Integrated Cross-Layer Design of Utility-Based Connection Admission Control in Packet-switched OFDM Wireless Networks

Hu Nan, Wand Xiaoxiang and Wu Weiling
Beijing University of Posts and Telecommunications, Beijing, China, e-mail: hunanbupt@gmail.com

Abstract—This paper proposes a new utility-based connection admission control algorithm called UCAC for real-time services in packet-switched OFDM wireless networks. The main idea of the algorithm is to periodically adjust the acceptance ratios according to the system information including channel condition in the physical layer, queue status in the radio link layer and traffic load in the network layer. By using the cross-layer approach, the algorithm takes the connection-level and packet-level performances of users into consideration. Furthermore, a user utility function is defined as the measure of how efficiently the bandwidth resource is used. The objective of the acceptance ratios adjusting is to maintain the maximum total user utility while guaranteeing the packet-level and call-level QoS requirements for real-time services. The performance of the algorithm is validated by extensive simulation results.

Keywords—Connection admission control, cross-layer, packet-switched.

I. INTRODUCTION

Mobile cellular networks are evolving into packet-switched wireless networks providing a wide range of high-data-rate services. To meet this requirement, in the physical layer, orthogonal frequency-division multiplexing (OFDM) is a promising wireless access technology due to its excellent performance of combating the multipath delay spread in frequency selective fading channels [1]. However, in the higher layer, the quality of service (QoS) performances especially for the real-time services (e.g., delay) can not be guaranteed without an efficient connection admission control (CAC) algorithm.

In recent years, considerable efforts have focused on the CAC problems and many schemes have been proposed in [2-4]. However, most of these studies just took the network-layer call-level performance into consideration. For packet-switched wireless networks, packet-level performance measures need to be considered in addition to the call-level performance measures in designing CAC.

In this paper, we propose a CAC algorithm simultaneously maximizing bandwidth utilization and guaranteeing QoS requirements for real-time services in packet-switched OFDM wireless networks. The main features of our scheme are described as follows:

1. Our scheme obtains the optimal point where the number of satisfied users in system is maximized and bandwidth resource is most efficiently used by maximizing total user utility defined in this paper.

2. By using cross-layer approach, the packet-level and call-level performance measures are both considered in our scheme.

3. The mechanism of our scheme is periodical. The optimal acceptance ratios are periodically determined according to the varying system conditions. Therefore, the scheme adapts a wide range of system conditions.

II. HELPFUL HINTS

A. Figures and Tables – Subsection Example

Position figures and tables at the tops and bottoms of columns. Avoid placing them in the middle of columns.

As shown in Fig. 1, an OFDM wireless cellular system with hexagonal cells of equal size is considered here and each cell contains a centrally located base station. The system serves multiple classes of traffic and a separate queue is maintained for each class. The CAC algorithm decides the admission of an incoming connection in that class.

![Fig. 1 System model.](image-url)

The connections in the same traffic class have the same QoS requirements, and therefore, the packets from the ongoing connections in the same class are buffered in the same queue [5][6]. The transmitter retrieves the head of line packets from each queue and transmits them to the destination. Different transmission rates and number of subcarriers are allocated to the packets according to the traffic QoS requirements.

We focus on analyzing one queue for a kind of real-time services and derive the CAC algorithm for the queue. Note that the method can be applied to any one of queues in the system. In the network layer, we assume that the connection arrival follows a Poisson process with arrival rate λ. The duration of a connection is assumed to be exponentially distributed with average 1/µ and the cell residency time of a connection is also assumed to be
exp(\ln(5) \cdot \ln(2) \cdot \ln(\frac{\ln(5)}{2} + 1) - \frac{1}{2} e^{\ln(5)}}

where $u_{\text{SNR}_m}$ is the average signal-to-noise ratio (SNR) of subcarrier $m$ and $P_{\text{SNR}_m}$ is the target bit error rate. The service rate $R_k$ of queue $k$ is:

$$R_k = \sum_{m \in M_k} R_m$$  (2)

where $M_k$ is the set of subcarriers allocated to queue $k$. The exact pdf of $R_k$ turns out to be intractable. However, since the $\{R_m | m \in M_k\}$ are i.i.d random variables, the pdf of $R_k$ can be approximated as Gaussian distribution using the central limit theorem, when the number of allocated subcarriers is large enough. The pdf of $R_m$, denoted by $f(R_m)$, can be expressed as:

$$f(R_m) = \frac{1}{\sqrt{2\pi \sigma_{R_m}^2}} \exp\left(-\frac{\left(R_m - \sum_{m \in M_k} u_{\text{SNR}_m}\right)^2}{2\sigma_{R_m}^2}\right)$$  (3)

where $u_{\text{SNR}_m}$ and $\sigma_{R_m}$ are the mean and variance of $R_m$, respectively.

### III. TOTAL UTILITY FUNCTION

If we consider the bandwidth resources as a public good, the best policy to share this good is the one that maximizes social welfare, which is the sum of consumers’ and producers’ surpluses. In order to maximize social welfare, the total user utility should be maximized [8]. In terms of economics, utility functions describe users’ level of satisfaction with the perceived Quality of Service. In this paper, we define the user utility $U_i$ for real-time services denoted by type-k as follows:

$$U_i = \begin{cases} 0, & \text{Pr}(W > LT) \leq p_{\text{pser}} \\ 1 - \exp\left(\frac{\text{Pr}(W > LT)}{p_{\text{pser}}} - 1\right), & 0 \leq \text{Pr}(W > LT) < p_{\text{pser}} \\ \end{cases}$$  (4)

where $W$ is the amount of time a packet spends waiting in queue and $LT$ is the delay bound of this kind of packets. $G$ determines the steepness (i.e., sensitivity of the utility function to delay requirement) and we set $G = 16$ according to simulation results. $U_i$ is a function of the packet delay-bound violation probability, which represents the main QoS metrics at packet-level in cellular networks. The total type-k user utility $G_{i,k}$ in cell $i$ is given by

$$G_{i,k} = N_{i,k} \cdot U_i$$  (5)

where $N_{i,k} \in \{0, 1, \ldots, D_k\}$ is the number of type-k connections in cell $i$ and $D_k$ is the maximum number of type-k connections that can be buffered in queue $k$.

According to the definition of total utility function, we find that with light traffic load, the satisfaction degradation caused by the increase of connections is not significant. In this case, the total utility increases when the number of connections increasing. In contrast, when the traffic load is high, the satisfaction degradation caused by the increase of connections becomes substantial. In this case, the total utility decreases when the number of connections increasing. Thus, there exists an optimal number of connections that maximizes the total user utility.

### IV. CALL ADMISSION CONTROL ALGORITHM

Based on the arguments described in Section 3, a call admission control algorithm is proposed. By periodically adjusting the new connection acceptance ratios according to system information, the proposed algorithm can maintain the number of connections at the optimal value that maximizes total user utility. Note that the system information includes channel condition in the physical layer, queue status in the radio link layer and traffic load in the network layer. Therefore, the cross-layer approach is needed for information transfer. To reduce the signaling overhead for information transfer, our algorithm has a periodic structure. All the information exchange and control parameter computations happen only once at the beginning of each control period. Specifically, at the beginning of a control period, each cell $i$ sends the information, including the number of active connections presented in the cell and the number of new connections and handoff connections admitted at the last control period, to its adjacent cells. Then, cell $i$ uses the received information and those available locally to decide the new connection acceptance ratios using the technique described in the next part of this section. Finally, the system uses the computed acceptance ratios to admit the connection requests into cell $i$.  

#### A. Computing Acceptance Ratio

There are three steps involved in computing the acceptance ratios.

Step 1) Derive the mapping between the number of connections $N_{i,k}$ buffered in queue $k$ of cell $i$ and the packet delay-bound violation probability by using a discrete time Markovian chain. Let $s_m$ denote the steady state probability of the queue where there are $m$ packets. Define $p_{\text{pser}}$ as the transition probability that the number of packets changes from $x$ in the current control period to $y$ in the next control period. When there are $N_{i,k}$ connections...
buffered in the queue, according to the distribution of the packet arrival we can obtain \( p_0 \) as follows:

\[
P_0 = \begin{cases}
\sum_{f=1}^{y-x} \frac{e^{-\lambda_f} (\lambda_f T)^x}{x!} \cdot \Pr(f), x > y \\
\sum_{f=1}^{y-x} \frac{e^{-\lambda_f} (\lambda_f T)^x}{x!} \cdot \Pr(f), x < y \\
\sum_{f=1}^{y-x} \frac{e^{-\lambda_f} (\lambda_f T)^x}{x!} \cdot \Pr(f), x = y
\end{cases}
\]  

(6)

\[x - y = \Delta \]

Here \( \nu \in \{0, 1, 2, \ldots, N_i V\} \) is the number of arriving packets in the control interval \( T \) and \( f \in \{0, 1, 2, \ldots, \min(F, x)\} \) is the number of transmitted packets in the control interval. \( V \) is the maximum number of packets that can arrive from one connection in \( T \) and \( F \) is the maximum number of packets that can be transmitted in \( T \). \( \Pr(f) \) is the probability that \( f \) packets are transmitted in \( T \) and according to (3), it can be given by

\[
\Pr(f) = \int_{x-(y-x)}^{x} \frac{1}{\sqrt{2\pi} \sum_{m=1}^{y} \sigma_m^2} \exp \left[ - \frac{\left( R - \sum_{m=1}^{y} u_m \right)^2}{2 \sum_{m=1}^{y} \sigma_m^2} \right] dR
\]

where \( R \) is the service rate in the queue and \( g \) is the size of a packet. For the steady state probability \( s_m \), we can get a set of equations as follows:

\[
s_y = \sum_{x=0}^{C} s_{y,x}, y \in \{0, 1, \ldots, C\}
\]

(8)

\[
\sum_{y=0}^{C} s_y = 1
\]

(9)

where \( C \) is the maximum number of packets that can be buffered in the queue. By solving the equations (8) and (9), we can obtain the steady state probability \( s_m \). Using the steady state probability \( s_m \), we can obtain the packet delay-bound violation probability when there are \( N_i,k \) connections buffered in the queue:

\[
\Pr(W > LT | N_i,k) = \sum_{y=0}^{C} s_y \cdot \Pr(j)
\]

(10)

\[
\Pr(j) = \int_{0}^{\frac{R}{\nu LT}} \frac{1}{\sqrt{2\pi} \sum_{m=1}^{y} \sigma_m^2} \exp \left[ - \frac{\left( R - \sum_{m=1}^{y} u_m \right)^2}{2 \sum_{m=1}^{y} \sigma_m^2} \right] dR
\]

(11)

It is noted that \( s_j \) is the function of \( N_i,k \) so we can find that \( \Pr(W > LT | N_i,k) \) is expressed as a function of \( N_i,k \).

Step 2) Estimate the time-dependent mean and variance of the number of connections in queue \( k \) of cell \( i \). Note that this step is the same as our previous work \([6]\) and readers can find more details in \([6]\). The number of connections in queue \( k \) of cell \( i \) at time \( t \), \( N_i,k(t) \), is the sum of two parts: 1) the number of existing connections which are already in queue \( k \) of cell \( i \) or its adjacent cells and 2) the number of new connections which will be admitted into queue \( k \) of cell \( i \) and its adjacent cells during the period \((0, t)(0 \leq t < T)\). Let \( LE_{i,k}(t) \) and \( LN_{i,k}(t) \) denote the number of existing and new connections in queue \( k \) of cell \( i \) at time \( t \), respectively. An existing connection in queue \( k \) of cell \( i \) will remain in cell \( i \) with the probability \( P_i(t) \) or will handoff to an adjacent cell \( j \) with probability \( P_j(t) \). A new connection admitted into queue \( k \) of cell \( i \) at time \( t \) will stay in cell \( i \) with probability \( Q_i(t) \) or will handoff to an adjacent cell \( j \) with probability \( Q_j(t) \). The mean of active connections in queue \( k \) of cell \( i \) at time \( t \) is given by

\[
E[N_{i,k}(t)] = E[LE_{i,k}(t)] + E[LN_{i,k}(t)]
\]

(12)

\[
E[LE_{i,k}(t)] = LE_{i,k}(0) P_i(t) + \sum_{j \in A} LE_{j,k}(0) P_j(t)
\]

(13)

\[
E[LN_{i,k}(t)] = \alpha_{i,k} \lambda_{i,k} Q_i(t) + \sum_{j \in A} \alpha_{j,k} \lambda_{j,k} Q_j(t)
\]

(14)

Here \( E[y] \) is the mean of \( y \). \( A \) means the set of adjacent cells of cell \( i \) and \( \alpha_{i,j} \) is the acceptance ratio of queue \( k \) in cell \( i \). Note that in our algorithm, the new connections are admitted according to the acceptance ratios, but handoff connections are always admitted except the new connection acceptance ratios are smaller than zero. The products \( \alpha_{i,j} \lambda_{i,j} \) in (14) are obtained by the estimation technique described in \([3]\). \( P_i(t) \), \( P_j(t) \), \( Q_i(t) \) and \( Q_j(t) \) can be calculated as follows:

\[
P_i(t) = \exp\left(-\frac{1}{\mu} + \frac{1}{h}\right) t
\]

(15)

\[
P_j(t) = \frac{1}{6} \left(1 - \exp\left(-\frac{t}{h}\right)\right) \exp\left(-\frac{t}{\mu}\right)
\]

(16)

\[
Q_i(t) = \frac{1}{t} \int_{0}^{t} P_j(t-x) dx
\]

(17)

\[
Q_j(t) = \frac{1}{t} \int_{0}^{t} P_j(t-x) dx
\]

(18)

Similarly, the variance can be obtained as follows:

\[
V[N_{i,k}(t)] = V[LE_{i,k}(t)] + V[LN_{i,k}(t)]
\]

(19)

\[
V[LE_{i,k}(t)] = LE_{i,k}(0) P_i(t)(1 - P_i(t)) + \beta_i \sum_{j \in A} LE_{j,k}(0) P_j(t)(1 - P_j(t))
\]

(20)

\[
V[LN_{i,k}(t)] = \alpha_{i,k} \lambda_{i,k} Q_i(t)(1 - Q_i(t)) + \beta_i \sum_{j \in A} \alpha_{j,k} \lambda_{j,k} Q_j(t)(1 - Q_j(t))
\]

(21)

where \( V[y] \) means the variance of \( y \). Because the number of the connections which can be served in OFDM system is large enough, it is reasonable to use the central limit theorem when the traffic load is high. Thus, the number of connections in queue \( k \) of cell \( i \) can be approximated by a Gaussian distribution\([3]\)[6][7] with mean \( E[N_{i,k}(t)] \) and variance \( V[N_{i,k}(t)] \). Note that \( E[N_{i,k}(t)] \) and \( V[N_{i,k}(t)] \) are the functions of the new connection acceptance ratios \( \alpha_{i,k} \).

Step 3) Compute the optimal acceptance ratio subject to the maximum the value of total user utility \( G_{i,k} \). According to
where the step 1 and 2, we can obtain the packet delay-bound violation probability (PDV) as:

\[
\Pr(W > LT) = \int_0^\infty \Pr(W > LT | N) \cdot \exp\left(-\frac{(N - E[N, t](t))^2}{2V[N, t](t)}\right) dN
\]  

(22)

According to (4), (5), (10), (11) and (22), we can obtain the mapping between the total type-k user utility of cell i during control period T and the new connection acceptance ratios for type-k traffic. We denote \(G_\alpha(a, t)\) as the total utility function of new connection acceptance ratios. Finally, we can obtain the optimal acceptance ratios using bisection method.

V. SIMULATION RESULTS

The performance results of the proposed algorithm are obtained by use of discrete event simulation. The network topology includes 49 cells surrounded by 84 additional virtual cells that wrap around and serve to avoid the border effect. The cell is hexagonal shape. The mobile user moves to any of its adjacent cells with equal probability and its moving direction remains the same during the whole connection. We consider one queue in a cell and 10 subcarriers are allocated for the queue. The bandwidth of each subcarrier is 80 KHz and the average SNR is 6 dB for all of the subcarriers. We set the target BER at 10^-3. The size of a packet is 800 bits and set LT=150 ms. The length of a time slot is 10 ms. The control period is 20 s. We assume that \(P_{QoS}=3\%\) to ensure the adequate service quality. We use connection load \(\rho\) in simulations which can be obtained as: \(\rho = \lambda_c \cdot \mu\).

Restated, the simulation parameters are the same as our previous work [6] for the purpose of comparison.

To evaluate the effectiveness of the proposed algorithm, the threshold-based admission control (TCAC) scheme proposed in [5] is also examined. For threshold-based admission control, a connection request is accepted when the number of ongoing connections is less than a predefined threshold, otherwise it is rejected. We employ the well-known guard channel policy in the threshold-based scheme for supporting the priority of handoff connections over new connections. Furthermore, the DCA scheme proposed in our previous work [6] is examined as well.

We first examine the performance of the CAC algorithms for a wide range of network-layer connection load. Fig. 2 (a) shows the connection blocking probability versus the connection load. When the connection load increases, the connection blocking probability increases for all of the CAC algorithms. And, it can be seen from Fig. 2 (a) that UCAC has lower connection blocking probability than TCAC and higher blocking probability than DAC. Fig. 2 (b) shows the handoff connection dropping probability under different connection loads.

It can be seen from this figure that all the algorithms can always satisfy QoS requirements of handoff connection dropping probability even with high connection loads. Fig. 2 (c) shows the PDV of different CAC algorithms versus the connection load. We can observe that all the algorithms can satisfy the requirement of PDV under different connection loads. Fig. 2 (d) illustrates the total user utility for a wide range of connection loads. As expected, the proposed UCAC has the highest total user utility. From these figures, it can be concluded that all the three algorithms can guarantee the connection-level and packet-level QoS requirements for varying network-layer circumstance and UCAC has the highest total user utility.

Then, we examine the performance of the CAC algorithm for varying the packet arrival rate in the data link layer. Fig. 3 (a) shows PDV versus the packet arrival rate. We can observe that UCAC and DAC can always satisfy QoS requirements of PDV no matter how the packet arrival rate is but TCAC fails to do so when the packet arrival rate is high. Due to the adaptive acceptance ratio adjusting, when the queue becomes congested because of the high packet arrival rate, UCAC and DAC admits lesser number of incoming connections in order to support the QoS requirements of ongoing connections, but TCAC fails to do so. It means that UCAC and DAC can guarantee the QoS under varying the data-link-layer conditions. Furthermore, UCAC has higher total utility than DAC, which is shown in Fig. 3 (b).

Fig. 2 The algorithm performances under different connection loads.

The impacts of physical-layer channel quality on the PDV and total user utility of the algorithms are shown in Fig. 4. From Fig. 4 (a), it can be observed that at lower SNR, TCAC cannot guarantee PDV. On the other hand, when SNR is high, TCAC keeps PDV in a very low value. But UCAC and DAC can always maintain the PDV in an acceptable level. This is due to the fact that UCAC and DAC can adaptively change the number of admitted
connections according to the channel quality. Furthermore, UCAC again obtains the highest total user utility of the three algorithms.

VI. CONCLUSION

In this paper, we propose a utility-based CAC algorithm for real-time services in packet-switched OFDM wireless networks. By using cross-layer approach, our algorithm adaptively adjusts the new connection acceptance ratio according to the varying parameters of physical, data link and network layers. The main advantages of the algorithm are its high bandwidth utilization and adaptability to varying system conditions. Simulation results have shown that the algorithm maximizes the total user utility and, at the same time, can guarantee QoS requirements of the users at both call level and packet level.

ACKNOWLEDGEMENT

The work is supported by the National High Technology Research and Development Program of China (863 Program: 2007AA01Z2652 and 006AA01Z258).

REFERENCES