A Neural Fuzzy Resource Manager for Hierarchical Cellular Systems Supporting Multimedia Services

Kuen-Rong Lo, Chung-Ju Chang, *Senior Member, IEEE*, and C. Bernard Shung

***Abstract*—Using intelligent techniques to perform radio resource management is an effective method. This paper proposes neural fuzzy control for radio resource management in hierarchical cellular systems supporting multimedia services. A *neural fuzzy resource manager* (NFRM) is designed, which mainly contains a *neural fuzzy channel allocation processor* (NFCAP), and NFCAP is in a two-layer architecture: a *fuzzy cell selector* (FCS) in the first layer and a *neural fuzzy call-admission and rate controller* (NFCRC) in the second layer. The FCS chooses not only the handoff failure probabilities and the resource availabilities in both microcell and macrocell but also the mobility of user as input linguistic variables. The NFCRC takes the handoff failure probability and the resource availability of the selected cell as input variables to perform call admission control and rate control for the call. Simulation results show that NFRM can always guarantee the quality of service (QoS) requirement of handoff failure probability for all traffic loads. Also, NFRM improves the system utilization by 31.1% while increasing the handoff rate by 2% over the overflow channel allocation (OCA) scheme [3]; it enhances the system utilization by 6.3% and 1.4%, and still reduces the handoff rate by 14.9% and 6.8%, as compared to the combined channel allocation (CCA) scheme [20], [21] and fuzzy channel allocation control (FCAC) scheme [9], respectively, under a predefined QoS constraint.**

***Index Terms*—Call admission, hierarchical cellular systems, macrocell, microcell, neural fuzzy, resource management.**

I. Introduction

THE future mobile communication system will provide not only voice and low-speed data services but also high-speed multimedia services [1], [2]. A way to provide a wide variety of services is to flexibly aggregate multiple channels (time-slot or spreading code), without changing the spectrum division, modulation, and burst structure. A mobile station (MS) specifies the required capacity and desired capacity in the setup request or handoff request message. The required and desired capacities characterize the service requirements of an application, or the patience of a user. Based on the availability of resources in a cell and the quality-of-service (QoS) requirement, the network gives the MS a number of channels between the required capacity and desired capacity. On the other hand, a hierarchical cellular structure, which contains overlaid microcells for high-teletraffic area and overlaying macrocells for low-teletraffic region, has merits of enhancing system capacity and improving coverage [3]–[6]. Such a wireless network must be adaptive and robust to support resource demands.

Nowadays, the intelligent techniques have been widely applied to nonlinear, time-varying, and complicated problems that were challenging using conventional algorithmic methods. These techniques such as fuzzy logics, neural networks, and neural fuzzy networks have been shown to outperform algorithmic methods. The advantages of intelligent techniques are numerous, most notably learning from experience and the scalability, adaptability, and ability to extract rules without the need for detailed or precise mathematical modeling [7]–[16].

In this paper, we propose a *neural fuzzy resource manager* (NFRM) for hierarchical cellular systems providing multimedia services. The NFRM utilizes the learning capability of the neural network to reduce the decision error of these conventional channel assignment schemes resulting from modeling, approximation, and unpredictable statistical fluctuations of the system. It also employs the control rule structure of fuzzy logic, which absorbs benefits of those conventional channel assignment schemes, to provide robust operation and to prevent operating errors due to the learning of incorrect training data. The NFRM contains a *neural fuzzy channel allocation processor* (NFCAP), a resource estimator, a performance evaluator, and base-station interface modules. NFCAP is a two-layer neural fuzzy logic controller that consists of a *fuzzy cell selector* (FCS) in the first layer and a *neural fuzzy call-admission and rate controller* (NFCRC) in the second layer. The FCS considers the handoff failure probability, the resource availability in both macrocell and microcell, and the mobility of users as input linguistic variables, and applies the *max-min* interference method to determine which cell, macrocell or microcell, to serve the call request; FCS intends to enhance the channel utilization by balancing utilization between macrocell and microcells. The NFCRC takes the handoff failure probability and the resource availability of the selected cell as input variables; NFCRC intends to guarantee the QoS and provides an appropriate rate for users. Simulation results show that NFRM can guarantee the QoS requirement of handoff failure probability for all traffic loads. NFRM improves the system utilization by 31.1% while increasing the handoff rate by 2% over the *overflow channel allocation* (OCA) scheme proposed in [3]; and it enhances the system utilization by 6.3% and 1.4%, and still reduces the handoff rate by 14.9% and 6.8%, as compared to the *combined channel allocation* (CCA) and *fuzzy channel allocation control* (FCAC) scheme proposed in [20], [21] and [9], respectively, under a defined QoS constraint.

Manuscript received May 17, 2001; revised December 31, 2002. This work was supported by the National Science Council, Taiwan, R.O.C., under Contract NSC 87-2218-E009-047 and NSC 88-2213-E009-127.

K.-R. Lo is with Telecommunication Laboratories, Chunghwa Telecom Co., Ltd., Chung-Li, Taiwan, R.O.C.

C.-J. Chang is with National Chiao Tung University, 300 Taiwan, R.O.C. (e-mail: cjchang@cc.nctu.edu.tw).

C. B. Shung is with Broadcom Corp., San Jose, CA 95134 USA.

Digital Object Identifier 10.1109/TVT.2003.816002

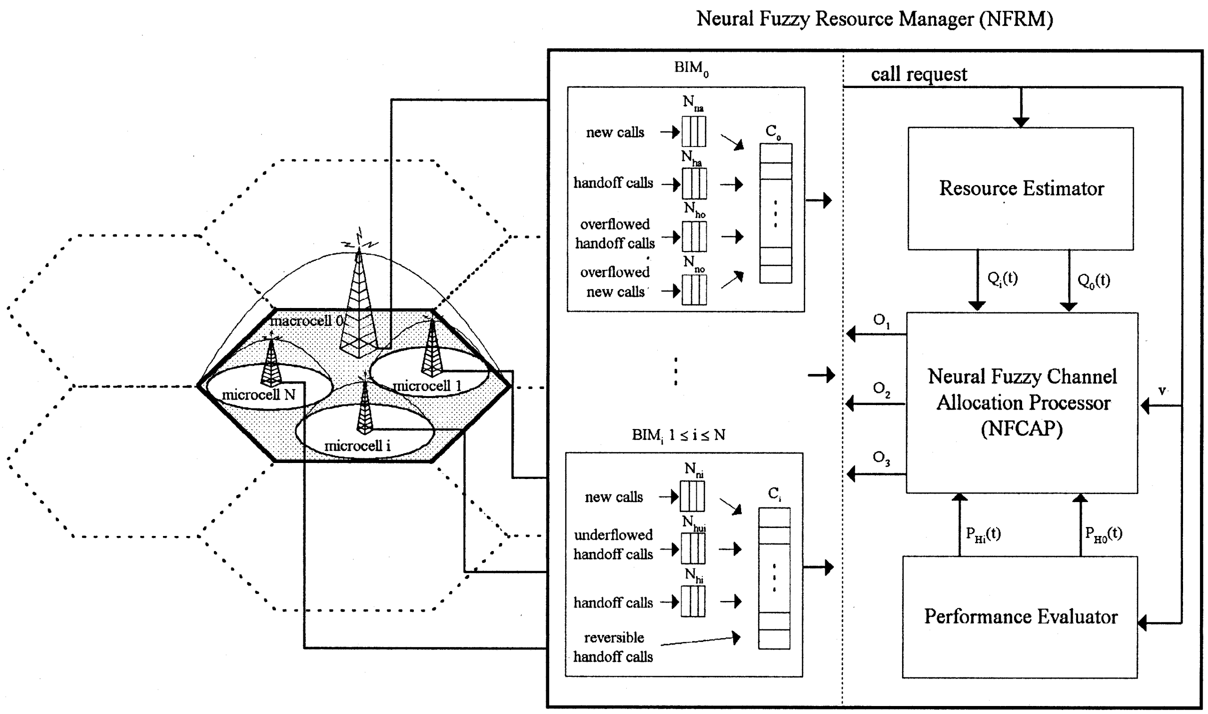

Fig. 1. The NFRM for hierarchical cellular systems.

The rest of this paper is organized as follows. Section II presents the functions of NFRM. Section III gives the design of NFCAP. Section IV shows simulation results and discussions. Conclusions are given in Section V.

II. Neural Fuzzy Resource Manager (NFRM)

Fig. 1 shows the functional block diagram of NFRM for hierarchical cellular systems supporting multimedia services, where the hierarchical cellular system contains a large geographical region tessellated by cells, referred to as macrocells, each of which overlays several microcells. The overlaying macrocell is denoted by cell 0 and its overlaid microcells are denoted by cell i, $1 \leq i \leq$ N. The NFRM contains functional blocks such as *base-station interface module* (BIM), *resource estimator, performance evaluator,* and *neural fuzzy channel allocation processor* (NFCAP). BIM is to interface with macrocell or microcell base stations. It is installed in the base-station controller (BSC) or mobile switching center (MSC). Note that for simplicity, NFRM is drawn to do the resource management for only one macrocell here.

Cell i in the macrocell is equipped with a pool of C_i independent communication channels, $1 \leq i \leq$ N. Assume all the channels are shared by new calls and handoff calls. The new call requests generated in each MS is modeled as a Poisson process with mean rate λ, in which the arrival rate of the new voice calls is $\lambda_{nv} = \gamma\lambda$ and the arrival rate of the new data calls is $\lambda_{nd} = (1-\gamma)\lambda$, $0 \leq \gamma \leq 1$. The call durations for the two streams are assumed to be exponentially distributed with averages equal to $1/\mu_v$ and $1/\mu_d$. We assume that all the data (voice) call requests have identical required capacity and desired capacity, denoted by R_{rd} and R_{dd} (R_{rv} and R_{dv}), respectively.

A. Base-Station Interface Modules (BIM)

Assume that each BIM provides complete-partitioning buffers for queueing new and handoff calls, which are originated in the corresponding cell and temporarily have no free channel to use. In the BIM for cell 0 (BIM_0) are a new-call buffer with capacity N_{na} for new calls originating in the macrocell-only region, a handoff-call buffer with capacity N_{ha} for handoff calls from adjacent macrocells, an overflowed handoff- call buffer with capacity N_{ho} for overflowed handoff calls from overlaid microcells, and an overflowed new-call buffer with capacity N_{no} for overflowed new calls from overlaid microcells. In the BIM for cell i (BIM$_i$), $1 \leq i \leq N$, there are a new-call buffer with capacity N_{ni} for new call originations, an underflowed handoff-call buffer with capacity N_{hui} for underflowed handoff calls from the overlaying macrocell, and a handoff-call buffer with capacity N_{hi} for handoff calls from adjacent microcells. No buffer is provided for the reversible handoff calls. Reneging of new calls and dropping of handoff calls are considered because of new calls' impatience and handoff calls' moving out the handoff area. The patience times are exponentially distributed.

Whenever BIM$_i$ receives a call request, $0 \leq i \leq N$, it sends the necessary calling information to the resource estimator, the performance evaluator, and the NFCAP. The calling information can distinctly indicate from which cell and in what type the call is originated. The k type of call is defined as: $k = 1$ denotes a new call originating in macrocell-only region; $k = 2$ denotes a handoff call from adjacent macrocell to macrocell-only region; $k = 3$ denotes a handoff call from microcell to macrocell-only region; $k = 4$ denotes new call originating in microcell; $k = 5$ denotes handoff call from adjacent macrocell to an overlaid microcell; $k = 6$ denotes handoff call from microcell to microcell; and $k = 7$ denotes reversible handoff. Note that the macro-

TABLE I
THE CALCULATION OF AVAILABLE RESOURCE

k	Q_0	Q_i
1	$C_0 + N_{na} - r_0(t) - b_{na}(t)$	0
2	$C_0 + N_{ha} - r_0(t) - b_{ha}(t)$	0
3	$C_0 + N_{ho} - r_0(t) - b_{ho}(t)$	0
4	$C_0 + N_{no} - r_0(t) - b_{no}(t)$	$C_i + N_{ni} - r_i(t) - b_{ni}(t)$
5	$C_0 + N_{ha} - r_0(t) - b_{ha}(t)$	$C_i + N_{hui} - r_i(t) - b_{hui}(t)$
6	$C_0 + N_{ho} - r_0(t) - b_{ho}(t)$	$C_i + N_{hi} - r_i(t) - b_{hi}(t)$
7	$C_0 - r_0(t)$	$C_i - r_i(t)$

cell-only region is the area inside macrocell 0 but outside all microcells. The first three types of calls are to use channels in macrocell, while other types of calls can use channels either in macrocell or in microcell.

B. Resource Estimator

The resource estimator calculates the available resources in macrocell 0 and microcell i when it receives calling information of type-k call from BIM$_i$ at time instant t, which are denoted by $Q_0(t)$ and $Q_i(t)$, respectively. The resource estimator knows system parameters of C_0, N_{na}, N_{ha}, N_{ho}, N_{no}, and C_i, N_{ni}, N_{hui}, N_{hi}, $1 \leq i \leq N$, and it obtains $Q_0(t)$ and $Q_i(t)$ by formulas shown in Table I.

In Table I, $r_0(t)(r_i(t))$ is the number of occupied channels in $C_0(C_i)$ at time t; $b_{na}(t)$ ($b_{ha}(t)$, $b_{ho}(t)$, $b_{no}(t)$) is the number of waiting calls in the new-call buffer (handoff-call buffer, overflowed handoff-call buffer, overflowed new-call buffer) of BIM$_0$, at time t; and $b_{ni}(t)$ ($b_{hui}(t)$, $b_{hi}(t)$) is the number of waiting calls in the new-call buffer (underflowed handoff-call buffer, handoff-call buffer) of BIM$_i$ at time t, $1 \leq i \leq N$.

C. Performance Evaluator

The performance evaluator is to calculate the handoff failure probability for NFCAP. The handoff failure probability in macrocell (microcells) at time t, denoted by $P_{H0}(t)(P_{Hi}(t))$, is defined as

$$P_{H0}(t) = \frac{HB_0(t) + HR_0(t)}{NH_0(t)}$$
$$\left(P_{Hi}(t) = \frac{HB_i(t) + HR_i(t)}{NH_i(t)}\right) \tag{1}$$

where $HB_0(t)(HB_i(t))$ is the number of blocked handoff calls in macrocell 0 (microcell i) at time t; $HR_0(t)(HR_i(t))$ is the number of dropped handoff calls in macrocell 0 (microcell i) at time t; and $NH_0(t)(NH_i(t))$ is the number of handoff calls in macrocell 0 (microcell i) at time t.

D. Neural Fuzzy Channel Allocation Processor (NFCAP)

NFCAP performs the channel allocation using neural fuzzy logic control to attain QoS guaranteed, high channel utilization, and good user satisfaction. In the neural fuzzy logic control, a reinforcement learning is designed to adjust the mean and the variance of the membership functions to cope with the input traffic fluctuation. The detailed design of NFCAP is described in the next section.

III. DESIGN OF NFCAP

NFCAP contains two functional blocks: FCS in the first layer and NFCRC in the second layer, as shown in Fig. 2.

A. Fuzzy Cell Selector (FCS)

NFCAP chooses five input linguistic variables for FCS: available resources in macrocell 0 $(Q_0(t))$ and in microcell i $(Q_i(t))$, handoff failure probabilities in macrocell 0 $(P_{H0}(t))$ and in microcell i $(P_{Hi}(t))$, and mobile speed (v), and has one output linguistic variable for FCS: the selection of macrocell or microcell (O_1). The available resource of cells can indicate the remaining capacity, the handoff failure probability can show the QoS, and the mobile speed can implicate the handoff rate. Term sets for both $Q_0(t)$ and $Q_i(t)$ are $T(Q_0(t)) = T(Q_i(t)) = T$ {More Enough, Slightly Enough, Not Enough} $= T\{ME,\ SE,\ NE\}$, term sets for both $P_{H0}(t)$ and $P_{Hi}(t)$ are $T(P_{H0}(t)) = T(P_{Hi}(t)) = T$ {Low, Medium, High} $= T\{L,\ M,\ H\}$, and the term set for v is $T(v) = T\{\text{Slow, Fast}\} = T\{S,\ F\}$. A trapezoidal function $g(x; x_0, x_1, a_0, a_1)$ is chosen to implement the membership function, which is given by

$$g(x; x_0, x_1, a_0, a_1) = \begin{cases} \frac{x-x_0}{a_0} + 1, & \text{for } x_0 - a_0 < x \leq x_0 \\ 1, & \text{for } x_0 < x \leq x_1 \\ \frac{x_1-x}{a_1} + 1, & \text{for } x_1 < x \leq x_1 + a_1 \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where $x_0(x_1)$ in $g(\cdot)$ is the left (right) edge of the trapezoidal function and $a_0(a_1)$ is the left (right) width of the trapezoidal function.

Denote $\mu_{ME}(Q_0(t))(\mu_{ME}(Q_i(t)))$, $\mu_{SE}(Q_0(t))(\mu_{SE}(Q_i(t)))$, and $\mu_{NE}(Q_0(t))(\mu_{NE}(Q_i(t)))$ as the membership functions for ME, SE, and NE in $T(Q_0(t))(T(Q_i(t)))$, respectively, and define $\mu_{ME}(Q_0(t))$, $\mu_{SE}(Q_0(t))$, $\mu_{NE}(Q_0(t))$, $\mu_{ME}(Q_i(t))$, $\mu_{SE}(Q_i(t))$, and $\mu_{NE}(Q_i(t))$ as

$$\mu_{ME}(Q_0(t)) = g(Q_0(t); E_{m0}, R_{m0}, E_{w0}, 0) \tag{3}$$
$$\mu_{SE}(Q_0(t)) = g(Q_0(t); S_0, S_0, S_{lw0}, S_{rw0}) \tag{4}$$
$$\mu_{NE}(Q_0(t)) = g(Q_0(t); 0, R_{n0}, 0, NE_{w0}) \tag{5}$$
$$\mu_{ME}(Q_i(t)) = g(Q_i(t); E_{mi}, R_{mi}, E_{wi}, 0) \tag{6}$$
$$\mu_{SE}(Q_i(t)) = g(Q_i(t); S_i, S_i, S_{lwi}, S_{rwi}) \tag{7}$$
$$\mu_{NE}(Q_i(t)) = g(Q_i(t); 0, R_{ni}, 0, NE_{wi}). \tag{8}$$

The maximum possible *"More Enough"* value of available resource $R_{m0}(R_{mi})$ would be the sum of buffer size and allocation channels, $E_{m0}(E_{mi})$ would be a safety margin of available resource in macrocell (microcells) in QoS requirement and traffic fluctuation, $R_{n0}(R_{ni})$ would be set to be a fraction of available resource in macrocell (microcells), $S_0 = (1/2)(E_{m0} + R_{n0})(S_i = (1/2)(E_{mi} + R_{ni}))$, and $E_{w0} = S_{rw0} = (E_{m0} - S_0)(E_{wi} = S_{rwi} = (E_{mi} - S_i))$ and $NE_{w0} = S_{lw0} = (S_0 - R_{n0})(NE_{wi} = S_{lwi} = (S_i - R_{ni}))$ are provided to tolerate the change of traffic in macrocell (microcells).

Denote $\mu_L(P_{H0}(t))$, $\mu_M(P_{H0}(t))$ and $\mu_H(P_{H0}(t))(\mu_L(P_{Hi}(t))$, $\mu_M(P_{Hi}(t))$, and $\mu_H(P_{Hi}(t)))$ to be the membership functions for L, M, and H in $T(P_{H0}(t))(T(P_{Hi}(t)))$, respectively, and

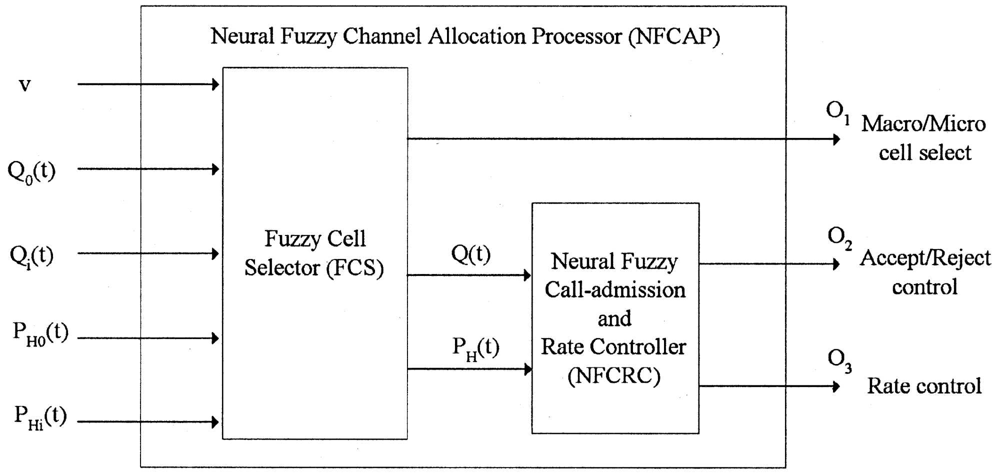

Fig. 2. The block diagram of NFCAP.

TABLE II
THE INFERENCE RULES FOR THE OVERLAY REGION

IF					THEN	IF					THEN
$Q_0(t)$	$Q_i(t)$	$P_{H0}(t)$	$P_{Hi}(t)$	v	O_1	$Q_0(t)$	$Q_i(t)$	$P_{H0}(t)$	$P_{Hi}(t)$	v	O_1
ME	SE	–	–	–	M_a	SE	ME	–	–	–	M_i
ME	NE	–	–	–	M_a	NE	ME	–	–	–	M_i
SE	NE	–	–	–	M_a	NE	SE	–	–	–	M_i
ME	ME	L	M	–	M_a	ME	ME	M	L	–	M_i
ME	ME	L	H	–	M_a	ME	ME	H	L	–	M_i
ME	ME	M	H	–	M_a	ME	ME	H	M	–	M_i
SE	SE	L	M	–	M_a	SE	SE	M	L	–	M_i
SE	SE	L	H	–	M_a	SE	SE	H	L	–	M_i
SE	SE	M	H	–	M_a	SE	SE	H	M	–	M_i
NE	NE	L	M	–	M_a	NE	NE	M	L	–	M_i
NE	NE	L	H	–	M_a	NE	NE	H	L	–	M_i
NE	NE	M	H	–	M_a	NE	NE	H	M	–	M_i
ME	ME	L	L	F	M_a	ME	ME	L	L	S	M_i
ME	ME	M	M	F	M_a	ME	ME	M	M	S	M_i
ME	ME	H	H	F	M_a	ME	ME	H	H	S	M_i

define $\mu_L(P_{H0}(t))$, $\mu_M(P_{H0}(t))$, $\mu_H(P_{H0}(t))$, $\mu_L(P_{Hi}(t))$, $\mu_M(P_{Hi}(t))$, and $\mu_H(P_{Hi}(t))$ as

$$\mu_L\left(P_{H0}(t)\right) = g\left(P_{H0}(t); 0, L_{e0}, 0, L_{w0}\right) \quad (9)$$

$$\mu_M\left(P_{H0}(t)\right) = g\left(P_{H0}(t); M_{e0}, M_{e0}, M_{lw0}, M_{rw0}\right) \quad (10)$$

$$\mu_H\left(P_{H0}(t)\right) = g\left(P_{H0}(t); H_{e0}, 1, H_{w0}, 0\right) \quad (11)$$

$$\mu_L\left(P_{Hi}(t)\right) = g\left(P_{Hi}(t); 0, L_{ei}, 0, L_{wi}\right) \quad (12)$$

$$\mu_M\left(P_{Hi}(t)\right) = g\left(P_{Hi}(t); M_{ei}, M_{ei}, M_{lwi}, M_{rwi}\right) \quad (13)$$

$$\mu_H\left(P_{Hi}(t)\right) = g\left(P_{Hi}(t); H_{ei}, 1, H_{wi}, 0\right). \quad (14)$$

$H_{e0}(H_{ei})$ would be set to be P_H^* provided to guarantee the QoS requirement in macrocell (microcells), $L_{e0}(L_{ei})$ would be set to be a safety margin of the handoff failure probability in QoS requirement in macrocell (microcells), $M_{e0} = (1/2)(H_{e0} + L_{e0})(M_{ei} = (1/2)(H_{ei} + L_{ei}))$, and $L_{w0} = M_{lw0} = M_{e0} - L_{e0}(L_{wi} = M_{lwi} = M_{ei} - L_{ei})$ and $H_{w0} = M_{rw0} = H_{e0} - M_{e0}(H_{wi} = M_{rwi} = H_{ei} - M_{ei})$ are provided to tolerate the dynamic behavior of the handoff failure probability in macrocell (microcells).

The membership functions for terms S and F in v, denoted by $\mu_S(v)$ and $\mu_F(v)$, are given by

$$\mu_S(v) = g(v; 0, S_e, 0, S_w) \quad (15)$$

$$\mu_F(v) = g(v; F_e, F_h, F_w, 0) \quad (16)$$

where $S_e(F_e)$ would be a fraction of slow (fast) speed of mobile user, $S_w(F_w)$ is provided to tolerate the change of slow (fast) speed, and F_h would be the fastest speed.

There are different call types in hierarchical cellular systems. For calls that can use only macrocell channels, FCS has to choose the macrocell, and send $P_{H0}(t)$ and $Q_0(t)$ to NFCRC. For calls that could use channels either in macrocell or microcell, FCS determines the serving cell according to input linguistic variables of $Q_0(t)$, $Q_i(t)$, $P_{H0}(t)$, $P_{Hi}(t)$, and v. The output linguistic variable $O_1 = M_a$ if the macrocell is assigned and $O_1 = M_i$ if the microcell is allocated. $T(O_1) = \{M_a, M_i\}$. The fuzzy rule base with dimension $|T(Q_0(t))| \times |T(Q_i(t))| \times |T(P_{H0}(t))| \times |T(P_{Hi}(t))| \times |T(v)|$ is shown in Table II, where $|T(\cdot)|$ denotes the number of terms in $T(\cdot)$.

The design idea of the fuzzy rule structure listed in Table II is described as follows. If the available resource in macrocell $(Q_0(t))$ is larger than that in microcell $(Q_i(t))$, the call would be directed toward the macrocell ($O_1 = M_a$), and vice versa ($O_1 = M_i$). If the available resources in both macrocell and microcell have the same fuzzy terms in the premises of the fuzzy rule, the call will be directed to a cell with *low* handoff failure probability ($P_{H0}(t)$ or $P_{Hi}(t)$), instead of the one with *high* handoff failure probability. If the available resource and handoff failure probability in both macrocell and microcell have the same fuzzy terms in the premises of the fuzzy rule, then the call is to be biased toward macrocell if the speed (v) is *fast*, for lessening frequent handoff, and vice versa. Note that the symbol "-" in the table represents no impact on the output of the fuzzy cell selector.

Membership functions for M_a and M_i in $T(O_1)$ are defined as

$$\mu_{Ma} = g(O_1; 0, 0, 0, 0) \quad (17)$$

$$\mu_{Mi} = g(O_1; 1, 1, 0, 0). \quad (18)$$

We adopt the *max-min* inference method and apply the *center-of-area* defuzzification method for output variable O_1 [9], which are not further described here. There are $P_H(t)$ and $Q(t)$ output to NFCRC determined by O_1: if the call is with channels in the macrocell $O_1 = 0$, $P_H(t) = P_{H0}(t)$ and $Q(t) = Q_0(t)$; otherwise, the call is with channels in the microcell $O_1 = 1$, and $P_H(t) = P_{Hi}(t)$, $Q(t) = Q_i(t)$.

B. Neural Fuzzy Call-Admission and Rate Controller (NFCRC)

The NFCRC takes the handoff failure probability $P_H(t)$ and available resource $Q(t)$ as input linguistic variables. The handoff failure probability shows the QoS, and the available resource implicates the traffic load intensity. This is a feedback control system that the handoff failure probability acts as a QoS index feedback to indicate how effectively the NFCRC is managing the radio resource.

We adopt a five-layer neural fuzzy controller to design the NFCRC. The best structure of NFCRC can be obtained via structure learning, which measures the degree of fuzzy similarity and decides the size of the fuzzy partition of the linguistic [17], [18]. Usually, a hybrid learning algorithm is applied to construct the NFCRC. The algorithm is a two-phase learning scheme. In phase I, a self-organized learning scheme is used to construct the presence of rules and to locate the initial membership functions; in phase II, a reinforcement learning scheme is used to optimally adjust the membership functions for desired outputs. To initiate the learning process, the size of the term set for each input/output linguistic variable, the fuzzy control rules, and training data must be provided. In the self-organized training phase, the initial structure of the controller could be constructed via Kohonen's feature-maps algorithm and the *N nearest neighbors* scheme [19] to provide a rough estimate of the structure, if the controller is not provided with an initial knowledge base. However, in this paper, we construct an initial form of the controller based on the domain knowledge obtained from the fuzzy channel allocation control scheme proposed in [9]. Only a slight modification for the structure is needed in the self-organized training phase.

The rule structure for NFCRC is shown in Table III. The design strategy in Table III is that if the handoff failure probability is *Low (L)* or *Medium (M)* and the available resource is not *Not Enough (NE)*, the call would have a chance to enter the system; if the available resource is *Not Enough (NE)*, the call would be *Rejected (R)* or *Weakly Rejected (WR);* if the handoff failure probability is *High (H)* and the available resource is *More Enough (ME)*, the call would be *Weakly Accepted (WA)* for increasing channel utilization. If the handoff failure probability is *High (H)* or available resource is *Not Enough (NE)*, the call would be allocated *Basic Rate (BR);* if the handoff failure probability is *Medium (M)* and the available resource is *Slightly Enough (SE)*, the call would be allocated *Low Medium Rate (LM);* if the handoff failure probability is *Medium (M) (Low (L))* and the available resource is *More Enough (ME) (Slightly Enough (SE))*, the call would be allocated *High Medium Rate (HM);* if the handoff failure probability is *Low (L)* and the available resource is *More Enough (ME)*, the call would be allocated *High Rate (HR)*.

TABLE III
THE INFERENCE RULES FOR NFCRC

RULE	IF		THEN	
	$P_H(t)$	$Q(t)$	O_2	O_3
1	H	ME	WA	BR
2	H	SE	WR	BR
3	H	NE	R	BR
4	M	ME	A	HM
5	M	SE	A	LM
6	M	NE	WR	BR
7	L	ME	A	HR
8	L	SE	A	HM
9	L	NE	WR	BR

The connectionist structure of the NFCRC is constructed in Fig. 3. The NFCRC has the nodes in layer one as input linguistic nodes. It has two pairs of nodes in layer five, where each pair of output nodes has two kinds of linguistic nodes. One is for feeding training data (desired output) into the net and the other is for pumping decision signals (actual output) out of the net. The nodes in layer two and layer four are term nodes, which act as membership functions of the respective linguistic variables. The nodes in layer three are rule nodes; each node represents one fuzzy rule and all nodes form a fuzzy rule base. The links in layer three and layer four function as an inference engine; layer-three links define preconditions of the rule nodes and layer-four links define consequences of the rule nodes. The links in layer two and layer five are fully connected between the linguistic nodes and their corresponding term nodes.

NFCRC has a net input function $f_i^{(k)}(u_{ij}^{(k)})$ and an activation output function $a_i^{(k)}(f_i^{(k)})$ for node i in layer k, where $u_{ij}^{(k)}$ denotes the input to node i in layer k from node j in layer $(k-1)$. The layers are described in the following.

Layer 1: In this layer, there are two input nodes with the respective input linguistic variables $P_H(t)$ and $Q(t)$.

Define

$$f_i^{(1)}\left(u_i^{(1)}\right) = u_i^{(1)} \text{ and } a_i^{(1)} = f_i^{(1)},\ 1 \le i \le 2 \quad (19)$$

where $u_1^{(1)} = P_H(t)$ and $u_2^{(1)} = Q(t)$.

Layer 2: The nodes in this layer are used as the fuzzifier. The term set used to describe the handoff failure probability is defined as $T(P_H(t)) =$ {Low (L), Medium (M), High (H)}. And the term set for the available resource is defined as $T(Q(t)) =$ {More Enough (ME), Slightly Enough (SE), Not Enough (NE)}. Thus we have six nodes in this

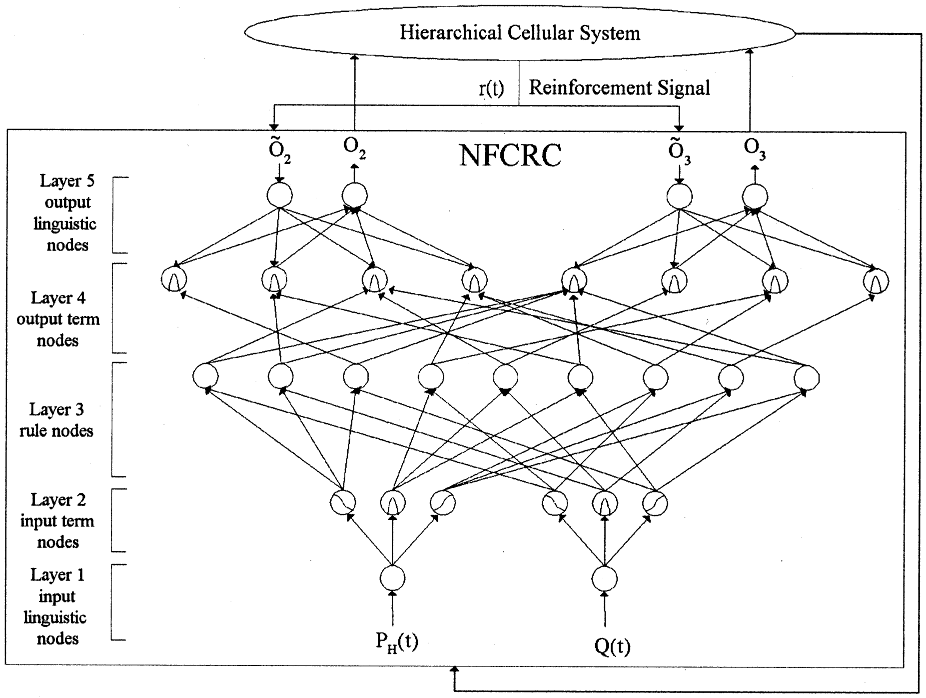

Fig. 3. The structure of the NFCRC controller.

layer. Each node performs a bell-shaped function defined as

$$f_i^{(2)}\left(u_{ij}^{(2)}\right) = -\frac{\left(u_{ij}^{(2)} - m_{jn}^{(I)}\right)^2}{\sigma_{jn}^{(I)2}} \text{ and } a_i^{(2)} = e^{f_i^{(2)}},\ 1 \le i \le 6 \tag{20}$$

where $u_{ij}^{(2)} = a_j^{(1)}$, $j = \lfloor (i+2)/3 \rfloor$, and $m_{jn}^{(I)}$ and $\sigma_{jn}^{(I)}$ are the mean and the standard deviation of the nth term of the input linguistic variable from node j in the input layer, respectively, $n = i$ if $i \le 3$ and $n = i - 3$ if $i > 3$.

Layer 3: The links perform precondition matching of fuzzy control rules. According to fuzzy set theory, the fuzzy rule base forms a fuzzy set with dimensions $|T(P_H(t))| \times |T(Q(t))|$. Thus, there are nine rule nodes in this layer. And each rule node performs the fuzzy AND operation defined as

$$f_i^{(3)}\left(u_{ij}^{(3)}\right) = \min\left(u_{ij}^{(3)}; \forall j \in P_i\right) \text{ and } a_i^{(3)} = f_i^{(3)},\quad 1 \le i \le 9 \tag{21}$$

where $u_{ij}^{(3)} = a_j^{(2)}$ and $P_i = \{j|$ all j that are precondition nodes of the ith rule.

Layer 4: There are two groups of output in this layer: one group for the output of admission control O_2 and the other group for the output rate control O_3. Nodes in this layer have two operating modes: *down-up* and *up-down*. In the down-up operating mode, the links perform consequence matching of fuzzy control rules. In order to provide a soft admission decision, the term set of the output linguistic variable O_2 is defined as $T(O_2) = \{$Reject (R), Weakly Reject (WR), Weakly Accept (WA), Accept (A)$\}$. Similarly, the term set of the output linguistic variable O_3 is defined as $T(O_3) =$ Basic Rate (BR), Low Medium Rate (LM), High Medium Rate (HM), High Rate (HR)$\}$. Thus, there are eight nodes in this layer. And each node performs a fuzzy OR operation, which integrates the fired strength of rules that have the same consequence and is defined as

$$f_i^{(4)}\left(u_{ij}^{(4)}\right) = \max\left(u_{ij}^{(4)} \forall j \in C_i\right) \text{ and } a_i^{(4)} = f_i^{(4)}\quad 1 \le i \le 8 \tag{22}$$

where $u_{ij}^{(4)} = a_j^{(3)}$ and $C_i = \{j|$all j that have the same consequence of the ith term in the term set of O_2 and$O_3\}$. The up-down operating mode is used during learning periods, which will be described later.

Layer 5: There are two pairs of nodes in this layer. One node in each pair performs the down-up operation for the decision signals O_2 and O_3. The node and its links act as the defuzzifier. The function used to simulate a center-of-area defuzzification method for O_2 signal is approximated by

$$\begin{aligned} f_i^{(5)}\left(u_{ij}^{(5)}\right) &= \sum_{j=1}^{4} m_j^{(O)} \sigma_j^{(O)} u_{ij}^{(5)} \text{ and } a_i^{(5)} \\ &= U\left(\frac{f_i^{(5)}}{\sum_{j=1}^{4} \sigma_j^{(O)} u_{ij}^{(5)}} - \theta\right),\ i = 1 \end{aligned} \tag{23}$$

where $u_{ij}^{(5)} = a_j^{(4)}$, θ is the decision threshold, and

$$U(x) = \begin{cases} 1, & \text{if } x \ge 0 \\ 0, & \text{otherwise.} \end{cases} \tag{24}$$

Clearly, $O_2 = a_1^{(5)}$, and a new connection will be accepted only if $O_2 = 1$. Similarly, the O_3 signal is also to simulate a center-of-area defuzzification method approximated by

$$f_i^{(5)}\left(u_{ij}^{(5)}\right) = \sum_{j=5}^{8} m_j^{(O)} \sigma_j^{(O)} u_{ij}^{(5)} \text{ and } a_i^{(5)} = \frac{f_i^{(5)}}{\sum_{j=5}^{8} \sigma_j^{(O)} u_{ij}^{(5)}} \times R_{dx},\ i = 2 \tag{25}$$

where $u_{ij}^{(5)} = a_j^{(4)}$; R_{dx} is the number of desired channels for a call; $x = v$ denotes the voice call; and $x = d$ denotes the data call. Clearly, $O_3 = \lceil a_2^{(5)} \rceil$ and a new connection is assigned to use a number of O_3 channels. The other node performs the up-down operation during the learning period.

The procedure to locate the mean m_i of the ith membership function for linguistic variable x, $1 \leq i \leq M$, is described below, given a set of training data x_j for x, $1 \leq j \leq N$. It employs the statistical clustering technique of Kohonen's feature-maps algorithm [19]. This is the initial definitions of membership functions required to drive the reinforcement learning algorithm.

```
Obtain m_i by using Kohonen's feature-maps algorithm
Step 1: Set initial values of m_i for all
  membership functions, 1 ≤ i ≤ M, such
  that

      min_{1≤i≤N} x_j ≤ m_i ≤ max_{1≤i≤N} x_j.

Set an initial learning rate α (0 < α < 1).
Step 2: Set j = 1.
Step 3: Present training data x_j and com-
  pute the distance d_i = |x_j − m_i|, 1 ≤ i ≤ M.
Step 4: Determine the kth membership func-
  tion that has the minimum distance d_k
  (d_k = min_{1≤i≤M} d_i). Update m_k by

      m_k = m_k + α(x_j − m_k).

Step 5: If j < N, j = j + 1, Goto Step 3
ELSE
  Decrease α and Goto Step 2.
EndIf
```

The above procedure will stop until $\alpha \leq 0$. The determination of which d_i is minimum at Step 4 can be quickly accomplished in constant time via a winner-take-all circuit [19]. The adaptive algorithm can be independently performed to obtain m_i for each input and output linguistic variable.

As for the corresponding standard deviations σ_i of the ith membership function of x, since m_i and σ_i will be finely tuned in the reinforcement learning phase, we just use a first nearest-neighbor heuristic to estimate σ_i, which is given by

$$\sigma_i = \frac{|m_i - m^*|}{\gamma} \tag{26}$$

where

$$m^* = \begin{cases} m_{i-1}, & \text{for } |m_i - m_{i-1}| < |m_i - m_{i+1}| \\ m_{i+1}, & \text{otherwise} \end{cases} \tag{27}$$

and γ is called an overlap parameter used to describe the degree of overlapping with two membership functions.

C. Reinforcement Learning Algorithm

Since there are no measurable output values fed back to instruct the NFCRC to learn, a reinforcement learning algorithm is adopted and an evaluative handoff failure probability is used as a reinforcement signal. Fig. 3 also shows the diagram of the reinforcement learning for NFCRC, where the hierarchical cellular system provides the reinforcement signal $r(t)$ as a desired output to NFCRC and receives the call admission control value O_2 and rate control value O_3 from NFCRC. The reinforcement signal is here defined as

$$r(t) = P_H^* - P_H(t) \tag{28}$$

where P_H^* denotes the QoS requirement of the desired handoff failure probability and $P_H(t)$ is the actually measured handoff failure probability at time t.

The reinforcement learning is applied to adjust parameters of input and output membership functions optimally, according to the input training data, the reinforcement signal, and the fuzzy logic rules. It derives updating rules for the mean and the standard deviation of the bell-shaped membership functions so as to minimize the error function, defined as

$$E(t) = \frac{1}{2} r^2(t) = \frac{1}{2}\left(P_H^* - P_H(t)\right)^2. \tag{29}$$

For each training data set, starting at the input nodes, the *down-up* operation can compute to obtain the actual outputs of call admission control O_2 and rate control O_3, and consequently $P_H(t)$ is measured. On the other hand, from the output node, the *up-down* operation is used to compute $(\partial E(t)/\partial w(t))$ for all hidden nodes, where $w(t)$ is the adjustable parameters such as the mean and the standard deviation for the input and output bell-shaped membership functions. We adopt the general learning rule

$$w(t+1) = w(t) + \eta \cdot \left(-\frac{\partial E(t)}{\partial w(t)}\right) \tag{30}$$

where η is the learning rate. In the following, we show the computations of $(\partial E(t)/\partial w(t))$ layer by layer, starting at the output nodes, and use the bell-shaped membership functions with centers m_i and the width σ_i as the adjustable parameters for these computations.

Layer 5: The updating rule for $m_j^{(O)}$ can be obtained by

$$m_j^{(O)}(t+1) = m_j^{(O)}(t) + \eta \cdot r(t) \cdot \frac{\sigma_j^{(O)} u_{ij}^{(5)}}{\sum_j \sigma_j^{(O)} u_{ij}^{(5)}} \tag{31}$$

the updating rule for $\sigma_j^{(O)}$ by

$$\sigma_j^{(O)}(t+1) = \sigma_j^{(O)}(t) + \eta \cdot r(t) \cdot \frac{m_j^{(O)} u_{ij}^{(5)} \left(\sum_j \sigma_j^{(O)} u_{ij}^{(5)}\right) - \left(\sum_j m_j^{(O)} \sigma_i^{(O)} u_{ij}^{(5)}\right) \left(u_{ij}^{(5)}\right)}{\left(\sum_j \sigma_j^{(O)} u_{ij}^{(5)}\right)^2}. \quad (32)$$

An error signal in this layer $\delta^{(5)}$, propagated to the proceeding layer, is given by

$$\delta^{(5)} = r(t). \quad (33)$$

Layer 4: In this layer, only the error signal $\delta_i^{(4)}$ needs to be computed. $\delta^{(4)}$ is derived as

$$\delta_i^{(4)} = r(t) \cdot \frac{m_i^{(O)} \sigma_i^{(O)} \left(\sum_i \sigma_i^{(O)} u_i^{(5)}\right) - \left(\sum_i m_i^{(O)} \sigma_i^{(O)} u_i^{(5)}\right) \left(\sigma_i^{(O)}\right)}{\left(\sum_i \sigma_i^{(O)} u_i^{(5)}\right)^2}. \quad (34)$$

Layer 3: As in layer 4, only the error signal $\delta_i^{(3)}$ needs to be computed as

$$\delta_i^{(3)} = \sum_i \delta_i^{(4)}. \quad (35)$$

Layer 2: The adaptive rule of m_i is derived as

$$m_{ij}^{(I)}(t+1) = m_{ij}^{(I)}(t) + \eta \cdot \delta_i^{(2)} \cdot e^{f_i} \cdot \frac{2\left(u_i^{(2)} - m_{ij}^{(I)}\right)}{\sigma_{ij}^{(I)2}} \quad (36)$$

and the adaptive rule of σ_{ij} becomes

$$\sigma_{ij}^{(I)}(t+1) = \sigma_{ij}^{(I)}(t) + \eta \cdot \delta_i^{(2)} \cdot e^{f_i} \cdot \frac{2\left(u_i^{(2)} - m_{ij}^{(I)}\right)^2}{\sigma_{ij}^{(I)3}} \quad (37)$$

where $\delta_i^{(2)} \triangleq -\left(\partial E(t)/\partial a_i^{(2)}\right) = -\sum_k q_k$. $q_k = -\delta_k^{(3)}$ if $a_i^{(2)}$ is minimum in the kth rule node's inputs; $q_k = 0$, otherwise.

IV. Simulation Results and Discussions

In the simulations, a hierarchical cellular system with $N = 9$ microcells constructed along the Manhattan streets is assumed, and the handoff behavior of users is characterized by a teletraffic flow matrix [3], defined as shown in the equation at the bottom of the page, where a_{ij}, $i \neq j$, represents the probability of a handoff call originated in cell i and directed to cell j, $1 \leq j \leq N$, and a_{id} denotes the probability of this handoff call directed to the adjacent macrocell, $0 \leq i \leq N$. $\sum_{j=0} a_{ij} = 1$ and a_{ii} would be zero.

The number of mobile stations in each cell is assumed to be 550, and $\lambda_{nv} = 0.8\lambda$, $\lambda_{nd} = 0.2\lambda$. Suppose $R_{rv} = 1$ and $R_{dv} = 1$ for voice calls and $R_{rd} = 1$ and $R_{dd} = 4$ for data calls. Low- and high-mobility users are generated in a ratio of 7:3, and the cell dwell time is exponentially distributed with mean 180 s (18 s) and 1440 s (144 s) for high- and low-mobility users in macrocell (microcells), respectively. The speed of mobile users is assumed to be uniformly distributed in the range of 0–40 km (40–80 km) for low- (high-) mobility users. We also assume that the mean unencumbered session duration is $1/\mu_v = 100\,\text{s}$ for voice call and $1/\mu_d = 60\,\text{s}$ for data call, and the patience time for queued voice (data) calls is in the range of 5–20 s. One hundred fifty channels are fixedly allocated to macrocell and microcells with a pattern of $(C_0, C_1, \ldots, C_N) = (42, 12, \ldots, 12)$. If the OCA and CCA schemes are applied, the system reserves a number of channels C_{ri} as guard channels for handoff calls in cell i, $0 \leq i \leq N$, which are denoted by $(C_{r0}, C_{r1}, \cdots, C_{rN})$. We do some simulations and obtain the appropriate $(C_{r0}, C_{r1}, \cdots, C_{rN}) = (8, 4, \cdots, 4)$ for OCA scheme and $(C_{r0}, C_{r1}, \ldots, C_{rN}) = (3, 2, \ldots, 2)$ for CCA schemes at $\lambda = 5 \times 10^{-4}$. Since the reneging (dropping) process is considered, it is not necessary to provide a large buffer size for new and handoff calls [9], [20], [21]; all buffer sizes in macrocell and microcells are assumed to be three. Note that in the following performance comparison, the OCA scheme provides no buffer and the CCA and FCAC schemes support the same buffering scheme and capacity as NFRM does.

$$A = \begin{bmatrix} a_{00} & a_{01} & a_{02} & \cdots & a_{0N} & a_{0d} \\ a_{10} & a_{11} & a_{12} & \cdots & a_{1N} & a_{1d} \\ a_{20} & a_{21} & a_{22} & \cdots & a_{2N} & a_{2d} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{N0} & a_{N1} & a_{N2} & \cdots & a_{NN} & a_{Nd} \end{bmatrix}$$

$$= \begin{bmatrix} 0.0 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.0 & 0.3 & 0.0 & 0.3 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.3 \\ 0.1 & 0.2 & 0.0 & 0.2 & 0.0 & 0.2 & 0.0 & 0.0 & 0.0 & 0.0 & 0.3 \\ 0.1 & 0.0 & 0.3 & 0.0 & 0.0 & 0.0 & 0.3 & 0.0 & 0.0 & 0.0 & 0.3 \\ 0.1 & 0.2 & 0.0 & 0.0 & 0.0 & 0.2 & 0.0 & 0.2 & 0.0 & 0.0 & 0.3 \\ 0.0 & 0.0 & 0.25 & 0.0 & 0.25 & 0.0 & 0.25 & 0.0 & 0.25 & 0.0 & 0.0 \\ 0.1 & 0.0 & 0.0 & 0.2 & 0.0 & 0.2 & 0.0 & 0.0 & 0.0 & 0.2 & 0.3 \\ 0.1 & 0.0 & 0.0 & 0.0 & 0.3 & 0.0 & 0.0 & 0.0 & 0.3 & 0.0 & 0.3 \\ 0.1 & 0.0 & 0.0 & 0.0 & 0.0 & 0.2 & 0.0 & 0.2 & 0.0 & 0.2 & 0.3 \\ 0.1 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.3 & 0.0 & 0.3 & 0.0 & 0.3 \end{bmatrix}$$

Based upon the QoS requirement and knowledge of CCA [20], [21] and FCAC [9] mechanisms, parameters of membership functions for input linguistic variables in the FCS are selected as follows: $L_{e0} = L_{ei} = 0$, $L_{w0} = L_{wi} = M_{e0} = M_{ei} = M_{lw0} = M_{lwi} = 0.01$, $H_{e0} = H_{ei} = 0.02$, and $M_{rw0} = M_{rwi} = H_{w0} = H_{wi} = 0.01$ for $\mu_L(P_{H0}(t))$, $\mu_M(P_{H0}(t))$, $\mu_H(P_{H0}(t))$, $\mu_L(P_{Hi}(t))$, $\mu_M(P_{Hi}(t))$, and $\mu_H(P_{Hi}(t))$ in (9)–(14); $E_{m0} = 12$, $R_{m0} = 45$, $E_{w0} = S_0 = S_{lw0} = S_{rw0} = NE_{w0} = 6$, and $R_{n0} = 0$, for $\mu_{ME}(Q_0(t))$, $\mu_{SE}(Q_0(t))$, and $\mu_{NE}(Q_0(t))$ in (3)–(5); $E_{mi} = 10$, $R_{mi} = 15$, $E_{wi} = S_0 = S_{lwi} = S_{rwi} = NE_{wi} = 5$, and $R_{ni} = 0$, for $\mu_{ME}(Q_i(t))$, $\mu_{SE}(Q_i(t))$, and $\mu_{NE}(Q_i(t))$ in (6)–(8).

In NFCAP, the initial values of membership functions of term sets for $P_H(t)$ are chosen according to QoS requirement and then properly adjusted via the learning algorithm. Thus, the mean value $m_{11}^{(I)}(m_{12}^{(I)}, m_{13}^{(I)})$ in the membership function of $H(M, L)$ of $P_H(t)$ is set to be 0.05 (0.02, 0), and let $\sigma_{11}^{(I)} = (1/2)\cdot(m_{11}^{(I)} - m_{12}^{(I)})(\sigma_{12}^{(I)} = \sigma_{13}^{(I)} = (1/2)\cdot(m_{12}^{(I)} - m_{13}^{(I)}))$. In order to utilize the resource as much as possible and to guarantee the QoS requirement, the initial values of membership functions of ME, SE, and NE for $Q(t)$ are set to be $m_{21}^{(I)} = 9$, $m_{22}^{(I)} = 3$, and $m_{23}^{(I)} = 0$ ($m_{21}^{(I)} = 7$, $m_{22}^{(I)} = 3$, and $m_{23}^{(I)} = 0$) if the call is assigned to use the channels in macrocell (microcell), and let $\sigma_{21}^{(I)} = (1/2)\cdot(m_{21}^{(I)} - m_{22}^{(I)})$ and $\sigma_{22}^{(I)} = \sigma_{23}^{(I)} = (1/2)\cdot(m_{22}^{(I)} - m_{23}^{(I)})$.

The initial membership functions of the mean $m_j^{(O)}$ of the term set $O_2(O_3)$ are set to be equally spaced in the range of [0,1], and let $\sigma_j^{(O)} = 0.01$. The decision threshold θ in (23) is set to be $\theta = 0$ for handoff call and $\theta = 0.5$ for new call because handoffs are given higher priority than new calls. The use of different η may drastically reduce the training time required in the learning phase. As for $P_H(t)$ and $Q(t)$, their initial membership functions were heuristically set and required further optimization in the learning phase. Thus, $\eta = 0.01$ was used.

Five performance measures such as the system utilization, the new-call blocking probability, the handoff failure probability, the forced termination probability, and the handoff rate are observed. The system utilization at time t, denoted by $U(t)$, is defined as

$$U(t) = \frac{K_0(t) + \sum_{i=1}^{N} K_i(t)}{C_0 + \sum_{i=1}^{N} C_i} \tag{38}$$

where $K_0(t)(K_i(t))$ is the average number of busy channels in macrocell 0 (microcell i) at time t and $C_0(C_i)$ is the channel capacity for macrocell 0 (microcell i). The new-call blocking probability at time t, denoted by $P_N(t)$, is defined as

$$P_N(t) = \frac{\sum_{i=0}^{N} (NB_i(t) + NR_i(t))}{\sum_{i=0}^{N} NN_i(t)} \tag{39}$$

where $NB_i(t)(NR_i(t))$ is the number of blocked (reneging) new calls in cell i and $NN_i(t)(NN_0(t))$ is the number of new calls originating in microcell i (macrocell-only region), at time t.Similarly, the handoff failure probability at time t, denoted by $P_H(t)$, is given by

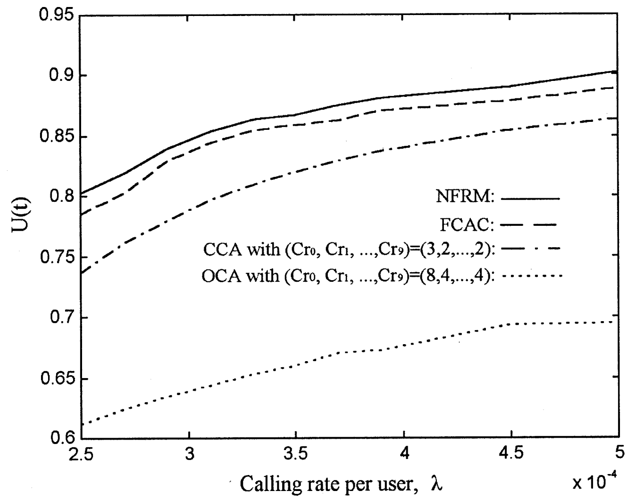

Fig. 4. $U(t)$ for NFRM, FCAC, OCA, and CCA schemes.

$$P_H(t) = \frac{\sum_{i=0}^{N} (HB_i(t) + HR_i(t))}{\sum_{i=0}^{N} NH_i(t)}. \tag{40}$$

A call will be forced termination if it is corrupted due to a handoff failure during its conversation time. The forced termination probability at time t, denoted by $P_F(t)$, is defined as

$$P_F(t) = \frac{\sum_{i=0}^{N} (HB_i(t) + HR_i(t))}{\sum_{i=0}^{N} NS_i(t)} \tag{41}$$

where $HB_i(t)(HR_i(t))$ is the number of blocked (dropped) handoff calls in cell i and $NS_i(t)$ is the number of admitted new calls originated in cell i, at time t. The handoff rate at time t, denoted by $R_H(t)$, is defined as

$$R_H(t) = \frac{\sum_{i=0}^{N} NH_i(t)}{\sum_{i=0}^{N} NS_i(t)}. \tag{42}$$

Fig. 4 shows the system utilization $U(t)$ versus the calling rate per user λ for schemes of NFRM, FCAC, OCA, and CCA at time $t = 10^8$. It reveals that the system utilization of NFRM gains 31.1%, 6.3%, and 1.4% improvement over the OCA, CCA, and FCAC methods, respectively. The superior performance of NFRM is because FCS in NFRM refers much more effective information than other conventional controllers, and it adopts fuzzy logic theory to balance traffic load between macrocell and microcells and provide a soft and accurate control during traffic fluctuation. In addition, NFCRC in NFRM possesses the learning capability of neural networks to reduce the decision error and the fuzzy logic theory to qualitatively represent control rules naturally in the neural network to overcome some uncertainty and imprecision, and NFCRC contains the rate control function, which has the flexibility of rate assignment.

Fig. 5 shows the new-call blocking probability $P_N(t)$ and the handoff failure probability $P_H(t)$ for schemes of NFRM,

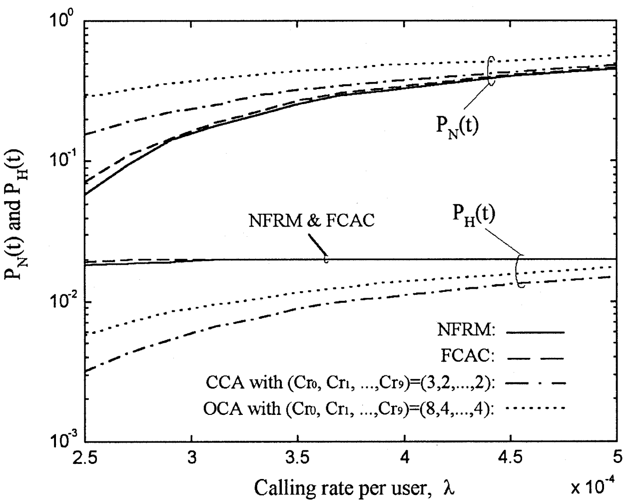

Fig. 5. $P_H(t)$ and $P_N(t)$ for NFRM, FCAC, OCA, and CCA schemes.

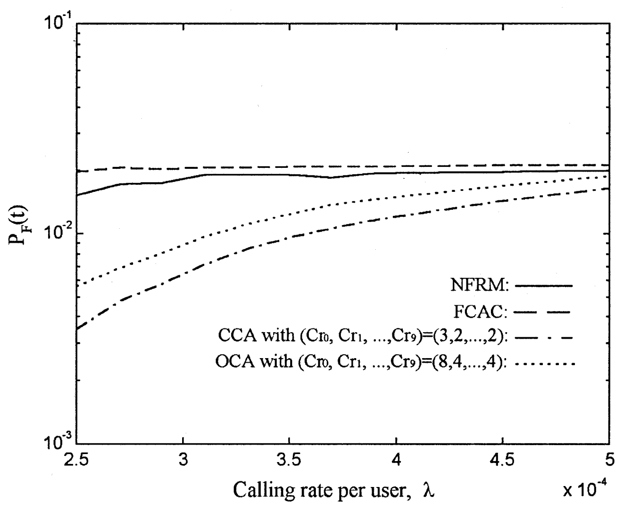

Fig. 6. $P_F(t)$ for NFRM, FCAC, OCA, and CCA schemes.

FCAC, OCA, and CCA versus the calling rate per user λ at time $t = 10^8$. It can be seen that, as λ varies, $P_H(t)$ of NFRM and FCAC remains constant at around $P_H^* = 2\%$, denoting that the system QoS requirement is guaranteed; and $P_N(t)$ of NFRM is minimal, denoting that the system utilization is maximum, compared to FCAC, OCA and CCA. This is because NFRM uses neural fuzzy control to the allocation of channels. Neural networks have merits of ability to learn from examples and to cope with incomplete input data. Fuzzy logic is a soft logic that is appropriate to represent in determining if a given requirement constraint is complied or violated. This in effect removes the imposition of worse case assumption from the decision-making of channel selection. The neural networks used in fuzzy call admission control and rate manager can effectively estimate the optimal call admission and appropriately allocate a number of channels for each call. The other schemes are inadaptive to determine the number of guard channels to maintain, but not to overprotect, the QoS requirement as the traffic load is fluctuating and the changing λ is unpredictable.

Fig. 6 shows the forced termination probability $P_F(t)$ versus the calling rate per user λ for schemes of NFRM, FCAC, OCA,

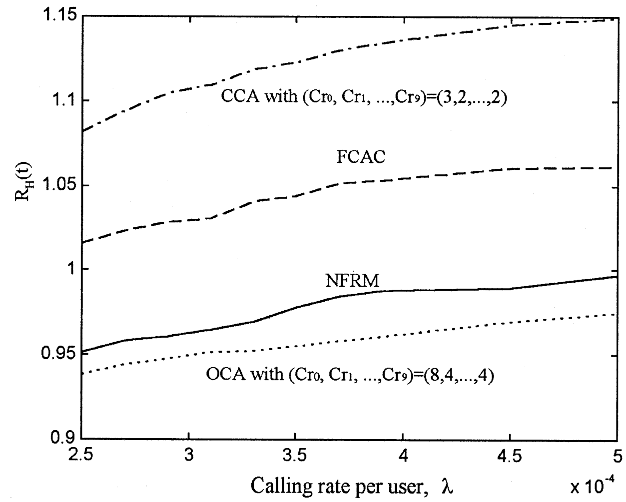

Fig. 7. $R_H(t)$ for NFRM, FCAC, OCA, and CCA schemes.

and CCA at time $t = 10^8$. It is found that $P_F(t)$ of NFRM has a flat curve under 2%. The reason is that NFRM obtains the unchanged $P_H(t)$, shown in Fig. 5.

Fig. 7 shows the handoff rate $R_H(t)$ versus the calling rate per user λ for schemes of NFRM, FCAC, OCA, and CCA at time $t = 10^8$. It reveals that NFRM has more handoff rate than OCA by an amount of 2%. The reason is that the design of NFRM is based on the knowledge of FCAC and CCA, which combines overflow, reversible, and underflow. Fortunately, the signaling overheads for these handoffs might not cost so much as those for conventional handoffs between macrocells since most of these handoffs occur in the same macrocell. It also reveals that NFRM achieves less handoff rate than CCA and FCAC by an amount of 14.9% and 6.8%, respectively. It is not only because of more information, such as the speed of mobile station considered in NFRM, but also because of the neural fuzzy logic control that can provide decision support and expert system with powerful reasoning and learning capabilities.

V. Concluding Remarks

In this paper, we propose a neural fuzzy resource manager for hierarchical cellular systems providing multimedia services. The NFRM mainly contains a neural fuzzy channel allocation processor, which is designed to be a two-layer controller. There is a fuzzy cell selector in the first layer and a neural fuzzy call-admission and rate controller in the second layer. The FCS uses soft logic to determine which cell, macrocell or microcell, to serve a call with channels. Then the NFCRC adopts a five-layer neural network with fuzzy logic control to determine whether the call is accepted or not and how many channels are allocated. Simulation results show that the proposed NFRM improves the overall channel utilization by an amount 31.1% higher than the OCA scheme, 6.3% better than the CCA scheme, and 1.4% larger than the FCAC scheme, while maintaining the QoS requirement. It still reduces the handoff rate by an amount of 14.9% under the CCA mechanism and 6.8% below the FCAC scheme, but increases the handoff rate by an amount of 2% over the OCA mechanism.

ACKNOWLEDGMENT

The authors thank the anonymous reviewers for their constructive suggestion for improving the presentation of this paper.

REFERENCES

[1] E. Dahlman, P. Beming, J. Knutsson, F. Ovesiö, M. Persson, and C. Roobol, "WDCMA—The radio interface for future mobile multimedia communications," *IEEE Trans. Veh. Technol.*, vol. 47, pp. 1105–1117, Nov. 1998.

[2] A. Samukic, "UMTS universal mobile telecommunications system: development of standards for the third generation," *IEEE Trans. Veh. Technol.*, vol. 47, pp. 1099–1104, Nov. 1998.

[3] S. S. Rappaport and L. R. Hu, "Microcellular communication systems with hierarchical microcell overlays: traffic performance models and analysis," *Proc. IEEE*, vol. 82, no. 9, pp. 1383–1397, Sept. 1994.

[4] ——, "Personal communication systems using multiple hierarchical cellular overlays," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 406–415, May 1995.

[5] C. Chang, C. J. Chang, and K. R. Lo, "Analysis of a hierarchical cellular system with reneging and dropping for waiting new and handoff calls," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 1080–1091, July 1999.

[6] C. L. I, L. J. Greenstein, and R. D. Gitlin, "A microcell/macrocell cellular architecture for low- and high-mobility wireless users," *IEEE J. Select. Areas Commun.*, vol. 11, pp. 885–891, Aug. 1993.

[7] E. Gelenbe, I. W. Habib, S. Palazzo, and C. Douligeris, Eds., *IEEE J. Select. Areas Commun. (Special Issue on Intelligent Techniques in High Speed Networks)*, Feb. 2000, vol. 18.

[8] K. Shum and C. W. Sung, "Fuzzy layer selection method in hierarchical cellular systems," *IEEE Trans. Veh. Technol.*, vol. 48, Nov. 1999..

[9] K. R. Lo, C. J. Chang, C. Chang, and B. Shung, "A QoS-guaranteed fuzzy channel allocation controller for hierarchical cellular systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1588–1598, Sept. 2000.

[10] P. R. Chang and B. C. Wang, "Adaptive fuzzy power control for CDMA mobile radio systems," *IEEE Trans. Veh. Technol.*, vol. 45, pp. 225–236, May 1996.

[11] J. S. Kim, S. H. Park, W. Dowd, and N. M. Nasrabadi, "Cellular radio channel assignment using a modified Hopfield network," *IEEE Trans. Veh. Technol.*, vol. 46, pp. 957–967, Nov. 1997.

[12] K. Smith and M Palaniswami, "Static and dynamic channel assignment using neural," *IEEE J. Select. Areas Commun.*, vol. 15, pp. 238–249, Feb. 1997.

[13] P. H. Chan, M. Palaniswami, and D. Everitt, "Neural network-based dynamic channel assignment for cellular mobile communication systems," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 279–288, May 1994.

[14] N. Funabiki and Y. Takefuji, "A neural network parallel algorithm for channel assignment problems in cellular radio networks," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 430–437, Nov. 1992.

[15] D. Kunz, "Channel assignment for cellular radio using neural networks," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 188–193, Feb. 1991.

[16] Z. Ali, A. Ghafoor, and C. S. G. Lee, "Media synchronization in multimedia web using a neuro-fuzzy framework," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 168–183, Feb. 2000.

[17] C. T. Lin and C. S. G. Lee, "Neural-network based fuzzy logic control and decision system," *IEEE Trans. Computers*, vol. 40, pp. 1320–1336, Dec. 1991.

[18] ——, "Reinforcement structure/parameter learning for neural-network-based fuzzy logic control systems," *IEEE Trans. Fuzzy Syst.*, vol. 2, pp. 46–63, Feb. 1994.

[19] ——, *Neural Fuzzy Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[20] K. R. Lo, C. J. Chang, C. Chang, and B. Shung, "A combined channel assignment strategy for hierarchical cellular systems," in *ICUPC 1997*, pp. 128–132.

[21] ——, "A combined channel assignment strategy for hierarchical cellular systems," *Comput. Commun.*, vol. 21, no. 13, pp. 1143–1152, Sept. 1998.

Kuen-Rong Lo was born in Hsinchu, Taiwan, R.O.C., in 1957. He received the B.S. degree in electronics engineering from National Taiwan University of Science and Technology, Taipei, Taiwan, in 1984, the M.S. degree in computer science from National Tsing Hua University, Hsinchu, Taiwan, in 1989, and the Ph.D. degree in electronics engineering from National Chiao Tung University, Taiwan, in 2000.

From 1984 to 1996, he was with Telecommunication Labs, Directorate General of Telecommunication (DGT), Ministry of Communications, Taiwan, as a Design Engineer,Task Leader, and then Project Leader. In July 1996, he joined the Telecommunication Laboratories Chunghwa Telecom Co., Ltd., Chung-Li, Taiwan, where he currently is in the Division of Wireless Communication Technologies. He has been a Project Manager since 1999. His research interests lie in the broad areas of mobile communication systems and mobile value-added terminal equipments. Presently, his research is focused on electronic fleet management systems and 3G terminal equipment technologies.

Chung-Ju Chang (S'81–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 National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 1972 and 1976, respectively, and the Ph.D degree in electrical engineering from National Taiwan University (NTU), Taiwan, in 1985.

From 1976 to 1988, he was with Telecommunication Laboratories, Directorate General of Telecommunications, Ministry of Communications, Taiwan, as a Design Engineer, Supervisor, Project Manager, and then Division Director. There, he was involved in designing digital switching system, RAX trunk tester, ISDN user-network interface, and ISDN service and technology trials in Science-Based Industrial Park. In the meantime, he also acted as a Science and Technical Advisor for the Minister of the Ministry of Communications from 1987 to 1989. In 1988, he joined the Faculty of the Department of Communication Engineering and Center for Telecommunications Research, College of Electrical Engineering and Computer Science, National Chiao Tung University, as an Associate Professor. He has been a Professor since 1993. He was Director of the Institute of Communication Engineering from August 1993 to July 1995 and Chairman of Department of Communication Engineering from August 1999 to July 2001. Now, he is the Dean of the Research and Development Office in NCTU. He was an Advisor for the Ministry of Education to promote the education of communication science and technologies for colleges and universities in Taiwan since 1995. He is also acting as a Committee Member of the Telecommunication Deliberate Body. His research interests include performance evaluation, wireless communication networks, and broadband networks.

Dr. Chang is a member of the Chinese Institute of Engineers (CIE).

C. Bernard Shung received the B.S. degree in electrical engineering from National Taiwan University, Taiwan, R.O.C., in 1981 and the M.S. and Ph.D. degrees in electrical engineering and computer science from the University of California, Berkeley, in 1985 and 1988, respectively.

He is currently a Senior Engineer Manager with Broadcom Corp., San Jose, CA. He was a Visiting Scientist with IBM Research Division, Almaden Research Center, San Jose, in 1988–1990, and later a Research Staff Member in 1998–1999. He was a Faculty Member in the Department of Electronics Engineering, National Chiao Tung University. Hsinchu, Taiwan, in 1990–1997. In 1994–1995, he was a Staff Engineer with Qualcomm Inc., in San Diego, CA, while taking a leave from NCTU. His research interests include VLSI architectures and integrated circuits design for communications, signal processing, and networking. He has published more than 50 technical papers in various research areas.