriate cost function to tell Node B whether the previous action is suitable or not.

To implement the advanced H-ARQ scheme with learning ability, first we should define the system state and the action. It is found that the signal to interference and noise ratio (SINR) is the most intuitional parameter to estimate the channel condition, and the block error rate (BLER) is another important parameter to be guaranteed. Therefore we choose the measurements of these parameters as the system state. On the other hands, the action is intuitively defined as the MCS since it is the major factor to influence the transmission rate. The last parameter we should define is the cost function. The cost function needs to reflect the

condition whether the applied action is suitable or not. We will propose it explicitly in later section.

Before the design of the FQL-based H-ARQ scheme, we briefly expound the concept of fuzzy Q-learning algorithm first.
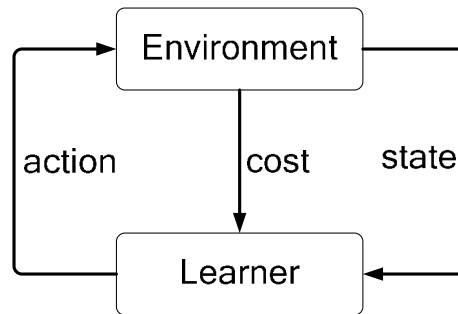

Figure 3.1 : Block diagram of a learning system
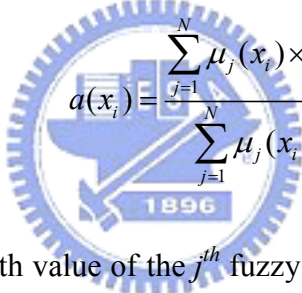
3.1 Fuzzy Q-learning Algorithm

The fuzzy Q-learning algorithm combines the fuzzy theorem and the learning ability of Q-learning algorithm to help the system adapting the channel states and the adopting actions. Since the FQL is based on the Q-learning, the learning processes of them are very similar. The major difference between these two algorithms is that the system state of Q-learning is described by the fuzzy system. We will illustrate the detail of the FQL in the following.

Owing to the property of the H-ARQ process, the input state information can be regarded as a Markov decision process (MDP). We denote **S** as the set of state vectors described by the fuzzy inference system (FIS) and $\mathbf{S} = \{S_j, j = 1, 2, \ldots, N\}$. Each state vector S_j is constituted by M fuzzy linguistic variables to describe the system. The set vector of the candidate output actions is denoted as **A,** and $\mathbf{A} = \{A_k, k = 1, 2, \ldots, K\}$. The input state vector received in i^{th} time period which is denoted as x_i. Then the rule representation of FQL for state S_j is in the following form：

$$\textbf{if } x_i \textbf{ is } S_j, \textbf{ then } A_k \textbf{ with } q(S_j, A_k), \ 1 \le j \le N \text{ and } 1 \le k \le K, \qquad (3\text{-}1)$$

where A_k is the kth candidate action that is possibly chosen by state S_j, and $q(S_j, A_k)$ is the Q-value for the rule-action pair (S_j, A_k). The number of rule-action pairs for each state S_j is equal to the number of the elements in the action set. It means that there are K possible consequence parts for the same antecedent and total $N \times K$ Q-values since N fuzzy rules. Every fuzzy rule has to choose an action from the action candidate set $\textbf{A}$ by a selection policy. In the FQL, the action selection policy for each fuzzy rule we adopted is *select–max* strategy. The action selection policy in FQL is called exploration/exploitation policy (EEP).

As to the defuzzification of the N fuzzy rules, the global inferred action $a(x_i)$ for the input vector x_i is expressed as

$$a(x_i) = \frac{\displaystyle\sum_{j=1}^{N} \mu_j(x_i) \times a_j}{\displaystyle\sum_{j=1}^{N} \mu_j(x_i)}, \qquad (3\text{-}2)$$

where $\mu_j(x_i)$ is the output truth value of the jth fuzzy rule obtained through the membership function if we have received the state information set x_i, and a_j is the action selected by EEP of jth fuzzy rule. Also, the Q-value for the state-action pair $(x_i, a(x_i))$ in i^{th} time period is

$$Q_i(x_i, a(x_i)) = \frac{\displaystyle\sum_{j=1}^{N} \mu_j(x_i) \times q_i(S_j, a_j)}{\displaystyle\sum_{j=1}^{N} \mu_j(x_i)}. \qquad (3\text{-}3)$$

Here $q_i(S_j, a_j)$ is the Q-value for rule-action pair (S_j, a_j) in the ith time period.

For the current system state x_i, after applying the chosen action $a(x_i)$, the next-stage system state is x_{i+1} and the system reinforcement signal is $r(x_i, a(x_i))$. To update the Q-value, the next-stage optimal Q-value, denoted by $Q^*(x_{i+1}, a(x_{i+1}))$, is defined as

15

$$Q_i^*\left(x_{i+1}, a(x_{i+1})\right) = \frac{\sum\limits_{j=1}^{N} \mu_j(x_i) \times q_i(S_j, a_j^*)}{\sum\limits_{j=1}^{N} \mu_j(x_i)}, \tag{3-4}$$

where $q_i\left(S_j, a_j^*\right)$ is the Q-value of rule-action pair $\left(S_j, a_j^*\right)$ and $a_j^* = \arg\max\limits_{A_k}\left\{q_i\left(S_j, A_k\right)\right\}$.

According to the Q-learning rule [11], the Q-value updated in the FQL can be expressed as

$$q_{i+1}\left(S_j, a_j\right) = q_i\left(S_j, a_j\right) + \eta \times \Delta q_i\left(S_j, a_j\right), \tag{3-5}$$

where η is the learning rate in the region [0, 1] and

$$\Delta q_i\left(S_j, a_j\right) = \frac{\mu_j(x_i)}{\sum\limits_{j=1}^{N} \mu_j(x_i)} \times \left\{r\left(x_i, a(x_i)\right) + \gamma \times Q_i^*\left(x_{i+1}, a(x_{i+1})\right) - Q_i\left(x_i, a(x_i)\right)\right\}. \tag{3-6}$$

Here γ is the discount factor.

In summary, the procedure of FQL algorithm is shown in Figure 3.2. The system will repeat Steps 2~7 iteratively if it receives state information in each time period. That means the FQL algorithm is implemented frame by frame.
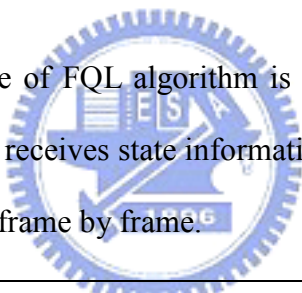
Step 1:	Initialize Q-values for all fuzzy rules
Step 2:	Receive the state information vector x_i
Step 3:	For each fuzzy rule, choose the rule consequence a_j by the EEP
Step 4:	Compute the global consequence $a(x_i)$ and its corresponding Q-value $Q_i(x_i, a(x_i))$ using (3-2) and (3-3)
Step 5:	Apply the action $a(x_i)$ to system
Step 6:	Receive the reinforcement $r(x_i, a(x_i))$
Step 7:	Update $q_i(S_j, a_j)$ for all j using (3-4), (3-5) and (3-6)

Figure 3.2 : Procedure of FQL algorithm

3.2 Design of FQL based H-ARQ

In this paper we incorporate the fuzzy Q-learning algorithm into the design of the H-ARQ scheme. The FQL combines the benefits of fuzzy system and reinforcement learning. The fuzzy system provides a good function approximation for the FQL and a prior knowledge can be easily applied to the system design. By applying FQL, the system can be managed with uncertain information. From the measured state information in each TTI, the H-ARQ mechanism should select a suitable modulation and coding scheme to transmit information data. In the design, the state information defined here means the channel quality indicator (CQI) and the block error rate (BLER). After we have made a decision, we can know how much information bits will be transmitted in this TTI. So the output signal of the H-ARQ scheme would be able to change the modulation and coding scheme and trim the number of information bits which will be transmitted in each TTI.

The H-ARQ scheme with FQL is shown as Figure 3.3. H-ARQ segments arrival packets according to the modulation order and coding rate of transmission that the FQL-based H-ARQ determines. The two-stage rate matching is to match the physical layer transmission rate based on the determined action and to buffer the retransmission data packet for IR request. The CRC attachment and the interleaving technique are used for the error detecting and against errors. The turbo coding with a minimum 1/3 code rate is employed in this paper.
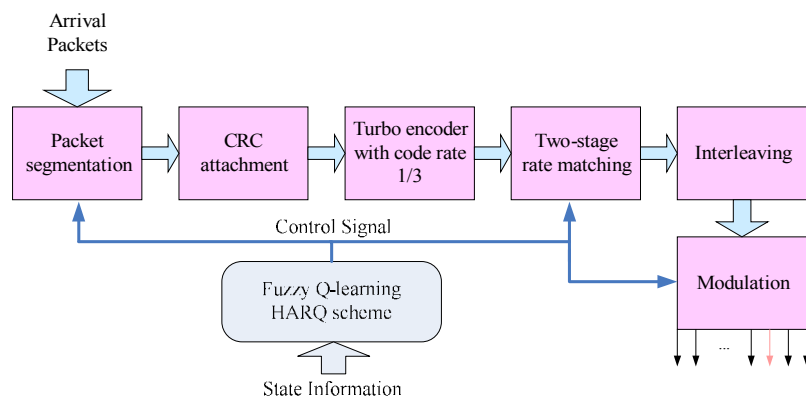

Figure 3.3 : The FQL-based H-ARQ for packet transmission in HS-DSCH

Figure 3.4 shows the block diagram of the FQL-based H-ARQ scheme. The FIS describes the system state using the channel information which is received from the HS-DPCCH. The EEP block decides the action of each fuzzy rule according to the Q-values of the fuzzy rules. With the membership values of the total fuzzy rules, the inferred action is output to the controller block, and the controller block translates it to the control signal to the H-ARQ procedure. After applying the global inferred action, there is a reinforcement signal to reflect the inferred action is suitable to the previous system state or not. Then the Q-learning block updates the Q-values of the fuzzy rules with the reinforcement value.

The state information in Figure 3.4 is the state vector which has been mentioned in the detail of FQL. In our proposed scheme, we choose the signal to interference and noise ratio (SINR) and the ratio of block error rate (BLER) to its requirement value as the input parameters to describe the system state. Instead of CQI value, we choose the measured SINR value as the input of the scheme to achieve more accurate calculation. We define $SINR_i$ is the measured value of SINR at the beginning of a TTI, and R_i as
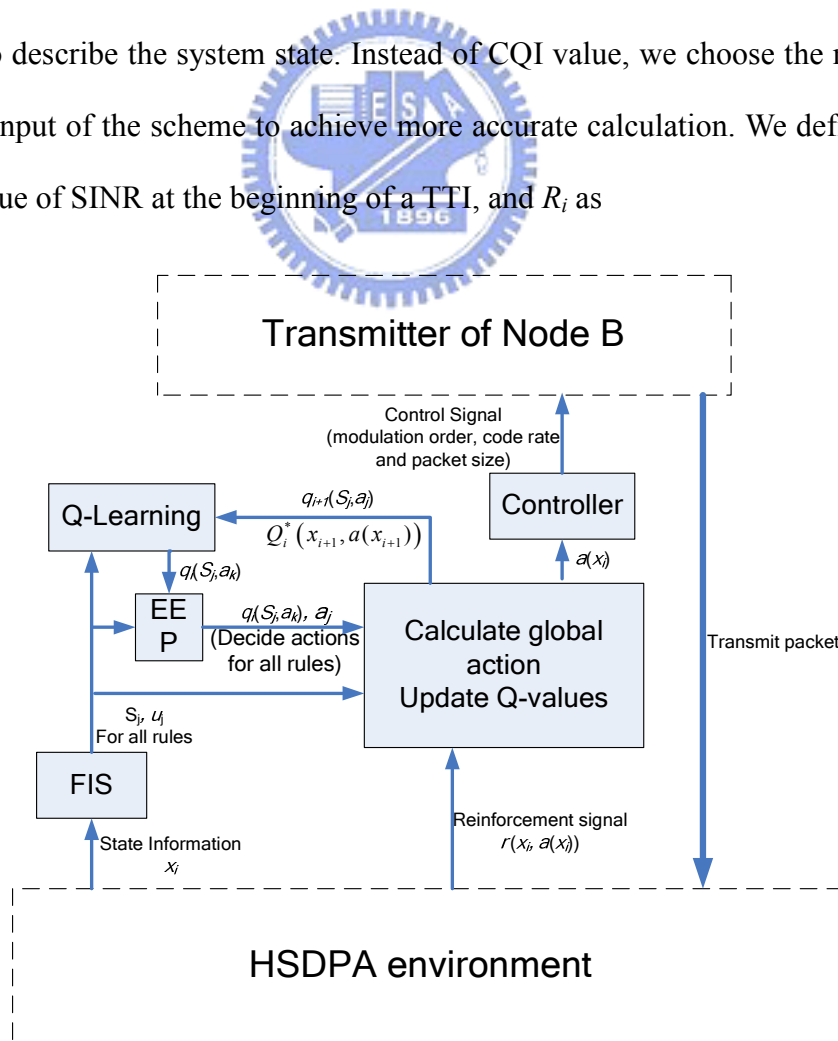

Figure 3.4 : Block diagram of the FQL-based H-ARQ scheme

$$R_i = \frac{BLER_i}{\widetilde{BLER}}. \tag{3-7}$$

Here the subscript i means the ith TTI and $\widetilde{BLER}$ is the requirement of the BLER value. The value R_i can clearly represent the degree of disparity between actual BLER and the guaranteed value.

We denote the linguistic variables of SINR and the ratio of BLER as L_{SINR} and L_R, respectively. Then the input state information set x_i is defined as

$$x_i = \left\{ L_{SINR_i}, L_{R_i} \right\}, \tag{3-8}$$

where the suffix i means the ith TTI. We use 9 terms for L_{SINR} and 2 terms for L_R. Here we define their fuzzy term sets as T(L_{SINR}) = {Extremely High, Very High, High, Slightly High, Medium, Slightly Low, Low, Very Low, Extremely Low } = {EH, VH, H, SH, M, SL, L, VL, EL}, and T(L_R) = {Medium, Small} = {M, S}. From the fuzzy set theory, the fuzzy rule base forms with dimensions $\left| T(L_{SINR}) \right| \times \left| T(L_R) \right|$. It means that the total number of fuzzy rules is equal to 18. Here the fuzzy term set of L_{SINR} is according to the desired SINR value of all action candidates under the guaranteed BLER condition.

On the other hand, the output action of FQL-based H-ARQ scheme is the modulation order and coding rate. From the combination of the two kinds of modulation order (QPSK and 16QAM) and five values of coding rate $\left\{ \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, 1 \right\}$, we can find 9 action candidates for each fuzzy rule, which has been shown in Table 2-1.

Since each modulation order with a coding rate corresponds to a specific peak data rate, we can build a table showed as Table 2-1, which shows the modulation and coding schemes that can be adopted for each state and their corresponding peak data rates. Here the fuzzy term set of the output action is equal to the action candidate set **A**, $\mathbf{A} = \left\{ A_k, k = 1, 2, \dots, 9 \right\}$, and A_k corresponds to Table 2-1 from lowest peak data rate to highest one..

The reinforcement signal should be an indicator to the system performance or related ones. Since the object is to keep the BLER under or equal to the requirement value, the reinforcement signal should be defined to reflect this condition. Based on this concept, we define the reinforcement signal from the function showed as follows:

$$
r\left(x_i, a(x_i)\right) = \begin{cases} \left(\dfrac{BLER\left(\widetilde{SINR}_{A0}, a(x_i)\right) - \widetilde{BLER}}{\widetilde{BLER}}\right)^2 & \text{if } SINR_{x_i} < \widetilde{SINR}_{A0} \\[3ex] \left(\dfrac{BLER\left(\widetilde{SINR}_{A9}, a(x_i)\right) - \widetilde{BLER}}{\widetilde{BLER}}\right)^2 & \text{if } SINR_{x_i} > \widetilde{SINR}_{A9} \\[3ex] \left(\dfrac{SINR_{x_i} - \widetilde{SINR}_{a(x_i)}^{3\times\widetilde{BLER}}}{\widetilde{SINR}_{a(x_i)}^{3\times\widetilde{BLER}}}\right)^2 & \text{if } \widetilde{SINR}_{A0} < SINR_{x_i} < \widetilde{SINR}_{A9} \ \& \ BLER < 0.9\times\widetilde{BLER} \\[3ex] \left(\dfrac{BLER\left(SINR_{x_i}, a(x_i)\right) - \widetilde{BLER}}{\widetilde{BLER}}\right)^2 & \text{otherwise} \end{cases}
$$

(3-9)

where $BLER\left(SINR_{x_i}, a(x_i)\right)$ is the theoretical BLER value of action $a(x_i)$ under the measured SINR value of state x_i. The notation $\widetilde{SINR}_{Ak}$ is the theoretical SINR value when applying action A_k under the requirement BLER, while the notation $\widetilde{SINR}_{a(x_i)}^{3\times\widetilde{BLER}}$ is the theoretical SINR value when applying global action $a(x_i)$ under the value of BLER three times the requirement value 0.1. The physical meaning of the reinforcement function is to adopt the actions which can reach the requirement BLER under different channel conditions. We define two boundary conditions to speed up the learning process since we already have the pre-knowledge of the system. If the BLER is smaller than the requirement value, we can venture to take an action with higher code rate to increase the system throughput. Under such condition, we design the reinforcement function to choose the action whose theoretical BLER value under the measured SINR value is three times the requirement value as far as possible. Notice that we use the normalized SINR value instead of normalized BLER value for the reason that the trend of the theoretical BLER versus SINR is not linear. It implies if SINR is

higher than a specified value, the normalized BLER values of different actions will make no difference. But the normalized BLER value is more effective than the normalized SINR value to guarantee the requirement value, so we still adopt it as the reinforcement for other cases. Hence we define that an action with lowest reinforcement value is most suitable. In other words, eventually the system would adaptive apply an action according to the channel conditions to satisfy the QoS and keep high throughput as far as it can.

The $BLER\left(SINR_{x_i}, a(x_i)\right)$ can be obtained by the following equation:

$$BLER\left(SINR_{x_i}, a(x_i)\right) = 1 - P_e^B.$$ (3-10)

Here the superscript B is the block size while the action $a(x_i)$ is applied, and P_e is the theoretical bit error rate of action $a(x_i)$ with the measured SINR value. For simplicity, we assume that a block would be regarded as a failure block as long as there is one error bit in this block. The theoretical uncoded bit error rate $P_{e,uncoded}$ is calculated from [2]

$$P_{e,uncoded} = \begin{cases} \dfrac{1}{2} erfc(\sqrt{SINR \times 16}) & , \quad \text{for QPSK} \\ \dfrac{3}{4} \times p - \dfrac{9}{16} \times p^2 & , \quad \text{for 16QAM} \end{cases}$$ (3-11)

and $p = \dfrac{1}{2} erfc(\sqrt{SINR \times \dfrac{32}{5}})$. To model the effect of the error correcting code, we make an assumption that the relation between the code rate and SINR is linear. In other words, the performance of a code rate 0.5 is equal to the uncoded performance with two times SINR value. Hence we can find the BLER values of transmissions using different modulation and coding schemes under any measured SINR value through this equation.

After received the system state information, the EEP would decide the output actions for all fuzzy rules based on the Q-values. Since each fuzzy rule represents a specific system state, the Q-values of all fuzzy rules represent the degree of suitability for their corresponding adopting actions and system states. After applying an action, the reinforcement signal is fed

back to be used to update the Q-values of those related fuzzy rules. Then the Q-value of an unsuited action would be decreased rapidly by the reinforcement signal. Eventually the most suitable action of a fuzzy rule would be with the largest Q-value and be always chosen by the EEP in that fuzzy rule.

3.3 Membership Function Definition

Finally we would define the membership functions of all fuzzy inputs. To illustrate the membership function, first we define a function $f(\cdot)$ which is expressed as

$$f(y; y_0, y_1, y_2) = \begin{cases} 1 + \dfrac{y - y_1}{y_1 - y_0} & \text{for } y_0 \leq y \leq y_1 \\ 1 + \dfrac{y - y_1}{y_1 - y_0} & \text{for } y_1 \leq y \leq y_2 \\ 0 & \text{otherwise} \end{cases} \tag{3-12}$$

where y_0 (or y_1, y_2) in $f(\cdot)$ is the left edge (or center, right edge) of the triangular function. The definition of this function is shown as Figure 3.5.
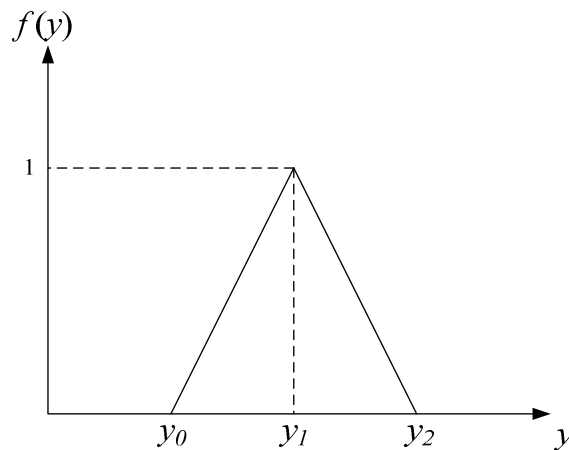

Figure 3.5 : Definitions for function $f(.)$

For the purpose of simplification, we set the membership function of the SINR input as triangle functions. Note that we build the input SINR levels according to the actions we can

adopt. We define a new notation as $\widetilde{SINR}_{A_k}^{0.5}$ to notice the theoretical SINR value reaching a target BLER value 0.5 with action A_k. Then the input SINR membership functions can be expressed as follow:

$$\mu_{EL}(SINR) = f\left(SINR; \ -\infty, \ \widetilde{SINR}_{A_1}, \ \widetilde{SINR}_{A_2}\right) \tag{3-13}$$

$$\mu_{VL}(SINR) = f\left(SINR; \ \widetilde{SINR}_{A_2}^{0.5}, \ \widetilde{SINR}_{A_2}, \ \widetilde{SINR}_{A_3}\right) \tag{3-14}$$

$$\mu_{L}(SINR) = f\left(SINR; \ \widetilde{SINR}_{A_3}^{0.5}, \ \widetilde{SINR}_{A_3}, \ \widetilde{SINR}_{A_4}\right) \tag{3-15}$$

$$\mu_{SL}(SINR) = f\left(SINR; \ \widetilde{SINR}_{A_4}^{0.5}, \ \widetilde{SINR}_{A_4}, \ \widetilde{SINR}_{A_5}\right) \tag{3-16}$$

$$\mu_{M}(SINR) = f\left(SINR; \ \widetilde{SINR}_{A_5}^{0.5}, \ \widetilde{SINR}_{A_5}, \ \widetilde{SINR}_{A_6}\right) \tag{3-17}$$

$$\mu_{SH}(SINR) = f\left(SINR; \ \widetilde{SINR}_{A_6}^{0.5}, \ \widetilde{SINR}_{A_6}, \ \widetilde{SINR}_{A_7}\right) \tag{3-18}$$

$$\mu_{H}(SINR) = f\left(SINR; \ \widetilde{SINR}_{A_7}^{0.5}, \ \widetilde{SINR}_{A_7}, \ \widetilde{SINR}_{A_8}\right) \tag{3-19}$$

$$\mu_{VH}(SINR) = f\left(SINR; \ \widetilde{SINR}_{A_8}^{0.5}, \ \widetilde{SINR}_{A_8}, \ \widetilde{SINR}_{A_9}\right) \tag{3-20}$$

$$\mu_{EH}(SINR) = f\left(SINR; \ \widetilde{SINR}_{A_9}^{0.5}, \ \widetilde{SINR}_{A_9}, \ \infty\right) \tag{3-21}$$

They are showed as Figure 3.6. As mentioned above, we can calculate the expected SINR values for all actions and set these values corresponding to linguistic levels respectively. Hence the $\widetilde{SINR}_{A_k}$ (or $\widetilde{SINR}_{A_k}^{0.5}$) is defined as the desired SINR value using action A_k to reach a statistic BLER equal to the requirement value (or 0.5). With the pre-knowledge of the relation between candidates of actions and the channel conditions, we set the theoretical SINR values of actions which can reach the specific BLER value 0.5 as the left boundaries of SINR levels. It is for the reason that an action should be adopted only if there is no other action more suitable than this action under any channel condition. Since for a SINR level there is a best suitable action corresponding to it, it is very intuitional that if the input SINR is very close to the desired SINR value corresponding to this action, the matched degree of this level

should be the highest value.

For the other input (the ratio of the measured BLER over the requirement value), we set the member function as Figure 3.7. We simply divide it into two regions for the reason that we expect to adopt a venturesome action or not. For the "Small" level, we expect to adopt an action with higher code rate to reach higher throughput. And for the "Medium" level, we prefer to adopt an action which can help the system satisfying the QoS requirement. So we define the membership functions of BLER ratio input as follow:

$$\mu_S(R) = f(R;\ 0,\ 0.8,\ 1) \tag{3-22}$$

$$\mu_M(R) = f(R;\ 0.8,\ 1,\ \infty) \tag{3-23}$$

The membership value of the output membership function for each rule is the mixed union of the fuzzy sets assigned to that output in a conclusion after cutting their degree of membership values at the degree of the membership for the corresponding premise. That is, the membership value of a fuzzy rule is defined as

$$\mu = \min\left(\mu_{L_{SINR}},\ \mu_{L_R}\right). \tag{3-24}$$

In our design, the output of each fuzzy rule is the action decided by the EEP which is regarded as the desired action in that scene. We adopt the center of area (CoA) defuzzification method to calculate the final output action. For simplification, we set the membership function of the output as an impulse function with unit length.

For the procedure of defuzzification, we adopt the max-min inferred method to get the truth value of each corresponding fuzzy rule. Through the center of gravity (COA) method, a crisp value is exported to decide which action will be applied. Since this crisp value usually not corresponds to any value of an action candidate, a quantized value is output to apply an action.
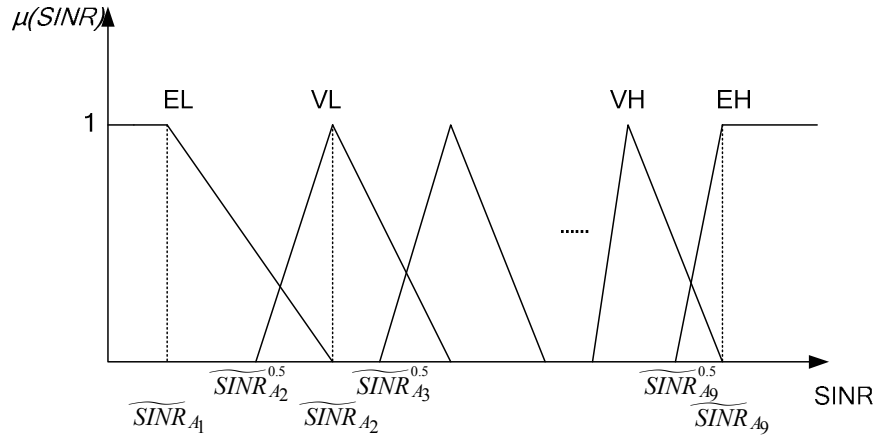
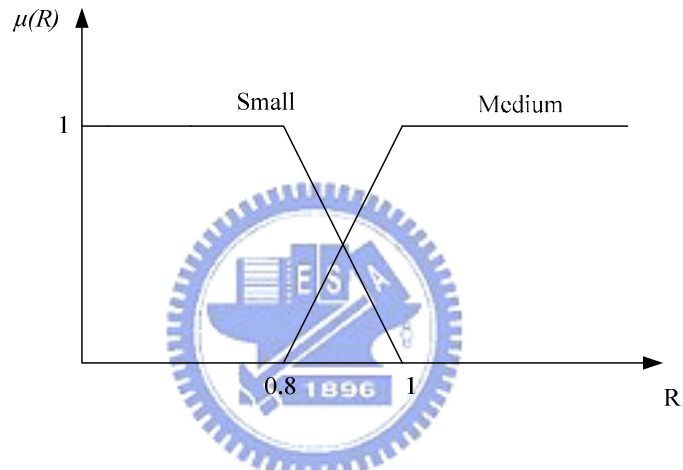Figure 3.6 : The membership function of the SINR input

Figure 3.7 : The membership function of the BLER ratio input

Chapter 4

Simulation Results

4.1 System Environment and Parameters

In our simulation, we consider a hexagonal grid cell structure. There are 19 base stations (BS) in the multi-cell system. For a HSDPA user, we assume that the HS-DSCH is allocated with total transmission power up to 80%. The power received from other cells is considered as the interference signal of the users in the own cell. For a HSDPA link, we consider the long-term fading and the short-term fading. The long-term fading consists of two components. They are shadow effect and path loss respectively. We consider the correlation in time for the same HSDPA link to describe the real environment since the mobility of the users is considered in our system. On the other hand, the short-term fading we adopt is the Jakes model. The total simulation parameters are shown as Table 4-1.

The total period of the propagation delay and the processing delay is assumed to be 6ms. It implies that the Node B requires two TTIs to receive an ACK/NACK from the UE after two additional TTIs. To simplify the simulation complexity and reach higher data rate, we assume the users always have data to be transmitted.

Table 4-1 : Simulation parameters

| Parameter	Assumption
Cellular layout	Hexagonal grid, 19 sites, 2000m cell radius
Path loss model $(\xi(r))$	$128.1 + 37.6\log_{10}(r)$ r is the base station separation in kilometers
Decorrelation length (d_{cor})	20m
σ_L	8.0
Mobility assignment	0, 20, 40, 60 km/hr, random distribution
Carrier frequency	2.0 GHz
Channel bandwidth	5.0 MHz
Chip-rate	3.84 Mcps
Spreading factor	16
Thermal noise density	-174 dBm/Hz
TTI length	2 ms
Number of UE in one cell	4, random distribution
Discounted factor (γ)	0.9
Learning rate (η)	0.2
BS total Tx power	Up to 44 dBm
Power for HSDPA data transmission	Maximum of 80% of total maximum available transmission power

4.2 Conventional Schemes

In the simulation, we will compare the proposed scheme with other conventional schemes. The purpose of the proposed scheme tries to adjust the initial transmission rate adaptive according to the received channel information. After the training process, the H-ARQ scheme could adopt suitable action to satisfy the QoS requirement and keep high throughput. The intuitive conventional schemes are incremental redundancy (IR) scheme and its advance schemes, which are proposed in [17]. From [17], the author proposes the two advanced schemes combined with link adaptation (LA) and IR protocol, and shows that they can reduce the number of transmission times effectively. This implies that the two schemes can adjust adaptively the initial

transmission rate to avoid retransmissions. We also compare our proposed scheme with Q-learning based scheme [12] since these two schemes are both learning schemes.

The two advanced IR schemes are described as follow:

- LA_IR_1: The code rate of the initial transmission is the one incremental version of the most recent code rate when a successfully decoding occurs without IR request. This scheme tries to make the initial code rate tracking the variation of the channel condition. It makes the initial code rate always retreat one version and then maintain the throughput of the conventional IR.

- LA_IR_2: The code rate of the initial transmission is to keep the current decoded code rate as the next initial code rate whenever a successful decoding occurs. If there is no IR request for two consecutive times, it takes the two incremental version of the current code rate as the next initial code rate.
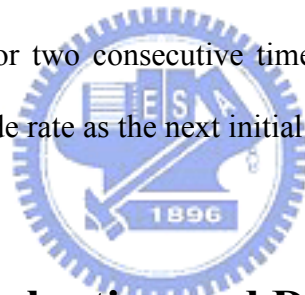
4.3 Performance Evaluation and Discussions

To show the simulation result, we define a parameter called signal to other services' interference ratio (SIOR) to describe the ratio of the transmitting power for HSDPA services to other services' signals. With the increase of SIOR value, the channel becomes better since the interference from other services decreases. When reaching a specific value, the SIOR would not affect the channel as much as lower value. It means that the interference from other services is very small. We take this parameter as the index to describe the channel condition. Notice that the SIOR value is the ratio of the transmitted power over the interference of other services, there should be a suitable range of SIOR value. We take 13dB as the upper bound for the reason that 13dB means the channel condition is good enough even for real

environment. On the other hand, we take 3dB as the lower bound since the channel condition is poor such that it cannot satisfy the QoS requirement even using the action with lowest code rate. In the thesis, we will compare the proposed scheme with other conventional schemes.

Figure 4.1 shows the BLER of different schemes versus SIOR values. From Figure 4.1, we find that the BLER of FQL is always near the requirement value 0.1. It shows that FQL can indeed satisfy the quality of service (QoS) requirement if defined well. On the other hand, the QL scheme also satisfies the QoS requirement for most of conditions except for the condition that SIOR equals to 5dB. It is for the reason that the QL scheme quantizes the system states, such that it cannot converge into the condition just with QoS requirement for quantization errors. Comparing the FQL scheme with the QL scheme, we find that the FQL scheme has better performance to exactly guarantee the QoS requirement than the QL scheme. The reason is that the system states described by the FQL scheme are more accurate and feasible than the QL scheme, such that the reinforcement function defined in the FQL scheme can be also more accurate and feasible than the cost function defined in the QL scheme.

The BLER values of LA_IR schemes decrease with the increase of the SIOR value. It is reasonable for the reason that while SIOR increases, the channel condition is better such that the probability of decoding errors decreases. Notice that the BLER values of the two LA_IR schemes using 16QAM are always below the requirement value. The reason is that under most conditions, the channel is bad for using 16QAM modulation order. Under the condition that SIOR less or equal to 5dB, the channel condition is even too bad for 16QAM modulation order such that the BLER values of them are vary high and even equal to 1. The LA_IR_2 scheme adopts conservative actions compared to LA_IR_1 scheme. That is why the BLER value of LA_IR_2 scheme is always less then LA_IR_1 scheme.
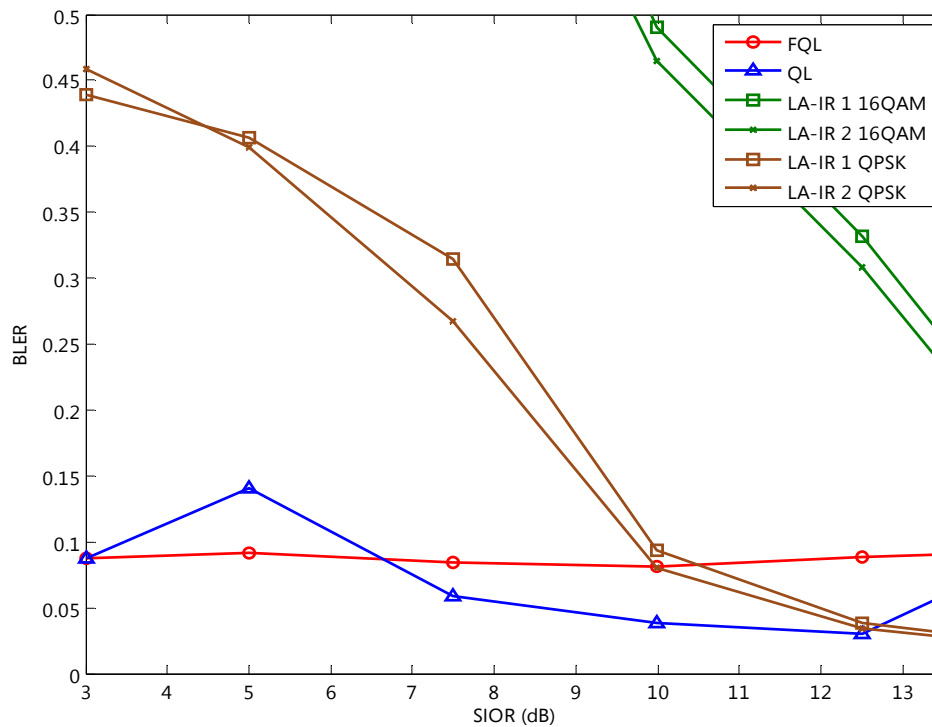
Figure 4.1: The BLER versus SIOR value

The BLER values of the two LA_IR schemes using QPSK are under the requirement value after SIOR equal to 10dB. It is because the channel condition after SIOR equal to 10dB is good enough for using QPSK modulation order, such that these two schemes can satisfy the QoS requirement. But it pays the cost that the throughputs of them are saturated after SIOR equal to 10dB. This fact is shown as Figure 4.2.

Figure 4.2 shows the throughputs of different schemes versus SIOR value with single code. We can see that the throughput of the FQL scheme is higher than the QL scheme after SIOR equal to 12.5 dB. From comparing with Figure 4.1 and Figure 4.2, we find that the throughput of FQL scheme is higher than that of QL scheme, but the trends of both schemes in Figure 4.1 are adverse. It is the reason that the FQL scheme tries to adopt the action with higher code rate if the BLER is much lower than the QoS requirement value, while the QL scheme cannot adapt such case because of the
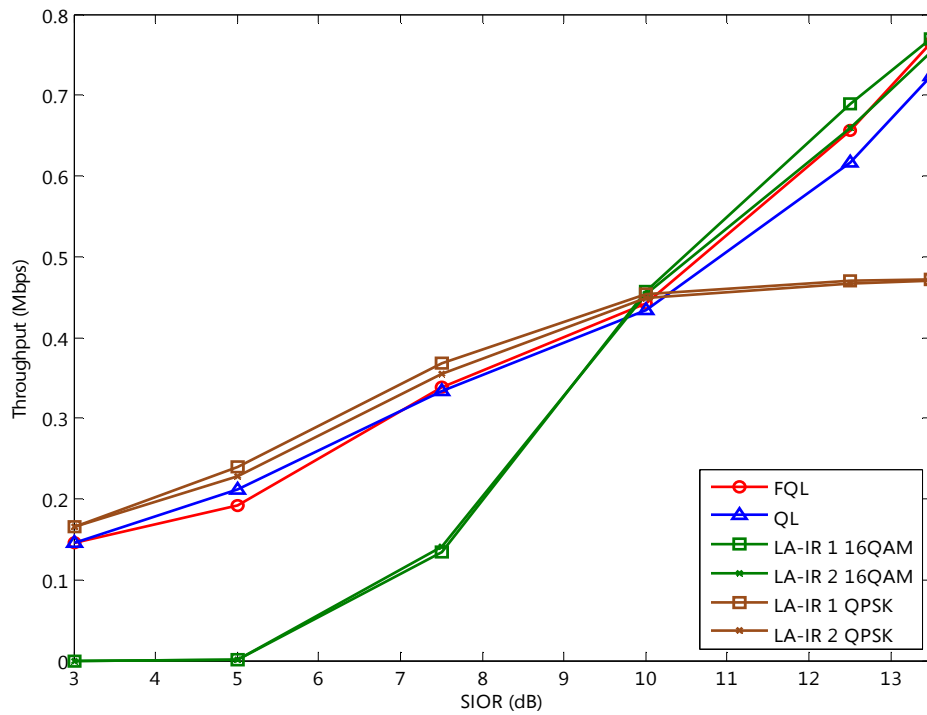
Figure 4.2: The system throughput versus SIOR value

quantization errors for system states. We also find that the trends of their throughputs roughly go along the trends of the throughputs of LA_IR schemes using QPSK before SIOR equal to 10dB, and LA_IR schemes using 16QAM modulation order after 10dB. Although the throughputs of FQL and QL schemes are a little less than these LA_IR schemes, the LA_IR schemes cannot guarantee the QoS requirement for all channel conditions. This phenomenon implies that the FQL scheme can satisfy the QoS and indeed improve the throughput under any channel condition at the same time.

For LA_IR schemes using QPSK modulation, the throughputs of them are saturated after SIOR equal to 10dB. It is for the reason that the channel condition is good enough for using QPSK modulation order. This means that even while the LA_IR schemes always use the action with highest code rate, the throughputs of them cannot increase any more.

On the other hand, for LA_IR schemes using 16QAM, the throughputs of them
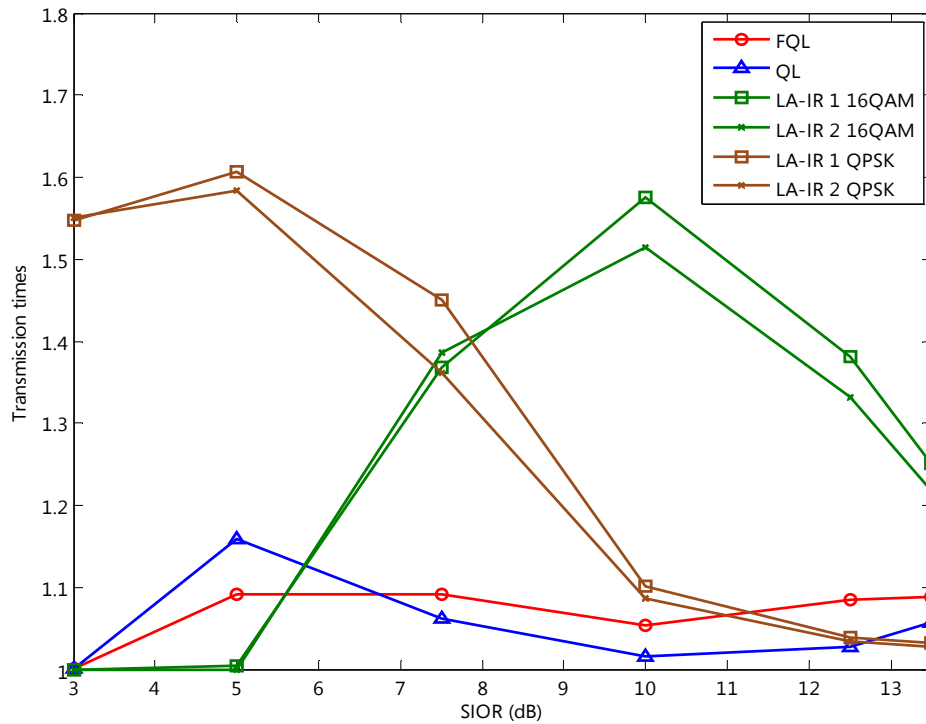
31

Figure 4.3: The average number of transmission times versus SIOR values

increase with the SIOR value. The worst case occurs when SIOR is less or equal to 5dB. For this case, the channel condition is even too bad for using 16QAM modulation order such that their throughputs are very low. Only after the case of 10dB, the throughputs of these two schemes outperform the other schemes using QPSK. But as mentioned before, they obviously cannot guarantee the QoS requirement.

Figure 4.3 shows the number of transmission times versus the SIOR value. From comparing with Figure 4.2 and Figure 4.3, we can see that the numbers of transmission times of the FQL and QL schemes are almost less than the LA_IR schemes whose throughput are roughly the same. The average value of total conditions is about 1.07 for the FQL scheme and 1.05 for the QL scheme. This implies that FQL and QL decide suitable action for initial transmissions than LA_IR schemes. The FQL scheme pays the cost of more times for retransmissions to reach higher throughput than the QL scheme since it still guarantees the QoS requirement.

The special case that the QL scheme needs more retransmissions than FQL occurs at SIOR equal to 5dB. This implies that the QL scheme adopts the action which are too venture for most cases. Hence it also results that the throughput of the QL scheme is higher than FQL under this condition, but it cannot satisfy the QoS at the same time. The average numbers of transmission times of LA_IR schemes using QPSK decrease with the increase of SIOR value. It is obviously for the reason that the channel condition is better. With a better channel condition, there are fewer cases for QPSK to need retransmissions. The average numbers of transmissions of LA_IR schemes using 16QAM increase first and then decrease with SIOR values after 10dB. The reason is that in the period between 3dB and 10dB, the channel condition is too bad for using 16QAM modulation oder such that the two schemes only need few retransmissions to reach the lowest code rate whenever a successful decoding occurs or not. Especially in the case that SIOR is less or equal to 5dB, the channel condition is very bad such that only the lowest code rate can be chosen as initial code rate for the two LA-IR schemes. After the case that SIOR is equal to 10dB, the channel condition is better and then the probability of successful decoding increases. Hence there are more conditions that the system can transmit with higher code rate if using 16QAM modulation order. LA_IR_2 needs less times of transmissions than LA_IR_1 for the reason that LA_IR_2 adopts more conservative actions to transmit data. It increases the code rate of an initial transmission only when two consequent ACKs occur, while LA_IR_1 scheme adopts an action with higher code rate as long as one successful decoding occurs.

Figure 4.4 shows the failure rate of different schemes versus SIOR value. The failure rate is different from the BLER. A failure in the simulation means that a packet is dropped since reaching the lowest code rate in transmissions or retransmissions. So a failure implies one or more than one NACK, but it is not vise versa. From the
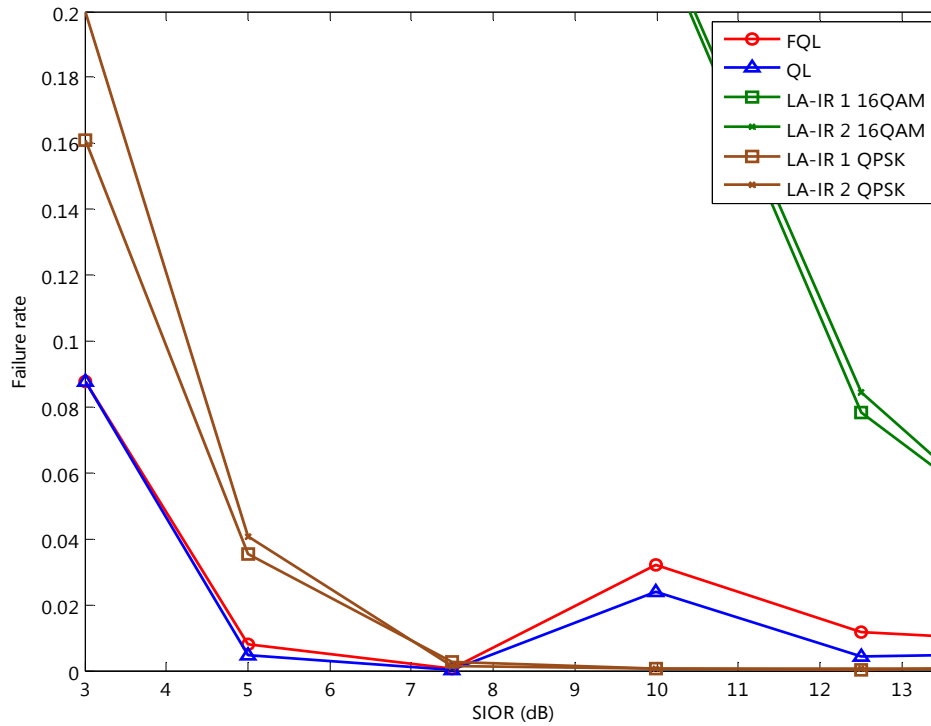
Figure 4.4: The failure rates versus SIOR values

comparison between Figure 4.4 and Figure 4.2, we can find that the failure rates of the

FQL and QL schemes are less than the LA-IR schemes whose throughput is roughly

the same. It is the reason that the two learning schemes adopt suitable actions for

initial transmissions. The failure rates of different schemes decrease with the increase

of SIOR value.

A special case of the FQL and QL schemes occurs at 10dB. It is the reason that

under this condition, the channel is good enough for using QPSK modulation order

but is not so good for using 16QAM modulation order. That means if using 16QAM

modulation order instead of QPSK modulation order under such condition, we can

expect the phenomenon that a failure occurs more easily. In other words, if the FQL

and QL schemes adopt 16QAM actions as initial transmissions, there are more

chances for them to drop packets since reaching the lowest code rate. This also

implies the increase of failures. Although comparing with the FQL and QL schemes,

the failure rate of the FQL scheme is slightly higher than the QL scheme, the throughput of the FQL scheme is higher than the QL scheme. This implies that under some condition the FQL scheme adopts actions using 16QAM modulation order while the QL scheme adopts actions using QPSK modulation order. It is the reason that the code rate of the adopted action using 16QAM with higher data rate is lower than the adopted action using QPSK with lower data rate, for example, 16QAM with code rate 2/3 and QPSK with code rate 1.

On the other hands, the failure rates of LA_IR schemes using 16QAM and QPSK decrease with the SIOR value. The reason is that with the increase of SIOR value, the channel condition is better such that the probability of successful decoding increases for both 16QAM and QPSK modulations. Notice that the failure rates of LA-IR schemes using 16QAM modulation order are even higher than 0.2 while the SIOR value is less than 10dB.

Chapter 5

Conclusion

In this thesis, we propose a fuzzy Q-learning (FQL) based hybrid ARQ scheme for HSDPA to achieve effect resource utilization under any channel condition. Since the hybrid ARQ procedure can be modeled as a discrete-time Markov decision process, we combine the fuzzy system and Q-learning algorithm to help the hybrid ARQ mechanism adjust the code rate and the modulation of the initial transmission in each system state. With suitable decisions of each initial transmission, the system can avoid retransmissions effectively. Fuzzy is used to describe the system state well and help the learning process more efficiently. Combined the fuzzy system and Q-learning, the system can easily reach the object that satisfy the quality of services and keep high throughput at the same time.

In the simulation, we compare the FQL based H-ARQ scheme with conventional schemes, such as two LA-IR schemes and Q-learning based hybrid ARQ scheme. From the simulation results, we find that the FQL based H-ARQ scheme can indeed satisfy the QoS requirement, which is specified in 3GPP Release 5. And the throughput of the proposed FQL based H-ARQ is higher than QL based H-ARQ under most conditions. On the other hand, although the two LA-IR schemes provide a little higher throughput than the proposed scheme, they cannot always satisfy the QoS

requirement. The residual simulation results show that the two learning based schemes can reduce the number of transmission times and the failure rate more effectively than the two LA-IR schemes. It implies that the performances of the FQL and QL schemes are comparable. Therefore, for the purpose that we want to maximize the throughput and guarantee the Qos requirement, the FQL scheme is more feasible than the QL scheme to reach this object.

Bibliography

[1] 3GPP web site, http://www.3gpp.org

[2] J. G. Proakis, "Digital Communications," 3rd Edition, McGraw-Hill.

[3] H. Holma and A. Toskala, Eds., *WCDMA for UMTS*, 3rd ed. JohnWiley & Sons, 2002.

[4] J. Zhang, W. Cao, M. Peng and W. Wang, "Investigation of hybrid ARQ performance for TDD CDMA HSDPA," *VTC 2003-Spring*. The 57th IEEE Semiannual Vol.4, pp. 2721 – 2724, April 2003

[5] A. Das, F. Khan and S. Nanda, "A^{2}IR – An Asynchronous and Adaptive HARQ scheme for 3G Evolution," *VTC 2001 spring*, IEEE VTS 53rd Vol.1, pp. 628 – 632, May 2001

[6] S. Semenov, "Modified maximal ratio combining HARQ scheme for HSDPA," *PIMRC 2004*. Vol.4, pp. 2451 – 2453, Sept. 2004.

[7] M. Peng and W. Wang, "Advanced HARQ and scheduler schemes in TDD-CDMA HSDPA systems," *The 5th International Symposium on Multi-Dimensional Mobile Communications Proceedings*, Vol.1, pp. 67 – 70, 29 Aug.-1 Sept. 2004.

[8] P. Frenger, S. Parkvall and E. Dahlman, **"**Performance comparison of HARQ with Chase combining and incremental redundancy for HSDPA," *VTC 2001 Fall*. IEEE VTS 54th Vol.3, pp. 1829 – 1833, Oct. 2001

[9] M. W. El Bahri, H. Boujernaa and M. Siala, "Performance comparison of type I, II

and III hybrid ARQ schemes over AWGN channels," *IEEE ICIT '04*. Vol.3, pp. 1417 – 1421, Dec. 2004

[10] A. Das, F. Khan, A. Sampath and H.-J. Su, "Performance of hybrid ARQ for high speed downlink packet access in UMTS," *VTC 2001 Fall*. IEEE VTS 54th Vol.4, pp. 2133 – 2137, Oct. 2001

[11] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Machine Learning*, Vol. 8.

[12] C. J. Chang, C. Y. Chang and F. C. Ren, "Q-learning-based Hybrid ARQ for High Speed Downlink Packet Access in UMTS," *VTC 2007 spring*, IEEE 65th , pp. 2610 – 2615 April 2007

[13] P. Y. Glorennec and L. Jouffe, "Fuzzy Q-Learning," *Proc. of the 6th IEEE Int. Conf. on Fuzzy Systems*, pp. 659 – 662, 1997

[14] P. Y. Glorennec, "Fuzzy Q-Learning and Dynamical Fuzzy Q-Learning," *Proc. of the 3th IEEE Int. Conf. on Fuzzy Systems*, Orlando, June 1994

[15] C.-H. Oh, T. Nakashima and H. Ishibuchi, "Initialization of Q-values by fuzzy rules for accelerating Q-learning," *Neural Networks Proceedings*, Vol. 3, pp. 2051 – 2056, May 1998

[16] P. Y. Glorennec, "Reinforcement Learning: an Overview," *European Symposium on Intelligent Techniques*, Sept. 2000.

[17] L. Zhao, J. W. Mark, and T. C. Yoon, "A combined link adaptation and incremental redundancy protocol for enhanced data transmission," in *Proc. IEEE Global Telecommunications Conference*, Vol. 2, Nov. 2001, pp. 1277-1281.

Appendix

Table A.1 : The notations used in FQL algorithm

notation	significance
S	the set of state vectors
S_j	j^{th} state vector, which is constituted by fuzzy linguistic variables
A	the set of candidate actions
A_k	k^{th} possibly action
x_i	input state vector received in i^{th} time period
$\mu_j(x_i)$	output truth value of jth fuzzy rule while received state vector x_i
a_j	the action of jth fuzzy rule selected by exploration/exploitation policy (EEP)
$a(x_i)$	the global inferred action after receiving state vector x_i
a_j^*	the action which corresponds to the maximal q-value

Table A.2 : The notations used in the definition of reinforcement function

| notation	significance
$SINR_{x_i}$	The measured SINR value in state x_i
$\widetilde{SINR}_{A_k}$	The desired SINR value using action A_k to reach requirement BLER value
$a(x_i)$	The applied action of state x_i
$\widetilde{SINR}_{a(x_i)}^{3\times\widetilde{BLER}}$	The desired SINR value using action $a(x_i)$ to reach three times the BLER requirement value
$\widetilde{BLER}$	The QoS requirement of BLER value 0.1
$BLER\left(SINR_{x_i}, a(x_i)\right)$	The theoretical BLER value using action $a(x_i)$ under the measured SINR value
$BLER\left(\widetilde{SINR}_{A_k}, a(x_i)\right)$	The theoretical BLER value using action $a(x_i)$ under the desired SINR value of action A_k
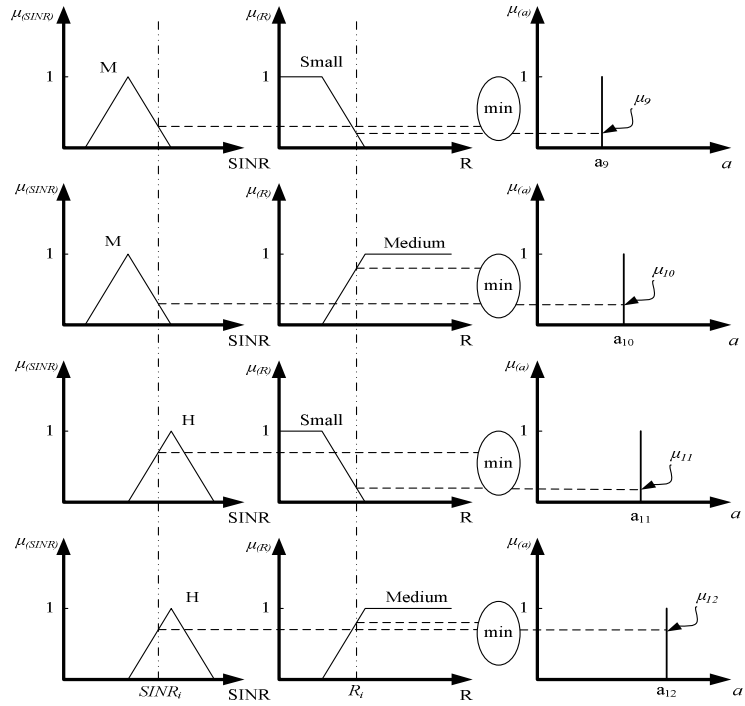
Figure A.1 : A sample of defuzzification in FIS

Vita

Shih-Hung Tsai was born on 1981 in Tainan, Taiwan. He received the B.E. degree in the department of electrical engineering from National Tsing-Hua University, Hsinchu, Taiwan, in 2003, and the M.E. degree in industrial technology R&D master program on communication engineering from Nation Chiao-Tung University, Hsinchu, Taiwan, in 2007. His research interests include radio resource management and wireless communication systems.