Optimal Priority-based Call Admission Control Scheme For QoS Provisioning in Heterogeneous Wireless Networks

Wen Chen, Jinming Yu, Feng Pan
College of Information Science and Technology
Engineering Research Center of Digitized Textile & Fashion Technology
Ministry of Education, Donghua University
Shanghai 201620, China
chenwen@dhu.edu.cn

Abstract—Next generation wireless networks will support a diverse set of access technologies. And the demanding breed of multimedia applications will even more considerably require Quality of Service (QoS) support. This paper proposes an optimal priority-based Call Admission Control (CAC) scheme in heterogeneous wireless networks, i.e., the generalized semi-Markov Decision Process based CAC (GCAC). Considering the limitations of the traditional Markov decision process for optimal CAC policy, we characterize the CAC problem through a GSMDP formulation that characterizes the network behaviors in the heterogeneous system more exactly. And the proposed scheme can guarantee the QoS constrains both in the physical and network layer. Numerical examples illustrate that the performance in the GCAC scheme is significantly better than that in the conventional GC scheme and an optimal CAC scheme based on MDP.

Index Terms—CAC; GSMDP; heterogenous; wireless

I. INTRODUCTION

Next generation wireless networks are promising to provide not only conventional voice services but also the efficiency and flexibility of multiplexing a wide variety of traffic from data to multimedia applications. And there are currently a large variety of wireless access networks, including wireless area networks (WLANs), 3G/beyond 3G (B3G) cellular networks, etc. Supporting multimedia services with different quality-of-service (QoS) requirements over such highly heterogeneous environments is a very challenging task, and efficient utilization of these available wireless resources by using efficient call admission control (CAC) policies has been one of the main concerns in mobile system design.

CAC for wired and wireless networks has been extensively studied in the past and many CAC schemes have been proposed [1] - [3]. The guard channel (GC) based scheme proposed in the literature (e.g., [1]) is one of the most popular priority-based (threshold-based) CAC schemes in providing preferential treatment. A set of guard channels are reserved for prioritized calls in GC schemes such as cutoff priority [1] [4] [5], fractional guard channel [6], and new call bounding [7] schemes. However, conventional GC-based CAC schemes do not suggest any optimal threshold to be used in a given traffic condition. To achieve better QoS, many schemes [8] - [12] apply the Markov decision process (MDP) and Semi-Markov Decision Process (SMDP) to derive an optimal CAC policy, for their classical mathematical formulation and solvability.

However, these methods are limited by some rather restrictive assumption that the corresponding discrete-event systems (DES) poss memoryless (i.e. Markov) property. Thus the optimal CAC schemes [8] - [12] use the following general assumptions: 1) Call arrivals for all the service classes are assumed to follow the Poisson process; 2) the cell dwelling time for the call is assumed to follow an exponential distribution. Only by the assumptions their CAC problems were characterized as MDP Models. Recent field studies [13] is shown that, except for the case of exponentially distributed cell residence times, the call holding time is not exponentially distributed, and the arrival of new/handoff traffic to a cell is not Poisson distributed yet. That is to say, Markov property is not maintained anymore. Using general traffic arriving pattern and general service time distributions may enable us to better model the real world of system behaviors [13].

In this paper, we propose an optimal priority-based CAC scheme for heterogeneous networks called the GSMDP-based CAC (GCAC), for Markov models do not adapt to characterize the current network behaviors [13]. The primary contribution is that we relax the Markov assumptions, and characterize the CAC problems through a generalized semi-Markov Decision Process (GSMDP, a model based on Generalized semi-Markov Process (GSMP)) formulation. GCAC allows the CAC operations to be performed in a general framework, even with various traffic and different access network. Besides, the GCAC system model takes into account several cross-layer QoS constrains in all the different access network, including the physical layer signal-to-interference (SIR) and the network layer blocking probabilities of calls. Thereby, the GCAC scheme maximums overall network revenue while satisfying the QoS constrains in the
heterogeneous network. And the performance is evaluated via simulation, and the results are compared with conventional GC scheme and an optimal CAC scheme based on SMDP (Semi-Markov decision process).

The rest of the paper is organized as follows. Section II describes the system model, the call admission procedures, and the important QoS metrics in the heterogeneous system. Then, GSMDP model for stochastic DDES is explained in Section III. Section IV formulates the optimal CAC problem in the next generation heterogeneous network as a GSMDP while satisfying QoS requirements. Subsequently, the numerical results are described in Section V. Finally, some concluding remarks are presented in Section VI.

II. SYSTEM MODEL

A. Description of Heterogeneous Network Model

In this paper, we consider a heterogeneous network architecture, which is comprising a core network and multiple heterogeneous access networks. Without loss of generality, we show two different access networks — 3G/beyond 3G (B3G) cellular networks and WLANs. Recent studies show that WLANs and 3G wide area networks, such as code division multiple access (CDMA) cellular networks, will coexist to offer Internet services to users [14] [15]. WLANs offer relatively high data rates. However, WLANs can only cover smaller areas (hotspots), such as hotels and airports. On the contrary, CDMA cellular networks support low data rates, but offer a much wider area of coverage that enables ubiquitous connectivity. The complementary characteristics of WLANs and CDMA cellular networks make it attractive to integrate these two wireless access technologies.

For WLAN/CDMA heterogeneous access network, the paper consider three types of calls for a certain class (e.g. voice or data or video) — horizontal handoffs, vertical handoffs, and new calls. When a Mobile System (MS) are moving between homogeneous systems (e.g., from a 3G cellular cell to another 3G cellular cell), horizontal handoff occurs. On the other hand, if the MSs are moving across different systems (e.g., from a 3G cellular cell to a WLAN system), vertical handoff calls occur. Handoff calls are commonly given a higher priority so as to provide a seamless connection. In general, for the same class call, we assign the priority order as horizontal handoff call > vertical handoff call > new call.

B. The CAC Problem in Heterogeneous Wireless Network

Here, the CAC policy resides in each BS of the cells or the access router in the radio access networks. The CAC process determines whether an incoming call can be admitted or not, and how to allocate the limited bandwidth resources.

To better model the real world of system behaviors, we use general traffic arriving pattern and general service time distributions [13]. We denote $G_k^i$ corresponding to the general interarrival time distribution of call arrival. So new mobile users arrive at a cell according to the general interarrival time distribution $G_k^i$, and handoff calls arrive according to a different one. When a mobile user (new call or handoff call) arrives at a cell, a CAC procedure is initiated; once the mobile user is accepted, a resource allocation procedure is further involved. This paper only focuses on call admission control on a call level. The probability that new calls are blocked is called new call blocking probability, which is denoted by $P_{nh}$. Similarly, the probability that handoff calls are blocked is called handoff dropping probability, denoted by $P_{bh}$ or $P_{hb}$, according as it belongs to horizontal handoff or vertical handoff.

Then, to simplify the system complexity, there are only three different types of incoming traffic of a class applied to the system. Let $P_{vh}, P_{vh}$, $P_{ss}$ denote the maximum allowed vertical call, horizontal call and handoff blocking probability. According to their priorities, we have the following relationships: $P_{hh} < P_{vh} < P_{nh}$. And $\omega_{vh}, \omega_{hh}, \omega_{nh}$ respectively represent the relative blocking cost rate of vertical call, horizontal call and new call with $\omega_{hh} > \omega_{vh} > \omega_{nh}$, expressing the fact that rejecting a handoff request is more undesirable than blocking a new call attempt. The proposed service model can be easily extended to $K$ priority classes with the homologous priority orders in the maximum allowed blocking probability and in the blocking cost rate.

Similarly, due to users mobility and the irregular geographical cell shapes, the cell residence times will also typically have a general distribution $G_k^i$ and therefore do not hold Markov property at all. So, Markov and semi-Markov models do not have sufficient modeling power to capture many of the complex discrete-event stochastic systems, especially the next generation heterogeneous wireless network. And we drop the classical assumption made in the past, so the GCAC problem can be formulated as a generalized semi-Markov process (GSMP) [16].

Besides, an optimal CAC should maximize the average network revenue and guarantee QoS constrains in both WLANs and CDMA networks. Since the charging method in WLANs is usually different from that in CDMA networks, we should consider the different revenue rates in the design of a CAC. For the QoS constraints in WLANs, throughput and packet delay are important metrics [11], [23]. Throughput should be kept above a target value and packet delay should be kept below a target value for a certain class of traffic. In CDMA networks, QoS requirements are usually characterized by SIR. However, guaranteeing SIRs of all classes of traffic at all time instants may result in low network utilization for bursty traffic. Therefore, besides the SIR, we use another QoS metric, namely, the SIR outage probability [11]. Instead of guaranteeing the SIR at all the time, we can guarantee the SIR outage probability below some target value. In addition, at the network layer of the integrated WLAN/CDMA systems, QoS metrics are blocking probabilities of new and
handoff calls in both networks, which should also be guaranteed.

C. QoS Metrics in Heterogeneous Network

The QoS metrics in WLANs and CDMA cellular networks will be used in the design of the GCAC scheme. So in the subsection we introduce the QoS metrics separately.

1) QoS Metrics in WLANs

Firstly, we describe important QoS metrics, throughput and packet delay, for multimedia traffic in WLANs. To support QoS in WLAN, a new medium access control (MAC) protocol, IEEE 802.11e, is proposed [22]. The IEEE 802.11e MAC employs a channel access function called the hybrid coordination function (HCF), which includes a contention-based HCF part and a contention-free part. The contention-based HCF part is also called the enhanced distributed coordination function (EDCF). EDCF provides differentiated access to the wireless medium for up to eight priorities. An access category (AC) mechanism is defined to support the priorities. There are four access categories (ACs) in the IEEE 802.11a specification [27].

For access category $k$, $k = 1, \ldots, 4$, there is a set of parameters, including transmission opportunity (TXOP) $\kappa_{TXOP}$, arbitration interframe space number $\kappa_{AIFSN}$, minimum contention window (CW) $\kappa_{CW,\text{min}}$, maximum contention window (CW) $\kappa_{CW,\text{max}}$, and maximum backoff stage $\kappa_{M}$, which are announced by the access point (AP).

The paper adopt the derivation of throughput and packet delay for IEEE 802.11e in [23], which are the main QoS Constrains of WLANs in the design of GCAC scheme. Let $L$ denote the length of a time slot, $M$ denote the average bit rate of WLAN, $T_{\text{SIFS}}$ denote the duration of a short interframe space (SIFS), $T_{\text{RTS}}, T_{\text{CTS}}, T_{\text{ACK}}, T_{\text{PHY}}, T_{\text{MAC}}$ denote the time required to transmit a request-to-send (RTS), a clear-to-send (CTS), an ACK, a physical layer head, and an MAC header. Assume that all packets have a constant propagation delay $\zeta$. There are $K$ classes of traffic with distinct QoS requirements in the network. $S_{w,k}$ denotes the number of calls for call classes $k$ ($k = 1, \ldots, K$). Assume that a class $k$, $k = 1, \ldots, K$ packet has a constant probability of collision, $pc_{k}$, all class $k$ packets have the same length, $L_k$. Then the average backoff counter of the class $k$ station $E[BO_k]$ [23]

$$E[BO_k] = (1 - pc_k)(CW_{k,\text{min}} - 1) + \frac{pc_k CW_{k,\text{min}} (1 - (2 pc_k)^M)}{2 (1 - 2 pc_k)}$$

The bandwidth for class $k$ traffic is [23]

$$B_k = \frac{k_k U_k L_k}{T_f + T_c + T_s}$$

where $k_k$ is the probability that a class $k$ packet is successfully transmitted during a transmission cycle, $U_k$ is the number of class $k$ packets that can be transmitted within $TXOP_k$, and $L_k$ is the packet length. So $k_k U_k L_k$ denote the number of bits successfully transmitted for class $k$ during a transmission cycle. And during a transmission cycle, $T_f$ is the average time of all idle periods, $T_c$ is the average time of all collision periods, $T_s$ is the average time of the successful transmission period. Their detailed calculation can be referred to [23].

The average packet delay of class $k$ traffic is [23]

$$D_k = b_k + \frac{a_k pc_k}{(1 - pc_k)^2}$$

where

$$b_k = \frac{1}{U_k}(4 T_{\text{SIFS}} + \epsilon AIFSN_k + T_{\text{RTS}} + 2 \zeta + T_{\text{CTS}} + T_{\text{PHY}} + T_{\text{MAC}} + L_k / M + T_{\text{ACK}} + \left( N_k - 1 \right)(2 T_{\text{SIFS}} + T_{\text{PHY}} + T_{\text{MAC}} + L_k / M + 2 \zeta + T_{\text{ACK}}))$$

and

$$a_k = \frac{1}{U_k}(2 T_{\text{SIFS}} + \epsilon AIFSN_k + T_{\text{RTS}} + \zeta)$$

And $N_k$ is the random variable representing the number of colliding periods needed in a transmission cycle.

Thus, if the vector $s_w = [s_{w,1}, s_{w,2}, \ldots, s_{w,K}]$ represents the number of calls for call classes $k$ ($k = 1, \ldots, K$), and the WLAN admissible set can be obtained by satisfying the throughput constrains (1) and average packet delay (2), i.e.

$$S_w = \{s_w \in \mathbb{Z}^+ : B_k \geq TB_k, D_k \leq TD_k, k = 1, 2, \ldots, K\}$$

where $B_k$ is defined in (1), $D_k$ is defined in (2), and $TB_k$ and $TD_k$ are the target throughput and average packet delay respectively, for class $k$ traffic.

2) QoS Metrics in CDMA Cellular Network

In CDMA cellular network, we should take into account both linear minimum-mean square error (LMMSE) receivers and varying the statistical characteristics of the packet traffic. Then, physical layer signal-to-interference (SIR) and SIR outage are the QoS constrains concerned in the paper, that are widely used in the CDMA CAC schemes [18] [24] [25].

Although SIR is an important performance measure in CDMA network, guaranteeing the SIR of all sessions at all time instants will result in low network utilization, especially for the bursty traffic. This is similar to the philosophy in wireline networks that allocating all sessions with their peak bandwidths guarantees no packet loss, but results in the lowest network utilization and no multiplexing gain. Most admission control schemes in wireline networks allow a small packet loss probability to
increase the network utilization. Similarly, since most applications in wireless networks can tolerate a small probability of SIR outage, we introduce the SIR outage probability in wireless packet CDMA networks similar to [26].

Consider a CDMA system with spreading gain $N$ and $Y$ users. Assumed that there are $K$ classes of traffic with distinct QoS requirements in the network. An important physical layer QoS metrics for class $k$ is the distribution function of clock SIRs, which should be kept above the target value $\alpha_k$. The signature sequences of all users are independent and randomly chosen. Due to multipath fading, each user appears as $L$ resolvable paths or components at the receiver. The path $l$ of user $y^j$ is characterized by its estimated average channel gain $\bar{h}_{y^j}$ and its estimation error variance $\varsigma_y^2$. Linear minimum mean square error (LMMSE) detectors are used at the receiver to recover the transmitted information. In a large system (both $N$ and $Y$ are large) with background noise $\sigma^2$, the SIR for the LMMSE receiver of some user $y^1$, $y^1 \in \{1,2,\cdots,Y\}$, can be expressed approximately as [25]

$$SIR_{y^1} = \frac{P_{y^1} \sum_{l=1}^L |\bar{h}_{y^1}|^2 \eta}{1 + P_{y^1} \varsigma_y^2 \eta} \quad (4)$$

Where $P_{y^1}$ is the attenuated transmitted power from user $y^1$, $\eta$ is the unique fixed point in $(0, \infty)$ that satisfies

$$\eta = \left[\sigma^2 + \frac{1}{N} \sum_{y,j} \left( (L-1)I(\varsigma_y^2, \eta) + \left( \sum_{l=1}^L |\bar{h}_{y,l}|^2 + \varsigma_y^2 \right) \right) \right]^{-1}$$

and

$$I(v, \eta) = \frac{v}{1 + v \eta}$$

Assume that all users in the same class have the same average channel gain $|\bar{h}_{y,k}|^2 = \sum_{l=1}^L |\bar{h}_{y,l}|^2$, $k = 1, \cdots, K$. It is proved in [24] that a minimum received power solution exists such that all users in the system meet their target SIRs if and only if

$$\alpha_k < \frac{|\bar{h}_{y,k}|^2}{\varsigma_y^2} \quad (5)$$

and

$$\frac{1}{N} \sum_{k=1}^K \sum_{l=1}^{n_{o,k}} R_{1,k} \Gamma_{y,k} < 1 \quad (6)$$

where $n_{o,k}$ is the number of class $k$ calls in the CDMA cell, $R_{1,k}$ is the number of signature sequences assigned to the $i$ th calls of class $k$, and the basic rate (obtained using the lowest spreading gain $N$) is

$$\Gamma_{y,k} = (L-1)\alpha_k \frac{\varsigma_y^2}{|\bar{h}_{y,k}|^2} + \frac{1 + \frac{\alpha_k^2}{\varsigma_y^2}}{1 + \alpha_k} \quad (7)$$

From (6), SIR outage probability can be expressed as

$$P_{out} = P \left\{ \frac{1}{N} \sum_{k=1}^K \sum_{l=1}^{n_{o,k}} R_{1,k} \Gamma_{y,k} \geq 1 \right\} \quad (7)$$

Therefore, the SIR outage probability constraints $P_{out} \leq TP_{out}$ ($TP_{out}$ is the target SIR outage probability) will be satisfied, if the vector $s = [s_{c,1}, s_{c,2}, \cdots, s_{c,K}]$ that represents the number of calls for call classes $k$ ($k = 1, \cdots, K$) in CDMA network lies within the admissible set

$$S = \left\{ s \in Z_+^K : P_{out} \leq TP_{out}, k = 1, \cdots, K \right\} \quad (8)$$

The several vital QoS constrains in both the WLAN and the CDMA network discussed above can be met in the design of the GCAC scheme. In the following, we introduce the GSDMP model, as which the admission control problem is formulated.

III. GENERALIZED SEMI-MARKOV DECISION PROCESS MODEL

Based on a GSMP model of discrete event systems, we add a decision dimension to the formalism by distinguishing a subset of the events as controllable and adding rewards, thereby obtaining a GSDMP [16]. Unlike an MDP/SMDP, a GSDMP remembers if an event enabled in previous states without triggering. It holds history dependence and therefore breaks the Markov property.

A. Description of a GSMP

A GSMP can be thought of as a normal mathematical model of a DES simulation. An excellent and more detailed description of this framework can be found in [16]. To specify a GSMP, we define the following elements:

- $S$ : a countable set of states;
- $E$ : a finite set of events;
- $E(s)$ : the set of events scheduled to occur in state $s \in S$. Of course, $E(s) \subseteq E$;
- $p(s', s, e^*)$ : the event $e^*$ with the smallest clock value triggers a transition to a state $s' \in S$ according to the probability;
- $G_{e^*}, e \in E$ : the distribution function of clock samples for event $e$.

At any point in time, each event $e \in E(s)$ has associated with a clock, governing the time until $e$ triggers if it remains enabled, namely sampling a residual
lifetime (or clock) $l_e$, and a next-state probability distribution $p(s', s, e)$. For $s \in S$, the set $C(s)$ of possible clock-reading vectors in state $S$ is defined as 

$$C(s) = \{t_{ee}, \cdots, t_{ee}, \cdots, t_{ee}\} | t_{ee} \geq 0 \text{ and } t_{ee} > 0 \iff e \in E(s)\}$$

where there are $k$ events being enabled in state $S$. For each new event $e$, a clock is initialized with the distribution $G_e$. And clocks for any old events, which remain enabled, continue to run in the new state. At each event epoch, in state $S$, the ensuing interevent time $\gamma(s)$ is given as

$$\gamma(s) = \min_{e \in h(s)} t_{ee}$$

However, it is well known that for GSMP there are not many quantitative results available [17] besides those based on Markov properties. The difficulty significantly prevents GSMP from modeling the realistic stochastic discrete systems. Thus we transform the GSMDP to a solvable problem in terms of a General State-Space Markov Chain (GSSMC), for GSSMC satisfies the Markov property.

B. General State-Space Markov Chain

In order to maintain the Markov property, the state space $S$ must be extended with the clock-readings of the enabled events. Thus, we obtain an extended state space $X = S \times R_{ee}^k$. We can define a Markov chain $(x, C(x)) : n \geq 0$ with state space $X [17]$. It is called as a GSSMC that corresponds to a GSMP with state space $S$.

Given that an extended state $x \in X$, and $f_e(t)$ is the probability density function of the distribution $G_e$ associated with event $e$, $f(x'| x)$ is defined as

$$f(x'| x) = p(s', s, e^k) \prod_{e \in E} F_e(t_{ee}) t_{ee},$$

where

$$F_e(t_{ee}) = \begin{cases} f_e(t_{ee}), & \text{if } e \in E(s) \cap (e^k \cup (E \setminus E(s))) \\ \delta(t_{ee} - t_{ee}), & \text{if } e \in E(s) \cap (E(s) \setminus e^k) \\ \delta(t_{ee} - \infty), & \text{if } e \in E(s) \end{cases}$$

Here, $\delta(t_{ee} - t_{ee})$ is the Dirac delta function with the property that $\int_0^\infty \delta(t_{ee} - t_{ee}) dt = 0$ for $x < t_{ee}$ and 1 for $x \geq t_{ee}$.

Afterwards, we will introduce two concepts used in the later development. One is Observation Model. Since the time that an event has been enabled is known to the system, this knowledge is sufficient to provide the system with a probability distribution over extended states. Then we set up an observation model based on GSSMC. Let $O = O_s \times O_e$ be the set of observations. An observation is $o = (s, u') \in O$, where a vector $u'$ with elements $u_e$ for each $e \in E$ being the time that the event $e$ has been enabled $(u_e = 0 \text{ if } e \notin E(s))$. These two components are both observable.

Another one is Function obs: $X \times O \times S \rightarrow O$. Given an extended state $x$, an observation of $x$, and the observable part $s'$ of a successor of $x$, it can obtain the observation of $x'$. Namely, $\text{obs}(x, o, s') = (s', \tilde{u}')$, where $\tilde{u}'$ consists of elements $u_e'$ for each $e \in E$, with

$$u_e' = \begin{cases} u_e + t_{ee}, & \text{if } e \in E(s') \cap (E(s) \setminus e^k) \\ 0, & \text{otherwise} \end{cases}$$

Later we characterize GSMDP model for the propose GCAC scheme, i.e. the optimal CAC problem in a multimedia heterogeneous network. The previous schemes for the problem are not competent for dealing with the next generation heterogeneous wireless network. Formulating the problem as a GSMDP model can sove the difficulty. Furthermore, it is convenient for discussing the evolution of a GSMDP model and proposing a feasible learning approach.

IV. GCAC FRAMEWORK

GCAC is a priority-based CAC scheme, which is designed based on the GSMDP model. The design discipline can derive the optimal CAC policy.

A. Formulating the Optimal CAC Problem as a GSMDP

In this section, the optimal CAC problem is formulated as a GSMDP. When a new, horizontal or vertical handoff call arrives, a decision must be made as to whether or not to admit. These time instants are called decision epochs and decisions are called actions in the GSMDP framework. The action chosen is based on the current state of the heterogeneous system. The state information includes the number of calls/sessions of each class of traffic in both the WLAN and the CDMA network. The QoS constrains in both networks discussed in Section II are incorporated by truncating the state space to those state that satisfy the constrains.

The optimality criterion for the GSMDP is the long-run average reward. Network layer blocking probability QoS constraints are accommodated by adding additional linear constraints to the reward function. Ultimately, our goal is to find a policy that for every system state $S$ chooses the correct control actions so that we maximize revenue subject to the QoS constrains.

Here, we consider the CAC process is observed at time instance $t = t_0 + t_1 + t_2 + \cdots$ for call arrivals or departures. In GCAC, the CAC decisions are made only at the occurrence of call-arrival events, which includes new call and handoff call arrivals. And no decisions are made at call departures.

The proposed GSMDP model includes the following five components: the state space, the events, the action space, the police and the reward function.
1) State Space: Let $S$ denotes the state space and $s \in S \subset Z_k^+$. Let $s(t)$ denotes the state of BS at time $t$, where $t \in R^+$ and $K$ is the total number of call types. Define row vector $s(t) = [s_{u,1}(t), s_{u,2}(t), \cdots, s_{u,K}(t)] \in Z_k^+$ represents the number of calls for call class $k$ ($k = 1, \cdots, K$) in the WLAN. Define row vector $s(t) = [s_{e,1}(t), s_{e,2}(t), \cdots, s_{e,K}(t)] \in Z_k^+$ that represents the number of calls for call class $k$ ($k = 1, \cdots, K$) in the CDMA networks. The state vector of the system at decision epoch $t$ is given by

$$s(t) = [s_{u}(t), s_{e}(t)] = [s_{u,1}(t), s_{u,2}(t), \cdots, s_{u,K}(t), s_{e,1}(t), s_{e,2}(t), \cdots, s_{e,K}(t)].$$

(13)

So, the state space $S$ of the GSMDP can be defined as a set of all possible combinations such that the throughput and average packet delay constrains in the WLAN cell and the SIR outage probability constrains in the CDMA cell can be met in (3) and (8), i.e.,

$$S = \{s(t), t \in R_+\}$$

is a finite-state stochastic process.

2) Events, Clocks: The set of all possible event $E$ that can occur in the system is defined as $E = \{a_{u,n,k}, a_{u,h,k}, a_{e,n,k}, a_{e,h,k}, k = 1, \cdots, K\}$, where $a_{u,n,k}$, $a_{u,h,k}$, $a_{e,n,k}$, $a_{e,h,k}$ denote the actions for new, horizontal and vertical class $k$ call arrival in the WLAN, and $a_{e,n,k}$, $a_{e,h,k}$ denote the actions for new, horizontal and vertical class $k$ call arrival in the CDMA cell. For the CAC decisions are made only at the occurrence of call-arrival events, we don’t take call-departure event into account.

In addition, $G_i$ corresponding to the general interarrival time distribution of call arrival, and $G_s$ corresponding to the general service time distribution respectively. In general, each general distribution is represented by two useful concepts consisting of the mean of a distribution $\mu$ and the variation $\sigma^2$ (their notation is uniform with their corresponding distribution).

3) Action Space: When a new or handoff call arrives, the GCAC makes an admission decision. The decision epochs for the GSMDP are the call arrival points. At each decision epoch, the system makes a decision for each possible call arrival that may occur in the time interval. These decisions are collectively referred to as an action $a$. So, the action space in the BS of a WLAN cell is a set of vectors consisting of $3K$ binary elements, i.e.,

$$a_{u,n} = [a_{u,n,1}, a_{u,n,2}, \cdots, a_{u,n,K}] \in \{0,1\}^K, a_{u,h} = [0,1]^K,$$

where $a_{u,n}$ is the action of call, and only take the binary values with 0 for rejecting and 1 for accepting the type of calls. In the same way, the action space $A_e$ in the BS of a CDMA cell is similar to $A_w$. The action space $A_e$ for state $s \in S$ can be characterized as

$$A_e = \begin{cases} \{a = (0,0,0)\}, & \text{if } s = (0,0,0) \leq \frac{B}{TDD, D_h < TD_h, P_{out} < TP_{out}}, k = 1,2, \cdots, K, \text{ otherwise} \\ \{a = (a_{u,n,1}, a_{u,n,2}, \cdots, a_{u,n,K})\}, & \text{if } P_{out} > TP_{out} \end{cases}$$

(15)

4) Police: A control policy $\pi$ determines which actions should be enabled at a given time in a state. An action being enabled in the state means that the system accepts the event. We allow the action choice to depend on the entire execution history of the process, which can be captured in an observation $o \in O$, as described above. Thus, a policy is a mapping from observations to sets of actions: $\Pi = \{\pi:O \rightarrow 2^A\}$, where $\Pi$ denote the set of admissible CAC policies. Given an observation $o$, the set of actions being enabled in state $s$ is $A(s) = \pi(o)$ for the policy $\pi$.

A GSMDP controlled by a policy $\pi$ is a GSSMC with $E(s)$ replaced by $E(\pi(o)) = A(s)$ in the definition of $e^*$, $f(x | o)$, and $ob(s, x, o, \pi)$. The probability density function $f(x | o)$ over $X$ is defined as

$$f(x | o) = \begin{cases} \prod_{\tau \in \tau} f_\tau(t_{\tau} | t_{\tau} > u_{\tau}, s), if s = (s, i) \\ 0, otherwise \end{cases},$$

(16)

where $f_\tau(t_{\tau} | t_{\tau} > u_{\tau}, s)$ is $f_\tau(t_{\tau} | t_{\tau} > u_{\tau})$ if $e \in E(s)$ and $\delta(t_{\tau} - \infty)$ otherwise. The next-state distribution is defined as

$$f(s' | x, o) = p(s', s, e) \prod_{\tau \in \tau} f_\tau(t'_{\tau} | s', x, o),$$

(17)
where \( \tilde{f}(t'|s',x,o) \) is defined as \( \tilde{f}(t'|s',x) \) with \( E^*(o) \) replacing \( E(s) \) and \( E'(obs(x,o,s')) \) replacing \( E(s') \).

5) Reward Function: Based on the actions being enabled in a state, the network earns revenue due to the carried traffic in the cell. So, let the being enabled actions in state \( S \) associates a continuous reward function \( r(s,\pi(o)) \), which is defined as the weighted average reward on pricing, i.e.,

\[
r(s,\pi(o)) = \sum_{k=1}^{K} w_k \cdot s_k
\]  

(18)

Where \( w_k \) is the weighting factor on pricing for each type of call. Notice that when all the weighting factors are equal to 1, the objective is to maximize the system utilization. To simplify the complexity, we only expatiate the reward function in one access network, for in the WLAN or CDMA network the design of the reward function is the same. Only the subscripts of the denotations are different, e.g. \( w_{w,k} \), \( w_{c,k} \) denote the weighting factor in the WLAN and the CDMA cell respectively. Here, the paper use \( w_k \) to represent one of them.

In addition, it is desirable for the network operator to characterize the penalty cost of a system. In the above paragraphs, by satisfying the physical layer QoS constraints in the system we restrict the state space in (14). Here, we also put constrains on the network layer blocking probabilities of certain class of traffic for the network operator. For example, in the case of network congestion, the network operator may want to have a small blocking for premium classes so as to first ensure their services, and a large blocking probability for economic classes. Therefore, we formulate call blocking probability constrains in the reward function by minimizing a linear objective function of the new and handoff call-blocking probabilities, which is widely used in the literature [13][16]. In this paper, we define a similar linear function of all call types,

\[
c(\gamma(s)) = \sum_{k=1}^{K} \omega_k \cdot \gamma(s) \cdot P_{k,obs}(o) + \sum_{k=1}^{K} \omega_k \cdot \gamma(s) \cdot P_{k,handoff}(o) + \sum_{k=1}^{K} \gamma(s) \cdot P_{k,handoff}(o).
\]  

(19)

where \( P_{k,obs}(o) \), \( P_{k,handoff}(o) \), \( P_{k,handoff}(o) \) denote the measured new call blocking and handoff dropping probability respectively during \( \gamma(s) \). By formulating the cost due to the blocking operation in the reward value function, we may control the blocking operation efficiently. For convenience, we denote the reward function,

\[
R(s,\pi(o)) = r(s,\pi(o)) - c(\gamma(s))
\]  

(20)

Thereupon, for a fixed policy \( \pi \), the expected discounted value of an observation \( o \) over an infinite-horizon is given by

\[
v^\pi(o) = \int_{\mathcal{X}} f(x|o) \left[ \int_{0}^{\infty} -\alpha \gamma(s) \right. \\
\left. \left( \int_{\mathcal{X}} f(x'|s) v^\pi(obs(x,o,s')) dx' \right) + R(s,\pi(o)) \right] dx
\]  

(21)

where the parameter \( \alpha \) \((0 \leq \alpha < 1)\) is the discount rate, and indicates the impact or not of future actions and their associates rewards [16]. Via (21), we give attention to both network rewards and the blocking costs; thereby we can find an optimal policy to balance them best.

In this paper, the controller objective is to determine a policy \( \pi^* \in \Pi \) that satisfies

\[
v^\pi^*(o) = \max_{o \in \mathcal{O}} v^\pi(o),
\]  

(22)

for a given observation \( o \). Here, \( \pi^* \) is the so-called optimal policy, and \( \pi(o) \) is the set of actions chosen to be enabled in state \( S \). By combining the QoS constraint with the objective of maximizing revenue, the chosen optimal policy can satisfy the interests of network managers and users both farthest.

B. Q-Learning Solution for the Optimal Policy

Although few available quantitative results for GSMP significantly prevent its applications, our GSMDP framework is defined in terms of GSSMC (it satisfies the Markov property). And the corresponding observation model also satisfies the Markov property. Moreover, a number of researchers have successfully explored many approaches to solve the optimal control problems in Markov environments, such as dynamic programming with a perfect model of the environment [18]. However, they require extremely large state space to model these problems exactly. Consequently, the numerical computation is intractable due to the curse of dimensionality. Also, a priori knowledge of state transition probabilities is required. Alternatively, many researchers turned to use the reinforcement learning (RL) [19] to approach an optimal solution online, which avoided the above disadvantages.

In this paper, we introduce a real time RL technique known as Watkin’s Q-learning [17] as an efficient method for computing \( v^\pi^* \) based on a reformulation of a Bellman equation. We denote Q-value, \( Q'(s,\pi(o)) \), as the expected return starting from \( s \), being enabled the actions \( \pi(o) \), and thereafter following policy \( \pi \):
\[ Q^*(s, \pi(o)) = \frac{1}{\alpha} e^{-\alpha t} R(s, \pi(o))dt + \]
\[ + \int e^{-\omega s} \frac{1}{\alpha} \left( 1 - e^{-\alpha t} \right) R(s, \pi(o)) + \]
\[ \sum_{s' \in S} p(s', s, \pi) e^{-\omega s} \left( obs(x, o, s') \right) ds' \]
\[ = \frac{1}{\alpha} \int \left( 1 - e^{-\alpha t} \right) R(s, \pi(o)) + \]
\[ e^{-\omega s} \sum_{s' \in S} p(s', s, \pi) e^{-\omega s} \left( obs(x, o, s') \right) ds' \]  

The object is to estimate Q-values for an optimal policy. Namely, given optimal Q-values, \( Q^*(s, \pi^*(o)) = \max_{\pi} Q^*(s, \pi(o)) \), the policy \( \pi^* \) defined by \( \pi^*(o) = \arg \max_{\pi} Q^*(s, \pi(o)) \) is optimal.

To learn optimal Q-values, we update our Q-value function recursively at each decision epoch as follows. Available information being used consists of current extended state \( x_t \), observation \( o_t \), selecting and being enabled action set \( \pi(o_t) \), the subsequent state \( s_{t+1} = (s_t, o_t) \), etc. The learning rule is:

\[ Q_{s_t}(s, \pi(o)) = \begin{cases} 
Q_{s_t}(s, \pi(o)) + \eta \left( s_t, \pi(o) \right) \Delta Q_{s_t}(s, \pi(o)), & \text{if } s = s_t \text{ and } \pi(o) = \pi(o_t) \\
Q_{s_t}(s, \pi(o)), & \text{otherwise.} 
\end{cases} \]

where

\[ \Delta Q_{s_t}(s, \pi(o)) = \int e^{-\alpha t} R(s, \pi(o_t))dt + \]
\[ \int e^{-\omega s} \max_{\pi(o_t) \in \pi \mid o(o_t) \in o_t} Q_{s_{t+1}}(s_{t+1}, \pi(s_{t+1}, o_{t+1}))ds'\]
\[ - Q_{s_t}(s, \pi(o)) \]

Here \( \eta \left( s_t, \pi(o) \right) \in [0,1] \) is the step size or learning rate, \( \eta \) is an integer variable to index successive updates.

It has been proved in detail [19][20] that if the Q-value of each admissible action-state pair is visited infinitely often, and if the learning rate is decayed appropriately, then as \( t \to \infty \), the above learning algorithm \( Q_{s_t}(s, \pi(o)) \) converges to \( Q^*(s, \pi^*(o)) \) with probability 1.

V. SIMULATION RESULTS

A. Algorithm Implementation

After the specification of the GSMDP model associated with GCAC scheme, we describe the implementation of the Q-learning algorithm for solving it.

In practice, an important issue is how to represent and store the Q-values. Currently, two different approaches [19][20], which are lookup table and neural network, are very popular to be used. In this paper, we only note the interesting states in which decisions need to be made are those associated with call arrivals. So we avoid the updates of Q-values at departure. It will reduce the amount of computation and storage of Q-values significantly so that the look-up table representation can be used.

When there is a call arrival (new, vertical or horizontal handoff call), the GCAC algorithm first determines if accepting the call will violate QoS. If this case, the call is rejected; else the action is chosen according to (24). Namely, when a call arrives, the Q-value of accepting and being assigned channels to serve this call are determined from the lookup table. If rejection has the higher value, the call is dropped. Otherwise, if acceptance has the higher value, the call is accepted and a chosen channel is assigned to it. To learn the optimal Q-values \( Q^*(s, \pi^*(o)) \), the value function is updated at each transition form state \( s \) to \( s' \) under action \( \pi(o) \) using (25). And training runs typically used a fixed learning rate \( \eta_0(s, \pi(o)) = 0.1 \), which seems to give results even though convergence theorems require decreasing learning rate with time.

B. Experimental Results

In order to evaluate the performance of our solution, we apply a test data set to compare it with the conventional GC scheme [1] and the optimal CAC scheme based on MDP model presented in [9] that we call MCAC.

The values of WLAN parameters used in the analytical and simulative models are summarized in Table I, which is similar to [23] . To reduce the look-up table size, one class of video traffic is considered in the system. And that the scheme only cares about three types of call arrival (new, vertical or horizontal handoff call). In this validation, each video packet has constant packet payload size, as shown in Table I, which is associated with default AC2 specified in the IEEE 802.11e draft. The values of parameters are assigned depending on the IEEE 802.11a specification [27]. For simulation efficiency, the average channel bit rate is assumed to be 11 Mbps.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Notation</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video packet payload size</td>
<td>V</td>
<td>1464 bytes</td>
</tr>
<tr>
<td>Propagation delay</td>
<td>( \zeta )</td>
<td>1 ( \mu s )</td>
</tr>
<tr>
<td>Average channel bit rate</td>
<td>M</td>
<td>11 Mbps</td>
</tr>
<tr>
<td>Slot time</td>
<td>( \varepsilon )</td>
<td>9 ( \mu s )</td>
</tr>
<tr>
<td>time required to transmit a MAC header</td>
<td>( T_{MAC} )</td>
<td>34 ( \mu s )</td>
</tr>
<tr>
<td>duration of a short interframe space (SIFS)</td>
<td>( T_{SIFS} )</td>
<td>16 ( \mu s )</td>
</tr>
<tr>
<td>time required to transmit a request-to-send (RTS)</td>
<td>( T_{RTS} )</td>
<td>20 ( \mu s )</td>
</tr>
<tr>
<td>time required to transmit a clear-to-send (CTS)</td>
<td>( T_{CTS} )</td>
<td>14 ( \mu s )</td>
</tr>
<tr>
<td>time required to transmit an ACK</td>
<td>( T_{ACK} )</td>
<td>14 ( \mu s )</td>
</tr>
<tr>
<td>time required to transmit a physical layer header</td>
<td>( T_{PHY} )</td>
<td>6 ( \mu s )</td>
</tr>
<tr>
<td>arbitration inter-frame space number (AIFS)</td>
<td>( AIFS )</td>
<td>1(25 ( \mu s ))</td>
</tr>
<tr>
<td>minimum contention window</td>
<td>( CW_{min} )</td>
<td>7</td>
</tr>
<tr>
<td>maximum contention window</td>
<td>( CW_{max} )</td>
<td>15</td>
</tr>
</tbody>
</table>
For CDMA network, the parameters used in the simulation are shown in Table II, which is similar to [24]. Specifically, the transmission rate for the video traffic in the CDMA network is 240 Kbps and corresponds to an equivalent spreading gain $N = 16$.

The most difficult parameters to choose in our experimentation are how to simulate the call arrival processes. For the limitations of the traditional Poisson model for network arrival processes have been demonstrated in recent studies, self-similar models corresponds better with the traffic. In our simulation, we generate synthetic traffic traces containing three call types, as described in section II. And the self-similar call arrival process was generated by multiplexing calls from 1000 ON/OFF-sources per class [21]. After some experimentation, the following parameter settings were chosen: The ON-period distribution had parameters $\theta_{ON} = 0.75$ and $\alpha_{ON} = 0.21$, and the OFF-period had parameters $\theta_{OFF} = 400$ and $\alpha_{OFF} = 1.9$. In the ON-periods, calls were generated at rate $1 \text{[s}^{-1}]$. This parameter combination resulted in a Hurst parameter $H = 0.8$. The inter-arrival times were linearly scaled by multiplication by $\lambda^{38.6}$ to get an approximate arrival rate of $\lambda$. Following similar notations defined in the previous section, let us denote $\lambda_{ON}$, $\lambda_{OFF}$, $\lambda_{total}$ to be the approximate call-arrival rates. In the simulation, the call-arrival rates range from 0.5 to 3.5 for each call type.

1) Average Reward Comparison

We firstly compare the average reward in (20) from the proposed scheme GCAC to that from two other scheme. schemes. Fig.1 shows the average rewards earned in different. We can see that the reward earned in the proposed scheme is always more than that in two other schemes, and the reward is the least in the traditional GC scheme without any optimal control.

2) QoS Guarantee in the Proposed Scheme

Besides the average reward represent better performance, the GCAC scheme can also guarantee QoS at physical layer and network layer. In the simulation, two upper bounds on the horizontal and vertical handoff dropping rates, i.e. $P_{hh}^* = 2\%$ and $P_{vh}^* = 3\%$, are set. And a target SIR outage probability in the CDMA network is 0.01. A target throughput probability in the WLAN is 1.1M.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Notation</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>data transmission rate for video traffic</td>
<td>$R$</td>
<td>240 kbps</td>
</tr>
<tr>
<td>target SIR for video traffic</td>
<td>$\omega$</td>
<td>10 DB</td>
</tr>
<tr>
<td>number of resolvable paths</td>
<td>$L$</td>
<td>5</td>
</tr>
<tr>
<td>channel estimation error variance for video traffic</td>
<td>$\xi^2$</td>
<td>0.05</td>
</tr>
<tr>
<td>estimated average channel gain for video traffic</td>
<td>$</td>
<td>h^2</td>
</tr>
</tbody>
</table>

Table II. CDMA Parameters Used in the Simulation

Figure 1. Average Reward for GCAC, MCAC and GC Schemes

(a) new call-blocking rates

(b) horizontal handoff dropping rates

(c) vertical handoff dropping rates

Figure 2. Performance Comparison for GCAC, MCAC and GC Schemes
Fig. 2(a) shows the performance of the system about the new call blocking rates. The associated horizontal and vertical handoff dropping rates can be guaranteed in the proposed GCAC scheme in Fig. 2(b) and Fig. 2(c). The simulation results shows that the GCAC scheme has the better performance than two other ones.

VI. CONCLUSION

In this paper, we have proposed an optimal priority-based CAC mechanism GCAC based on the GSMDP and Q-learning solutions for various heterogeneous systems. The proposed scheme considers the QoS constraints in physical and network layer, namely, throughput and packet delay, SIR outage probability and call blocking probability. To better model the real world of system behaviors, we relax the Markov assumptions and formulate a priority-based CAC mechanism as a GSMDP model, so as to remain optimal. We have shown that the use of the GSMDP can determine the optimal CAC policy to achieve the predetermined goals, whereas the use of Q-learning can greatly reduce the effort of unnecessary recomputation. And we have evaluated the performance of the proposed GCAC scheme and compared it with the performance of traditional schemes (i.e. GC and MCAC schemes) by extensive simulations. Results have shown that the proposed scheme can not only obtain best performance, but also successfully maintain the QoS metrics at the required level.

REFERENCES


Wen Chen received the Ph.D. degree from the Computer Science and Engineering Department, Shanghai Jiao Tong University, Shanghai, China, in 2006. Since July 2006, she has been an instructor in the College of Information Science and Technology Engineering, Donghua University, Shanghai, China. Her principal research interests include QoS support, wireless sensor network, and network optimization. Her other research interests include network management and network security. She has more than 12 international conference and five international journal publications.

Jinming Yu received the BS degree from the Electrical Engineering Department, Shanghai University of Science and Technology, Shanghai, China, in 1991. He is currently an associate professor in the College of Information Science and Technology Engineering, Donghua University, Shanghai, China. His main research interests include design, analysis, and optimization of wireless and optical communication networks.

Feng Pan received the Ph.D. degree from the Computer Science and Engineering Department, Shanghai Jiao Tong University, Shanghai, China, in 2007. Since 2007, he has been an instructor in the College of Information Science and Technology Engineering, Donghua University, Shanghai, China. His main research interests include QoS support, energy-efficient routing in wireless sensor network, and resource management in wireless multimedia networks.