A QoS-Provisioning Neural Fuzzy Connection Admission Controller for Multimedia High-Speed Networks

Ray-Guang Cheng, Chung-Ju Chang, Senior Member, IEEE, and Li-Fong Lin, Student Member, IEEE

Abstract—This paper proposes a neural fuzzy approach for connection admission control (CAC) with QoS guarantee in multimedia high-speed networks. Fuzzy logic systems have been successfully applied to deal with traffic-control-related problems and have provided a robust mathematical framework for dealing with real-world imprecision. However, there is no clear and general technique to map domain knowledge on traffic control onto the parameters of a fuzzy logic system. Neural networks have learning and adaptive capabilities that can be used to construct intelligent computational algorithms for traffic control. However, the knowledge embodied in conventional methods is difficult to incorporate into the design of neural networks. The proposed neural fuzzy connection admission control (NFCAC) scheme is an integrated method that combines the linguistic control capabilities of a fuzzy logic controller and the learning abilities of a neural network. It is an intelligent implementation so that it can provide a robust framework to mimic experts’ knowledge embodied in existing traffic control techniques and can construct efficient computational algorithms for traffic control. We properly choose input variables and design the rule structure for the NFCAC controller so that it can have robust operation even under dynamic environments. Simulation results show that compared with a conventional effective-bandwidth-based CAC, a fuzzy-logic-based CAC, and a neural-net-based CAC, the proposed NFCAC can achieve superior system utilization, high learning speed, and simple design procedure, while keeping the QoS contract.

I. INTRODUCTION

HIGH-SPEED network supporting multimedia services have to be capable of handling bursty traffic and satisfying various quality-of-service (QoS) and bandwidth requirements. Therefore, a multimedia high-speed network must have an appropriate connection admission control (CAC) scheme not only to guarantee QoS for existing calls but also to achieve high system utilization. As is known asynchronous transfer mode (ATM) is one of the technologies that can integrate multimedia services for high-speed networks.

Conventional CAC schemes [1]–[5] that utilize either capacity estimation or buffer thresholds suffer from some fundamental limitations. One of the limitations is the difficulty of obtaining complete statistics on input traffic to a network. As a result, it is not easy to accurately determine the equivalent capacity or effective thresholds for multimedia high-speed networks in various bursty traffic flow conditions. Besides, these conventional schemes provide optimal solutions only under a steady state. A control scheme that dynamically regulates traffic flows according to changing network conditions, however, requires understanding of network dynamics. The rationale and principles underlying the nature and choice of thresholds or equivalent capacity under dynamic conditions are unclear [6]. Networks are forced to make decisions based on incomplete information [6] so that the decision process is full of uncertainty. Thus, because of unpredictable statistical fluctuations of the system, these control schemes will always be subject to decision error, which degrades performance.

Fuzzy logic systems have been widely employed to deal with CAC-related problems in ATM networks [7]–[9]. Fuzzy set theory appears to provide a robust mathematical framework for dealing with real-world imprecision, and the fuzzy approach exhibits a soft behavior, which means a greater ability to adapt itself to dynamic, imprecise, and bursty environments [7]–[9]. Bonde and Ghosh [7] used fuzzy mathematics to provide a flexible high-performance solution to queue management in ATM networks. Ndousse [9] proposed a fuzzy logic implementation of the leaky bucket mechanism that used a channel utilization feedback to improve performance. In [8], a fuzzy traffic controller which simultaneously incorporates CAC and congestion control was proposed. It is a fuzzy implementation of the two-threshold congestion control method and the equivalent capacity admission control method extensively studied in the literature. Comparative studies have shown that the proposed fuzzy approaches significantly improve system performance compared with conventional approaches. However, no clear and general technique has been presented to map existing knowledge on traffic control onto the design parameters of the fuzzy logic controller. Self-learning capability should be incorporated into the fuzzy logic controller to simplify the design procedure and obtain better control results.
The self-learning capability of neural networks has been applied to characterize the relationship between input traffic and system performance [10]–[13]. In [11], Hiramatsu used a neural network as a CAC. In [12], Tran-Gia and Gropp investigated the possible use of a neural network to perform CAC. In [13], Youssef, Habib, and Saadawi proposed a call admission controller for ATM networks. A neural network is trained to compute the effective bandwidth required to support MPEG-I VBR video calls with different QoS requirements. They showed that the adaptability of the neural network controller to new traffic situations had been achieved by adopting a hierarchical approach to the design. However, in most of the proposed neural-net approaches for CAC, the numbers of users for each kind of service were selected as input parameters. The dimension of neural network and the learning time would increase as the number of traffic types grows. The system complexity would increase for system upgrade. Therefore, the application of neural network to CAC is limited to a simplistic traffic environment, such as limited traffic type, simplified traffic source, etc.

Conventional, fuzzy-logic-based, and neural-net-based CAC schemes all have various benefits in handling CAC. Conventional CAC, based on mathematical analysis, provides robust solutions for different kinds of traffic environments but suffers from estimation error (due to modeling) and approximation error (due to the need to complete calculations in real time), so is not suitable for dynamic environments. Fuzzy-logic-based CAC is excellent in dealing with real-world imprecision and has a greater ability to adapt itself to dynamic, imprecise, and bursty environments, but lacks the learning capability needed to automatically construct its rule structure and membership functions so as to achieve optimal performance. Neural-net-based CAC provides learning and adaptation capabilities which reduce the estimation error of conventional CAC and achieve performance similar to that of a fuzzy logic controller. However, the knowledge embodied in conventional methods is difficult to incorporate into the design of a neural network.

This paper proposes a neural fuzzy connection admission control (NFCAC) scheme, which absorbs benefits of the three approaches while minimizing their drawbacks, for multimedia high-speed networks. The NFCAC scheme utilizes the learning capability of the neural network to reduce decision errors of conventional CAC policies resulted from modeling, approximation, and unpredictable traffic fluctuations of the system. It also employs the rule structure of the fuzzy logic controller to prevent operating errors, due to incorrect learning, and to decrease training time. Furthermore, the neural fuzzy network is a simple structured network. Here we properly choose input variables and design the rule structure for the NFCAC scheme so that it not only provides a robust framework to mimic experts’ knowledge embodied in existing traffic control techniques but also constructs intelligent computational algorithm for traffic control. Simulation results reveal that the NFCAC scheme achieves superior system utilization and high learning speed while keeping the QoS contract, compared with the effective-bandwidth-based CAC (EBCAC) [3], the fuzzy-logic-based CAC (FLCAC) [8], the neural-net-based CAC (NNCAC) [14], and the radial-basis-function-based CAC (RBFAC) schemes.

The rest of this paper is organized as follows. In Section II, the basic concepts behind a neural fuzzy controller are introduced, and an NFCAC scheme is proposed to cope with CAC-related problems in multimedia high-speed networks. Section III presents simulation results comparing the proposed NFCAC scheme with the existing effective bandwidth approach, the fuzzy logic approach, and neural net approach. Finally, some concluding remarks are given in Section IV.

II. NFCAC CONTROLLER

A. Neural Fuzzy Controller

A fuzzy set \( F \) in a universe of discourse \( U \) is characterized by a membership function which takes values in the interval \([0, 1]\). A linguistic variable \( x \) in \( U \) is defined by \( T(x) = \{T_1, T_2, \ldots, T_k\} \) and \( M(x) = \{M_1, M_2, \ldots, M_k\} \), where \( T(x) \) is a term set of \( x \), i.e., a set of terms \( T_k \) with membership function \( M_k \) defined on \( U \), and \( M(x) \) is a semantic rule for associating each term with its meaning.

A fuzzy logic controller, as shown in Fig. 1, has three functional blocks: a fuzzifier, a defuzzifier, and an inference engine containing a fuzzy rule base [15]. The fuzzifier is a mapping from an observed \( m \)-dim input \( x_i \) to fuzzy set \( T_{x_i} \) with degree \( M_{x_i} \), \( i = 1, \ldots, m \). The fuzzy rule base is a control knowledge-base characterized by a set of linguistic statements in the form of “if–then” rules that describe a fuzzy logic relationship between \( m \)-dim inputs \( x_i \) and \( n \)-dim outputs \( y_j \). The inference engine contains the decision-making logic; it acquires the input linguistic terms of \( T(x_i) \) from the fuzzifier and uses an inference method to obtain the output linguistic terms of \( T(y_j) \). The defuzzifier adopts a defuzzification function to convert \( T(y_j) \) into a nonfuzzy value that represents decision \( y_j \). The fuzzy-logic controller can incorporate domain knowledge from existing techniques.

A multilayer feedforward neural network is a layered network that consists of an input layer, an output layer, and at least one hidden layer. Each hidden layer consists of nonlinear processing elements, called nodes. Nodes in two adjacent layers are fully interconnected with variable link weights. The output of a node in one layer multiplied by the link weight becomes the input of a node in the next layer. Each node forms a weighted sum of its inputs and generates an output according to a predefined activation function \( g(\cdot) \). Consider a feedforward network \( NN(X, W) \) with input vector \( X \) and a set of weight vectors \( W \) which will be updated by some learning rules. It needs to train \( NN(X, W) \) (actual output) to approximate a desired output function \( g(X) \) as close as possible. The Stone–Weierstrass theorem [16] shows that for any continuous function \( g \in C(D) \) with respect to \( X \) and a
compact metric space $C(D)$, an NN$(X,W)$ with appropriate weight $W$ can be found so that $||NN(X,W) - q(X)||_X < \varepsilon$ for an arbitrary $\varepsilon > 0$, where $||e||_X = \sum_{X \in D} ||e(X)||^2$ and $|| \cdot ||$ is a vector norm. The neural network is a nonstructured network, which cannot incorporate knowledge about system.

A neural fuzzy network integrates a fuzzy logic system with a neural network. The integration brings the low-level learning and computational power of the neural network into the fuzzy logic system, and provides the high-level human-like thinking and reasoning of fuzzy logic system for the neural network. The neural fuzzy network generally takes the form of a multilayer network to realize a fuzzy logic system [17]. It is a structured network that can incorporate domain knowledge from conventional policies.

### B. NFCAC Controller

We adopt a five-layer neural fuzzy architecture to design the NFCAC controller. As shown in Fig. 2, the NFCAC controller has nodes in layer one as input linguistic nodes. It has two kinds of output linguistic nodes used in layer five. 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.

The NFCAC controller adopts three linguistic inputs of an available capacity $C_A$, a congestion indicator $\gamma$, and a cell loss ratio $p_l$ and outputs a decision signal $\hat{\gamma}$ to indicate acceptance or rejection of the new call request (shown in Fig. 2). The available capacity $C_A$ is the amount of remaining capacity, that is, the total capacity subtracted from those required by the new call and all of the existing calls, and is widely employed in the equivalent-bandwidth-based CAC schemes. The congestion indicator $\gamma$ is the degree of congestion currently in the network that provides more insight information of the system. And the cell loss ratio $p_l$ is the system performance feedback which can be used to provide a closed-loop control system capable of adjusting itself to provide stable and robust operation. In order to generate these input linguistic variables, some peripheral processors are designed for the NFCAC controller.

Fig. 3 shows an NFCAC controller with its peripheral processors for multimedia high-speed networks. The peripheral processors are a congestion controller, a bandwidth estimator, and a network resource estimator. The congestion controller generates a congestion indicator $\gamma$ according to the measured system statistics, such as the queue length $q_t$, the change rate of the queue length $q_t$, and the cell loss ratio $p_l$. Different congestion control algorithms could be employed to implement the congestion controller. For example, rate-based feedback congestion control approaches, which commonly take the queue length and the cell loss ratio into account, could be used. One of the most frequently used congestion control methods is the buffer threshold method, where a congestion alarm occurs whenever the queue length exceeds some predefined thresholds. Here, we adopt a fuzzy congestion controller [8] which is a fuzzy implementation of the two-threshold congestion control scheme proposed in [18]. Network congestion is then averted by regulating the traffic flow of the incoming sources according to the traffic load adjustment parameter generated by the fuzzy congestion controller. The bandwidth estimator estimates the required capacity $C_e$ for a new connection from its traffic description parameters such as the peak cell rate, sustainable cell rate, and peak cell rate duration, denoted by $R_{pp}$, $R_{cm}$, and $T_p$, respectively. It employs the equivalent-capacity-based algorithm proposed in the literature. The equivalent capacity method [1, eq. (2)] transforms the traffic characteristics (usually described by three traffic parameters: peak cell rate, sustainable cell rate, and peak cell rate duration) of a new call into a unified metric, called the equivalent bandwidth, to reduce the dependence of

![Fig. 2. The architecture of the NFCAC controller.](image)
the proposed control mechanism on the traffic type. Such a transformation can greatly reduce the number of dimensions of the NFCAC scheme and save a large percentage of learning time. Here, we adopt a fuzzy bandwidth estimator [8], which is a fuzzy implementation of the equivalent capacity method in [1]. The network resource estimator does the accounting for system-resource usage. When a new connection with bandwidth $C_a$ is accepted, the value of $C_a$ is updated by subtracting $C_a$ from the original value of $C_a$. Conversely, when an existing connection with bandwidth $C_a$ is disconnected, the value of $C_a$ is updated by adding $C_a$ to the original value of $C_a$. $\alpha$ is initially set to 1. The NFCAC controller takes the available capacity $C_a$, the congestion indicator $y$, and the system performance feedback of cell loss ratio $\ell$ as input linguistic variables to handle the CAC procedure and sends a decision signal $\hat{z}$ back to the new connection to indicate acceptance or rejection of the new call request.

In general, the NFCAC controller (shown in Fig. 2) has a net input function $f_i^k(u_{ij}^k)$ and an activation output function $\delta_i^k(f_i^k)$ for node $i$ in layer $k$, where $u_{ij}^k$ denotes a possible input to node $i$ in layer $k$ from node $j$ in layer $(k-1)$. The layers are described below.

Layer 1: In this layer, there are three input nodes with respective input linguistic variables $C_a$, $y_i$, and $p_i$. Define

$$f_i^1(u_{ij}^1) = u_{ij}^1 \quad \text{and} \quad \delta_i^1 = f_i^1 \quad (1)$$

where $u_{i1} = C_a$, $u_{i2} = y_i$, $u_{i3} = p_i$, and $1 \leq i \leq 3$.

Layer 2: The nodes in this layer are used as the fuzzifier. As in the CAC methods in the literature [1], [8], the term used to describe the remaining capacity available for a new connection is either “Enough” or “Not Enough.” Thus the term set for the available capacity is defined as $T(C_a) = \text{Not Enough (NE)}, \text{Enough (E)}$. The system is in either a congestion state (“$y$ is Negative”) or a congestion-free state (“$y$ is Positive”), so the term set for the traffic load adjustment parameter is defined as $T(y) = \text{Negative (N)}, \text{Positive (P)}$. The term used to describe the cell loss ratio, which is one of the dominant QoS requirements, is either “Satisfied” or “Not Satisfied,” and thus the term set for the cell loss ratio is defined as $T(\ell) = \text{Satisfied (S)}, \text{Not Satisfied (NS)}$. In all, we have six nodes in this layer. Each node performs a bell-shaped function, defined as

$$f_i^2(u_{ij}^2) = \left(\frac{u_{ij}^2 - m_{ij}^2}{\sigma_{ij}^2}\right)^2 \quad \delta_i^2 = e_i^2 \quad (2)$$

where $u_{ij}^2 = a_{ij}^{(2)}, 1 \leq i \leq 6, j = \left\lceil\frac{i+1}{2}\right\rceil$, and $m_{ij}^2$ and $\sigma_{ij}^2$ are the mean and the standard deviation of the $n$th term of the input linguistic variable from node $j$ in input layer, respectively. $n = 1$ if $i$ is the odd node and $n = 2$ if $i$ is the even node.

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 $[T(C_a)] \times [T(y)] \times [T(p_i)]$ (the membership function of the number of terms in $T(x)$). Consequently, there are eight rule nodes in this layer. Each rule node performs the fuzzy AND operation defined as

$$f_i^3(u_{ij}^3) = \min(u_{ij}^3; \forall j \in P_i)$$

$$\delta_i^3 = f_i^3 \quad (3)$$

where $u_{ij}^3 = a_{ij}^{(2)}$ and $P_i = \{ j \mid \text{all that are precondition nodes of the } i\text{-th rule} \}, 1 \leq i \leq 8$.

Layer 4: The 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, not only “Accept” (A) and “Reject” (R) but also “Weak Accept” (WA) and “Weak Reject” (WR) are employed to describe the accept/reject decision. Therefore, NFCAC controller may have an alternative choice for calls which fall into the area around the call acceptance/rejection decision boundary. Thus, the term set of the output linguistic variable $\hat{z}$ is defined as $T(\hat{z}) = \text{R, WR, WA, A}$. There are four nodes in this layer. Each node performs a fuzzy operation to integrate the fired strength of rules that have the same consequence. Thus, we define

$$f_i^4(u_{ij}^4) = \max(u_{ij}^4; \forall j \in C_i)$$

$$\delta_i^4 = f_i^4 \quad (4)$$

where $u_{ij}^4 = a_{ij}^{(3)}$ and $C_i = \{ j \mid \text{all that have the same consequence of the } i\text{-th term in the term set of } \hat{z}, 1 \leq i \leq 4 \}$. The up-down operating mode is used during the training period. The nodes in this layer and the links in layer five have functions similar to those in layer two. Each node performs a bell-shaped function defined as

$$f_i^5(u_{ij}^5) = \left(\frac{u_{ij}^5 - m_{ij}^5}{\sigma_{ij}^5}\right)^2 \quad \delta_i^5 = e_i^5 \quad (5)$$

where $u_{ij}^5$ is set to be $a_{ij}^{(2)}$ obtained from the up-down operating nodes in layer five, and $m_{ij}^{(O)}$ and $\sigma_{ij}^{(O)}$ are the mean and the standard deviation of the $j$-th term of $\hat{z}$, respectively. $1 \leq i \leq 4$, $j = 1$.

Layer 5: There are two nodes in this layer. One node performs the down-up operation for the actual decision signal $\hat{z}$. The node and its links act as the defuzzifier. The function used to simulate a center-of-area
defuzzification method is approximated by

\[ f_i^{(5)}(u_{ij}^{(5)}) = \sum_{j=1}^{4} m_j^{(O)} \alpha_j^{(O)} u_{ij}^{(5)} \]

and

\[ a_i^{(5)} = U \left( \frac{f_i^{(5)}(u_{ij}^{(5)})}{\sum_{j=1}^{4} \alpha_j^{(O)} u_{ij}^{(5)}} - z_0 \right) \]

where \( u_{ij}^{(5)} = \alpha_j^{(5)}, \ i = 1, \ z_0 \) is the decision threshold, and

\[ U(x) = \begin{cases} 1, & \text{if } x \geq 0, \\ 0, & \text{otherwise}. \end{cases} \]

Clearly, \( \hat{z} = \alpha_i^{(5)} \) and a new connection will be accepted only if \( \hat{z} = 1 \). The other node performs the up-down operation during the training period. It feeds the desired decision signal \( \tilde{z} \) into the controller to adjust the link weights optimally. For this kind of node

\[ f_i^{(5)}(u_{ij}^{(5)}) = u_{ij}^{(5)} \quad \text{and} \quad a_i^{(5)} = f_i^{(5)} \]

where \( i = j = 1 \) and \( u_{11}^{(5)} = z \).

C. Hybrid Learning Algorithm

A hybrid learning algorithm is applied in the design of the NFCAC controller. The algorithm is a two-phase learning method. In phase one, a self-organized learning scheme is used to construct the rules and to locate the initial membership functions. In phase two, a supervised learning scheme is adopted to optimally adjust the membership functions for desired outputs. Training data must be provided for the learning process, in addition to the size of the term set for each input/output linguistic variable and the fuzzy control rules. The procedure for constructing the set of training data is described below.

\[
\text{Construction of Training Data:} \\
\text{For a new connection request with traffic parameters of } R_m, R_p, \text{ and } P_r \\
\text{Estimate the required capacity } C_e \text{ by using fuzzy bandwidth estimator} \\
\text{Count the available capacity } C_a \text{ by using network resource estimator} \\
\text{Generate a congestion indicator } y \text{ by using fuzzy congestion controller} \\
\text{Get the cell loss ratio } P_l \text{ measured from system information statistics} \\
\text{If } P_l > \text{QoS} \ \\
\text{Then Reject the request and set the desired output } z = 0 \\
\text{Else Accept the request and set the desired output } z = 1 \\
\text{Verification of the Acceptance Decision:} \\
\text{Continue the simulation for a predefined time interval, without accepting any new connection requests} \\
\text{Obtain the statistics of cell loss ratio } P_l^t \\
\text{If } P_l^t > \text{QoS (acceptance decision is failed), then} \ \\
\text{Set } z = 0 \text{ EndIf} \ \\
\text{EndIf} \\
\text{Store training data of } C_a, y, P_l, \text{ and } z \\
\text{Using the input training data } C_a, y, P_l, \text{ the desired output } z, \text{ the fuzzy partitions } [C_a], [y], [P_l], [z], \text{ and the desired shape of the membership functions, the self-organized training would locate the membership functions and find the fuzzy control rules. If an initial knowledge base is employed to help constructing an initial structure of the fuzzy control rules, a number of possible rule structures can be formed by slight modification of rules. Among all of the possible structures, the one that yields the minimum square error } E \text{ for the training data is selected. } E \text{ is defined as} \\
\text{Obtain the statistics of cell loss ratio } P_l^t \\
\text{If } P_l^t > \text{QoS (acceptance decision is failed), then} \ \\
\text{Set } z = 0 \text{ EndIf} \ \\
\text{EndIf} \\
\text{With respect to } E \text{ defined as} \\
\text{where } N \text{ is the number of training data and } z(t_j) \text{ and } \tilde{z}(t_j) \text{ are the desired output and the actual output obtained at time } t_j, \text{ respectively.} \\
\text{If an initial knowledge base is not provided, the initial locations of membership functions are estimated by using Kohonen’s self organizing feature-maps algorithm and the N-nearest-neighbors scheme [19], and the initial rule structure is constructed via genetic algorithms (GA’s) [20].} \\
\text{The procedure to locate the means } m_i \text{ of the } i \text{th membership function for linguistic variable } x_i, 1 \leq i \leq M, \text{ given a set of training data } x_j \text{ for } x_i, 1 \leq j \leq N, \text{ is described below.} \\
\text{It employs the statistical clustering technique of Kohonen’s feature-maps algorithm [17].} \\
\text{Obtain } m_i \text{ by using Kohonen’s Feature-Maps Algorithm:} \\
\text{Set initial values of } m_i \text{ for all membership functions, } \ 1 \leq i \leq M, \text{ such that} \\
\text{Set an initial learning rate } \alpha \ (0 < \alpha < 1). \\
\text{Step 1: Set } j = 1. \\
\text{Step 2: Present training data } x_j \text{ and compute the distance} \\
\text{Step 3: Determine the } k \text{th membership function which has} \\
\text{Step 4: Update } m_k \text{ by} \\
\text{Step 5:} \\
\text{Step 6:} \\
\text{The above procedure will stop until } \alpha \leq 0. \text{ The determination of which } d_k \text{ is minimum at Step 4 can quickly be accomplished}
in constant time via a winner-take-all circuit [17]. The adaptive 
algorithm can be independently performed to obtain \( m_i \) for 
each input and output linguistic variables.

As for the corresponding standard deviation \( \sigma_i \) of the \( i \)th membership function of \( x_i \), since \( m_i \) and \( \sigma_i \) will be finely 
tuned in the supervised learning phase, we just use a first-
nearest-neighbor heuristic to estimate \( \sigma_i \), which is given by

\[
\sigma_i = \frac{|m_i - m^*|}{\gamma}
\]

where

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

and \( \gamma \) is called an overlap parameter used to describe the 
degree of overlapping for the two membership functions.

GA’s are search algorithms based on the mechanics of 
natural selection and natural genetics [21, pp. 1–22]. They 
combine the survival of the fittest and some of the innovative 
flair of human search. According to the fittest values among 
those randomly selected string structures, a structured but 
randomized information exchange is defined to form a search 
algorithm. Although the randomized generating procedure is 
used, GA’s are not simple random walks. They efficiently 
make use of the historical information to speculate on new 
search points with expected improved performance [21]. The 
input/output rule structure is encoded into a gene string \( G(t) \) 
defined as

\[
G(t) = [g_1(t), g_2(t), \ldots, g_n(t)]
\]

where \( n \) is the total number of rule nodes, and \( g_i(t) \) (\( 1 \leq i \leq n \)) denotes the \( i \)th gene in \( G(t) \). For example, if the \( i \)th rule 
node in Layer 3 is connected to the \( j \)th node (\( 1 \leq j \leq \lceil T(\hat{z}) \rceil \)) in 
Layer 4 at time \( t \), then \( g_i(t) \) is set to \( j \). Initially, the rules 
\( g_i(0) \) are integers and are randomly assigned within the 
range of \([1, \lceil T(\hat{z}) \rceil] \). \( G(t) \) is then updated by genetic operators of 
crossover and mutation according to the value of fitness 
function, which is defined as the inverse of the error \( E \) defined 
in (9). The structure that provides the minimum value of \( E \) 
will be chosen as the optimal structure.

After the self-organized training phase, the NFCAC con-
troller then enters the supervised learning phase. The aim of the 
supervised learning is to further minimize \( E \) for the training 
data using a back-propagation learning algorithm. Starting at 
the output node, a backward pass is used to compute \( \partial E / \partial w \) 
for all the hidden nodes in Layer 4 and Layer 2. Assuming that 
\( w \) is an adjustable parameter in a node (i.e., the mean or the 
standard deviation of the membership function), the general 
learning rule is

\[
w^{new} = w^{old} + \eta \frac{\partial E}{\partial w}
\]

where \( \eta \) is the learning rate and

\[
\frac{\partial E}{\partial w} = \frac{\partial E}{\partial f} \frac{\partial f}{\partial w} = \frac{\partial E}{\partial a} \frac{\partial a}{\partial f} \frac{\partial f}{\partial w},
\]

\( f \) and \( a \) were defined in the previous subsection. Here, 
different values of \( \eta \) could be used in Layer 2 and Layer 
4 to provide different learning rates for input and output 
variables. Different values of \( \eta \) represent different adoption 
rates for these variables. If the membership function of a 
specific linguistic variable is not intended to be modified, then 
\( \eta = 0 \) is used.

### III. Simulation Results

Simulations were performed to test the effectiveness of the 
proposed NFCAC scheme. Before discussing the results of the 
simulations, we will first describe the simulation environment.

#### A. Simulation Environment

Assume that an ATM network is chosen to be the high-speed 
network supporting multimedia services. The input traffic is 
categorized into two types: real-time (type-1) and nonreal-
time (type-2) traffic. Video and voice services are examples 
of type-1 traffic, while data services are examples of type-2 
traffic. The network system provides two separate finite buffers 
with size \( K_i \), in order to support different QoS requirements 
for type-\( i \) traffic, \( i = 1 \) and 2. When the buffer is full, 
incoming cells are blocked and lost. The system reserves \( C_t \) 
portion of its capacity for type-1 traffic and the remaining 
\((1 - C_t)\) portion for type-2 traffic. When there is unused type-
1 or type-2 capacity, it is used for the other type of traffic. 
In the simulations described here, \( K_1 = K_2 = 100 \) cells 
and \( C_t = 0.8 \). Also, the QoS requirement for type-1 traffic 
\( QoS_1 \) is \( 10^{-5} \) and that for type-2 traffic \( QoS_2 = 10^{-6} \).

The cell-generation process for a video coder is assumed to 
have two motion states: one is the low motion state for the 
rate of interframe coding and the other is the high motion state 
for the rate of intraframe coding [22]. The rate of intraframe 
coding is further divided into two parts: the first part has 
the same rate as the intraframe coding and the second part, 
called difference coding, is the difference between the rates 
of intraframe coding and interframe coding. The interframe 
coding and the difference coding are all modeled as discrete-
state Markov-modulated Bernoulli processes (MMBP) with 
basic rates \( A_p \) and \( A_q \). The state-transition diagram is shown 
in Fig. 4(a) and (b). Let \( \lambda_p(t), \lambda_q(t), \) and \( X_p(t) \) denote 
the cell generation rates for intraframe coding, interframe coding, 
and difference coding at time \( t \), respectively, from the video 
coder. Clearly, \( \lambda_0(t) = \lambda_p(t) + \lambda_q(t) \). The process of \( \lambda_q(t) \) 
is an \( (M_p + 1) \)-state birth–death Markov process. The 
state-transition diagram for \( \lambda_q(t) \) uses the label \( m_{p+1} \), to 
indicate the cell generation rate of interframe coding of a state 
and uses the labels \( (M_p - m_p) \gamma \) and \( m_p \psi \) to denote the transition 
probabilities from state \( m_p.A_p \) to state \( (m_p + 1).A_p \) and from 
state \( m_a.A_p \) to state \( (m_a - 1).A_p \). Similarly, 
the process for \( X_p(t) \) is an \( (M_a + 1) \)-state birth–death Markov 
process. The state-transition diagram for \( X_p(t) \) uses the label 
\( m_a.A_a \) to indicate the additional cell generation rate of a state 
due to intraframe coding and uses the labels \( (M_a - m_a) \psi \) 
and \( m_a \psi \) to denote the transition probability from state 
\( m_a.A_a \) to state \( (m_a + 1).A_a \) and from state \( m_a.A_a \) to state \( (m_a - 1).A_a \), 
respectively. One should note that the long-term correlation 
behavior of a video source is resulted from the process 
\( \lambda_0(t) \). The video source will alternate between interframe and
Fig. 4. Level transition diagram for (a) interframe coding \( \lambda_i(t) \), (b) difference state \( \lambda'_i(t) \), (c) interframe and intraframe alternate model, and (d) voice source.

TABLE I

<table>
<thead>
<tr>
<th>Rule</th>
<th>( C_a )</th>
<th>( y )</th>
<th>( \psi )</th>
<th>( \phi )</th>
<th>( R_p )</th>
<th>( R_m )</th>
<th>( T_p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NE</td>
<td>N</td>
<td>NS</td>
<td>R</td>
<td>5</td>
<td>E</td>
<td>N</td>
</tr>
<tr>
<td>2</td>
<td>NE</td>
<td>N</td>
<td>S</td>
<td>WR</td>
<td>6</td>
<td>E</td>
<td>N</td>
</tr>
<tr>
<td>3</td>
<td>NE</td>
<td>P</td>
<td>NS</td>
<td>R</td>
<td>7</td>
<td>E</td>
<td>P</td>
</tr>
<tr>
<td>4</td>
<td>NE</td>
<td>P</td>
<td>S</td>
<td>WR</td>
<td>8</td>
<td>E</td>
<td>P</td>
</tr>
</tbody>
</table>

The behavior of the congestion indicator \( y \) could be monitored from the congestion and congestion-free states during a long-term simulation of the network operation. Thus, the membership functions of \( y \) could be initially optimized based on the obtained information. The mean value \( m_{11}^{(L)} \) of the membership function of \( N \) would be set to be the mean value of the queue-length change rate during congestion-free periods, and let \( \alpha = 1.2 \). These parameters could be further off-line optimized via GA by simulation.

The initial membership functions of the cell loss ratio \( p_L \) were set according to the QoS requirement. The mean value \( m_{32}^{(L)} \) of the membership function of \( S \) would be set to be the QoS requirement, and the standard deviations would be set to be \( \sigma_{31}^{(L)} = \sigma_{32}^{(L)} = m_{32}^{(L)} - m_{31}^{(L)} \). As a result, there exists a safety margin between the membership functions of terms \( S \) and \( R_m = 7.36 \times 10^{-3} \), and \( T_p = 3.14 \times 10^{-2} \) s, which would give \( \theta_1 = 0.1 \), and for low-bit-rate data sources, it is assumed that \( R_p = 3.68 \times 10^{-2} \), \( R_m = 7.36 \times 10^{-4} \), and \( T_p = 2.88 \times 10^{-2} \) s, which give \( \theta_2 = 0.02 \). The mean holding time is 60 min for a video service, 3 min for a voice service, and 18 s for both high- and low-bit-rate data services. Notice that the values of \( R_p \) and \( R_m \) have been normalized by the network capacity.

Two kinds of cell loss ratios for type-\( i \) traffic are considered: the source loss ratio due to selective discarding at the customer side \( p_{bs} \) and the node loss ratio due to blocking at the network side \( p_{bs} \). The overall cell loss ratio for type-\( i \) traffic \( p_{ki} \) is defined as

\[
p_{ki} = \kappa p_{bs} + p_{bs}, \quad i = 1, 2
\]

where \( \kappa \) is used to indicate the significance of the node loss ratio over the source loss ratio. \( \kappa = 0.8 \) is assumed here because selectively discarding cells at the source should have less effect on information retrieval than blocking cells at the node. In the simulations, the cell loss ratio is estimated as the total loss cells divided by the arriving cells during the whole simulation interval.

B. Simulation Results and Discussion

On the basis of prior knowledge concerning CAC, the rule structure and parameters of the NFCAC controller can be initially set and then properly adjusted via the learning algorithm. The membership functions of the linguistic variables for type-1 and type-2 traffic were initially specified in the left-hand side of Fig. 5(a) and 5(b), respectively. As we know, the available capacity \( C_0 \) deduced from the equivalent capacity \( C_e \) of the existing calls, may possess estimation errors. In order to utilize the network as much as possible, we may employ an idea of “budget deficit” to over-assign the capacity. Thus, the mean value \( m_{11}^{(L)} \) of the membership function of \( NE \) was set to be a negative value and the mean value \( m_{32}^{(L)} \) of the membership function of \( E \) was set to be a value close to zero.

As for the data source, there are high-bit-rate and low-bit-rate data services. The generation of high-bit-rate and low-bit-rate data cells is characterized by Bernoulli processes with rates \( \theta_1 \) and \( \theta_2 \), respectively. Also, the distributions of the holding times for video, voice, high-bit-rate data, and low-bit-rate data are assumed to be exponentially distributed.

In the simulations, for the arrival process of a video source, it is assumed that \( R_p = 3.31 \times 10^{-2} \), \( R_m = 1.10 \times 10^{-2} \), and \( T_p = 0.5 \) s, which would give \( M_p = M_a = 20 \), \( A_p = 1.34 \times 10^{-3} \), \( A_a = 3.15 \times 10^{-1} \), \( \gamma = 3.77 \times 10^{-6} \), \( \omega = 5.65 \times 10^{-6} \), \( \phi = \psi = 2.83 \times 10^{-7} \), \( c = 5.65 \times 10^{-6} \), and \( d = 5.09 \times 10^{-5} \); for the arrival process of a voice source, it is assumed that \( R_p = 4.71 \times 10^{-2} \), \( R_m = 2.12 \times 10^{-4} \), and \( T_p = 1.35 \) s, which would give \( A_p = 4.71 \times 10^{-1} \), \( \alpha = 1.71 \times 10^{-6} \), and \( \beta = 2.00 \times 10^{-6} \); for high-bit-rate data sources, it is assumed that \( R_p = 7.36 \times 10^{-2} \),
NS provided to tolerate the dynamic behavior of the network operation and insure the QoS requirement.

Here, little information about the setting of initial values for the mean $m_{ij}^{(O)}$ of the term set $T(\hat{z})$ could be employed; therefore, the values of $m_{ij}^{(O)}$ are set to be equally spaced in the range of $[0, 1]$. Based on the initial membership functions, an optimal rule structure shown in Table I was obtained by using GA in the self-organized learning phase. When the fuzzy logic rules were found, the NFCAC controller entered the supervised learning phase, in which the membership functions were adjusted optimally.

Three different values of $\eta$ were used for the variables $C_a$, $y$, $p$, and $\hat{z}$. $\eta$ was set to zero for $p$ because the membership functions were specified by the QoS constraint and should not be modified. $\eta = 0.001$ was used for $y$ because the membership functions of $y$ were initially optimized. As for $C_a$ and $\hat{z}$, their initial membership functions were heuristically set and required further optimization in the supervised learning phase. Thus, $\eta = 0.01$ was used. The use of different $\eta$ may drastically reduce the training time required in the supervised learning phase. The learned membership functions of the linguistic variables for type-1 and type-2 traffic were shown in the right-hand side of Fig. 5(a) and Fig. 5(b), respectively.

For type-1 traffic in Fig. 5(a), it can be found that the differences of the membership functions before and after learning are as follows. For the membership functions of $C_a$, the mean value $m_{11}^{(I)}$ of the membership function of $NE$ was properly modified from $0.4$ to $0.27$. Similarly, the mean value $m_{12}^{(I)}$ of the membership function of $E$ was properly modified from $0.16$ to $0.02$. There is a drastically change for membership functions of $C_a$, and the phenomenon can also be found in the membership functions of $y$. It is because we heuristically set their initial values and we used only two terms to describe $C_a$ or $y$. The change of the position of one term of $C_a$ and $y$ will squeeze the other term but receive less counteraction from the other one term (compared to $\hat{z}$ described later). Membership functions of $p$ are not changed since $\eta$ for $p$ was chosen to be zero. For the membership function of $\hat{z}$, however, the mean $m_{11}^{(O)}$ of the membership function of $R$ is slightly increased from $0$ to $0.05$, representing that the effect of “Reject” is decreased. Also, the mean $m_{3}^{(O)}$ of the membership function of $WA$ is slightly increased from $0.67$ to $0.72$, representing that the effect of “Weak Accept” is increased. The small change is because we used four terms to describe $\hat{z}$. The change of the position of one term of $\hat{z}$ will squeeze the other three terms but receive more counteraction from the three terms. Therefore, the change of position would be confined in a smaller range. The changes of membership functions of $\hat{z}$ imply that the NFCAC controller prefers to accept new calls. This phenomenon demonstrates that the NFCAC controller intends to recover some system bandwidth which the equivalent capacity method wastes due to over-estimation, while keeping the QoS contract. It may be the reason for the utilization improvement of the proposed NFCAC controller, which will be shown below. Similar results could be found for type-2 traffic in Fig. 5(b).

We compare the NFCAC scheme with the effective-bandwidth-based CAC (EBCAC) scheme proposed in [3], the fuzzy-logic-based CAC (FLCAC) scheme proposed in [8], the neural-net-based CAC (NNCAC) scheme proposed in [14], and the radial-basis-function-based CAC (RBFCAC) scheme from the aspects of the cell loss ratio (CLR), the system utilization, and/or the training time under the constraint of QoS guarantee. The EBCAC scheme is a hybrid technique combining the conventional techniques of the Gaussian approximation and the bufferless analysis; it is an improved version of the...
equivalent capacity method [1]. Simulation of the EBCAC scheme is simply to calculate the required bandwidth of a new connection. The new connection request is accepted if the total bandwidth required by the new connection and the existing connection is less than the system capacity. Otherwise, it is rejected. The FLCAC scheme is a fuzzy implementation of the equivalent capacity admission control method; details for the FLCAC scheme can be referred to [8]. The NNCAC and RBFCAC schemes are neural-net implementation of the equivalent capacity admission control method, where the NNCAC adopts the multilayer perceptron (MLP) structure with 30 hidden nodes, while the RBFCAC uses radial basis function network (RBFN) with 30 hidden nodes. Details for the NNCAC scheme can be referred to [14]. In the simulations, the NNCAC, NNCAC, or RBFCAC controller is equipped with the same three peripheral processors as those used in the NFCAC controller shown in Fig. 3. The sizes of training set and test set are all equal to 200, the number of repeated experiments is 20, and the standard deviation is less than 5%.

Fig. 6 shows the CLR’s of an ATM traffic controller employing the NFCAC scheme, and the EBCAC, FLCAC, NNCAC, RBFCAC schemes. It is found that the QoSs for both types of traffic are indeed guaranteed for all of these control schemes. Fig. 7 shows that the system utilization of the NFCAC scheme and the four schemes. We can find that the utilization of the NFCAC scheme is slightly greater than that of the NNCAC and the RBFCAC schemes; the system utilizations of NFCAC, NNCAC, and RBFCAC are 91%, 90.5%, and 89%, respectively; and the NFCAC scheme offers about 32% and 11% greater system utilization than the EBCAC scheme and the FLCAC scheme. It is because NFCAC can incorporate the domain knowledge obtained from both the analytical-based method (the equivalent capacity scheme [1] is employed in the bandwidth estimator) and the measurement-based method (the system statistics of the queue length, the change rate of the queue length, and the CLR are considered in the congestion controller). Also, the reason for the performance improvement is that NFCAC possesses the learning capability of the neural network.

Fig. 8 shows the training time required for the NFCAC scheme and the NNCAC, RBFCAC schemes. Here, a widely used back-propagation learning algorithm was employed to adjust the membership functions (i.e. represented in terms of weights) of the multilayer neural fuzzy network and neural network for the NFCAC and NNCAC schemes, while the RBFCAC scheme is basically trained by the hybrid learning rule: unsupervised learning in the input layer and supervised learning in the output layer. It is found that NFCAC has training time of 7 (4) epochs, while RBFCAC and NNCAC have training time of $10^3$ (40) and $5 \times 10^3$ ($6 \times 10^2$), respectively, for type-1 (type-2) traffic. The NFCAC has higher learning speed than the RBFCAC and NNCAC. One reason is that the neural fuzzy network is a structured network, thus the NFCAC controller can easily adopt the domain knowledge of conventional control methods to construct the initial rule structure and the parameters of the membership functions, providing an excellent initial guess in adjusting its weights; on the contrast, the neural network is a nonstructured network, which cannot incorporate domain knowledge about system. The other reason is that the neural fuzzy network has simpler structure than the neural network; the number of tunning parameters used in the neural fuzzy network is quite small, as compared to the neural network such as MLP and RBFN considered here. In this paper, there are only 16 weighting parameters used in NFCAC, while there are 150 and 480 weighting parameters required for the RBFCAC and NNCAC, respectively. It is also noted that the RBFCAC scheme has less learning time than the NNCAC scheme. This is because the RBFCAC scheme can have the proper initial setting of means
order and the variances for the Gaussian activation functions during unsupervised learning according to the prior knowledge, and it has only one layer of connection needed to be trained by supervised learning.

As usually noted, RBFCAC can have faster training speed than NNCAC but cannot achieve the same accuracy as the back-propagation NNCA. In the simulations, we first adopted the same set of data used to train NFCAC and NNCAC for RBFCAC. However, it was found that RBFCAC finally violated the QoS contracts due to its error decision of accepting more users than it should be. In order to provide QoS guarantee for RBFCAC, we have to prepare much more training data, especially those around the acceptance/rejection boundary. This will increase the training time of RBFCAC in each epoch than those required by NFCAC and NNCAC. Moreover, the overall processing time of RBFCAC is greater than that needed by either NFCAC or NNCAC because RBFCAC uses more nodes (compared with NFCAC) and a more complicated activation function (compared with NNCAC). All of these would degrade the performance of RBFCAC in real application.

IV. CONCLUSION

This paper proposes a neural fuzzy approach for connection admission control in high-speed multimedia networks. The NFCAC scheme combines the linguistic control capability of a fuzzy logic controller and the learning ability of a neural network. This type of integrated neural fuzzy system can automatically construct a rule structure by learning from training examples and can self-calibrate parameters of membership functions. It not only provides a robust framework to mimic experts’ knowledge embodied in existing traffic control techniques but also constructs intelligent computational algorithms for traffic control. It can be easily trained and enhances system utilization. Simulation results show that the proposed NFCAC scheme provides system utilization about 32% and 11% higher than the EBCAC and FLCAC schemes proposed in [3] and [8], respectively, and the NFCAC scheme requires only a fraction of the $10^3$ order and the $10^4$ order of training cycles, consumed by the NNCAC scheme proposed in [14] and RBFCAC scheme, respectively. An NFCAC scheme such as the one introduced here may be the answer to the problem of designing a coherent call admission controller for ATM systems.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments, which led to a substantial improvement in the paper.

REFERENCES


Ray-Guang Cheng received B.E., M.E., and Ph.D. degrees in communication engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1991, 1993, and 1996, respectively. Since 1997, he has been a Researcher in the Advanced Technology Center, Computer and Communications Research Laboratories, Industrial Technology Research Institute, Taiwan, where he was involved in the designing of the medium access control protocol for the third-generation wireless networks. His current research interests include performance analysis, broadband wireless communication networks, and asynchronous transfer mode networks.

Dr. Cheng is a member of Phi Tau Phi.

Chung-Ju Chang (S’81-M’85-SM’94) received the B.E. and the M.E. degrees in electronics engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1972 and 1976, respectively, and the Ph.D. degree in electrical engineering from National Taiwan University in 1985.

From 1976 to 1988, he was with Telecommunication Laboratories, Directorate General of Telecommunications, Ministry of Communications, Republic of China, as a Design Engineer, Supervisor, Project Manager, and then Division Director. There, he was involved in designing digital switching system, ISDN user–network interface, and ISDN service and technology trials. He has also been a Science and Technical Advisor for Minister of the Ministry of Communications from 1987 to 1989. In August 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. He served as 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 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).

Li-Fong Lin (S’98) received the B.E. degree in communication engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1996, where he is currently working toward the Ph.D. degree in communication engineering.

His current research interests include performance analysis, connection admission control over asynchronous transfer mode networks, fuzzy systems, and neural networks.