# QoS-aware Adaptive Call Admission Control in Multiuser OFDM Wireless Network

YU, Xi



A Thesis in Partial Fulfillment

of the Requirements for the Degree of Master of Philosophy

in

Information Engineering



September 2008

# Acknowledgement

I would like to thank my supervisor, Professor Zhang Yingjun, for continuous and patient guidance in this work. This work would not have been completed without her useful discussion, revision and comments. During the study period, I learned how to innovative research work. Thank all the fellow postgraduate students in my lab. I will remember the help and joy brought by you for ever.

Last but not least I also would like to my parents who support me unconditionally in any situation. Without their love, I can not have the courage to step forward on the way I chose.

# 摘要

雖然動態資源分配被廣泛應用在無線網路中，以期望在保障用戶服務品質的同時滿足某些系統性目標，但是在沒有適當的接入控制機制的條件下，資源分配問題無法實現。在一種特定的資源分配方案 一 動態子載波分配的情況下，我們設計了一個在多用戶共存正交頻分複用無線網路中，關注用戶服務品質的適應性接入控制策略。

通過分析動態子載波分配策略的條件，並且根據物理通道及載波分配的統計資訊，我們設計的接入控制策略能夠有效地保障系統的服務品質。因爲動態子載波分配策略充分利用了多載波網路中用戶的多樣性和衰落通道的時變性，這個接入控制策略設計同樣得益於這樣的分配方式。爲了進一步提高系統性能，根據用戶帶給系統收益的能力，用戶被有選擇性的接入。 這個控制接入方案通過利用已有的接入記錄和對未來的預測迅速的適應網路的變化，從而做出接入控制決定。

在整篇論文中，我們嚴格的分析推導以得到這樣關注用戶服務品質的適應性接入控制策略。並且通過大量的仿真證實，我們的接入控制方案可以以較低的複雜程度同時實現系統服務品質保障和最大化系統平均收益。

# Abstract

Though adaptive resource allocation is proposed in wireless networks for obtaining both QoS guarantee and some system-wide utility target, the problem may not be feasible without an appropriate call admission control (CAC) scheme. With respect to a specific adaptive resource allocation scheme—dynamic sub-carrier allocation (DSA), we design a QoS-aware adaptive call admission control strategy for multi-user OFDM wireless networks.

By analyzing the optimal criteria of dynamic sub-carrier allocation, our CAC strategy effectively support QoS provisioning in system-wide through incorporating the information of physical channel and the dynamic sub-carrier allocation. Since the sub-carrier allocation scheme makes use of multi-user diversity in multi-carrier system and the time diversity of the fading channel, the CAC strategy can also share the benefits through closely following the sub-carrier allocation feature. For better system performance, the users are selectively admitted according to their revenue contribution potential to the whole system. This call admission control strategy adapts to multi-user OFDM wireless networks quickly with learning how to do decision based on history decisions and future predictions.

In the thesis, we demonstrate the way how the QoS-aware adaptive call admission control strategy is derived rigorously, and the wide simulation results validate that the CAC strategy can achieve QoS provisioning and average system revenue maximum simultaneously with low complexity.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction and Background

The wireless communication system has been a hot topic in both research and commercial applications for several decades. With the advances of wireless technology, modern communication systems such as WLAN, WiMax, and so on are able to support high data rate for users. For example, WiMax can achieve a highest data rate of 75 Mbps per channel as proposed in the standard. Orthogonal Frequency Division Multiplexing (OFDM) is adopted in these standards as the physical transmission technique, because it can effectively combat multi-path fading to improve the achievable service rate. What is more attractive is that multiple users can be served simultaneously without interfering with each other in the OFDM system. Various adaptive resource allocation schemes are designed in the OFDM context for achieving certain system objective and satisfying each user's demand in the meanwhile. However, adaptive resource allocation alone is not a strong guarantee to obtain the two targets well, for the network may be overloaded from time to time without an effective call admission control (CAC) scheme. Being a critical component in the wireless network, CAC endeavors to prevent the network from being overloaded, and in the meantime maximizes the average system revenue by selectively admitting users. In wireless networks, the CAC component is more important than its counterpart in wired networks because of the more volatile communication environment.

In this thesis, we focus on the design of an effective CAC strategy in the multi-user OFDM wireless network where the dynamic sub-carrier allocation (DSA) is adopted in the lower layer. Typically there are a large number of orthogonal sub-carriers available in OFDM system. By assigning sub-carriers to users according to the instantaneous channel conditions of the users, DSA can greatly enhance the spectral efficiency of wireless networks and support multiple transmission sessions simultaneously. Since various users suffer different degrees of fading, the possibility that all users go into deep fading over all the sub-carriers is pretty small. It implies more users can be accommodated, if the multi-user diversity of the wireless fading channel is fully exploited by the DSA component.

On the other hand, unlike dynamic sub-carrier allocation, CAC predicts the influence of admitting a newly arrived user on the performance of the system before the user is actually enrolled. Most existing CAC schemes assume that the total amount of resource as well as the resource needed by each type of user is constant and known. Consequently, only the dynamic arrival and departure process of users need to be considered. In contrast, the design of CAC is much more challenging in OFDM systems with dynamic sub-carrier allocation. This is because the actual data rate achieved by each user is a function of the channel statistics altered by the DSA as well as the channel correlation between different users. Therefore, it is hard to quantify the amount of resource needed by a newly arrived user through simple calculation as the assumption in existing CAC schemes. To predict the performance of the system with a tolerable complexity is the major obstacle.

In this work, we design an adaptive CAC strategy for multiple user OFDM wireless networks. The proposed CAC strategy is composed of two stages: at the first stage, the strategy aims to decide the admissibility of a newly arrived user. While at the second stage, it aims to maximize the average system revenue by only admitting the "best" users from those are already considered to be admissible in the first stage. The strategy is derived from the analysis of the specific problems, and validated through simulations. The strategy effectively protects existing user's QoS provisioning, and then it performs very close to the optimal result in terms of revenue collected by the network system with low complexity.

In the rest of this chapter, we briefly review the development of the CAC scheme design over the past few decades, then the emerging new challenges when CAC is applied in the OFDM wireless network with dynamic sub-carrier allocation, and finally the organization of the thesis.

# 1.1 Background

## 1.1.1 Brief Review of CAC

The ultimate goal of CAC is to protect the service quality for existing users from being interrupted or degraded too much upon the admission of a new one. It defines the capacity of the network in terms of user number, and the system will corrupt if the user number exceeds the boundary given by the call admission control. Furthermore, besides providing the admission region, CAC plays an important role in obtaining some system-wide objective. For example, service providers will charge various users at diverse rates according to their different applications. Without violating the admission boundary, the CAC scheme can select the "best" users to maximize the average revenue for the network system.

If the resource left in the network is not enough for a new coming user, it will be blocked by the network, and then an indicator for each type of user called blocking probability can be obtained in the long run. The blocking probability is one of the important metrics to evaluate how good the CAC scheme is. In traditional wireless cellular networks [1], [2], [3], the limited number of links is taken as the only resource that the CAC schemes should pay attention to. It is because once a mobile terminal grasps a transmission link, the service rate on the link is presumed to meet the mobile terminal's demand during its entire service session. As a result, the performance of the network is tractable by modeling the evolution of user number in the system as a time-continuous Markov chain bounded by the number of total links. The best call admission control scheme can be found out for any blocking probability target off-line, given the link capacity of the network. Usually the target blocking probability for each user is determined by their priorities. For example, in the cellular communication system, the calls are classified as new calls in the cell and handoff calls from other cells. The new call may be rejected even the admission region boundary is not exceeded, because some resource is reserved by the CAC policy for handoff calls that have higher priority.

However, due to the time-varying nature of wireless channel, there is a possibility that the service rate on one link may dissatisfy one subscriber from time to time. In order to make fully use of the scarce resource in wireless networks, the future communication system standards as 802.11, 802.16 intend to replace link-based communication with packet-based communication. The wireless channel is modeled as a finite state Markov chain [4], [5], [6], so correspondingly the service from physical layer is denoted as an aggregate of amount of packets that can be transmitted time slot by time slot. With the user dynamic and the channel fluctuation, the point to point transmission in wireless channel is like a queueing system with either multiple-class service or single-class service. Through controlling some parameters of the queueing system, call admission control schemes are designed for constraining each type user's blocking probability below a predefined target [7], [8], [9].

Interference is another problem of the wireless network, which is caused by the broadcasting characteristic of the wireless channel. One user's performance may be degraded due to interferences from users in the same network cell or adjacent cells, so the signal to interference ratio (SIR) of on-going users is denoted as a QoS metric for wireless networks. The wireless channel varies with time as a stochastic process, so the SIR of each user fluctuates accordingly. The outage concept is introduced to evaluate the performance of the wireless network, which means when the SIR of one user drops below the predefined threshold, we can assert one outage event occurs. By adjusting threshold according to the distribution of SIR in ergodic channel, the threshold based mechanism is applied as connection admission control directly [10], [11]. With respect to dynamic behavior of users and the interferences between different users in the physical layer, some work [12] gives a Super Markov model to track the fluctuation of SIR of each type of user in multiservice wireless networks. Though it builds up a rigorous theoretical frame, unfortunately the problem is computational prohibitive because of an over large state space of this model constructed for the network. Under this condition, QoS can be roughly guaranteed through worst case analysis, and the high complexity forbids the method to be implemented as a good on-line admission control scheme. Since adaptive resource allocation is employed to improve the spectral efficiency, the call admission control schemes are improved by incorporating information from adaptive resource allocation obviously [13], [14].

In addition to making provision for QoS demands of all the users, the CAC schemes can differentiate users to maximize the system revenue. Different kinds of users bring different amount of revenue or cause different cost to the network. If the network is modeled as a multi-state dynamic system, it is better to develop a CAC mechanism to make the network has a higher probability to stay in the high revenue states by admitting the "better" user. A good CAC scheme can always lead the system to the high revenue states by rejecting an improper coming user and reserving the resource for a "better" user coming later.

For the sake of tractability, the evolution of the network is modeled as a Semi-Markov Decision Process (SMDP) in which the state space is determined by the type of user that can give rise to various rewards or different costs to the system. This kind of mathematical model has been investigated theoretically [15]. [16] demonstrates how to apply those methodologies, e.g. policy iteration, value iteration and linear programming, to call admission control scheme design in the network. In most cases, with these methods an optimal policy can be found out for either maximizing the long term revenue or minimizing the long term cost of the network. These methods suffer from curse of dimension, which means the computing burden increases exponentially with the state space of SMDP; meanwhile the state space is usually large in a typical communication system. Instead of solving the SMDP problem directly, some threshold-based CAC schemes calculate the thresholds with iterative algorithms that are derived from the feature of optimal algorithms for SMDP [17], [18], [19]. In fact, to solve the SMDP problem is a procedure of searching optimal solution in a specific solution space, so it is very natural to borrow the ideas from artificial intelligence that concentrates on how to search optimal solution efficiently in a large solution space. Besides genetic algorithm that is applied to optimal admission control design, given a full scope of the system [20], [21], a class of algorithms called reinforcement learning is introduced to reduce the complexity greatly through searching with adjustments based on previous trials [22], [23], [24], [25]. Though the call admission control schemes proposed by these works can achieve the maximum average revenue or minimum average cost, these CAC design methods either cost too much time to collect enough the information of the system or take too long time to converge to optimal CAC schemes.

### 1.1.2 Dynamic Sub-carrier Allocation in Multi-user OFDM Wireless Network

Dynamic sub-carrier allocation is employed as the resource allocation component in our system. The DSA problem is usually formulated as an optimization problem like that in [26], [27], [28], [29], [30], [31]. Since the sub-carrier allocation by DSA affects the way to design our CAC strategy, it is worth building up the DSA module first. The DSA scheme takes the advantage of multi-user diversity in fading channel to improve the system performance.

*A.* Multi-user frequency-selective fading channel

Excluding pilot sub-carriers for various estimation and synchronization and null sub-carriers for guard bands, there are $N$ sub-carriers for data transmission with $M$ users served in the OFDM network. Taking multi-path fading effect into account, we get the impulse response over one sub-carrier for user $i$

$$h_i(t,\tau)=\sum_l \gamma_{il}(t)\delta(t-\tau_{il}) \tag{1}$$

where $\tau_{il}$ is the delay of $l$th path for user $i$ and $\gamma_{il}(t)$ is the corresponding complex amplitude. Then the frequency response can be expressed as

$$H_i(f,t)=\int_{-\infty}^{+\infty} h_i(t,\tau)e^{-j2\pi f\tau}d\tau=\sum_l \gamma_{il}(t)e^{-j2\pi f\tau_{il}} \tag{2}$$

Each receiver gets different distorted information from the same source in Fig. 1,

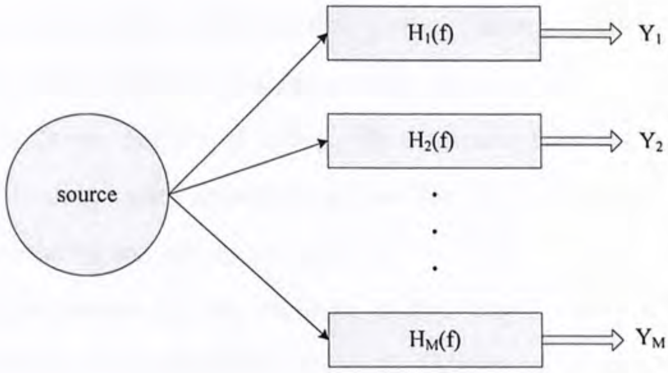


Fig. 1 Downlink Channel Model

where the time factor $t$ is omitted in frequency response $H_i(f)$ for only instantaneous channel conditions are considered. It keeps as a constant during one OFDM frame

$$SNR_i(f) = |H_i(f)|^2 / N_i(f) \tag{3}$$

Since the sub-carriers are orthogonal to each other perfectly, there is no interference between different transmission sessions. The signal to noise ratio (SNR) over each sub-carrier is enough to characterize the feature of the wireless fading channel.

*B.* Rate adaptation

In both research and commercial systems, rate adaptation is widely adopted to improve the spectral efficiency of wireless network. Whether we can get accurate Channel State Information (CSI) has great impact on both rate adaptation and the sub-carrier allocation scheme. In the time division duplex (TDD) system, the CSI can be estimated by the transmitter through the feedback from the receiver because of the reciprocal characteristics of channel. In addition, the base station can obtain channel state information directly from receiver over another independent channel in the frequency division duplex (FDD) system. Above all, we assume instantaneous channel state information can be gotten without any error. The base station changes modulation and coding schemes from time to time according to the channel condition, e.g. QPSK, 16-QAM and 64-QAM modulations are alternatively used in 802.16 standard for both fixed and mobile WiMAX implementation. Finite-state Markov channel model is proposed [4], [5], [6] for guiding how to implement the adaptive modulation and coding (AMC) in the wireless fading channel. Especially in a multi-carrier wireless network, each sub-carrier may travel through different paths and then suffer different degrees of fading. We can make fully use of scarce wireless resource by loading varied amount of information bits on different sub-carriers with adaptive modulation and coding schemes.

The rate adaptation can be realized in two ways: fast adaptation and slow adaptation. The fast rate adaptation arouses the large overhead problem for it changes the modulation and coding scheme once the channel condition changes, while the slow one performs too poor in a fast fading channel. In the thesis, we assume channel only changes after one frame transmission finishes. Instead of concerning about what kind of AMC scheme is used, we care about the achievable data rate over one sub-carrier through continuous adaptation theoretically. The achievable data rate is determined by the signal to noise ratio and the predefined bit error ratio (BER) [32], [33], [34]:

$$c = W \log_2(1 + \beta SNR) \tag{4}$$

where $W$ is the bandwidth and $\beta$ is the SNR gap in this form

$$\beta = -\frac{1.5}{\ln(5BER)} \tag{5}$$

Since all the sub-carriers have the same bandwidth, $W$ can be omitted without changing the nature of the rate adaptation. The channel state information is assumed to be sent by a separate error free control channel, hence we can always obtain the instantaneous SNR of user $i$ over $k$th sub-carrier, which is denoted as $\gamma_{ik}$. If without the SNR gap, we can get the data rate for user $i$ on sub-carrier $k$ as

$$c_{ik} = \log_2(1 + \gamma_{ik}) \tag{6}$$

where the continuous rate adaptation is perfectly implemented.

*C.* Dynamic sub-carrier allocation problem

If there are $N$ users served by $M$ sub-carriers in an OFDM wireless network as in Fig. 2:

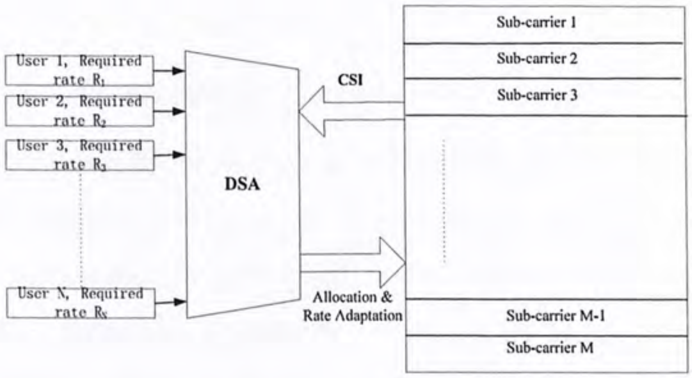


Fig. 2 Transmission from Base Station to Mobile Node

If each user is associated with a system utility $U_i(A)$, the dynamic sub-carrier allocation problem in the wireless network is formulated as a utility-based optimization problem mathematically. For instance, the utility of the network may be proportional to service rate received by users, e.g. the total throughput of the network $\sum_i \sum_k c_{ik}\rho_{ik}$. The objective function can be a concave function $\sum_{i=1}^{N} U_i(A)$ with respect to the sub-carrier allocation matrix $A$

$$A = [\rho_{ik}]_{N\times M} \tag{7}$$

In the matrix, each element $\rho_{ik}$ should be either 1 or 0, which means the $k$th sub-carrier is assigned to user $i$ or not. One sub-carrier can only be assigned to one user exclusively within a frame. In other words, for sub-carrier $k$, if $\rho_{ik} = 1$, $\rho_{jk} = 0$ for all $j \neq i$.

Therefore the DSA problem can be formulated as follows:

$$\max_{\rho} \sum_{i=1}^{N} U_i(A(\rho_{ik})) \tag{8}$$

$$s.t. \quad \sum_{k=1}^{M} c_{ik}\rho_{ik} \geq R_i \qquad \forall i \in \{1,2,3,\ldots,N\} \tag{9}$$

$$\sum_{i=1}^{N} \rho_{ik} = 1 \qquad \forall k \in \{1,2,3,...,M\} \tag{10}$$

$$\rho_{ik} \in \{0,1\} \tag{11}$$

where $R_i$ is the constant data rate demand of the $i$th user. It is the actual rate in unit of *b/sec/Hz* and denotes the user's QoS more precisely than other indicators. The optimization problem (8)-(11) is an integer programming problem that is too complicated to be solved. Moreover, parameters in problem (8)-(11) change with time according to equation (6), because the channel varies with time. We assume the channel is ergodic, and each user can go through all the possible channel states. In order to reduce the complexity and implement it online, $\rho_{ik}$ is relaxed as a real number varying between 0 and 1, which denotes the fraction that one user will occupy the sub-carrier in the transmission frame. The sub-carriers of the OFDM symbols are assigned to different users proportionally by $\rho_{ik}$. Thus the integer programming problem is replaced by a lower-costing convex optimization problem:

$$\max_{\rho} \sum_{i=1}^{N} U_i(A(\rho_{ik})) \tag{12}$$

$$s.t. \quad \sum_{k=1}^{M} c_{ik}\rho_{ik} \geq R_i \qquad \forall i \in \{1,2,3,...,N\} \tag{13}$$

$$\sum_{i=1}^{N} \rho_{ik} = 1 \qquad \forall k \in \{1,2,3,...,M\} \tag{14}$$

$$0 \leq \rho_{ik} \leq 1 \tag{15}$$

Every user's QoS can be satisfied by the sub-carrier allocation scheme under some channel condition. However, the scheme may not work when there are lots of users in the network, and in the meantime most of the sub-carriers drop into deep fading. It means that the optimiaztion problem cannot be solved without violating any constraints. In fact, if users can get into the network without any control, the violation will occur more and more frequently with the increase of users. So it is necessary to design a CAC strategy adapting to the time-changing environment to guarantee the feasibility of the DSA problem in multiuser wireless networks.

# 1.2 Problem Statement

When various adaptive resource allocation strategies are adopted in OFDM wireless networks, the distribution of service rate received by each user from physical layer no longer simply conforms to the fading channel SNR distribution, i.e. exponential distribution in Rayleigh fading channel. Traditional CAC modules that are decoupled completely from resource allocation cannot adapt to the environment well. The CAC strategy can be improved if the information from resource allocation module is incorporated. Though the structure of threshold-based call admission control scheme is simple, to find a proper threshold requires searching all the possible thresholds for the system. It is hard to obtain the optimal threshold efficiently in a high dynamic wireless network. We use the achievable data rate for each user as the QoS indicator directly, and design the CAC strategy to find out a trade-off between the complexity and performance.

With respect to maximizing the average system revenue, the optimal CAC schemes can be derived from the algorithms for SMDP problem. However, a little change of the system will cause serious modification of the call admission control scheme, for the system model has to be reconstructed. The computing burden of reconstructing the system model overwhelms the benefit brought by the derived CAC schemes. Sometimes it is impossible to obtain all the information of a wireless network before operating of the system, so the CAC scheme should have the ability of adapting to the system without any knowledge of the network beforehand. If the performance of the CAC scheme is very close to the optimal scheme like the standard methods in SMDP, it can be claimed a good scheme.

# 1.3 The Organization of The Thesis

The rest of the thesis is organized as follows. Chapter 2 describes the scope of the multi-user OFDM wireless network and the function of the CAC module in the system in detail. A framework of our CAC strategy that adapts to the multiple user OFDM wireless networks is illustrated.

After decoupling the CAC strategy into two-stage realization, we demonstrate stage one -- how to guarantee each user's QoS requirement with limited resource in Chapter 3. Considering the benefit aroused by adaptive resource allocation and time-varying channel, we develop an algorithm to measure the throughput achieved by each user in the worst condition. Based on the algorithm, the QoS-provisioning CAC scheme is proposed to determine the admissibility of users.

In Chapter 4, the second stage of the CAC strategy is proposed for maximizing long term revenue brought by different kinds of users. In a real network nobody can predict what kind of application will appear, so traditional algorithms of SMDP are hard to be applied to the CAC scheme design directly. With the idea of making use of history record and prediction of the network, we derive a low complexity average revenue maximization CAC scheme to do admission without setting up the system model beforehand.

Finally, we come to the conclusion and discuss the possible future work in Chapter 5.

# Chapter2

# System Model and Call Admission Control Framework

Our CAC strategy is expected to find an application in next generation broadband wireless communication system like WiMax, where there is a base station coordinating all the communication and controlling the admission of users. As mentioned in Chapter 1, we propose a CAC strategy that can guarantee the QoS in system-wide as well as maximize average system revenue. We show how the two stages constitute an integrated CAC component for multiple user OFDM wireless networks.

## 2.1 System Setup

In OFDM wireless networks, the number of sub-carriers is very large. For example, 192 sub-carriers are used for data transmission in the fixed WiMax system. Due to the correlation between the channel conditions of adjacent sub-carriers, the sub-carriers are usually grouped into some sub-bands for allocation. For simplicity, we assume that there are 32 independent sub-carriers for data transmission in the system, and that power is uniformly distributed over all the sub-carriers. For a specific user, it has the same mean signal to noise ratio (SNR) across all the sub-carriers, but the instantaneous SNR on different sub-carriers may be different. The DSA scheme of problem (12)-(15) is applied in this OFDM wireless network, which affects the design of our CAC strategy.

The new-coming user is featured by $\{\gamma, R, r, \lambda, \mu\}$, where $\gamma$ is the mean SNR, $R$ is a constant data rate demand denoting the minimum QoS requirement of this user,

$r$ is the rate at which revenue is generated by serving the user, and $\lambda$ and $\mu$ represent the mean arrival rate and average service rate respectively. It is assumed that the inter-arrival time and service time of each user follow the exponential distribution. Likewise, we assume that the duration of a transmission session is much longer than the channel coherence time, so that a user will go through all the possible channel states once it is admitted into the network.

## 2.2 The CAC Strategy Framework

Unlike traditional schemes, our CAC module exchanges information with dynamic sub-carrier allocation module so as to incorporate the physical channel information. Though best performance can be achieved by integrating the two modules together completely, the complexity will be forbidden.

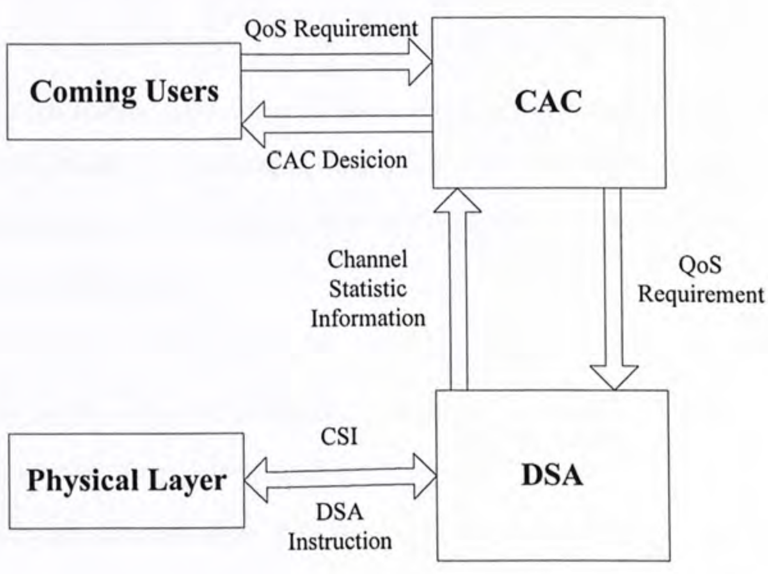


Fig. 3 Call Admission Control (CAC) Framework with Dynamic Sub-carrier Allocation (DSA)

The system diagram is illustrated in Fig. 3 above: The DSA component allocates bandwidth resource to admitted users according to problem (12)-(15) formulated with channel state information (CSI) and QoS requirements of users. According to the channel statistic information provided by the DSA module, the CAC module determines the admissibility of a new-coming user.

There are two dynamics in a multi-user OFDM wireless network that need to be considered in the CAC design: the dynamic of wireless channel condition and the dynamic arrival and departure of users. Our CAC strategy is composed of two CAC stages to handle the dynamics for a balance between performance and complexity.

With respect to the dynamic of wireless channel, the first stage mainly deals with the admissibility of new-coming users. Whether a new user can be admitted to the system depends on the wireless channel statistics and QoS requirements of itself as well as existing users. This strategy is carried out by checking the feasibility of the DSA problem in (12)-(15) with taking the new user into account.

Among the users that are considered to be admissible by the first stage, the second stage selectively admit those that will lead to maximal average system revenue. We take two-type user case as an example to illustrate how different kinds of users give rise to different revenue contribution to the system:

1) Case 1: $r_1 >> r_2$ and $\lambda_1/\mu_1 < \lambda_2/\mu_2$ . Though the traffic load of type-1 user is lower than that of type-2 user, it has a much higher revenue rate. There is a possibility that type-1 user can contribute more to the network due to the much higher revenue rate, thus the system prefer to accept type 1 user whenever it appears in the network.
2) Case 2: $r_1 > r_2$ and $\lambda_1/\mu_1 \ll \lambda_2/\mu_2$ . In this condition, type-2 user compensate for the low revenue rate with huge quantity, it is a waste of resource to reject type-2 user for the possible coming of type-1 user.

So the second stage accepts users with different priorities by taking $\{ r, \lambda, \mu \}$ into consideration under the work of first stage.

Any new-coming user needs to go through the two stages of our CAC strategy before it is admitted. The two-stage CAC strategy is illustrated as Fig. 4 below:

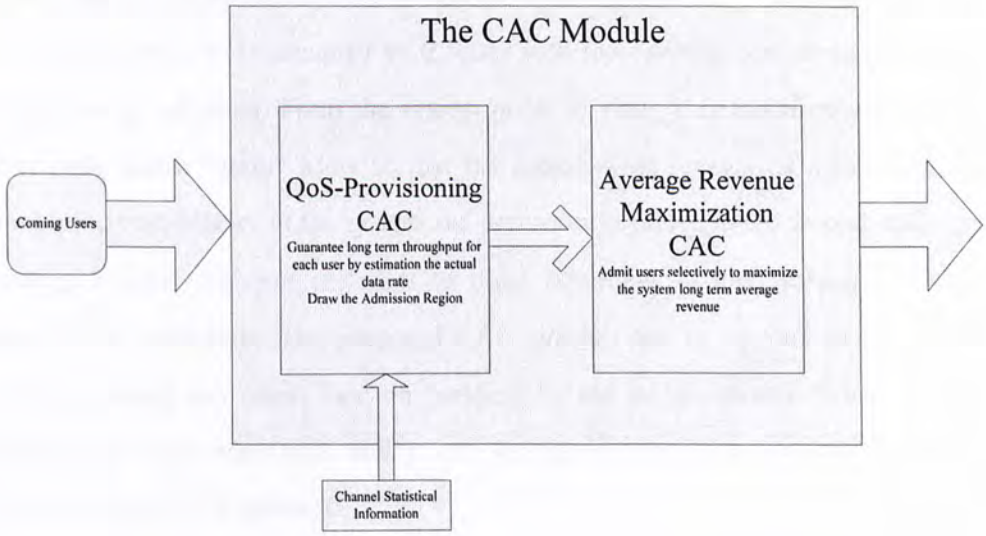


Fig. 4 The Structure of CAC Strategy

*A.* The QoS-Provisioning CAC

The target of this stage is to prevent the QoS of existing users from degrading to an unacceptable level on the admission of a new-coming user. Specifically, it is to make sure that the DSA problem in (12)-(15) is feasible with high probability. Due to the time-varying nature of the wireless channel, it is impossible to predict the exact channel conditions during the whole transmission session before the user is admitted. In this work, we propose a novel methodology to estimate the average throughput achievable by each user under the framework of adaptive OFDM system. Based on the estimation, we can decide to admit/ reject a user if the outage (infeasible) probability of the system is under/beyond a tolerable small value. The performance is validated through simulations later. As a matter of fact, this scheme can work independently to ensure the system operates smoothly, which is the prime motivation of CAC design.

*B.* The Average revenue maximization CAC

Once the admission region is determined, various adaptive resource allocation schemes with different system utility functions can be applied. However, feasibility alone does not imply high system revenue in (12)-(15). In other words, it is possible that system resource is occupied by the user with low revenue contribution rate, if a "bad" user is admitted. From the system point of view, it is therefore desirable to selectively admit "good" users so that the system-wide revenue is maximized. By combining both history of the system and prediction together, in the second stage, our strategy assigns different priorities to users according to their different revenue contribution potentials. The proposed CAC strategy can be applied to the system without costing too much time on building up the system model. What is more, besides the quick adaptation ability, our strategy performs close to the theoretical optimal results in simulations.

# Chapter 3

# QoS-aware Adaptive Call Admission Control—Stage One: The QoS-Provisioning CAC

In this stage, we tailor a QoS-Provisioning CAC scheme for OFDM wireless networks with dynamic sub-carrier allocation. Instead of using bandwidth or SNR to denote the QoS for each user, we present QoS in terms of outage probability of the system by estimating the actual data rates received by the users in the long-run. Specifically, our contribution is three-folded:

1) We formulate a linear programming problem concerning about system outage probability in multiuser OFDM wireless networks based on the DSA problem (12)-(15).

2) An effective estimation algorithm is proposed to calculate the mean data rate for each user accurately. This algorithm follows the linear programming problem formulated for CAC scheme design to predict the influence on the whole network caused by the entry of a new-coming user.

3) Based on the estimation algorithm developed above, the QoS-provisioning call admission control scheme can be applied in two ways. If the system is small, we can draw the small outage-based admission region by the CAC scheme. For a large system, the call admission control scheme can control users on-line through calculation of this algorithm, which exploits the advantage brought by DSA under multiuser OFDM wireless network condition.

# 3.1 Problem Formulation

In each sub-carrier allocation procedure, an outage event occurs once any constraint in optimization problem (12)-(15) is violated. The frequency of outages recorded in a long enough system time is taken as the outage probability of the network which heavily depends on users' conditions and the resource allocation algorithm adopted. Mathematically our CAC scheme is to check the feasibility of optimization problem (12)-(15), if admitting a new user.

Let $\varepsilon$ be the tolerable outage probability that should be a reasonable small value for guaranteeing the system-wide QoS. If constraint (14) and (15) are always satisfied, the outage probability can be expressed mathematically:

$$P(\bigcap_{i=1}^{N}(\sum_{k=1}^{M} c_{ik}\rho_{ik} \geq R_i)) \geq 1-\varepsilon \;\; \forall i \in \{1,2,3,...,N\} \tag{16}$$

It can be asserted that an outage event occurs inevitably if the worst user's data requirement cannot be met after all the available resource and allocation patterns are tried. With respect to the constraints in problem (12)-(15), we formulate an optimization problem that balances the distribution of resource so that at least the worst user can perform as well as other users with available resource. If the worst user's data rate requirement can be satisfied, all the users' QoS in the network are guaranteed. The problem can be formulated as follows:

$$\max \min_{i \in \{1,2,3,...,N\}} \left( \sum_{k=1}^{M} c_{ik}\rho_{ik} \right) \Big/ R_i \tag{17}$$

$$s.t. \quad \sum_{i=1}^{N} \rho_{ik} = 1 \qquad \forall k \in \{1,2,3,...,M\} \tag{18}$$

$$0 \leq \rho_{ik} \leq 1 \tag{19}$$

In fact, such a max-min optimization problem can be reformed as a standard linear programming problem without any dual gap:

$$\max_{i} C \tag{20}$$

$$s.t. \quad C \leq \frac{\sum_{k=1}^{M} \rho_{ik} c_{ik}}{R_i} \qquad \forall i \in \{1,2,3,...,N\} \tag{21}$$

$$\sum_{i=1}^{N} \rho_{ik} = 1 \qquad \forall k \in \{0,1,2,...,M\} \tag{22}$$

$$0 \leq \rho_{ik} \leq 1 \tag{23}$$

When the channel state changes with time, the achievable data rate over different sub-carriers will change correspondingly by rate adaptation, and then the sub-carrier allocation has to be altered accordingly. If the value of $\max_i C$ is much smaller than 1, it implies constraint (13) is violated, for at least the acquired data rate of one user cannot meet its demand. Problem (20)-(23) is logically equivalent to problem (12)-(15) in terms of outage probability of the network, which is illustrated in a simple simulation as shown in Fig. 5.

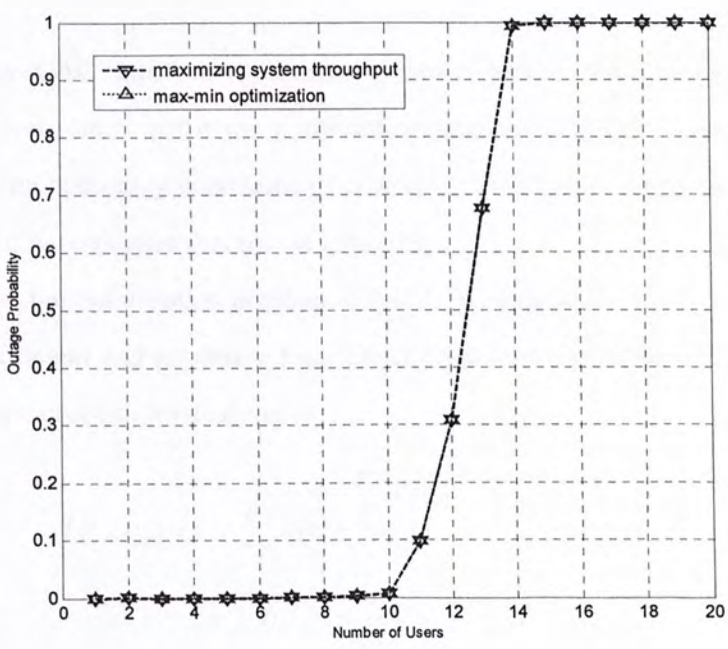


Fig. 5 Comparison of Dynamic Sub-carrier Allocation and Linear Programming Problem for CAC in terms of Outage Probability

In the simulation, 32 sub-carriers are assigned to users according to their instantaneous channel state information to maximize system throughput by DSA. The outage probability of the system is recorded by counting the outage events, if the system runs a long enough time. Compared with the outage probability in simulation, the probability that the objective of the linear programming problem (20)-(23) drops below 1 coincides to the same curve. Our formulation for the outage problem is validated through this simulation, so we can study the outage problem by checking the probability that the objective value of the linear programming problem (20)-(23) drops below 1. Another interesting observation is that the outage probability jumps from a very low level to nearly 1 with the increase of user number. The slope of the curve is so steep that the call admission control scheme may be simplified by this feature in the multi-carrier network.

# 3.2 Optimality Condition Analysis

Since the CAC must do admission decision before the resource allocation procedure, we cannot solve the optimization problem (20)-(23) directly to make decisions. The optimality conditions of problem (20)-(23) give some hints about the estimation of the objective function in long-run.

Turning to the optimization problem (20)-(23), we exploit the KKT condition first, which is sufficient and necessary for a linear programming problem [35]. We write down the dual problem for analysis:

$$\begin{aligned} L(C,\rho_{ik}) = C - \sum_{i=1}^{N} \lambda_i \Big(C - \frac{\sum_{k=1}^{M} \rho_{ik} c_{ik}}{R_i}\Big) \\ - \sum_{k=1}^{M} \eta_k \Big(\sum_{i=1}^{N} \rho_{ik} - 1\Big) + \sum_{i=1}^{N}\sum_{k=1}^{M} \mu_{ik}\rho_{ik} \end{aligned} \tag{24}$$

$$\lambda_i \Big(C - \frac{\sum_{k=1}^{M} \rho_{ik} c_{ik}}{R_i}\Big) = 0 \qquad \lambda_i \ge 0 \tag{25}$$

$$\eta_k(\sum_{i=1}^{N}\rho_{ik}-1)=0 \qquad \eta_k \geq 0 \tag{26}$$

$$\mu_{ik}\rho_{ik}=0 \qquad \mu_{ik}\geq 0 \tag{27}$$

By differentiating $L(C,\rho_{ik})$ with respect to $C$ and $\rho_{ik}$, we get the KKT condition:

$$\frac{\partial L}{\partial C}=1-\sum_{i=1}^{N}\lambda_i=0 \qquad \lambda_i \geq 0 \tag{28}$$

$$\frac{\partial L}{\partial \rho_{ik}}=\begin{cases}\frac{\lambda_i c_{ik}}{R_i}-\eta_k+\mu_{ik}=0 & \rho_{ik}=0\\ \frac{\lambda_i c_{ik}}{R_i}-\eta_k=0 & \rho_{ik}>0\end{cases} \qquad \eta_k\geq 0,\mu_{ik}\geq 0 \tag{29}$$

Specifically, we can lead to the conclusion from condition (29) that for sub-carrier *k*, it will be allocated to the *i*th user if $\frac{\lambda_i c_{ik}}{R_i}=\eta_k$; and other users with $\frac{\lambda_j c_{jk}}{R_j}=\eta_k-\mu_{jk}$, $j\neq i$ are excluded from this sub-carrier. As a result, sub-carrier *k* will be allocated to the user satisfying $\frac{\lambda_i c_{ik}}{R_i}>\frac{\lambda_j c_{jk}}{R_j}$ for optimality.

# 3.3 Throughput Estimation Algorithm

With the tolerance of small violation in the real network, the mean throughput over a long period may be a quite good indicator for call admission control scheme. If the average service rate of any user is far below its requirement, the system performance cannot be guaranteed upon the admission of the new-coming user. Since the achievable data rate $c_{ik}$ varies with time by rate adaptation in the wireless channel, it is prohibitive to repeat solving problem (20)-(23) to get the mean value. Alternatively, we try to estimate the mean service rate received by each user through calculating the fraction that each user will occupy under the optimization problem (20)-(23), if the newly arrived user is admitted into the system.

According to above analysis of the optimality conditions, if one user occupies the sub-carrier, $\frac{\lambda_i c_{ik}}{R_i} > \frac{\lambda_j c_{jk}}{R_j}$ must hold for all $j \neq i$. In a time-varying channel, the Lagrangian multiplier $\lambda$ is an unknown dual variable, which varies with channel state in each allocation cycle. In order to do the long term estimation without repeating allocation procedure many times, another variable $w$ is introduced to replace $\lambda$ as a weight for each user, which can help measure the mean data rates of users. Let $P_{ij}(\frac{w_i c_{ik}}{R_i} > \frac{w_j c_{jk}}{R_j})$ denote the probability that the $i$th user has a priority to the $j$th user on sub-carrier $k$. The probability that user $i$ can occupy sub-carrier $k$ is $\prod_{\substack{j=1 \\ j\neq i}}^{N} P_{ij}(\frac{w_i c_{ik}}{R_i} > \frac{w_j c_{jk}}{R_j})$ , for all the users are mutually independent. We can derive such a probability combing the SNR distribution and the optimality conditions with the alternative weight $w$.

With an error free control channel, the instantaneous SNR matrix of the whole network under Rayleigh fading environment is given by

$$SNR = [\gamma_{ik}]_{M\times N} \tag{30}$$

where each row denotes all the SNR values over different sub-carriers for one user, and they are independently identically distributed. Each element of this matrix is an exponential distributed random variable with different mean SNR depending on user in the Rayleigh fading channel. $\gamma_{ik}$ of the SNR matrix conforms to a probability density function (PDF)

$$f(\gamma_{ik}) = \frac{1}{\bar{\gamma}_i}\exp(\frac{-\gamma_{ik}}{\bar{\gamma}_i}) \tag{31}$$

By employing rate adaptation, we have the corresponding data rate matrix as follows

$$Rate = [c_{ik}]_{M\times N} \tag{32}$$

Based on (6) and (31), the PDF of data rate on sub-carrier $k$ for user $i$ is

$$f(c_{ik}) = \frac{2^{c_{ik}} \ln 2}{\bar{\gamma}_i} \exp\left(\frac{-(2^{c_{ik}} - 1)}{\bar{\gamma}_i}\right) \tag{33}$$

Then it is easy to derive that

$$\begin{aligned} P_{ij}\left(\frac{w_i c_{ik}}{R_i} > \frac{w_j c_{jk}}{R_j}\right) &= P_{ij}\left(c_{jk} < \frac{w_i c_{ik} R_j}{w_j R_i}\right) \\ &= \int_0^{\frac{w_i c_{ik} R_j}{w_j R_i}} f(c_{jk}) dc_{jk} \\ &= 1 - \exp\left( -(2^{\frac{w_i c_{ik} R_j}{w_j R_i}} - 1) \Big/ \bar{\gamma}_j \right) \end{aligned} \tag{34}$$

so the average data rate user $i$ can achieve over sub-carrier $k$ is given by an integral:

$$\begin{aligned} T_{ik} &= \int_0^{+\infty} c_{ik} f(c_{ik}) \prod_{\substack{j=1 \\ j \neq i}}^{N} P_{ij}\left(\frac{w_i c_{ik}}{R_i} > \frac{w_j c_{jk}}{R_j}\right) dc_{ik} \\ &= \int_0^{+\infty} c_{ik} \frac{2^{c_{ik}} \ln 2}{\bar{\gamma}_i} \exp[-(2^{c_{ik}} - 1)/\bar{\gamma}_i] P(i) dc_{ik} \end{aligned} \tag{35}$$

$$P(i) = \prod_{j=1, j \neq i}^{N} [1 - \exp(-(2^{\frac{w_i c_{ik} R_j}{w_j R_i}} - 1)/\bar{\gamma}_j)] \tag{36}$$

One significant characteristic of the optimization problem (20)-(23) is to balance the traffic among all the users proportional to their demands. Upon this nature, we propose an iterative algorithm in Fig. 6 to find out the weight for each user as well as estimate the mean throughput for each user heuristically.

Initialize $\mathrm{w}_i \leftarrow 1/N$ $N$-*user number* for all $i$

$\mathbf{T} \leftarrow \mathbf{0}$ vector of the estimated data rate for all the users

$\mathbf{R}$ vector of required data rate for all users

$\mathbf{E} \leftarrow \mathbf{1}$ vector whose element is $\underset{i}{mean}(T_{ik}/R_i)$ for all $i$

mean *SNR* for each user over one sub-carrier

While $(\max_i |E_i - T_{ik}/R_i|) > \sigma$

for user $\mathrm{i} = 1:N$

$$\prod_{\substack{j=1\\ j\neq i}}^{N} P_{ij}\left(\frac{w_i c_{ik}}{R_i} > \frac{w_j c_{jk}}{R_j}\right)$$

$$T_{ik} \leftarrow \int_0^{+\infty} c_{ik} f(c_{ik}) \prod_{\substack{j=1\\ j\neq i}}^{N} P_{ij}\left(\frac{w_i c_{ik}}{R_i} > \frac{w_j c_{jk}}{R_j}\right) dc_{ik}$$

end;

$E_i \leftarrow \underset{i}{mean}(T_{ik}/R_i)$ for all $i$

update $\mathrm{w}_i \leftarrow w_i + step \times (E_i - T_{ik})$ for all $i$

End

Fig. 6 Throughput Estimation Algorithm

We can estimate mean data rate for all users on one sub-carrier, then the mean throughput for each user can be calculated $T_i = MT_{ik}$ since all the sub-carriers have the same statistical characteristics.

# 3.4 QoS-Provisioning CAC

Based on the throughput estimation algorithm developed above, we implement the QoS-provisioning call admission control in two ways for different conditions:

If there is a finite user set, it is easy to construct an admission region which shows that how many users can be supported potentially in a network. With the estimation algorithm and given user types, the boundary of admission region can be constructed as the maximum number of users, when the estimation value $\min_{i=1,2,\ldots N}(M \times T_{ik}/R_i)$ approaches 1. In the simulation, we show the mean data rate based admission region can control the outage probability of the network within a tolerable small value, for the variance is so small that it does not affect the system performance too much.

On the other hand, in a real network, we have little knowledge about what type of user will appear in the network before system operation. It is a very time-consuming task to draw an admission region before deciding whether to accept the new-coming user or not. When a new user comes to request for service, instead of allocating resource actually, we take it into the network virtually and calculate the value of the indicator through the throughput estimation algorithm. If $\min_{i=1,2,\ldots N}(M \times T_{ik} / R_i)$ is much smaller than 1, it means that at least one user cannot be satisfied and the outage probability of the system will be too high to afford. In this case, for the sake of the undergoing users, the system rejects the new coming user rather than degrades other users to allow it. On the contrary, if the indicator is equal to or greater than 1, it means that the system can perform well after admitting this user, and hence it is admitted into the network. Such an admission scheme is much like a complete sharing method, except that the admission region may change since the actual capacity of the network cannot be simply described by its bandwidth.

# 3.5 Performance Evaluation

In this section, we set up a central controlled OFDM network to illustrate the performance of the QoS-provisioning CAC scheme based on the throughput estimation algorithm. There are 32 sub-carriers transmitting through multiple paths in the network, the sub-carriers are statistically identical to one specific user. The user type is defined by both its mean SNR and data rate requirement, which are determined by the distance of the user from the base station and the application demand from upper layer separately.

In order to validate that our algorithm can predict the mean throughput for each user, we randomly choose four kinds of users in the network, which means four pairs of $\{\gamma, Rate\}$. Comparing the numerical result of our estimation algorithm with simulation result of problem (20)-(23) that changes with channel condition for a long enough time, we can see in Fig. 7 our predicted result is very close to the real condition. The small gap is caused by the data rate variance that is too hard to be predicted.

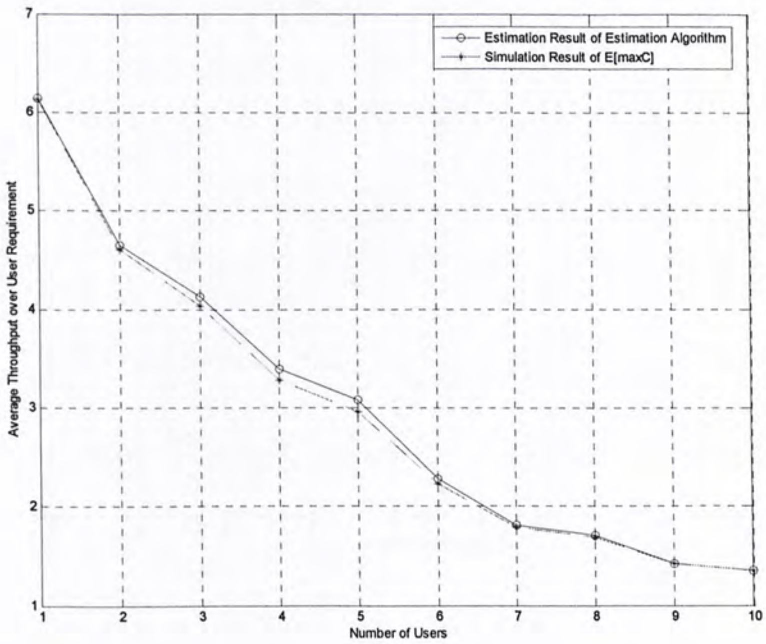


Fig. 7 Comparison of Numerical Result from Estimation Algorithm with Statistical Mean from Problem (20)-(23) in Simulating OFDM Wireless Network

It is worth noting that the step size in the throughput estimation algorithm should be carefully chosen for fast convergence. In our simulation shown in Fig. 8, a small step size is selected with a precision parameter $\sigma$ in scale $10^{-4}$, the speed is acceptable for online implementation. In each round of iteration in the throughput estimation algorithm, the weight for each user is adjusted. The iteration number decreases due to less and less resource difference among users in allocation, when more and more users are admitted into the network. This curve can be smoothed given long enough system operation time.

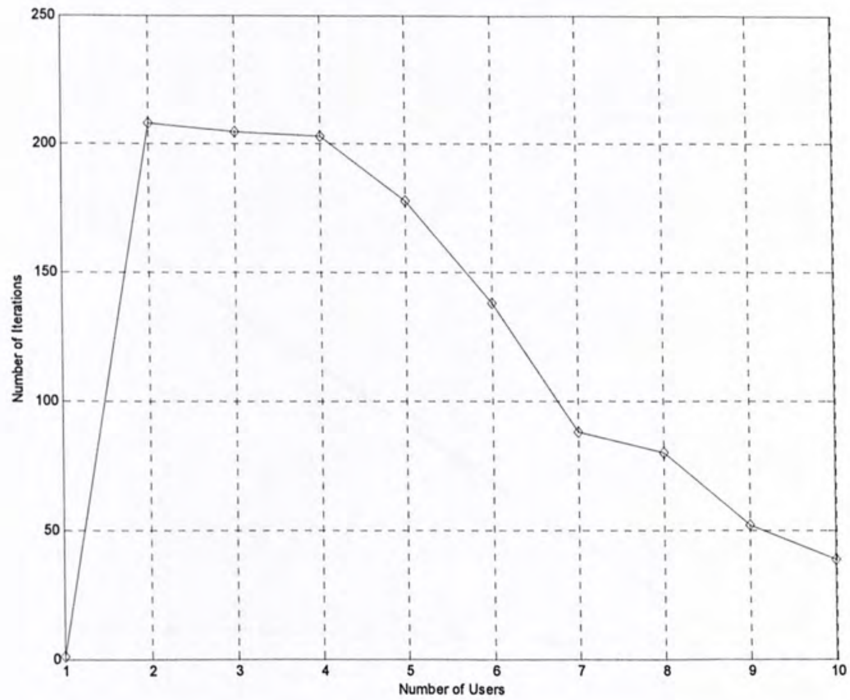


Fig. 8 Iteration Speed vs. Number of users

Furthermore, we specify two kinds of users in the network: type 1 with mean SNR 12 dB, constant data rate demand 18 *b/sec/Hz* and type 2 with mean SNR 6 dB, constant data rate demand 6 *b/sec/Hz* respectively. Under this situation, the admission region with small outage probability is shown in Fig. 9. The outage probability counted for the whole network is around 0.1, which is reasonable for a practical system. Though we lack the simulation result of non-adaptive condition, the admission region of non-adaptive one should be smaller than ours due to each user is assumed to consume more resource without making use of multi-user diversity by DSA.

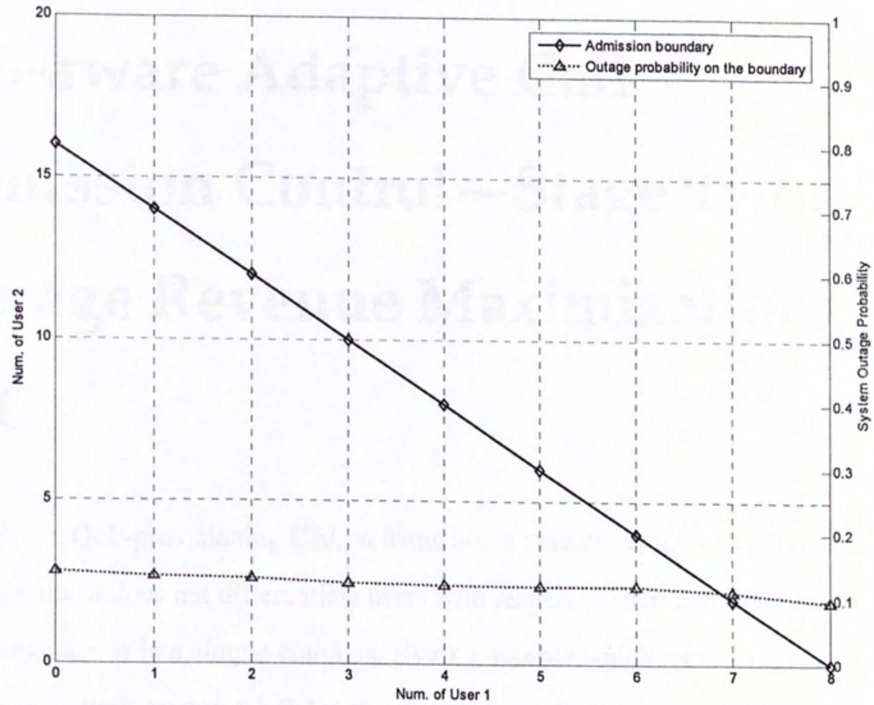


Fig. 9 Admission Region with Outage Probability in Two Types of Users Case

# Chapter 4

# QoS-aware Adaptive Call Admission Control—Stage Two: Average Revenue Maximization CAC

Though the QoS-provisioning CAC scheme alone guarantees the smooth operation of the system, it does not differentiate users with respect to their contributions to the system revenue. It is a simple complete sharing scheme which admits users as long as there is enough resource left for the new comers. This scheme can achieve the maximum average system revenue, when there is only one type of user or one kind of user dominates the traffic in the network. In order to enhance the performance of our CAC strategy, the second CAC stage is proposed under the framework for maximizing average system revenue. This scheme is model-free, implying that it can adapt to the network automatically, and the performance of this scheme is very close to optimal result in the simulations.

## 4.1 Semi-Markov Decision Process

From a service provider point of view, it collects revenue from the served users, and the charge rate may be different across different types of users. Each type of user is associated with its revenue rate, mean arrival rate and mean service rate as $\{r_i, \lambda_i, \mu_i\}$. For the ease of demonstration, we take the two-type user case as example in the following analysis and simulations, and the multiple-user case follows the same principle.

Given the capacity of network in terms of user number, the evolution of the network without any control over the users can be modeled as a time-continuous Markov chain. We use a vector $s=\{n_1,n_2\}$ to denote the state of the Markov chain, where $n_1$ is the number of type-1 user and $n_2$ represents the number of type-2 user.

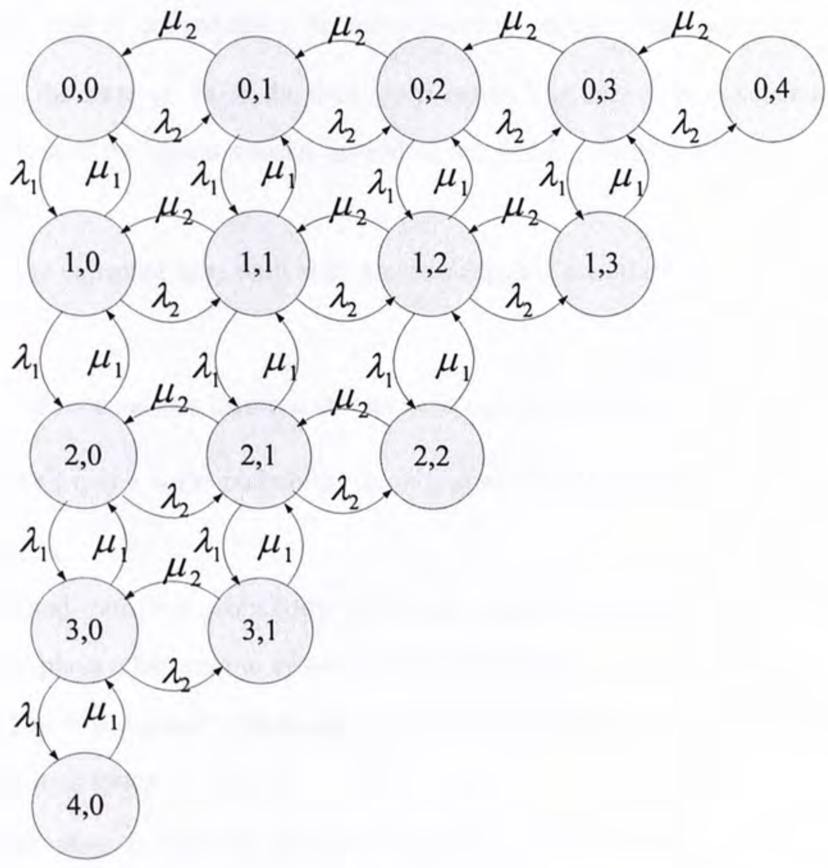


Fig. 10 The Time-continuous Markov Chain Model for Uncontrolled Network

As shown in Fig.10, the current state of the Markov chain just relies on previous state and transition rate. The transition only happens between two adjacent states, which means only one arrival or departure event occurs within a small time interval. The chain is confined to an area bounded by the capacity region of the network, which is drawn by the QoS-provisioning CAC scheme in Chapter 3.

If the decision is made when new events occur in the system, the transition in the

Markov chain is re-shaped according to our will. The new process is referred as a Semi-Markov Decision Process (SMDP), because the decision is made at any time instant contrasted with the Markov decision process where the decision is made at discrete time slots. The SMDP characterizes a dynamic system that is observed at random time points from the beginning and classified into a number of possible states. After observing the state of the system and the coming event, the controller makes a decision to lead a consequence of some revenue change. The action set $A(s)$ depends on the state at which the decision is made. The next state is dependent on current state and the action chosen instead of the whole history. Following is some key factors:

$\tau_s(a)$ ---- the expected time until next decision epoch if action $a$ is chosen at present state $s$.

$r_s(a)$----- the revenue rate incurred at state $s$ with action $a$ is chosen.

$p_{ss'}(a) = p(s' \mid s, a)$----- the probability that the system will transit from state $s$ to $s'$ given action $a$.

The original transition probability in Markov chain is modified by the decision of actions that plays a key role in system control. A SMDP model is built up to analyze the behaviors of a communication network, which is featured by following elements:

1) The state space

   Each state is denoted as a vector $s = \{n_1, n_2\}$ representing the number of different users served in the network. $S$ is used to denote the state space for the possible evolution of the network, and it is bounded by the capacity of the network. Regardless of the types of users admitted, we assume the system can support $N$ users simultaneously at most. It implies $n_1 + n_2 \le N$ must hold for the system.

2) The action set

   When a new user arrives, we should make a decision to select an action from current action set at this time instance. For most states, the action set is

$A(s)=\{(0,0),(1,0),(0,1)\}$. $a=(0,0)$ is to reject any arrived users, in the mean time $a=(1,0)$ means accepting type-1 user and $a=(0,1)$ means accepting type-2 user. Within a very small time interval, only one event occurs, so that action $(1,1)$ is discarded. However, in some extreme cases, the action set is different from the normal states, e.g. in state $s=\{0,N\}$, the action set is $A(s_0)=\{(0,0)\}$ for no more users can be admitted on the boundary states.

3) The expected sojourn time

When the system is in state $s$ with action $a$, the expected sojourn time until entering a new state is given by:

$$\tau_s(a)=[\sum_{i=1}^{2} n_i\mu_i+\sum_{i=1}^{2}\lambda_i a_i]^{-1} \tag{37}$$

We assume that user arrival is a Poisson stochastic process and the service time conforms to exponential distribution with different mean time. $n_i\mu_i$ is the average rate at which calls of type-$i$ user terminate and $\lambda_i a_i$ represents the mean rate at which new calls come into the network with permission.

4) The transition probability matrix

The major difference between the SMDP and the time-continuous Markov chain is the transition probability. In our system, all the transition conditions are given by:

$$p((n_1-1,n_2)\,|\,(n_1,n_2),(0,0))=n_1\mu_1/\tau_{(n_1,n_2)}(a=(0,0)) \quad n_1\neq 0 \tag{38}$$

$$p((n_1-1,n_2)\,|\,(n_1,n_2),(1,0))=n_1\mu_1/\tau_{(n_1,n_2)}(a=(1,0)) \quad n_1\neq 0, n_1+n_2<N \tag{39}$$

$$p((n_1-1,n_2)\,|\,(n_1,n_2),(0,1))=n_1\mu_1/\tau_{(n_1,n_2)}(a=(0,1)) \quad n_1\neq 0, n_1+n_2<N \tag{40}$$

$$p((n_1,n_2-1)\,|\,(n_1,n_2),(0,0))=n_2\mu_2/\tau_{(n_1,n_2)}(a=(0,0)) \quad n_2\neq 0 \tag{41}$$

$$p((n_1,n_2-1)\,|\,(n_1,n_2),(1,0))=n_2\mu_2/\tau_{(n_1,n_2)}(a=(1,0)) \quad n_2\neq 0, n_1+n_2<N \tag{42}$$

$$p((n_1,n_2-1)\,|\,(n_1,n_2),(0,1)) = n_2\mu_2 / \tau_{(n_1,n_2)}(a=(0,1)) \qquad n_2 \neq 0, n_1+n_2 < N \tag{43}$$

$$p((n_1+1,n_2)\,|\,(n_1,n_2),(1,0)) = \lambda_1 / \tau_{(n_1,n_2)}(a=(1,0)) \qquad n_1+n_2 < N \tag{44}$$

$$p((n_1,n_2+1)\,|\,(n_1,n_2),(0,1)) = \lambda_2 / \tau_{(n_1,n_2)}(a=(0,1)) \qquad n_1+n_2 < N \tag{45}$$

$$p(s'\,|\,s,a) = 0 \qquad \textit{else} \tag{46}$$

5) The revenue rate

The revenue rate of state $s$ with action $a$ chosen is given by

$$r_s(a) = \sum_{i=1}^{2} r_i n_i \tag{47}$$

where $r_i$ is the revenue rate brought by type-$i$ user. Further, the expected lump sum revenue received by the system is

$$R_s(a) = r_s(a)\tau_s(a) \tag{48}$$

# 4.2 Investigation of Algorithms for SMDP

There has been a large amount of interest in finding optimal decision policy to achieve maximum average revenue of a SMDP system. Generally, they can be categorized into model-based algorithms and model-free algorithms.

The three standard algorithms for SMDP: policy iteration, value iteration and linear programming formulation are model-based algorithm. In most cases, these algorithms achieve maximum average revenue, provided all the parameters of a SMDP model. Let us take the linear programming formulation as an example to show how to derive the optimal policy. This method will be used in simulations later as the benchmark.

In order to find the optimal policy for maximizing average revenue, the SMDP problem can be formulated in this form:

$$\max \sum_{s\in S} \sum_{a\in A(s)} r_s(a)\tau_s(a)z_s(a) \tag{49}$$

$$s.t. \sum_{s\in S}\sum_{a\in A(s)} \tau_s(a)z_s(a) = 1 \quad (50)$$

$$\sum_{a\in A(s')} z_{s'}(a) = \sum_{s\in S}\sum_{a\in A(s)} p(s' \mid s,a)z_s(a) \quad (51)$$

$$z_s(a) \geq 0 \qquad s\in S, a\in A(s) \quad (52)$$

In fact, the linear programming formulation is introduced for the sake of easy extra constraint adding. The objective (49) is the average revenue of the system in a long term, $z_s(a)$ is the decision variable that we try to search for optimization. $\tau_s(a)z_s(a)$ is the long-run fraction of decision epochs, at which the system is in state $s$ and action $a$ is chosen. Equation (50) is the normalized condition of system operation time, and equation (51) shows the balance between different decision epochs from the long-run point of view. It can be applied as a probabilistic call admission control scheme, at each state $s$, the action $a$ is selected with probability $p_s(a) = z_s(a) \Big/ \sum_{a\in A(s)} z_s(a)$.

These algorithms for SMDP problem can be applied to call admission control design in the small scale network, but in a real network system which is usually large some new challenges emerge:

1) The three standard methods all need complete information of the system such as transition probability matrix. Even a little shift of any parameters, e.g. one type user's revenue rate changes from $r_i$ to $r_i'$, will force us to take the burden to rebuild the system model and search the optimal solution again. Unfortunately in a real network, especially the wireless network, some parameters, e.g. the system capacity, keep varying from time to time due to the channel nature.
2) There is no guarantee that the policy iteration algorithm and value iteration algorithm can converge to optimal solution under any condition. This prevents them from being applied to real systems, if the stability of the system is a concern.
3) If there are more than two types of users appearing in the system, the SMDP will be more than three dimensional and not mention to express all the transition probabilities. This is called curse of dimension, which is hard to be solved.

Above all, the standard optimal algorithms cannot adapt to any change of environment easily, though they are applied to off-line CAC scheme design under some condition. On the other hand, another set of algorithms named Reinforcement Learning (RL) trains the interaction between agent and dynamic environment through learning. A typical model of these algorithms is described as follows:

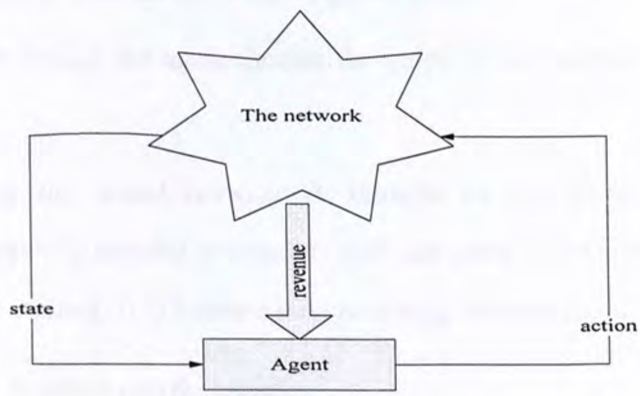


Fig. 11 The Interaction between Agent and Network of RL

As depicted in Fig. 11, an agent makes an action decision for the network considering feedbacks from the network: what state the network stays in now and the revenue collected by the network with the history decisions of the agent. After repeating the trial-and-error process for a long enough time, the agent learns an admission control scheme of choosing actions that can lead to optimal average revenue for the system. Some of the RL algorithms are model-based, because the agent tries to build the model through past experience first, and then derives an optimal admission control scheme for this specific model. Other RL algorithms are model-free, for the agent is trained how to react to the network best through long term learning without building up the system model.

The comparative lower computing model-free algorithm is more attractive. A representative model-free algorithm called Q-learning [36] is proposed for its easier implementation. There is a value $Q(s,a)$ for each state $s$ with action $a$ in the system,

$R$ is the actual lump sum revenue the system collects until next decision epoch, if action $a$ is chosen. The optimal action set is proved to be achieved by updating these $Q$ values recursively:

$$Q(s,a) := Q(s,a) + \alpha(R + \gamma \max_{a'} Q(s',a') - Q(s,a)) \tag{53}$$

This algorithm runs in this way:

a) $Q(s,a)$ is randomly initialized as a non-negative value.

b) When staying in state $s$, the agent chooses the action $a = \arg\max_a Q(s,a)$ as the decision.

c) After receiving the actual revenue $R$ brought by this decision, the corresponding $Q$ value is updated as equation (53) that takes future expectation $\max_{a'} Q(s',a')$ into account. It is because the uncertainty of expectation that one discount factor $\gamma$ is added into the equation.

If this value updating iteration runs an infinite long time and the learning speed $\alpha$ decays appropriately, this procedure will lead to an optimal policy $\pi^*(s) = \arg\max_a Q^*(s,a)$ and the revenue contribution of each state to the whole system $V^*(s) = \max_a Q^*(s,a)$. However, there is an obstacle keeping it from an efficient CAC scheme in a large scale network: this algorithm needs a huge space for storing the $Q$ value as well as a pretty long time to converge to the optimal stationary policy.

# 4.3 The Average Revenue Maximum CAC

From above brief introduction of model-based algorithms and the model-free counterparts, we find there are some common characters in them: both search the optimal policy in a greedy way. Both the revenue contributed by history decision and an imprecise discounted prediction of future caused by the decision are recorded as a key factor, and the system becomes more and more stable with the value updating procedure after a long enough time. Starting from the same basic idea, we propose a greedy algorithm for on-line call admission control scheme design in real networks, no matter it is a large one or small one. As described in Section 4.1, each type of users can be defined by $\{r_i, \lambda_i, \mu_i\}$. We keep a record of the contributed revenue for each type of user as $H(i)$, which denotes the history revenue contribution. Whenever there is an arrival event or departure event occurring in the system, $H(i)$ needs to be updated:

$$H(i) := H(i) + r_i n_i t_i \tag{54}$$

where $n_i$ is the number of type-*i* user served in the network during two events interval $t_i$. There is another parameter $D(i)$ for each type of user, when making admission decision. Only type-*i* user's $D(i)$ is in following form, if the coming user type is $i$ :

$$D(i) = H(i) + \gamma r_i \tau_i \tag{55}$$

Otherwise $D(j) = H(j)$ for all $j \neq i$. The type-*i* user will be admitted with probability $P_a = D(i) / \sum_i D(i)$. $\gamma r_i \tau_i$ characterizes the expected revenue caused by admission of this user, and it is discounted due to the uncertainty of future. Our scheme can be described as following flow chart:

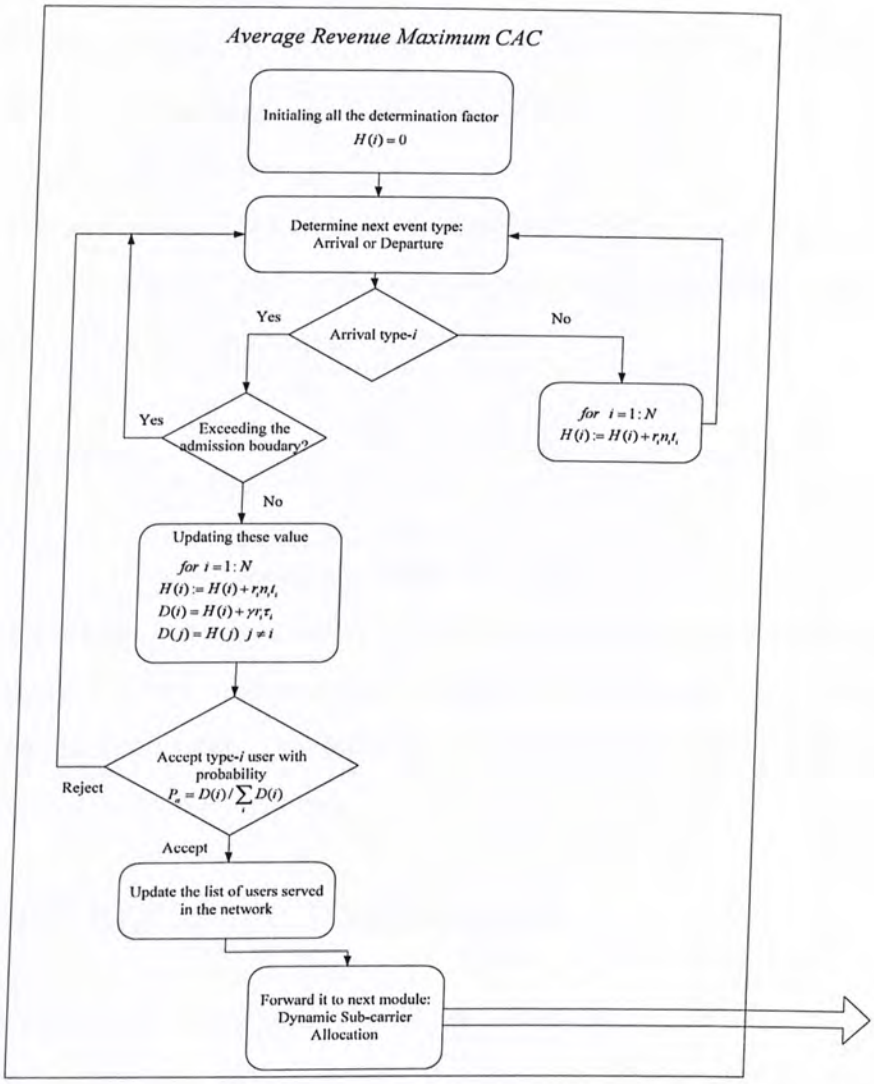


Fig. 12 Average Revenue Maximum CAC (ARM-CAC)

The admission region boundary in Fig.12 is determined by the first stage, the average revenue maximum CAC (ARM-CAC) scheme deals with the arrival and departure dynamic of users from the service provider point of view. We propose this probability control scheme mainly due to two reasons: first, the revenue contributed by one user is proportional to the admission chance of this type of users; second, the probabilistic scheme adapts to varied environment better than a fixed one in correcting any small violation on the system.

Compared with Q-learning algorithm that is also model-free, our scheme saves a great amount of storage space: generally, the Q-learning algorithm needs a storage space $|S|\times|A|$, where $|S|$ and $|A|$ are the cardinality of the SMDP state space and action space respectively. However, our scheme only requires $N$ units of space for $N$ types of users.

If there are totally 2 types of users in the network, and the capacity of the network is 6 in terms of total number of users. We compare the storage space needed by Q-learning algorithm and our scheme:

| | Storage space |
|---|---|
| Q-learning Algorithm | 7×8×3/2-7×2-1=69 |
| ARM-CAC | 2 |

Table 1 Comparison of storage space of two schemes

In fact, our scheme is not constrained in wireless networks. Once given the capacity of one network, it can work to obtain maximum average revenue in any kind of systems. In the next section, we compare the performance of our scheme with linear programming solution that is optimal.

# 4.4 Performance Evaluation

In this section, we simulate several discrete event systems controlled by our ARM-CAC scheme and compare it with the long-term average system revenue obtained by the linear programming CAC method. Assume that the system can support 6 users in total, regardless of the user type.

If there is just single type of user, complete sharing policy is global optimal. Under this condition, our scheme converges to the complete sharing policy naturally. The QoS-provisioning CAC scheme is good enough for the whole system. When the traffic load increases, the average revenue will converge to a constant due to the limited capacity of the network. In fact, this system is a M/M/6 queueing system with no buffer, so the incoming user is either served or dropped immediately.

We have two kinds of users: one type is associated with $\{r_1=3,\lambda_1=10,\mu_1=10\}$ and the revenue rate of the other type is $\{r_2=5,\lambda_2,\mu_2=10\}$, where the arrival rate

$\lambda_2$ varies from 1 to 10. We compare our scheme with the optimal results in the multiple user conditions.

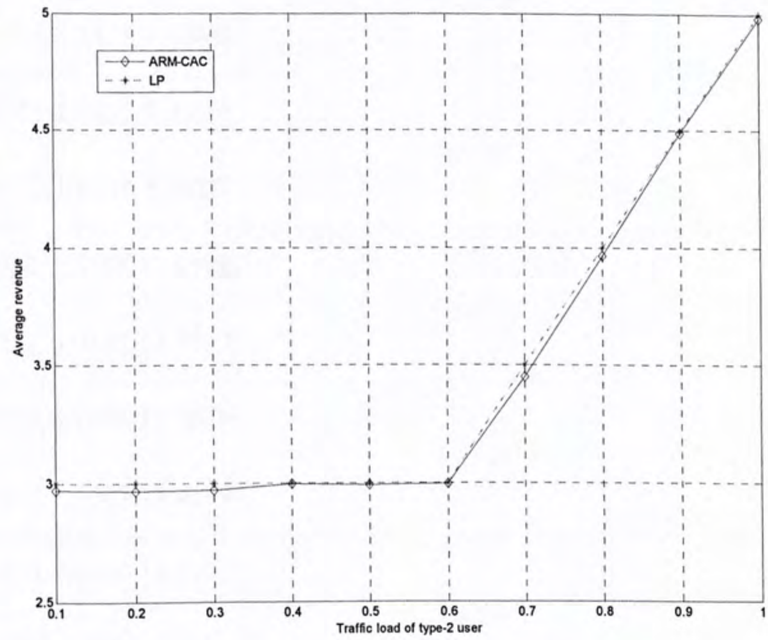


Fig. 13 Performance Comparison under Multiple User Case

In Fig. 13, when the traffic load of type-2 user is low, most of the revenue is contributed by type-1 user, and the average revenue received keeps a constant. As the traffic load of type-2 user increases, the system will collect more and more revenue from type-2 user. In the case that one user's arrival rate is changing, the simulation results of our scheme can keep close to the optimal results. The gap between results of our scheme and the optimal results is very small in this simulation.

At last, we randomly generate users to validate our scheme by more simulation results:

| Two types of users $\{r, \lambda, \mu\}$ | Theoretical optimal result from LP | Result of ARM-CAC |
|---|---|---|
| $\{2.0211, 3.0673, 4.8320\}$ $\{5.0134, 8.7455, 4.2808\}$ | 10.1074 | 10.1032 |
| $\{3.3025, 3.5856, 7.0525\}$ $\{5.3172, 1, 6119, 5.3327\}$ | 1.6790 | 1.6520 |
| $\{5.9950, 2.5798, 17.0372\}$ $\{3.3232, 5.0385, 13.7799\}$ | 1.2151 | 1.1450 |
| $\{9.7914, 5.7906, 13.7678\}$ $\{6.3437, 7.5607, 19.4799\}$ | 4.1181 | 4.0928 |
| $\{3.7423, 8.6500, 13.2786\}$ $\{7.5665, 5.1483, 5.2405\}$ | 7.4299 | 7.3894 |
| $\{8.6994, 7.4821, 8.8156\}$ $\{6.9850, 3.5817, 10.5295\}$ | 7.3818 | 7.3569 |
| $\{6.7420, 1.1386, 14.9578\}$ $\{0.4180, 6.0942, 6.7551\}$ | 0.5132 | 0.5128 |

Table 2 Comparison under multiple user case

It is clear that no matter how the condition changes, our scheme can always guarantee a close to optimal result. Since our scheme is derived for any multiple-user case, it is proved to be a good CAC scheme with little storage space and computing burden in wide simulations at least.

Cooperating with previous QoS-provisioning CAC that determines the admission region, the average maximization CAC can achieve the maximum long-term average revenue for the multiple user OFDM wireless networks. Furthermore, it is not constrained in the OFDM wireless network. Given the system capacity, it can be extended to any network concerning about system revenue maximization.

# Chapter 5

# Conclusion and Future Work

We propose a QoS-aware adaptive call admission control strategy in the context of multi-user OFDM wireless networks. This strategy realizes call admission control in two stages: the first stage aims to protect the service quality of existing users. By estimating the average throughput of both existing and in-coming users, the strategy admits the new coming user only when the minimum data rate requirements of all users, including the new one, can still be satisfied. The second stage aims to maximize the average system revenue. This strategy selectively admits the "admissible" users according to the revenue the users contribute to the system.

There are several advantages of our CAC strategy:

1) We analyze the achievable data rate of each user in the OFDM wireless system with a specific resource allocation scheme—dynamic sub-carrier allocation to design the CAC strategy, which fully exploits the multi-user diversity and time diversity in this environment. The analysis facilitates us to construct an admissible region for multi-type users.
2) Without any knowledge of the network, this strategy can make use of the history revenue record and discounted expected revenue of each type of users to make an admission decision. Our results show that the scheme achieves close-to-maximum average system revenue.
3) The proposed CAC strategy performs pretty well both in QoS guarantee and revenue maximization with low complexity. It is possible to implement it as an on-line call admission control strategy in future broadband wireless communication system.

Though it is expected to be applied in the future systems like LTE, WiMax and the like, some related problems should be solved before the implementation such as how accurate physical channel information we can achieve will affect the performance severely.

As a future work, the throughput estimation algorithm can be improved by taking into considering of the user arrivals and departures, if users do not stay in the network for a long enough time so that they do not experience all the possible states of the wireless fading channel. On the other hand, though nearly all the current works treat the revenue rate of users as a constant related to the user only, many communication systems concern about the system utility related to the throughput. How to incorporate this kind of utility in the call admission control strategy is another possible direction for future work.

# Bibliography

[1] R. Ramjee, R. Nagarajan and D. Towsley, "On optimal call admission control in cellular networks", in *Proc. IEEE INFOCOM1996*, pp. 43-50, Mar., 1996

[2] S. Wu, K. Y. Wong and B. Li, "A dynamic call admission policy with precision QoS guarantee using stochastic control for mobile wireless networks", *IEEE/ACM Trans. Network.*, vol. 10, no. 2, pp. 257-271, Apr., 2002

[3] M. Grossglauser and D. Tse, "A framework for robust measurement-based admission control", *IEEE/ACM Trans. Network.*, vol. 7, no. 3, pp293-pp309, June, 1999

[4] H. S. Wang and N. Moayeri, "Finite-state Markov channel – a useful model for radiio communication channels", *IEEE Trans. Veh. Tech.*, vol. 44, no. 1, pp. 163-171, Feb., 1995

[5] Q. Zhang and S. A. Kassam, "Finite-state Markov model for Rayleigh fading channels", *IEEE Trans. Comm.*, vol. 47, no. 11, pp. 1688-1692, Nov., 1999

[6] M. Hassan, M. M. Krunz and I. Matta, "Markov-based channel characterization for tractable performance analysis in wireless packet networks", *IEEE Trans. Wireless Comm.*, vol. 3, no. 3, pp. 821-831, May, 2004

[7] D. Niyato and E. Hossain, "Connection admission control algorithms for OFDM wireless networks", in *Proc. IEEE GLOBECOM 2005*, vol. 6, pp. 2455-2459, Nov.-Dec., 2005

[8] L. B. Le, E. Hossain and A. S. Alfa, "Queueing analysis and admission control for multi-rate wireless networks with opportunistic scheduling and ARQ-based error control", in *Proc. IEEE ICC2005*, vol. 5, pp. 3329-3333, May, 2005

[9] S. L. Spitler and D. C. Lee, "Optimal call admission control under packet and call level QoS constraints and effect of buffering", *Int, J. Commun. Syst.*, 16, pp. 647-662, May, 2003

[10] J. P. Linnartz, "Exact analysis of the outage probability in multiple-user mobile radio", *IEEE Trans. Comm.*, vol. 40, no. 1, pp. 20-23, Jan., 1992

[11]J. Zhang, L. Li, B. Li and P. Wu, "QoS-based connection admission control in MC-CDMA cellular networks", In *Proc. IEEE ICC2004*, vol. 7, Jun., 2004

[12]T. Javidi and D. Teneketzis, "An approach to connection admission control in single-hop multi service wireless networks with QoS requirements", *IEEE Trans. Veh. Tech.*, vol. 52, no. 4, pp. 1110-1124, Jul., 2003

[13]B. Rong, Y. Qian and K. Liu, "Integrated downlink resource management for multiservice WiMAX networks", *IEEE Trans. Mobile Comput.*, vol. 6, no. 6, Jun., 2007

[14]X. Yu, W. Xie and Y. J. Zhang, "QoS-aware adaptive call admission control in multiuser OFDM wireless network", *Technical Report 2008*

[15]H. C. Tijms, Stochastic Modelling and Analysis: A Computational Approach, New York, Wiley, 1986

[16]K. W. Ross and D. Tsang, "Optimal circuit access policies in an ISDN environment: a Markov decision approach", *IEEE Trans. Comm.*, vol. 37, no. 9, pp.934-939, Sep., 1989

[17]J. Ni, D. Tsang, S. Tatikonda and B. Bensaou, "Threshold and reservation based call admission control policies for multiservice resource-sharing systems", in *Proc. IEEE INFOCOM2005*, vol. 2, pp. 773-783, Mar., 2005

[18]J. Ni, D. Tsang, S. Tatikonda and B. Bensaou, "Optimal and structured call admission control policies for resource-sharing systems", *IEEE Trans. Comm.*, vol. 55, no. 1, pp. 158-170, Jan., 2007

[19]E. Altman, T. Jiménez and G. Koole, "On optimal call admission control in a resource-sharing system", *IEEE Trans. Comm.*, vol. 49, no. 9, Sep., 2001

[20]Y. Xiao, C. L. Chen and Y. Wang, "A near optimal call admission control with genetic algorithm for multimedia services in wireless/mobile networks", in *Proc. NAECON2000*, pp. 787-792, Oct., 2000

[21]D. Karabudak, C. C. Hung and B. Bing, "A call admission control scheme using genetic algorithms", in *Proc. ACM Symp. Applied Computing 2004*, pp. 1151-1158, Mar., 2004

[22] L. P. Kaelbling, M. Littman and A. Moore, "Reinforcement learning: a survey", *J. Artificial Intelligence Research*, 4, pp. 237-285, May, 1996

[23] S. Singh, "Learning to solve Markovian Decision Process", Ph.D. Dissertion, University of Massachusetts Amherst, 1994

*[24]* T. Kwon, Y. Choi and M. Naghshineh, "Optimal distributed call admission control for multimedia services in mobile cellular network", In *Proc. 5th Intl. Workshop on Mobile Multimedia Communications MoMuc'98*

[25] S. M. Senouci, A. L. Beylot and G. Pujolle, "Call admission control in cellular networks: a reinforcement learning solution", *Int. J. Network Mgmt*, 14, pp. 89-103, 2004

[26] C. Y. Wong, R. Cheng, K. B. Khaled and R. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation", *IEEE J. Select. Areas Comm.*, vol. 17, no. 10, pp. 1747-1758, Oct. 1999

[27] G. Song and Y. Li, "Cross-layer optimization for OFDM wireless networks—part I: theoretical framework", *IEEE Trans. wireless Comm.*, vol. 4, no. 2, pp. 614-624, Mar., 2005

[28] G. Song and Y. Li, "Cross-layer optimization for OFDM wireless networks—part II: algorithm development", *IEEE Trans. wireless Comm.*, vol. 4, no. 2, pp. 625-634, Mar., 2005

[29] Y. J. Zhang and K. B. Letaief, "Cross layer adaptive resource management for wireless packet networks with OFDM signaling," *IEEE Trans. Wireless Comm.*, vol. 5, no. 11, pp. 3244-3254, Nov. 2006.

[30] Y. J. Zhang and K. B. Letaief, "Multiuser adaptive subcarrier-and-bit allocation with adaptive cell selection for OFDM systems" *IEEE Trans. Wireless Comm.*, vol. 3, No. 5, pp. 1566-1575, Sep. 2004.

[31] K. B. Letaief and Y. J. Zhang, "Dynamic multiuser resource allocation and adaptation for wireless systems" *IEEE Wireless Communications (invited paper)*, vol. 13, no. 4, pp. 38-47, Aug. 2006.

[32] A. J. Goldsmith and S. G. Chua, "Variable-rate variable-power MQAM for fading channels", *IEEE Trans. Comm.*, vol. 45, no. 10, pp. 1218-1230, Oct. 1997.

[33]M. S. Alouini and A. J. Goldsmith, "Adaptive M-QAM modulation over Nakagami fading channels", in *Proc. GLOBECOM1997*, pp. 218-223, Nov., 1997

[34]T. Keller and L. Hanzo, "Adaptive multicarrier modulation: a convenient framework for time-frequency processing in wireless communications", in *Proc. IEEE*, vol. 88, no. 5, pp. 611-640, May, 2000

[35]S. Boyd and L. Vandenberghe, Convex Optimization. New York: Cambridge University Press, 2006

[36]Watkins, C. J. C. H. and Dayan, "Q-learning", *Machine Learning*, 8(3), pp. 279-292, 1992