行政院國家科學委員會專題研究計畫　成果報告

子計畫一：異質多接取網路之資源管理技術(2/2)

計畫類別：整合型計畫
計畫編號：NSC93-2219-E-009-011-
執行期間：93 年 08 月 01 日至 94 年 07 月 31 日
執行單位：國立交通大學電信工程學系(所)

計畫主持人：張仲儒
共同主持人：廖維國，王蒞君
計畫參與人員：陳詠翰、楊煖玉、顏志明、郭立忠、李宗軒

報告類型：完整報告
報告附件：出席國際會議研究心得報告及發表論文
處理方式：本計畫可公開查詢

中　華　民　國 94 年 10 月 27 日

摘要

為實現 B3G 中異質性多接取網路的高效率資源配置以及服務品質（qualtiy-of-service，QoS）保證，網路端實體層（PHY）與媒體控制層（MAC）如何共同相互配合達到高系統使用效率、與通訊環境高適應性，將直接決定多接取網路的整體效能。在本子計畫中，我們的研究利用各項智慧型技術，設計 B3G 系統中關鍵的無線多接取資源管理機制，包括應用於 B3G 的智慧型資料存取管理技術、智慧型多媒體排程機制，以及自動偵測通訊狀況的動態細胞配置之研究。其中智慧型資料存取管理採用了 fuzzy Q-learning 的技術，有效地監控系統通訊品質，主要包括來自同細胞與鄰近細胞的干擾狀況，並且設計 fuzzy Q-learning residual capacity estimator（FQ-RCE）剩餘容量估測器準確地估測系統剩餘容量，以及傳輸速率排程（data rate scheduler，DRS），達到最有效率的存取估測與控制，而由 FQ-RCE 與 DRS 所獲得的即時系統資訊就能作為資源配置與排程之用。有了這些資訊，我們更進一步混和利用蜂巢式神經網路（cellular neural network，CNN）以及效能函式（utility functions）進行排程控制器（CNNU-based scheduler）的設計研究。在我們所提出的排程控制機制中，CNNU-based scheduler 能決定無線資源使用狀況，並且根據系統動態變化進行資源配置，以達最大系統使用效率與最佳傳輸效能之目的，而且能維持 QoS 保證。而通訊系統動態變化控制則可透過動態細胞配置（dynamic cell configuration）的機制達成。在這項設計中，我們同時考慮到 soft handoff、link power allocation 以及 admission control 範圍的條件，採用了 reinforcement-learning 的技術動態調整系統實體層 pilot 的大小，也就是動態調整通訊覆蓋範圍，以達最佳的功率配置效能，並且進一步平衡系統負載。

關鍵字：異質性多接取網路，服務品質，多媒體排程，動態細胞配置，fuzzy Q-learning，剩餘容量估測器，傳輸速率排程，蜂巢式神經網路，效能函式，reinforcement-learning

Abstract

In order to achieve high-efficiency resource allocation and quality-of-service (QoS) guarantee in B3G heterogeneous multiple access networks, the collaboration of resource control in PHY and MAC layers will directly effect the system performance. In the subproject, we take advantage of intelligent technologies to design three critical mechanisms, which includes intelligent data access management, intelligent multimedia scheduling, and situation-aware dynamic cell configuration. The intelligent data access management adopts fuzzy Q-learning technology to monitor the communication situations such as inter-cell and intra-cell interference. The fuzzy Q-learning residual capacity estimator (FQ-RCE) and data rate scheduler (DRS) are proposed to efficiently estimate and control system resources. The real-time system information from FQ-RCE and DRS can further supports radio resource allocation and scheduling. Then we propose a cellular neural network utility (CNNU)-based scheduler, which combines the technologies of cellular neural network (CNN) and utility function. The CNNU-based scheduler decides the radio resource situations and allocations according to the system changes. With QoS guarantee, the CNNU-based scheduler can achieve maximum system utilization and throughput. And the system changes can be controlled by advanced dynamic cell configuration. In the design, we consider soft handoff, link power allocation, and admission control ranges to dynamically adjust the power of pilot in PHY by using reinforcement-learning technology. This will change the coverage of the controlled cell to maximize the performance of power allocation and load-balancing.

Keywords: heterogeneous multiple access network, QoS, multimedia scheduling, dynamic cell configuration, fuzzy Q-learning, FQ-RCE, DRS, CNN, utility function, reinforcement-learning

Contents

List of Figures

Chapter 2

Chapter 3

Chapter 4

List of Tables

Chapter 1

Project Overview

The applications of multimedia services over wideband communication networks increase dramatically in recent years. In order to support a diverse of multimedia applications, the next-generation broadband networks have been required to satisfy the Quality of Service (QoS) requirements. Real-time and precise traffic control and scheduling mechanisms are essential to achieve the QoS guarantee and maximum utilization. Major topics about traffic control and resource management include link situation awareness, capacity estimation, rate allocation, traffic scheduling, call admission, resource monitoring, and cell configuration. According to the required services, users can access the network through the beyond third generation (B3G) mobile systems, which could be composed of heterogeneous networks. The radio resource management (RRM) in the heterogeneous network is essential, and the

performance of RRM design affects the utilization directly. But it is always challenging to find the ways of achieving best system utilization while maintaining QoS of every service in such B3G systems. In order to maximize the utilization in B3G systems, we should focus on some essential elements of RRM to efficiently allocate, manage, and monitor the radio resources. Therefore the key technologies of RRM in B3G systems are the most critical points to provide comprehensive and satisfactory mobile communication experience.

In the subproject, we propose a set of intelligent RRM schemes to reach our goals. In the second chapter, we propose 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, which considers the effect of interference on adjacent cells.

In the third chapter, a cellular neural network and utility (CNNU)-based scheduler is proposed for multimedia CDMA cellular networks supporting differentiated quality-of-service (QoS). The cellular neural network is powerful for complicated optimization problems and has been proved that it can rapidly converge to a desired equilibrium; the utility-based scheduling algorithm can efficiently utilize the radio resource for system and provide QoS requirements and fairness for connections. A relevant utility function for each connection is here defined as its radio resource function further weighted by both a QoS requirement deviation function and a fairness compensation function. The CNNU-based scheduler determines a radio resource assignment vector for all connections so that the overall system utility is maximized and the system throughput can be achieved as high as possible. At the same time, the performance measures of all connections are kept closed to their QoS requirements in an efficient way.

The fourth chapter presents a novel dynamic cell configuration scheme in next-generation situation-aware CDMA networks. To balance the time-varying traffic load between cells,

caused by user mobility and diverse applications, it is crucial for next-generation CDMA cellular networks to configure cell coverage and capacity dynamically. In this chapter, we show that pilot power allocation is highly coupled to other facets of radio resource management. We propose a novel dynamic cell configuration scheme for multimedia CDMA cellular networks, based on reinforcement-learning, which takes into account pilot, soft handoff, and maximum link power allocations as well as call admission control mechanisms. Simulation results demonstrate the effectiveness of the proposed scheme in situation-aware CDMA networks.

Chapter 2

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

I. Introduction

The 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 over-loading, 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 become 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 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 good 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 part 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 link capacity definition. 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 this chapter 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.

II. System Model

This part 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 $\text{SF} = 256/2^k, k = 0, 1, \cdots, 6$, 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), 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) [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 **S** the set of state vectors for the system, **S**=$\{S_i,\ i = 1, 2, \cdots, M\}$; each state vector S_i comprises L fuzzy linguistic variables selected to describe the system. Denote **A** the set of actions possibly chosen by system states, **A**=$\{A_j,\ j = 1, 2, \cdots, N\}$. For an input state vector **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 $q[S_i, A_j]$, $1 \leq i \leq M$ and $1 \leq j \leq N$,

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 **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 **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, $0 \leq \eta \leq 1$, and

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

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

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

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

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

$$RC = I_{h2} + P_{inc}, \quad (9)$$

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

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

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

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

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

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

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

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

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 up-link 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 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.

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,” *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,” *IEEE VTC2000-Spring*, Tokyo, May 2000, 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,” *17th International Teletraffic Congress* (ITC-17), Sep., 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, Sep. 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. Comm.*, no. 11, Nov. 1993, pp. 1626-1634.

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

[20] S. Haykin, Neural Networks 2nd. Prentice Hall, 1999.

[21] 3rd Generation Partnership Project, (Sep. 2002) Spreading and Modulation (FDD), 3GPP TS 25.213 [On-line]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, I, and A. Barriga, ”Integrated circuit implementation of fuzzy controllers,” *IEEE J. Solid-State Circuits*, vol. 31, no. 7 , pp. 1051-1058, Jul. 1996.

[24] B.M. Wilamowski, ”Neuro-fuzzy systems and their applications,” *IECON ’98*, vol. 1 pp. 35-49.

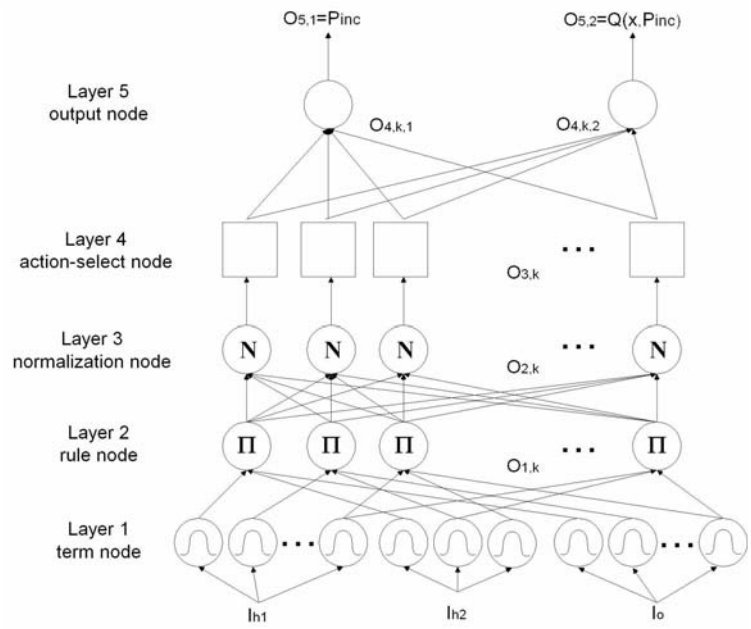

Fig. 1. Structure of FQ-RCE

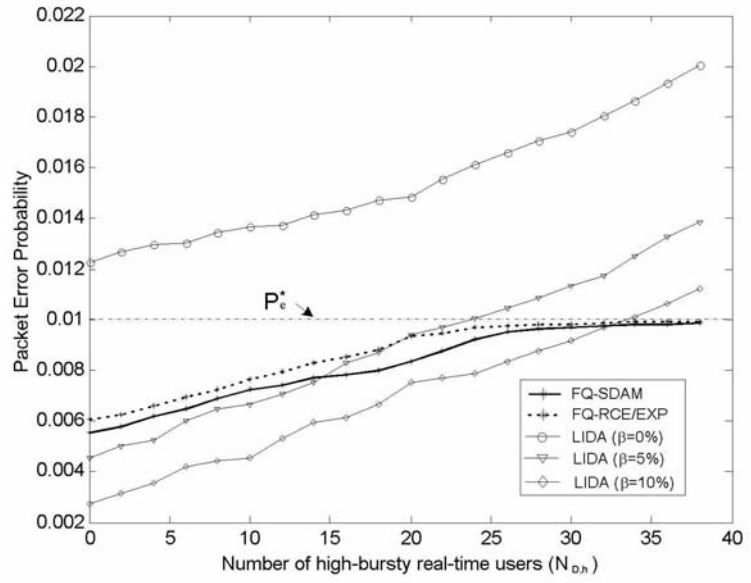

Fig. 2. Packet error probabilities: homogeneous case

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

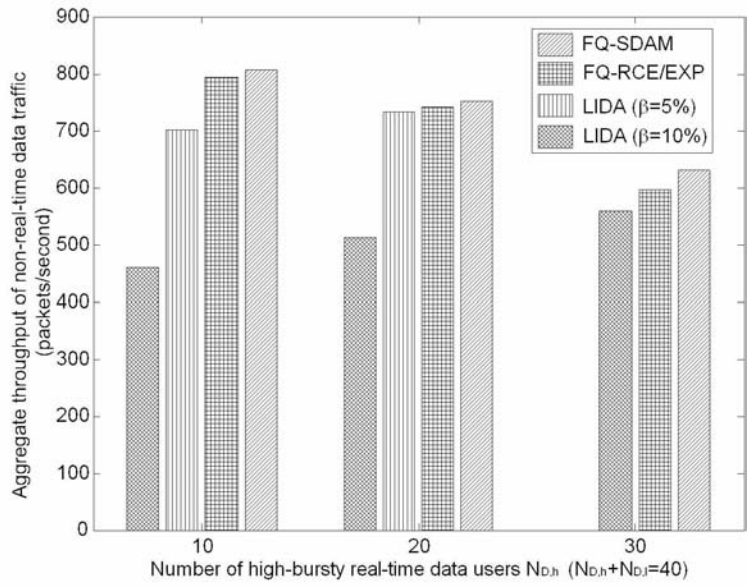

Fig. 3. Aggregate throughput of non-real-time data traffic: homogeneous case

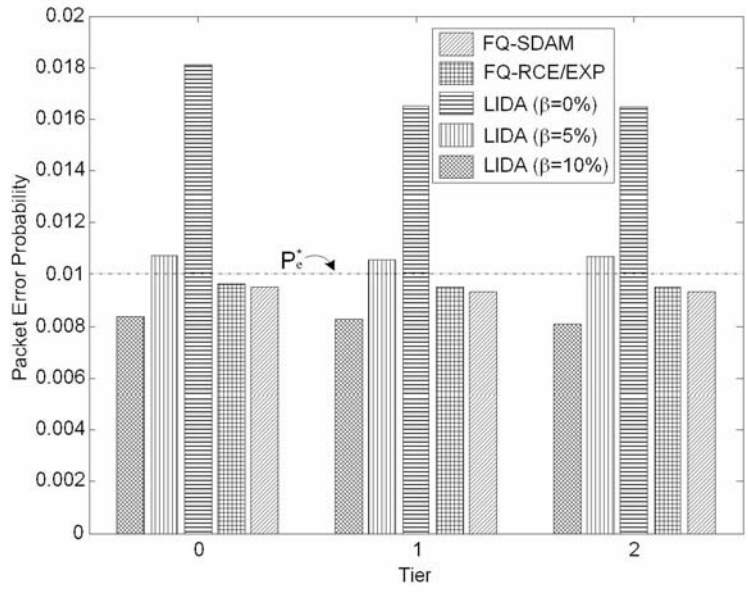

Fig. 4. Packet error probabilities: non-homogeneous case

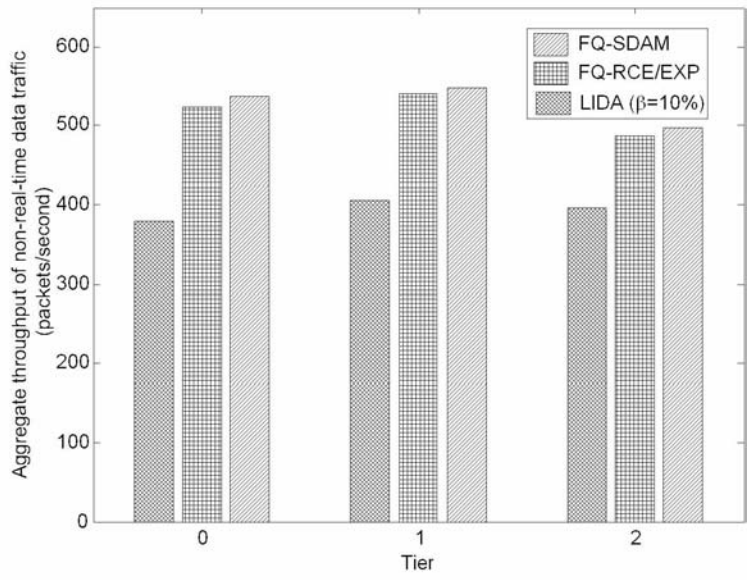

Fig. 5. Aggregate throughput of non-real-time data traffic: non-homogeneous case

Chapter 3

A Cellular Neural Network and Utility-based Scheduler for Multimedia CDMA Cellular Networks

I. Introduction

In future wireless networks, heterogeneous and customized services with diverse traffic characteristics and QoS requirements are expected to be provided via a number of air interfaces. Also, multimedia applications are commonly accepted as enabling services, which are categorized into several classes [1]. To meet various traffic characteristics and QoS requirements of these potential

applications, a sophisticated scheduling algorithm plays an essential role so that the system resource allocation is optimal, while retaining a pre-defined QoS requirements and fairness among them.

Many scheduling algorithms have been widely studied for wireline networks [2]-[3]. In the wireless communication networks, the radio channel have quite different characteristics from those in wireline networks. The transmission error probability is by several order greater than that in wireline links, and the available maximum transmission rate to each connection is location-dependent and time-varying due to link loss, shadowing, and multi-path fading. The QoS requirements and the weighted fairness among all connections should be modified.

The literature studied the resource scheduling and allocation among connections with consideration of physical layer processing, power control range, and link conditions [4]-[5]. Bhargharvan, Lu, and Nandagopal [6] proposed a framework to achieve long-term fairness in wireless network. Varsou and Poor [7] proposed another class of scheduling algorithm from EDF concept in wireless environment. This class of scheme considers delay bound as its QoS requirement. In [8], a throughput-optimal scheduling algorithm for delay bounded system was proposed and proved. Shakkottai and Stolyar [10] considered both link quality and QoS requirements as the criteria and derived the exponential form of scheduling function via fluid Markovian techniques. Many of these scheduling algorithms above, [4]-[5], [8]-[10], were formulated in utility-based approaches.

The utility-based scheduling algorithm over radio channels, is usually formulated as a complicated constrained optimization problem with real time requirement. To solve this optimization problem, the class of generalized HNN has been adopted for real-time tasks with several inherent defficiencies. A special type of Hopfield neural networks (HNN), named *cellular neural network* (CNN) proposed in [11], has been proved that it can rapidly converge to desired equilibrium on vertex along the prescribed trajectories by proper design [12]. The CNN was widely applied in image processing field and was suitable for VLSI implementation. However, to adopt the CNN technique for the scheduling optimization problem, modifications of its architecture are necessary.

In the paper, we propose a CNN and utility (CNNU)-based scheduler for downlink in multimedia CDMA cellular networks. The CNNU-based scheduler contains a utility function (UF) preprocessor, a radio-resource range (RR) decision maker, and a CNN processor. Noticeably, the

utility function for each connection, adopted in the UF preprocessor, jointly considers radio resource efficiency, diverse QoS requirements, and fairness. It is a radio resource function weighted by both its QoS requirement deviation function and its fairness compensation function. The UF preprocessor generates a matrix of normalized utility functions of all connections. On the other hand, the RR decision maker determines a matrix showing the upper limit of radio resource assignment for each connection. The CNN processor receives the two matrix as inputs and determines an optimal normalized radio resource assignment vector for connections in multimedia CDMA cellular systems, by minimizing the system cost function which is in terms of the overall system utility function under system constraints of maximum transmission power, minimum spreading factor, and remaining queue length. The architecture of the CNN is constructed via the energy-based approach [13]-[14]. by mapping the system cost function to a proper energy function. It is designed in a two-layered configuration, which consists of a decision layer and an output layer, to reduce the number of inter-connections in the CNN. It can be shown that the stable equilibriums locate in the desired state space and the stability exists. The performance of the proposed CNNU-based scheduler is investigated by comparing with *Exponential Rule* [10] for systems using both dedicated and shared channel. Results show that the CNNU-based scheduler is efficient and effective for multimedia CDMA cellular networks.

The rest of the paper is organized as follows. Section II presents the features and the operations of the considered system. In section III, an relevant utility function is then proposed. In section IV, the architecture of CNNU-based scheduler and the structure of CNN are discussed.

II. System Model

Assume that there are N real-time (RT) and non-real-time (NRT) connections (users) in the downlink transmissions of the multimedia CDMA cellular system with chip rate W. RT connections transmit on dedicated channels and NRT connections transmit on shared channels. For every active connection using either dedicated or shared channels, a fixed number of code channels with their corresponding spreading factors are given in the connection setup phase. A minimum spreading factor SF_i is therefore associated with the assigned code channels for connection i. The system radio resource is here defined to be the transmission power. It is limited by a maximum power budget denoted by P^*_{max} and scheduled to all connections every frame

time period T_f.

For a downlink connection i, there are four QoS requirements defined in either the packet level, such as BER_i^*, or the call level, such as delay bound D_i^*, packet dropping ratio $P_{D,i}^*$, and minimum transmission rate $R_{m,i}^*$. For RT connections, hard delay bound D_i^* exists and $P_{D,i}^*$ can be larger than zero; while for NRT connections, no explicit delay bound is imposed, but $R_{m,i}^* > 0$ should be satisfied for interactive connections and $R_{m,i}^* = 0$ be set for best effort connections.

For a RT connection i, a transmission suspension in a soft fashion is carried out by allocating zero transmission power when its utility calculated by the scheduler is lower than those of NRT connections. At that moment, its link gain $\zeta_i(t)$ is lower than the averaged mean link gains of all NRT connections $\overline{\zeta}_{NRT}$ by a relative margin, and this relative margin should be considered to restrict the probability of transmission suspension below $P_{D,i}^*$ due to the delay-sensitive nature. Denote by ζ_i^* the suspension threshold of connection i, which is obtained by $\mathbf{P}\{\zeta_i(t) \leq \zeta_i^*\} \leq P_{D,i}^*$. Then the relative margin of $\zeta_i(t)$ is a function of $\overline{\zeta}_{NRT}$ and ζ_i^*, and is dependent on the design of scheduling algorithm. For NRT connections, their transmissions are scheduled so that NRT connections will be allocated with proper radio resource to achieve high system utilization and keep the fairness and the QoS requirements fulfilled as much as possible.

Assume that the link-gain $\zeta_i(t)$ and the interference $\mathcal{I}_i(t)$ for connection i at time t can be measured at the user side and perfectly signaled to the base station. The $\zeta_i(t)$ consists of the mean path loss, long-term fading, and short-term fading, and is given by $\zeta_i(t) = d_i^{-4} \cdot 10^{\frac{\zeta_i^L(t)}{10}} \cdot \zeta_i^S(t)$, where d_i is the distance between the user i and its base station, $\zeta_i^L(t)$ is the log-normal shadowing component, and $\zeta_i^S(t)$ is the Rayleigh-fading component. The adaptive QAM modulation is adopted and the modulation order M_{κ_i} with index κ_i for connection i is determined according to the link gain quality and interference. The traffic source of connection i generates packets and packets are queued in its individual buffer. The buffer size is infinite. The source models are assumed to be on-off for RT connections, Perato for NRT interactive (NRT-I) connections, and batch Poisson with truncated geometrical batch size for NRT best-effort (NRT-B) connections.

The proposed CNNU-based scheduler determines an optimal normalized radio resource assignment vector $\vec{c}^*(t) = (c_1^*(t), \ldots, c_N^*(t))$ to N connections via maximizing an overall system utility function. The transmission rate for connection i at t-th frame , denoted by $r_i(t)$, is then allocated according to $c_i(t)$ of connection i.

III. Formulation of The Utility Function

The utility function for connection i, $\mathcal{U}_i(t)$, is defined as the radio resource function of connection i, $\mathcal{R}_i(t)$, weighted by its QoS requirement deviation function $\mathcal{A}_i(t)$ and its fairness compensation function $\mathcal{F}_i(t)$. It can be expressed as

$$\mathcal{U}_i(t) = \mathcal{R}_i(t) \cdot \mathcal{A}_i(t) \cdot \mathcal{F}_i(t). \tag{1}$$

A. Radio Resource Function $\mathcal{R}_i(t)$

With the modulation order M_{κ_i} of the adaptive QAM modulation scheme and the corresponding $(E_b/N_0)^*_{\kappa_i}$ to satisfy the BER_i^* requirement for connection i, the following inequality should hold

$$\frac{W}{R_{s,i}(t)} \cdot \frac{c_i(t) \cdot P^*_{max} \cdot \zeta_i(t)}{\mathcal{I}_i(t)} \geq \left(\frac{E_b}{N_o}\right)^*_{\kappa_i}, \tag{2}$$

where $R_{s,i}(t)$ is its symbol rate and $c_i(t)$ is its normalized radio resource assignment at time t. The $\mathcal{I}_i(t)$ in (2) is given by $[(1-\alpha)P^*_{max} \cdot \zeta_i(t) + \sum_b P^*_{max} \cdot \zeta_{i,b}(t) + N_0 W]$, where α is the orthogonality factor for downlink, b is the index referring to the adjacent base stations, $\zeta_{i,b}(t)$ is the link gain from base station b to connection i, and the $(E_b/N_0)^*_{\kappa_i}$ in (2) is given by $\frac{-(M_{\kappa_i}-1)\cdot ln\{5BER_i^*\}}{1.5}$. We denote the *maximum achievable symbol rate* that can fulfill the $(E_b/N_0)^*_{\kappa_i}$ at $c_i(t) = 1$ by $R^*_{s,i}(t)$. Clearly, $R^*_{s,i}(t) = \frac{W}{(E_b/N_0)^*_{\kappa_i}} \cdot \frac{P^*_{max}\cdot\zeta_i(t)}{\mathcal{I}_i(t)}$. The $R^*_{s,i}(t)$ is further limited by $\frac{W}{SF_i}$ for a given spreading factor SF_i of the allocated code channel. Thus the $R^*_{s,i}(t)$ can be obtained by

$$R^*_{s,i}(t) = \min\left\{\frac{W}{(E_b/N_0)^*_{\kappa_i}} \cdot \frac{P^*_{max} \cdot \zeta_i(t)}{\mathcal{I}_i(t)},\ \frac{W}{SF_i}\right\}. \tag{3}$$

According to (3), the most efficient modulation order M_{κ_i} is selected by the following inequality,

$$M_{\kappa_i} \leq \frac{SF_i \cdot P^*_{max} \cdot \zeta_i(t)}{\mathcal{I}_i(t) \cdot \left(\frac{-ln\{5BER_i^*\}}{1.5}\right)} + 1 \leq M_{(\kappa_i+1)}. \tag{4}$$

Since the information bit of one symbol is $log_2 M_{\kappa_i}$, consequently the radio resource function of connection i, $\mathcal{R}_i(t)$, can be obtained by

$$\mathcal{R}_i(t)\frac{1.5W \cdot log_2 M_{\kappa_i}}{(M_{\kappa_i}-1) \cdot \left[ln\frac{1}{BER_i^*} - ln5\right]} \cdot \frac{P^*_{max} \cdot \zeta_i(t)}{\mathcal{I}_i(t)}. \tag{5}$$

Note that if the assignment of radio resource for connection i, $c_i(t)$, is allocated, the transmission rate $r_i(t)$ is therefore equal to $c_i(t) \cdot \mathcal{R}_i(t)$.

B. The QoS Requirement Deviation Function $\mathcal{A}_i(t)$

The QoS requirement deviation function $\mathcal{A}_i(t)$ is used to indicate how much extent the connection i deviates from its call-level QoS requirements. For a RT connection i, a hard delay bound D_i^* is imposed on each packet. Since QoS over wireless interface can be provided in a soft fashion, the QoS guarantee of packet dropping ratio due to excess delay is expressed by $P_{rob}\{D_i(t) > D_i^*\} < P_{D,i}^*$, where $D_i(t)$ is the waiting time delay for head-of-line packet at time t. For an NRT interactive (NRT-I) connection i, a different notion of QoS requirement is that a minimum transmission rate must be guaranteed by $\mathbf{E}[r_i(t)] \geq R_{m,i}^*$. As for an NRT best-effort (NRT-B) connection i, no call level QoS requirements are guaranteed and the $R_{m,i}^*$ is set to be 0.

From [18], the proposed *Modified Largest Weighted Delay First* (M-LWDF) algorithm suggests that an exponential rule [10] be the form with throughput optimal for the above call level QoS requirement constraints. Therefore, the *QoS requirement deviation function* $\mathcal{A}_i(t)$ is defined as

$$\mathcal{A}_i(t) = \begin{cases} exp\left\{\frac{\frac{-log(P_{D,i}^*)}{D_i^*}\cdot D_i(t) - \overline{D}(t)}{1+\left[\overline{D}(t)\right]^{1/2}}\right\}, & \text{if } i \in \{\text{RT}\}, \\ exp\left\{\frac{\hat{L}_i(t)-\overline{L}(t)}{1+\left[\overline{L}(t)\right]^{1/2}}\right\}, & \text{if } i \in \{\text{NRT-I}\}, \\ 1, & \text{if } i \in \{\text{NRT-B}\}, \end{cases} \tag{6}$$

where $\overline{D}(t) = \frac{1}{N}\sum_i \left(\frac{-log(P_{D,i}^*)}{D_i^*}\right) \cdot D_i(t)$ is the average weighted delay, $\hat{L}_i(t) = \hat{L}_i(t-1) + \left(\frac{R_{m,i}^* - r_i(t)}{R_{m,i}^*}\right)$ is the normalized measurement on the difference of guaranteed minimum transmission rate and the assigned rate, and $\overline{L}(t) = \frac{1}{N}\sum_i \hat{L}_i(t)$. For the RT connections, if the weighted delay is more than the average weighted delay of all connections, the $\mathcal{A}_i(t)$ will be exponentially increased, and more resource will be scheduled; on the other hand, if the weighted delay is less than the average weighted delay, the $\mathcal{A}_i(t)$ will dramatically decayed, and less resource will be allocated. Similarly, for the NRT-I connections, if the accumulated difference of the guaranteed minimum transmission rate and the assigned rate is greater than the average value, more resource is assigned. As to the NRT-B connections, this function is simply bypassed.

C. The Fairness Compensation Function $\mathcal{F}_i(t)$

The fairness compensation function is to ensure that RT connections using dedicated channels have the relative priority over NRT connections using shared channels. It is also the way that

the radio resource shared by all NRT connections is assigned according to a predefined target weighting factor. With the pre-defined target weighting factor w_i for NRT connections i for $1 \leq i \leq N$, the radio resources are here expected to be allocated to any two any connections, i and k, so that their average assigned transmission rates, $\mathbf{E}[r_i(t)]$ and $\mathbf{E}[r_k(t)]$, can be achieved by $\frac{\mathbf{E}[r_i(t)]}{\mathbf{E}[r_k(t)]} = \frac{w_i}{w_k}$ [6]. The *fairness compensation function* for connection i till time t, $\mathcal{F}_i(t)$, is defined by,

$$\mathcal{F}_i(t) = \begin{cases} \beta_i, \text{ if } i \in \{\text{RT}\}, \\ \beta_0 \cdot \left[(w_i - \bar{w}_i(t) - 1)^+ + 1\right], \text{ if } i \in \{\text{NRT}\}, \end{cases} \tag{8}$$

where β_i is the priority bias for RT connections to differentiate with NRT connections, β_0 is the basic reference value set for NRT connections, $(x)^+ = max\{x, 0\}$, and $\overline{w}_i(t)$ is the moving-average of $r_i(t)$. For RT connections, the weighted fairness is not considered and only priority bias is set due to their QoS-driven nature. For NRT connections, $[(w_i - \bar{w}_i(t) - 1)^+ + 1] = max\{(w_i - \bar{w}_i(t)), 1\}$, where $(w_i - \bar{w}_i(t))$ indicates the unfairness of connections i, and the $\mathcal{F}_i(t)$ is limited by β_0 after $\bar{w}_i(t)$ is above w_i and will make no further effects on the utility. The more extent of the unfairness of connection i is, the larger the $\mathcal{F}_i(t)$ will be; then more resource will be scheduled to connection i, and the $(w_i - \bar{w}_i(t))$ will be smaller afterwards. In the stationary situation, the unfairness of all NRT connections should be almost the same via the linear feedback control.

The target weighting factor w_i is defined as the target average transmission rate of NRT connection i. It is considered to be a function of its equivalent traffic source rate s_i^*, mean link gain $\bar{\zeta}_i$, mean interference level $\bar{\mathcal{I}}_i$, and its guaranteed minimum transmission rate, $R_{m,i}^*$. The w_i is given by $w_i = max\left\{\frac{P_{max}^* \cdot \bar{\zeta}_i}{\bar{\mathcal{I}}_i \cdot (\frac{E_b}{N_0})_i^*} \cdot \frac{s_i^*}{\sum_k s_k^*},\ R_{m,i}^*\right\}$,where $(\frac{E_b}{N_0})_i^*$ is the required E_b/N_0 to achieve BER_i^* of connection i using the least-order modulation scheme. Note that $R_{m,i}^* = 0$ for the best-effort connection, and the target weighting factor of the best-effort connection is usually less than that of the interactive connection. The priority bias β_i for RT connection i is a relative margin for $\zeta_i(t)$ over the link gains of NRT connections, and is a function of its transmission suspension threshold ζ_i^*, the average of the mean link gains of all NRT connections $\overline{\zeta}_{NRT}$, and the E_b/N_0 requirements. The β_i is given by $\beta_i = \left(\frac{\bar{\zeta}_{NRT}}{\zeta_i^*} \cdot \frac{(\frac{E_b}{N_0})_i^*}{(\frac{E_b}{N_0})_{NRT}^*}\right) \cdot \beta_0$.

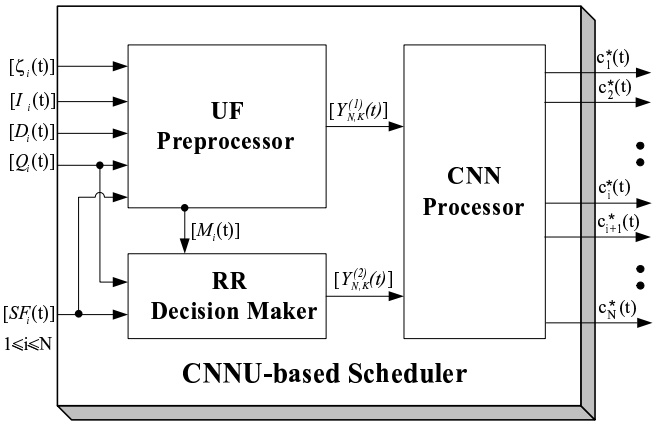

Fig. 1. The block diagram of CNNU-based scheduler.

IV. Design of the CNNU-based Scheduler

Fig. 1 shows the block diagram of the CNNU-based scheduler. It contains a *utility function (UF) preprocessor*, a *radio-resource range (RR) decision maker*, and a *CNN processor*. The proposed CNNU-based scheduler takes the link information, interference, delay, queue length, and spreading factor of all connections as inputs, and finally outputs an optimal normalized radio resource assignment vector $\vec{c}^{\,*}(t) = (c_i^*(t), \ldots, c_N^*(t))$, where $c_i^*(t),\ 1 \le i \le N$, is expressed by K bits.

The *UF preprocessor* first calculates the utility function $\mathcal{U}_i(t)$ given in (1) , $1 \le i \le N$. Then it normalizes $\mathcal{U}_i(t)$ by a compression function $(1 - e^{-\sigma \mathcal{U}_i(t)})$, expresses $(1 - e^{-\sigma \mathcal{U}_i(t)})$ to be an $1 \times K$ vector given by

$$\left((1 - e^{-\sigma \mathcal{U}_i(t)}) \cdot 2^{-1}, \ldots, (1 - e^{-\sigma \mathcal{U}_i(t)}) \cdot 2^{-k}, \ldots, (1 - e^{-\sigma \mathcal{U}_i(t)}) \cdot 2^{-K}\right),$$

and finally constructs an $N \times K$ input matrix $\left[Y_{i,k}^{(1)}\right]$ for the CNN processor, where $Y_{i,k}^{(1)} = (1 - e^{-\sigma \mathcal{U}_i(t)}) \cdot 2^{-k}$. Notice that σ is a constant related to the slope and the linear region of the compression function. The compression function $(1 - e^{-\sigma \mathcal{U}_i(t)})$ normalizes $U_i(t) \in [0, \infty)$ into the unit range of $[0, 1)$. A good compression function is the one with broad linear range so that the individual utility function is normalized linearly within a reasonable range. *The UF preprocessor* also determines a vector of modulation order $[M_{\kappa_i}]$ for all connections and outputs to the *RR decision maker*. The *RR decision maker* determines the upper limit for the radio resource assignment for every connection i and expresses it by an $1 \times K$ vector which is the

upper limit multiplied by the bit-weighted vector $(2^{-1},\ldots,2^{-k},\ldots,2^{-K})$, $1 \le i \le N$. Note that, with given spreading factor SF_i and queue length $Q_i(t)$ for connection i, its radio resource assignment $c_i(t)$ is upper limited by $min\left\{\frac{W\cdot\log_2 M_{\kappa_i}}{SF_i\cdot\mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)}\right\}$, which will be further discussed in section IV-B. Then the *RR decision maker* constructs the second input matrix $\left[Y_{i,k}^{(2)}\right]$, $1 \le i \le N$, $1 \le k \le K$, of which the element $Y_{i,k}^{(2)}$ is given by $\left(min\left\{\frac{W\cdot\log_2 M_{\kappa_i}}{SF_i\cdot\mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)}\right\}\right)\cdot 2^{-k}$. The *CNN processor* receives input matrices, $\left[Y_{i,k}^{(1)}\right]$ and $\left[Y_{i,k}^{(2)}\right]$, and determines the optimal radio resource assignment vector $\vec{c}^{\,*}(t)$. During the computation process, denote by τ the instantaneous time index of the CNN and by $\vec{c}(t,\tau)$ the instantaneous radio resource assignment vector at time τ during the frame t. For each $c_i(t,\tau)$, $1 \le i \le N$, it is represented by K bits, $X_{i,k}(\tau)$, $1 \le k \le K$, and $c_i(t,\tau)$ can be expressed by

$$c_i(t,\tau) \cong \sum_{k=1}^{K} X_{i,k}(\tau)\cdot 2^{-k}. \tag{8}$$

When the CNN processor arrives at an equilibrium, the output will converge to the optimal radio resource assignment vector, ie., $\lim_{\tau\to\infty}\vec{c}(t,\tau) = \vec{c}^{\,*}(t)$.

In the following, the design of the CNN processor for the CNNU-based scheduler is described. Characteristics of the original CNN proposed in [11] is first briefed, and a cost function corresponding to the system utility function with system constraints is formulated. A modified architecture for *CNN processor* is then presented, based on the Lyapunov method. The stability and convergence of the neural network is briefly discussed.

A. Preliminaries for Cellular Neural Networks

Consider a neural network with $N \times K$ neurons arranged in a rectangular array, where neuron (i,k) is denoted by $z_{i,k}$. The output of $z_{i,k}$ at time τ, denoted by $X_{i,k}(\tau)$, can be expressed by $X_{i,k}(\tau) = f\left(X_{i,k}^{(s)}(\tau)\right)$, where $f(x) = \frac{1}{2}\left[|x| - |x-1|\right] + \frac{1}{2}$ is an activation function of $z_{i,k}$ and $X_{i,k}^{(s)}(\tau)$ is the state variable of $z_{i,k}$ at time τ. $X_{i,k}^{(s)}(\tau)$ consists of the recurrent inputs, external inputs, and a bias current. For each neuron $z_{i,k}$, it connects with all other neurons within its neighborhood, denoted by $Z_n(i,k)$. The area of $Z_n(i,k)$ is determined according to the design of the neural network. Generally, the dynamics of the CNN at time τ is represented by

$$\begin{aligned}\frac{dX_{i,k}^{(s)}(\tau)}{d\tau} &= -\frac{X_{i,k}^{(s)}(\tau)}{\nu} + A_{i,k;i,k}\cdot X_{i,k}(\tau) + B_{i,k;i,k}\cdot Y_{i,k} \\ &\quad + \sum_{z_{j,m}\in Z_n(i,k)} A_{i,k;j,m}\cdot X_{j,m}(\tau) + \sum_{z_{j,m}\in Z_n(i,k)} B_{i,k;j,m}\cdot Y_{j,m} + V_{i,k},\end{aligned} \tag{9}$$

where ν is a time constant for all neurons, $A_{i,k;j,m}$ is the recurrent interconnection weight from neuron $z_{j,m}$ to $z_{i,k}$, $B_{i,k;j,m}$ is the control weight of external input from $z_{j,m}$ to $z_{i,k}$, $Y_{j,m}$ is the external input to the neuron $z_{j,m}$, and $V_{i,k}$ is the bias current to $z_{i,k}$, which is usually a fixed value V. It is worth mentioning that $A_{i,k;i,k} > 1/\nu$ is hold so that the neuron $z_{i,k}$ will eventually enter into a saturation region [11]. Also, the interconnection weights are assumed to be symmetric, that is, $A_{i,k;j,m} = A_{j,m;i,k}$, thus the CNN is stable [11].

An energy function at time τ which decreases along the trajectories of (9) is generally expressed by [11]

$$\begin{aligned} E(\tau) &= -\frac{1}{2}\sum_{i=1}^{N}\sum_{k=1}^{K} A_{i,k;i,k}X_{i,k}^2(\tau) - \frac{1}{2}\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{j=1}^{N}\sum_{m=1}^{K} A_{i,k;j,m}X_{j,m}(\tau)X_{i,k}(\tau) \\ &\quad -\frac{1}{2}\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{j=1}^{N}\sum_{m=1}^{K} B_{i,k;j,m}Y_{j,m}X_{i,k}(\tau) - \sum_{i=1}^{N}\sum_{k=1}^{K} V_{i,k}\cdot X_{i,k}(\tau). \end{aligned} \tag{10}$$

At the stable state, outputs of neurons will arrive at an equilibrium of that the energy function is minimized. If the energy function is properly designed and acts as a cost function, such an optimization problem can be solved via the Lyapunov method [12], [13]-[14]. By the Lyapunov method, the CNN can be designed with a set of prescribed trajectories. The trajectories are described by the gradient of the Lyapunov function $E(\tau)$ which is the energy of the CNN network at time τ. With an appropriate energy function designed according to the cost function, the minimization of the cost can be achieved along the designed trajectories. In the mean time, it can be proved that the architecture of the designed CNN can be related with the energy function by

$$\frac{dX_{i,k}^{(s)}(\tau)}{d\tau} = -\frac{X_{i,k}^{(s)}(\tau)}{\nu} - \frac{\partial E(\tau)}{\partial X_{i,k}(\tau)}. \tag{11}$$

Using (11), the desired system parameters of inter-connection weights, control weights, and bias currents can be found from the trajectories of energy function.

B. Cost Function for CNN Processor

The cost function of CNN [13] for frame t at time τ, denoted by $\mathcal{H}(t,\tau)$, consists of a cost function for the utility function, denoted by $\mathcal{H}_u(t,\tau)$, in conjunction with cost functions for system constraints Ψ_1 and Ψ_2, denoted by $\mathcal{H}_{\Psi_1}(t,\tau)$ and $\mathcal{H}_{\Psi_2}(t,\tau)$, respectively. The constraint $\Psi_1 = \{\vec{c}(t) : \sum_n c_n(t) \leq 1\}$ is because the system transmission power is limited by a

maximum power budget P^*_{max}. Notice that the assigned transmission rate $r_i(t)$ of connection i for frame t is determined according to both the $c_i(t)$ and the modulation order M_{κ_i}; the $r_i(t)$ is further limited by the minimum spreading factor SF_i and the waiting queue length $Q_i(t)$; and too large $c_i(t)$ with excess allocated power makes no effects on $r_i(t)$. Also the constraint $\Psi_2 = \left\{ \vec{c}(t) : c_i(t) \leq min \left\{ \frac{W \cdot \log_2 M_{\kappa_i}}{SF_i \cdot \mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)} \right\}, \forall i \right\}$, where $\mathcal{R}_i(t)$ is the radio resource function indicating the maximum achievable transmission rate for connection i at time t. The constraint Ψ_2 indicates no further utility can be gained if $r_i(t)$ exceeds the supported rate which is the rate when the power ratio $c_i(t)$ equals $(\frac{W \cdot \log_2 M_{\kappa_i}}{SF_i \cdot \mathcal{R}_i(t)})$, or the necessary rate to transmit all remaining packets in $Q_i(t)$, where the necessary rate is the rate when $c_i(t)$ equals $(\frac{Q_i(t)/T_f}{\mathcal{R}_i(t)})$. The $\mathcal{H}(t,\tau)$ has the form of

$$\mathcal{H}(t,\tau) \;\; = \;\; \mathcal{H}_u(t,\tau) + \mathcal{H}_{\Psi_1}(t,\tau) + \mathcal{H}_{\Psi_2}(t,\tau). \tag{12}$$

The $\mathcal{H}_u(t,\tau)$ is defined to be the difference between an overall normalized utility function and its maximum, and the overall normalized utility function is defined as $\sum_{i=1}^{N} c_i(t,\tau) \cdot (1 - e^{-\sigma \mathcal{U}_i(t)})$. When $\sum_{i=1}^{N} c_i(t,\tau) \leq 1$, $\sum_{i=1}^{N} c_i(t,\tau) \cdot (1 - e^{-\sigma \mathcal{U}_i(t)})$ is bounded by 1. Thus the $\mathcal{H}_u(t,\tau)$ is given by

$$\mathcal{H}_u(t,\tau) = \eta_0 \left[1 - \sum_{i=1}^{N} c_i(t,\tau) \left(1 - e^{-\sigma \mathcal{U}_i(t)} \right) \right], \tag{13}$$

where η_0 is the coefficient for the $\mathcal{H}_u(t,\tau)$.

The $\mathcal{H}_{\Psi_1}(t,\tau)$ is defined as

$$\mathcal{H}_{\Psi_1}(t,\tau) = \eta_1 \left[\sum_{i=1}^{N} c_i(t,\tau) - 1 \right]^2, \tag{14}$$

where $\eta_1 = \eta_1^+ \cdot u \left(\sum_{i=1}^{N} c_i(t,\tau) - 1 \right) + \eta_1^- \cdot \left(1 - u \left(\sum_{i=1}^{N} c_i(t,\tau) - 1 \right) \right)$, $u(\cdot)$ is the unit-step function, η_1^+ is the slope constant for the cost increment when the total radio resource is greater than the maximum, and η_1^- is the slope constant for the cost increment otherwise. The ranges of η_1^+ and η_1^- are further investigated in the next section to ensure the stability and the desired output pattern of the *CNN processor*.

The $\mathcal{H}_{\Psi_2}(t,\tau)$ is defined to be proportional to the difference $\left[c_i(t,\tau) - min \left\{ \frac{W \cdot \log_2 M_{\kappa_i}}{SF_i \cdot \mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)} \right\} \right]$ if $c_i(t,\tau) > min \left\{ \frac{W \cdot \log_2 M_{\kappa_i}}{SF_i \cdot \mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)} \right\}$; otherwise, no cost will be incurred because

the radio resource will be allocated to other connections for efficiency. It is given by

$$\mathcal{H}_{\Psi_2}(t,\tau) = \eta_2 \left[\sum_{i=1}^{N} \left(\left(c_i(t,\tau) - min \left\{ \frac{W \cdot \log_2 M_{\kappa_i}}{SF_i \cdot \mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)} \right\} \right)^+ \right)^2 \right], \tag{15}$$

where η_2 is the coefficient for the $\mathcal{H}_{\Psi_2}(t,\tau)$.

Consequently, the cost function $\mathcal{H}(t,\tau)$ is given by

$$\begin{aligned} \mathcal{H}(t,\tau) &= \eta_0 \left[1 - \sum_{i=1}^{N} c_i(t,\tau) \left(1 - e^{-\sigma \mathcal{U}_i(t)} \right) \right] + \eta_1 \cdot \left[\sum_{i=1}^{N} c_i(t,\tau) - 1 \right]^2 \\ &+ \eta_2 \left[\sum_{i=1}^{N} \left(\left(c_i(t,\tau) - min \left\{ \frac{W \cdot \log_2 M_{\kappa_i}}{SF_i \cdot \mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)} \right\} \right)^+ \right)^2 \right]. \end{aligned} \tag{16}$$

C. The Architecture of CNN Processor

According to the cost function $\mathcal{H}(t,\tau)$ at time t, the energy function $E(\tau)$ can be designed for the CNN processor in the CNNU-based scheduler. However, some modifications on $\mathcal{H}(t,\tau)$ for $E(\tau)$ should be made to ensure the correctness of the desired output and the stability of the CNN processor. The $E(\tau)$ is given by

$$\begin{aligned} E(\tau) &= -\eta_0 \left[\sum_{i=1}^{N} (\sum_{k=1}^{K} X_{i,k}(\tau) \cdot 2^{-k}) \cdot (1 - e^{-\sigma \cdot \mathcal{U}_i(t)}) \right] \\ &+\eta_1 \left[\left(\sum_{i=1}^{N} \frac{1}{2} (\sum_{k=1}^{K} X_{i,k}(\tau) \cdot 2^{-k}) - 1 \right) \cdot \left(\sum_{i=1}^{N} (\sum_{k=1}^{K} X_{i,k}(\tau) \cdot 2^{-k}) \right) \right] \\ &+\eta_2 \left[\sum_{i=1}^{N} \left(\left(\frac{1}{2} (\sum_{k=1}^{K} X_{i,k}(\tau) \cdot 2^{-k}) - min \left\{ \frac{W \cdot \log_2 M_{\kappa_i}}{SF_i \cdot \mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)} \right\} \right)^+ \cdot \right. \right. \\ &\left. \left. (\sum_{k=1}^{K} X_{i,k}(\tau) \cdot 2^{-k}) \right) \right] + \eta_3 \left[\sum_{i=1}^{N} \left(\sum_{k=1}^{K} X_{i,k}(\tau) \left(1 - X_{i,k}(\tau) \right) \cdot 2^{-k} \right) \right], \end{aligned} \tag{17}$$

where η_3 is a constant for additional auxiliary terms. The first item differs from the corresponding cost in (16) in that the scalar 1 is ignored. The $\sum_{i=1}^{N} (\sum_{k=1}^{K} X_{i,k}(\tau) \cdot 2^{-k}) \cdot (1 - e^{-\sigma \cdot \mathcal{U}_i(t)})$ is bounded above by 1 and has the same minimum as in the cost function. For the second and third terms, the quadratic forms in (16) are replaced by convex functions which merely contain state variable $X_{i,k}(\tau)$ without any scalar. The local minimums would be the same; the resulting energy at any equilibrium would be shifted by a constant value, compared to the cost in (16) and independent of the inputs and the output pattern. The last term of (17) is an auxiliary

factor which emerges to ensure the convergence because the energy function due to this auxiliary term approaches zero only when every state variable outputs approaches either one or zero.

By (11) and (17) , the dynamics of each neuron in the proposed CNN for the CNNU-based scheduler can be expressed by

$$\begin{aligned}
\frac{dX_{i,k}^{(s)}(\tau)}{d\tau} &= -\frac{X_{i,k}^{(s)}(\tau)}{\nu_k} + \eta_0\left(1-e^{-\sigma\mathcal{U}_i(t)}\right)\cdot 2^{-k} - \eta_1\left(\sum_{i=1}^{N}\sum_{k=1}^{K}X_{i,k}(\tau)\cdot 2^{-k}-1\right)\cdot 2^{-k} \\
&\quad -\eta_2\left(\sum_{m=1}^{K}X_{i,m}(\tau)\cdot 2^{-m} - min\left\{\frac{W\cdot\log_2 M_{\kappa_i}}{SF_i\cdot\mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)}\right\}\right)^{+}\cdot 2^{-k} \\
&\quad -\eta_3\left(1-2X_{i,k}(\tau)\right)\cdot 2^{-k} \\
&= -\frac{X_{i,k}^{(s)}(\tau)}{\nu_k} + \left[2\eta_3\cdot 2^{-k} - \eta_1\cdot 2^{-2k} - \eta_2\cdot u\left(\sum_{k=1}^{K}X_{i,k} - min\left\{\frac{W\cdot\log_2 M_{\kappa_i}}{SF_i\cdot\mathcal{R}_i(t)},\right.\right.\right. \\
&\quad \left.\left.\left.\frac{Q_i(t)/T_f}{\mathcal{R}_i(t)}\right\}\right)\cdot 2^{-2k}\right]\cdot X_{i,k}(\tau) + \eta_0\left(1-e^{-\sigma\mathcal{U}_i(t)}\right)\cdot 2^{-k} \\
&\quad +\eta_2\cdot u\left(\sum_{m=1}^{K}X_{i,m}(\tau)\cdot 2^{-m} - min\left\{\frac{W\cdot\log_2 M_{\kappa_i}}{SF_i\cdot\mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)}\right\}\right)\cdot \\
&\quad min\left\{\frac{W\cdot\log_2 M_{\kappa_i}}{SF_i\cdot\mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)}\right\}\cdot 2^{-k} - \left(\sum_{j=1,j\neq i}^{N}\sum_{m=1}^{K}\eta_1\cdot 2^{-(m+k)}\cdot X_{j,m}(\tau)\right) \\
&\quad -\left(\sum_{m=1,m\neq k}^{K}\left[\eta_1 - \eta_2\cdot u\left(\sum_{m=1}^{K}X_{i,m}(\tau)\cdot 2^{-m} - min\left\{\frac{W\cdot\log_2 M_{\kappa_i}}{SF_i\cdot\mathcal{R}_i(t)},\right.\right.\right.\right. \\
&\quad \left.\left.\left.\left.\frac{Q_i(t)/T_f}{\mathcal{R}_i(t)}\right\}\right)\right]\cdot 2^{-(m+k)}\cdot X_{i,m}(\tau)\right) + \eta_1\cdot 2^{-k} - \eta_3\cdot 2^{-k}, \qquad (18)
\end{aligned}$$

where ν_k is modified to 2^k to retain the stability and desired output pattern of the designed CNN.

From (9) and (18) , the recurrent interconnection weights, the external control weights, and the bias current can be determined by

$$\begin{cases}
A_{i,k;i;k} = -\eta_1\cdot 2^{-2k} - \eta_2\cdot u\left(\sum_{k=1}^{K}X_{i,k} - min\left\{\frac{W\cdot\log_2 M_{\kappa_i}}{SF_i\cdot\mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)}\right\}\right)\cdot 2^{-2k} + 2\eta_3\cdot 2^{-k}, \\
B_{i,k;i;k}^{1} = +\eta_0, \\
B_{i,k;i;k}^{2} = +\eta_2\cdot u\left(\sum_{k=1}^{K}X_{i,k} - min\left\{\frac{W\cdot\log_2 M_{\kappa_i}}{SF_i\cdot\mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)}\right\}\right), \\
A_{i,k;j;m} = -\eta_1\cdot 2^{-(k+m)}\cdot(1-\delta_{k,m})\delta_{i,j} - \eta_1\cdot 2^{-(k+m)}\cdot(1-\delta_{i,j}) \\
\qquad\qquad -\eta_2\cdot u\left(\sum_{k=1}^{K}X_{i,k} - min\left\{\frac{W\cdot\log_2 M_{\kappa_i}}{SF_i\cdot\mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)}\right\}\right)\cdot\delta_{i,j}\cdot 2^{-(k+m)}, \\
V_{i,k} = \eta_1\cdot 2^{-k} - \eta_3\cdot 2^{-k},
\end{cases} \qquad (19)$$

where $B^1_{i,k;i;k}$ and $B^2_{i,k;i;k}$ are the external control weights for first and second external inputs, $Y^{(1)}_{i,k} = \left(1 - e^{-\sigma \mathcal{U}_i(t)}\right) \cdot 2^{-k}$ and $Y^{(2)}_{i,k} = min\left\{\frac{W \cdot \log_2 M_{\kappa_i}}{SF_i \cdot \mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)}\right\} \cdot 2^{-k}$, respectively, and $\delta_{x,y} = 1$ if $x = y$, $\delta_{x,y} = 0$ otherwise.

The range of coefficients η_0, η_1 (η_1^+, η_1^-), η_2, and η_3 must be properly selected to ensure the stability and the desired response. For a tolerant error level ε, which is the maximum difference between stable output $\lim_{\tau \to \infty} \vec{c}(t, \tau)$ and the optimum $\vec{c}^*(t)$, the range of these coefficients are obtained as following : $0 < \eta_0 < \eta_3$, $\eta_1^+ > 2^K$, $\eta_1^- \geq \frac{\eta_0 \cdot 2^{-3}}{\varepsilon}$, $\eta_2 \geq 2^K$, $\eta_3 > \frac{1}{2} + \frac{\eta_1^-}{2}$.

We have proved that with a matrix of given utility function and a matrix of radio resource assignment ratio upper limits, the proposed CNN architecture will converge to the neighborhood of the optimal pattern $\vec{c}^*(t)$ within the difference ε, with the range of these coefficients. If $\varepsilon \leq 2^{-K}$, the CNN converges to $\vec{c}^*(t)$ exactly .

However, the complexity of the interconnection is in the order of $[N \times K]^{N \times K}$, which is almost infeasible for practical implementation. We propose an equivalent two-layer structure for the CNN processor in the next subsection to efficiently reduce the complexity of interconnection.

D. The Two-Layer Structure for CNN Processor

Fig. 2 shows the architecture of the two-layered CNN processor. The equivalent two-layer structure for the CNN processor involves the first decision layer, $\left[z^1_{i,k}\right]$, with state variable output $X_{i,k}(\tau)$, and the second output layer, $\left[z^2_{i,k}\right]$, with state variable output $c_i(t, \tau)$. The decision layer consists of $N \times K$ neurons; the output layer is with an $(N + 1) \times 1$ array, where the output of the first neuron is the summation of all the others. The interconnections between the neurons of decision layer and those of output layer are defined by

- For the first decision layer to the second output layer, the connection weight between $X_{i,k}(\tau)$ and $c_j(t, \tau)$ is 2^{-k}, $\forall k$ if $j = i$; is zero if $j \neq i$.
- For the second output layer feedback to the first decision layer, only the first neuron output is connected to the $X_{i,k}(\tau)$ of the decision layer with the interconnection weight $\eta_1 \cdot 2^{-k}$ for $\forall i$.

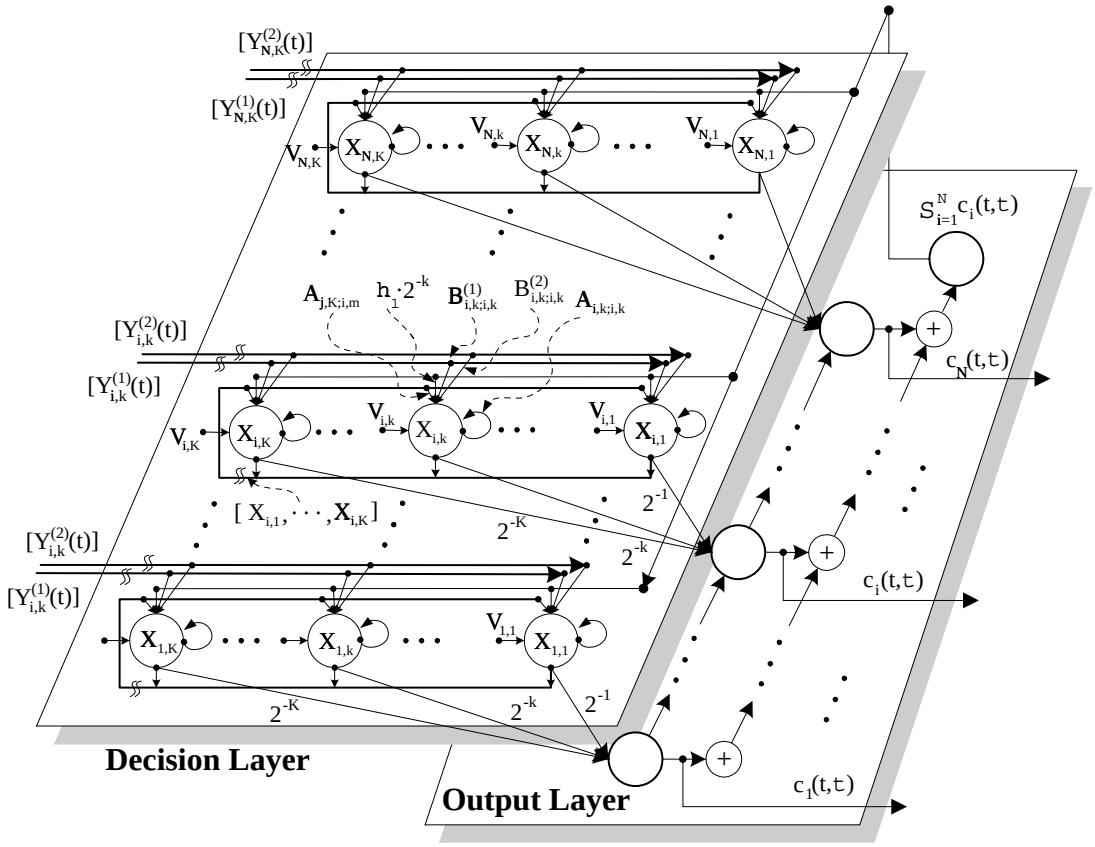

Fig. 2. The two-layer structure of CNN processor.

The recurrent interconnection weights and the external control weights for the first decision layer defined in (19) are then modified to be

$$\begin{cases} A_{i,k;i;k} = -\eta_1 \cdot 2^{-2k} - \eta_2 \cdot u\left(\sum_{k=1}^{K} X_{i,k} - min\left\{\frac{W\cdot\log_2 M_{\kappa_i}}{SF_i\cdot\mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)}\right\}\right)\cdot 2^{-2k} + 2\eta_3, \\ B^1_{i,k;i;k} = \eta_0, \\ B^2_{i,k;i;k} = \eta_2 \cdot u\left(\sum_{k=1}^{K} X_{i,k} - min\left\{\frac{W\cdot\log_2 M_{\kappa_i}}{SF_i\cdot\mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)}\right\}\right), \\ A_{i,k;j;m} = -\eta_2 \cdot u\left(\sum_{k=1}^{K} X_{i,k} - min\left\{\frac{W\cdot\log_2 M_{\kappa_i}}{SF_i\cdot\mathcal{R}_i(t)}, \frac{Q_i(t)/T_f}{\mathcal{R}_i(t)}\right\}\right)\cdot \delta_{i,j}\cdot 2^{-(k+m)}, \\ V_{i,k} = \eta_1 \cdot 2^{-k} - \eta_3, \end{cases} \tag{20}$$

For the second output layer, there are no external inputs, and only recurrent interconnection weights exist. The interconnection weight between $c_i(t,\tau)$ and $c_j(t,\tau)$ is given by $\delta_{0,j}$ with $i = 0$.

It can be shown that the two-layer structured CNN processor has the same energy function and the local minimum as the single-layer one defined by (19) . However, the complexity of

interconnections in the two-layer one is proportional to $[3N \times K + N]$, which is significantly lower than $[N \times K]^{N \times K}$ in the single-layer one.

V. Simulation Results and Discussion

In the simulations, a scenario with five types of services in three classes is assumed. Type-1 service is a real-time class of traffic with peak rate 15kbps, activity factor 0.57, $P_D^* = 0.05$, $D^* = 40$ms, and $BER^* = 10^{-3}$. Type-2 (type-3) service is a non-real-time interactive class of traffic with Pareto process [19] of which the mean rate is 8kbps (12kbps), $R_{m,i}^*$=7.2kbps ($R_{m,i}^*$=11kbps), and $BER^* = 10^{-5}$ ($BER^* = 10^{-5}$). And type-4 (type-5) service is a non-real-time best effort class of traffic in batch Poisson distribution with mean rate 6kbps (15kbps) and mean batch size 1.2k bits (1.2k bits), and $BER^* = 10^{-5}$. The proportion in the number of connections from type-1 to type-5 is kept at 1:1:1:1:1. Also, four modulation schemes, BPSK, QPSK, 16QAM, and 64QAM, are available for transmission as long as the BER requirement can be fulfilled and the remaining queue is enough.

We compare the proposed CNNU-based scheduler with the *exponential rule* (EXP) scheduling scheme. The performance measures are such as the average system throughput, the average packet dropping ratio of RT connections, $\overline{P}_D$, the average transmissio rate of NRT interactive connections, $\overline{R}_m$, the ratio of RT connections in which their packet dropping ratio requirement is not guaranteed, ϕ_{P_D}, the ratio of NRT interactive connections in which their minimum transmission rate requirement is not guaranteed, ϕ_{R_m}, and the fairness variance index of NRT connections, F_v. The F_v is defined for measuring the variance of fairness to share the radio resource among all NRT connections. It is given by

$$F_v = \frac{1}{N_{NRT}} \sum_{i}^{N_{NRT}} \left| \frac{E\left[r_i(t)\right]}{\sum_j^{N_{NRT}} E\left[r_j(t)\right]} - \frac{w_i}{\sum_j^{N_{NRT}} w_j} \right|^2,$$

where N_{NRT} is the number of NRT connections. The fairness variance index shows the variance of the normalized radio resource allocated and the normalized proportion of resource desired to share.

Fig. 3 shows the average system throughput. It can be found that the CNNU-based scheduler can always have a higher system capacity than the EXP scheduling scheme in all traffic load conditions; it achieves the improvement of the system throughput over the EXP scheduling scheme by more than 9% as the number of connections is greater than 200, and by higher

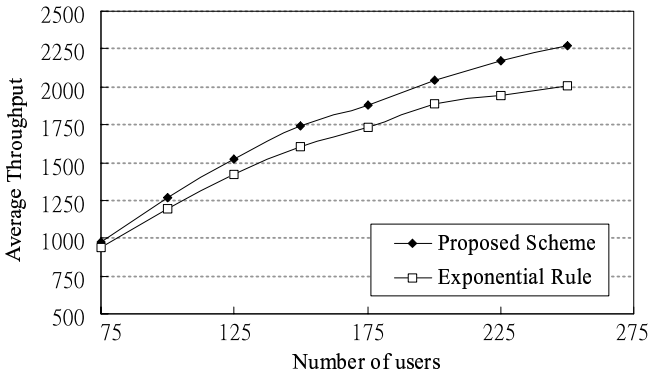

Fig. 3. The average system throughput

than 15% as the number of connections increases up to 250. This is because the radio resource funciton makes CNN processor adapt to the link variation and allocate radio resource in an efficient way. Both RT and NRT connections with relatively worse link conditions have lower probability to be scheduled as long as their QoS requirements can be achieved in a long term sense. The fairness compensation function makes the NRT connections share the radio resource according to the location dependent fairness and thus a higher radio resource efficiency can be achieved. Also, the CNN processor can determine an optimal radio resource assignment vector in the sense that the allocation of downlink power by CNNU-based scheduler is the most efficient one, with given utilities and upper limits of the radio resource assignment. Additionally, beyond the point of 250 connections, the throughput of the EXP scheme is almost saturated, while the throughput of the CNNU-based scheduler continues to grow up but with a slightly lowering slope. It is because the CNNU-based scheduler can achieve utilization of multiuser diversity gain better than the EXP scheduling scheme.

Fig. 4 depicts performance measures of the average packet dropping ratio of type-1 RT connections $\overline{P}_D$ and the average transmission rate of type-2 and type-3 NRT interaction connections $\overline{R}_m$. It can be found that the $\overline{P}_D$ of the CNNU-based scheduler is larger than that of the EXP scheme and it violates the P_D^* requirement as the number of users is at about 250; on the other hand, all the $\overline{R}_m$ of type-2 and type-3 connections of the CNNU-based scheduler are greater than that of the EXP scheme as the number of connections is greater than 125 but the EXP scheme violates the R_m^* requirements as the number of users is at about 170. These indicate that the QoS

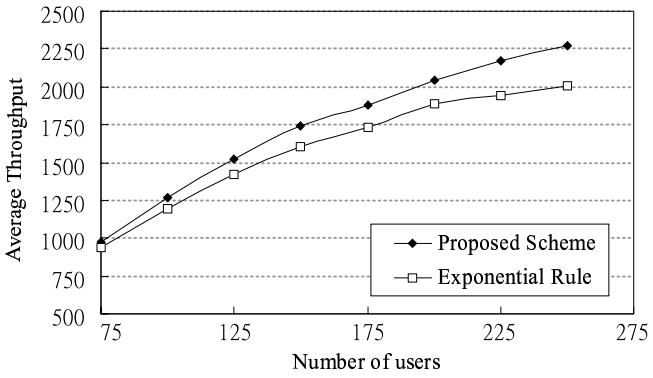

Fig. 4. QoS performance measures of $\overline{P}_D$ and $\overline{R}_m$

guaranteed region by the CNNU-based scheduler is as the number of connections is less than 250, while that by the EXP scheme is as the number of connections is less than 175. The QoS quaranteed region achieved by the CNNU-based scheduler is larger than that given by the EXP scheme. This is beacause the QoS deviation function together with the priority bias designed in the CNNU-based scheduler can balance the extent of deviation of every performance measure from the QoS requirement. The worse the QoS performance measure is, the more the radio resource will be scheduled. Besides, since the CNNU-based scheduler has higher throughput performance, the more number of connections can be served in the QoS region. Moreover, if we define the maximum throughput achievable in QoS guaranteed region to be the average system throughput, the CNNU-based scheduler can have the average system throughput equal to 2160Mbps at 245 connections, while the EXP scheme can have the average system throughput equal to 1600Mbps at 175 connections. The former attains the average system throughput greater than the latter by an amount of 25%.

Fig. 5 shows the ratio of RT connections of which the packet dropping ratio requirement is not graranteed, ϕ_{P_D}, and the ratio of NRT interactive connections of which the minimum transmission rate requirement is not guaranteed, ϕ_{R_m}. It can be seen that the total ratio of connections with QoS requirements un-guaranteed for the CNNU-based scheduler is about 0.0435, while that for the EXP scheme is greater than 0.18 in heavy loaded situations as the number of connections is greater than 225. The total ratio of connections with QoS requirements un-guaranteed is here defined as $\frac{1}{3}\phi_{P_D} + \frac{2}{3}\phi_{R_M}$, which is weighted by the number of RT and

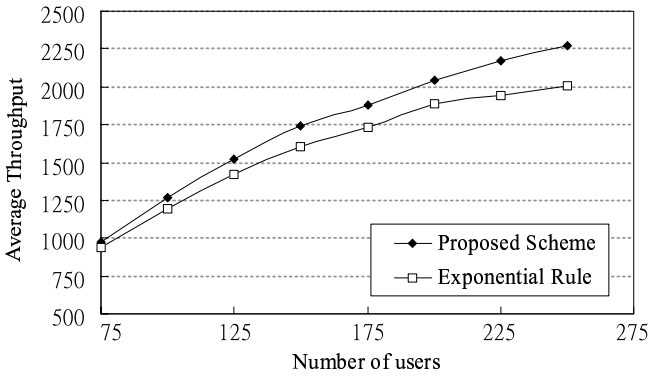

Fig. 5. The ratio ϕ_{P_D} for RT connections and the ratio ϕ_{R_m} for NRT interactive connections

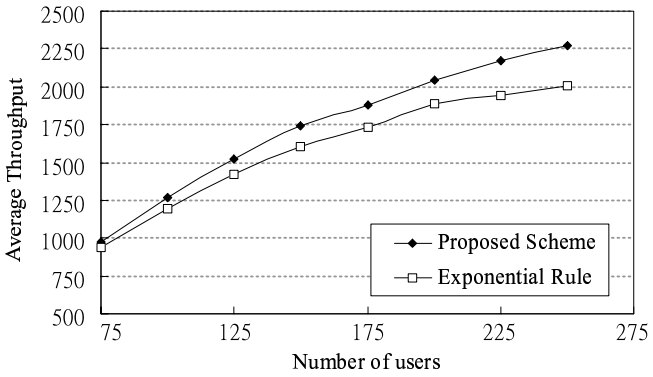

Fig. 6. The fairness variation index for NRT connections

NRT interactive connections. This indicates that CNNU-based scheduler can achieve lower total ratio of connections in all traffic types of which the corresponding QoS requirements are not guaranteed than the EXP scheme does. The reason is that the CNNU-based scheduler can balance the allocation of radio resources among traffic types and avoid allocating excess radio resource to connections with bad link condition, while the EXP scheme prefers RT connections and overprotects them so that the QoS guaranteed region is reduced. Note that the ratios of ϕ_{P_D} and ϕ_{R_m} are greater than zero at any traffic load conditions due to the existence of connections with very bad link quality. These results imply that the CNNU-based scheduler will not guarantee all the QoS requirements all the time, and a properly designed call admission control is required to reject the connections with very bad link quality in terms of the current traffic load conditions.

Fig. 6 shows the fairness variance index of NRT connections. It can be found that the fairness

variance index of the CNNU-based scheduler retains within 1 in almost all simulation cases, and grows up slowly as the number of connections increases; the fairness variance index of the EXP scheme, on the other hand, increases with slightly higher slope compared to CNNU-based scheduler. This is because the fairness compensation function of the CNNU-based scheduler considers the location dependent information and aims to share the radio resource fairly as long as the minimum rate is guaranteed, while the design of the EXP scheme ignores the location dependent information and allocates rate fairly to all connections. The fairness compensation function, considering location dependent information, also facilitates the higher capacity for the CNNU-based scheduler shown in Fig. 3.

REFERENCES

[1] 3rd Generation Partnership Project Jun. 2001, QoS Concept and Architecture, *3GPP TS 23.107*, [Online] http://www.3gpp.org.

[2] A. Parekh and R. G. Gallager, "A generalized processor sharing approach to flow control in integrated services networks: the single node case," *IEEE/ACM Trans. Networking*, vol. 1, no. 3, pp. 344-357, JUN. 1993.

[3] D. Ferrari and D. Verma, "A scheme for real-time channel establishment in wide-area networks," *IEEE J. Select. Area Commun.*, vol. 8, no. 3, pp. 368-379, Apr. 1990.

[4] Y. M. Lu and R. W. Brodersen, "Integrating power control, error correction coding, and scheduling for a CDMA downlink system," *IEEE J. Select. Area Commun.*, vol. 17, no. 5, pp. 978-989, May 1999.

[5] V. K. N. Lau, and Y. K. Kwok, "On generalized optimal scheduling of high data-rate bursts in CDMA systems," *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 261-266, Feb 2003.

[6] V. Bharghavan, S. W. Lu, and T. Nandagopal, "Fair queuing in wireless networks: issues and approaches," *IEEE Personal Commun.*, vol. 6, no.1, pp. 44-53, Feb. 1999.

[7] A. C. Varsou and H. V. Poor, "HOLPRO: a new rate scheduling algorithm for the downlink of CDMA networks," *Proc. IEEE VTC'2000*, pp. 948-954.

[8] A. L. Stolyar and K. Ramanan, "Largest weighted delay first scheduling: large deviations and optimality," *Annal Appl. Prob.*, vol. 11, no. 1, pp. 1-48, 2001.

[9] A. C. Kam, T. Minn, and K. Y. Siu, "Supporting rate guarantee and fair access for bursty data traffic in W-CDMA," *IEEE J. Select. Area Commun.*, vol. 19, no. 11, pp. 2121-2130, 2001.

[10] S. Shakkottai and A. L. Stolyar, "Scheduling algorithms for a mixture of real-time and non-real-time data in HDR," *Bell Lab Reports*, 2000.

[11] L. O. Chua and L. Yang, "Cellular neural networks: Theory," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 1257-1272, Oct. 1988.

[12] J. A. Nossek, "Design and learning with cellular neural networks," *IEEE Proc. CNNA-1994,* Rome, pp. 137-146, Dec. 1994.

[13] "On sovling constrained optimization problems with neural networks: a penalty method approach," *IEEE Trans. Neural Net.,* vol. 4, no. 6, pp. 931-940, Nov. 1993.

[14] "Energy function based dynamic programming", *IEEE Trans. Neural Net.*, vol. no. pp., 1991.

[15] A. J. Goldsmith and S. G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218-1230, Oct. 1997.

[16] C. S. Chang and J. A. Thomas, "Effective Bandwidth in High-Speed Digital Networks," *IEEE J. Select. Area Commun.,* vol. 13, no. 6, AUG. 1995.

[17] L. S. Chou and C. S. Chang, "Experiments of the theory of effective bandwidth for markov sources and video traces," *Proc. IEEE INFOCOM'96*, pp.

[18] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, "Providing quality of service over a shared wireless link," *IEEE Commun. Mag.*, vol. 39, no. 2 , pp. 150-154, Feb. 2001.

[19] Universal Mobile Telecommunications System (UMTS), Selection procedures for the choice of radio transmission technologies of the UMTS, *UMTS 30.03*, version 3.2.0, 1998.

Chapter 4

A Novel Dynamic Cell Configuration Scheme in Next-Generation Situation-Aware CDMA Networks

I. Introduction

The growing popularity of multimedia Internet applications is a strong driving force for future cellular mobile systems. Due to user mobility and wide range of applications, the traffic pattern of each cell can vary dynamically. Thus, the current practice of engineering cell coverage and capacity based on pre-defined traffic patterns before a code division multiple access (CDMA) cellular network is deployed may lead to poor utilization of radio resources.

Due to asymmetric traffic and the interdependence of traffic capacity and coverage, this problem could be exacerbated in next-generation CDMA cellular networks, especially over the capacity-limited downlink [1]−[4].

To adapt to the variations of traffic load, tradeoffs between coverage and capacity in CDMA cellular systems have been considered [3]−[7]. For example, to guarantee the coverage of a cell, more power is used to reach mobile stations (MSs) near cell boundaries under power control. However, in interference-limited systems, the resulting higher inter-cell interference will reduce the system capacity significantly. Furthermore, under large traffic variations, power control may not be effective [3]−[5]. A uniform network layout with equal-sized cells, while optimal under uniform traffic, suffers significant capacity degradations if traffic loads are not balanced among all the cells [6]. To accommodate traffic load variations between cells, it is crucial for next-generation CDMA cellular networks to be aware of system situations and configure cell coverage and capacity dynamically [1], [7].

Several schemes for dynamic cell configuration (DCC) have recently been proposed [8]−[16]. Optimization of pilot power, and downlink capacity and coverage planning were considered in [8], [9]. In [10], a DCC scheme for circuit-switched micro-cellular CDMA systems was proposed to enhance the uplink performance. In [11], the competitive characteristics of network coverage and capacity were analyzed for a simple network. Only one class of service was considered in [8]−[11], and it may be difficult to extend these schemes to multiple classes of service. Some techniques based on heuristics have also been proposed for dynamic pilot power allocation (DPPA) to balance downlink traffic load while assuring service coverage [12], [13]. However, these schemes may cause “coverage failure regions” between cells where pilot signals are too weak to serve a MS [14], [15]. Moreover, a common shortcoming of the previous work [8]−[15] is that only pilot power is adjusted dynamically in the time-varying environment, without adjusting other parameters critical to radio resource management (RRM).

In fact, pilot power allocation and other RRM parameters are tightly coupled. In our previous work [16], we have shown that system performance can be improved significantly by a self-organized DCC scheme with coordinated call admission control (CAC), compared to fixed pilot power allocation (FPPA) and DPPA without taking CAC into account. Other work has shown that signal quality degradation can be prevented by configuring cell areas adaptively and setting power levels appropriately [4], [17], and soft handoff has significant impacts on the system

capacity and cell coverage [18], [19]. Therefore, an effective mechanism, link proportional power allocation (LPPA), was proposed for downlink soft handoff in [20], [21]. It was shown that LPPA can enhance system capacity in CDMA cellular systems with mixed-size cells, compared to conventional site-selection diversity transmissions (SSDT) scheme [22].

In this chapter, we show that DPPA without changing other related RRM parameters accordingly can result in performance degradations. To address this problem, we propose a novel DCC scheme based on reinforcement-learning called DCC-RL. The novelties are as follows. 1) DPPA is linked with soft handoff power and maximum link power allocations as well as CAC mechanisms. 2) Reinforcement-learning efficiently tackles optimization problems with large state spaces and action sets [23] in realistic CDMA multimedia cellular networks, which were previously deemed intractable [24]. 3) Our method does not require a priori knowledge of the state transition probabilities associated with the cellular network, which are very difficult to estimate in practice due to the varied propagation environment, diverse multimedia services, and random user mobility. 4) DCC-RL can be implemented in a distributed manner in each base station (BS), minimizing signaling overhead between BSs and radio network controllers, and the number of system states involved in computations.

We compare DCC-RL with fixed cell configuration (FIX) employing FPPA, and DPPA without changing other RRM parameters. Simulation results show that DCC-RL outperforms the others by increasing the total throughput, decreasing the frame error probability, blocking probability, and handoff forced termination probability with the price of slightly increasing the size of the active set.

The rest of this chapter is organized as follows. DCC issues are discussed in Section II. Section III describes the system model. Section IV formulates the DCC problem taking into account of RRM, and presents the proposed DCC-RL scheme. Simulation results are presented and discussed in Section V.

II. Issues of Dynamic Cell Configuration

A. Effects of Pilot Power Allocation Schemes

Since each BS has a finite transmit power, the pilot and traffic channels have to share the total power. Pilot power allocation can be either fixed or dynamic. In FPPA schemes, which is used in current CDMA systems, about 10-15% of the total power is allocated to the common

pilot channel and is not changed after the deployment of a cellular network, as shown in Fig. 1(a). When the traffic load is too high to allow allocation of sufficient power for all MSs, the system performance can degrade severely. Some strategies have to be employed to balance power between cells, e.g., by DPPA as illustrated in Fig. 1(b). The pilot power can be adjusted between the maximum and minimum constraints based on various traffic situations. When traffic is light, the pilot power can be increased to extend cell coverage to more MSs. On the other hand, when traffic is heavy and there is insufficient power for allocation to all traffic channels, the pilot power can be decreased to shrink cell coverage. This explains the interdependence of coverage and capacity in CDMA cellular systems.

Moreover, in future CDMA networks, diverse multimedia traffic and random user mobility will make preplanning of coverage and capacity difficult to manage. To achieve load balance whenever traffic congestion occurs, DCC through DPPA will be necessary.

B. Effects of Soft Handoff Power Allocation Schemes

The soft handoff mechanism can provide seamless connections and better signal qualities for MSs near the cell boundaries. Since the limited power available for traffic channels in each BS is shared between non-handoff and soft handoff MSs, there are tradeoffs between coverage and capacity. For example, a BS may shrink the cell coverage to serve less handoff MSs near the cell boundary, leaving more power available for allocation to non-handoff MSs with higher transmission rates. As soft handoff mechanisms have direct impacts on cell coverage and capacity, RRM for soft handoff MSs is a challenging issue in CDMA cellular systems with *mixed-size* cells formed by different levels of pilot power [20], [21].

C. Effects of New/Handoff Call Admission Control

In downlink CDMA systems, since each BS has finite power resource to be shared among MSs, the allocated pilot power and traffic channel power are directly related to the coverage and capacity of the cell. To achieve load balance whenever traffic congestion occurs, DCC through DPPA is necessary. However, it is necessary to consider the pilot power allocation and strategies of new/handoff CAC jointly in order to design an effective DCC scheme that improves the system performance while minimizing the undesirable effects.

For new call arrivals near cell boundaries, the pilot power determines the MSs initial access cells. Therefore, reducing the pilot power of a congested cell causes the MS to request a traffic channel from an adjacent cell. If the MS fails initially to detect a BS with enough signal strength, it cannot make a call request to the system. This is referred as a *coverage failure*. As a consequence, although the new call blocking probability of the congested cell could be decreased, the coverage failure probability might be increased.

For ongoing calls near cell boundaries, decreasing or increasing the pilot power of a BS can force some of the MSs to handoff into other cell(s) or vice versa. Therefore, the average size of the active set and handoff rates would be increased. In addition, if a MS suffers a bad signal quality and fails to to execute the handoff in time, a handoff forced termination occurs.

III. System Model

The system block diagram of our proposed DCC-RL scheme is shown in Fig. 2. DCC-RL can be implemented in a distributed manner in each BS, which adjusts its pilot power periodically to adapt to the variations of system situation through the dynamic pilot power controller. Based on the determined pilot power level, the maximum link power constraint and CAC criterion are adjusted accordingly. Then, the traffic channel power allocator adjusts its maximum link power constraint that is obtained from the maximum link power estimator. After applying all updates for RRM to the entire cellular network, the reinforcement signal is input to the dynamic pilot power controller to aid its decision for the next pilot power level. In this section, we describe the signal model and the link budget model in CDMA systems. An initial cell coverage design for the CDMA cellular system is provided to illustrate the interrelation between capacity and cell coverage.

A. Signal Model

Assume the total allocated power of BS b is P_b, including pilot channel power P_b^I and traffic channel power P_b^T, where P_b is smaller than or equal to the BS's maximum transmit power $\widetilde{P}_b$. The pilot power of BS b is given by $P_b^I = f_b \times \widetilde{P}_b$, where $f_b \in [f_{min}, f_{max}]$ is the fraction of the pilot power relative to BS b's maximum transmit power, constrained between minimum fraction f_{min} and maximum fraction f_{max}. For the traffic channel of MS m served by BS b, the allocated transmit power from BS b is $p_{b,m} = \phi_m P_b^T \leq \widetilde{p}_b$, where ϕ_m is the fraction of traffic

channel power allocated for transmission to MS m; $\widetilde{p_b}$ is the maximum link power of BS b. Thus, $\sum_{m \in \mathrm{U}_b} \phi_m = 1$, where the U_b represents the set of all MSs served by BS b.

B. Initial Cell Coverage Design

The initial design of cell coverage can be obtained by link budget analysis. The equivalent isotropic radiated power (EIRP) at a BS's transmitter, E_P, of each traffic channel can be calculated by $E_{P[\mathrm{dBm}]} = \widetilde{p_b}_{[\mathrm{dBm}]} + G_{B[\mathrm{dBi}]} - L_{C[\mathrm{dB}]}$, where G_B and L_C are the antenna gain and cable loss of the BS, respectively. Note that the units of the parameters are given in brackets [1]. On the other hand, the EIRP, E_T, measured at a MS's receiver, taking into account the soft gain G_S, the antenna gain G_M, and the body loss L_D of the MS, is $E_{T[\mathrm{dBm}]} = E_{P[\mathrm{dBm}]} + G_{M[\mathrm{dB}]} - L_{D[\mathrm{dB}]} + G_{S[\mathrm{dB}]}$. Moreover, assume that the interference margin (maximum planned noise rise) is Ω_I, and the received noise power (product of thermal noise density, chip rate, and noise figure) is η_o. The receiver sensitivity of the MS given service rate r is $H_R(r)_{[\mathrm{dB}]} = H_S(r)_{[\mathrm{dB}]} + \Omega_{I[\mathrm{dB}]} + \eta_{o[\mathrm{dBm}]}$, where $H_S(r)_{[\mathrm{dB}]}$ is the required signal-to-interference-plus-noise (SINR) value for service rate r, which is equal to the required bit-energy-to-noise ratio (Eb/No), $\gamma^*(r)_{[\mathrm{dB}]}$, minus the processing gain $G_P(r)_{[\mathrm{dB}]}$. From the link budget, the maximum allowable path loss for service rate r is

$$PL(r)_{[\mathrm{dB}]} = E_{T[\mathrm{dBm}]} - H_R(r)_{[\mathrm{dB}]} - \Omega_{L[\mathrm{dB}]}, \tag{1}$$

where Ω_L is the margin for log-normal fading. When a MS is near the cell boundary, the received chip-energy-to-interference ratio, E_c/I_o, should not fall below the minimum requirement $\Upsilon(r)$ for service rate r, given by $\Upsilon(r)_{[\mathrm{dB}]} = P_b^I{}_{[\mathrm{dBm}]} - PL(r)_{[\mathrm{dB}]} - \Omega_{I[\mathrm{dB}]} - \eta_{o[\mathrm{dBm}]} \geq \widetilde{\Upsilon}$. In general, pilot power P_b^I is around 1-4 watts, which is about $5\% - 20\%$ of the maximum total transmit power of the BS, $\widetilde{P_b}$.

Based on the allowable maximum path loss and the applied channel model, the resultant cell radius $R(r)$ is different with different service rate r. For $r \in [\, r_{min},\, r_{max}\,]$, since $H_S(r_{min}) < H_S(r_{max})$, therefore $PL(r_{min}) > PL(r_{max})$ and $R(r_{min}) > R(r_{max})$. This phenomenon raises the issue of fairness for different service rates in terms of service coverage and transmit power. If the same transmit power is allocated to MSs with different service rates, the higher service rate results in a smaller service coverage. Alternately, in order to support the same service

[1] In this chapter, a variable is linear if its unit is not specified.

coverage for different service rates, more transmit power is needed to support MSs with higher service rates near cell boundaries. Note that since total downlink transmit power of each BS is limited, system capacity is directly related to transmit power management. Based on the above concerns, in order to optimize system capacity, cell radius can be determined in terms of a suitable *reference service rate* r^*, where $r^* \in [r_{min}, r_{max}]$. The corresponding cell radius $R(r^*)$ is determined by the maximum allowable path loss $PL(r^*)$. Therefore, the required E_c/I_o of the system, $\widetilde{\Upsilon}$, is equal to

$$\widetilde{\Upsilon}_{[\text{dB}]} = P_{b\,[\text{dBm}]}^{I} - PL(r^*)_{[\text{dB}]} - \Omega_{I\,[\text{dB}]} + \eta_{o\,[\text{dBm}]}, \tag{2}$$

where $\widetilde{\Upsilon}$ is within the range from -16 [dB] to -20 [dB].

IV. Proposed DCC-RL Scheme

We formulate the DCC problem as a Markov decision process (MDP) [26]. However, traditional model-based solutions of MDP, such as policy iteration and linear programming, require a prior knowledge of the state transition probabilities. Due to the diverse multimedia traffic and random user mobility, these conventional solutions suffer from the *curses of dimensionality and modeling*. As described below, we propose a novel reinforcement-learning-based DCC scheme, DCC-RL, to find an optimal policy for pilot power allocation that takes RRM into account (see Fig. 2).

A. Problem Formulation as a Markov Decision Process

In DCC-RL, the BS pilot power is periodically adjusted to adapt to changing conditions. These time instants are called *decision epochs* and the adjustments of pilot power are called *actions* in the MDP formulation. The chosen action is based on the current *state* of the system. Depending on the action taken by the system, the system can earn *rewards*. The objective is to optimize the sequence of actions to maximize the accumulated rewards. The detailed formulation is as follows:

[Decision epochs]: In CDMA systems, the pilot signal is broadcasted from each BS periodically [30]. The state of the system changes accordingly. Therefore, we adjust the pilot power every M frames, where M is a design parameter.

[States]: Define the state vector of the system as $s = (\omega_M, \omega_V) \in \mathbb{R}_+^2$, where ω_M denotes the mean power of the BS and ω_V denotes the variance of the power load. Assume there are N samples from the measurements, where N is also a design parameter. Also, ω_M and ω_V can be obtained from the sample mean and variance, respectively, as follows.

$$\omega_M = \frac{1}{N}\sum_{n=1}^{N} P_b^T(n), \tag{3}$$

$$\omega_V = \frac{1}{N-1}\sum_{n=1}^{N}(P_b^T(n) - \omega_M)^2. \tag{4}$$

The decision process can be implemented in each BS in a distributed manner because the variation of the BS's power load can implicitly reveal the load information about all cells.

[Actions]: At each decision epoch, the BS makes a decision to choose a suitable fraction of the pilot power based on state s. The action $a(s) \in A$ of BS b is defined as the fraction, $f_b \in [f_{min}, f_{max}]$, of the pilot power relative to the maximum transmit power.

[Rewards function]: Based on the action $a(s)$ in a state s, the system earns a reward $\varphi(s, a(s))$. We choose the total throughput as the reward:

$$\varphi(s, a(s)) = \sum_{m} r_m, \tag{5}$$

where $r_m \in [r_{min}, r_{max}]$ is the transmission rate of MS m.

B. MDP Solution by Reinforcement-Learning

The objective of the decision process is to find an optimal policy π^* for each state s, which minimizes the cumulative measure of the reward $\varphi_t = \varphi(s_t, a(s_t))$ that is received over time, where the subscript represents the time instant t. The total expected discounted reward over an infinite time horizon can be represented by the value function with policy π, $V^\pi(s) = E\left\{\sum_{t=0}^{\infty} \lambda^t \cdot \varphi(s_t, \pi(s_t)) | s_0 = s\right\}$, with discount factor $0 \leq \lambda < 1$. Let $P(s'|s, a(s))$ be the transition probability from state s to s'. The value function can be rewritten as

$$V^\pi(s) = U(s, \pi(s)) + \lambda \sum_{s' \in S} \Pr(s'|s, \pi(s)) V^\pi(s'), \tag{6}$$

where $U(s, \pi(s)) = E\left\{\varphi(s, \pi(s))\right\}$. Define a Q-function of state-action pair with policy π as $Q^\pi(s, a(s)) = U(s, a(s)) + \lambda \sum_{s' \in S} \Pr(s'|s, \pi(s)) V^\pi(s')$. The optimal value function V^{π^*} with

the optimal policy π^* satisfies Bellman's optimality criterion [27]

$$Q^*(s, a(s)) = V^{\pi^*} = \max_{b \in A} Q^{\pi^*}(s, b(s)). \tag{7}$$

Thus, the optimal Q-function $Q^*(s, a(s))$ can be obtained from finding an optimal policy of Q-function $Q^{\pi^*}(s, a(s))$. Without knowing $U(s, a(s))$ and Pr$(s'|s, a(s))$, the Q-learning process can still find an optimal policy π^* through updating $Q(s, a(s))$ to find $Q^*(s, a(s))$ in a recursive manner using the information of current state s_t, action a_t, reward φ_t, and next state s'_t. Watkins [28] has shown that if the Q-value of each feasible state-action pair $(s, a(s))$ is visited infinitely often, and if the learning rate is decreased to zero in a suitable way, then $Pr\{Q(s, a(s)) \rightarrow Q^*(s, a(s))$ as $t \rightarrow \infty\} = 1$. The Q-values of the state-action pairs are usually stored in a look-up table. However, this approach is not suitable for problems with continuous state spaces as in multimedia CDMA systems, where the curse of dimensionality is hard to tackle. It has been shown [29] that fuzzy Q-learning is an efficient technique for the approximation of continuous system states by adapting Watkins's Q-learning [28] technique such that a fuzzy inference system (FIS) is incorporated into reinforcement-learning to generalize Q-learning by inferring both the actions and Q-functions from fuzzy rules. Taking advantage of the Q-learning technique, the universal approximation property of the FIS makes the representation of Q-values with large state-action space possible, and a priori knowledge can be integrated in the learning procedure [16].

Furthermore, in DCC-RL, a simple strategy for feature abstraction, exploitation and exploration is applied to speed up the learning procedures (and shorten the convergence time) for obtaining the optimal solution. A policy feasible action set $\mathbf{A_s} \subset \mathbf{A}$ can be obtained based on the current state s. State ω_M can be adopted as an indicator to classify the feasible action sets. For example,

$$\mathbf{A_s} = \begin{cases} \{f_{min}, \cdots, f_\Theta\}, & \text{if } \omega_M \geq \Theta \\ \{f_\Theta, \cdots, f_{max}\}, & \text{otherwise} \end{cases} \tag{8}$$

where f_Θ is the cutting value of the action set, $f_\Theta \in [f_{min}, f_{max}]$, and Θ is the threshold of the mean power as the quality of service (QoS) constraint. Since a greedy policy can easily cause the system to converge to locally optimal solutions, it is necessary to visit all the sets of possible actions for all states to find the globally optimum solution. This is the so-called exploration/exploitation dilemma. An action $a(s)$ of state s is selected from the feasible action

set $\mathbf{A_s}$ using an exploitation and exploration policy. Here, a pseudo-exhaustive policy is applied, in which the action with the best Q-value is chosen with a selection probability based on the Boltzmann distribution. Otherwise an action that is the least visited will be chosen. The resulting action is converted to the pilot power of the BS. The reward can be measured from the system, and fed back to update the Q-function.

C. Dynamic Maximum Link Power Constraint Design

The main purpose of the adjustment of maximum link power constraint is to couple the pilot power into the design. Note that pilot power adjustment affects the cell coverage, while the maximum link power of a cell affects the service coverage for MSs with different service rates near the cell boundary. In order to match cell coverage and service coverage, based on the maximum path loss (1) and the receiver sensitivity in terms of referenced service rate r^*, the total EIRP of pilot power E_T^I should be

$$E_{T^{[\mathrm{dBm}]}}^I = PL(r^*)_{[\mathrm{dB}]} + H_{R^{[\mathrm{dB}]}}^I + \Omega_{L^{[\mathrm{dB}]}}, \tag{9}$$

where H_R^I is the receiver sensitivity of the pilot signal such that $H_{R^{[\mathrm{dB}]}}^I = H_{S^{[\mathrm{dB}]}}^I + \Omega_{I\,[\mathrm{dB}]} + \eta_{o^{[\mathrm{dBm}]}}$, where H_S^I is the required SINR value of the pilot signal, which is equal to the required Eb/No of the pilot signal, $\gamma_{I\,[\mathrm{dB}]}^*$, minus the processing gain of the pilot signal $G_{P^{[\mathrm{dB}]}}^I$. Then, substituting (1) into (9), we obtain $E_{T^{[\mathrm{dBm}]}}^I = E_{T^{[\mathrm{dBm}]}} - H_R(r^*)_{[\mathrm{dB}]} + H_{R^{[\mathrm{dB}]}}^I$. Hence, as soon as the pilot power of BS b, P_b^I, has been adjusted dynamically, the maximum link power of cell b should be

$$\widetilde{p}_{b^{[\mathrm{dBm}]}} \leq P_{b\ [\mathrm{dBm}]}^I + H_R(r^*)_{[\mathrm{dB}]} - H_{R^{[\mathrm{dB}]}}^I. \tag{10}$$

The maximum link power constraint is thus coupled with pilot power accordingly. Note that the same constraint of the maximum link power for different service rates is adopted in this chapter because the processing gain can be regarded as a priority index for different service rates.

D. Dynamic CAC Criterion Design

In DCC-RL, as soon as the optimal pilot power P_b^I has been determined by the dynamic pilot power controller, as shown in Fig. 2, the corresponding maximum link power $\widetilde{p}_b$ can be updated by (10). The SINR threshold Λ_b for call admission in cell b becomes

$$\Lambda_{b^{[\mathrm{dB}]}} = \widetilde{p}_{b^{[\mathrm{dBm}]}} - PL(r^*)_{[\mathrm{dB}]} - \Omega_{I\,[\mathrm{dB}]} - \eta_{o^{[\mathrm{dBm}]}}. \tag{11}$$

For CAC of new calls, MS m originating a new call measures and reports its received SINR, $\widehat{H}_S$. The BS accepts the new call if $\widehat{H}_{S[\text{dB}]} > \Lambda_{b[\text{dB}]}$, otherwise the new call is blocked. For CAC of handoff calls, the soft handoff algorithm [30] is implemented, in which maximal ratio combining is used to obtain the overall SINR of MS h, $\widehat{H}_S$, from all serving BSs in the active set D_h. A handoff request is issued to BS b whenever an add event occurs. The BS accepts the handoff request if $\widehat{H}_{S[\text{dB}]} > \Lambda_{b[\text{dB}]}$, and the admitted handoff MS adds BS b into its active set D_h. Otherwise the handoff call request is blocked. On the other hand, if the blocked handoff call has not yet exceeded the handoff delay time, the MS can make a handoff request again as long as the link quality does not fall below the E_c/I_o requirement, $\widetilde{\Upsilon}_{[\text{dB}]}$ (2).

V. Simulation Results and Discussions

A simulation model is set up to examine the performance of the DCC-RL scheme in a CDMA cellular system. We first describe the simulation platform, and then the simulation results are presented and discussed.

A. Simulation Model

1) Cell Model: We consider a hexagonal cellular system with 19 wrap-around cells, in which the central cell is a hotspot cell with a high traffic load. As before, the load ratio ρ is defined as the ratio between the call arrival rates in the hotspot cell and in each surrounding cell. Geographically, the cellular deployment is homogenous, and the default cell radii can be determined by the link budget design in subsection III-C. The link budget parameters are as follows: $\widetilde{P}_b = 43_{[\text{dBm}]}$, $G_B = 2_{[\text{dBi}]}$, $L_C = 3_{[\text{dB}]}$, $G_M = 2_{[\text{dBi}]}$, $L_D = 3_{[\text{dB}]}$, $G_S = 3_{[\text{dB}]}$, $\Omega_I = 5_{[\text{dB}]}$, and $\eta_o = -127.24_{[\text{dBm}]}$.

2) Mobility Model: Assume MSs are uniformly distributed in each cell, and their initial speeds are uniformly distributed between 0 and the maximum speed. The maximum speeds for MSs in the hotspot cell, 1st-tier cells, and 2nd-tier cells are assumed to be 30, 60, and 60 km/hr, respectively. Whenever a MS moves into a different cell tier, a new speed is chosen according to the above distribution. Each MS is subject to correlated shadowing effect based on the Gudmundson model [30], in which the decorrelation length is 20 m in a vehicular environment. The shadowing effect is updated according to the correlated shadowing model, with coverage probability 95%. During each shadowing effect update, with probability 0.2 the

moving direction of the MS is changed and a new direction is selected at random among ± 45 degrees [30].

3) Channel Model: For the channel model [30], the path loss is obtained by $40 \times (1 - 0.004h_b) \times \text{log10}(d) - 18 \times \text{log10}(h_b) + 21 \times \text{log10}(f_d) + 80$, , where d is the distance between the BS and the MS; h_b and f_d are the antenna height of the BS and the downlink frequency, respectively. In our simulations, the downlink frequency is 2.4 Ghz and the antenna height is 20 m.

4) Traffic Model: Poisson call arrivals are assumed. Three service classes including real-time voice, real-time data, and non-real-time data, are considered in the system. In the simulations, the fractions of voice, real-time data, and non-real-time data traffic are 60%, 35%, and 5%, respectively. A 2-level Markov modulated Poisson process (MMPP) is used to model voice traffic while a 5-level MMPP is used to model real-time data traffic. The mean duration of each state in the 5-level MMPP is 1 second. The call holding times of real-time voice and data services are exponentially distributed with means 60 and 30 s, respectively. The transmission rate and required Eb/No of the voice traffic are 12.2 kbps and 5 dB, respectively. The service rates of the data traffic are 16, 32, 64, and 144 kbps and the corresponding Eb/No requirements are 5, 4, 3, and 2 dB. Note that adaptive rate transmission is applied whenever the power resources are not enough to support the existing MSs. For the non-real-time data service, variable length data bursts are assumed to be geometrically distributed with a mean burst size of 200 frames. Moreover, there are 6 different service rates: 16, 32, 64, 144, 384, and 512 kbps, which require Eb/No of 5, 4, 3, 2, 1.5, and 1 dB, respectively. The transmissions are on a burst-by-burst basis.

B. Performance Measurements and Discussions

We compare the performance of four schemes: FPPA with SSDT (FIX-SSDT), FPPA with LPPA (FIX-LPPA), DCC-RL with SSDT (DCC-SSDT), and DCC-RL with LPPA (DCC-LPPA). For FPPA, the default pilot power, P_b^I, is set at 2.5 w (12.5% of the maximum transmit power) for each cell. The maximum link power, $\widetilde{p}_b$, and the CAC threshold, H_I, are fixed and calculated from (10) and (11), respectively. For DCC-RL, P_b^I, $\widetilde{p}_b$, and H_I are adjusted dynamically as described in Section IV. Assume the arrival rate is 1.6 calls/s, and the traffic load ratio, ρ, is varied from 1 to 5. For the design parameters of DCC-RL, maximum and minimum fractions of pilot power are $f_{min} = 0.05$ and $f_{max} = 0.2$, respectively; decision period N is 10 frames;

total number of measurement samples M is 100 frames; and total simulation time is 10^6 frames (10^5 learning times).

Figs. 3(a) and (b) show the average pilot power distribution of the hotspot, 1st-tier, and 2nd-tier cells using DCC-LPPA and DCC-SSDT schemes, respectively. We can see that the DCC-RL schemes adjust the pilot power in each cell according to various system situations. When the traffic load ratio is increased, the pilot power of the hotspot cell is reduced aggressively so as to balance traffic load with adjacent cells, but the coverage is shrunk accordingly. In this way, the BS of the hotspot cell can save its transmit power to serve new call arrivals. Besides, adjustments of the pilot power can make the existing MSs near the cell boundary enter soft handoff mode so as to balance traffic load. Furthermore, for the hotspot cell, the slope of the pilot power level versus traffic load ratio for DCC-SSDT is sharper than that for DCC-LPPA. This is because both DCC and LPPA strategies are helpful for power-balancing, so that the pilot power of the DCC-LPPA scheme does not have to be adjusted aggressively.

Figs. 4(a) and (b) show the new call blocking probability of real-time and non-real-time services, respectively. We can see that the DCC-RL schemes improve the blocking probability of both real-time and non-real-time services relative to the FIX schemes. In order to achieve power-balance between cells, DCC-RL adjusts pilot power and coordinates other RRM mechanisms dynamically. This is the reason why the DCC-RL schemes can save more power resource to accommodate new call requests. Performance results of DCC-LPPA and DCC-SSDT schemes without adapting other RRM parameters are also presented for comparison. We can see that the DCC-RL schemes with fixed RRM parameters have worse new call blocking performance than the FIX schemes, as explained in Section IID. Similarly impaired results in handoff forced termination occur when a MS fails to add new BSs into its active set and suffers degraded channel quality, as shown in Fig. 5. This is because existing MSs near the cell boundaries often suffer bad transmission quality, and they may be dropped when power is not enough for admitting handoff requests. On the other hand, compared to the FIX schemes, the proposed DCC-RL schemes can improve handoff forced termination probabilities greatly.

Fig. 6 shows the total throughput of the system. In the FIX cases, FIX-LPPA outperforms FIX-SSDT. When the traffic load ratio is higher, the throughput of FIX-SSDT degrades sharply because of the inefficient handoff power allocation strategy. With FIX-LPPA, the average throughput keeps fairly constant when traffic load ratio is less than 4. Compared to the FIX

schemes, the DCC-RL schemes improve the average throughput when the traffic load ratio is increased. This is because DCC-RL can dynamically balance traffic load between cells through pilot power adjustments based on system situations as well as CAC criterion and the maximum link power constraint.

Furthermore, Fig. 7 compares the average frame error rates. We observe that DCC-RL can keep the frame error rate roughly under the requirement of 0.01 by the simple feature abstraction design. A more sophisticated design of the feature abstraction can guarantee the QoS requirement of the frame error rate strictly. It is noteworthy that the frame error rates of the DCC-RL schemes are worse than those of FIX-LPPA in some cases. This is because the DCC-RL schemes can make more efficient use of the total power resource to provide MSs with a good enough QoS that is just within the system requirement of a 0.01 frame error rate. Though FIX-LPPA can provide a better frame error rate than DCC-RL schemes when the traffic load ratio is high, the corresponding system throughput is lower resulting in poor new call blocking probability and handoff forced termination probability. The complementary results for system performance as described above give important insights in the design of downlink CDMA cellular systems.

In order to balance traffic loads between cells, DCC-RL can reduce or increase pilot power aggressively. Power-balancing can be achieved by forcing MSs near the cell boundary into handoff mode. Therefore, the average size of the active set and handoff rates can be increased, as shown in Fig. 8. It is found that the DCC-RL schemes cause slight increases in soft handoff events. Furthermore, Table I shows the coverage failure probability. A coverage failure occurs when a MS starting a new call fails to detect a good enough signal from a BS. The DCC-SSDT and DCC-LPPA schemes cause slightly higher coverage failure probabilities than the FIX schemes. This is because even though DCC works to balance traffic load through pilot power adjustments so as to reduce the interference of the hotspot cell, MSs near the cell boundary may suffer bad signal strengths from all BSs in the active set. Because of the tradeoff between capacity and coverage, we stress that coverage failure is an inevitable downside of any kind of DCC-RL scheme. The goal is to reduce the impact of this drawback through performance gains in system throughput, new call blocking probability, and handoff forced termination rate, etc. Due to the maximum power constraint in each BS, the system shows a performance tradeoff between coverage failure and call admission blocking. In a cellular system under heavy traffic

load, a new call request could fail either due to coverage failure, or due to blocking by CAC. Since our results show that the performance gain in reduced call blocking more than offsets the performance loss in increased coverage failure, our proposed DCC-RL can give an overall gain in system performance, and the goal stated above is successfully achieved.

REFERENCES

[1] S. Sharma, A. R. Nix, and S. Olafsson, "Situation-aware wireless networks," *IEEE Commun. Mag.*, pp. 44-50, July 2003.

[2] J. S. Lee, and L. E. Miller, *CDMA Systems Engieering Handbook*, Artech House, pp. 1111-1186, 1998.

[3] J. Laiho, A. Wacker, and T. Novosad, Eds, *Radio Network Planning and Optimization for UMTS*, Wiley and Sons, pp. 280-290, 2002.

[4] W. W. Lu, *Broadband Wireless Mobile: 3G and Beyond*, John Wiley & Sons Ltd, pp. 307-315, 2002.

[5] V. V. Veeravalli and A. Sendonaris, "The coverage-capacity tradeoff in cellular CDMA systems," *IEEE Trans. Veh. Technol.*, vol. 48, no. 5, pp. 1443-1450, Sept. 1999.

[6] R. G. Akl, M. V. Hegde, M. Naraghi-Pour, and P. S. Min, "Multicell CDMA network design," *IEEE Trans. Veh. Technol.*, vol. 50, no. 3, pp. 711-722, May 2001.

[7] S. Sharma and A. R. Nix, "Situation awareness based automatic basestation detection and coverage reconfiguration in 3G systems," in *Proc. IEEE PIMRC'02*, Lisbon, Portugal, pp. 16-20, Sept. 2002.

[8] Y. Ishikawa, T. Hayashi, and S. Onoe,"W-CDMA downlink transmit power and cell coverage planning," *IEICE Trans. Commun.*, vol. E85-B, no. 11, pp. 2416-2426, Nov. 2002.

[9] S. J. Park; D. Kim; C. Y. Kim,"Optimal power allocation in CDMA forward link using dependency between pilot and traffic channels," in *Proc. VTC'99 Fall*, Amsterdam, Netherlands, pp. 223-227, Sept. 1999.

[10] K. Mori, and H. Kobayashi, "Dynamic cell configuration scheme for common channel communications in CDMA cellular packet systems," in *Proc. IEEE ICC'04*, Paris, France, pp. 159-163, June 2004.

[11] G. Hampel, K. L. Clarkson, J. D. Hobby, and P. A. Polakos, "The tradeoff between coverage and capacity in dynamic optimization of 3G cellular networks," in *Proc. IEEE VTC'03 Fall*, Orlando, FL, pp. 927-932, Sept. 2003.

[12] R. T. Love, K. A. Beshir, D. Schaeffer, and R. S. Nikides, "A pilot optimization technique for CDMA cellular systems," in *Proc. IEEE VTC'99 Fall*, Amsterdam, Netherlands, pp. 2238-2242, Sept. 1999.

[13] K. Valkealahti, A. Hoglund, J. Parkkinen, and A. Flanagan, "CDMA common pilot power control with cost function minimization," in *Proc. IEEE VTC'02 Fall*, Vancouver, BC, pp. 2244-2247, Sept. 2002.

[14] D. Kim, Y. Chang, and J. W. Lee, "Pilot power control and service coverage support in CDMA mobile systems," in *Proc. IEEE VTC'99 Spring*, Amsterdam, Netherlands, pp. 1464-1468, May 1999.

[15] A. D. Smith, "Designing for coverage availability with different data rates - an improved methodology," in *Proc. IEEE VTC'02 Fall*, Vancouver, BC, pp. 1821-1824, Sept. 2002.

[16] C. Y. Liao, F. Yu, V. C. M. Leung, and C. J. Chang, "Reinforcement-learning-based self-organization for cell configuration in multimedia mobile networks," to appear in European Trans. on Telecomm., vol. 6, no. 5, Oct. 2005.

[17] A. Hoglund, and K. Valkealahti, "Quality-based tuning of cell downlink load target and link power maxima in CDMA," in *Proc. IEEE VTC'02, Fall*, Vancouver, BC, pp. 2248-2252, Sept. 2002.

[18] B. Hashem, and E. L. Strat, “On the balancing of the base stations transmitted powers during soft handoff in cellular CDMA systems,” in *Proc. IEEE ICC’00*, New Orleans, LA, pp. 1497 - 1501, June 2000.

[19] J. A. Flanagan, and T. Novosad, ”CDMA network cost function minimization for soft handover optimization with variable user load,” in *Proc. IEEE VTC’02 Fall*, Vancouver, BC, pp. 2224-2228, Sept. 2002.

[20] C. Y. Liao, “Downlink soft handoff mechanisms and cell reconfiguration planning in mixed-size WCDMA cellular networks,” *Ph.D. Dissertation*, National Chiao Tung University, Taiwan, 2004.

[21] C. Y. Liao, L. C. Wang, and C. J Chang, “Power Allocation Mechanisms for Downlink Handoff in the CDMA System with Heterogeneous Cell Structures,” to appear in *ACM/Kluwer WINET*, Sept. 2005.

[22] H. Furukawa, K. Hamabe, and A. Ushirokawa, “SSDT – Site Selection Diversity Transmission Power Control for CDMA Forward Link,” *IEEE J. Selected Areas Commun.*, vol. 18, no. 8, pp. 1546-1554, Aug. 2000.

[23] R. S. Sutton, and A. G. Barto, *Introduction to Reinforcement Learning*, MIT Press/Bradford Boos, Cambridge, MA. 1998.

[24] D. Bertsekas and J.N. Tsitsiklis, *Neuro-Dynamic Programming*, Athena Scientific, 1996.

[25] 3GPP Technical Specification 25.942, RF System Scenarios, page 26, December, 1999.

[26] M. L. Putterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley & Sons, Inc., 1994.

[27] R. E. Bellman, *Dynamic Programming*, Princeton, New Jersey: Princeton University Press, 1957.

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

[29] L. Jouffe, “Fuzzy inference system learning by reinforcement methods,” *IEEE Trans. SMC*, Part C, vol. 8, no. 3, pp. 338-355, Aug. 1998.

[30] Unviersal Mobile Telecommunication System (UMTS); “Selection procedures for the choice of radio transmission technologies of the UMTS,” *UMTS 30.03*, version 3.2.0, TR 101 112, pp. 54-55, 1998.

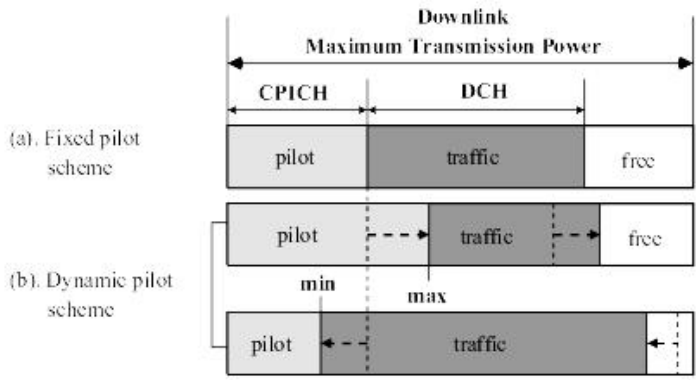

Fig. 1. Power allocation in downlink CDMA systems.

TABLE I

AVERAGE COVERAGE FAILURE PROBABILITY

Scheme/Traffic load ratio	1.0	2.0	3.0	4.0	5.0
FIX-SSDT	0.0	0.0	0.0	0.0	0.0
FIX-LPPA	0.0	0.0	0.0	0.0	0.0
DCC-SSDT	0.0	0.0	4.2e-05	1.0e-04	4.6e-04
DCC-LPPA	0.0	0.0	0.0	3.2e-05	9.2e-05

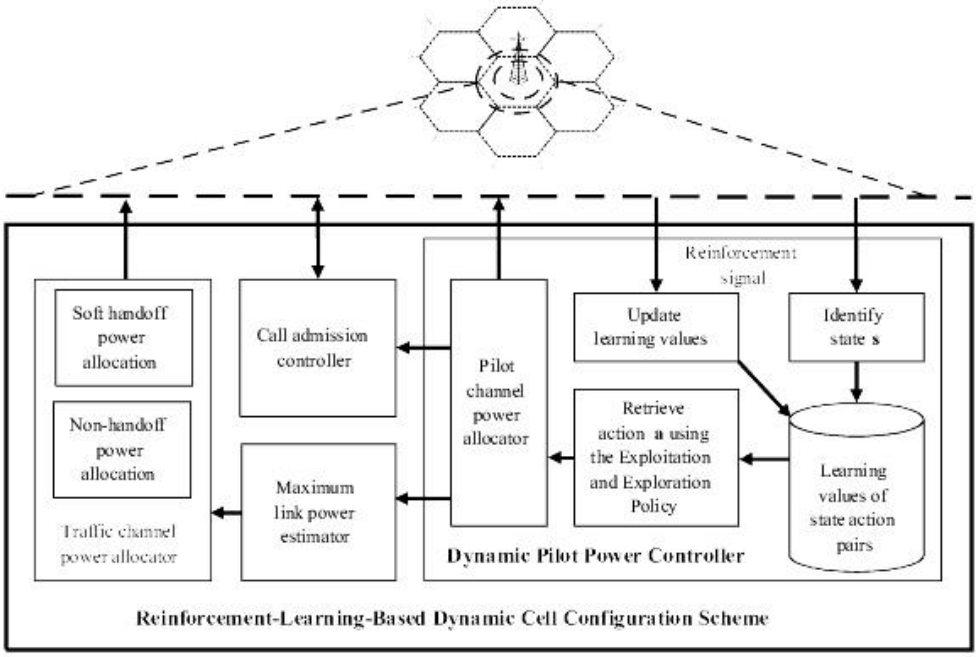

Fig. 2. System block diagram of proposed DCC-RL scheme.

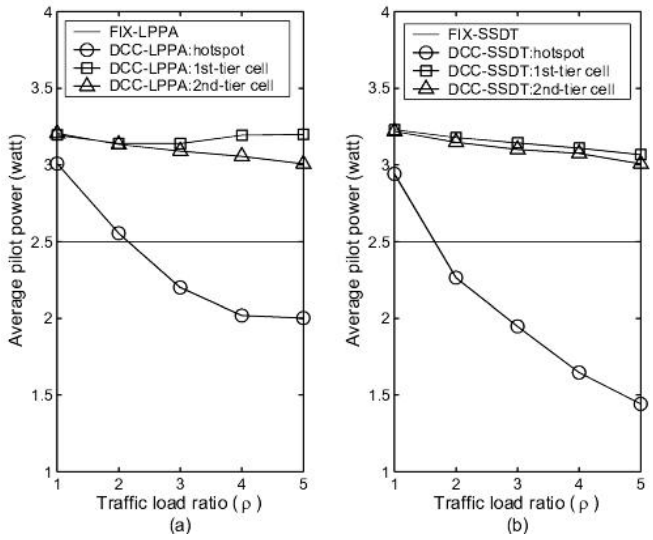

Fig. 3. Average pilot power of hotspot, 1st-tier, and 2nd-tier cells for (a) LPPA scheme and (b) SSDT scheme under FIX and DCC-RL.

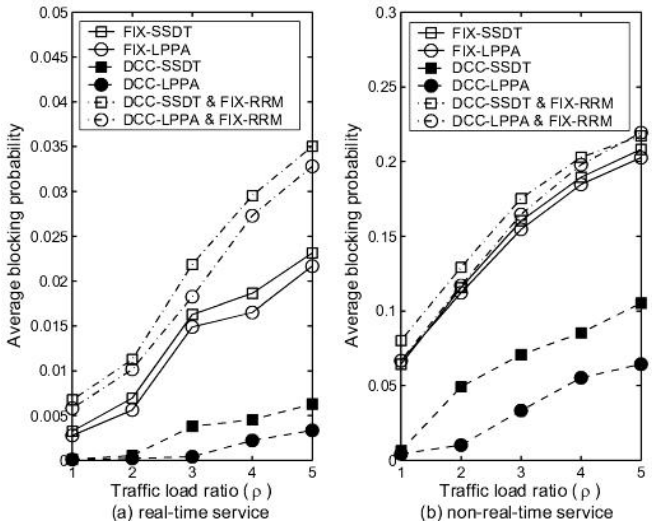

Fig. 4. Comparison of blocking probability of (a) real-time and (b) non-real-time services.

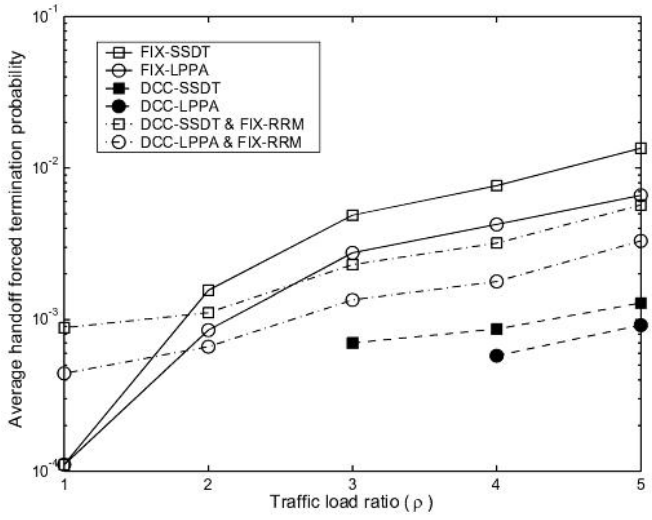

Fig. 5. Comparison of handoff forced termination probability.

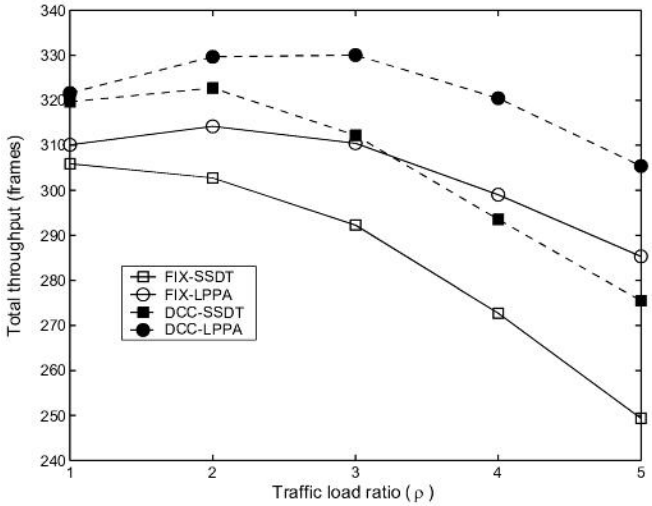

Fig. 6. Comparison of average total throughput.

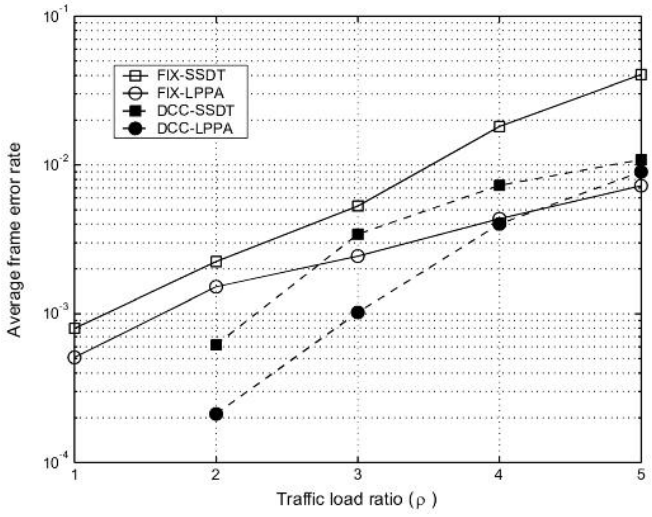

Fig. 7. Comparison of frame error probability.

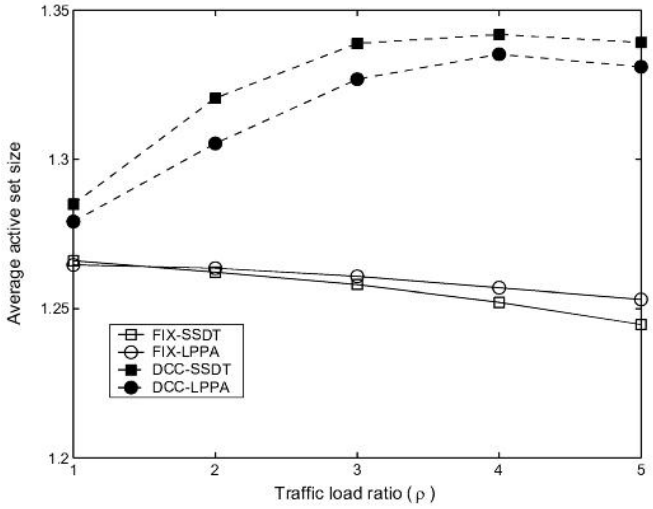

Fig. 8. Comparison of size of the active set.

Chapter 5

Concluding Remarks

In order to achieve high-efficiency resource allocation and quality-of-service (QoS) guarantee in B3G heterogeneous multiple access networks, we propose three critical schemes of resource management in this subproject. First, we present 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, 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. In the previous work, 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.

Then, we present a cellular neural network and utility (CNNU)-based scheduler, which jointly considers its radio resource efficiency, diverse QoS requirements, and fairness, to schedule the radio resource for connections in multimedia CDMA cellular systems. The utility function is defined to be the radio resource function properly weighted by the QoS requirement deviation function and the fairness compensation function. Also, the cellular neural network (CNN) is adopted to solve the constrained optimization problem defined for the radio resource scheduling in a real time fashion. Simulation results show that the CNNU-based scheduler can efficiently allocate the radio resource to achieve higher throughput than the EXP scheduling scheme. It can also effectively support differentiated QoS requirements for connections with variant traffic characteristics. Moreover, the CNNU-based scheduler can enlarge the QoS guaranteed region under the complicated QoS requirements environments. The CNNU-based scheduler is effective for multimedia CDMA cellular systems with diverse of QoS requirements when both dedicated and shared channels are adopted. However, call admission control to accept/reject the good/bad user to ensure the operation of the CNNU-based scheduler should be further considered.

Finally, we have studied the DCC problem in next generation CDMA networks, and proposed a model-free reinforcement-learning solution, DCC-RL, to solve the problem. DCC-RL can dynamically configure cell coverage and capacity based on the varying situations of the system. Simulation results show that pilot and soft handoff power allocations, maximum power constraint design, and the admission control criterion are highly coupled and should be considered jointly. Results also show that DCC-RL significantly increases the system throughput compared to conventional fixed pilot schemes. Furthermore, combining DCC-RL with LPPA gives the advantage of power-balancing for soft handoff so that the system capacity of the DCC-LPPA scheme outperforms conventional FIX-SSDT scheme significantly. The proposed DCC-RL solution gives a design framework suitable not only for the next-generation CDMA networks, but future cellular systems employing any signaling and multiple access techniques that take advantage of power control.