Situation-Aware Data Access Manager Using Fuzzy Q-learning Technique for Multi-cell WCDMA Systems

Yih-Shen Chen, *Member, IEEE,* Chung-Ju Chang, *Fellow, IEEE,* and Fang-Ching Ren, *Member, IEEE*

***Abstract*— This paper proposes a novel situation-aware data access manager using fuzzy Q-learning technique (FQ-SDAM) for multi-cell WCDMA systems. The FQ-SDAM contains a fuzzy Q-learning-based residual capacity estimator (FQ-RCE) and a data rate scheduler (DRS). The FQ-RCE can accurately estimate the situation-dependent residual system capacity, and appropriately chooses the received interference powers from the home-cell and adjacent-cell as input linguistic variables, which simplifies the multi-cell environment into a single-cell environment by applying a perceptual coordination mechanism. The DRS can effectively allocate the resource for non-real-time terminals by modifying the exponential rule [10], which considers the effect of interference on adjacent cells. Simulation results show that, compared to the link and interference-based demand assignment (LIDA) scheme proposed in [7], FQ-SDAM can effectively reduce the packet error probability and improve aggregate throughput of the non-real-time services in both the homogeneous and non-homogeneous multi-cell WCDMA environments. Additionally, the modified exponential rule achieves better system performance than the original exponential rule.**

***Index Terms*— Data access control, fuzzy Q-learning, modified exponential rule, and multi-cell WCDMA system.**

I. INTRODUCTION

WIDEBAND CDMA (WCDMA) cellular system supports integrated services with mixed QoS (quality of services) requirements: real-time services require continuous transmission and is intolerant to time delay, while non-real-time services require bursty transmission and tolerate moderate time delay. An adequate radio resource management (RRM) is required to maximize the system capacity and fulfill the complementary QoS requirements. Among many traffic engineering techniques for the RRM, a *call admission control* method is applied to prevent system overloading, based on the long-term availability of radio resources. On the other hand, a *data access control scheme* provides bursty transmission permission for non-real-time services, based on the short-term availability of radio resources.

The main purpose of the *data access control scheme* in WCDMA systems supporting integrated services is to maximize the throughput of non-real-time services while maintaining the transmission quality of real-time services [1]-[5]. To achieve this goal, dynamic access probability schemes [2]-[4] and a base station-controlled scheduling scheme [5] have proposed. In these schemes, the residual system capacity for non-real-time services is first estimated and then shared to non-real-time terminals. A single-cell environment was considered in [2]-[4], while a multi-cell environment was studied in [5]. The multi-cell scheme [5] treats the interference generated from other-cell terminals as if from several home-cell terminals, and consequently the multi-cell environment is regarded as a single-cell environment. However, the mutual-affected behavior of radio resource allocation in the multi-cell environment is still not considered. Notably, in the multi-cell WCDMA system, the increment of data transmission power in one cell would cause the interference level to rise in the adjacent cells. If each cell allocates the entire residual capacity for bursty transmission without considering the interference influence from adjacent cells, then the system becomes overloaded.

The over-loading phenomenon could be alleviated by an appropriate coordination method among cells [6]. Knowing the radio resources of all cells, a centralized data access method for the multi-cell WCDMA system can maximize the system throughput by applying a global optimization method. Unfortunately, the coordination procedure takes a long time to transact the resource information between cells, making practical implementation infeasible. Usually, the data access control scheme operates in the short-term time scale, *e.g.* frame time, making distributed schemes preferable. Kumar and Nanda [7] proposed a distributed scheme called load and interference-based demand assignment (LIDA). The LIDA is a resource reservation-based scheme which reserves some resources in each cell against the interference variation. Additionally, LIDA uses the concept of burst admission threshold for high-rate transmission in a cell to avoid excess interference power to adjacent cells, allowing bursty transmission only when the strength difference between the received pilot signals from the home cell and adjacent cells is larger than the threshold. The effectiveness of this scheme relies on the selection of the reservation threshold, which should be dynamically chosen according to the system loading and the received interference power level.

Additionally, a rate scheduling scheme is also embedded in the data access control scheme to allocate the residual

Manuscript received June 16, 2004; revised May 6, 2005; accepted January 5, 2006. The associate editor coordinating the review of this paper and approving it for publication was J. Evans. This work was supported by National Science Council, Taiwan, under Grant NSC 93-2219-E-009-011.

Y. S. Chen and C. J. Chang are with the National Chiao Tung University, Taiwan (email: yihshenchen@gmail.com, cjchang@cc.nctu.edu.tw).

Fang-Ching Ren is with the Industrial Technology Research Institute, Taiwan (email: Frank_Ren@itri.org.tw).

Digital Object Identifier 10.1109/TWC.2006.04394.

capacities for non-real-time terminals according to a service principle. Ramakrishna and Holtzman adopted a *maximization throughput* criterion for the scheduling scheme [8]. This criterion can maximize the system throughput, but may cause the low-class users to suffer from starvation. Alternatively, Jalali, Padovani, and Pankai proposed a *proportional fairness* criterion [9] for a down link scheduling scheme in a CDMA-HDR (high data rate) system. Their proposed scheme defines a utility function as a ratio of the supported and the average data rates. The supported data rate is determined by the channel condition, while the average data rate is calculated as the window average of the transmitted throughput. The terminal with the highest utility value transmits data in the next frame time. This algorithm may lead to large transmission delay for some terminals. Additionally, Shakkottai and Stolyar proposed an *exponential rule* criterion [10] for the another definition of the utility function to strike a balance between the system throughput and the transmission delay. However, applying the exponential rule to the uplink transmission should consider the terminal's location factor minimize interference with adjacent cells.

This paper proposes a situation-aware data access manager using fuzzy Q-learning technique (FQ-SDAM) for multi-cell WCDMA systems. The proposed FQ-SDAM scheme consists of two parts: *fuzzy Q-learning-based residual capacity estimator* (FQ-RCE) and *data rate scheduler* (DRS). The FQ-RCE, by fuzzy Q-learning, estimates the appropriate situation-dependent residual system capacity, in terms of interference power, for non-real-time services, while the DRS assigns transmission rates for non-real-time terminals by a modified exponential rule.

The fuzzy inference system (FIS) and the reinforcement learning technique have been separately applied to solve network resource management problems [11]-[14]. A fuzzy resource allocation controller was proposed in [12], where the FIS method was adopted to estimate the resource availability. A reinforcement learning technique, Q-learning, was applied respectively to handle dynamic channel assignment in [13] and multi-rate transmission control problems in [14] for wireless communication systems. By learning from the system environment, the Q-learning technique can converge to a pre-defined optimal control target. In [15], Jouffle proposed a reinforcement learning technique for FIS, called fuzzy Q-learning (FQL). The FQL technique combines the advantages of FIS and reinforcement learning. The FIS provides a good function approximation for the FQL, which enables *a priori* knowledge to be applied to the system design. Additionally, the reinforcement learning provides a model-free approach to obtain a control target. By applying the FQL technique, the radio resource can be managed under partial, uncertain information, and the optimal resource management can be reached incrementally.

FQ-RCE uses interference measures from three sources as input linguistic variables to estimate the situation-dependent residual capacity in the multi-cell environment: the received interference power from real-time terminals at the home cell, the received interference power from non-real-time terminals at the home cell and the received interference power from the adjacent cells. Notably, the received interference power from adjacent cells is regarded as a different variable from the received interference power from home cell to distinguish the interference variations. Therefore, by the linguistic variable of the adjacent-cell interference power, the FQ-RCE at the home cell can *perceive* the radio resource allocation by those FQ-SDAMs in adjacent cells, or say, be aware of the loading of adjacent cells, and precisely estimate the residual resource in a distributed fashion. Thus, the multi-cell WCDMA environment does not require an explicit action coordination.

On the other hand, the DRS modifies the exponential rule in [10] to assign the transmission rates for non-real-time terminals, based on the residual capacity estimated by FQ-RCE. The modified exponential rule is a utility-function-based scheduling algorithm which considers the transmission delay, average transmission rate, and link capacity. The modified rule differs from the original exponential rule [10] in the definition of link capacity. For the modified exponential rule, the link capacity is defined as the maximum available rate where the interference influence on adjacent cells by the transmission power is below a guard threshold, considering location awareness. The modified exponential rule is most suitable for applications in the uplink transmission of multi-cell WCDMA systems, which is explained later. Simulation results show that the proposed FQ-SDAM outperforms the LIDA scheme since it can effectively reduce the packet error probability and improve the aggregate throughput in both homogeneous and non-homogeneous multi-cell WCDMA environments. Additionally, the modified exponential rule can achieve better system performance than the original exponential rule. In the homogeneous case, FQ-SDAM achieves higher aggregate throughput by 75.3% (53.3%) than LIDA with β=10%, under high-bursty (low-bursty) real-time traffic. In the nonhomogeneous case, FQ-SDAM achieves greater aggregate throughput by 31.53%, 35.5%, and 34.2% for the cells in the central, first-tier, and second-tier, respectively, than LIDA with β=10%.

The rest of the paper is organized as follows. The system model is described in Section II. Section III briefly describes the concept of fuzzy Q-learning and proposes the design of FQ-SDAM. Simulation results are presented in Section IV, which compares the performance of the FQ-SDAM and a conventional LIDA scheme. Finally, concluding remarks are given in Section V.

II. System Model

This paper considers a multi-cell WCDMA system containing N cells, where each cell has a base station using FQ-SDAM to allocate the radio resource for real-time and non-real-time terminals within its coverage area. An uplink supporting slotted transmission is adopted. All terminals transmit at the same frequency band and are distinguished by their own spreading codes. Each terminal holds two communication channels, the dedicated physical data channel (DPDCH) and the dedicated physical control channel (DPCCH). The DPDCH carries data generated by layer 2 protocol, while the DPCCH carries control information. A channel has a frame-based structure, where the frame length $T_f = 10$ ms is divided into 15 slots with length $T_{\text{slot}} = 2560$ chips, each slot corresponding to one power control period. Hence, the power control frequency is 1500 Hz. The spreading factor (SF) for DPDCH can vary between $4 \sim 256$ by SF $= 256/2^k, k = 0, 1, \cdots, 6$, carrying 10×2^k bits per slot, and the SF for DPCCH is fixed at 256, carrying 10 bits per slot.

Two types of traffic are considered: real-time (type-1) traffic and non-real-time (type-2) traffic. The system provides continuous transmission for real-time traffic and bursty transmission for non-real-time traffic. Here, the real-time terminal is the terminal supporting real-time services, and the non-real-time terminal is the terminal supporting non-real-time services. The real-time terminals may transmit at any possible data rate while necessary; on the other hand, the transmission of non-real-time terminals is controlled by the data access manager at the base station. Considering the terminal's link gain and the received interference power from both the home and adjacent cells, the data access manager assigns an appropriate data rate for each non-real-time terminal. For the bursty transmission, the available data transmission rates are 1X, 2X, 4X and 8X, and 1X transmission rate is called the basic rate. A strength-based power control scheme is assumed such that the required transmission power of a mobile is directly proportional to the transmission rate [18]. Additionally, the overall capacity is set by the upper bound of the total received interference power, and the residual capacity is defined as the allowable received interference power from the non-real-time terminals.

The link gain between terminal i to base station j, denoted by h_{ij}, is usually determined by the long-term fading FL_{ij} and the short-term fading FS_{ij} [19], which is given by

$$h_{ij} = FL_{ij} \times FS_{ij}. \tag{1}$$

The long-term fading FL_{ij}, combining the path loss and shadowing, is modelled as

$$FL_{ij} = k \times r^{-\alpha} \times 10^{\eta/10}, \tag{2}$$

where k is constant, r is distance from mobile i to base station j, α is path loss exponent usually lying between 2 and 5 for a mobile environment (α = 4 in this paper), and η is normal-distributed random variable with zero mean and variance σ_L^2. The parameter σ_L is affected by the configuration of the terrain and ranges from 5 to 12 (σ_L^2=10 in this paper) [19]. The short-term fading FS_{ij} is mainly caused by multi-path reflections, and is modelled by Rayleigh distribution.

The real-time service is modelled as an ON-OFF Markov process with a transition rate μ from ON to OFF and λ from OFF to ON. The non-real-time service is modelled as a batch Poisson process, in which the arrival process of the data burst is in Poisson distribution and the data length is assumed to have a geometric distribution. The measure of the packet error probability, denoted by P_e, is regarded as the system performance index. The maximum tolerable packet error probability, denoted by P_e^*, is defined as the system QoS requirement. Additionally, the measure of packet transmission delay is used as a parameter for the data rate scheduler.

III. Design of FQ-SDAM

The FQ-SDAM contains two functional blocks of a fuzzy Q-learning-based residual capacity estimator (FQ-RCE) and a data rate scheduler (DRS). The FQ-RCE estimates the residual interference power budget, and then the DRS allocates the resource for the non-real-time terminals. The following section describes the fuzzy Q-learning and the detailed design of the two function blocks.

A. The Fuzzy Q-Learning

(FQL) Denote $\mathbf{S}$ the set of state vectors for the system, $\mathbf{S}$=$\{S_i,\ i = 1, 2, \cdots, M\}$; each state vector S_i comprises L fuzzy linguistic variables selected to describe the system. Denote $\mathbf{A}$ the set of actions possibly chosen by system states, $\mathbf{A}$=$\{A_j,\ j = 1, 2, \cdots, N\}$. For an input state vector $\mathbf{x}$ containing the L linguistic variables, the rule representation of FQL for state S_i is in the form by

if x is S_i, **then** A_j with $q[S_i, A_j]$,
$1 \le i \le M$ and $1 \le j \le N$,

where A_j is the jth action candidate that is possibly chosen by state S_i, and $q[S_i, A_j]$ is the Q-value for the state-action pair (S_i, A_j). The number of state-action pairs for each state S_i equals the number of the elements in the action set; *i.e.*, each antecedent has N possible consequences. Every fuzzy rule needs to choose an action A_i from the action candidates set $\mathbf{A}$ by an action selection policy. In the FQL, the action selection policy for each fuzzy rule may be *select-max* or another exploration strategy. To defuzzify the M fuzzy rules, the inferred action $a(\mathbf{x})$ for the input vector $\mathbf{x}$ is expressed as

$$a(\mathbf{x}) = \frac{\sum_{i=1}^{M} \alpha_i \times A_i}{\sum_{i=1}^{M} \alpha_i}, \tag{3}$$

where α_i is the truth value of the rule representation of FQL for state S_i. Additionally, the Q-value for the state-action pair $(\mathbf{x}, a(\mathbf{x}))$ is given by

$$Q(\mathbf{x}, a(\mathbf{x})) = \frac{\sum_{i=1}^{M} \alpha_i \times q[S_i, A_i]}{\sum_{i=1}^{M} \alpha_i}. \tag{4}$$

For the current system state $\mathbf{x}$ after applying the chosen action $a(\mathbf{x})$, the next-stage system state is assumed at $\mathbf{y}$, and the system reinforcement signal is given by $c(\mathbf{x}, a(\mathbf{x}))$. To update the Q-value, the next-stage optimal Q-value, $Q^*(\mathbf{y}, a(\mathbf{y}))$, is defined as

$$Q^*(\mathbf{y}, a(\mathbf{y})) = \frac{\sum_{i=1}^{M} \alpha_i \times q[S_i, a_i^*]}{\sum_{i=1}^{M} \alpha_i}, \tag{5}$$

where $q[S_i, a_i^*]$ is the Q-value of state-action pair (S_i, a_i^*) and

$a_i^* = \underset{A_j}{\operatorname{argmax}}$ {q[S_i, A_j]}. According to the Q-learning rule [17], the Q-value update in the FQL can be expressed as

$$q[S_i, a_i] = q[S_i, a_i] + \eta \Delta q[S_i, a_i], \tag{6}$$

where η is the learning rate, $0 \le \eta \le 1$, and

$$\Delta q[S_i, a_i] = \{c(\mathbf{x}, a(\mathbf{x})) + \gamma Q^*(\mathbf{y}, a(\mathbf{y})) - Q(\mathbf{x}, a(\mathbf{x}))\} \times \frac{\alpha_i}{\sum_{k=1}^{M} \alpha_k}. \tag{7}$$

$c(\mathbf{x}, a(\mathbf{x}))$ in (7) is the reinforcement signal.

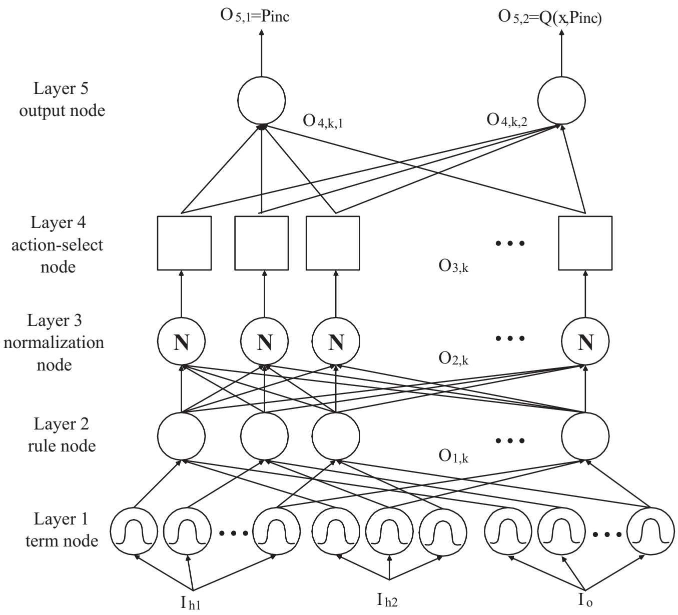

Fig. 1. Structure of FQ-RCE.

B. Fuzzy Q-learning-based Residual Capacity Estimator

(FQ-RCE) The FQ-RCE selects three interference measures as input linguistic variables: the received interference power from real-time terminals at the home cell (I_{h1}), the received interference power from non-real-time terminals at the home cell (I_{h2}), and the received interference power from adjacent cells (I_o). Notably, the received interference power in the WCDMA system is a good indicator of system loading because the system capacity is interference-limited; moreover, the interference generated from the home cell can be identified by PN codes and the interference from adjacent cells can be distinguished by long scrambling codes [21]. Accordingly, the system state vector **x** containing the three linguistic variables input to FQ-RCE is defined as

$$\mathbf{x} = (I_{h1}, I_{h2}, I_o). \tag{8}$$

Comprehensive experiments found that five terms for both I_{h1} and I_o, and three terms for I_{h2} were proper. Hence, their fuzzy term sets are T(I_{h1})={Largely High, HiGh, MeDium, LoW, Largely Low}={LH, HG, MD, LW, LL}, T(I_{h2})={HiGh, MeDium, LoW}={HG, MD, LW}, and T(I_o)={Largely High, HiGh, MeDium, LoW, Largely Low}={LH, HG, MD, LW, LL}. From the fuzzy set theory, the fuzzy rule base forms have dimensions |T(I_{h1})|×|T(I_{h2})|×T|(I_o)|. Accordingly, M=75. On the other hand, the step-wise incremental/decremental action of the interference power budget for the non-real-time services, denoted by P_{inc}, is selected as the output linguistic variable. Here, seven levels of increment actions (N=7) are given, and the corresponding fuzzy term set is T(P_{inc})={PI_1, PI_2, PI_3, PI_4, PI_5, PI_6, PI_7}. After the interference increment is estimated by the FQ-RCE, the residual system capacity (RC) being allocated for the non-real-time services is defined as

$$RC = I_{h2} + P_{inc}, \tag{9}$$

where I_{h2} is the capacity previously assigned to the non-real-time services. Additionally, the reinforcement learning signal $c(\mathbf{x}, a(\mathbf{x}))$ is defined as

$$c(\mathbf{x}, a(\mathbf{x})) = [\frac{P_e(\mathbf{x}, P_{inc}) - P_e^*}{P_e^*}]^2, \tag{10}$$

where $P_e(\mathbf{x}, P_{inc})$ is the packet error probability of real-time services for the state-action pair $(\mathbf{x}, P_{inc})$, which is a performance measure of the system, and P_e^* is the QoS requirement of real-time packet error probability.

Figure 1 shows the structure of FQ-RCE as a five-layer adaptive-network-based implementation of a fuzzy inference system. In the FQ-RCE, *layer 1* to *layer 3* are the antecedent components of the FIS, while *layer 4* and *layer 5* represent the consequent components. The node function in each layer is described as follows.

Layer 1: Every node k, $1 \leq k \leq 13$, in this layer is a term node which represents a fuzzy term of an input linguistic variable, where k= 1, $\cdots$, 5 (6, 7, 8) (9, $\cdots$, 13) denotes that node k is the kth (($k-5$)th) (($k-8$)th) term in $T(I_{h1})$ ($T(I_{h2})$) ($T(I_o)$). The node function is defined as the membership function with a bell shape for the term. Thus, for an input linguistic variable x, the output $O_{1,k}$ is given by

$$O_{1,k} = b(x; m^k, \sigma^k) = e^{-\frac{(x-m^k)^2}{\sigma^{k2}}}, \tag{11}$$

where $b(\cdot)$ is the bell-shaped function, and m^k and σ^k is the mean and the variance of the node k, respectively.

Layer 2: Every node k, $1 \leq k \leq 75$, in this layer is a rule node which represents the truth value of kth fuzzy rule; it is a *fuzzy*-AND operator. Here, the product operation is employed as the node function. Since each fuzzy rule has three input linguistic variables, the node output $O_{2,k}$ is the product sum of three fuzzy membership values corresponding to the inputs. Therefore, $O_{2,k}$ is given by

$$O_{2,k} = \prod\{O_{1,l}\}, \forall l \in P_k, \tag{12}$$

where P_k={l| all ls that are the pre-condition nodes of the k-th fuzzy rule}.

Layer 3: Every node k, $1 \leq k \leq 75$, in this layer is a normalization node which performs a normalization operation so that all the truth values sum to unity. After the normalization, the output of this node $O_{3,k}$ is given by

$$O_{3,k} = \frac{O_{2,k}}{\sum_{l=1}^{75} O_{2,l}}. \tag{13}$$

Layer 4: Every node k, $1 \leq k \leq 75$, in this layer is an action-select node which represents the consequence part of kth fuzzy rule. Based on the action selection policy and Q-values of the possible action candidates (PI_j, $j = 1, 2, \cdots 7$), the node needs to choose an appropriate action. Since improper initial fuzzy parameters settings would lead to a bad learning result, the Boltzmann-distributed exploration strategy in [20] is employed to explore the set of all the possible action candidates. In the Boltzmann-distributed exploration, the node chooses the state-action pair (S_k, a_k), $a_k \in T(P_{inc})$, for the kth rule, with the probability $\xi(S_k, a_k)$ given by

$$\xi(S_k, a_k) = \frac{e^{q[S_k, a_k]/T}}{\sum_{j=1}^{7} e^{q[S_k, PI_j]/T}}, \tag{14}$$

where T is the *temperature* which reflects the randomness of action selection. After the action is chosen, the node sends two outputs $O_{4,k,1}$ and $O_{4,k,2}$ to the action node and Q-value node in layer 5, respectively. Outputs $O_{4,k,1}$ and $O_{4,k,2}$ are represented by

$$O_{4,k,1} = O_{3,k} \times a_k, \tag{15}$$

and

$$O_{4,k,2} = O_{3,k} \times q[S_k, a_k]. \tag{16}$$

Layer 5: This layer has two output nodes, action node $O_{5,1}$ and Q-value node $O_{5,2}$, which represent the fuzzy defuzzification of FQ-RCE. Herein, the center of area method is applied for defuzzification. Since layer 3 normalizes the truth value of the antecedent part of the ith fuzzy rule, the node functions in layer 5 are summation of the inputs from layer 4. Hence, $O_{5,1}$ and $O_{5,2}$ are given by

$$O_{5,1} = P_{inc} = \sum_{k=1}^{M=75} O_{4,k,1}, \tag{17}$$

and

$$O_{5,2} = Q(\mathbf{x}, P_{inc}) = \sum_{k=1}^{M=75} O_{4,k,2}. \tag{18}$$

After the action is performed, the FQ-RCE calculates the reinforcement signal $c(\mathbf{x}, a(\mathbf{x}))$ by (10) and updates the Q-value of each state-action pair according to (6).

Notably, the convergence property of Q-learning is held for the single-agent (learner) case and may not be held for multiple-agent cases. Additionally, the convergence of Q-learning in multi-cell WCDMA systems would be a difficult task because decision policies of all cells concurrently change during the learning phase. To handle this difficulty, the perceptual coordination mechanism [16] is applied to FQ-RCE by designing the input linguistic variables, which incorporate two parts: I_{h1} and I_{h2} represent the current state of the radio resource usage in home cell and I_o represents the radio resource allocations performed in adjacent cells. Therefore, by measuring the adjacent-cell interference, the FQ-RCE at home cell can implicitly *perceive* the situation of radio resource allocation (action) in adjacent cells. The multi-cell learning environment can then be simplified as a single-cell environment, and the convergence property for the FQ-RCE can be held as a result.

C. The Data Rate Scheduler

(DRS) The DRS modifies the exponential rule scheduling algorithm in [10]. The formula of the modified exponential rule is given by

$$j = \operatorname*{argmax}_i \{\frac{r_i}{\bar{r}_i} \times e^{\frac{W_i - \overline{\mathbf{W}}}{1+\sqrt{\overline{\mathbf{W}}}}}\}, \tag{19}$$

where r_i, $\bar{r}_i$, and W_i are the link capacity, the average transmission rate, and the waiting time, of the ith data terminal, respectively, and $\overline{\mathbf{W}}$ is the average waiting time of all the data terminals. The main difference between the modified and the original exponential rules is in the definition of the link capacity. The original exponential rule was proposed for downlink transmission in the CDMA HDR system [9], where the link capacity was defined as the maximum transmission rate under the current link condition. However, in the multi-cell WCDMA environment, the uplink transmission power would interfere with adjacent cells. The closer the terminal's location near the cell boundary, the larger the interference power. Therefore, the modified exponential rule algorithm sets a guard threshold of adjacent-cell interference for the uplink transmission power such that its incurred adjacent-cell interference is lower than the pre-defined level. Then, the location-dependent link capacity r_i is defined as the maximum transmission rate available for a radio link, which must satisfy the following condition:

$$P(r_i) \times h_i^a \leq P_d, \tag{20}$$

where $P(r_i)$ is the transmission power of terminal i with rate r_i, h_i^a is the maximum link gain between the terminal i and adjacent cells, and P_d is the guard threshold of the adjacent-cell interference. In the strength-based power control scheme, the transmission power $P(r_i)$ is given by

$$P(r_i) = \frac{r_i \times (E_b/N_0)^* \times I_{max}}{PG \times h_i}, \tag{21}$$

where $(E_b/N_0)^*$ is the signal-to-noise requirement, I_{max} is the maximum received interference power, PG is the processing gain, and h_i is the link gain between the terminal and its home cell. Additionally, h_i and h_i^a can be measured by monitoring the received pilot strength from the home and adjacent cells. Hence, the modified exponential rule states that *the terminal with higher maximum available transmission rate, lower average transmitted rate and longer delay obtains higher transmission priority*. As the terminal moves toward the cell boundary, the emission power to the adjacent cells increases, the transmission priority falls, and the waiting time accumulates. However, if the terminal's waiting time is long, the transmission priority is high. Therefore, the modified exponential rule can strike a balance among the link gain, the location and the waiting time of terminals.

The DRS performs the rate allocation according to the terminal's priority. The terminal with the highest priority is given the rate allocation first, and the other terminals are given the allocation in priority order. The operation of the DRS stops when all the data power budget is used out. Its procedure is described below:

[**The DRS Algorithm**]

Step 1 Obtain the residual system capacity (RC) for non-real-time services from FQ-RCE.

Step 2 Choose the highest-priority terminal, j, out of data terminals that are not allocated yet, by (19).

Step 3 Compute the remaining RC by

$$RC = RC - P(r_j)/PG.$$

If the remaining RC is larger than 0, go back to **Step 2**. Otherwise, go to **Step 4**.

Step 4 Inform terminals of the assigned data rate via the signaling channel. **End**

IV. Simulation Results and Discussion

In the simulations, a concatenated 19-cell (N=19) environment was configured as the multi-cell WCDMA system. The central cell was labelled as cell 1, the cells in the first tier were cell 2 $\sim$ cell 7, and the cells in the second tier were cell 8 $\sim$ cell 19. Three kinds of real-time traffic were considered: voice traffic, high-bursty real-time data traffic and low-bursty real-time data traffic. The voice traffic assumed 2-level transmission rate traffic which is modelled by a 2-level MMDP (Markov modulated deterministic process) [22]. The real-time data traffic was modelled by an ON/OFF traffic stream with specific burstiness $1/\rho_h$ ($1/\rho_l$) and peak rate $R_{p,h}$ ($R_{p,l}$) for high-bursty (low-bursty) real-time traffic. The two real-time data traffic flow had the same mean rate but different burstiness level. The non-real-time data traffic was considered to have a Poisson arrival process with data burst length in geometric distribution. Table I shows all the detailed traffic parameters. A basic rate in the WCDMA system is assumed to be a physical channel with SF=256. For each connection, DPCCH is always active to maintain the connection reliability. To reduce the overhead cost of interference produced by DPCCHs, the transmitting power of a DPCCH was assumed to be lower than its respective DPDCH by an amount of 3 dB. The QoS requirement of the packet error parameter, P_e^*, is set to be 0.01.

The conventional resource reservation scheme proposed in [7], LIDA (load and interference demand assignment), was used as a benchmark for performance comparison. The basic concept of the LIDA scheme is two-folded: firstly, a portion of interference power budget, β, is reserved to avoid overloading, and second, a burst-mode admission is applied for the high-rate traffic. Additionally, the allocation of the incremental of transmission power, P_{inc}, to the non-real-time data traffic in the LIDA scheme is given by

$$P_{inc} = (1-\beta)I_{max} - I_{h1} - I_{h2} - I_o. \qquad (22)$$

The performance of the LIDA scheme relies heavily on the choice of reservation threshold, β. The simulations considered three reservation threshold, β = 0%, 5%, and 10%, and the modified exponential rule with P_d=2dB was applied for the LIDA scheme. Moreover, a scheme which combines the FQ-RCE with the original exponential rule, called FQ-RCE/EXP, was considered to further evaluate the effectiveness of the modified exponential rule. Notably, all the considered schemes were applied only to non-real-time terminals, and all the real-time terminals initiated data transmission whenever they had packets in queues.

A. Homogeneous Case

In the homogeneous case, all cells are assumed to contain 22 voice terminals, 40 real-time data terminals and 20 non-real-time data terminals. The 40 real-time data terminals consist of $N_{D,h}$ high-bursty and $N_{D,l}$ low-bursty data users, where $N_{D,h}$+$N_{D,l}$=40.

Figure 2 shows the packet error probabilities versus the number of high-bursty real-time data users. The packet error probability of the LIDA scheme was found to violate the QoS requirement, and the LIDA scheme without reservation (β=0%) had the largest packet error probability. The results demonstrate the necessity to precise residual capacity estimation to avoid overloading in the multi-cell WCDMA environment. The packet error probabilities of the FQ-SDAM and FQ-RCE/EXP schemes always fulfill the QoS requirement because the FQ-RCE adopts the FQL, which inherently possesses the capability of reinforcement learning. Thus, the FQ-RCE can precisely determine the residual system capacity by monitoring the loading status of the home cell and the interference variation of adjacent cells. Additionally, regardless of the value of $N_{D,h}$, FQ-SDAM scheme always achieves lower packet error probabilities than the FQ-RCE/EXP because the uplink transmission powers emitted from terminals interfere with users at the home cell and adjacent cells in the multi-cell environment. With the awareness of location of users, the modified exponential rule in FQ-SDAM effectively curbs the interference influence on adjacent cells within a sustainable level and consequently reduces the packet error probabilities.

Figure 3 shows the aggregate throughput of non-real-time data traffic versus three numbers of high-bursty real-time users: $N_{D,h}$=10, 20 and 30. The three cases of different real-time data users were used to simulate the low-bursty, medium-bursty and high-bursty scenarios. Here, the performance of the LIDA scheme with β=0% was not considered due to its QoS violation. FQ-SDAM was found to achieve

TABLE I

Traffic Parameters in The Multi-cell WCDMA System

Traffic Type	Traffic Parameters
2-level real-time voice	Mean talkspurt duration: 1.00 seconds Mean silence duration: 1.35 seconds
High-bursty real-time data traffic	Peak rate ($R_{p,h}$): 4-fold of basic rate Mean rate: 1-fold of basic rate ρ_h: 0.25
Low-bursty real-time data traffic	Peak rate ($R_{p,l}$): 2-fold of basic rate Mean rate: 1-fold of basic rate ρ_l: 0.5
Non-real-time data traffic	Mean data burst size: 200 packets $r_{\min}$: 1-fold of basic rate $r_{\max}$: 8-fold of basic rate

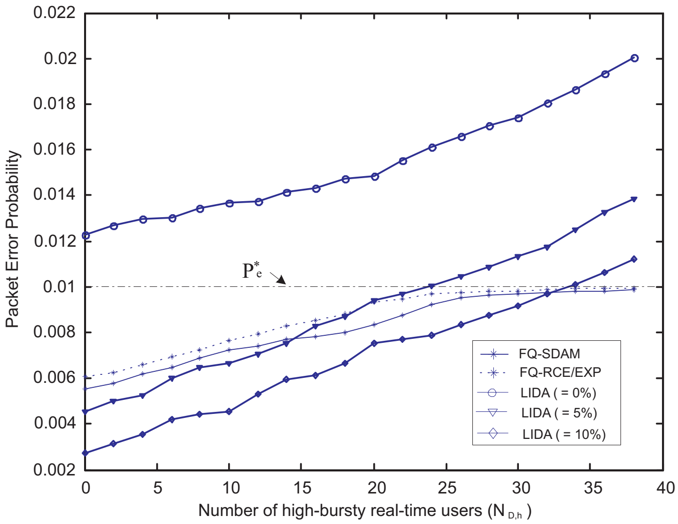

Fig. 2. Packet error probabilities: homogeneous case.

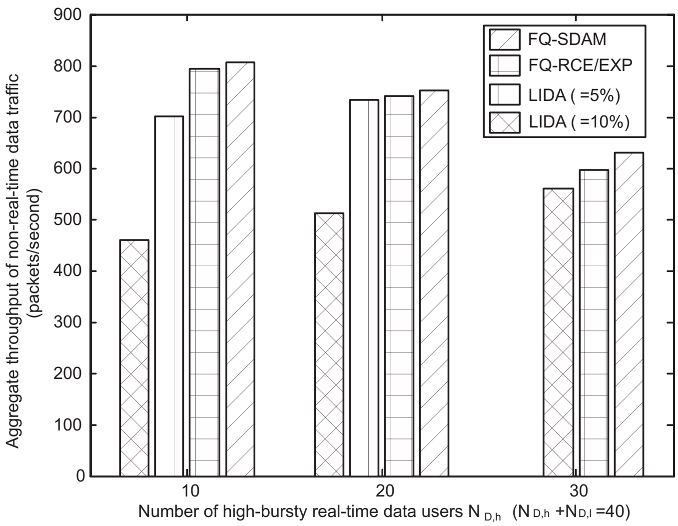

Fig. 3. Aggregate throughput of non-real-time data traffic: homogeneous case.

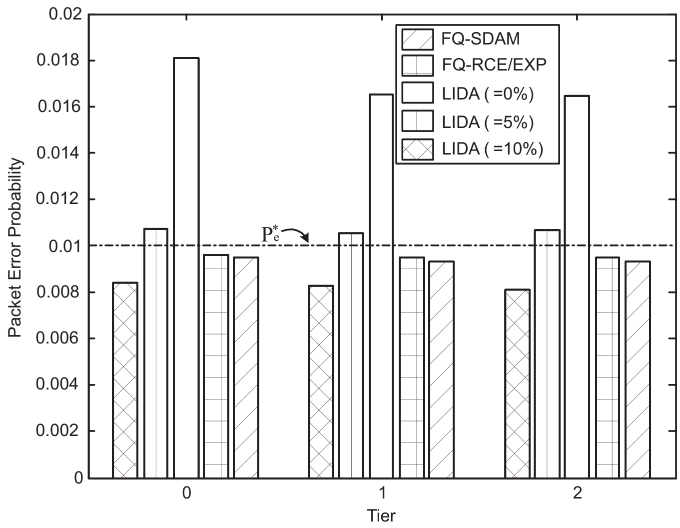

Fig. 4. Packet error probabilities: non-homogeneous case.

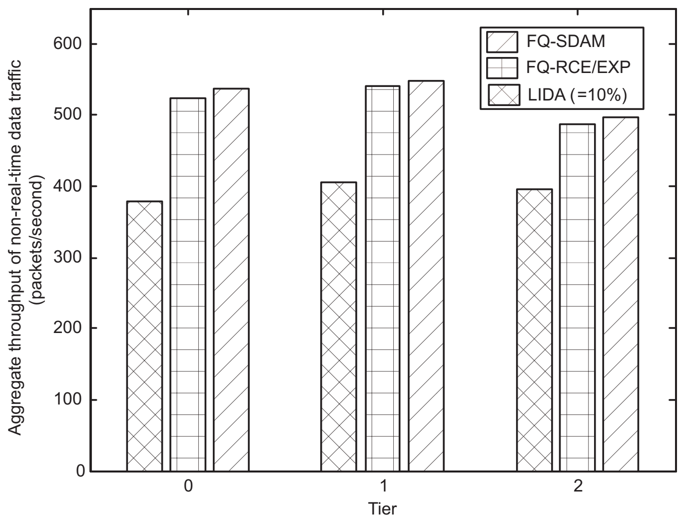

Fig. 5. Aggregate throughput of non-real-time data traffic: non-homogeneous case.

the highest data throughout for non-real-time services, while LIDA with β=10% produced the lowest throughput. Compared with the LIDA scheme with β=10%, the FQ-SDAM, FQ-RCE/EXP, and LIDA with β=5% improved the throughput by 75.3%, 73.3% and 52.9% (53.3%, 51.1% and 49.2%), respectively, in the low-bursty (medium-bursty) case. In the high-bursty case, under QoS constraint, FQ-SDAM and FQ-RCE/EXP schemes improved the throughput over the LIDA with β=10% by 16.8% and 10.7%, respectively, because FQ-SDAM approaches the desired transmission target (P_e^*=0.01) by fuzzy Q-learning. According to the definition of reinforcement signal $c(\mathbf{x}, a(\mathbf{x}))$, FQ-SDAM would try to allocate the maximum possible resource under the QoS requirement. By contrast, LIDA with β=10% is a conservative scheme, which has the lowest packet error probability at the expense of capacity waste. Additionally, in the three cases, the FQ-SDAM achieved a higher aggregate throughput than FQ-RCE/EXP by 1.4%, 1.43% and 5.5%, respectively. As the number of high-bursty real-time users goes up, the performance gain rises because the modified exponential rule considers the terminal's interference influence on adjacent cells and accordingly cuts the packet error probability in the multi-cell WCDMA environment. With a reinforcement signal containing a lower packet error probability, the FQ-RCE tends to allocate more capacity in the next-turn decision during the fuzzy Q-leaning period; consequently, the data throughput increases as more packets are successfully transmitted.

B. Non-homogeneous Case

In the non-homogeneous case, the real-time data terminals for the first-tier cells (cell 2 to cell 8) are: $N_{D,h} = 25 - 2 * (i-1)$ and $N_{D,l}$ = $40 - N_{D,h}$, i=2, $\cdots$, 8, while for the central and second-tier cells, the real-time data terminals are: $N_{D,h}$ = $N_{D,l}$ = 20.

Figure 4 shows the packet error probabilities of the three tiers in the multi-cell WCDMA system. As the figure reveals, only FQ-SDAM, FQ-RCE/EXP, and LIDA with β=10% meet the QoS requirement because FQ-SDAM and FQ-RCE/EXP consider the received adjacent-cell interference power as an input parameter for resource estimation. The resource allocation in the adjacent-cells is perceived by observing the interference fluctuation. Consequently, the resource allocations between cells can be conceptually coordinated implicitly. Additionally, compared to Fig. 2 at $N_{D,h} = 20$, the packet error probability in the non-homogeneous case is larger than that in the homogeneous case because the fluctuation of received adjacent-cell interference ,in the non-homogeneous case, differs from cell to cell when the cells compete for the residual capacity in the multi-cell environment. Without coordination, each cell allocates myopically, causing the system to over-loading.

Fig. 5 shows the aggregate throughputs of non-real-time data traffic in the three tiers of the multi-cell WCDMA system. Here, the aggregate throughputs of the LIDA with β=0% and β=5% are not considered due to their QoS violation. The aggregate throughput in the non-homogeneous case is smaller than that in the homogeneous case due to the higher interference fluctuation. Also, the FQ-SDAM and FQ-RCE/EXP schemes still achieves higher aggregate throughput by an amount of 31.53% and 28.346% (35.5% and 33.63%) (34.2% and 32%) for the cells in the central (first-tier) (second-tier) than the LIDA with $\beta = 10\%$ scheme does.

V. Concluding Remarks

This paper presents a novel situation-aware data access manager using fuzzy Q-learning technique (FQ-SDAM) for multi-cell WCDMA systems. The proposed scheme is designed with a fuzzy Q-learning-based residual capacity estimator (FQ-RCE) and a data rate scheduler (DRS). Through perceptual coordination, FQ-RCE considers the received home-cell interference power and adjacent-cell interference power as two separate linguistic variables such that it can adaptively determine the residual capacity according to the current loadings in the home and adjacent cells. Simulation results show that, compared to the LIDA scheme [7], the proposed FQ-SDAM can effectively reduce the packet error probability and improve the aggregate throughput of the non-real-time services in both the homogeneous and non-homogeneous multi-cell WCDMA environment because, using FQL, the FQ-RCE monitors the radio resource allocation in the adjacent cells, perceives the partial and uncertain information, and incrementally improves the residual system capacity estimation. Additionally, the DRS effectively allocates the resource for the non-real-time terminals with a modified exponential rule which considers the impact of the interference from terminals.

In practical implementation, compared to LIDA, FQ-SDAM requires additional computation complexity, which mainly comes from the operational calculation of fuzzy Q-learning. However, the additional complexity can be resolved by using some emerging fuzzy VLSI techniques [23], [24]. In [24], a VLSI fuzzy controller was designed for Sugeno fuzzy inference system. The VLSI fuzzy controller considers the weighting sum method in the defuzzification to prevent problems of the limited accuracy and stability problems. Therefore, the computation of the FQ-SDAM can be finished within a frame time, making the realization feasible and cost-effective.

Acknowledgment

The authors would like to give thanks to anonymous reviewers for their comments to improve the correctness and presentation of the paper. Especial thank goes to Dr. Scott Shen for his valuable comments and discussions.

References

[1] K. Das and S. D. Morgera, "Interference and SIR in integrated voice/data wireless DS-CDMA networks - a simulation study," *IEEE J. Select. Areas. Commun.*, vol. 15, no. 8, pp. 1527-1537, Oct. 1997.

[2] T. K. Liu and J. A. Silvester, "Joint admission/congestion control for wireless CDMA systems supporting integrated services," *IEEE J. Select. Areas. Commun.*, vol. 16, no. 6, pp. 845-857, Aug. 1998.

[3] C. Comaniciu and N. B. Mandayam, "Delta modulation based prediction for access control in integrated voice/data CDMA systems," *IEEE J. Select. Areas. Commun.*, vol. 18, no. 1, pp. 112-122, Jan. 2000.

[4] A. Sampath and J. M. Holtzman, "Access control of data in integrated voice/data CDMA systems: benifits and tradeoffs," *IEEE J. Select. Areas. Commun.*, vol. 15, no. 8, pp. 1511-1526, Oct. 1997.

[5] C. Comaniciu, N. B. Mandayam, D. Famolari, and P. Agrawal, "Wireless access to the world wide web in an integrated CDMA system," *IEEE J. Select. Areas. Commun.*, vol. 2, pp. 472-483, May 2003

[6] L. Chen, H. Kayama, and N. Umeda, "Power resource cooperation control considering wireless QoS for CDMA packet mobile communication systems," in *Proc. IEEE Int'l Symposium on Personal, Indoor and Mobile Radio Communications* (PIMRC 2002), vol. 3, pp. 1092-1096.

[7] S. Kumar and S. Nanda, "High data-rate packet communication for cellular network using CDMA: algorithm and performance," *IEEE J. Select. Areas. Commun.*, vol. 17, no. 3, pp. 472-492, Mar. 1999.

[8] S. Ramakrishna and J. M. Holtzman, "A scheme for throughput maximization in a dual-class CDMA system, " *IEEE J. Select. Areas. Commun.*, vol. 16, no. 6, pp. 830-844, Aug. 1998.

[9] A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," in *Proc. IEEE Vehicular Technology Conference (VTC2000-Spring)*, pp. 1854-1858.

[10] S. Shakkottai and A. L. Stolyar, "Scheduling algorithms for a mixture of real-time and non-real-time data in HDR," in *Proc. 17th International Teletraffic Congress (ITC-17)*, Sept. 2001.

[11] L. Wang (Ed), *Soft Computing in Communications*, Springer, 2003.

[12] Y. S. Chen and C. J. Chang, "A resource allocation scheme using adaptive-network based fuzzy control for mobile multimedia network," *IEICE Trans. Commun.*, vol. E85-B, no. 2, pp. 502-513, Feb. 2002.

[13] J. Nie and S. Haykin, "A Q-learning-based dynamic channel assignment technique for mobile communication systems," *IEEE Trans. Veh. Technol.*, vol. 48, no. 5, pp. 1676-1687, Sept. 1999.

[14] Y. S. Chen, C. J. Chang, and F. C. Ren, "A Q-learning-based multi-rate transmission control scheme for RRM in multimedia WCDMA systems," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp 38-48, Jan. 2004.

[15] L. Jouffle, "Fuzzy inference system learning by reinforcement methods," *IEEE Trans. Syst. Man. Cybern.*, vol. 8, no. 3, pp. 338-355, Aug. 1998.

[16] O. Abul, F. Polat, and R. Alhajj, "Multiagent reinforcement learning using function approximation," *IEEE Trans. Syst. Man. Cybern.*, vol. 30, no. 4, pp. 485-497, Nov. 2000.

[17] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, pp. 279-292, 1992.

[18] S. Ariyavistakul and L. F. Chang, "Signal and interference statistics of a CDMA system with feedback power control," *IEEE Trans. Commun.*, vol. 41, no. 11, pp. 1626-1634, Nov. 1993.

[19] G. L. Stüber, *Priniciple of Mobile Communication*, Kluwer Academic Publishers, 1996.

[20] S. Haykin, *Neural Networks, Second Edition*. Prentice Hall, 1999.

[21] 3rd Generation Partnership Project, (Sept. 2002), Spreading and Modulation (FDD), 3GPP TS 25.213 [online]http://www.3gpp.org.

[22] P. T. Brady, "A model for on-off speech patterns in two-way conversation," *Bell Syst. Tech. J.*, vol. 48, pp. 2445-2472, Jan. 1969.

[23] J. L. Huertas, S. Sanchez-Solano, I. Baturone, and A. Barriga, "Integrated circuit implementation of fuzzy controllers," *IEEE J. Solid-State Circuits*, vol. 31, no. 7 , pp. 1051-1058, July 1996.

[24] B. M. Wilamowski, "Neuro-fuzzy systems and their applications," in *Proc. IECON'98*, vol. 1, pp. 35-49.

Yih-Shen Chen was born in Taiwan, Republic of China, in September 1973. He received the B.E., M.E. and Ph. D degree in communication engineering from Nation Chiao-Tung University, Hsinchu, Taiwan, in 1995, 1997 and 2004, respectively. In 2005, he joined a 3G protocol software project in Sunplus Technology, Taiwan. His research interests include performance analysis, protocol design, and PCS (personal communication service) networks.

Chung-Ju Chang was born in Taiwan, R.O.C., in August 1950. He received the B.E. and M.E. degrees in electronics engineering from National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 1972 and 1976, respectively, and the Ph.D degree in electrical engineering from National Taiwan University (NTU), Taiwan, in 1985. From 1976 to 1988, he was with Telecommunication Laboratories, Directorate General of Telecommunications, Ministry of Communications, Taiwan, as a Design Engineer, Supervisor, Project Manager, and then Division Director. There, he was involved in designing digital switching system, RAX trunk tester, ISDN user-network interface, and ISDN service and technology trials in Science-Based Industrial Park. In the meantime, he also acted as a Science and Technical Advisor for the Minister of the Ministry of Communications from 1987 to 1989. In 1988, he joined the Faculty of the Department of Communication Engineering and Center for Telecommunications Research, College of Electrical Engineering and Computer Science, National Chiao Tung University, as an Associate Professor. He has been a Professor since 1993. He was Director of the Institute of Communication Engineering from August 1993 to July 1995, Chairman of Department of Communication Engineering from August 1999 to July 2001, and the Dean of the Research and Development Office from August 2002 to July 2004. Also, he was an Advisor for the Ministry of Education to promote the education of communication science and technologies for colleges and universities in Taiwan during 1995 - 1999; he is acting as a Committee Member of the Telecommunication Deliberate Body,Taiwan. Moreover, he serves as Editor for *IEEE Communications Magazine* and Associate Editor for *IEEE Transactions on Vehicular Technology*. His research interests include performance evaluation, wireless communication networks, and broadband networks. Dr. Chang is a member of the Chinese Institute of Engineers (CIE).

Fang-Chin Ren was born in Hsinchu, Taiwan. He received B.E., M.E., and Ph.D. degrees in communication engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1992, 1994, and 2001, respectively. Since 2001, he has been an Engineer in the Protocol Design Technology Department, Computer and Communications Research Laboratories, Industrial Technology Research Institute, Taiwan, where he was involved in the designing of the radio resource managemnt for the third-generation wireless networks. His current research interests include performance analysis, protocol design, and mobile radio network.