

An Optimal Admission Control Protocol for Heterogeneous Multicast Streaming Services

Tae Ok Kim, Chamara N. Devanarayana, *Student Member, IEEE*,
Attahiru S. Alfa, *Member, IEEE*, and Bong Dae Choi, *Member, IEEE*

Abstract—We investigate optimal call admission control (CAC) policy for multicast streaming services (MSS) in 3rd generation (3G) and beyond wireless mobile networks. Several MSS sessions are supported simultaneously in a bandwidth-limited network. Active sessions are those that are currently serving some users, and inactive sessions are those that are currently not serving any users. An admission decision in MSS is required only when an inactive session is requested, unlike in unicasting. For this reason, if a user request for an inactive MSS session arrives, we should make an admission decision in anticipation of (i) the possible reward earned based on users served during a session active time generated by accepting it, and (ii) the influence of the session active time upon the future status of network bandwidth and admission decisions. Our objective is to determine when to admit or block a user asking an inactive MSS session to achieve the optimality in rewards. We formulate this problem as a semi-Markov decision process (SMDP), and a value iteration algorithm is used to obtain an optimal stationary deterministic policy. We also derive the user blocking probability of the optimal policy by analyzing an embedded Markov chain induced by it.

Index Terms—Multicast service, admission control policy, Markov decision policy, performance evaluation, Markov chain.

I. INTRODUCTION

IN recent years, the demand on multimedia streaming applications such as mobile IPTV (Internet Protocol Television) and location-based applications (location-based advertisement and queries for location-based services) over wireless mobile networks has been dramatically increasing. Unlike traditional voice services, multimedia streaming services are characterized by large bandwidth requirements. However, unicasting is not

designed to support such requirements of multimedia streaming services. Typical standardization bodies for the 3rd generation (3G) wireless mobile networks, such as 3GPP (3G partnership project) [1] and IEEE 802.16e [2], addressed this problem by introducing multicast/broadcast services, so-called MBMS (multimedia broadcast and multicast services) and MBS (multicast and broadcast services), respectively. As multicast streaming service has the point-to-multipoint nature in which a single copy of data is conveyed to multiple receivers asking the same service, it significantly improves the scalability of multimedia streaming services. For example, Yota Telecom [3] has started mobile TV services with 25 live TV channels over its 10 Mbps mobile WiMAX networks.

Although multicast streaming services (MSS) are already deployed in industrial fields, optimal admission control is still a crucial design issue to be investigated in bandwidth-limited wireless mobile networks. In unicast services, a certain amount of bandwidth is allocated to each admitted user and it is released right after the user terminates the service, so that an admission decision is made at every user arrival. However, in MSS, once a session becomes active, all the arriving users are accepted and the bandwidth occupation lasts until there are no users and then the session goes inactive. Consequently, an admission decision in MSS is required only when an inactive session is requested, and the ripple effect of the admission control in MSS on the future status of network bandwidth and admission decisions is greater than that of unicast services. With these different features from unicast services, a proper admission control scheme for MSS should be carefully designed to optimize the overall network operation, and popularity of each session must be a key factor in the design of the scheme.

Huang *et al.* [4] investigated an efficient admission control policy for multicast and unicast integrated services. However, the way to differentiate multicast services with different popularity was not covered in that work. Lohmar *et al.* [5] and Catrein *et al.* [6] have proposed a hybrid broadcast-unicast scheme for multimedia streaming services to differentiate between popular and unpopular services, but they did not consider multicast service. In multicast service, users have to declare their interest for a specific multicast session so that the session needs to be activated only when there exist users who are interested in it, while sessions are always activated without knowing whether there is any user in broadcast service. It was established that multicasting is a more efficient method to support group streaming service than unicasting or broadcasting since it allows packet transmission/routing with fewer network resource [7]. However admission control to differentiate MSS services

Manuscript received August 20, 2014; revised February 13, 2015; accepted April 1, 2015. Date of publication April 10, 2015; date of current version June 12, 2015. This work was supported by the NSERC fund and National Research Foundation of Korea grant (No: 2012-008099). A part of this work was presented at the IEEE International Conference on Communications 2012 held in Ottawa, ON, Canada. The associate editor coordinating the review of this paper and approving it for publication was B. Liang.

T. O. Kim was with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB R3T 5V6, Canada (e-mail: taekkim98@gmail.com).

C. N. Devanarayana is with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB R3T 5V6, Canada (e-mail: umdevana@myumanitoba.ca).

A. S. Alfa is with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB R3T 5V6, Canada, and also with the Department of Electrical, Electronic and Computer Engineering, University of Pretoria, Pretoria 0002, Africa (e-mail: Attahiru.Alfa@umanitoba.ca).

B. D. Choi is with the Research Institute for Information and Communication Technology, Korea University, Seoul 136-701, Korea (e-mail: queue@korea.ac.kr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCOMM.2015.2421896

according to their popularity is not yet investigated to the best of our knowledge. This work is more essential considering the download frame structure of IEEE 802.16e standard, in which MSS service is supported on a separate bandwidth zone from unicast service [8].

The Markov decision process (MDP) [9] is a well-known stochastic control technique, and widely used to solve optimization problems in extensive practical applications. The greatest advantage of MDP-based optimal control policy is that each decision is made by taking into account both present and future outcomes. As a result, MDP has been adopted in numerous research dedicated to admission control schemes in bandwidth-limited wireless networks [10]–[16]. Ho *et al.* [10], Choi *et al.* [11] and Ni *et al.* [12] investigated a MDP-based optimal call admission control policy with multiple traffic classes to maximize system utilization/revenue. Yu *et al.* [13], Stevens-Navarro *et al.* [14], Chen *et al.* [15] and Bühler *et al.* [16] applied the MDP-based approach to integrated heterogeneous wireless networks to find an optimal call admission control mechanism. However, none of the quoted references dealt with the admission control for MSS.

In this paper, we present a generic framework for the optimization on admission control for MSS systems that supports heterogeneous sessions in terms of the popularity and the mean user service time. Our objective is to achieve the optimal expected infinite-horizon discounted reward where admitting a user request brings a reward, rejecting it incurs some penalty, and setting up service of an inactive MSS session involves some cost. This problem is formulated as a semi-Markov decision process (SMDP). What makes admission control for MSS distinguishable from unicasting is that, (i) the decision is required only when an inactive session is requested and (ii) that decision should be made by considering in advance the entire reward obtained by all the users arriving during a session active time. We estimate the entire reward by analyzing the stochastic behavior of user arrival/departure during a session active time. With a finite state space and a finite action space, a value iteration algorithm is used to obtain an optimal stationary deterministic policy. The optimal policy is proven to have a structural tendency to reject more user requests for inactive sessions when the system is getting congested. Given the optimal stationary deterministic policy, we derive the user blocking probability of each MSS session/overall system by analyzing an embedded Markov chain induced by the policy. We demonstrate how to implement the optimal policy using a lookup table, and also numerically evaluate the expected reward and user blocking probability under the optimal policy. Furthermore, we ascertain the gain on the expected reward of the optimal policy over the complete sharing policy, in which every user requesting an inactive MSS session is accepted as long as the bandwidth is available.

The remainder of this paper is organized as follows. In Section II, the system model and the problem statement are described. In Section III, we analyze the stochastic behavior of user arrival/departure to estimate the entire reward during a session active time. The SMDP model formulation is presented in Section IV. Based on this model, the optimality equations and the value iteration algorithm to determine the optimal sta-

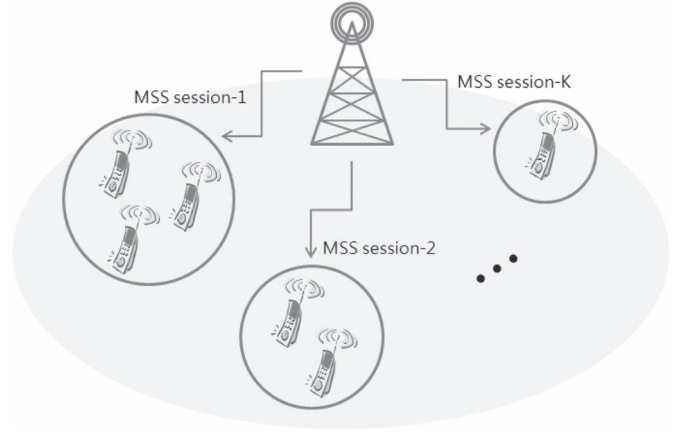


Fig. 1. Multicast services.

tionary deterministic policy are developed in Section V. Then we analyze the user blocking probability in Section VI. We present numerical results in Section VII, and give concluding remarks in Section VIII.

II. SYSTEM MODEL AND PROBLEM STATEMENT

A. System Model

We consider MSS system in a single-hop wireless cellular network as shown in Fig. 1. Specifically, the base station is responsible for offering K MSS sessions where each MSS session is transmitted to multiple end users who have interests in it. The system bandwidth is C and each session- k consumes a bandwidth w_k where $\sum_{k=1}^K w_k > C$. User requests arrive according to a Poisson process with rate λ over time and an arriving user request belongs to a specific MSS session- k with probability p_k , so that the users requests for MSS session- k also arrive according to a Poisson process with rate $\lambda_k (= \lambda p_k)$. We refer to p_k as the popularity of the session, and assume that MSS sessions are sorted in descending order of popularity, i.e., $p_1 \geq p_2 \geq \dots \geq p_K$. The popularity of a session is assumed to follow a Zipf-like distribution [17]–[19], so that the relative probability of a request for session- k is proportional to $1/k^\psi$ where ψ determines the skewness in the Zipf-like distribution. Then p_k is given by Ω/k^ψ where $\Omega = (\sum_{k=1}^K (1/k^\psi))^{-1}$. The service time of a user with MSS session- k is assumed to be exponentially distributed with rate μ_k . For the MSS session- k , the number of concurrent users is restricted to N_k due to fading environments [20], and so the user arriving when N_k users are already with the session is rejected. This truncation is essential for modelling the session holding time of the users as a Phase-Type distribution. The value for N_k can be calculated based on the minimum SNR requirement β_k and the outage probability ϵ to ensure sufficient Quality of Service (QoS) for the users. If we assume the users around the base station to be distributed according to the Poisson point process with intensity λ , the path loss exponent to be γ and the channel to experience flat Rayleigh fading, the value for N_k can be represented according to Equation (1), where $\rho = \frac{P_{\text{total}}}{C}$ is the transmission power density, $\Pi = \frac{22}{7}$ and N_0 is the noise power

density. The decision to block users when there are N_k users already in the session can be treated as a dummy decision which is there to take into account the cost of unsatisfied users. This truncation is further required to express the active time of the sessions as Phase-Type and to make the problem complexity low.

$$N_k = \left\lfloor p_k \lambda \Pi \left(-\frac{\log(1-\epsilon)\rho}{\mathcal{N}_0 \beta_k} \right)^{\left(\frac{2}{\gamma}\right)} \right\rfloor \quad (1)$$

B. Problem Statement

1) *Objective*: We describe the state s of the MSS system at time t by the vector $\mathbf{n}(t) \in \mathcal{N}$ which is defined by

$$\mathbf{n}(t) = (n_1(t), n_2(t), \dots, n_K(t)), \quad (2)$$

where $n_k(t)$ denotes the number of users who are already accepted (concurrent users) in session- k at time t and the state space \mathcal{N} is given in Equation (3), shown at the bottom of the page. A table of notations used can be found in Table I.

In Equation (3), $\mathbf{1}_{\{\cdot\}}$ denotes the indicator function, *i.e.*, it is 1 if the statement inside the braces is true; otherwise, it is 0. A policy π specifies the decision rules $d_i (i = 1, 2, \dots)$ to be used at every decision epoch σ_i . In our MSS admission control problem, decisions are required only when a user requests an inactive session, *i.e.*, decision epochs σ_i correspond to the time immediately following those user arrivals. At each decision epoch σ_i , the admission decision should be made considering the current system state and any possible future system states.

For each policy π , let $v_\alpha^\pi(s)$ be the expected infinite-horizon discounted reward per unit bandwidth usage with the discount rate α , given that the process occupies state s when the system is initiated. This parameter α plays a role in the convergence rate of the algorithm. The smaller the value of α faster will be the convergence. We leave this as a parameter to be decided by the operator, which is the norm in the literature concerning MDP. Here onwards the term reward represents the reward per unit bandwidth usage. Our problem is to construct a policy π that maximizes the expected infinite-horizon discounted reward $v_\alpha^\pi(\mathbf{n}(0))$ for every initial state $\mathbf{n}(0)$. The decision rule d_i at the epoch σ_i specifies the action “accept/reject” when a user requesting the inactive MSS session- k arrives at the system in state $\mathbf{n}(i-1)$. Thus, the objective function is given in Equation (4), shown at the bottom of the page, where $e_i \in \{A_k : k \text{ s.t. } n_k = 0\}$ stands for the event at the decision epoch

TABLE I
TABLE OF NOTATIONS USED

Parameter	Description
$\mathbf{n}(t)$	Array containing the number of users using each session at time t
\mathcal{N}	State space
N_k	Maximum number of users in Session k
p_k	Probability of a user arrival to Session k
π	The Policy
$v_\alpha^\pi(s)$	Expected infinite horizon discounted reward per unit band width
α	Discount rate
$r(s, a)$	Received reward by taking action a at state s
A_k	Arrival to an inactive Session k
D	Departure of user
a_A	Action to accept an arriving user
a_R	Action to reject an arriving user
a_C	Action to continue
$\mathbb{E}_\alpha(U_k)$	Discounted number of MSS Session k users served
$\mathbb{E}_\alpha(D_k)$	Discounted number of MSS Session k users rejected
\mathbf{T}	Transition matrix of a Ph-Type Distribution
τ	Initial distribution of the Ph-Type Distribution
\mathbf{T}^0	Absorption Matrix of the Ph-Type Distribution
λ	Session arrival rate
μ_k	Service rate of Session k
$q(s' s, a_s)$	State transition probability from s to s' when action a_s is taken
$d(s)$	Decision taken at State s
P_k	Instantaneous penalty of rejecting a user for Session k
R_k	Instantaneous reward of accepting a user for Session k
C_k	Setup cost for Session k
w_k	Bandwidth usage for session k

σ_i (A_k denotes a user arrival for the inactive session- k) and $r(s, e, a)$ represents the received reward by taking the action a when the event e arises in the state s .

2) *Reward*: From the service provider’s perspective, we consider that (i) C_k units of cost per unit bandwidth are needed to setup MSS session- k (ii) every single user request for MSS session- k would contribute a payoff of R_k units to the revenue per unit bandwidth usage (*e.g.*, monetary value) when it is successfully accepted and served, and conversely a penalty of P_k units per unit bandwidth usage indicating the negative value (*e.g.*, potential damage from blocked users’ dissatisfaction) imparted to the system when a user is rejected [21]. These values are measured and given by service providers based on their statistical data and service policies. The specific mechanism to determine these is beyond the scope of the paper. Thus we focus on optimizing the expected infinite-horizon discounted reward under the assumption that the values are known *a priori*.

$$\mathcal{N} := \left\{ (m_1, m_2, \dots, m_K) : \forall k \ 0 \leq m_k \leq N_k, \sum_{k=1}^K w_k \cdot \mathbf{1}_{\{m_k > 0\}} \leq C \right\}. \quad (3)$$

$$v_\alpha^\pi(\mathbf{n}(0)) = \mathbb{E}_{\mathbf{n}(0)}^\pi \left\{ \sum_{i=1}^{\infty} e^{-\alpha \sigma_i} \cdot r(\mathbf{n}(i-1), e_i, d_i(\mathbf{n}(i-1), e_i)) \right\}. \quad (4)$$

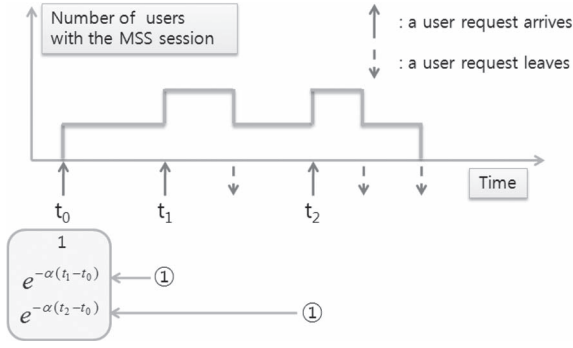


Fig. 2. Discounted number of users served/rejected users during a session active time.

Once the service of the MSS session- k starts, all the arriving users are accepted as long as the number of concurrent users with the session is less than N_k . For this reason, the reward that the system would obtain by taking the action “accept” is the entire reward during a session active time, which consists of (i) the setup cost, (ii) the payoff generated by all the served users and (iii) the penalty generated by all the blocked users due to the limitation on the maximal number N_k of concurrent users. Thus we estimate the expected reward arising from accepting the user for the inactive MSS session- k with the discount rate α as $\{R_k \mathbb{E}_\alpha(U_k) - C_k - P_k \mathbb{E}_\alpha(D_k)\}$, where $\mathbb{E}_\alpha(U_k)$ and $\mathbb{E}_\alpha(D_k)$ are the expectations for the discounted numbers of served users and rejected users during a session active time, respectively, which are derived in Section III.

III. DISCOUNTED NUMBERS OF SERVED/REJECTED USERS DURING A SESSION ACTIVE TIME

In this part, we derive (i) the expectation $\mathbb{E}_\alpha(U_k)$ for the discounted number of MSS session- k users served during a session active time and (ii) the expectation $\mathbb{E}_\alpha(D_k)$ for the discounted number of MSS session- k users rejected due to the limitation on the maximal number N_k of concurrent users during a session active time, where the discount rate α is given. The term *discounted number* means the converted number of the events during a session active time to the value at the moment that the session service starts. In Fig. 2, we display an example for the discounted number of users served during a session active time of a tagged MSS session. In this example, there are three user request arrivals at times $\hat{t}_0, \hat{t}_1, \hat{t}_2$ during the session active time where \hat{t}_0 is the moment that the session service starts. It is trivial that the first arrival does not suffer any discount on the value. With the discount rate α , the discounted values for the second and the third arrivals are $e^{-\alpha(\hat{t}_1-\hat{t}_0)}$ and $e^{-\alpha(\hat{t}_2-\hat{t}_0)}$, respectively. Thus the discounted number of served users in this example is $\{1 + e^{-\alpha(\hat{t}_1-\hat{t}_0)} + e^{-\alpha(\hat{t}_2-\hat{t}_0)}\}$. The

discounted number of rejected users during a session active time is described analogously. The discounted expectations $\mathbb{E}_\alpha(U_k)$ and $\mathbb{E}_\alpha(D_k)$ are used in the estimation of the reward exclusive of setup cost and penalty incurred by accepting a user arriving while the session is inactive. The estimated value is used in the reward function for the MDP describing the MSS admission control problem (see Section IV-E). In this section, we drop the index k from all the parameters and the variables to simplify notations.

Let L be the number of total steps at which the number of concurrent users changes in MBS session- k , and $\hat{t}_1, \hat{t}_2, \dots, \hat{t}_L$ be the (random) time points of user arrivals or departures. Then the embedded stochastic process $\{X(\hat{t}_l) : l = 1, 2, \dots, L\}$, which describes the number of concurrent users with the MSS session- k , is a discrete time Markov process with one absorbing state 0 and the transient states $1, 2, \dots, N$. Further the transition probability matrix \mathbf{P} is given by:

$$\mathbf{P} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{T}^0 & \mathbf{T} \end{bmatrix}, \quad (5)$$

where (i) \mathbf{T} is a square matrix of dimension N which corresponds to the transition of the Markov process $\{X(\hat{t}_l)\}$ over the transient states, (ii) $\mathbf{T}^0 + \mathbf{T}\mathbf{1} = \mathbf{1}$ (here, $\mathbf{1}$ is a column vector of ones with order N), and (iii) $\mathbf{0}$ denotes a row vector of zeros with proper order. The column vector \mathbf{T}^0 has the first component $\lambda/(\lambda + \mu)$, which is the only non-zero one. The components $\gamma_{i,j}$, ($1 \leq i, j \leq N$), of the matrix \mathbf{T} are defined in Equation (7), shown at the bottom of the page. Here, $\lambda/(\lambda + \mu)$ gives the probability that a user arrives before another user departs when there are i users currently in the system and $i\mu/(\lambda + i\mu)$ gives the probability that a user departs before another one arrives. Then the number of total steps L follows a discrete phase-type (PH) distribution denoted by $\mathbf{DPH}(\boldsymbol{\tau}, \mathbf{T})$, where $\boldsymbol{\tau}$ is the initial distribution. Note that $\boldsymbol{\tau} = [1, 0, \dots, 0]$ since a session active time is initiated by one user. The probability mass function of L is obtained by

$$\mathbb{P}(L = \eta) = \boldsymbol{\tau} \mathbf{T}^{\eta-1} \mathbf{T}^0 \quad \text{for } \eta \geq 1. \quad (6)$$

Now we derive the discounted expectations $\mathbb{E}_\alpha(U)$ and $\mathbb{E}_\alpha(D)$ by transforming the transition probability matrix \mathbf{P} as occasion demands. To do that, we consider the matrices \mathbf{P}_α , $\mathbf{P}_\alpha^{(U)}(z)$ and $\mathbf{P}_\alpha^{(D)}(z)$ corresponding to the matrix \mathbf{P} which are represented as follows.

- 1) The matrix \mathbf{P}_α takes the discount factor into account by:

$$\mathbf{P}_\alpha = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{T}^0 & \mathbf{T}_\alpha \end{bmatrix}, \quad (8)$$

$$\gamma_{i,j} = \begin{cases} \lambda/(\lambda + i\mu), & \text{if } (1 \leq i < N \text{ and } j = i + 1) \text{ or } (i = j = N), \\ i\mu/(\lambda + i\mu), & \text{if } 2 \leq i \leq N \text{ and } j = i - 1, \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

where the matrix \mathbf{T}_α has the same size as the matrix \mathbf{T} whose components $(\gamma_\alpha)_{i,j}$ are represented as in Equation (9), where $\delta_\alpha(i)$ is given in Equation (10).

$$(\gamma_\alpha)_{i,j} = \gamma_{i,j} \cdot \delta_\alpha(i), \quad (9)$$

for all i and j . In (9), $\delta_\alpha(i)$ denotes the discount factor to convert one unit of reward at the next event epoch to the present epoch, given that there are presently i concurrent users. The next event can either be an arrival or a departure from the set of i users which ever occurs first. Therefore the time to the next event follows an exponential distribution with rate $\lambda + i\mu$ and $\delta_\alpha(i)$ is obtained as in Equation (10).

$$\begin{aligned} \delta_\alpha(i) &= \int_0^\infty e^{-\alpha t} (\lambda + i\mu) e^{-(\lambda + i\mu)t} dt \\ &= \frac{\lambda + i\mu}{\alpha + \lambda + i\mu}. \end{aligned} \quad (10)$$

- 2) From the matrix \mathbf{P}_α , the matrices $\mathbf{P}_\alpha^{(U)}(z)$ and $\mathbf{P}_\alpha^{(D)}(z)$ additionally adopt a dummy variable z as follows:

$$\mathbf{P}_\alpha^{(U)}(z) = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{T}^0 & \mathbf{T}_\alpha^{(U)}(z) \end{bmatrix}, \quad (11)$$

$$\mathbf{P}_\alpha^{(D)}(z) = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{T}^0 & \mathbf{T}_\alpha^{(D)}(z) \end{bmatrix}, \quad (12)$$

where the matrices $\mathbf{T}_\alpha^{(U)}(z)$ and $\mathbf{T}_\alpha^{(D)}(z)$ have the same size as the matrix \mathbf{T}_α (or \mathbf{T}) whose components $(\gamma_\alpha^{(U)})_{i,j,z}$ and $(\gamma_\alpha^{(D)})_{i,j,z}$ are represented by equations (13) and (14) respectively (see equations at the bottom of the page). Note that, in $(\gamma_\alpha^{(U)})_{i,j,z}$, the entries corresponding to accepting a user for a particular active session are multiplied by a dummy variable z . Similarly in $(\gamma_\alpha^{(D)})_{i,j,z}$ the entries corresponding to rejecting a user for the session are multiplied by z . Here the dummy variable z acts as a mark of the corresponding event. By adopting it accordingly in $\mathbf{P}_\alpha^{(U)}(z)$ and $\mathbf{P}_\alpha^{(D)}(z)$, we are allowed to extract the expected discounted number of served and rejected users from matrix multiplications in (15) and (16), using differentiation with respect to z as in (17) and (18).

Consider the set \mathcal{P}_η of all the possible paths of the variation in the number of concurrent users, given that the session becomes inactive after η step transitions. For example, if $\eta = 5$ when $N = 3$ then $\mathcal{P}_5 = \{(1) \rightarrow 2 \rightarrow 1 \rightarrow 2 \rightarrow 1 \rightarrow$

$0\}, [(1) \rightarrow 2 \rightarrow 3 \rightarrow 2 \rightarrow 1 \rightarrow 0]\}$. We define $\varphi_\alpha^{(U)}(\eta; z)$ and $\varphi_\alpha^{(D)}(\eta; z)$ as follows:

$$\varphi_\alpha^{(U)}(\eta; z) = \tau \left\{ \sum_{l=1}^{\eta-1} (\mathbf{T}_\alpha)^{l-1} \mathbf{T}_\alpha^{(U)}(z) \mathbf{T}^{\eta-l-1} \right\} \mathbf{T}^0, \quad (15)$$

$$\varphi_\alpha^{(D)}(\eta; z) = \tau \left\{ \sum_{l=1}^{\eta-1} (\mathbf{T}_\alpha)^{l-1} \mathbf{T}_\alpha^{(D)}(z) \mathbf{T}^{\eta-l-1} \right\} \mathbf{T}^0. \quad (16)$$

Then, we observe that right-hand sides of both equations consist of a first-order term and a constant term only. Further, each first-order term, which can be obtained by differentiating (15) and (16), is rewritten as

$$\frac{d}{dz} \varphi_\alpha^{(U)}(\eta; z) = \sum_{p \in \mathcal{P}_\eta} \tilde{\delta}_\alpha^{(U)}(p) \cdot g(p), \quad (17)$$

$$\frac{d}{dz} \varphi_\alpha^{(D)}(\eta; z) = \sum_{p \in \mathcal{P}_\eta} \tilde{\delta}_\alpha^{(D)}(p) \cdot g(p). \quad (18)$$

Here the factors $g(p)$, $\tilde{\delta}_\alpha^{(U)}(p)$ and $\tilde{\delta}_\alpha^{(D)}(p)$ in (17) and (18) are interpreted as follows:

- $\tilde{\delta}_\alpha^{(U)}(p)/\tilde{\delta}_\alpha^{(D)}(p)$: the discounted number of users served/rejected during a session active time, given that the number of concurrent users varies according to the path p , respectively,
- $g(p)$: the probability that the number of concurrent users varies according to the path p .

In other words, each first term in (15) and (16) represents the expected discounted number of served (or rejected) users during a session active time, given that the session becomes inactive after η step transitions. Thus we finally obtain the expectations $\mathbb{E}_\alpha(U)$ and $\mathbb{E}_\alpha(D)$ as follows:

$$\mathbb{E}_\alpha(U) = 1 + \sum_{\eta=1}^{\infty} \frac{d}{dz} \varphi_\alpha^{(U)}(\eta; z), \quad (19)$$

$$\mathbb{E}_\alpha(D) = \sum_{\eta=1}^{\infty} \frac{d}{dz} \varphi_\alpha^{(D)}(\eta; z). \quad (20)$$

Note that the first term 1 on the righthand side of (19) corresponds to the first arrival which generates the session active time.

IV. MARKOV DECISION PROCESS FORMULATION

The admission control decision problem described in Section II is naturally formulated as a semi-Markov decision process (SMDP) [9] with decision epochs each time a user requests an inactive MSS session. At each decision epoch,

$$(\gamma_\alpha^{(U)})_{i,j,z} = \begin{cases} (\gamma_\alpha)_{i,j} \cdot z, & \text{if } 1 \leq i \leq N-1 \text{ and } j = i+1, \\ (\gamma_\alpha)_{i,j}, & \text{otherwise,} \end{cases} \quad (13)$$

$$(\gamma_\alpha^{(D)})_{i,j,z} = \begin{cases} (\gamma_\alpha)_{i,j} \cdot z, & \text{if } i = j = N, \\ (\gamma_\alpha)_{i,j}, & \text{otherwise.} \end{cases} \quad (14)$$

if the user requesting an inactive MSS session is accepted then the reward of $\{R_k \mathbb{E}(U_k) - C_k - P_k \mathbb{E}(D_k)\}$ is gained; otherwise, the penalty of P_k is paid. However, this SMDP is hardly tractable analytically due to the complexity in the transition probabilities. Thus we further take redundant decision epochs, which correspond to user arrivals for MSS session in service and user departures, to exploit the Markovian structure. The intervals between consecutive decision epochs becomes exponentially distributed in this way, since the Poisson user arrival process for each MSS session and the the exponential service time are assumed. Consequently, the admission control decision process is converted to a continuous-time Markov decision process (CTMDP) [9]. The memoryless property of the exponential distribution plays a key role in the analytical tractability. A CTMDP model is uniquely identified by the following five components: *the state space, the action space, the decision epochs, the transition probabilities and the reward function*. The notation we use follow those described in [9].

A. State Space

In the CTMDP, we replace the state space \mathcal{N} in the SMDP with the state space \mathcal{S} in which each state s consists of the number of concurrent users with each MSS session and the observation of user arrival/departure events, *i.e.*,

$$\mathcal{S} = \left\{ s \triangleq (\mathbf{n}, e) : \mathbf{n} \in \mathcal{N}, e \in \{A_k, D\} \right\} \quad (21)$$

where e stands for the most recent event, *i.e.*, an arrival A_k of the MSS session- k user or a departure D of any MSS session user. Note that the arrival A_k in this CTMDP denotes any user arrival for the MSS session- k , whereas we count a user request of the inactive session- k as A_k .

B. Action Space

For a state of s , let a_A , a_R and a_C be the action to accept an arriving user, the action to reject it and the action to continue, respectively, right after the event of user arrival/departure. When a user requesting the MSS session- k arrives, we have two possible actions we can take; a_A and a_R if there are no current users using the session and the bandwidth remains enough to serve the requested session. For other cases, we do not have alternative actions and the action is automatically decided as follows. If the bandwidth is fully occupied, a user requesting an

inactive session is rejected. When the requested MSS session is in service, the only action is to accept as long as the number of concurrent users with the session is less than N_k . For each state $s \in \mathcal{S}$ with $e = A_k$ ($k = 1, 2, \dots, K$), the action space is therefore represented as Equation (22), shown at the bottom of the page.

When a user of any MSS session leaves the system, the only action is to continue and the current state is just kept until next transition occurs. Thus, $\mathcal{A}_s = \{a_C\}$ if $e = D$. Thus when an arrival to session k occurs n_k does not include the user who requested the session till the decision is taken but when a departure from session k occurs it is already reflected in n_k since it occurs immediately before the decision epoch.

C. Decision Epochs

A decision is made at which point a user arrives to or leaves from the network over all sessions. Thus the next decision epoch follows whichever comes first between a user arrival or a user departure of any session. Since the inter-arrival time and the service time of a user are all exponentially distributed, each inter-decision time is also exponentially distributed but the inter-decision-epoch rate changes depending on the number of concurrent users over all sessions when a previous decision is made. If the action a_R or the action a_C is chosen at the current state s , the number of concurrent users is not changed for all the MSS sessions; otherwise (*i.e.*, the action a_A is chosen), the number of concurrent users for the requested MSS session is increased by 1. Thus, the inter-decision-epoch rate $\beta(s, a_s)$ for each state $s \in \mathcal{S}$ and action $a_s \in \mathcal{A}_s$ is driven by Equation (23), shown at the bottom of the page.

Further the probability distribution function of the inter-decision time to the next decision epoch is obtained by:

$$F(t|s, a_s) = 1 - e^{-\beta(s, a_s)t} \quad \text{for } t \geq 0. \quad (24)$$

D. Transition Probabilities

Let $q(s'|s, a_s)$ be the state transition probability that the system moves to state s' by a user arrival or departure, given that, at the previous decision epoch, the action a_s is taken at the state s . Note that s and s' denote the state before an action is taken. While a departure itself affects the number of concurrent users (being served), an arrival does not affect the number of concurrent users before an action is taken. For any

$$\mathcal{A}_s = \begin{cases} \{a_A, a_R\}, & \text{if } n_k = 0 \text{ and } \sum_{j=1}^K w_j \cdot \mathbf{1}_{\{n_j \geq 1\}} \leq C - w_k, \\ \{a_R\}, & \text{if } n_k = 0 \text{ and } \sum_{j=1}^K w_j \cdot \mathbf{1}_{\{n_j \geq 1\}} > C - w_k, \\ \{a_A\}, & \text{if } 1 \leq n_k < N_k, \\ \{a_R\}, & \text{if } n_k = N_k. \end{cases} \quad (22)$$

$$\beta(s, a_s) = \begin{cases} \lambda + \sum_{k=1}^K n_k \mu_k, & \text{if } a_s \in \{a_R, a_C\}, \\ \lambda + \sum_{k=1}^K n_k \mu_k + \mu_{k'}, & \text{if } a_s = a_A \text{ and } e = A_{k'}. \end{cases} \quad (23)$$

pair of $\{s = (\mathbf{n}, A_k), a_s = a_R\}$ or $\{s = (\mathbf{n}, D), a_s = a_C\}$, the state probability $q(s'|s, a_s)$ is obtained by Equation (25), shown at the bottom of the page, where \mathbf{I}_k is the row vector of order K that contains a one at the k^{th} component and zeros elsewhere. Here, $s' = (\mathbf{n}, A_{k'})$ and $s' = (\mathbf{n} - \mathbf{I}_{k'}, D)$ denote an arrival and a departure of an MSS session- k' user, respectively.

Meanwhile, admitting an arriving user request migrates the system state immediately, and so the pair $\{s = (\mathbf{n}, A_k), a_s = a_A\}$ implies that the number of concurrent user is changed to $\mathbf{n} + \mathbf{I}_k$ right after the action a_s is taken and the behavior thereafter is the same with the case $\{s = (\mathbf{n} + \mathbf{I}_k, A_k), a_s = a_C\}$. Thus, for any pair of $\{s = (\mathbf{n}, A_k), a_s = a_A\}$, we have

$$q(s'|s, a_s) = q(s' | (\mathbf{n} + \mathbf{I}_k, D), a_C). \quad (26)$$

Here, $q(s' | (\mathbf{n} + \mathbf{I}_k, D), a_C)$ is calculated by (25).

E. Reward Function

A substantial admission control decision is needed only when an inactive MSS session is requested. Therefore a reward is only acquired at such a point. As discussed in Section II-B2, if an inactive MSS session is requested by a user and it is accepted, then the network obtains the amount $\{R_k \mathbb{E}_\alpha(U_k) - C_k - P_k \mathbb{E}_\alpha(D_k)\}$ of reward; otherwise, the amount P_k of penalty is incurred. The reward for accepting the user for an inactive session is determined by considering all the revenue and the penalty generated by ensuing users arriving in the interval where the session is already active. Thus we do not need to consider the reward for users arriving after the session has already started service at decision epochs when it is accepted/rejected anymore. Therefore, we have the expected reward function $r(s, a)$ as in Equation (27), shown at the bottom of the page, for each state $s \in S$ and action $a \in \mathcal{A}_s$.

V. OPTIMAL STATIONARY DETERMINISTIC POLICY

A policy is said to be stationary if the decision is independent of the decision epoch σ_t (*i.e.*, $d_t = d$) which can be denoted by d_∞ . For our admission control problem, the action space \mathcal{A}_s is finite for all s since $\mathcal{A}_s \subset \{a_A, a_R, a_C\}$ so that there exists an optimal stationary deterministic policy as a result of Theorem 11.3.2 of [9]. Thus our problem can be reduced to finding the stationary deterministic decision policy d_∞ to maximize the expected infinite-horizon discounted reward given in (4). Letting $\pi' = (d_2, d_3, \dots)$ and $s_2 \in S$, it follows from (4) that $v_\alpha^\pi(s)$ can be defined as in Equation (28), shown at the bottom of the page. For the stationary decision rule d_∞ , $\pi = \pi' = d_\infty$, therefore the infinite-horizon discounted reward for the stationary policy can be presented as in Equation (29), shown at the bottom of the next page.

Now we do a uniformization [9] of our CTMDP, in which the CTMDP is ultimately converted to a discrete-time Markov chain, so that we can analyze the process easily. Let $c = \lambda + \zeta \left(\max_{1 \leq k \leq K} N_k \mu_k \right)$, where ζ is the maximum number of sessions that can be supported concurrently and it is obtained by

$$\zeta = \arg \max_j \left\{ \sum_{k=1}^j w_{[k]} \leq C \right\}, \quad (30)$$

for the array of bandwidth usage of the sessions $\mathbf{w} = (w_{[1]}, w_{[2]}, \dots, w_{[K]})$ in ascending order. Then the condition of *Assumption 11.5.1* of [9] stated below holds for our process:

$$\{1 - q(s|s, a)\} \beta(s, a) \leq c, \quad \text{for all } s \in S \text{ and } a \in \mathcal{A}_s. \quad (31)$$

Note that $q(s|s, a) = 0$ for all $s \in S$ in our CTMDP. A uniformization process of our CTMDP follows the following steps where we use components denoted by:

- 1) State space and action space: Set $\tilde{S} = S$ and $\tilde{\mathcal{A}}_s = \mathcal{A}_s$.

$$q(s'|s, a_s) = \begin{cases} \lambda p_{k'} / \left(\lambda + \sum_{k=1}^K n_k \mu_k \right), & \text{if } s' = (\mathbf{n}, A_{k'}), \\ n_{k'} \mu_{k'} / \left(\lambda + \sum_{k=1}^K n_k \mu_k \right), & \text{if } s' = (\mathbf{n} - \mathbf{I}_{k'}, D), \end{cases} \quad (25)$$

$$r(s, a) = \begin{cases} R_k \mathbb{E}_\alpha(U_k) - C_k - P_k \mathbb{E}_\alpha(D_k), & \text{if } n_k = 0, e = A_k \text{ and } a = a_A, \\ -P_k, & \text{if } n_k = 0, e = A_k \text{ and } a = a_R, \\ 0, & \text{otherwise.} \end{cases} \quad (27)$$

$$\begin{aligned} v_\alpha^\pi(s) &= r(s, d_1(s)) + \mathbb{E}_s^\pi \left\{ e^{-\alpha \sigma_1} v_\alpha^{\pi'}(s_2) \right\} \\ &= r(s, d_1(s)) + \sum_{j \in S} \int_0^\infty e^{-\alpha t} q(j|s, d_1(s)) F(dt|s, d_1(s)) v_\alpha^{\pi'}(j). \end{aligned} \quad (28)$$

- 2) Transition probabilities: In this step, we transform the CTMDP to a process observed at random times following an exponential time with rate c . Since c is bigger than any $\beta(s, a)$, the transformed process is observed more often than the original CTMDP. Thus fictitious transitions from each state to itself are introduced, and the transition probabilities determined depending on if the transition to another state occurs at each new observation point as follows:

$$\tilde{q}(j|s, a) = \begin{cases} 1 - \frac{\beta(s, a)}{c}, & \text{if } j = s, \\ \frac{q(j|s, a)\beta(s, a)}{c}, & \text{if } j \neq s. \end{cases} \quad (32)$$

- 3) Reward function: Since the transformed process in the above step has constant transition rate c , it can be directly transformed to a discrete-time discrete model in which the reward function is modified as below so that $\tilde{v}_\alpha^{d_\infty} = v_\alpha^{d_\infty}$ as proven in *Proposition 11.5.1* of [9]:

$$\tilde{r}(s, a) = r(s, a) \frac{\alpha + \beta(s, a)}{\alpha + c}. \quad (33)$$

Using this uniformization, the optimality equation in (29) would change to the form represented in Equation (34), shown at the bottom of the page.

For states with $\mathbf{n} = (n_1, n_2, \dots, n_K)$ and $e = D$, the only possible action is a_C and the state after taking the action remains at \mathbf{n} . Thus using Equation (34) we can derive the reward at time epoch $i + 1$ using reward at epoch i for state

$\{\mathbf{n}, D\}$ as in Equation (35), shown at the bottom of the page, where i is an iteration index and $\beta_0 = \{1 - q(\mathbf{n}, D|\mathbf{n}, D, a_C)\}$.

Since admitting a call migrates the system state immediately, also from Equation (34), we have the following optimality equations for states with $e = A_k$ ($1 \leq k \leq K$):

- 1) if $n_k = 0$ and $\sum_{i \neq k} w_i \cdot \mathbf{1}_{\{n_i \geq 1\}} \leq C - w_k$, then the action to take is a_R or a_A and the state after choosing one of them transit to \mathbf{n} or $\mathbf{n} + \mathbf{I}_k$, respectively, and so we have $v^{i+1}(\mathbf{n}, A_k)$ as given in Equation (36), shown at the bottom of the page.
- 2) if $n_k = 0$ and $\sum_{i \neq k} w_i \cdot \mathbf{1}_{\{n_i \geq 1\}} > C - w_k$, the arriving user should be rejected and the state remains at \mathbf{n} , and so we have

$$v^{i+1}(\mathbf{n}, A_k) = -P_k + v^i(\mathbf{n}, D), \quad (37)$$

- 3) if $1 \leq n_k < N_k$, then the arriving user is automatically accepted since the session is already in service, and so we have

$$v^{i+1}(\mathbf{n}, A_k) = v^i(\mathbf{n} + \mathbf{I}_k, D), \quad (38)$$

- 4) if $n_k = N_k$, then the arriving user is rejected due to the limitation on the number of concurrent user, and so we have

$$v^{i+1}(\mathbf{n}, A_k) = v^i(\mathbf{n}, D). \quad (39)$$

$$\begin{aligned} v_\alpha^{d_\infty}(s) &= r(s, d_1(s)) + \sum_{j \in S} \int_0^\infty e^{-\alpha t} q(j|s, d_\infty(s)) F(dt|s, d_\infty(s)) v_\alpha^{d_\infty}(j) \\ &= r(s, d_1(s)) + \sum_{j \in S} \left(\int_0^\infty e^{-\alpha t} \beta(s, d_\infty(s)) e^{-\beta(s, d_\infty(s))t} dt \right) q(j|s, d_\infty(s)) v_\alpha^{d_\infty}(j) \\ &= r(s, d_1(s)) + \frac{\beta(s, d_\infty(s))}{\alpha + \beta(s, d_\infty(s))} \sum_{j \in S} q(j|s, d_\infty(s)) v_\alpha^{d_\infty}(j). \end{aligned} \quad (29)$$

$$v(s) \triangleq v_\alpha^{d_\infty}(s) = \tilde{v}_\alpha^{d_\infty} = \max_{a \in \mathcal{A}_s} \left\{ \tilde{r}(s, a) + \frac{c}{\alpha + c} \sum_{j \in S} \tilde{q}(j|s, a) v(j) \right\}. \quad (34)$$

$$v^{i+1}(\mathbf{n}, D) = \frac{1}{\alpha + c} \left[\sum_{k=1}^K \{ \lambda p_k v^i(\mathbf{n}, A_k) + n_k \mu_k v^i(\mathbf{n} - \mathbf{I}_k, D) \} + (c - \beta_0) v^i(\mathbf{n}, D) \right], \quad (35)$$

$$v^{i+1}(\mathbf{n}, A_k) = \max \{ -P_k + v^i(\mathbf{n}, D), R_k \mathbb{E}_\alpha(U_k) - C_k - P_k \mathbb{E}_\alpha(D_k) + v^i(\mathbf{n} + \mathbf{I}_k, D) \}, \quad (36)$$

From Equation (36), we have the decision rule $d(\mathbf{n}, A_k)$ as given in Equation (40), shown at the bottom of the page, for the cases of $n_k = 0$ and $\sum_{i \neq k} w_i \cdot \mathbf{1}_{\{n_i \geq 1\}} \leq C - w_k$, where $\Delta v_k(\mathbf{n}, D) = v^i(\mathbf{n} + \mathbf{I}_k, D) - v^i(\mathbf{n}, D)$. Note that the decisions for other cases are already given in (22).

To solve the optimization problem given by (34), there are various algorithms such as the value iteration, policy iteration, and action elimination algorithms [9]. With a finite state space and a finite action space, we use the following value iteration algorithm (MSSAC-VIA) for our optimal MSS admission control derived from (35)–(39) to determine the stationary deterministic optimal policy and the corresponding expected infinite-horizon discounted reward:

MSSAC-VIA:

- 1) set $v^0(s) = 0$ for every state s . Specify $\epsilon > 0$, and set $i = 0$,
- 2) for each state s with $e = A_k$, compute $v^i(s)$ by (36)–(39),
- 3) for each state s with $e = D$, compute $v^i(s)$ by (35),
- 4) if $\|v^{i+1} - v^i\| < \epsilon$, go to step 5. Otherwise, increase i by 1 and return to step 2.
- 5) for each s with $n_k = 0$ and $\sum_{i \neq k} w_i \cdot \mathbf{1}_{\{n_i \geq 1\}} \leq C - w_k$, determine the stationary optimal policy $d_\infty(s)$ by (40) and stop,

where the norm function $\|\cdot\|$ is defined as $\|v\| = \max_{s \in \mathcal{S}} |v(s)|$.

The complexity of the value iteration algorithm increases exponentially with the number of sessions K . Which is one of the limitations of this algorithm. But this can be used as a benchmark to test heuristic algorithms with lower complexity. In the following proposition, we discuss a fundamental structure of the optimal policy derived above. The result in this proposition implies that there is a tendency to reject more user requests for inactive MSS sessions as the system is getting congested.

Proposition 1: For any k such that $n_k = 0$ and $\sum_{i \neq k} \mathbf{1}_{\{n_i \geq 1\}} < C$, $\Delta v_k(\mathbf{n}, D)$ is a non-increasing function on any n_i ($i \neq k$), i.e.,

$$v(\mathbf{n} + \mathbf{I}_k, D) - v(\mathbf{n}, D) \geq v(\mathbf{n} + \mathbf{I}_i + \mathbf{I}_k, D) - v(\mathbf{n} + \mathbf{I}_i, D), \quad (41)$$

so that if $d(\mathbf{n}, A_k) = a_R$ then $d(\mathbf{n} + \mathbf{I}_i, A_k) = a_R$ for any $i \neq k$.

Proof: Using a similar approach as in the proof of Theorem 1 in [22], we can prove that both $v(\mathbf{n}, D)$ and $v(\mathbf{n}, A_k)$ are concave and so (41) is verified by value iteration method. Hence we skip the proof of this part. By (40), $d(\mathbf{n}, A_k) = a_R$ implies

$$\Delta v_k(\mathbf{n}, D) < -R_k \mathbb{E}_\alpha(U_k) + \mathbf{C}_k + \{\mathbb{E}_\alpha(D_k) - 1\} P_k, \quad (42)$$

so that we can see the fact below:

$$\Delta v_k(\mathbf{n} + \mathbf{I}_i, D) \leq \Delta v_k(\mathbf{n}, D) < -R_k \mathbb{E}_\alpha(U_k) + \mathbf{C}_k + \{\mathbb{E}_\alpha(D_k) - 1\} P_k. \quad (43)$$

By (40) again, it is proved that $d(\mathbf{n} + \mathbf{I}_i, A_k) = a_R$. ■

VI. PERFORMANCE MEASURES

In this section, we construct an embedded Markov chain (EMC) induced by the Markov decision process in Section IV and its optimal stationary deterministic policy obtained in Section V, so as to derive the user blocking probability of each MSS session and the Average number of users served in the optimal policy. We first obtain the user blocking probability of the overall MSS system by considering the ratio of user arrival rate. The user blocking probability is defined as the ratio of the number of blocked users to total number of arriving users.

We consider the stochastic process $\{X_i : i = 1, 2, \dots\}$ that describes the state $s_i = (\mathbf{n}_i, e_i)$ at each decision epoch σ_i . Given the optimal stationary deterministic admission policy d , we observe that (i) each decision in a stationary deterministic policy is determined by the current state s_i and (ii) the future state s_{i+1} is determined by the current state s_i and the selected action a_i . Thus the stochastic process $\{X_i\}$ becomes an EMC with the decision epochs $\{\sigma_i\}$ as the embedded points and the one-step transition probabilities $q(s'|s, d(s))$ obtained in Section IV-D. The state space $\Omega_{s(0)}$ of the EMC consists of accessible states from the initial state $s(0) = (n_1(0), n_2(0), \dots, n_K(0))$ of the MSS system at time 0 by the optimal policy d . A state $s \in \mathcal{S}$ is said to be accessible from the initial state $s(0)$ if there exists a n -tuple $(s_1, \dots, s_{n-1}, s_n (= s))$ in \mathcal{S}^n such that

$$\prod_{i=1}^{n-1} q(s_{i+1}|s_i, d(s_i)) > 0, \quad (44)$$

for $s_1 = (s(0), e)$ with any possible event e that occurs when the state of the MSS system is $s(0)$. Then we can obtain the steady-state probabilities $\{\pi_s\}$, $s \in \mathcal{S}$, of being in state s from the balance equation and the normalization condition of the EMC.

A. Blocking Probability

Now we derive the user blocking probability \mathbb{P}_{B_k} of each MSS session, which is defined as the ratio of the number of blocked users to total number of users requesting the session, with the steady-state probabilities $\{\pi_s\}$. The observation on user arrival/departure in each state s makes the tracking of user blocking events for the MSS session- k available. Let Ω_{A_k} be

$$d(\mathbf{n}, A_k) = \begin{cases} a_A, & \text{if } \Delta v_k(\mathbf{n}, D) \geq -R_k \mathbb{E}_\alpha(U_k) + C_k + \{\mathbb{E}_\alpha(D_k) - 1\} P_k, \\ a_R, & \text{if } \Delta v_k(\mathbf{n}, D) < -R_k \mathbb{E}_\alpha(U_k) + C_k + \{\mathbb{E}_\alpha(D_k) - 1\} P_k, \end{cases} \quad (40)$$

TABLE II
EXPERIMENTAL SCENARIOS

	Scenario 1	Scenario 2	Scenario 3
Mean service time (min)	(5,5,5,5)	(5,10,15,20)	(5,10,15,20)
Reward	(20,20,20,20)	(20,20,20,20)	(20,20,24,30)

TABLE III
LOOKUP TABLE

busy channels	State s		Action a_s		
	n	e	Scn 1	Scn 2	Scn 3
0	(0,0,0,0)	A_1	a_A	a_A	a_A
		A_2	a_A	a_A	a_A
		A_3	a_A	a_A	a_A
		A_4	a_A	a_R	a_A
1	(n ₁ ,0,0,0)	A_2	a_A	a_A	a_A
		A_3	a_A	a_A	a_A
		A_4	a_A	a_R	a_A
		A_1	a_A	a_A	a_A
	(0,n ₂ ,0,0)	A_3	a_A	a_A	a_A
		A_4	a_A	a_R	a_A
		A_1	a_A	a_A	a_A
		A_2	a_A	a_A	a_A
	(0,0,n ₃ ,0)	A_4	a_A	a_R	a_A
		A_1	a_A	a_A	a_A
		A_2	a_A	a_A	a_A
		A_3	a_A	a_R	a_A
2	(n ₁ ,n ₂ ,0,0)	A_3	a_A	a_A	a_A
		A_4	a_A	a_R	a_A
		A_2	a_A	a_A	a_A
		A_4	a_A	a_R	a_R
	(n ₁ ,0,n ₃ ,0)	A_2	a_A	a_A	a_A
		A_4	a_A	a_R	a_R
		A_1	a_A	a_A	a_A
		A_3	a_A	a_R	a_R
	(n ₁ ,0,0,n ₄)	A_2	a_A	a_A	a_A
		A_3	a_A	a_R	a_R
		A_4	a_A	a_R	a_R
		A_1	a_A	a_A	a_A
	(0,n ₂ ,n ₃ ,0)	A_4	a_R	a_R	a_R
		A_1	a_A	a_A	a_A
		A_2	a_A	a_A	a_A
		A_3	a_A	a_R	a_R
	(0,n ₂ ,0,n ₄)	A_1	a_A	a_A	a_A
		A_4	a_R	a_R	a_R
		A_2	a_A	a_A	a_A
		A_3	a_A	a_R	a_R

TABLE IV
ACTIONS FOR THE STATE $s = (0, 0, n_3, n_4, A_2)$ IN TABLE III

		n_3									
		1	2	3	4	5	6	7	8	9	10
n_4	1	a_A	a_A	a_A	a_A	a_A	a_A	a_A	a_A	a_A	a_A
	2	a_A	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R
	3	a_A	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R
	4	a_A	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R
	5	a_A	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R
	6	a_A	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R
	7	a_A	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R
	8	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R
	9	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_B
	10	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R	a_R

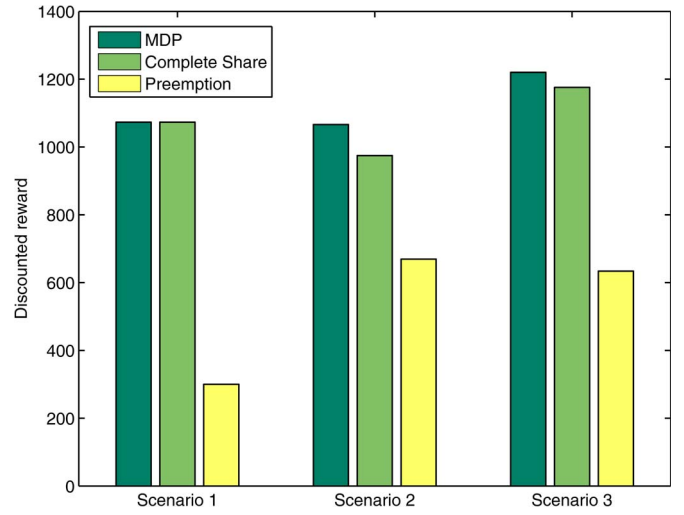


Fig. 3. Expected infinite-horizon discount reward.

the subspace of $\Omega_{s(0)}$, which consists of the states with $e = A_k$, and Ω_{B_k} be the further subspace of the states at which the user is blocked, *i.e.*, $\Omega_{A_k} = \{s \in \Omega : e = A_k\}$ and $\Omega_{B_k} = \{s \in \Omega_{A_k} : d(s) = a_R\}$. Then the user request blocking probability \mathbb{P}_{B_k} of MSS session- k is calculated by

$$\mathbb{P}_{B_k} = \frac{\sum_{s \in \Omega_{B_k}} \pi_s}{\sum_{s \in \Omega_{A_k}} \pi_s}. \quad (45)$$

Furthermore we derive the user blocking probability \mathbb{P}_B of the overall MSS system, which is defined as the ratio of the number of blocked users to total number of users arriving to the MSS system by

$$\mathbb{P}_B = \frac{\sum_{k=1}^K \left(\sum_{s \in \Omega_{B_k}} \pi_s \right)}{\sum_{s \in \Omega_A} \pi_s}, \quad (46)$$

where $\Omega_A = \bigcup_{k=1}^K \Omega_{A_k}$. By substituting (45) into (46), \mathbb{P}_B is obtained as the weighted mean reflecting the probability p_k of popularity for each MSS session as follows:

$$\mathbb{P}_B = \sum_{k=1}^K \frac{\left(\sum_{s \in \Omega_{A_k}} \pi_s \right) \mathbb{P}_{B_k}}{\sum_{s \in \Omega_A} \pi_s} = \sum_{k=1}^K p_k \mathbb{P}_{B_k}. \quad (47)$$

B. Average Number of Users Served

As we discussed earlier, each state label consists of two parts; the first is the number of users occupying each session and the second is the event that occurred just before that decision epoch. For the calculation of the average number of users we are interested in number of users at each state $s \in \mathcal{S}$. Let us define a function $SU_k(s)$ which returns the number of users in session k at State s . Then the expected number of users in Session k , $\mathbb{E}(n_k)$, is given by Equation (48).

$$\mathbb{E}(n_k) = \sum_{s \in \mathcal{S}} SU_k(s) * \pi_s \quad (48)$$

VII. NUMERICAL EXAMPLES

A. Actions in MDP-Based Admission Control and Performance Evaluation

In this section, we numerically investigate the MDP-based optimal admission control policy for MSS and evaluate the performance such as expected reward and user blocking probability. We set the parameters as $K = 4$, $C = 3$, $w_k = 1, \forall k = \{1, 2, 3, 4\}$, $\lambda = 1$, $\psi = 0.47$, $\delta = 0.001$ (*i.e.*, $(p_1, p_2, p_3, p_4) = (0.3521, 0.2542, 0.2101, 0.1836)$), $N_1 = N_2 = N_3 = N_4 = 30$ and $\alpha = 0.01$. Both setup cost and reject penalty are set to 10 for all MSS sessions. We consider three scenarios with

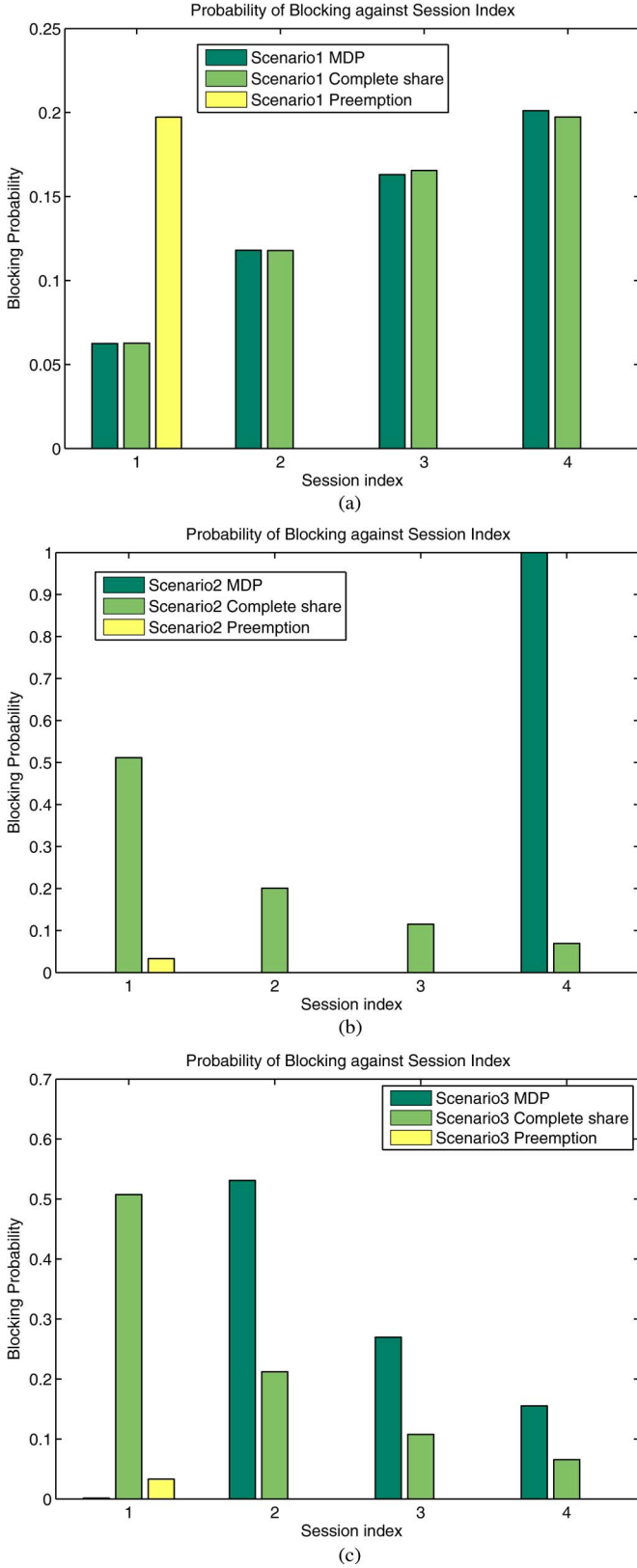


Fig. 4. User blocking probability. (a) Scenario 1. (b) Scenario 2. (c) Scenario 3.

different combinations of expected user service time μ_k and the reward R_k of each MSS session as Table II.

1) *Admission Control Actions:* In Table III, we present a lookup table listing actions at each user arrival for inactive

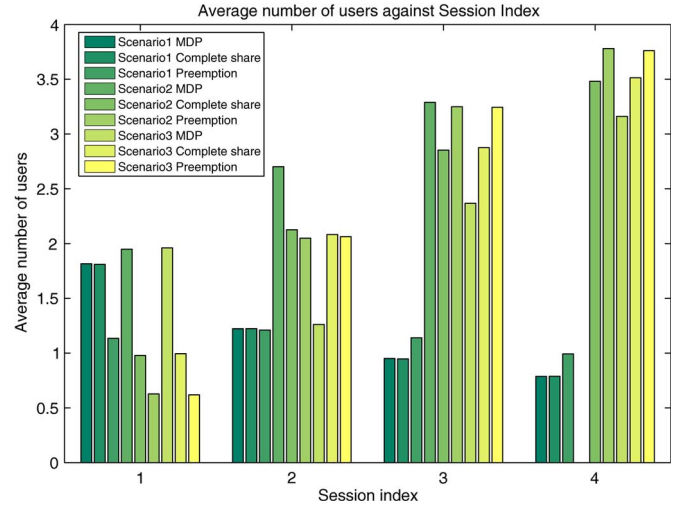


Fig. 5. Expected number of users in each session.

MSS sessions in the MDP-based optimal policy. A dot \cdot in this table indicates that the corresponding state can not be reached, so the situation to take an action at the stage never happens. For example if we consider the state $((0, 0, 0, n_4), A_k)$ in Scenario 2, it can be only reached if the previous state was $((0, 0, 0, 0), A_4)$ and the session is accepted, but in the optimal policy it is not accepted thus the state is not reached. A user for any inactive session is mostly accepted in Scenario 1, since the user service time for all MSS sessions is short so that the overall offered load is still affordable in the system. Only if MSS session-1 is inactive while 2 sessions are already in service (i.e., $s = (0, n_2, n_3, 0)$, $(0, n_2, 0, n_4)$, $(0, 0, n_3, n_4)$), the users requesting the other inactive session experience blocking. For $s = (0, n_2, n_3, 0)$ and $(0, n_2, 0, n_4)$, the remaining bandwidth to serve one more session is reserved for the MSS session-1 of the highest popularity, and prevented from the session-3 or the session-4. Meanwhile, when $s = (0, 0, n_3, n_4)$, an admission control decision for the MSS session-2 is affected by the number of users with ongoing sessions as listed in Table IV. Also, it is verified that if $d((0, 0, n_3, n_4), A_2) = a_R$ then $d((0, 0, n_3 + 1, n_4), A_2) = a_R$ and $d((0, 0, n_3, n_4 + 1), A_2) = a_R$ from Table IV. This coincides with the theoretical conclusion in Proposition 1.

In Scenario 2, the user service time is modified so as to make the offered load $\lambda p_k / \mu_k$ the same for all sessions. The session active time of unpopular MSS sessions (especially, session-4) becomes longer without enough contribution to the system in terms of revenue since the user arrival rate is comparatively too low. In this case, the MSS session-4 is completely blocked and it implies that we might as well not serve the session-4 to maximize the expected infinite-horizon discounted reward. In Scenario 3, users for MSS session-4 are allowed to enter the system by raising the revenue for unpopular sessions to compensate for the long bandwidth occupation.

2) *Performance Evaluation:* In Fig. 3, we investigate the expected infinite-horizon discounted reward $v_\alpha^\pi(\mathbf{n}(0))$ in (4) for the MDP-based optimal policy, given that the MSS system initiates without any users (i.e., $\mathbf{n}(0) = (0, 0, 0, 0) = \vec{0}$). Then we

TABLE V
SKEWNESS AND POPULARITY

ψ	0.5	1	1.5
(p_1, p_2, p_3, p_4)	(0.3591, 0.2539, 0.2073, 0.1796)	(0.48, 0.24, 0.16, 0.12)	(0.5984, 0.2116, 0.1152, 0.0748)

compare it to that of the complete sharing policy and Complete sharing with preemption policy. For the preemption policy we have made the sessions 3 and 4 high priority. Therefore when all the channels are occupied, if a user requests Session 3 or 4, we drop all the users currently using either session 1 or 2 and admit the user requesting the session with high priority. These high priority sessions although having larger or equal revenue to the others they are not that efficient in revenue generation as explained above. Therefore they earn less reward as shown in the Fig. 3. Note that if the action “ a_R ” widely appear in a MDP-based optimal policy then the gap on the expected reward becomes bigger between the two policies. In Scenario 1, the optimal policy has the same reward as the complete sharing policy since the users for inactive sessions are mostly accepted. In Scenario 2, however, the improvement increases up to 8.56% by preventing unpopular MSS session users from occupying bandwidth without enough contribution. Raising the revenue for unpopular sessions in Scenario 3 allows more MSS session-4 users, so that the improvement is reduced to 3.61%.

We also compare the user blocking probabilities of the MDP-based optimal policy, the complete sharing policy and the complete sharing policy with preemption. Note that the user blocking probability in both the complete sharing policy and the preemption policy is not affected by the revenue vector, and so it is the same for Scenario 2 and Scenario 3. In Fig. 4(a), the user blocking probabilities for relatively popular MSS sessions display tendencies to decrease in the optimal policy for Scenario 1. In Scenario 2 one can see that the Session 4 is completely blocked in the optimal policy. Since the optimal policy take account of the revenues in accepting users the blocking probabilities of the scenarios 2 and 3 are different for the optimal policy while they are being same for the other two schemes.

Finally we compare the average number of users in each session in Fig. 5. For Scenario 1 Optimal policy and Complete sharing policy have a common trend in the number of users while preemptive policy gives preference to Session 4 over Session 1. Both the Complete sharing and the Preemptive policy have higher number of users on Session 4 at Scenario 4 while the optimal policy has none. Optimal policy maintains higher numbers on other sessions.

B. Gain on the Expected Infinite-Horizon Discount Reward

In this section, we demonstrate the gain of the MDP-based optimal policy over the complete sharing policy from extensive experiments, to confirm how effectively the MDP-based optimal policy operates. For three scenarios in Section VII-A, the parameters λ and ψ vary by 0.5 from 1 to 2 and from 0.5 to 1.5, respectively, and the corresponding popularity vector of each skewness ψ is listed in Table V.

Fig. 6 displays the gain on the expected infinite-horizon discount reward. In Scenario 1, the gain is mostly small (below

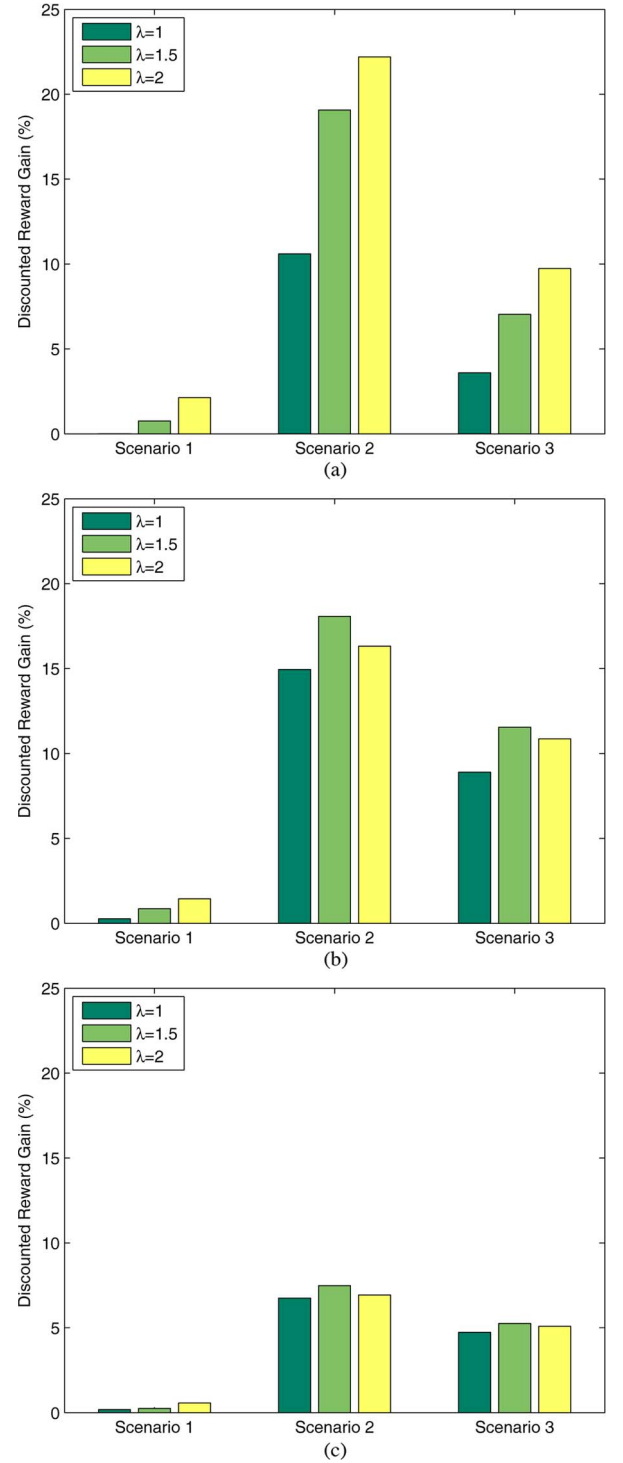


Fig. 6. Performance gain: expected infinite-horizon discount reward. (a) $\psi = 0.5$. (b) $\psi = 1.0$. (c) $\psi = 1.5$.

2.5%) for all the values of λ and ψ , but slightly increases as the value of λ is increased from 1 to 2. In Scenario 2 and 3, the gain ranges from 6.75% (when $\lambda = 1$ and $\psi = 1.5$) to 22.2% (when $\lambda = 2$ and $\psi = 0.5$). It implies that the MDP-based optimal

policy effectively controls the service of less popular sessions. The gain decreases without exception in Scenario 3 as raising the revenue for unpopular sessions cause a reduction in the probability of rejection of the unpopular sessions, and ranges from 3.59% (when $\lambda = 1$ and $\psi = 0.5$) to 11.55% (when $\lambda = 1.5$ and $\psi = 1.0$). Fig. 6 also shows that the gain is generally small when $\psi = 1.5$ in comparison with $\psi = 0.5$ and $\psi = 1.0$. This results just from that the arrival rates of unpopular sessions are too low so that the influence of admission control on these sessions becomes less significant.

C. Numerical Evaluation for Higher Number of Sessions

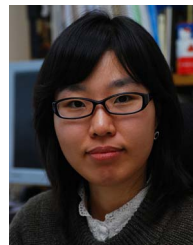
Having used the MDP, this algorithm is not immune to the curse of dimensionality. Therefore the time complexity increases with the number of available streams. One approach we plan to explore, in the future, for reducing this is the stochastic learning and optimization approach which was introduced by Cao [23].

VIII. CONCLUSION

In this paper, we addressed the MDP-based optimal admission control policy for MSS systems to maximize the expected infinite-horizon discounted reward, by taking into account the point-to-multipoint nature in multicast service. To determine the expected reward of accepting a user for an inactive MSS session, we analyzed the stochastic behavior of user arrival/departure in a session active time. The optimal policy has a structural tendency to reject more user requests for inactive sessions when the system is getting congested. We also mathematically derive the user blocking probability and the average number of users by analyzing an embedded Markov chain induced by the optimal policy. It is demonstrated that the proposed optimal policy remarkably improves the expected reward when there is a bias between the bandwidth occupation by each session service and the contribution to the system revenue.

REFERENCES

- [1] *Multimedia Broadcast/Multicast Service; Stage 1 (Release 6)*, 3GPP TS 22.146, V6.7.0, Mar. 2006.
- [2] *IEEE standard for Local and Metropolitan Area Networks-Part 16: Amendment for Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1*, IEEE Std. 802.16e-2005 and IEEE Std. 802.16-2004/Cor 1-2005, Nov. 8, 2005.
- [3] Yota Mobile WiMAX Home Page. [Online]. Available: <http://www.yota.ru/en/info/main/>
- [4] Y. Huang *et al.*, "Call admission control scheme for multicast service enabled cellular networks," in *Proc. 72nd VTC Fall*, 2010, pp. 1–5.
- [5] T. Lohmar and U. Horn, "Hybrid broadcast-unicast distribution of mobile tv over 3G networks," in *Proc. IEEE LCN*, 2006, pp. 850–851.
- [6] D. Catrein, J. Huschke, and U. Horn, "Analytic evaluation of a hybrid broad-unicast TV offering," in *Proc. IEEE VTC Spring*, 2008, pp. 2864–2868.
- [7] A. Alexiou and C. Bouras, "Multicast in UMTS: Evaluation and recommendations," *Wireless Commun. Mobile Comput.*, vol. 8, no. 4, pp. 463–481, May 2008.
- [8] J. M. Lee, H.-J. Park, S. G. Choi, and J. K. Choi, "Adaptive hybrid transmission mechanism for on-demand mobile IPTV over WiMAX," *IEEE Trans. Broadcast.*, vol. 55, no. 2, pp. 468–477, Jun. 2009.
- [9] M. L. Puterman, *Markov Decision Process: Discrete Stochastic Dynamic Programming*. Hoboken, NJ, USA: Wiley-Interscience, 2005.
- [10] C.-J. Ho and C.-T. Lea, "Finding better call admission policies in wireless networks," in *Proc. 48th IEEE VTC*, 1998, pp. 2135–2139.
- [11] J. Choi, T. Kwon, Y. Choi, and M. Naghshineh, "Call admission control for multimedia services in mobile cellular networks: A Markov decision approach," in *Proc. 5th IEEE ISCC*, 2000, pp. 594–599.
- [12] W. Ni, W. Li, and M. Alam, "Determination of optimal call admission control policy in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 1038–1044, Feb. 2009.
- [13] F. Yu and V. Krishnamurthy, "Optimal joint session admission control in integrated WLAN and CDMA cellular networks with vertical handoff," *IEEE Trans. Mobile Comput.*, vol. 6, no. 1, pp. 126–139, Jan. 2007.
- [14] E. Stevens-Navarro, Y. Lin, and V. W. S. Wong, "An MDP-Based vertical decision algorithm for heterogeneous wireless networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 2, pp. 1243–1254, Mar. 2008.
- [15] H. Chen, C.-C. Cheng, and H.-H. Yeh, "Guard-channel-based incremental and dynamic optimization on call admission control for next-generation QoS-aware heterogeneous systems," *IEEE Trans. Veh. Technol.*, vol. 57, no. 5, pp. 3064–3082, Sep. 2008.
- [16] J. Bühler and G. Wunder, "Traffic-aware optimization of heterogeneous access management," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1737–1747, Jun. 2010.
- [17] L. Berslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, "Web caching and Zipf-like distributions: Evidence and implications," in *Proc. IEEE INFOCOM*, Mar. 1999, pp. 126–134.
- [18] E. Veloso, V. Almeida, W. Meira, A. Bestavros, and J. Shudong, "A hierarchical characterization of a live streaming media workload," *IEEE/ACM Trans. Netw.*, vol. 14, no. 1, pp. 133–146, Feb. 2006.
- [19] K. Sripanidkulchai, B. Maggs, and H. Zhang, "An analysis of live streaming workloads on the Internet," in *Proc. 4th ACM SIGCOMM Conf. IMC*, New York, NY, USA, pp. 41–54.
- [20] C. Suh and J. Mo, "Resource allocation for multicast services in multicarrier wireless communications," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 27–31, Jan. 2008.
- [21] I.-R. Chen and T.-H. Hsi, "Performance analysis of admission control algorithms based on reward optimization for real-time multimedia servers," *Perform. Eval.*, vol. 33, no. 2, pp. 89–112, Jul. 1998.
- [22] S. A. Lippman, "Applying a new device in the optimization of exponential queueing systems," *Oper. Res.*, vol. 23, no. 4, pp. 687–710, Jul./Aug. 1975.
- [23] X. R. Cao, *Stochastic Learning and Optimization - A Sensitivity-Based Approach*. New York, NY, USA: Springer-Verlag, 2007.



Tae Ok Kim received the B.S., M.S., and Ph.D. degrees in mathematics from Korea University, Seoul, Korea, in 2003, 2006, and 2009, respectively. She worked as a Research Assistant Professor in Telecommunication Mathematics Research Center and as a Postdoctoral Researcher in the Department of Electrical and Computer Engineering at University of Manitoba. Her research interests include queueing theory, MAC performance analysis, and resource scheduling for wireless communication networks.



Chamara N. Devanarayana received the B.Sc. (Hons.) degree in electronics and telecommunications engineering from the University of Moratuwa, Sri Lanka, in 2008 and the M. Sc. degree in electrical and computer engineering from the University of Manitoba, Canada, in 2012. He is currently pursuing the Ph.D. degree at the University of Manitoba. His research interests include queueing theory, Markov decision process, and resource allocation for wireless communication networks.



Attahiru S. Alfa is a Professor of telecommunication systems at the Department of Electrical and Computer Engineering University of Manitoba, and also a SARCHI Chair Professor at the Department of Electrical, Electronic and Computer Engineering, University of Pretoria. His research covers, but is not limited to, the following areas: performance analysis and resource allocation in telecommunication systems, modeling of communication networks, queueing theory, optimization, analysis of cognitive radio networks, modeling and analysis of wireless sensor networks, developing efficient decoding algorithms for LDPC codes, channel modeling, traffic estimation for the Internet, and cross layer analysis. He also works in the application of queueing theory to other areas such as transportation systems, manufacturing systems and healthcare systems. He was NSERC Chair for tele-traffic from 2004 to 2012. He has carried out applied research for Nortel Networks, Bell-Northern Research, TRILabs (now TRTech), Bell Canada, Winnipeg Regional Health Authority, Motor-coach Industries, and several other industries. Dr. Alfa teaches courses on telecommunication networks, queueing theory, optimization, and network theory. He has authored a book, *Queueing Theory for Telecommunications: Discrete Time Modelling of a Single Node System* (Springer 2010). He is currently working on another book, *Applied Discrete Time Queueing Theory* (Springer 2015).



Bong Dae Choi received the B.S. and M.S. degrees in mathematics from Kyungpook University, and the Ph.D. degree in mathematics from Ohio State University, in 1980. He is a Professor at Research Institute for Information and Communication Technology, Korea University, Seoul, Korea. He had worked as a Professor at Kyungpook University (1972–1983), at KAIST (1983–1999), at Korea University (1999–2012), and at Sungkyunkwan University (2012–2015), Korea. He was a Director of Telecommunication Mathematics Research Center (2001–2012) at Korea University. He received best paper award from IEE in 2000 and he was awarded an academic award from Korean Mathematical Society in 1997 and Seoul Culture Prize in Science in 2001. He is a Fellow of Korea Academy of Science and Technology. He was an Editor of JOURNAL OF COMMUNICATIONS AND NETWORKS, an Associate Editor of *Queueing Systems* (Springer), and is an Associate Editor of *Telecommunication Systems* (Springer). His areas of interest include queueing theory and its applications to the communication systems. His recent interests are performance evaluation of IEEE 802.11, 15.4, 16e, power saving scheme, cognitive radio networks and IEEE 802.11p/1609.4 WAVE. He has published about 115 papers in referred journals and 21 patents.