Scalable Video Multicasting: A Stochastic Game Approach With Optimal Pricing

Chih-Yu Wang, *Member, IEEE*, Yan Chen, *Senior Member, IEEE*, Hung-Yu Wei, *Member, IEEE*, and K. J. Ray Liu, *Fellow, IEEE*

***Abstract*—Heterogeneous multimedia content delivery over wireless networks is an important yet challenging issue. One of the challenges is maintaining the quality of service due to scarce resources in wireless communications and heavy loadings from heterogeneous demands. A promising solution is combining multicasting and scalable video coding (SVC) techniques via cross-layer design, which has been shown to effectively enhance the quality of multimedia content delivery service in the literature. Nevertheless, most existing works on SVC multicasting system focus on the static scenarios, where a snapshot of user demands is given and remains the same. In addition, the economic value of the SVC multicasting system, which is an important issue from the service provider's perspective, has seldom been explored. In this paper, we study a subscription-based SVC multicasting system with stochastic user arrival and heterogeneous user preferences. A stochastic framework based on the multidimensional Markov decision process (M-MDP) is proposed to study the negative network externality existing in the proposed system and theoretically evaluate the corresponding system efficiency. A game-theoretic analysis is conducted to understand the rational demands from heterogeneous users under different subscription pricing schemes. By transforming the original dynamic and complex M-MDP revenue optimization problem into a traditional average-reward MDP problem, we show that the optimal pricing strategy that maximizes the expected revenue of the service provider can be derived efficiently. Moreover, the overall user's valuation on the system, e.g., social welfare, is maximized under such an optimal pricing strategy. Finally, the efficiency of the proposed solutions is evaluated through simulations.**

***Index Terms*—Markov decision processes, game theory, scalable video coding, multicasting, pricing.**

I. Introduction

With the development of multimedia compression and the advance of wireless networking techniques, multimedia content delivery over wireless networks becomes more and more popular, e.g., mobile video download/upload, live video streaming [1] and Internet Protocol TV (IPTV) [2]. One challenging issue in such wireless multimedia delivery systems is how to maintain the quality of service due to scarce resources in wireless networks and heavy loadings from heterogeneous demands.

Multicasting in wireless communications is a natural solution to the overloading problem in wireless live video streaming and IPTV. When multiple wireless users within a certain range request for the same multimedia content, the transmitter can simply broadcast one copy of the content to all receivers, which is called **multicasting**. The service provider may categorize users into several multicast groups according to their demands in contents, and then perform multicasting in delivering. However, challenges still exist in such a group-based multicast approach since each multicast group still requires resources to function. How the limited resources should be allocated to each group is an important issue. Moreover, heterogeneous users may use devices with different computational capabilities. Those devices with low computation capabilities may not be able to decode high-quality contents. Therefore, the service provider needs to deliver the same content in multiple qualities, such as in standard resolution (SD) and high resolution (HD), to satisfy these heterogeneous devices' requirements. This introduces serious redundancy in the delivery and therefore further aggravates the overloading issue.

A. SVC Multicasting Service

Scalable video coding (SVC) is a promising technique to resolve the content redundancy issue [3]. It provides a flexible design to encode the videos into a series of data streams, each of which represents a layer of the video. The base layer (layer 1) can be decoded independently without the information stored in other data streams. It also has a low decoding requirement in computation capability. Other layers are called enhancement layers, which contain extra information to reconstruct a higher quality video. A receiver may derive a higher quality video by decoding the base layer and subsequent enhancement layers. Therefore, the redundancy in delivery can be greatly reduced with such a technique. By combining multicasting and SVC techniques, the cross-layer design shows great potentials in enhancing the quality of multimedia delivery service. An example is illustrated in Fig. 1, where a video multicasting server is offering two SVC videos, each with two layers. Four users are already using the service, while subscriber 2∼4 request for

Manuscript received February 22, 2014; revised August 13, 2014; accepted December 8, 2014. Date of publication December 24, 2014; date of current version May 7, 2015. The work of H.-Y. Wei was supported by the Ministry of Science and Technology under Grants 102-2221-E-002-077-MY2 and 103-2221-E-002-086-MY3. The associate editor coordinating the review of this paper and approving it for publication was G. Xing.

C.-Y. Wang is with the Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617, Taiwan, and also with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: tomkywang@gmail.com).

Y. Chen and K. J. R. Liu are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: yan@umd.edu; kjrliu@umd.edu).

H.-Y. Wei is with the Graduate Institute of Communication Engineering and Department of Electrical Engineering, National Taiwan University, Taipei 10617, Taiwan (e-mail: hywei@cc.ee.ntu.edu.tw).

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/TWC.2014.2385773

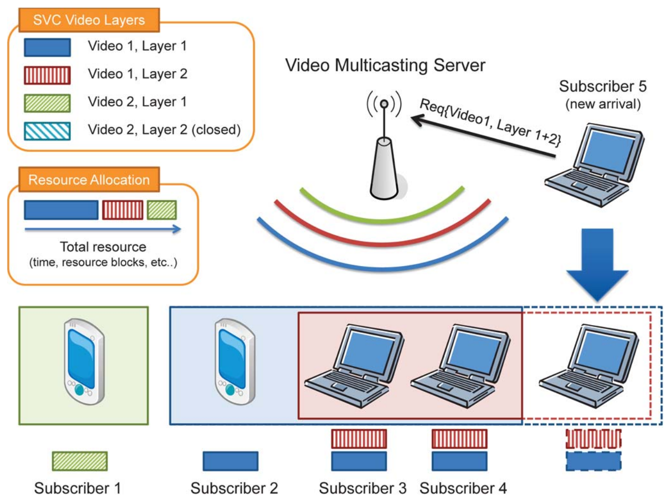


Fig. 1. A SVC multicasting platform offering two 2-layer videos.

video 1 and subscriber 1 requests for video 2. Three multicasting groups, each for a specific layer of a video, are formed according to their requests for the video and layers. Notice that the transmission of video 2's layer 2 is suspended because no users are requesting it. Nevertheless, when subscriber 5 arrives the system, she sends a request to the video server for the video 1 with both layers. If the video server accepts the request, subscriber 5 will join the multicasting groups for video 1's layer 1 and layer 2. The resource allocation of each multicasting group will then be determined according to the channel conditions and demands of users.

B. Related Works

The cross-layer design of SVC multicasting system has been discussed in the literature. A popular approach is the utility-based approach [4], where users are assumed to have some utilities if they receive and decode the demanded videos correctly. Under such an approach, the objective of the system is maximizing all users' utilities given the current demands and the channel conditions. Generally, given a snapshot of the multicasting system, the multiple multicast group resource allocation problem can be reduced to a 0–1 Knapsack problem if the selectable resource allocation pattern is not continuous, e.g., finite set of modulation and coding schemes. Since the 0–1 Knapsack problem is NP-hard, approximated algorithms are more desired to efficiently solve the resource allocation problem. Several approximated algorithms with different objectives for the snapshot optimization have been proposed in the literature. A straightforward greedy algorithm is proposed in [5] while a pseudo-polynomial algorithm for minimizing the total energy consumption is proposed in [6]. In [2], the authors proposed an envelope-based approximated algorithm and gave a tight error bound. Dynamic programming approaches, which make use of the sequential utility increment structure in SVC technique, are introduced in [1], [7], and [8]. This problem becomes more challenging when it comes to cognitive radio [9]. In [10] the authors illustrate how to jointly consider the availability of channels with video multicasting. A similar opportunistic multicasting approach is shown in [11] to enhance the quality of service in WiMAX system. Pricing is a powerful tool to regulate the subscribers' actions in a streaming service [12]. In [13], a pricing approach is proposed to transform the utility into the immediate revenue with different priorities in layers.

However, most existing SVC multicasting works focus on the static case, i.e., considering the current snapshot scenario where the demands from users are given. In practice, the SVC multicasting system will be highly dynamic in a long run due to the changes of channel conditions, user demands, and the reserved resource for the service. In a long-term SVC multicasting service, the resource allocation should not only address the current demand but also compensate the loss from previous transmission failures. Scheduling, which is naturally considered when we consider an OFDMA system with time-frequency multiplexing [14], need to be expanded from single frame scale to multiple frame scale. One workaround is using an average window on the objective function to track the utility of each users in the long-term SVC multicasting service [15], [16]. When a user experiences transmission failures in previous slots, her degraded utility will be reflected in the objective function, which leads to a higher priority in the resource allocation mechanism. The deadline of each packet can be further addressed by adding the corresponding weights to the utility of the users in the objective function [15]. Additionally, different scheduling strategies like Max-Sum, Max-Product, and Round-Robin, can be applied in such an approach [16]. Despite the average window, a gradient-based objective function can also be applied to track the changes in utility and make the corresponding compensation in resource allocation to users [17]. Nevertheless, most existing works addressed the compensation in resource allocation to previous transmission failures, while the potential loss in the future transmissions is not of their concerns. Additionally, they assume that the number of users and their demands remain static in the process of SVC multicasting service. This could be a problem when we consider a dynamic system with users arrival and departure, which is commonly observed in real world systems.

A stochastic analysis and design on SVC multicasting system is more desirable in this case, since it helps us to not only consider the past history of the system performance but also estimate the future outcome of the system. Through the analysis, we can predict the system performance in a dynamic system where users arrive and depart stochastically. The analysis can also help us to estimate the expected long-term utility of users if they subscribe to certain services. In the literature, there are works investigating the policy design of stochastic multicasting system in wired networks [18]. However, to the best of our knowledge, there is no existing work on the stochastic analysis on the SVC multicasting system over wireless networks.

In addition, from the service provider's perspective, the economic value of a SVC multicasting system may be the most important factor. A commercial service provider will provide such a service only when it is beneficial in profits [12]. Taking into account such a factor, we discuss a subscription-based economic model with different pricing schemes, where the selfish nature of rational users is involved. Researchers have discovered that the rational behaviors of users should be seriously considered in designing any system with users applying actions and making decisions [19], [20]. A centralized-control process designed to optimize the overall system performance may fail in a user-oriented system since the rational users focus on their own utilities instead of overall system performance and therefore may disobey the instructions from the control process. The selfish behaviors of users may eventually degrade the system efficiency or service provider's revenue [12]. Some incentive designs may be required to regulate the behaviors of users in order to improve the system efficiency [21]. By considering rational users' selfish nature, we propose to use game theory to analyze how users react under different pricing schemes.

Game theory has been applied in the multimedia multicast transmission systems for enhancing the service quality. Peer-to-peer cooperative communication, for instance, is proposed in [22] to enhance the transmission reliability of multi-view videos. In their proposed system, user who experience packet loss may recover the packet through the redundant coding information (RCI) contained in packets for other users. A game-theoretic monetary-based process is implemented to promote the cooperation among the users. Another relay-based cooperation scheme is proposed in [23], where users form multiple coalitions for two-hop multicast transmission. The head of each coalition takes the responsibility to transmit the contents from central server through multicast to all members of the coalition. A coalition game is formulated here for analyzing the system, and a distributed coalition formulation algorithm is proposed to find the optimal coalition. For non-cooperation system, the competition for resource is discussed in [24], where IEEE 802.11-based stations compete for TXOP for their own video transmissions. Given that each station may have different channel quality and video requirement, it is necessary to demand all the stations to reveal their true requirements in order to determine the efficient TXOP allocation. By inducing the additional transmission cost to each station, a VCG-based mechanism is proposed to ensure 1) all stations truthfully reveal their information and requirements, and 2) the optimal TXOP allocation is guaranteed. Nevertheless, to best of our knowledge, none of these game-theoretic works address the SVC multicasting system. Additionally, a stochastic analysis on the demands from rational users in long-term multimedia transmission system is also lacking in this area.

C. Contributions

In light of these concerns, we study a dynamic SVC multicasting system with stochastic user arrival and heterogeneous user demand in this paper. Specifically, we propose a Multi-dimensional Markov Decision Process (M-MDP) framework to analyze the optimal load balancing and economic efficient policies for the dynamic SVC multicasting system. The M-MDP framework is a stochastic extension [25], [26] to the Chinese restaurant game [27], [28], in which the authors investigate how the negative network externality and social learning influence the decisions of rational users. Network externality represents the effect of number of users who make the same decision/action on the utilities of these users. The effect could be positive or negative, as the utility function is an increasing or decreasing function of user amounts, respectively. The network externality effect also exists in our framework, as the amount of demands on each video influences the resource allocation and thus differentiates the quality of services to the users requesting different videos.

In this paper, we introduce a subscription system to help the service provider to regulate the subscription requests from the heterogeneous users. We consider two subscription pricing schemes: one-time charge and per-slot charge scheme. We prove that both pricing schemes achieve the same optimal revenue. In addition, we prove that the complex M-MDP optimization problem can be reduced to the traditional average-reward MDP problem when the optimal pricing strategy is adopted, and the solution can be effectively derived through the proposed algorithms.

In summary, our main contributions are shown as follows:

1) We develop a stochastic framework to analyze the resource allocation in a SVC multicasting system with heterogeneous user demands. By considering the stochastic user arrival, such a framework is more general than the existing snapshot-based approaches in the literature.
2) We propose a game-theoretic model to analyze the behaviors of heterogeneous users. We study how rational and intelligent users submit their demands, i.e., subscriptions, under two pricing schemes: one-time charge scheme and per-slot charge scheme, and derive the equilibrium conditions of the game. To the best of our knowledge, this paper is the first work bringing game theoretic analysis to the SVC multicast system.
3) We theoretically evaluate the economic value of the SVC multicasting system. Specifically, we investigate the revenue-maximized policy and pricing strategies in both one-time charge and per-slot charge schemes, which are hard to derive due to the coupling effects in both terms. By proving that the maximum revenue under both schemes is equivalent under all policies, we transform the complex M-MDP revenue maximization problem to a traditional average-reward MDP problem, due to which we can derive the optimal policy and pricing strategies in an efficient way. Both theory and simulation results confirm that the derived solution not only maximizes the expected revenue but also optimizes the social welfare.

Our previous work [29] includes the basic system model, per-slot charge pricing scheme, and a naive algorithm to identify the optimal pricing strategy. The per-slot charge scheme is not practical in real system. The convergence of the naive algorithm is not guaranteed neither. In this paper, we propose one-time charge pricing scheme as a practical and efficient pricing scheme for the subscription system. Based on the new pricing scheme, we transform the original M-MDP problem into an average-reward MDP problem, and therefore propose a novel algorithm which identifies the optimal pricing strategy in a significantly higher speed. The convergence of the proposed algorithm to the optimal strategy is also guaranteed. Extensive simulations with real data are also provided in this paper for evaluating the performance of the proposed algorithm and both pricing schemes, comparing with two other traditional pricing schemes.

The rest of the paper is organized as follows. In Section II, we describe a general SVC multicasting system where users have heterogeneous preferences and computation capabilities. In Section III, we formulate the system as a game-theoretic M-MDP framework for investigating the rational behaviors of users under such a subscription-based SVC multicasting system. Then, the equilibrium conditions of the system are derived in Section IV, where the expected subscription requests from users under certain pricing scheme strategies are analyzed. In Section V, the optimal pricing strategies that maximize the revenue of the service provider under both pricing schemes are derived. By applying optimal pricing strategies, we reduce the complex M-MDP problem into a traditional average MDP problem and derive the revenue-maximized subscription regulation policy in Section VI. Finally, simulation results are discussed in Section VII. We draw conclusions in Section VIII.

II. System Model

We consider a video multicasting service with one video server and multiple potential users who arrive and depart stochastically. The service is offering multiple choices of SVC-encoded videos. Given the limited resource (time, bandwidth, etc.), the service provider needs to determine the video server's resource allocation given the current demands for videos. In addition, the provider also determines the prices of video multicasting services, which is dynamically adjusted.

Users make request for the service when they arrive the system. Their requests are based on their preferences, computation capabilities, and the price of the service. The objective of the service provider is maximizing her revenue, while the users aim to seek for best experience, i.e., highest utility, in the video multicasting service.

A. Video Server

Let us consider a video server which is capable of serving at most N subscribers.[1] The server provides multiple videos denoted by $\mathcal{J} = \{1, 2, \ldots, J\}$. Each video is encoded by SVC into a video stream. An encoded video steam contains K layers. The server transmits the layers of each video periodically on the same channel, and all packets of layers are transmitted with the same interval.

A subscriber can decode a video if she receives at least the base layer (layer 1), while the quality of the video will be enhanced if she successfully decodes more subsequent layers. Let layer k be the k-th layer of the video, we assume that a subscriber may decode the video up to layer k only when all layer $1 \sim k$ are successfully received. In Fig. 1, we illustrate a SVC multicasting server with 5 subscribers, while the server is offering $J = 2$ videos, each with $K = 2$ layers.

[1]Notice that here we impose a constraint on the number of subscribers to simplify the analysis. It is also feasible to apply the subscriber limitations based on the amount of consumed/required resource. In such a case, different boundary conditions will apply in the proposed model, while the main structure and results of the proposed model will be unaffected.

The reception of the layer is determined by two factors: the supported modulation and coding scheme (MCS) at the subscriber side and the MCS applied on the layer at the video server side. On the subscriber side, let $g_i \in \mathcal{G}$ be the maximum MCS supported by subscriber i, where $\mathcal{G}$ is the universe set of the MCSs. In this paper, we assume that the channel quality g_i is a random variable with a common probability density function $f(g)$. We consider a slotted time wireless system where the slot time T_s is equal to the coherence time of the channel quality. Additionally, the channel distribution we considered is the prior estimation on the channel quality of average users within the service area. We believe that it is reasonable to assume that the service provider has collected such information in advance since the service provider must predict the potential service quality in advance so they can estimate the profitability of the service before deploying it.

On the video server side, the video server needs to determine the MCS applied on each layer of each video. Let $\underline{g_{j,k}}$ be the MCS applied on layer k of video j. When the layer k of video j's data stream is transmitted with the MCS $\underline{g_{j,k}}$, all users with channel quality $g_i \geq \underline{g_{j,k}}$ can receive this stream.

The applied MCS $\underline{g_{j,k}}$ determines the required resource (transmission time) to transmit the layer. We denote $C_{j,k}(\underline{g_{j,k}})$ as the required resource to transmit layer k of video j to users with MCS $\underline{g_{j,k}}$. The $C_{j,k}(\underline{g_{j,k}})$ should be a decreasing function of $\underline{g_{j,k}}$ since a layer can be transmitted in a shorter time if a higher-level MCS with higher throughput is applied. The video server may choose to stop transmitting layer k. We define $C_{j,k}(\underline{g^c}) \equiv 0$. The $\underline{g^c}$ represents the case that the transmission of this layer is disabled and therefore no users can receive this layer. Finally, let C^{total} be the total available resource, which is the total transmission time in our system. As the overall resource is limited to C^{total}, we have a firm constraint that $C^{total} \geq \sum_{j,k} C_{j,k}(\underline{g_{j,k}})$.

Notice that this formulation can be used to capture the rate and delay constraints of specific video streams. Let the resource C be the transmission time in the system in a period. In such a setting, the required transmission time for layer k of video j in each period should be the number of bits in each period divided by the supported transmission rate under the applied MCS $\underline{g_{j,k}}$. When the required transmission time is larger than the delay requirement of video j, the video quality will becomes unacceptable. In such a case, the required resource can be defined as infinity at these channel qualities to indicate that the targeting video quality can never meet under these channel qualities. Specifically, let $\bar{g}$ be the minimum required MCS for layer k of video j to meet the delay constraint, we have $C_{j,k}(\underline{g'_{j,k}}) = \infty$ if $\underline{g'_{j,k}} < \bar{g}$. Through this formulation, we capture the rate and delay constraints in our framework.

Notice that the resource should be dynamically allocated according to the current realization of users' channel qualities in order to maximize the delivery efficiency. Let $\mathbf{s} = \{n^s_{j,k}\}$ be the current **system state**, where $n^s_{j,k}$ be the number of users requesting video j's layer k stream. The state $\mathbf{s}$ represents the current loading of the video server. Then, we denote the dynamic resource allocation rule as $\mathbb{D}(\mathbf{s}) = \{\underline{g_{j,k}} | \forall j, k\}$. The allocation rule takes the current system state $\mathbf{s}$ as input and outputs the corresponding MCS for each layer. The $\mathbb{D}(\mathbf{s})$ may be implemented with different objectives, such as efficiency maximization or fairness constraints. In a utility-based system [2], [5], a common and reasonable choice of $\mathbb{D}$ is the overall utility maximization, that is,

$$\mathbb{D}(\mathbf{s}) = \arg\max_{\underline{g_{j,k}}} \mathbb{E}\left[u_i\left(\{\underline{g_{j,k}}\}\right)\right] \tag{1}$$

under the resource constraint $C^{total} \geq \sum_{j,k} C_{j,k}(\underline{g_{j,k}})$, where u_i is the utility of user i. In a wireless system with finite choices of the lowest channel quality, e.g., limited choices of MCS in WiMAX, this problem has been shown to be NP-hard [2]. Therefore, heuristic approaches are required and can be found in the literature [1], [2], [5], [7], [8].

B. User Valuation

In our system, users with different preferences arrive stochastically. In general, there are some users who have strong preference on a certain type of videos, e.g., sport fans always subscribe to the sport news channels. On the other hand, there are also some users that may not have a strong preference on the type of videos. They can enjoy all videos they successfully receive and decode.

Users with similar preferences usually have similar valuations on video content and quality. However, their devices may have different capabilities in receiving and decoding the videos. For instance, users who prefer the sport news may subscribe to the same channels, but some of them are only equipped with mobile phones. With such a limited capability device, only base layer can be decoded and displayed correctly. Therefore, users' abilities to have better video quality is limited.

We model all the aforementioned properties with the following notations. Users are categorized into types, which is denoted by $t \in \{1,2,..,T\} = \mathcal{T}$. A type $t = (\mathcal{J}^t, k^t)$ user prefers videos $j^t \in \mathcal{J}^t$ and is equipped with a device capable of decoding the SVC-encoded video up to layer k^t. We denote the valuation function on video $j \in \mathcal{J}$ with maximum decoded layer k as $v_j(k)$. The valuation function represents the amount of utility the user gains if she decodes the video correctly in this slot. Then, a type t user's valuation on video j with maximum consecutively received layer k is denoted as

$$v^t(j,k) = \begin{cases} v_j(k), & j \in \mathcal{J}^t, k \leq k^t; \\ v_j(k^t), & j \in \mathcal{J}^t, k > k^t; \\ 0, & \text{else.} \end{cases} \tag{2}$$

Note that a user has positive valuations on the service only if she receives and decodes her preferred video $j^t \in \mathcal{J}^t$ successfully. Finally, we assume users with different types arrive independently. We describe the arrival process of type t users with a Poisson process with the average arrival rate $\bar{\lambda}^t$. Notice that the popularity of a video can be characterized by the arrival rate of the corresponding user type which prefers such a video. A higher arrival rate of a specific type of users means a higher popularity of the video these users preferred.

Additionally, a user may leave the system at any time. We assume the time period of a user staying in the system is an exponential process with the average departure rate $\bar{\mu}$.

C. Payment System

We consider a subscription-based payment system as the revenue source of the service provider. We assume that the video service is private and all transmissions are encrypted. Therefore, users who enter the system should subscribe to one of the videos in order to correctly decrypt the corresponding data streams. A subscription contains two terms: the subscribed video j and the desired maximum layer k. When a subscription is accepted by the video server, the decryption keys of video j's layer $1 \sim k$ streams are delivered to the user. However, the reception of these streams is not guaranteed due to the randomness of the channel quality g_i, which is characterized by $f(g_i)$. We assume the channel qualities of all users are independent from each other and among time slots.

The price for a subscription, which is determined by the service provider, should be properly chosen in order to maximize its expected revenue. We model it as a function of the state $\mathbf{s} = (n^s_{1,1}, n^s_{1,2}, \ldots, n^s_{J,K})$. In this paper, two pricing schemes are considered.

1) One-time charge: when a user arrives the system in state $\mathbf{s}$, a payment $P^e_{j,k}(\mathbf{s})$ is charged as soon as the user's subscription (j,k) is accepted, and no further payments are required. In this pricing scheme, the $\mathbf{s}$ denotes the system state when the user arrives at the system.
2) Per-slot charge: At each time slot, as long as the user stays in the system with a valid subscription, she is continuously charged with an payment of $P_{j,k}(\mathbf{s})$, where state $\mathbf{s}$ is the system state at that time slot. Notice that the price depends on the current state $\mathbf{s}$, which may change over time.

Additionally, the subscription is canceled immediately when the user leaves the system, which is described by the departure process we mentioned in previous section.

III. Game Theoretic Formulation

In our model, we assume that users are rational and thus naturally selfish. Therefore, we need to consider users' selfish behaviors when evaluating and designing the pricing strategies, and game theory is powerful tool that analyzes the strategic interactions among selfish decision makers [20].

We consider a subscription game where the players are the subscribers and the service provider. The service provider determines the service price, $\{P^e_{j,k}(\mathbf{s})\}$ for one-time charge scheme or $\{P_{j,k}(\mathbf{s})\}$ for per-slot charge scheme, at the beginning of the game. Subscribers then submit their requests according to the prices when they arrive the system. Notice that each subscriber may arrive at different time given that this is a stochastic system.

The objective of the service provider is to maximize the expected revenue in the system. Notice that users are rational and desire to maximize their own utility given the prices determined by the service provider. As a result, the service provider needs to carefully determine the pricing strategy by taking into account the response of selfish users.

A user's objective is to maximize her own utility by choosing the best subscription. As described in Section II, users arrive stochastically. When a type t user arrives, she determines whether to subscribe to a specific video at certain layers. To subscribe, the user sends a subscription request to the server. She will receive the corresponding data streams if the server accepts the request. Note that since the system state and the channel quality are changing over time, the service quality is dynamic.

Although users only make decisions (choose the video/ layers) at the time they arrive. The key point for them to choose proper subscriptions is to calculate the expected utility given their initial states and choices on the video/layers. Nevertheless, the system we consider involves multiple subscribers, who may arrive and depart at different time. The choices of later subscribers on the video/layers will influence the system loading, change the resource allocation, and therefore affect the video reception quality experienced by early subscribers. Additionally, their decisions not only change the server loading at that moment they arrive but also the transition of the system state. It becomes necessary for a rational subscriber to predict how other subscribers in the future will choose the video/layers in order to make correct estimations on the expected utility.

The expected utility of the user with subscription (j,k) is conditioned on the system state $\mathbf{s}$ when she arrives. Let the system state at time slot l be $\mathbf{s}^l$, and let the realized channel quality be g_i^l. A type t user with a valid subscription (j,k) has an immediate valuation on the service, $v^t(j,\overline{k}^l)$, where $\overline{k}^l(g_i^l) \le k$ is the maximum successfully decoded layer at current time slot l. Notice that the reception of layer k' depends on the realized channel quality g_i^l and the MCS $\underline{g_{j,k'}^l}$ applied on the stream at that slot. Additionally, SVC can only be incrementally decoded. Therefore, we have

$$\overline{k}^l\left(g_i^l\right) = \arg \max_{k' \le k, \forall k'' < k', \underline{g_{j,k''}^l} \le g_i^l} k' \tag{3}$$

In addition, there is a cost of using the service, which is the charged payment determined by the pricing scheme. Given the state and costs, the expected utility of a type t user under subscription (j,k) is

$$\begin{aligned} \mathbb{E}\left[u^t(j,k)|\mathbf{s}\right] &= -c(\mathbf{s},j,k,0) \\ &+ \mathbb{E}\left[\sum_{l=l_a}^{l_d}\left(\sum_{g_i^l \in \mathcal{G}} v^t\left(j,\overline{k}^l\left(g_i^l\right)f\left(g_i^l\right) - c(\mathbf{s}^l,j,k,1)\right)\right|\mathbf{s}^{l_a} = \mathbf{s}\right], \end{aligned} \tag{4}$$

where $\overline{k}^l$ is given by (3), $c(\mathbf{s},j,k,e)$ is a common pricing function depending on the pricing scheme, $\mathbf{s}$ is the state when the user arrives the system, and l_a and l_d are the user's arrival and departure time indices, respectively. Specifically, $c(\mathbf{s},j,k,0)$ is the entrance fee to request a subscription (j,k) before using the service, which will be zero under the per-slot charge scheme. The $c(\mathbf{s},j,k,1)$ is the per-slot charge when the user is in the system, which will be zero under the one-time charge scheme. A rational user will choose the subscription that maximizes (4) when arriving at the system.

IV. Equilibrium Conditions

Nash equilibrium is a solution concept for predicting the outcomes of a game with the assumption that all players are fully-rational. Nash equilibrium describes an action profile, where each player's action is the best response to other players' actions in the profile. Since all players apply their best responses, none of them has the incentive to deviate from their actions described in the profile.

The Nash equilibrium of the proposed video subscription game can be analyzed through the following procedures. We first model the selfish users' behaviors through a multi-dimensional Markov decision process by fixing the pricing function $c(\cdot)$. The steady state and the expected utilities can then be calculated, and therefore the users' equilibrium conditions can be derived. With the equilibrium conditions of users, we then derive the equilibrium conditions for the service provider to maximize the revenue.

A. Users' Behavior Modeling Using Multi-Dimensional Markov Decision Process

The video subscription game, when the pricing function $c(\cdot)$ is given, can be formulated as a multi-dimensional Markov decision process (M-MDP) [26]. A Markov decision process describes a stochastic system where the transition between states is partially or fully determined by a decision maker [30]. The objective of the decision maker, the subscriber in our model, is to maximize her expected reward. In a traditional MDP, there is only one decision maker and the optimal solution that maximizes the unique expected reward can be found using dynamic programming [30]. However, since there are multiple decision makers with their own utility functions in our game, the traditional MDP cannot be directly applied here. On the contrary, in M-MDP we have multiple decision makers, each with different sets of actions to choose from. The actions applied by one decision maker may influence the expected utility of other decision makers, and therefore alter their willingness to choose certain actions.

In the proposed M-MDP framework, we consider a discrete-time Markov system where each time slot has a duration of T_s. The arrival and departure probability of each type t of subscribers can be approximate to $\lambda^t = \bar{\lambda}^t T_s$ and $\mu = \bar{\mu} T_s$, respectively, when T_s is sufficient small and therefore it is very unlikely to have more than two subscribers arrive at the same time slot. The system state is $\mathbf{s} = (n_{j,k}^s | j \in \{1 \dots J\}, k \in \{1 \dots K\}) \in \mathcal{S}$, where $n_{j,k}^s$ denotes the number of users subscribing video j with maximum subscribed layer k. The server can serve up to N users, therefore we have the boundary constraints $\sum_{j,k} n_{j,k}^s \le N$ on the states.

The action of a user is the subscription request $a = (j,k)$. Different types of users may have different action space due to the limitation in computation capability and their preferences. The action space of a type t user is $\mathcal{A}^t = \mathcal{J}^t \times \mathcal{K}^t \bigcup \{(0,0)\}$, where (0,0) represents that she does not subscribe to any video and leaves the system immediately. Note that users will not subscribe the unpreferred videos since their valuations on those videos are zero. We denote $V(j,k) \equiv v_j(k)$ to describe any subscriber's valuation on the video if she indeed submits a request (j,k) and receive the video at this slot. Therefore, after taking the action $a = (j,k)$, the user can obtain an immediate reward as follows:

$$R(\mathbf{s}, j, k) = \mathbb{E}\left[U(j,k)|\mathbf{s}\right] = V_{j,k}(\mathbf{s}) - c(\mathbf{s}, j, k, 1), \tag{5}$$

where $U(j,k)$ is the common utility function of any subscriber in the system if she indeed requests for (j,k). In general, the immediate reward is the expected valuation of the successfully decoded video $V_{j,k}(\mathbf{s})$ minus the subscription payment, which happens to be the expected utility of a type t user with subscription $(j \in \mathcal{J}^t, k)$ in state $\mathbf{s}$. For the case that $(j,k) = (0,0)$, we let $R(\mathbf{s}, 0, 0) = 0,\ \forall \mathbf{s}$.

Policy is an important concept in Markov decision process. A policy is an action profile which describes a decision of the decision maker at a certain state in the stochastic system. In our model, a policy is denoted as a function $\pi(\mathbf{s},t) : \mathcal{S} \times \mathcal{T} \mapsto \mathcal{A}^t$, which describes the subscription decision of type t user when she arrives in state $\mathbf{s}$.

In M-MDP, multi-dimensional Bellman equations are necessary to understand the expected utilities of all possible decision makers, given their types and states. That is, the Bellman equations help a decision maker to know what the expected utility is when certain actions (by herself and other decision makers) described in a given policy apply to the stochastic system. In addition, each rational decision maker will choose the action that maximizes her expected utility given her type and state, which is described by the optimal policy equations. It should be noted that although it may be intuitive for a decision maker to choose the optimal action given the currently predicted policy, other decision maker's choices on the action might change accordingly, alter the policy, and therefore change the expected utility given by the multi-dimensional Bellman equations. Therefore, both bellman equations and policy should be balanced in order to reach the equilibrium state.

B. Expected Reward Under Transition Probability

A rational user will make the decision to maximize the expected reward defined as follows:

$$W(\mathbf{s}, j, k) = \mathbb{E}\left[\sum_{l=l^e}^{\infty} (1-\mu)^{l-l^e} R(\mathbf{s}^l, j, k)|\mathbf{s}\right], \tag{6}$$

where μ is the departure rate, and thus $1-\mu$ is the probability that the user will stay at next time slot. Therefore, a user who enters the system at slot l^e will receive the reward at slot l if and only if he stays in the system for subsequent $l - l^e$ slots, which happens with the probability $(1-\mu)(l-l^e)$. Note that $1-\mu$ can also be considered as the discount factor of the future utility. In the steady state, the Bellman equation of the expected reward can be written as follows:

$$W(\mathbf{s}, j, k) = R(\mathbf{s}, j, k) + (1-\mu)\sum Pr(\mathbf{s}'|\mathbf{s}, \pi, j, k) W(\mathbf{s}', j, k), \tag{7}$$

where $Pr(\mathbf{s}'|\mathbf{s}, \pi, j, k)$ denotes the transition probability from $\mathbf{s}$ to $\mathbf{s}'$ when the user takes the action (j,k) under the policy π. Notice that this transition probability is conditioned on the fact that she is not leaving the system at next time slot. Otherwise she will not receive the rewards from next time slot. Let $\mathbf{e}_{j,k}$ be a standard basis vector. Then, the transition probability of a user with a subscription (j,k) is as follows:

$$Pr(\mathbf{s}'|\mathbf{s}, \pi, j, k) = \begin{cases} \sum_{t \in \mathbf{T}, \mathbf{s}' = \mathbf{s} + \mathbf{e}_{\pi(\mathbf{s},t)}} \lambda^t, & \exists t \in \mathbf{T}, \mathbf{s}' = \mathbf{s} + \mathbf{e}_{\pi(\mathbf{s},t)} \text{ and } \pi(\mathbf{s},t) \neq (0,0); \\ \left(n^s_{j',k'}\right)\mu, & \mathbf{s}' = \mathbf{s} - \mathbf{e}_{j',k'}; \\ \left(n^s_{j,k} - 1\right)\mu, & \mathbf{s}' = \mathbf{s} - \mathbf{e}_{j,k}, n^{s'}_{j,k} \neq 0; \\ 1 - \mu(N(\mathbf{s}) - 1) - \lambda(\mathbf{s}, \pi), & \mathbf{s} = \mathbf{s}', n^{s'}_{j,k} \neq 0; \\ 1 - \mu N(\mathbf{s}) - \lambda(\mathbf{s}, \pi), & \mathbf{s} = \mathbf{s}', n^{s'}_{j,k} = 0 \\ 0, & \text{else.} \end{cases} \tag{8}$$

where $N(\mathbf{s}) = \sum_{j,k} n^s_{j,k}$ and $\lambda(\mathbf{s}, \pi) = \sum_{t \in \mathcal{T}} \mathbf{1}(\pi(\mathbf{s},t) \neq (0,0))\lambda^t$.

Since users are assumed to be rational, a type t user should choose the subscription that maximizes her expected utility when she arrives at the system. Recalling that $c(\mathbf{s}, j, k, 0)$ is the initial subscription fee for (j,k), which will be zero under the per-slot charge scheme and non-negative under the one-time charge scheme, the optimal policy under the expected reward $W(\mathbf{s}, j, k)$ is given by

$$\pi(\mathbf{s}, t) \in \arg\max_{(j,k) \in \mathcal{A}^t} W(\mathbf{s} + \mathbf{e}_{j,k}, j, k) - c(\mathbf{s}, j, k, 0), \tag{9}$$

Notice that when the server is full, new users cannot enter, i.e., $\forall N(\mathbf{s}) = N, \pi(\mathbf{s}, t) = (0,0),\ \forall t \in \mathcal{T}$.

The Nash equilibrium is achieved when (7) and (9) are satisfied, which are denoted as the equilibrium conditions. When the equilibrium conditions are met, each type of users has chosen the subscription that maximizes the expected utility. Therefore, they have no incentive to deviate.

C. Average Revenue Maximization for Service Provider

The objective of the service provider is to maximize her revenue in the system under the rational response of subscribers. Let π and W be the policy and expected rewards derived in (7) and (9) under the pricing function $c(\cdot)$. Then, let $Q(\mathbf{s})$ be the expected revenue in state $\mathbf{s}$. The best response of the service provider is the solution to the following optimization problem:

$$\max_{c(\cdot)} \lim_{L \to \infty} \frac{1}{L} \mathbb{E}\left[\sum_{l=1}^{L} Q(\mathbf{s}^l)\right], \tag{10}$$

under the constraints

$$W^*\left(\mathbf{s} + \mathbf{e}_{\pi^*(\mathbf{s},t)}, \pi^*(\mathbf{s},t)\right) - c(\mathbf{s}, j, k, 0) \geq 0,\ \forall \mathbf{s}, t \tag{11}$$

where π^* and W^* satisfy (7) and (9).

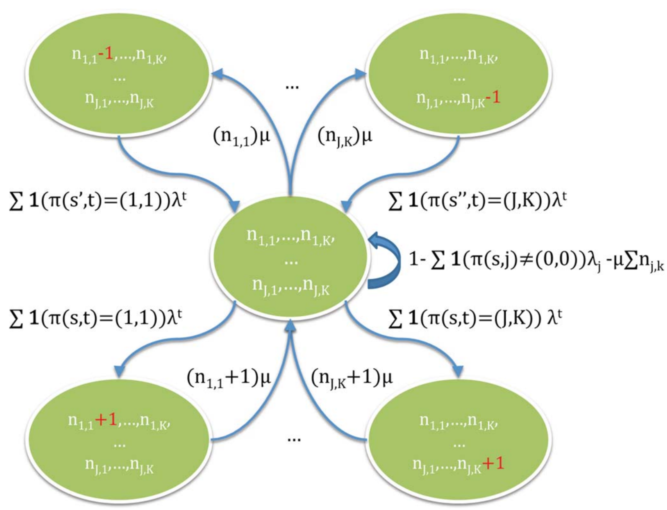


Fig. 2. An illustration of state transition in the proposed M-MDP system.

Note that the state transition probability in this problem is different from the one observed by users in (8). It should be:

$$Pr(\mathbf{s}'|\mathbf{s},\pi) = \begin{cases} \sum_{t\in\mathbf{T},\mathbf{s}'=\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)}} \lambda^t, & \exists t\in\mathbf{T}, \mathbf{s}'=\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)} \\ & \text{and } \pi(\mathbf{s},t)\neq(0,0); \\ \left(n^s_{j,k}\right)\mu, & \mathbf{s}'=\mathbf{s}-\mathbf{e}_{j,k}; \\ 1-N(\mathbf{s})\mu-\lambda(\mathbf{s},\pi), & \mathbf{s}=\mathbf{s}'; \\ 0, & \text{else.} \end{cases} \tag{12}$$

An illustration of the state transition is shown in Fig. 2.

With primal-dual transformation [30], the expected average revenue can be given as $\sum_{\mathbf{s}\in\mathcal{S}} Pr(\mathbf{s}|\pi)Q(\mathbf{s})$, where $Pr(\mathbf{s}|\pi)$ is the stationary distribution of the states under policy π.

V. Optimal Pricing Strategies

In this section, we would like to discuss the optimal pricing strategy that maximizes the expected revenue under a given policy while satisfying the user equilibrium conditions in (7) and (9). Since the policy is given and fixed, the optimization problem is simplified. This helps us to have an initial understanding on how the service provider should determine the price in order to maximize its expected revenue given a known pattern of the requests from users and constraints from the corresponding equilibrium conditions.

A. Optimal Pricing in One-Time Charge Scheme

In one-time charge scheme, a user is charged with a state-dependent subscription fee $P^e_{j,k}(\mathbf{s})$ when her subscription (j,k) is accepted, and no further payments are required. Therefore, we have $c(\mathbf{s},j,k,1)=0$ and $c(\mathbf{s},j,k,0)=P^{\mathbf{e}}_{j,k}(\mathbf{s})\geq 0$ for all $\mathbf{s}$, j,k. Let $W(\mathbf{s},j,k)$ be the expected reward derived by solving (7) through dynamic programming or matrix operations. Since the policy π is fixed, the transition probability is fixed. Therefore, the original revenue optimization problem can be solved in a state-wise way.

Since the server can serve at most N users, we have $N(\mathbf{s})\leq N$ for all $\mathbf{s}$. When $N(\mathbf{s})=N$, all requests will be rejected, which leads to zero revenue. When $N(\mathbf{s})<N$, the revenue maximization problem can be written as follows:

$$\max_{\left\{P^e_{j,k}(\mathbf{s})\right\}} Pr(\mathbf{s}|\pi) \sum_{t\in\mathcal{T},\pi(\mathbf{s},t)\neq(0,0)} \lambda^t P^e_{\pi(\mathbf{s},t)}(\mathbf{s}), \tag{13}$$

under the constraint in (7) and (9). We first relax the constraint set by letting all subscribers derive non-negative expected rewards if they follow the policy $\pi(\mathbf{s},t)$ and ignore (9). We then have the following relaxed constraint set

$$W\left(\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)},\pi^*(\mathbf{s},t)\right) - P^e_{\pi(\mathbf{s},t)}(\mathbf{s}) \geq 0, \quad \forall \mathbf{s},t. \tag{14}$$

Clearly, the solution to the relaxed optimization problem is $\forall t\in\mathcal{T}, P^e_{\pi(\mathbf{s},t)}(\mathbf{s}) = W(\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)},\pi^*(\mathbf{s},t))$. By applying the solution to all $\mathbf{s}\in\mathcal{S}$, we have the optimal pricing strategy for the relaxed problem:

$$\forall \mathbf{s},t,N(\mathbf{s})<N, P^e_{\pi(\mathbf{s},t)}(\mathbf{s}) = W\left(\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)},\pi^*(\mathbf{s},t)\right). \tag{15}$$

Then, we derive the following theorem.

Theorem 1: A policy $\pi(s,j)$ with the pricing strategy $P^e_{j,k}(\mathbf{s}) = W(\mathbf{s}+\mathbf{e}_{j,k},j,k)$ $\forall \mathbf{s},N(\mathbf{s})<N, j\in\mathcal{J},k\in\mathcal{K}$ satisfies (9) for all $\mathbf{s}\in\mathcal{S}$.

Proof: Notice that the expected utility from a subscription (j,k), which is $W(\mathbf{s}+\mathbf{e}_{j,k},j,k)-P^e_{j,k}(\mathbf{s})$, becomes zero under every state, every type, and every subscription when $P^e_{j,k}(\mathbf{s}) = W(\mathbf{s}+e_{j,k},j,k)$ $\forall \mathbf{s},N(\mathbf{s})<N, j\in\mathcal{J},k\in\mathcal{K}$. Therefore, the (9) is always satisfied. □

From Theorem 1, we can see that there is always a solution to the optimization problem in (13) and the solution can be described as

$$P^e_{j,k}(\mathbf{s}) = W(\mathbf{s}+\mathbf{e}_{j,k},j,k), \quad \forall \mathbf{s}\in\mathcal{S},N(\mathbf{s})<N. \tag{16}$$

B. Optimal Pricing in Per-Slot Charge Scheme

In per-slot charge scheme, a user is charged with a state-dependent $P_{j,k}(\mathbf{s})$ at each slot she stays, and no entrance fee is required. Therefore, we have $c(\mathbf{s},j,k,1)=P_{j,k}(\mathbf{s})\geq 0$ and $c(\mathbf{s},j,k,0)=0$ for all $\mathbf{s}$, j, k. We would like to derive the optimal pricing strategy under the constraints in (9).

Let $Pr(\mathbf{s}|\pi)$ be the stationary probability that the system is in state $\mathbf{s}$ under policy π. Since the policy π is fixed, the transition probability is fixed. The revenue maximization problem becomes

$$\max_{\{P_{j,k}(\mathbf{s})\}} \sum_{\mathbf{s}\in\mathcal{S}} Pr(\mathbf{s}|\pi) \sum_{j\in\mathcal{J},k\in\mathcal{K}} n^s_{j,k}P_{j,k}(\mathbf{s}) \tag{17}$$

under the constraints in (9).

Let $\mathcal{R}_{j,k}$, $\mathcal{W}_{j,k}$, $\mathcal{V}_{j,k}$, and $\mathcal{P}_{j,k}$ be the $1-$by$-|\mathcal{S}|$ matrix representation of $R(\mathbf{s},j,k)$, $W(\mathbf{s},j,k)$ $V_{j,k}(\mathbf{s})$, and $P_{j,k}(\mathbf{s})$ over $\mathcal{S}$. Then, let $\mathcal{L}(\pi)$ be the state transition probability matrix under policy π, which is a $|\mathcal{S}|-$by$-|\mathcal{S}|$ matrix with terms $Pr(\mathbf{s}'|\mathbf{s},\pi)$ over $\mathcal{S}$. From (5) and (7), we have

$$(I-(1-\mu)\mathcal{L}(\pi))\,\mathcal{W}_{j,k} = \mathcal{R}_{j,k} = \mathcal{V}_{j,k} - \mathcal{P}_{j,k}, \quad \forall j,k$$

Therefore, the constraints in (9) can be re-written as

$$(I-(1-\mu)\mathcal{L}(\pi))^{-1}(\mathcal{V}_{j,k}-\mathcal{P}_{j,k})=\mathcal{W}_{j,k},\ \forall j,k \tag{18}$$

$$W\left(\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)},\pi(\mathbf{s},t)\right)\geq 0,\ \forall\pi(\mathbf{s},t)\neq(0,0) \tag{19}$$

$$W\left(\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)}),\pi(\mathbf{s},t)\right)-W(\mathbf{s}+\mathbf{e}_{j,k},j,k)\geq 0,$$
$$\forall\pi(\mathbf{s},t)\neq(0,0),(j,k)\in\mathcal{A}^t \tag{20}$$

$$W\left(\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)},\pi(\mathbf{s},t)\right)\leq 0,$$
$$\forall\pi(\mathbf{s},t)=(0,0),(j,k)\in\mathcal{A}^t \tag{21}$$

which is a set of linear constraints over $\mathcal{P}=\{P_{j,k}(\mathbf{s})|\mathbf{s}\in\mathcal{S},j\in\mathcal{J},k\in\mathcal{K}\}$. Therefore, the original maximization problem is equivalent to the following linear programming problem:

$$\max_{\mathcal{P}}\sum_{\mathbf{s}\in\mathcal{S}}Pr(\mathbf{s}|\pi)\sum_{j\in\mathcal{J},k\in\mathcal{K}}n^s_{j,k}P_{j,k}(\mathbf{s}) \tag{22}$$

with the constraints in (18)–(21), which can be solved by standard linear programming methods.

VI. Revenue-Maximized Policy

Finding revenue-maximized policy is challenging since the policy, revenue function and the pricing strategy couple together as discussed in previous sections. In such a case, the traditional MDP solvers cannot be directly applied for the revenue-maximization MDP problem. Therefore, there is a need for a more efficient approach. In this section, we will first prove that the maximum revenue under both pricing schemes is equal for any given policy. Then, we will show that there exists an optimal pricing strategy in per-slot charge scheme which always satisfies the equilibrium conditions while maximizing the expected revenue. Moreover, such an optimal pricing strategy makes the per-state revenue independent from the policy. In such a case, we can reduce the original complex revenue maximization problem to a traditional average-reward MDP problem that can be solved efficiently.

A. Revenue-Maximization Problem

In this subsection, we discuss how to formulate the revenue-maximization problem. We will use the one-time charge scheme for illustration. The per-slot charge scheme can be formulated in a similar way. From (16), we can see that the optimal price is equal to users' expected reward, which means that the revenue is maximized when the expected rewards are maximized. Let $\mathcal{L}(\pi)$ be the state transition matrix and $Q^*(\mathbf{s},\pi)$ be the revenue over states, which is

$$Q^*(\mathbf{s},\pi)=\sum_{t\in\mathcal{T},\pi(\mathbf{s},t)\neq(0,0)}\lambda^t W(\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)},\pi(\mathbf{s},t)), \tag{23}$$

where $W(\mathbf{s},j,k)$ satisfies (7). Then, the revenue optimization problem becomes

$$\max_{\pi}\lim_{L\to\infty}\frac{1}{L}\sum_{l=1}^{L}\mathcal{L}(\pi)^{l-1}Q^*(\mathbf{s^0},\pi), \tag{24}$$

which is a Markov decision process concerning the average expected reward. Unfortunately, the immediate reward $\mathcal{V}^*$ depends on the policy π, due to which, the linear programming or the traditional iteration-based algorithms cannot be directly applied here. A dynamic approach may be applied by iteratively updating the $\mathcal{V}^*$ and π. However, the convergence cannot be guaranteed.

B. Equality in Optimal Revenue and Policy

Here we state one of our main results in the revenue optimization problem.

Theorem 2: Let $Rev^{one,*}$ and $Rev^{per,*}$ be the optimal revenue of the proposed system under one-time charge and per-slot charge schemes. Then, $Rev^{one,*}=Rev^{per,*}$.

Proof: The proof contains two parts. In the first part, we prove that given any pricing strategy under the one-time charge scheme with a given policy π, we can always find a feasible solution under the per-slot charge scheme reaching the same revenue. In the second part, we prove vice versa. When these two directions hold for any policy, we can conclude that the optimal revenue under per-slot charge scheme is never larger or lower than under the one-time charge scheme and therefore $Rev^{one,*}=Rev^{per,*}$.

Part I: Let $\{P^e_{j,k}(\mathbf{s})\}$ be the optimal solution to the one-time charge price optimization problem in (13), with the expected revenue as $Rev^{one,*}$. From (16), we have

$$P^e_{\pi(\mathbf{s},t)}(\mathbf{s})=W^{one}\left(\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)},\pi(\mathbf{s},t)\right),\ \forall\pi(\mathbf{s},t)\neq(0,0).$$

Then we construct the per-slot charge prices that reaches the same revenue. Let $P'_{j,k}(\mathbf{s})=V_{j,k}(\mathbf{s}),\ \forall\mathbf{s},j,k$ be the constructed per-slot charge price, where $\mathcal{P}'$ is its matrix form. With such a price, we have the immediate reward $R^{per}(\mathbf{s},j,k)=V_{j,k}(\mathbf{s})-P'_{j,k}(\mathbf{s})=0$ and therefore the expected reward of per-slot charge scheme $W^{per}(\mathbf{s},j,k)=0$, which satisfies (7) and (9) under any given policy. Moreover, as $\mathcal{W}^{per}$ is a linear transform of $R^{per}(\mathbf{s},j,k)=V_{j,k}(\mathbf{s})-P'_{j,k}(\mathbf{s})$, there exists a transform matrix $A(\pi)$ with the expected reward matrix of per-slot charge scheme $\mathcal{W}^{per}=A(\pi)\mathcal{R}^{per}=A(\pi)\mathcal{V}-A(\pi)\mathcal{P}'=0$. Notice that since the expected reward matrix of one-time charge scheme $\mathcal{W}^{one}$ is also a linear transform of $\mathcal{V}$ with the same transform matrix $A(\pi)$, we have $\mathcal{W}^{one}=A(\pi)\mathcal{V}$, which leads to $\mathcal{W}^{one}=A(\pi)\mathcal{P}'$.

The expected total payment by a user with subscription (j,k) under the per-slot charge scheme can be written using Bellman equation as follows:

$$Rev^{per}_{j,k}(\mathbf{s})=P'_{j,k}(\mathbf{s})+(1-\mu)\sum_{\mathbf{s}'\in\mathcal{S}}Pr(\mathbf{s}'|\mathbf{s},\pi)Rev^{per}_{j,k}(\mathbf{s}') \tag{25}$$

Let $\mathcal{E}^{per}$ be the matrix form of $P^{total,per}_{j,k}(\mathbf{s})$. The solution $\mathcal{E}^{per}$ to the above Bellman equation function is $\mathcal{E}^{per}=A(\pi)\mathcal{P}'=\mathcal{W}^{one}$. Finally, a type t user that arrives in state $\mathbf{s}$ submits the subscription $\pi(\mathbf{s},t)$ to the system. The per-slot charge payment starts as soon as she enters the system, where the state becomes

$\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)}$. Therefore, the expected revenue generated from the type t user with $P'_{j,k}(\mathbf{s})$ is

$$Rev^{per,t}_{\pi(\mathbf{s},t)} = Rev^{per}_{\pi(\mathbf{s},t)}\left(\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)}\right) = W^{one}\left(\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)}, \pi(\mathbf{s},t)\right), \quad (26)$$

which is exactly the same as the optimal one-time charge price $P^e_{j,k}(\mathbf{s})$. Therefore, the expected revenue Rev^{per} under the per-slot charge scheme with price $\{P'_{j,k}(\mathbf{s}) = V_{j,k}(\mathbf{s})\}$ should be equal to the one under the one-time charge scheme with price $\{P^e_{j,k}(\mathbf{s})\}$. As a result, $Rev^{per,*} \geq Rev^{per} = Rev^{one,*}$.

Part II: Let $\{P_{j,k}(\mathbf{s})\}$ be a feasible solution to the per-slot charge pricing optimization problem in (22) and $\mathcal{P}$ is its matrix form. Since it is a feasible solution, it should satisfy (18)–(21). Recalling the Bellman equation in (25) and let $\mathcal{E}^{per,t}$ be the matrix form of the expected revenue generated by a type t user under the per-slot charge price $\{P_{j,k}(\mathbf{s})\}$ and $\mathcal{P}$ be $\{P_{j,k}(\mathbf{s})\}$'s matrix form, we have $\mathcal{E}^{per} = A(\pi)\mathcal{P}$. In addition, the expected reward $W^{per}_{j,k}(\mathbf{s})$ in matrix form is

$$\mathcal{W}^{per} = A(\pi)\mathcal{R}^{per} = A(\pi)\mathcal{V} - A(\pi)\mathcal{P} = W^{one} - \mathcal{E}^{per}. \quad (27)$$

Recalling $Rev^{per,t}_{\pi(\mathbf{s},t)}$ be the expected revenue generated from the type t user under π and state $\mathbf{s}$. According to the above discussion and (19), we have $Rev^{per,t}_{\pi(\mathbf{s},t)} = Rev^{per}_{\pi(\mathbf{s},t)} \leq W^{one}(\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)}, \pi(\mathbf{s},t))$, $\forall \pi(\mathbf{s},t) \neq (0,0)$, which means that the expected revenue from user t is bounded by $W^{one}(\mathbf{s}+e_{\pi(\mathbf{s},t)}, \pi(\mathbf{s},t))$. Therefore, the expected overall revenue is

$$\begin{aligned} Rev^{per} &= Pr(\mathbf{s}|\pi)\sum_{t\in\mathcal{T}}\lambda^t Rev^{per,t}_{\pi(\mathbf{s},t)} \\ &\leq Pr(\mathbf{s}|\pi)\sum_{t\in\mathcal{T}}\lambda^t W^{one}\left(\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)}, \pi(\mathbf{s},t)\right) \\ &\leq Rev^{one,*}. \end{aligned} \quad (28)$$

The last inequality comes from the fact that $Rev^{one,*}$ is the optimal revenue under one-time charge scheme. Therefore, we have $Rev^{one,*} \geq Rev^{per}$ for all feasible $\{P_{j,k}(\mathbf{s})\}$, which means $Rev^{one,*} \geq Rev^{per,*}$. Combining the results in Part I and II, we conclude that $Rev^{one,*} = Rev^{per,*}$. □

From Theorem 2, we can see that both pricing schemes have the same maximum expected revenue when the optimal pricing strategies are applied. In other words, these two schemes are equivalent in terms of revenue optimization. Moreover, in the Part I of the proof, we observe that a simple pricing strategy under the per-slot charge scheme leads to optimal under the one-time charge scheme, which is

$$P^*_{j,k}(\mathbf{s}) = V_{j,k}(\mathbf{s}), \ \forall \mathbf{s} \in \mathcal{S}, j \in \mathcal{J}, k \in \mathcal{K}. \quad (29)$$

Since the revenue under such a pricing strategy reaches the same optimal value of the one-time charge scheme, it is also the optimal solution to the per-slot charge scheme. Such a pricing strategy is very useful since the price and per-state revenue is now independent from the policy. The optimal policy can then be derived by formulating the problem as a traditional average-reward MDP with $V(\mathbf{s}) = \sum_{j,k} n^s_{j,k} P^*_{j,k}(\mathbf{s})$ as its immediate reward.

C. Algorithm for Finding Revenue-Maximized Policy

In Algorithm 1, we propose a value iteration method to find the ε-optimal solution. Since both pricing schemes share the same optimal revenue under the same policy, the policy found through Algorithm 1 is optimal under both pricing schemes.

Algorithm 1 Value Iteration for Revenue-Maximization Solution

```
Initialize π^o, W^{v,o}
while 1 do
  for all s ∈ S do
    π^n ← argmax_{π(s)} {V(s) + Σ_{s'∈S} Pr(s'|s, π(s)) W^{v,o}(s')}
    W^{v,n}(s) ← V(s) + Σ_{s'∈S} Pr(s'|s, π^n) W^{v,o}(s')
  end for
  W^{v,d} ← W^{v,n} − W^{v,o}
  if max W^{v,d} − min W^{v,d} < ε then
    Break
  else
    W^{v,o} ← W^{v,n}
  end if
end while
Output π^n and W^{v,n}
```

The convergence and optimality of such an pseudo-polynomial algorithm is guaranteed as shown in [30].

Social welfare, which is a common measurement of the system efficiency in game theory [20], is the sum of all utilities of the players. Given that the utility of the service provider is exactly the payment from the users, the social welfare of the proposed framework is equal to the total valuation of users to the service. Notice that under the optimal pricing strategy $P^*_{j,k}(\mathbf{s})$, the per-state revenue is

$$V(\mathbf{s}) = \sum_{j\in\mathcal{J},k\in\mathcal{K}} n^s_{j,k} P^*_{j,k}(\mathbf{s}) = \sum_{j\in\mathcal{J},k\in\mathcal{K}} V_{j,k}(\mathbf{s}), \quad (30)$$

which is the overall user's valuation on the system, i.e., social welfare, in state $\mathbf{s}$. Therefore, the solution to the revenue-maximization problem also maximizes the social welfare of the system. As a result, we have the following corollary.

Corollary 1: When $\varepsilon \to 0$, the revenue-maximized policy derived in Algorithm 1 is equivalent to a socially-optimal policy.

Nevertheless, the value-iteration algorithm 1 converges slowly as $\varepsilon \to 0$ since the revenue-optimization problem is an average-reward MDP system. Next, we will propose an approximate algorithm based on a discounted MDP model.

D. Policy Iteration γ-Optimal Algorithm

Here we propose a discounted MDP model as an approximation to the average-reward MDP. By modeling the revenue-optimization problem as a discounted MDP system, we propose an approximate algorithm which converge to the optimal policy in significantly less rounds.

We first model the revenue-optimization problem as a γ-discounted Markov decision process. We introduce γ as a discounted factor for the service provider on the future revenue, where $1 > \gamma > 0$. Then, let $W^{\gamma}(\mathbf{s})$ be the expected total revenue if the current state is $\mathbf{s}$, then we have

$$W^{\gamma}(\mathbf{s}) = \sum_{l=0}^{\infty} \mathbb{E}\left[(1-\gamma)^l V(\mathbf{s^l}) | \mathbf{s^0} = \mathbf{s}\right], \quad (31)$$

Notice that when γ is close to 0, the expected reward is arbitrarily close to the average expected reward. Given the γ, we can derive the γ-optimal policy through the following process. According to the Bellman equation, the expected total revenue can be written as follows:

$$W^{\gamma}(\mathbf{s}) = V(\mathbf{s}) + (1-\gamma) \sum_{\mathbf{s}' \in \mathcal{S}} Pr(\mathbf{s}'|\mathbf{s}, \pi) W(\mathbf{s}'), \quad (32)$$

where transition probability $Pr(\mathbf{s}'|\mathbf{s}, \pi)$ is given in (12).

For the optimal policy, the service provider should choose the action that maximizes its expected total revenue at every state, which is as follows:

$$\pi^{\gamma}(\mathbf{s}) = \arg\max_{\pi} \sum_{t \in \mathcal{T}} \lambda^t W^{\gamma}\left(\mathbf{s} + e_{\pi(s,t)}\right). \quad (33)$$

Notice that both (32) and (33) are coupling together. These two equations describes the optimality conditions of the proposed γ-discounted MDP. As shown in literature [30], the γ-discounted MDP can be solved by policy iteration algorithm. We provide such an algorithm in Algorithm 2.

Algorithm 2 Policy Iteration for γ-Optimal Solution

1: Initialize π^o
2: **while** 1 *do*
3: Solve W^n using (32) with $\pi = \pi^o$
4: Solve π^n using (33) with $W = W^n$
5: **if** $\pi^n = \pi^o$ **then**
6: Break
7: **else**
8: $\pi^o \leftarrow \pi^n$
9: **end if**
10: **end while**
11: Output π^n and W^n

Notice that the convergence of Algorithm 2 is guaranteed as this is a traditional discounted MDP problem with a fixed discounted factor [30]. In addition, the algorithm terminates in polynomial time, which is much faster than Algorithm 1. In all of our simulations, the approximated algorithm converges in less than ten rounds with the resulting revenue loss less than 1% when the discounted factor γ is 0.01. We provide further discussions on this through simulations in Section VII.

VII. Simulation Results

We evaluate the efficiency of the proposed approach through simulations. We consider a SVC multicasting service over a WiMAX network. The wireless system parameters follow the WiMAX standard, in which 7 level of MCSs are chosen and given in Table I [2]. Then, let each MCS's lowest required signal to noise and interference ratio (SINR) be quantized to seven levels $\mathcal{G} = \{1, 2, 3, 4, 5, 6, 7\}$. A user with the channel quality $g \in \mathcal{G}$ can receive data streams transmitted by up to g-th MCS. In all simulations, we assume g is equally, independently, and randomly chosen from these seven levels in every time slot. Then, the transmission time of a specific data stream will be the function of the channel quality. The reason that we model the transmission time as a function of the channel quality is due to the fact that when the channel quality decreases, the supported MCS and the corresponding data rate decreases, which suggests a longer transmission time for the same data stream. Let the data rates offered by the MCSs be $b_1 \sim b_7$, where the exact values are given in Table I. A data stream with a bit-rate of B requires B/b_m time to transmit at each slot if MCS m is chosen.

TABLE I
Transmission Throughput

Quality	Modulation	Data Rate (Mbps)
1	BPSK 1/2	3.8768
2	QPSK 1/2	7.7552
3	QPSK 3/4	11.6336
4	QAM32 1/2	15.512
5	QAM32 3/4	23.2688
6	QAM64 2/3	31.0256
7	QAM64 3/4	34.904

TABLE II
User Specifications

Type	Preferred Video(s)	Maximum Layer
1	MOBCAL	2
2	STOCKHOLM	2
3	STOCKHOLM	3
4	MOBCAL,STOCKHOLM	3

We simulate a SVC multicasting server which provides two videos, MOBCAL and STOCKHOLM, recorded by SVT Sveriges Television AB [31]. We use JSVM reference software [32] to encode each video into a three-layer spatial-scaled H.264 SVC video stream. The cumulative bit-rates (CBR), resolutions, quantization parameters (QP), and peak-signal-to-noise-ratio (PSNR) of each encoded video with different layers can be found in Table III.

The server can serve up to N users, while the total available service time per second is a ratio between 0% and 100%. There are four types of users with difference preferences on videos and different computation capabilities, which are specified in Table II. The user's valuations on layers and videos are shown in Table III. Finally, the user arrival and departure parameters are set to be 0.04 and 0.01.

The resource allocation rule $\mathcal{D}$, which determines the applied MCS $\{\underline{g_{j,k}}\}$ in each state, is maximizing the expected overall valuation given the current demands of videos and corresponding layers from the users. The optimal solution is derived through exhaustive search in the simulations. Notice that most approach to the snapshot-based optimization problem in the literature is applicable to our system.

We evaluate the system efficiency through two performance metrics: average social welfare, which is the total users' valuations on the service, and average revenue of the proposed

TABLE III
VIDEO SPECIFICATIONS

Layer	CBR (kbps)	Resolution	QP	PSNR	User Valuation
MOBCAL					
1	315.33	360i	40	29.74	0.5
1,2	1660.83	720i	36	33.38	0.75
1,2,3	10719.41	1080i	32	33.80	1
STOCKHOLM					
1	319.75	360i	40	29.70	0.8
1,2	1480.32	720i	36	33.68	0.9
1,2,3	6806.29	1080i	32	34.01	1

SVC multicasting system. For the social welfare performance, we compare two policies: revenue-maximized policy and free subscription policy. The revenue-maximized policy is given by Algorithm 1 with $\varepsilon = 0.0001$. The free subscription policy is the solution to (7) and (9) without payment, i.e., $P^e_{j,k} = P_{j,k} = 0,\ \forall j,k$. Free subscription represents the case that all users are free to subscribe any video without any payment. By comparing these two policies, we can evaluate the efficiency loss when no pricing scheme is applied.

For the performance of the average revenue, we compare four pricing schemes: optimal one-time charge pricing, optimal per-slot charge pricing, maximum fixed entrance fee, and differentiate price. The first two schemes are the proposed pricing schemes as discussed in Section VI, while the third one is the pricing scheme with a fixed entrance fee, which is widely adopted in current video subscription services. In the simulation, the entrance fee for the maximum fixed entrance fee scheme is chosen as $P^e = \min_{\mathbf{s},t,\pi(\mathbf{s},t)\neq(0,0)} W^{indv}(\mathbf{s}+\mathbf{e}_{\pi(\mathbf{s},t)}, \pi(\mathbf{s},t))$, where W^{indv} is the expected reward under the free subscription policy. In the differentiate price scheme, we provide differentiated price of each layer and each video in the system. The exact price of a video j up to layer k is determined by the minimum expected utility a subscriber may receive from subscribing such a video in free subscription scheme. This promises that the users still have a non-negative incentive to subscribe the video under this scheme.

We first study how the server capacity influences the social welfare and the revenue by adjusting the server capacity N from 4 to 14 while keeping other settings unchanged. The results are shown in Fig. 3.

From Fig. 3(a), we can see that the proposed revenue-maximized policy achieves higher average social welfare than the free subscription policy under all server capacities. This result verifies that there will be an efficiency loss if no pricing scheme is applied. Notice that under both policies the social welfare per capacity decreases with the expansion of the server capacity. This is due to the negative network externality effect in this system. When the server capacity increases, there will be more users with different demands simultaneously, which means more multicasting groups simultaneously. Given the fixed amount of service time ratio, each multicasting group receives less service time. This generally impairs all users' utilities in this system and therefore reduces the social welfare per user.

From Fig. 3(b), we can see that both optimal per-slot charge pricing and one-time charge pricing result in the same revenue

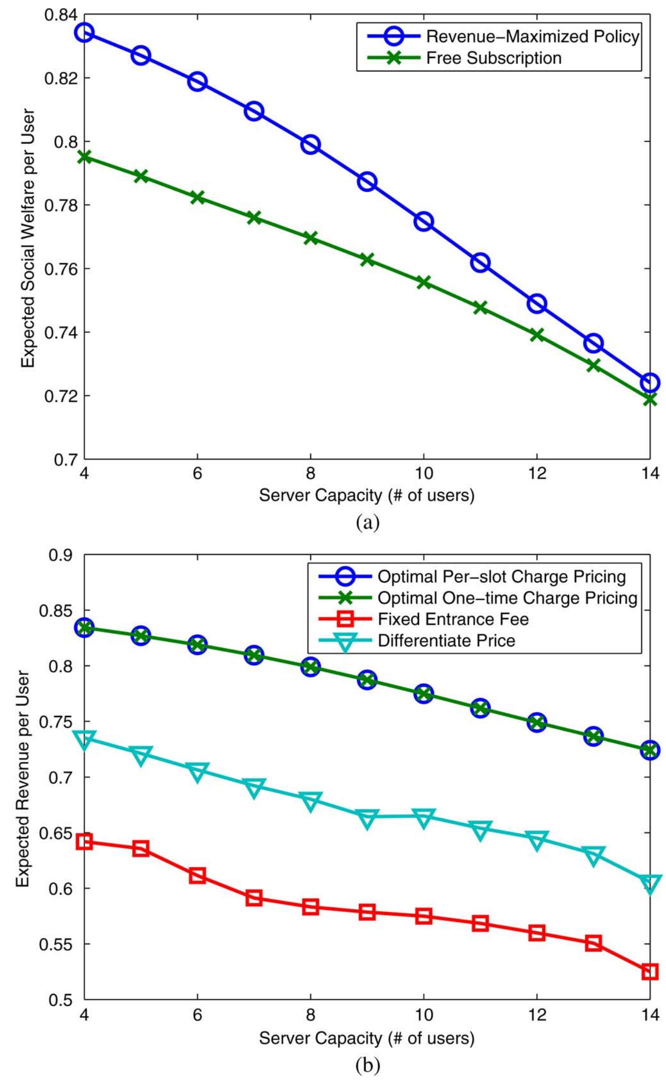


Fig. 3. System performance under different server capacity. (a) Social welfare per user. (b) Revenue per user.

under all scenarios, which verifies our conclusion in Theorem 2. Moreover, the revenue under both optimal pricing schemes is significantly higher than the ones under fixed entranced fee and differentiate price schemes. The later ones lose a large amount of revenue from users with higher valuations on the service. We also observe that differentiate price scheme provides a higher revenue than the fixed-entrance fee scheme since it utilizes the differentiated interests of subscribers on the videos. A more popular, high-quality, and high-value video or layer will be priced with a higher price in the differentiate price scheme. Still, the achieved revenue is still lower than the ones under optimal per-slot charge pricing and optimal one-time charge pricing schemes. Nevertheless, both optimal pricing schemes have lower revenue per capacity when the server capacity expands. The negative network externality in this system also has a negative effect on the revenue.

Then, we investigate how the amount of service time affects the efficiency of the SVC multicasting system under different policies. We control the service time ratio in the system from 0% to 100% with other settings unchanged. The simulations results are shown in Fig. 4.

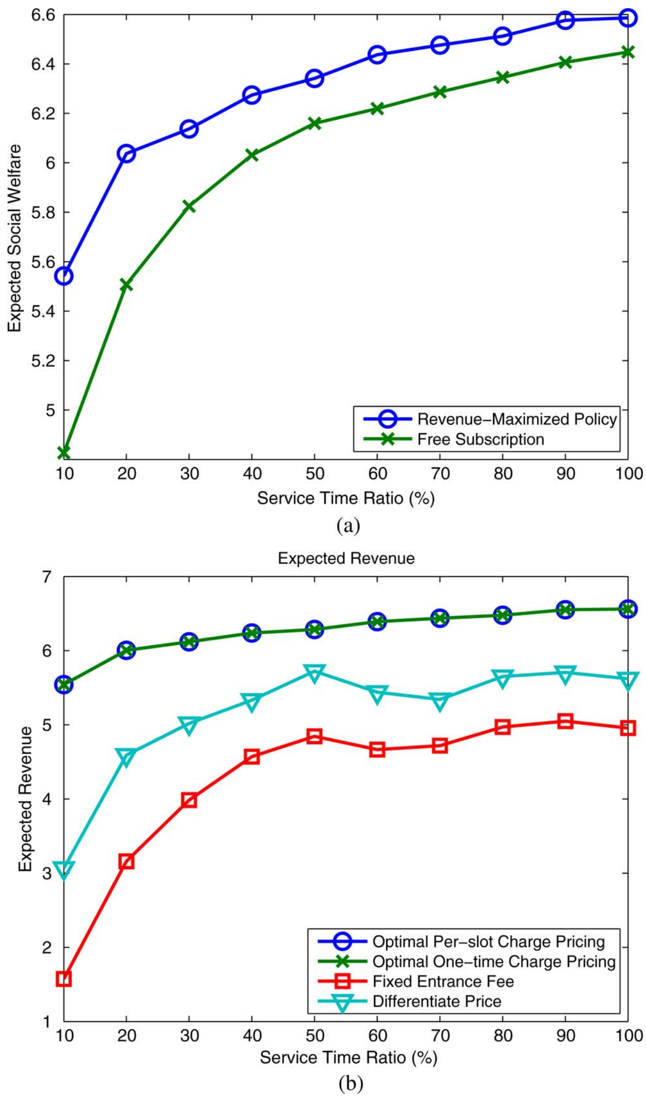


Fig. 4. System performance under different available service time ratio. (a) Total social welfare. (b) Total revenue.

From Fig. 4(a), we observe that the revenue-maximized policy always achieves better performance than the free subscription policy. For the performance of the average revenue shown in Fig. 4(b), we again observe that both optimal pricing schemes have higher revenue than other two schemes. Notice that the revenues under fixed entrance fee and differentiate price schemes do not always increase when the service time ratio increases. This phenomenon comes from the fact that the optimal resource allocation problem in (1) is nonlinear. When the service time increases, it is possible that some streams eventually get less resource in order to improve the transmission quality of other streams along with the increased resource under the optimal allocation. In such a case, the fixed entrance fee scheme has a lower revenue since the price under this scheme is constrained by the video stream with lowest expected reward, which is also the case of differentiate price scheme. Nevertheless, the proposed pricing schemes are resistant to this effect since the corresponding expected revenues are equal to the overall social welfare, which is maximized and nondecreasing with increased service time when the optimal solution in (1) is applied.

Finally, we evaluate the convergence speed and solution quality of the proposed γ-optimal approximate algorithm. We set the maximum number of users as 8 and the service time ratio as 10% with other settings unchanged. Then, we control the discounted factor γ in Algorithm 2 from 0.1 to 0.001, and compare it with the ε-optimal solution derived in Algorithm 1 with $\varepsilon = 0.00001$. The results are in Table IV.

We observe that when the discounted factor γ decreases, the resulting expected revenue with Algorithm 2 is closer to the optimal one from Algorithm 1. When γ is 0.01, there is only 0.01% revenue loss from Algorithm 2. When γ decreases to 0.001, the derived policy is the same as the one from Algorithm 1, and the expected revenue is equal under both algorithms. In addition, the convergence round never exceeds 5 under any simulated γ. In contrasts, Algorithm 1 requires 1080 rounds to converge. We conclude that the proposed γ-optimal algorithm provides efficient approximate solutions to the revenue-maximization problem with much less computational complexity.

VIII. Conclusion and Future Directions

In this paper, we proposed a novel framework to study a general stochastic SVC multicasting system. This framework supports stochastic user arrival and heterogeneous user preferences. It is compatible to existing resource allocation algorithms in the literature. A subscription-based payment system was studied in this framework for exploring the economic value of the system with rational users. The responses of selfish users under two pricing schemes, one-time charge scheme and per-slot charge scheme, were discussed. The equilibrium conditions were derived as the constraints of the corresponding revenue optimization problem for the service provider. We theoretically proved that both pricing schemes reach the same optimal revenue under all policies, and the optimal pricing strategies and policies which maximize the expected revenue of the system can be efficiently derived by reducing the M-MDP problem to a traditional average-reward MDP problem. Moreover, we showed that the revenue-maximized policy is a socially-optimal policy, which means that the proposed optimal policy also maximizes the social welfare.

The resource allocation rule we proposed in (1) is considered as the expectation on the applied MCS in each data stream when a state s is given and the channel distribution $f(g)$ is expected, while the real channel quality is unknown. In the operating stage, the applied MCSs on each stream should depend on the real channel quality each subscriber experienced according to the instantaneous feedback. The original resource allocation rule in (1) is considered as an approximate rule to the one applied in the operating stage. A more accurate resource allocation rule with instantaneous channel quality as input is also possible by including the lowest channel quality experienced by subscribers in each data stream in the state s. The corresponding state transition probability should also be refined according to the channel quality distribution $f(g)$. Nevertheless, such a model will significantly increase the complexity

TABLE IV
EXPECTED REVENUE UNDER DIFFERENT DISCOUNTED FACTOR γ

γ	0.09	0.07	0.05	0.03	0.01	0.001	0
Converg. round	4	4	5	5	4	3	1080
Revenue	5.316	5.389	5.400	5.499	5.542	5.542	5.543
Efficiency	0.959	0.972	0.974	0.992	0.999	1	1

of the proposed framework and will be difficult to address all the concerns within the page limitation.

Readers may also notice that the number of states $|\mathbf{S}|$ grows exponentially when the server capacity N grows. This may bring some concerns on the computation complexity when the server expands. Nevertheless, there are several approximate methods we may apply in the proposed model to reduce the complexity. One method is considering the server loading in blocks instead of in exact numbers when defining the states. That is, the server loading range 1 N is segmented into K blocks. Through this approximation, we reduce the complexity to the desired degree.

REFERENCES

[1] P. Li, H. Zhang, B. Zhao, and S. Rangarajan, "Scalable video multicast with adaptive modulation and coding in broadband wireless data systems," *IEEE-ACM Trans. Netw.*, vol. 20, no. 1, pp. 57–68, Feb. 2012.

[2] W.-H. Kuo, W. Liao, and T. Liu, "Adaptive resource allocation for layer-encoded IPTV multicasting in IEEE 802.16 WiMAX wireless networks," *IEEE Trans. Multimedia*, vol. 13, no. 1, pp. 116–124, Feb. 2011.

[3] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

[4] S.-F. Chang and A. Vetro, "Video adaptation: Concepts, technologies, open issues," *Proc. IEEE*, vol. 93, no. 1, pp. 148–158, Jan. 2005.

[5] S. Deb, S. Jaiswal, and K. Nagaraj, "Real-time video multicast in WiMAX networks," in *Proc. IEEE INFOCOM*, Phoenix, AZ, USA, Apr. 2008, pp. 2252–2260.

[6] Y. Yu, P. Hsiu, and A. Pang, "Energy-efficient video multicast in 4G wireless systems," *IEEE Trans. Mobile Comput.*, vol. 11, no. 10, pp. 1508–1522, Oct. 2012.

[7] S. Chuah, Z. Chen, and Y. Tan, "Energy-efficient resource allocation and scheduling for multicast of scalable video over wireless networks," *IEEE Trans. Multimedia*, vol. 14, no. 4, pp. 1324–1336, Aug. 2012.

[8] W. Ji, Z. Li, and Y. Chen, "Joint source-channel coding and optimization for layered video broadcasting to heterogeneous devices," *IEEE Trans. Multimedia*, vol. 14, no. 2, pp. 443–455, Apr. 2012.

[9] Y. Chen, Y. Wu, B. Wang, and K. J. R. Liu, "Spectrum auction games for multimedia streaming over cognitive radio networks," *IEEE Trans. Commun.*, vol. 58, no. 8, pp. 2381–2390, Aug. 2010.

[10] P. Polacek, T.-Y. Yang, and C.-W. Huang, "Joint opportunistic spectrum access and scheduling for layered multicasting over cognitive radio networks," in *Proc. IEEE MMSP*, Hangzhou, China, Oct. 2011, pp. 1–6.

[11] C.-W. Huang, S.-M. Huang, P.-H. Wu, S.-J. Lin, and J.-N. Hwang, "OLM: Opportunistic layered multicasting for scalable IPTV over mobile WiMAX," *IEEE Trans. Mobile Comput.*, vol. 11, no. 3, pp. 453–463, Jan. 2012.

[12] W. Lin and K. J. R. Liu, "Game-theoretic pricing for video streaming in mobile networks," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2667–2680, May 2012.

[13] P.-H. Wu and Y. Hu, "Optimal layered video IPTV multicast streaming over mobile WiMAX systems," *IEEE Trans. Multimedia*, vol. 13, no. 6, pp. 1395–1403, Dec. 2011.

[14] K. Sundaresan and S. Rangarajan, "Scheduling algorithms for video multicasting with channel diversity in wireless OFDMA networks," in *Proc. ACM MobiHoc*, Paris, France, May 2011, Art. ID. 40.

[15] X. Ji, J. Huang, M. Chiang, G. Lafruit, and F. Catthoor, "Scheduling and resource allocation for SVC streaming over OFDM downlink systems," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 10, pp. 1549–1555, Oct. 2009.

[16] V. Vukadinovic and G. Dán, "Multicast scheduling for scalable video streaming in wireless networks," in *Proc. ACM MMSys*, San Jose, CA, USA, Feb. 2010, pp. 77–88.

[17] E. Maani, P. V. Pahalawatta, and R. Berry, "Scalable video coding and packet scheduling for multiuser video transmission over wireless networks," in *Proc. SPIE Opt. Photon.*, San Diego, CA, USA, 2009.

[18] A. Dan, D. Sitaram, and P. Shahabuddin, "Dynamic batching policies for an on-demand video server," *Multimedia Syst.*, vol. 4, no. 3, pp. 112–121, Jun. 1996.

[19] Y. Chen and K. J. R. Liu, "Understanding microeconomic behaviors in social networking: An engineering view," *IEEE Signal Process. Mag.*, vol. 29, no. 2, pp. 53–64, Mar. 2012.

[20] B. Wang, Y. Wu, and K. J. R. Liu, "Game theory for cognitive radio networks: An overview," *Comput. Netw.*, vol. 54, no. 14, pp. 2537–2561, Oct. 2010.

[21] W. Lin, H. Zhao, and K. J. R. Liu, "Incentive cooperation strategies for peer-to-peer live multimedia streaming social networks," *IEEE Trans. Multimedia*, vol. 11, no. 3, pp. 396–412, Apr. 2009.

[22] V. Zhao and G. Cheung, "Game theoretical analysis of wireless multiview video multicast using cooperative peer-to-peer repair," in *Proc. IEEE ICME*, Barcelona, Spain, Jul. 2011, pp. 1–6.

[23] L. Al-Kanj, W. Saad, and Z. Dawy, "A game theoretic approach for content distribution over wireless networks with mobile-to-mobile cooperation," in *Proc. IEEE PIMRC*, Toronto, ON, Canada, Sep. 2011, pp. 1567–1572.

[24] Y. Su and M. der van Schaar, "Multiuser multimedia resource allocation over multicarrier wireless networks," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 2102–2116, May 2008.

[25] Y.-H. Yang, Y. Chen, C. Jiang, C.-Y. Wang, and K. Liu, "Wireless access network selection game with negative network externality," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5048–5060, Oct. 2013.

[26] C. X. Jiang, Y. Chen, Y. Yang, C. Wang, and K. J. R. Liu, "Dynamic Chinese restaurant game: Theory and application to cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 1960–1973, Apr. 2014.

[27] C.-Y. Wang, Y. Chen, and K. J. R. Liu, "Chinese restaurant game," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 898–901, Dec. 2012.

[28] C.-Y. Wang, Y. Chen, and K. J. R. Liu, "Sequential Chinese restaurant game," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 571–584, Feb. 2013.

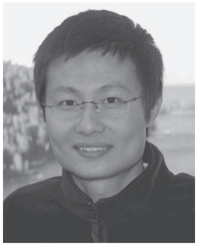
[29] C.-Y. Wang, Y. Chen, H.-Y. Wei, and K. Liu, "Optimal pricing in stochastic scalable video coding multicasting system," in *Proc. IEEE INFOCOM*, Apr. 2013, pp. 540–544.

[30] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Hoboken, NJ, USA: Wiley, 1994.

[31] L. Haglund, "SVT video test sequence," SVT Sveriges Television AB, 2001. [Online]. Available: ftp://ftp.ldv.e-technik.tu-muenchen.de/pub/test_sequences/

[32] H. A. Q. Maarif, T. S. Gunawan, O. O. Khalifa, JSVM Reference Software. [Online]. Available: http://www.hhi.fraunhofer.de/de/kompetenzfelder/image-processing/research-groups/image-video-coding/svc-extension-of-h264avc/jsvm-reference-software.html

Chih-Yu Wang (S'07–M'13) received the B.S. and Ph.D. degrees in electrical engineering and communication engineering from National Taiwan University, Taipei, Taiwan, in 2007 and 2013, respectively. He has been a Visiting Student at the University of Maryland, College Park, MD, USA, since 2011. He is currently an Assistant Research Fellow with the Research Center for Information Technology Innovation, Academia Sinica, Taipei. His research interests include game theory, wireless communications, social networks, and data science.



Yan Chen (SM'14) received the Bachelor's degree from the University of Science and Technology of China, Hefei, China, in 2004, the M.Phil. degree from the Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, in 2007, and the Ph.D. degree from the University of Maryland, College Park, MD, USA, in 2011. His current research interests are in data science, network science, game theory, social learning and networking, as well as signal processing and wireless communications.

Dr. Chen was a recipient of multiple honors and awards, including the Best Paper Award from IEEE GLOBECOM in 2013; Future Faculty Fellowship and Distinguished Dissertation Fellowship Honorable Mention from the Department of Electrical and Computer Engineering in 2010 and 2011, respectively; Finalist of Dean's Doctoral Research Award from A. James Clark School of Engineering, University of Maryland, in 2011; and Chinese Government Award for Outstanding Students Abroad in 2011.

Hung-Yu Wei received the B.S. degree in electrical engineering from National Taiwan University (NTU) Taipei, Taiwan, in 1999 and the M.S. and Ph.D. degrees in electrical engineering from Columbia University, New York, NY, USA, in 2001 and 2005, respectively. He is currently an Associate Professor with the Department of Electrical Engineering and Graduate Institute of Communication Engineering, NTU. His research interests include wireless networking and game theoretic models for networking. He was a Consulting Member of the Acts and Regulation Committee of the National Communications Commission during 2008–2009. He actively participates in wireless network standardization activities and is a voting member of the IEEE 802.16 Working Group. He was the recipient of the Recruiting Outstanding Young Scholar Award from the Foundation for the Advancement of Outstanding Scholarship in 2006 and the NTU Excellent Teaching Award in 2008.

K. J. Ray Liu (F'03) was named a Distinguished Scholar-Teacher of the University of Maryland, College Park, MD, USA, in 2007, where he is Christine Kim Eminent Professor of Information Technology. He leads the Maryland Signals and Information Group, conducting research encompassing broad areas of signal processing and communications with recent focus on cooperative and cognitive communications, social learning and network science, information forensics and security, and green information and communications technology.

Dr. Liu is a Fellow of AAAS. He was the President of the IEEE Signal Processing Society (2012–2013), where he has served as Vice President—Publications and Board of Governor. He was the Editor-in-Chief of *IEEE Signal Processing Magazine* and the founding Editor-in-Chief of *EURASIP Journal on Advances in Signal Processing*. He was a Distinguished Lecturer, recipient of IEEE Signal Processing Society 2009 Technical Achievement Award, 2014 Society Award, and various best paper awards. He also received various teaching and research recognitions from the University of Maryland, including university-level Invention of the Year Award; and Poole and Kent Senior Faculty Teaching Award, Outstanding Faculty Research Award, and Outstanding Faculty Service Award, all from A. James Clark School of Engineering. He is also an ISI Highly Cited Author.