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

Ching-Yu Liao, *Member, IEEE*, Fei Yu, *Member, IEEE*, Victor C. M. Leung, *Fellow, IEEE*, and Chung-Ju Chang, *Senior Member, IEEE*

Abstract—To balance the time-varying traffic load between cells, caused by user mobility and diverse applications, it is crucial for next-generation code-division multiple-access (CDMA) cellular networks to configure cell coverage and capacity dynamically. In this paper, 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.

Index Terms—Code-division multiaccess, land mobile radio cellular systems, Markov processes.

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 predefined 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 intercell 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] and [9]. In [10], a DCC scheme for circuit-switched microcellular 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 with 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] and [21]. It was shown that LPPA can enhance system capacity in CDMA cellular systems with mixed-size cells, compared with conventional site-selection diversity transmissions (SSDT) scheme [22].

In this paper, 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.
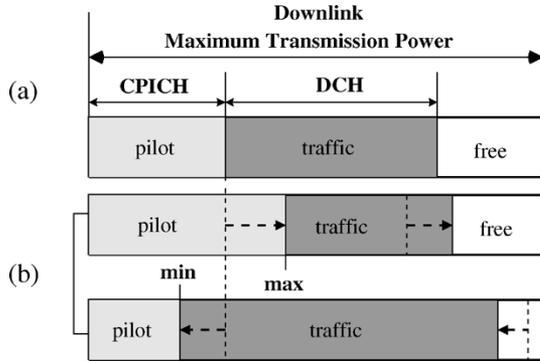
Fig. 1. Power allocation in downlink CDMA systems. (a) Fixed pilot scheme. (b) Dynamic pilot scheme.

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 paper 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. Section VI concludes the paper.

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 nonhandoff 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 nonhandoff 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 execute the handoff in time, a handoff forced termination occurs.

D. Simulations Illustrating Impacts of Pilot Power Control

In this section, we present simulation results to show that pilot power allocation and other RRM parameters are highly coupled. The simulation environment and parameters are given
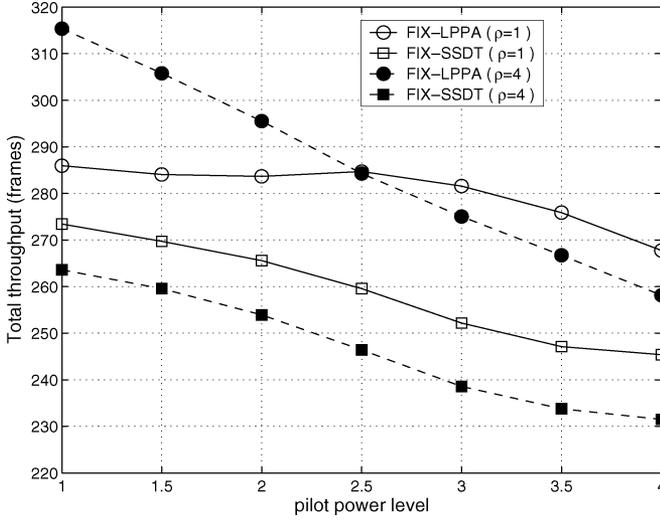
Fig. 2. Total throughput of fixed cell configuration for SSDT and LPPA schemes under uniform ($\rho = 1$) and nonuniform ($\rho = 4$) cell load cases.

in Section V. Fig. 2 shows the system throughput of fixed cell configuration with different pilot power levels under uniform ($\rho = 1$) and nonuniform ($\rho = 4$) cell traffic load for both SSDT and LPPA soft handoff schemes (FIX-SSDT, FIX-LPPA), where ρ is the ratio of traffic load between the hotspot cell in the center and the surrounding cells. In the simulations, the pilot channel power is set to 1 W in the FPPA scheme, and all other RRM algorithms are optimized according to this pilot power setting and then fixed. Subsequently, we simulate DPPA by adjusting the pilot power level, without changing the other RRM parameters.

Fig. 2 shows that the system throughput degrades whenever a pilot power other than 1 W (used in the fixed scheme) is used in DPPA. This is because increasing pilot power also increases interference to adjacent cell MSs; the larger the pilot power, the larger the interference, and the lower the throughput. Another reason is that the CAC criterion and the maximum link power constraint remain the same when the pilot power changes cell coverage. For example, when new or handoff calls issue requests to cells with light or heavy traffic, the tight or loose criteria of CAC may result in new call blockings or handoff forced terminations, respectively. Thus, uncoordinated design of pilot power and other RRM strategies can degrade the system performance severely. Fig. 2 also shows that soft handoff power allocation and pilot power allocation are highly coupled, both affecting system throughput. LPPA has larger throughput than SSDT with different pilot power and traffic load distributions. The throughput difference between SSDT and LPPA is larger when the traffic load is nonuniformly distributed ($\rho = 4$).

III. SYSTEM MODEL

The system block diagram of our proposed DCC-RL scheme is shown in Fig. 3. 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 $p_{b,m} = \phi_m P_b^T \le \widetilde{p}_b$, 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, $\Sigma_{m \in U_b} \phi_m = 1$, 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 $E_P[\text{dBm}] = \widetilde{p}_b[\text{dBm}] + G_B[\text{dBi}] - L_C[\text{dB}]$, 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 $E_T[\text{dBm}] = E_P[\text{dBm}] + G_M[\text{dB}] - L_D[\text{dB}] + G_S[\text{dB}]$. 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 $H_R(r)[\text{dB}] = H_S(r)[\text{dB}] + \Omega_I[\text{dB}] + \eta_o[\text{dBm}]$, where $H_S(r)[\text{dB}]$ 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), $\gamma^*(r)[\text{dB}]$, minus the processing gain $G_P(r)[\text{dB}]$. From the link budget, the maximum allowable path loss for service rate r is

$$PL(r)[\text{dB}] = E_T[\text{dBm}] - H_R(r)[\text{dB}] - \Omega_L[\text{dB}] \quad (1)$$

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 $\Upsilon(r)[\text{dB}] = P_b^I[\text{dBm}] - PL(r)[\text{dB}] - \Omega_I[\text{dB}] - \eta_o[\text{dBm}] \ge \widetilde{\Upsilon}$. In general, pilot power P_b^I is around 1–4 W, 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 $H_S(r_{\min}) <$

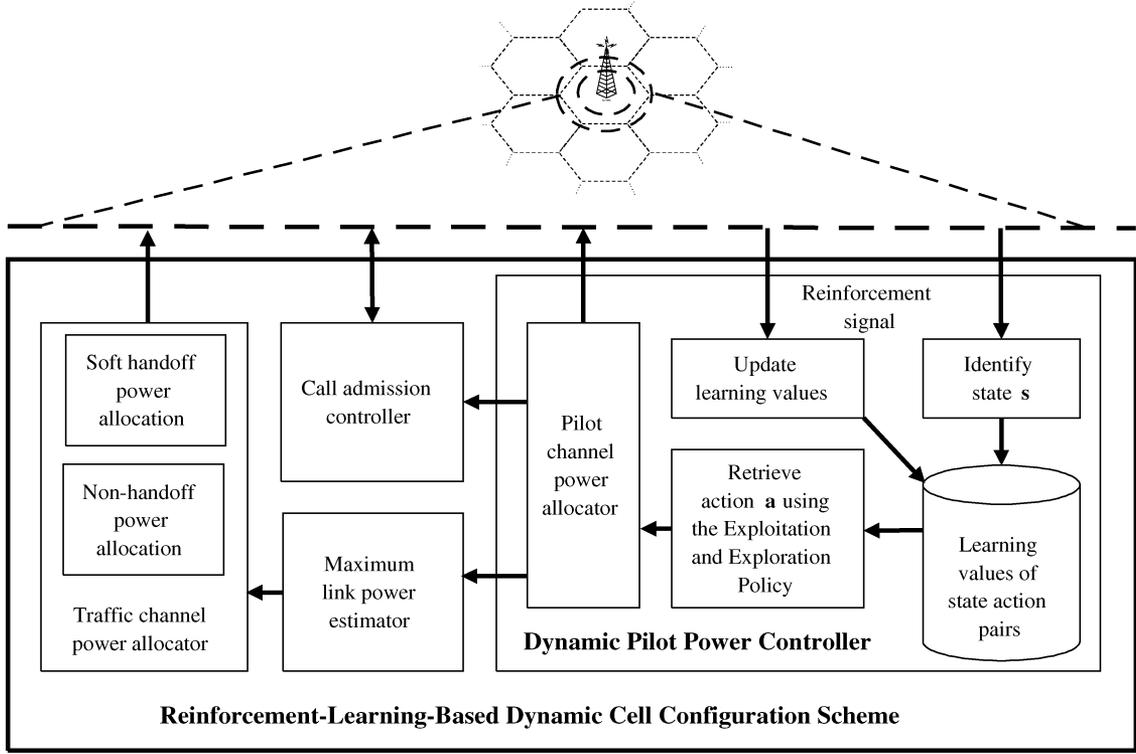[1]In this paper, a variable is linear if its unit is not specified.

Fig. 3. System block diagram of proposed DCC-RL scheme.

$H_S(r_{\max})$, 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 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

$$\widetilde{\Upsilon}[\mathrm{dB}] = P_b^I[\mathrm{dBm}] - PL(r^*)[\mathrm{dB}] - \Omega_I[\mathrm{dB}] + \eta_o[\mathrm{dBm}] \quad (2)$$

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

Based on the link budget and the channel model, the cell radii can be calculated in terms of different reference service rates $r_0 < r_1 < \cdots < r_6$ and the results represented as $\{\mathrm{ref}_0, \mathrm{ref}_1, \cdots, \mathrm{ref}_6\}$ in Fig. 4. Assume that 1 W power is allocated to the pilot channel for the FIX configuration with FPPA. Fig. 4 shows results of the total throughput (system capacity) in the cellular network by applying SSDT and LPPA schemes in terms of different referenced service coverage under uniform ($\rho = 1$) and nonuniform ($\rho = 4$) traffic load cases, where ρ is the traffic load ratio between a central cell and its surrounding cells. It is observed that, in all cases, a smaller cell coverage can
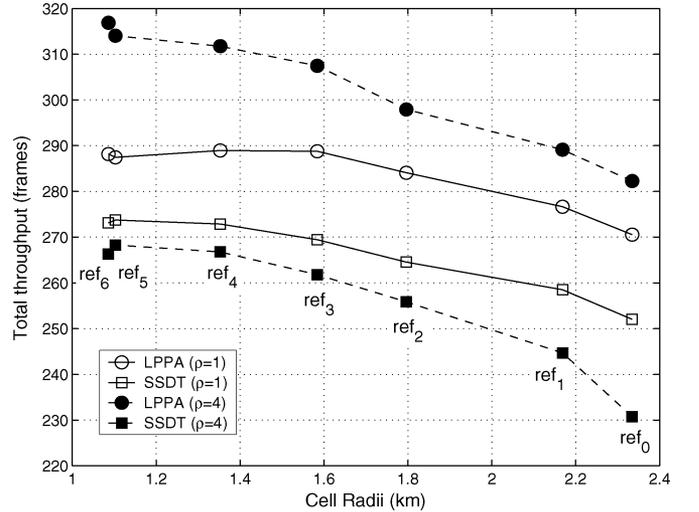

Fig. 4. Capacity versus referenced service coverage under fixed pilot power for SSDT and LPPA schemes under uniform ($\rho = 1$) and nonuniform ($\rho = 4$) cell loads.

increase the total throughput. This is because the lower propagation loss in a small cell provides a better signal quality for the MSs. Also, we can see that under nonuniform load ($\rho = 4$), the total throughput increases at a decreasing rate as the reference service rate increases. This means that the interrelation of cell coverage and system capacity becomes more sensitive to the reference service rate when there is unbalanced load between cells. Moreover, the curves flatten when the reference service rate exceeds r_4. This implies that the system has reached its capacity limit. Hence, the initial cell coverage in this simulation platform is set with a cell radius corresponding the reference service rate $r^* = r_4 = 144$ kb/s.

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 next, we propose a novel rein-forcement-learning-based DCC scheme, DCC-RL, to find an optimal policy for pilot power allocation that takes RRM into account (see Fig. 3).

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.

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

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

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 deci-sion 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, $V^\pi(s) = E\{\sum_{t=0}^{\infty} \lambda^t \cdot \varphi(s_t, \pi(s_t))|s_0 = s\}$, 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 $U(s, \pi(s)) = E\{\varphi(s, \pi(s))\}$. 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 crite-rion [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 $Pr\{Q(s, a(s)) \to Q^*(s, a(s)) \text{ as } t \to \infty\} = 1$. The Q-values of the state-action pairs are usually stored in a lookup table. However, this approach is not suitable for problems with con-tinuous 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 ap-proximation 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 pos-sible, and *a priori* knowledge can be integrated in the learning procedure [16].

Furthermore, in DCC-RL, a simple strategy for feature ab-straction, exploitation, and exploration is applied to speed up the learning procedures (and shorten the convergence time) for ob-taining 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 pseudoexhaustive 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

$$E_T^I[\text{dBm}] = PL(r^*)[\text{dB}] + H_R^I[\text{dB}] + \Omega_L[\text{dB}] \quad (9)$$

where H_R^I is the receiver sensitivity of the pilot signal such that $H_R^I[\text{dB}] = H_S^I[\text{dB}] + \Omega_I[\text{dB}] + \eta_o[\text{dBm}]$, 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 $\gamma_I^*[\text{dB}]$, minus the processing gain of the pilot signal $G_P^I[\text{dB}]$. Then, substituting (1) into (9), we obtain $E_T^T[\text{dBm}] = E_T[\text{dBm}] - H_R(r^*)[\text{dB}] + H_R^I[\text{dB}]$. 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

$$\widetilde{p}_b[\text{dBm}] \leq P_b^I[\text{dBm}] + H_R(r^*)[\text{dB}] - H_R^I[\text{dB}]. \quad (10)$$

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 paper 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. 3, 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

$$\Lambda_b[\text{dB}] = \widetilde{p}_b[\text{dBm}] - PL(r^*)[\text{dB}] - \Omega_I[\text{dB}] - \eta_o[\text{dBm}]. \quad (11)$$

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 wraparound 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 Section III-C. The link budget parameters are as follows: $\widetilde{P}_b = 43$ [dBm], $G_B = 2$ [dBi], $L_C = 3$ [dB], $G_M = 2$ [dBi], $L_D = 3$ [dB], $G_S = 3$ [dB], $\Omega_I = 5$ [dB], and $\eta_o = -127.24$ [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, first-tier cells, and second-tier cells are assumed to be 30, 60, and 60 km/h, 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 $\pm 45°$ [30].

3) Channel Model: For the channel model [30], the path loss is obtained by $40 \times (1 - 0.004 h_b) \times \log 10(d) - 18 \times \log 10(h_b) + 21 \times \log 10(f_d) + 80$, 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 nonreal-time data, are considered in the system. In the simulations, the fractions of voice, real-time data, and nonreal-time data traffic are 60%, 35%, and 5%, respectively. A two-level Markov modulated Poisson process (MMPP) is used to model voice traffic, while a five-level MMPP is used to model real-time data traffic. The mean duration of each state in the five-level MMPP is 1 s. 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 kb/s and 5 dB, respectively. The service rates of the data traffic are 16, 32, 64, and 144 kb/s, 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 nonreal-time data service, variable length data bursts are assumed to be geometrically distributed with a mean burst size of 200 frames. Moreover, there are six different service rates: 16, 32, 64, 144, 384, and 512 kb/s, 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 $\tilde{p}_b$ and the CAC threshold H_I are fixed and calculated from (10) and (11), respectively. For DCC-RL, P_b^I, $\tilde{p}_b$, 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 ten frames; total number of measurement samples M is 100 frames; and total simulation time is 10^6 frames (10^5 learning times).

The comparison between FIX-LPPA and FIX-SSDT in terms of capacity and coverage is shown in Fig. 4. We see that the FIX-LPPA scheme achieves a higher total throughput than the FIX-SSDT scheme for both uniform and nonuniform cell load cases. The throughput of FIX-LPPA is about 20% higher than that of the FIX-SSDT scheme in the nonuniform cell load case. This is because FIX-LPPA successfully releases congested cell's load through a power-balance strategy, whereas the FIX-SSDT scheme lacks the flexibility to adapt to nonuniform cell load situations.

Fig. 5(a) and (b) shows the average pilot power distribution of the hotspot, first-tier, and second-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 hot spot 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.

Fig. 6(a) and (b) shows the new call blocking probability of real-time and nonreal-time services, respectively. We can see that the DCC-RL schemes improve the blocking probability of both real-time and nonreal-time services relative to the FIX schemes. In order to achieve power balance between cells,
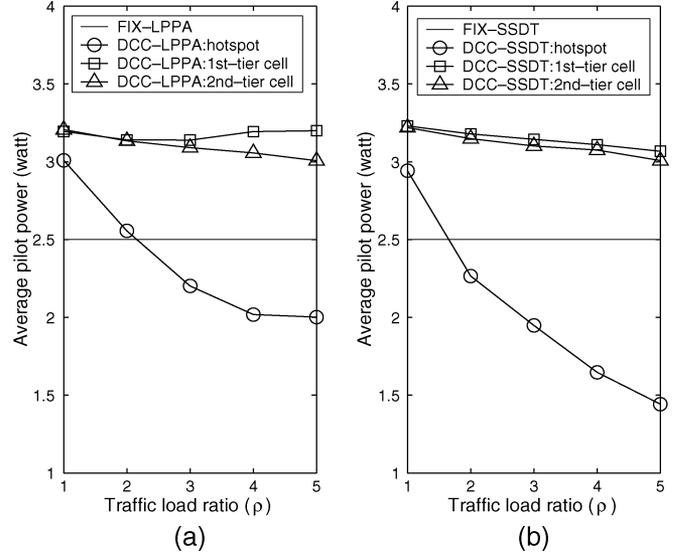

Fig. 5. Average pilot power of hotspot, first-tier, and second-tier cells for (a) LPPA scheme and (b) SSDT scheme under FIX and DCC-RL.
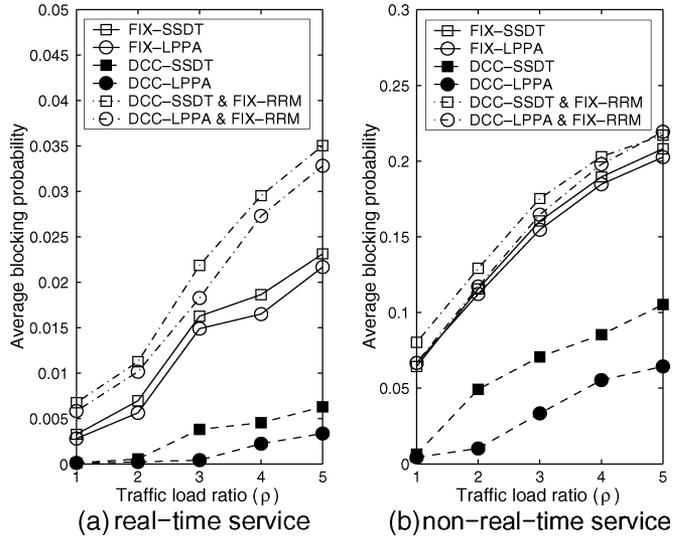

Fig. 6. Comparison of blocking probability of (a) real-time and (b) nonreal-time services.

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 II-D. 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. 7. 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 with the FIX schemes, the proposed DCC-RL schemes can improve handoff forced termination probabilities greatly.
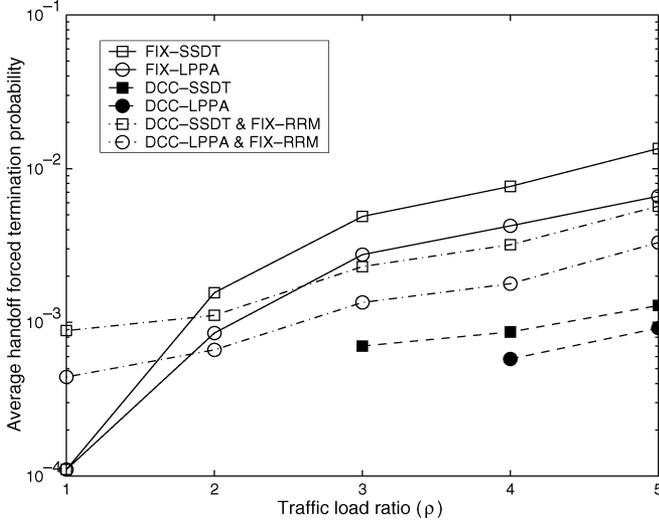
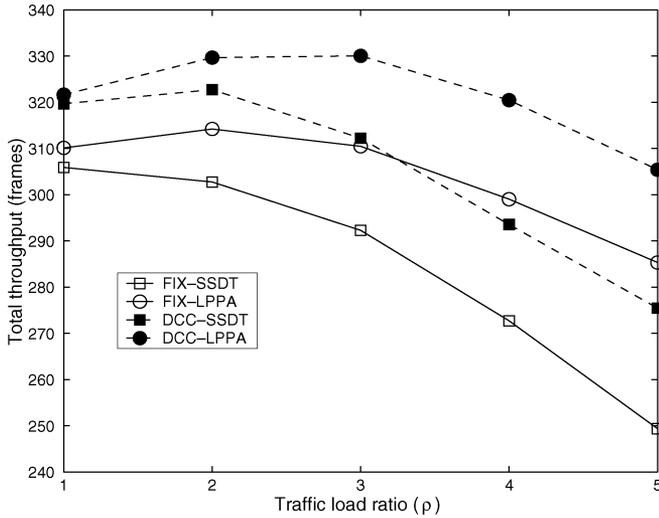Fig. 7.   Comparison of handoff forced termination probability.

Fig. 9.   Comparison of frame error probability.

Fig. 8.   Comparison of average total throughput.
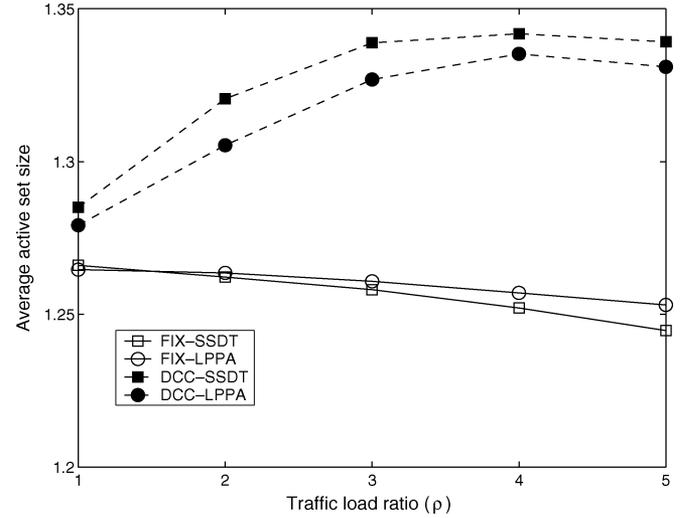
Fig. 10.   Comparison of size of the active set.

Fig. 8 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 with 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. 9 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
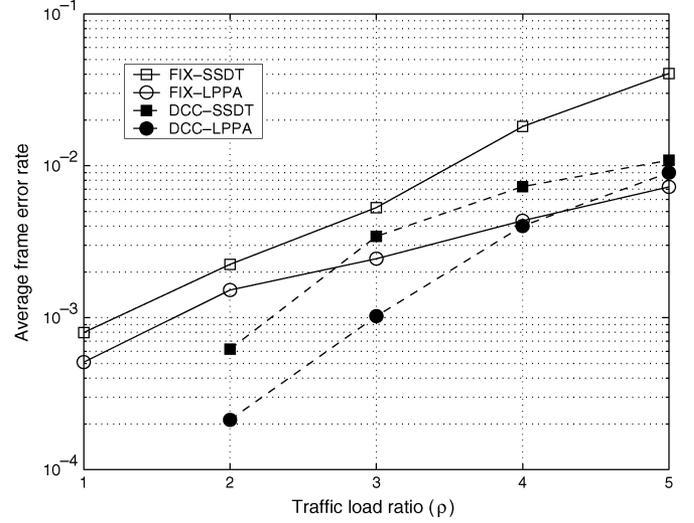
efficient use of the total power resource to provide MSs with a good enough QoS that is just within the system requirement of a 0.01 frame error rate. Though FIX-LPPA can provide a better frame error rate than DCC-RL schemes when the traffic load ratio is high, the corresponding system throughput is 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. 10. 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

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

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.

VI. CONCLUSION

In this paper, 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 with 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.

REFERENCES

[1] S. Sharma, A. R. Nix, and S. Olafsson, "Situation-aware wireless networks," *IEEE Commun. Mag.*, vol. 41, no. 7, pp. 44–50, Jul. 2003.
[2] J. S. Lee and L. E. Miller, *CDMA Systems Engineering Handbook*. Norwood, MA: Artech House, 1998, pp. 1111–1186.
[3] J. Laiho, A. Wacker, and T. Novosad, Eds., *Radio Network Planning and Optimization for UMTS*. New York: Wiley, 2002, pp. 280–290.
[4] W. W. Lu, *Broadband Wireless Mobile: 3G and Beyond*. New York: Wiley, 2002, pp. 307–315.
[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, Sep. 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 base station detection and coverage reconfiguration in 3G systems," in *Proc. IEEE PIMRC*, Lisbon, Portugal, Sept. 2002, pp. 16–20.
[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, and C. Y. Kim, "Optimal power allocation in CDMA forward link using dependency between pilot and traffic channels," in *Proc. Veh. Technol. Conf., Fall*, Amsterdam, The Netherlands, Sep. 1999, pp. 223–227.
[10] K. Mori and H. Kobayashi, "Dynamic cell configuration scheme for common channel communications in CDMA cellular packet systems," in *Proc. IEEE ICC*, Paris, France, Jun. 2004, pp. 159–163.
[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 Veh. Technol. Conf., Fall*, Orlando, FL, Sep. 2003, pp. 927–932.
[12] R. T. Love, K. A. Beshir, D. Schaeffer, and R. S. Nikides, "A pilot optimization technique for CDMA cellular systems," in *Proc. IEEE Veh. Technol. Conf., Fall*, Amsterdam, The Netherlands, Sep. 1999, pp. 2238–2242.
[13] K. Valkealahti, A. Hoglund, J. Parkkinen, and A. Flanagan, "CDMA common pilot power control with cost function minimization," in *Proc. IEEE Veh. Technol. Conf., Fall*, Vancouver, BC, Canada, Sep. 2002, pp. 2244–2247.
[14] D. Kim, Y. Chang, and J. W. Lee, "Pilot power control and service coverage support in CDMA mobile systems," in *Proc. IEEE Veh. Technol. Conf., Spring*, Amsterdam, The Netherlands, May 1999, pp. 1464–1468.
[15] A. D. Smith, "Designing for coverage availability with different data rates—An improved methodology," in *Proc. IEEE Veh. Technol. Conf., Fall*, Vancouver, BC, Canada, Sep. 2002, pp. 1821–1824.
[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," *Eur. Trans. Telecomm.*, vol. 16, no. 5, pp. 385–397, Sep. 2005.
[17] A. Hoglund and K. Valkealahti, "Quality-based tuning of cell downlink load target and link power maxima in CDMA," in *Proc. IEEE Veh. Technol. Conf., Fall*, Vancouver, BC, Canada, Sep. 2002, pp. 2248–2252.
[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*, New Orleans, LA, Jun. 2000, pp. 1497–1501.
[19] J. A. Flanagan and T. Novosad, "CDMA network cost function minimization for soft handover optimization with variable user load," in *Proc. IEEE Veh. Technol. Conf., Fall*, Vancouver, BC, Canada, Sep. 2002, pp. 2224–2228.
[20] C. Y. Liao, "Downlink soft handoff mechanisms and cell reconfiguration planning in mixed-size wcdma cellular networks," Ph.D. dissertation, National Chiao Tung Univ., Hsinchu, 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," *ACM/Kluwer WINET*, vol. 11, no. 5, pp. 593–605, Sep. 2005.
[22] H. Furukawa, K. Hamabe, and A. Ushirokawa, "SSDT—Site selection diversity transmission power control for CDMA forward link," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 8, pp. 1546–1554, Aug. 2000.
[23] R. S. Sutton and A. G. Barto, *Introduction to Reinforcement Learning*. Cambridge, MA: MIT Press, 1998.
[24] D. Bertsekas and J. N. Tsitsiklis, *Neuro-Dynamic Programming*. Belmont, MA: Athena Scientific, 1996.
[25] 3GPP Tech. Specification 25.942, RF System Scenarios, p. 26, Dec. 1999.
[26] M. L. Putterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York: Wiley, 1994.
[27] R. E. Bellman, *Dynamic Programming*. Princeton, NJ: Princeton Univ. Press, 1957.
[28] C. J. C. H. Watkins and P. Dayan, "*Q*-learning," *Mach. Learn.*, vol. 8, pp. 279–292, 1992.
[29] L. Jouffe, "Fuzzy inference system learning by reinforcement methods," *IEEE Trans. Syst., Man, Cybern.*, pt. C, vol. 8, no. 3, pp. 338–355, Aug. 1998.
[30] Universal 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, 1998.

Ching-Yu Liao (S'02–M'05) received the Ph.D. degree in communication engineering from the National Chiao Tung University (NCTU), Hsinchu, Taiwan, R.O.C., in 2004.

From 2004 to 2005, she was a Visiting Scholar in the Department of Electrical Engineering, University of British Columbia (UBC), Vancouver, BC, Canada. She is currently a Senior Research Scientist at Telcordia Applied Research Center, Taiwan Company (TARC-TW), Taipei. Her research interests include handoff techniques, radio resource management, heterogeneous wireless networks, etc.

Victor C. M. Leung (S'75–M'89–SM'97–F'03) received the B.A.Sc. (Honors) and Ph.D. degrees from the University of British Columbia (UBC), Vancouver, BC, Canada, in 1977 and 1981, respectively, both in electrical engineering.

From 1981 to 1987, he was a Senior Member of Technical Staff and Satellite Systems Specialist at MPR Teltech, Ltd. In 1988, he was a Lecturer in Electronics at the Chinese University of Hong Kong. He returned to UBC as a faculty member in 1989, where he is a Professor and holder of the TELUS Mobility Research Chair in Advanced Telecommunications Engineering in the Department of Electrical and Computer Engineering. His research interests are in mobile systems and wireless networks.

Dr. Leung is member of the Association for Computing Machinery (ACM). He is an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and an Associate Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.

Fei Yu (S'00–M'04) received the Ph.D. degree in electrical engineering from the University of British Columbia (UBC), Vancouver, BC, Canada, in 2003.

From 1998 to 1999, he was a System Engineer at China Telecom, P.R. China, working on the planning, design, and performance analysis of national SS7 and GSM networks. From 2002 to 2004, he was a Research and Development Engineer at Ericsson Mobile Platforms, Sweden, where he worked on dual-mode UMTS/GPRS handsets. He is currently a Research Associate at the UBC. His research interests are QoS, cross-layer design, and mobility management in wireless networks.

Chung-Ju Chang (S'84–M'85–SM'94) was born in Taiwan, R.O.C., in August 1950. He received the B.E. and M.E. degrees in electronics engineering from the National Chiao Tung University (NCTU), Hsinchu, Taiwan, R.O.C., in 1972 and 1976, respectively, and the Ph.D degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 1985.

From 1976 to 1988, he was with Telecommunication Laboratories, Ministry of Communications, Taiwan, as a Design Engineer, Supervisor, Project Manager, and then Division Director. In 1988, he joined the Faculty of the Department of Communication Engineering, NCTU, as an Associate Professor. He has been a Professor since 1993. He was Director of the Institute of Communication Engineering from August 1993 to July 1995, Chairman of the Department of Communication Engineering from August 1999 to July 2001, and Dean of the Research and Development Office from August 2003 to July 2004, at NCTU. His research interests include performance evaluation, wireless communication networks, and broadband networks.

Dr. Chang is a member of the Chinese Institute of Engineers (CIE). He serves as an Editor for the *IEEE Communication Magazine* and an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.