Community Clinic: Economizing Mobile Cloud Service Cost via Cloudlet Group

Jin Zhang, Tao Xiong and Wei Lou
Department of Computing, The Hong Kong Polytechnic University, Hong Kong
Email: {csjzhang, cstxiong, csweilou}@comp.polyu.edu.hk

***Abstract*—The explosive growth of mobile applications causes the mobile traffic to easily exceed the capacity of the cloud service due to the bandwidth limits of last mile connections to the cloud and legacy backhauls to macrocells' base stations. It degrades mobile applications' quality of service since the mobile devices have to spend more time and thus consume more battery power for data transmissions. It also enforces the cloud provider to put a huge investment to update its infrastructure and the induced cost is inevitably borne by all mobile users. To resolve this issue, in this paper we propose a so-called *community clinic* solution, which embeds the cloudlet group between the cloud and mobile users, to cut down the cost introduced by the massive deployment of the cloud's data centers and save the battery power consumed by the mobile devices. We firstly show that the mobile devices can consume less energy by choosing the service provided by the cloudlet group. We then model the system with and without the cloudlet group as two types of supply chain and prove that the cloudlet group can increase the cloud's profit without putting additional cost on mobile users. We also propose the real-time group-buying auction for the cloudlet group to promote its service to its nearby mobile users with a lower price and maximize its profit. The community clinic can result in a win-win-win outcome among the cloud, cloudlet group and mobile users. Numerical experiments are further conducted to demonstrate the effectiveness of our scheme.**

***Index Terms*—Mobile cloud, cloudlet group, supply chain, group-buying auction, game theory.**

I. INTRODUCTION

In recent years, mobile devices such as smart phones and tablets have made the information at fingertips whenever and wherever. However, the mobile services provided are greatly degraded by the available resources, including storage and battery power, of the mobile devices. Fortunately, the cloud computing paradigm has fundamentally changed the resource-insufficient situation of mobile devices because the resource-demanding tasks can be offloaded from the mobile devices to the cloud, which provides a tremendous amount of resources and services to the mobile users directly through Internet (Fig. 1), and thus expedited the explosive growth of mobile data traffic. According to the prediction of Cisco [1], by 2018 the mobile data traffic will surpass 15 exabytes per month, a 11-fold growth from 2013. This will bring a huge burden on the cloud provider since the service demands from mobile users can easily exceed the capacity of current cloud infrastructure due to the following two bottlenecks, last mile connections to the cloud and legacy backhauls to macrocells' base stations [2]. The bandwidth of the last mile connections to the cloud is a major obstacle for providing effective cloud services, which both limits the number of users to access the cloud

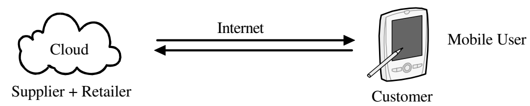


Fig. 1. The bipartite model with cloud and mobile users.

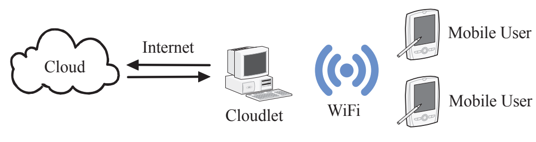


Fig. 2. The model with cloud, cloudlet and mobile users.

and increases the access delay. The upgrading of backhauls to cellular networks is lagged far behind the deployment of cellular networks themselves, e.g., 4G networks can achieve maximal data rates at 100~1000Mbps while their backhuals can support the bandwidth at only 3~8Mbps (with 2 to 4 T1/E1 lines). However, solutions of either building up massive data centers to increase the bandwidth of the last mile connections to the cloud or enlarging the scale of backhauls to the macrocells will incur a remarkable deployment cost for the cloud provider, which is inevitably borne by all mobile users.

Another critical cost that dominates a mobile user's attitude toward the mobile cloud services is directly related to data rate the mobile user can experience. As the explosive growth of mobile service demands will unfavorably have each mobile user suffer a lower per-service data rate, the mobile user has to spend a longer time in transmitting and receiving the data, consequently causing her mobile device to have a higher battery power consumption. As a consequence, a mobile user will prefer to access the mobile cloud services via a wireless network that provides high per-service data rate connections, which can minimize the energy consumption of her mobile device.

Recently, M. Satyanarayanan proposed the cloudlet-based mobile computing model [3] in which the cloudlet is a credible resource-rich computer (or computer cluster) with well-connected Internet, which aims to provide services for proximate mobile users (Fig. 2). As the cloudlet is close to the mobile devices, which enables a high speed transmission between them, it can serve the mobile users effectively by firstly downloading the contents from the cloud to the cloudlet and then distributing them to the mobile users. Obviously, using the cloudlet is a better solution to overcome the above issues. However, it will raise the following concerns from the viewpoints of the cloud, the cloudlet and the mobile user:

1) *The cloud concerns about the reallocation of the profit*: As a separate entity that serves for the mobile users, the cloudlet should be driven by certain incentives. However, it should not increase the spending of the mobile users; otherwise, the users would unlikely choose the services provided by the cloudlet. Under this condition, will the cloudlet take away some profit from the cloud when it partakes in serving the mobile users, which sacrifices the cloud's profit if all the services should be originally served by the cloud?

2) *The cloudlet intends to attract more mobile users to use its services*: The cloudlet can get payoff by lessening the burden of the cloud for serving the mobile users. As the mobile users are rational, how can the cloudlet motivate the mobile users to choose its services? Obviously, a most effective way to attract as many mobile users as possible to use the cloudlet's services is to cut down the service price provided by the cloudlet. Then, how can the cloudlet cut down the service price even when the mobile users grow explosively?

3) *The mobile users concern about the energy consumption of the services*: Since leveraging the cloudlet to transmit and receive the data instead of accessing the cloud directly changes the data flow path, will this change introduce more energy consumption to the mobile devices in the process of data transmission and reception? As any change that may cause the mobile devices to consume more energy would likely be abandoned by the mobile users, how can the cloudlet assist the mobile devices to consume less energy?

To address these concerns, we propose a novel solution called as *community clinic* in this paper, which embeds a *cloudlet group* component between the cloud and the mobile users. The relationship among the cloud, cloudlet group and the mobile users is analogical to that among the central hospital, community clinic and the patients. The community clinics provide convenient medical treatments for the regional patients at a lower price, which lessens the burden of the central hospitals effectively. To easily describe our scheme, we choose the videos as the examples of digital contents because the mobile videos, which take about 53% of the mobile traffic in 2013, exceeds the total mobile traffic in 2012 [1]. Such mobile videos are paid by the end users directly by purchasing data access services or indirectly by viewing advertisements. We leverage the community clinic scheme to resolve all the above concerns with its unique economic and technic characteristics:

1) We evaluate the profit of the cloud by modeling the whole system as a tripartite supply chain where the cloud plays as the supplier, the cloudlet group plays as the retailer, and the mobile users play as the consumer. We prove that, without putting extra cost on the mobile users, the cloudlet group is capable of raising the profit for the cloud through satisfying more demands of mobile users, comparing with the bipartite supply chain model where the cloud plays the dual roles as both supplier and retailer, and the mobile user plays as the consumer. The cloudlet group will make the cloud provider save the cost for deploying massive data centers, which is the premise for lowering the price of any cloud service.

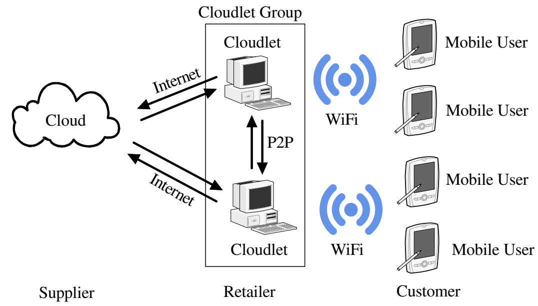


Fig. 3. The tripartite model with cloud, cloudlet group and mobile users.

2) We introduce the real-time group-buying auction mechanism to greatly reduce the price of the digital contents so as to attract more mobile users to choose the services provided by the cloudlet group instead of the cloud. Based on the group-buying auction, the more users bid for the digital contents, the lower price they get. Interestingly, lowering down the price will not sacrifice the profit of the cloudlet group; on the contrary, it will maximize the expected profit of the cloudlet group. The members of the cloudlet group will cooperate with each other to maximize their benefits. We also analyze the expected profit of the cloudlet group and the rationality of the group-buying auction in our system.

3) We build a theoretical model to analyze the energy saving of mobile devices due to the use of the cloudlet group. We demonstrate that the mobile devices can effectively cut down the energy consumption with the service of the cloudlet group.

To the best of our knowledge, our solution makes the first effort on employing the supply chain and group-buying auction to provide efficient mobile cloud services. Though in this paper we focus on the design and analysis of the supply chain and group-buying auction approaches but do not go into the technical details about how to realize them, we believe that our system model can shed some light on providing practical mobile cloud services in the real world.

The rest of this paper is organized as follows: In Section II, we describe the system components of community clinic and analyze the energy consumptions. In Section III, we analyze the profit of the cloud in the bipartite supply chain and tripartite supply chain. In Section IV, we propose a real-time group-buying auction for the cloudlet group to promote its service to the mobile users. In Section V, we detail the performance of the community clinic. In Section VI, we briefly introduce some prior related work. Lastly, in Section VII we give the conclusion of the community clinic scheme.

II. System Model

For the sake of lessening the burden of the cloud to distribute digital contents, we propose a new system model that introduces the cloudlet group component between the cloud and mobile users (Fig. 3). To compare the difference between the system with and without the cloudlet group, we denote the system model with the cloud and mobile users as bipartite model (Fig. 1), and that with the cloud, the cloudlet group and mobile users as tripartite model (Fig. 3). We describe the system components of the two models and analyze their energy consumptions.

A. Bipartite Model

The bipartite model is composed of the cloud and mobile users which are described as following:

- *Cloud*: The cloud is an entity with abundant digital contents. The access to a digital content is controlled by its digital right license. When a mobile user wants to view a digital content, she has to purchase the corresponding digital right license before she can access the digital content;
- *Mobile users*: Mobile users are the consumers of digital contents. They must own the digital right licenses before they can access the digital contents; otherwise, they could not view the corresponding digital contents.

The procedure that a mobile user purchases a digital content from the cloud directly has two rounds, one is the purchasing round in which the mobile user searches the resources on the cloud, chooses the interested digital content, and pays the cloud for the digital content; the other is the downloading round, in which the mobile user downloads the digital content and corresponding digital right license from the cloud.

We further analyze the energy consumption of the mobile device in the process of acquiring the digital content under the bipartite model. To simplify our discussion, all the variables used in this paper are in the meaning of average. We believe this simplification does not fundamentally alter the results obtained in the paper. In this model, we assume that purchasing the digital content requires I_C instructions from the cloud, and the mobile device can process M instructions per second. Then it will take $\frac{I_C}{M}$ seconds for the purchasing round. We further assume that the available bandwidth of the mobile device when it connects to the cloud is B_C bytes per second. It takes $\frac{D}{B_C}$ seconds to download D bytes of digital content from the cloud. The total amount of the energy to acquire the digital contents from the cloud, E_C, is:

$$E_C = P_I \frac{I_C}{M} + P_D \frac{D}{B_C}. \tag{1}$$

Here P_I and P_D are the power consumptions of the mobile device for processing instructions and downloading the digital contents, respectively.

B. Tripartite Model

The tripartite model introduces a new entity, cloudlet group, between the cloud and mobile users. The cloudlet group is defined as follows:

- *Cloudlet group*: A cloudlet group is comprised of a number of cloudlets which are geographically close to each other and well inter-connected in a peer-to-peer way. Each cloudlet is able to provide services to proximate mobile devices efficiently. The digital contents can be firstly pre-downloaded from the cloud to the cloudlet, and then distributed to the mobile users. The cloudlet group can further meet the diversified demands of mobile users by providing enough resources to accommodate various digital contents.

This model describes an alternative scenario that the mobile users can purchase the digital contents via the cloudlet group. The cloudlet group firstly pre-downloads the digital contents and corresponding digital right licenses from the cloud. When a mobile user wants to purchase a digital content, she can acquire the digital content and corresponding digital right license from the cloudlet group if the digital content can be found from the cloudlet group. Otherwise, she will purchase the digital content from the cloud directly.

Since mobile users are rational, they prefer to purchase the QoS-guaranteed digital contents with low price from the cloud or cloudlet. To motivate the mobile users to use the cloudlet group to acquire digital contents, in the tripartite model the cloudlet group offers a low price service through a group-buying auction mechanism. Within the cloudlet group, there is a cloudlet leader that manages all the members' information. This cloudlet leader is called as *auction information publisher* (AIP) as it collects all the auction information in this cloudlet group and publicizes such information to the mobile users when they want to purchase the digital contents. For a valid group-buying auction, it is initiated by a cloudlet and lasts for a given period of time, and it can be accessed by mobile users via the nearest cloudlet within the same cloudlet group. We call the cloudlet that initiates the group-buying auction as *auction sponsor* (AS) and the cloudlet that assists the mobile user to link up to the auction sponsor as *auction assistant* (AA). The process that a mobile user purchases a digital content from the cloudlet group is as follows: The mobile user first connects to the cloudlet group via the nearest available cloudlet and then visits the AIP and reviews the information about the digital contents in this cloudlet group. After the mobile user selects an interested digital content, she will link up to the AS that owns the digital content. After that, the mobile user can make a bid for the digital content. If the mobile user bids the digital content successfully, the AS will deliver the digital content to her device. If the mobile user directly connects to the AS, she will obtain the digital content and its digital right license directly; otherwise, if the mobile user connects to an AA, she will obtain the digital content together with its digital right license from the AS via the AA.

Similar to the bipartite model, we could analyze the energy consumption of the mobile device in the process of group-buying auction under the tripartite model. We still use M as the mobile device's instruction processing rate. Assume that purchasing the digital content requires I_{CG} instructions from the AS, the available bandwidth between the mobile device and AS is B_{CG} bytes per second. In order to download D bytes of digital content from the AS, the energy consumption of the mobile device could be calculated as:

$$E_{CG} = P_I \frac{I_{CG}}{M} + P_D \frac{D}{B_{CG}}. \tag{2}$$

P_I and P_D are the same as those in the bipartite model.

We compare the energy consumption of the mobile device for viewing the video programs under these two models. As the energy consumed by the mobile device to process the instructions of purchasing the videos from either the cloud or AS is negligible in contrast to the energy consumed by downloading the videos, the difference of the energy consumption in the purchasing process is close to zero, that is, $P_I \frac{I_{CG}}{M} - P_I \frac{I_C}{M} \approx 0$. The ratio of the energy saving could be

$$\frac{E_C - E_{CG}}{E_C} \approx 1 - \frac{B_C}{B_{CG}}. \tag{3}$$

As the cloudlet group supports the mobile device with a high data rate connection, we denote g to be the bandwidth gain ratio between B_{CG} and B_C, i.e., $g = \frac{B_{CG}}{B_C}$. Then, Eq. (3) will be

$$\frac{E_C - E_{CG}}{E_C} \approx 1 - \frac{1}{g}. \tag{4}$$

Considering that the videos pre-downloaded to the cloudlet group may not always meet the mobile user's request, the ratio of energy saving should be adjusted by a hit ratio α, which denotes the probability that the videos on the cloudlet group could meet the mobile users' requests, that is, the ratio of energy saving of the cloudlet group, η_T, should be:

$$\eta_T = \alpha(1 - \frac{1}{g}). \tag{5}$$

From Eq. (5), we can see that the cloudlet group can assist the mobile device to save its energy consumption by providing high hit ratio α and bandwidth gain ratio g. In Section V we further conduct numerical experiments to demonstrate its performance improvement on energy saving.

III. The Supply Chain Model

In this section, we formulate the bipartite model and tripartite model as the bipartite supply chain and tripartite supply chain, respectively and analyze the profit of the cloud on these two supply chains based on the classic *Newsvendor problem model* [4]. For the bipartite supply chain, the cloud plays dual roles as both the supplier and retailer, the mobile user who purchases the digital content plays as the customer. Different from the bipartite supply chain, in the tripartite supply chain, the cloud only plays a sole role as the supplier to provide the digital contents to the cloudlet group, and the cloudlet group acts as the retailer to sell the digital contents to the mobile users who act as the customers. We compare the profit of the cloud before and after the cloudlet group participates in the supply chain and draw the conclusion that the cloudlet group can enlarge the profit of the cloud. To simplify our supply chain model, we analyze the profit of the cloud based on the process of selling one video programm through these two supply chains.

A. Bipartite Supply Chain

Before discussing the bipartite supply chain, we briefly describe the Newsvendor problem. The Newsvendor problem is to find the optimal order quantity of newspapers for the newsvendor to maximize the expected average profit in the condition that the demand distribution and cost parameters are known. Mathematically, assuming that the order quantity of the newspapers is Q, and the uncertain demand of the newspaper is a random variable D defined by the demand distribution density function $f(D)$. If the newspapers are over-ordered, the unsold newspapers will have to be thrown away or sold as scrap papers at a very low price at the end of the day, that is, $min(Q, D)$ units are sold and $(Q - D)^+$ units are residual (Here, a^+ is defined as $max(a, 0)$). With the per-unit salvage value of residual newspapers S where the salvage value defines the residual value of unsold newspapers, the value of the salvaged newspapers is $S \cdot (Q - D)^+$. If the newspapers are not enough ordered, some customers will be disappointed at the unmet demands for the newspaper, that is, $min(Q, D)$ units are sold and $(D - Q)^+$ units are unmet. We denote G as the per-unit goodwill cost of newspapers, then the goodwill cost for all unmet newspapers is $G \cdot (D - Q)^+$. Let P and C denote the per-unit price and cost of newspapers respectively, then the profit of the newsvendor is calculated as follows:

$$\Pi(Q) = P \cdot min(Q, D) + S \cdot (Q - D)^+ - G \cdot (D - Q)^+ - CQ. \tag{6}$$

Generally speaking, most video programs are seasonal goods and their popularity (click-through rate) declines dramatically after the season is passed. With the reason that the sale of video programs can be controlled by selling their digital right licenses and each digital right license cannot be duplicated, the retailer must predict the demands of the video programs and order a proper quantity of digital right licenses to maximize its profit. For the bipartite supply chain model, the cloud plays as the supplier and retailer to provide the mobile video service and the mobile users play as the consumers. Based on the Newsvendor problem model, we can get the profit of the cloud. Let P denote the per-unit price of videos, S is the per-unit salvage value of residual digital right license, G is the per-unit goodwill cost for the unmet digital right license, Q_B is the quantity of the videos ordered by the cloud and D is a random variable that represents the mobile users' demand distribution for the videos. The cost for the cloud to provide the mobile video service includes two parts: one part is the expense on purchasing the digital right licenses of the video, denoted as C_L; the other part is the system cost for distributing the videos from the cloud to the mobile devices, denoted as $C_C(\mu)$ where μ is the mean of mobile users' demand. Note that the system cost is mainly spent on the deployment of system infrastructure and its operation cost such as power draw. This cost is closely related to the mobile users' demand. When the users' demand exceeds the capacity of the cloud provider, this cost will become higher as the mobile users have to spend more time on downloading the videos. Moreover, though the cloud, playing as the supplier and retailer at the same time, can have unlimited digital right licenses and will leave no residual digital right licenses in the end, its service capacity is capped by its network bottleneck between the cloud and mobile users. As a result, the cloud will also suffer a goodwill loss G for per-unit unmet video and the profit of the cloud in the bipartite supply chain, $\Pi^B(Q_B)$, is

$$\Pi^B(Q_B) = P \cdot min(Q_B, D) - G \cdot (D - Q_B)^+ - (C_C(\mu) + C_L)Q_B. \tag{7}$$

We assume that the service capacity of the cloud is to satisfy the service orders up to $\widetilde{Q}$. If the orders are more than $\widetilde{Q}$, that is, the users' demand exceeds the capacity of the cloud, the system cost will be much higher. Thus, $C_C(\mu)$ can be denoted as a segment function with the mean of demand μ and the capacity of the cloud $\widetilde{Q}$, that is,

$$C_C(\mu) = \begin{cases} C_C & , \quad \mu \le \widetilde{Q}; \\ C_C \cdot H(\mu/\widetilde{Q}) & , \quad \mu > \widetilde{Q}. \end{cases} \tag{8}$$

Here $H(\mu/\widetilde{Q})$ is the cost gain factor when the mean of mobile users' demand exceeds the capacity of the cloud. In Eq. (7), the profit of the cloud in the bipartite supply chain, $\Pi^B(Q_B)$, is relative to order quantity Q_B. We could acquire the optimal order quantity to maximize the the profit of the cloud by the following proposition:

Proposition 1: The optimal order quantity Q_B^* and the maximum expected profit of the cloud $E\Pi^B(Q_B^*)$ should be

$$Q_B^* = F^{-1}(\frac{P+G-C_C(\mu)-C_L}{P+G}), \tag{9}$$

$$E\Pi^B(Q_B^*) = (P+G)\int_{D=0}^{Q_B^*} Df(D)dD - \mu G. \tag{10}$$

Here $f(D)$ and $F(D)$ are density function and cumulative distribution function of demand D, $x = F^{-1}(y)$ is the inverse function of y=F(x), and the mean of demand $\mu = \int_{D=0}^{\infty} Df(D)dD$.

Proof: We assume that

$$C_B(Q_B) = (P+G-C_C(\mu)-C_L)(D-Q_B)^+ - (C_C(\mu)+C_L)(Q_B-D)^+,$$

then

$$\Pi^B(Q_B) = (P-C_C(\mu)-C_L)D - C_B(Q_B).$$

The first derivative of $\Pi^B(Q_B)$ is

$$\frac{d\Pi^B(Q_B)}{dQ_B} = -\frac{dC_B(Q_B)}{dQ_B}.$$

According to the properties of the cumulative distribution, the expected function of $C_B(Q_B)$, denoted as $EC_B(Q_B)$, is

$$\begin{aligned} EC_B(Q_B) &= \int_{D=0}^{\infty} C_B(Q_B)f(D)dD \\ &= (P+G-C_C(\mu)-C_L)\int_{D=Q_B}^{\infty}(D-Q_B)f(D)dD \\ &\quad +(C_C(\mu)+C_L)\int_{D=0}^{Q_B}(Q_B-D)f(D)dD. \end{aligned}$$

In order to obtain the optimal order quantity Q_B^*, we calculate the first derivative of $EC_B(Q_B)$ and set it to zero:

$$\begin{aligned} \frac{dC_B(Q_B)}{dQ_B} &= (P+G)F(Q_B) - (P+G-C_C(\mu)-C_L) \\ &= 0. \end{aligned}$$

Then, we can get:

$$Q_B^* = F^{-1}(\frac{P+G-C_C(\mu)-C_L}{P+G}).$$

As the second derivative of $EC_B(Q_B)$ is

$$\begin{aligned} \frac{d^2EC_B(Q_B^*)}{d(Q_B^*)^2} &= (P+G)f(Q_B^*) \\ &\ge 0, \end{aligned}$$

and the optimal order quantity Q_B^* is

$$Q_B^* = F^{-1}(\frac{P+G-C_C(\mu)-C_L}{P+G}),$$

then,

$$\begin{aligned} E\Pi^B(Q_B^*) &= EC_B(Q_B) + (P-C_C(\mu)-C_L)\int_{D=0}^{\infty} Df(D)dD \\ &= (P+G)\int_{D=0}^{Q_B^*} Df(D)dD - \mu G. \end{aligned}$$

■

B. Tripartite Supply Chain

In what follows, we consider the profit of the cloud in the tripartite supply chain. As we have mentioned in Section II.B, in the tripartite model, both the cloud and cloudlet group can sell the videos to the mobile users. We consider the scenario that the users' demand on a video exceeds the capacity of the cloud in which the service of the cloudlet group is critical. Let Q_T ($Q_T > \widetilde{Q}$) be the quantity units of the videos consumed by the mobile users in the tripartite supply chain model, among which the quantity units Q, where $Q < \widetilde{Q}$, are directly provided by the cloud and the remaining quantity units $Q_T - Q$ are provided by the cloudlet group. For the videos directly provided by the cloud, the total cost can be calculated as $(C_C + C_L)Q$ since $C_C(\mu)$ equals to C_C when $Q < \widetilde{Q}$. For the remaining videos that are provided by the cloudlet group, they are first purchased by the cloudlet group and then sold to the mobile users. Thus, the cloudlet group needs to pre-download one unit of the video and $Q_T - Q$ units of the digital right licenses from the cloud, which induce the total cost of $C_C + (Q_T - Q)C_L$. Besides, the cost for the cloudlet group to distribute the video from AS to each end user, $C_{CG}(\mu)$, should also be considered. By introducing the cloudlet group, the capacity to serve the mobile users in the tripartite model will be enlarged by the bandwidth gain ratio g compared to that in the bipartite model, that is, the service capacity of the tripartite model is $g\widetilde{Q}$. Then the cost for the cloudlet group to distribute per-unit video could be expressed as:

$$C_{CG}(\mu) = \begin{cases} C_{CG} & , \quad \mu \le g\widetilde{Q}; \\ C_{CG} \cdot H(\mu/g\widetilde{Q}) & , \quad \mu > g\widetilde{Q}. \end{cases} \tag{11}$$

In this paper we only consider the scenario that the mobile users' demand does not exceed the service capacity of the cloud with the cloudlet group, that is, $C_{CG}(\mu) = C_{CG}$. The cost for the cloudlet group to distribute the videos to all mobile users is $(Q_T - Q) \cdot C_{CG}$. As we have mentioned, the cloudlet group should order a proper quantity of digital right licenses of the video from the cloud in advance so that it can sell the video to the end users. There are several reasons why the cloudlet group should do this: (1) The digital right license has its value and will not be ordered freely; (2) The cloud will offer the video to the cloudlet group with a lower wholesale price if more quantity units are ordered; (3) The residual digital right licenses could not be returned to the cloud with its original price because it will reduce the profit of the cloud. Thus, the cloudlet group must order a proper quantity of digital right licenses to maximize its profit. The unsold digital right licenses are of salvage value S which is much less than the lowest wholesale price. The cloudlet group also suffers goodwill cost G when the order quantity cannot totally meet the mobile users' demand. With the same variables P defined in the bipartite supply chain, the profit of both the cloud and cloudlet group in the tripartite supply chain is:

$$\begin{aligned}\Pi^T(Q_T) &= P\cdot min(Q_T, D) + S\cdot(Q_T - D)^+ - G\cdot(D - Q_T)^+ \\ &\quad -(C_C + C_L)Q - (C_{CG} + C_L)(Q_T - Q) - C_C. \quad (12)\end{aligned}$$

For the tripartite supply chain model, we have the following proposition:

Proposition 2: The optimal order quantity Q^*_T and the maximum profit of both the cloud and cloudlet group, $E\Pi^T(Q^*_T)$, should be

$$Q^*_T = F^{-1}(\frac{P + G - C_{CG} - C_L}{P + G - S}), \quad (13)$$

$$E\Pi^T(Q^*_T) = (P+G-S)\int_{D=0}^{Q^*_T} Df(D)dD - \mu G + QC_{CG} - (Q+1)C_C. \quad (14)$$

Proof: We assume that

$$C_T(Q_T) = (P + G - C_{CG} - C_L)(D - Q_T)^+ - (C_{CG} + C_L - S)(Q_T - D)^+,$$

then

$$\Pi^T(Q_T) = (P - C_{CG} - C_L)D - C_T(Q_T) - C_C.$$

The first derivative of $\Pi^T(Q_T)$ is

$$\frac{d\Pi^T(Q_T)}{dQ_T} = -\frac{dC_T(Q_T)}{dQ_T}.$$

According to the properties of the cumulative distribution, the expected function of $C_T(Q_T)$, denoted as $EC_T(Q_T)$, is

$$\begin{aligned}EC_T(Q_T) &= \int_{D=0}^{\infty} C_T(Q_T)f(D)dD \\ &= (P + G - C_{CG} - C_L)\int_{D=Q_T}^{\infty}(D - Q_T)f(D)dD \\ &\quad +(C_{CG} + C_L - S)\int_{D=0}^{Q_T}(Q_T - D)f(D)dD.\end{aligned}$$

In order to obtain the optimal order quantity Q^*_T, we calculate the first derivative of $EC_T(Q_T)$ and set it to zero:

$$\begin{aligned}\frac{dC_T(Q_T)}{dQ_T} &= (C_{CG} + C_L - S)F(Q_T) - (P + G - C_C - C_L)(1 - F(Q_T)) \\ &= 0.\end{aligned}$$

Then, we can get:

$$Q^*_T = F^{-1}(\frac{P + G - C_{CG} - C_L}{P + G - S}).$$

As the second derivative of $EC_T(Q_T)$ is

$$\begin{aligned}\frac{d^2EC_T(Q^*_T)}{d(Q^*_T)^2} &= (P + G - S)f(Q^*_T) \\ &\geq 0,\end{aligned}$$

and the optimal order quantity Q^*_T is

$$Q^*_T = F^{-1}(\frac{P + G - C_{CG} - C_L}{P + G - S}),$$

then,

$$\begin{aligned}E\Pi^T(Q^*_T) &= EC_T(Q_T) + (P - C_{CG} - C_L)\int_{D=0}^{\infty} Df(D)dD + C_C \\ &= (P + G - S)\int_{D=0}^{Q^*_T} Df(D)dD - \mu G + QC_{CG} - (Q+1)C_C.\end{aligned}$$

■

Comparing the expected profit of the bipartite supply chain (Eq. (10)) with that of tripartite supply chain (Eq. (14)), we can obtain the difference between the two supply chains:

$$\begin{aligned}\Delta_{E\Pi} &= E\Pi^T(Q^*_T) - E\Pi^B(Q^*_B) \\ &= (P + G)\int_{D=Q^*_B}^{Q^*_T} Df(D)dD - S\int_{D=0}^{Q^*_T} Df(D)dD + QC_{CG} \\ &\quad -(Q+1)C_C.\end{aligned}$$

As $\Delta_{E\Pi}$ is the surplus profit of the tripartite supply chain in contrast to the bipartite supply chain, it can be shared between the cloud and cloudlet group. If the cloudlet group only takes partial surplus profit, i.e., $\delta\Delta_{E\Pi}$ where $0 < \delta < 1$, the cloud will also benefit from the tripartite supply chain because the expected profit of the cloud will be $E\Pi^B(Q^*_B) + (1 - \delta)\Delta_{E\Pi}$, which shows that the cloud obtains more profit in the tripartite supply chain. From the discussion above, we can see that the cloudlet group could earn extra income in the tripartite supply chain, which can motivate multiple parties, e.g., home gateway providers and cloud providers, to deploy cloudlet groups.

IV. Real-time Group-Buying Auction in Cloudlet Group

The previous section sums up that both the cloud and cloudlet group would get more profits in the tripartite supply chain when the order quantity increases. In this section, we propose a real-time group-buying auction for the cloudlet group to promote its service to the mobile users. Based on this strategy, the more videos are sold, the lower price it is. The lower price would consequently attract more and more users to choose the service provided by the cloudlet group.

A. Real-time Group-buying Auction

The group-buying auction [6], [7] is a dynamic pricing mechanism which outperforms the fixed price mechanism in the scenario with economies of scale, because the group-buying auction can automatically set up a higher price for a product when its unit cost increases and a lower price for the same product when its unit cost decreases. It is a variant of the double auction in which the trading price is affected by both the seller's offer and the buyer's bid. The group-buying auction process has two rounds, offer round and bidding round. In the offer round, the vendor sets up a group-buying auction for a product with quantity N, price curve Q and auction period T which are open to all bidders. In the bidding round, the bidders bid the goods orderly based on their arrival times, and the auction price will change in accordance with the price curve as the number of the successful bidders increases. The auction ends

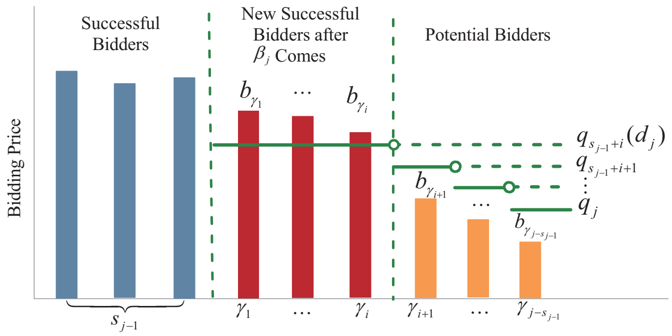


Fig. 4. The process of real-time group-buying auction.

when the amount of successful bidders reaches N or the auction time T expires. The successful bidders will acquire the products and the final auction price becomes the final deal price for all the bidders. However, this group-buying auction should not be directly adopted by the cloudlet group to promote its service to the mobile users due to the reason that it does not make the product available in real-time, that is, the mobile user has to wait for the end of the group-buying auction to acquire the product. Otherwise, different bidders may obtain the product with different deal prices since their bids success at different auction prices.

To deal with this issue, we design the *real-time group-buying auction* which allows each successful bidder to obtain the video in real-time with identical deal price. We define the mobile users in the real-time group-buying auction as two types of bidders: *successful bidder* and *potential bidder*. If a mobile user bids a video successfully, she becomes a successful bidder. She can download the video together with its digital right license and view the video immediately. If the mobile user's bidding is unsuccessful, she has two choices: one choice is to quit the auction, another is to stay in the auction until the auction price of the video turns to be no higher than the bidding price. When the mobile user chooses to stay in the group-buying auction and waits for a lower auction price, she becomes a potential bidder. She can download the video first and acquire the corresponding digital license later when she becomes a successful bidder, for the sake that she can view the video as soon as possible. We denote that the price curve Q is $Q = (q_1, q_2, \cdots, q_N)$, where q_i $(1 \leq i \leq N)$ is the auction price for the ith copy of the video and $q_1 \geq q_2 \geq \cdots \geq q_N$. Let the jth $(j \leq N)$ mobile user be β_j, her bidding price and deal price are b_j and d_j. We denote the number of the successful bidders are s_{j-1} before the jth mobile user arrives (assuming that $s_0 = 0$). The lists of the successful bidders and the potential bidders are defined as Λ and Γ, the orders of the bidders in successful bidders list and potential bidders list are determined by the deal price and the bidding price, respectively. Then the real-time group-buying auction process can be described as follows (Fig. 4):

1: The AS starts a group-buying auction with initial auction price q_1, quantity N and auction time T, and waits for mobile users to bid the video during the active auction time.
2: Suppose there are already s_{j-1} successful bidders in Λ and $j-1-s_{j-1}$ potential bidders in Γ. When a new mobile user β_j comes to bid the video with the bidding price b_j, the mobile user β_j will be inserted in Γ in descending order according to the bidding price b_j.
3: The AS continues comparing each potential bidder's bidding price with each auction price in the price curve starting from the last bidder $\gamma_{j-s_{j-1}}$ in Γ until it finds the first bidder γ_i whose bidding price b_{γ_i} is not less than the auction price $q_{s_{j-1}+i}$ in the price curve.
4: If such bidder γ_i exists, all the potential bidders $(\gamma_1, \gamma_2, \cdots, \gamma_i)$ become the successful bidders and are inserted into Λ according to the auction price $q_{s_{j-1}+i}$. The potential bidders list turns to be $\Gamma = (\gamma_{i+1}, \cdots, \gamma_{j-s_{j-1}})$.
5: When the auction ends, the AS refunds $d_j - d$ (d is the final deal price) to the jth successful bidder for the fair treatment.

The process of real-time group-buying auction can be illustrated more clearly by the following examples:

Example 1: Assume that the price curve is $Q = (100, 90, 90, 85, 80)$, and the AS has five pieces of license. The AS can attract five bidders to purchase the digital contents, and the bidders bid the digital contents one by one with the price (85, 100, 90, 90, 80). The detail process of bidding (Table I) is as follows:

TABLE I
THE PROCESS OF GROUP-BUYING DEMO

Auction price	Bidder	Bidding price	Successful bidder list	Deal price	Potential bidder List
100	$\beta 1$	85	$\emptyset$	-	$\beta 1$
100	$\beta 2$	100	$\beta 2$	100	$\beta 1$
90	$\beta 3$	90	$\beta 2$, $\beta 3$	90	$\beta 1$
90	$\beta 4$	90	$\beta 2$, $\beta 3$, $\beta 1$, $\beta 4$	85	$\emptyset$
80	$\beta 5$	85	$\beta 2$, $\beta 3$, $\beta 1$, $\beta 4$, $\beta 5$	80	$\emptyset$

1) The AS starts group-buying auction and sets the initial auction price as 100 based on the price curve;
2) The first bidder $\beta 1$ arrives and his bidding price is $b_1 = 85$, and he becomes a potential bidder because his bidding price is less than the current auction price $(b_1 < q_1)$. Then, there are no successful bidders ($s_1 = 0$ and $\Lambda = \emptyset$) but one potential bidder $\beta 1$ ($\Gamma = (\beta 1)$);
3) The second bidder $\beta 2$ comes with the bidding price $b_2 = 100$, and β_2 is inserted in Γ based on the bidding price and $\Gamma = (\beta_2, \beta_1)$. The second bidder is a successful bidder and the deal price of $\beta 2$ is $d_2 = 100$. Still there are no successful bidders ($s_2 = 1$, $\Lambda = (\beta 2)$) but one potential bidder $\beta 1$ ($\Gamma = \{\beta 1\}$);
4) The third bidder $\beta 3$ joins the auction with bidding price $b_3 = 90$. β_3 is inserted to the potential bidder list ($\Gamma = (\beta_3, \beta_1)$). $\beta 3$ is a successful bidder through the above algorithm, and the deal price is $d_3 = 90$. $\beta 2$ and $\beta 3$ are successful bidders ($\Lambda = (\beta 2, \beta 3)$) while $\beta 1$ is still a potential bidder ($\Gamma = (\beta 1)$);
5) The forth one $\beta 4$ bids with the price $b_4 = 90$, then the potential bidder list turns to be $\Gamma = (\beta_4, \beta_1)$. Due to the auction price $q_4 = 85$, both $\beta 4$ and $\beta 1$ join the group of the successful bidders, and their deal price is $d_1 = d_4 = 85$. The first four bidders are successful bidders

($s_4 = 4$, $\Lambda = (\beta2,\beta3,\beta1,\beta4)$) and there is no potential bidder ($\Gamma = \emptyset$);

6) The fifth bidder $\beta5$ bids the digital content with the price $b_5 = 85$, and he becomes a successful bidder and his deal price is $d_5 = 80$;
7) The auction closes, the AS determines the final deal price as $q_5 = 80$. As the deal price for ($\beta2,\beta3,\beta1,\beta4,\beta5$) are (100, 90, 85, 85, 80), the AS will refund the spare payment (20, 10, 10, 5, 0) to these five bidders respectively;
8) The group-buying auction ends.

We can see that the bidder (either a new comer or a potential bidder) will become a successful bidder once her bidding price reaches the auction price. A successful bidder can obtain the product immediately, that is, she can watch the video in real-time. Besides, the AS will refund the extra payment to all the successful bidders when the final deal price is determined, which makes all the successful bidders purchase the video at the same price. Based on the price curve Q, the more the successful bidders are, the lower the final deal price is. This strategy can effectively motivate more and more mobile users to choose the service provided by the cloudlet group.

B. The Price Curve of the Auction Sponsor

The group-buying auction can lower the price of the video and attract more mobile users to choose the cloudlet group to purchase the video. However, lowering the price of the video may reduce the profit for the AS, even though much more copies of the video may be sold in the group-buying auction. Therefore, we need to determine a proper price curve to ensure that the AS can maintain a maximum expected profit even when the price of the video is lowered in the group-buying auction.

We model the determination of the price curve as a multi-stage game between the AS and the bidders. The sequence of events in this game is listed as follows:

1: The AS determines the video's quantity and picks a price curve;
2: The bidders offer their bidding prices for the video and become successful bidders or potential bidders;
3: The AS can achieve the optimal price curve and maximum profit when the multi-stage game reaches the equilibrium.

Now we discuss how the price curve is designed in detail. As mentioned, the price curve is denoted as a vector $Q = (q_1, q_2, \cdots, q_N)$ with $q_1 \geq q_2 \geq \cdots \geq q_N$, which is mainly determined by the cost of this video. The cost includes two parts: one part is the expense for the AS to distribute the video from the cloud, i.e., C_C; the other is the cost to sell the video from the cloudlet group to each mobile user, i.e., $C_{CG} + C_L$. The AS would evaluate the cost to distribute the video before bidding because such factors will affect the AS's profit significantly. Suppose that the cost to serve v successful bidders is c_v, that is, $c_v = (C_{CG} + C_L)v + C_C$. Then we can get $\frac{c_v}{v} - \frac{c_{v+1}}{v+1} \leq \frac{c_{v-1}}{v-1} - \frac{c_v}{v}$. This suggests that the marginal cost for each bidder is decreasing. We also assume that each bidder's bidding price is independent and drawn from a uniform distribution with the unit interval [0,1]. The AS needs to consider the scenario that, when the auction ends, there is a total of m mobile users who bid the video, among which v mobile users bid the video successfully and these v successful bidders' bidding prices are not less than the auction price q_v. When $m \geq N$, the probability of this scenario is $C_m^v q_N^{m-N} q_N!(1-q_v)^v/q_v!$, where $q_v! = q_v q_{v-1} \cdots q_1$; when $m < N$, the probability of the scenario is $C_m^v q_m!(1-q_v)^v/q_v!$. Then the expected profit for the AS is

$$E\Pi^{AS}(Q) = \begin{cases} \sum_{v=1}^{N} C_m^v q_N^{m-N} q_N!(1-q_v)^v(vq_v - c_v)/q_v! & , \quad m \geq N; \\ \sum_{v=1}^{m} C_m^v q_m!(1-q_v)^v(vq_v - c_v)/q_v! & , \quad m < N. \end{cases}$$

We maximize the expected profit for the AS $E\Pi^{AS}$, which is a nonlinear programming problem:

$$\begin{aligned} max \quad & E\Pi^{AS}(Q), \\ s.t. \quad & q_1 \geq q_2 \geq \cdots \geq q_N. \end{aligned} \tag{15}$$

We solve this programming problem using the Karush Kuhn Tucker (KKT) condition and get a price curve as the optimal solution [8]. The optimal solution is able to impel the multi-stage game to reach the equilibrium. Thus, we can obtain the following proposition:

Proposition 3: The optimal solution of the nonlinear programming problem in Eq. (15) reaches the unique subgame-perfect equilibrium for the real-time group-buying auction.

Proof: The nonlinear programming problem in Eq. (15) is equivalent to the following nonlinear programming problem:

$$\begin{aligned} min \quad & -E\Pi^{AS}(Q), \\ s.t. \quad & q_1 \geq q_2 \geq \cdots \geq q_N. \end{aligned} \tag{16}$$

To illustrate above nonlinear programming meeting the KKT condition, we need to prove that $-E\Pi^{AS}(Q)$ is a convex function.

When $m < N$, we assume that

$$E\Pi_v^{AS} = -q_m!(1-q_v)^v(vq_v - c_v)/q_v!. \tag{17}$$

The Hessian Matrix of $E\Pi_v^{AS}$ is

$$H(E\Pi_v^{AS}) = \begin{pmatrix} \frac{\partial^2 E\Pi_v^{AS}}{\partial q_1^2} & \cdots & \frac{\partial^2 E\Pi_v^{AS}}{\partial q_1 \partial q_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 E\Pi_v^{AS}}{\partial q_m \partial q_1} & \cdots & \frac{\partial^2 E\Pi_v^{AS}}{\partial q_m^2} \end{pmatrix}. \tag{18}$$

We can verify that the leading principle minor of $H(E\Pi_v^{AS})$ is not less than 0. Therefore, $H(E\Pi_v^{AS})$ is a positive-semidefinite matrix and $E\Pi_v^{AS}$ is a convex function [8], which causes $-E\Pi^{AS}(Q)$ to be a convex function. Therefore, the nonlinear programming problem in Eq. (15) satisfies the KKT condition and the solution of this nonlinear programming problem reaches the unique subgame-perfect equilibrium for the real-time group-buying auction.

Similar result could be proved when $m \geq N$. ■

From the above discussion, we achieve a price curve composed of a non-increasing sequence of auction prices, which is the optimal solution for Eq. (15). Based on such price curve, the AS can maximize its expected payoff from the real-time group-buying auction.

C. Cooperation or not in the Cloudlet Group

In the group-buying auction, AS needs to distribute the video and its digital right license to the mobile users, and the AA is assumed to help the AS complete this process. If the AA is willing to assist the AS to complete the auction, such auction is called "cooperation". However, an AA may also start a new auction for the same video, causing both cloudlets (AS and AA) to have the auctions for the same video at the same time. The two cloudlets will compete with each other and none of them is willing to be the AA of its competitor, which causes both of them not to cooperate with each other. Fortunately, our group-buying auction mechanism can avoid such competition and prompt the AA to cooperate with the AS, because the cooperation can raise the expected profit for every cloudlet and mobile user.

Assume that the cloudlet group has M homogeneous cloudlets and the expected profit of this cloudlet group is $E\Pi_{CG}$, then the expected profit of every cloudlet in this cloudlet group is $\frac{E\Pi_{CG}}{M}$, which implies that each cloudlet's expected profit is only related to the expected profit of its cloudlet group. We then estimate the expected profit of the cloudlet group under the condition whether the two cloudlets cooperate or compete with each other, respectively. We denote the two cloudlets as A and B, and the quantities of videos to be sold by A and B are G_A and G_B. Without loss of generality, we assume $G_A \geq G_B$. For the cloudlet A, the price curve is $Q^A = (q_1, q_2, \cdots, q_{G_A})$, where $q_1 \geq q_2 \geq \cdots \geq q_{G_A}$; the cost for cloudlet A to sell i units is c_i $(i = 1, 2, \cdots, G_A)$. In the same way, the price curve and the cost for cloudlet B are $Q^B = (q_1, q_2, \cdots, q_{G_B})$ where $q_1 \geq q_2 \geq \cdots \geq q_{G_B}$, and c_i $(i = 1, 2, \cdots, G_B)$. Then, the following proposition can be derived:

Proposition 4: Given that the expected profit of the cloudlet group, under the condition whether the two cloudlets cooperate or compete with each other, are $E\Pi_{CG}^{coop}$ and $E\Pi_{CG}^{comp}$ respectively, then $E\Pi_{CG}^{coop} > E\Pi_{CG}^{comp}$.

Proof: Both cloudlets A and B are homogenous, and they adopt the same price curve to maximize their own profits as they compete with each other. We combine the price curve of cloudlets A and B together to be a new vector Q' and arrange the elements of Q' in the descending order. For the case that the two cloudlets cooperate with each other, we achieve the price curve Q^* through solving Eq. (15).

When $m \geq G_A + G_B$ (m is the number of the mobile users bidding for the video), $Q' = (q'_1, q'_2, \cdots, q'_{G_A+G_B})$ and $Q^* = (q^*_1, q^*_2, \cdots, q^*_{G_A+G_B})$. Q^* reaches the unique subgame-perfect equilibrium, then $E\Pi^{AS}(Q^*) > E\Pi^{AS}(Q')$ according to Proposition 3. Cloudlets A and B purchase the video from the cloud twice and the cost is more than purchasing the video once, which leads to $E\Pi^{AS}(Q') > E\Pi_{CG}^{comp}$. Therefore, $E\Pi_{CG}^{coop} = E\Pi^{AS}(Q^*) \geq E\Pi^{AS}(Q') > E\Pi_{CG}^{comp}$.

When $m < G_A + G_B$, $Q' = (q'_1, q'_2, \cdots, q'_m)$ and $Q^* = (q^*_1, q^*_2, \cdots, q^*_m)$. Through Proposition 3, $E\Pi^{AS}(Q^*) > E\Pi^{AS}(Q')$. Cloudlets A and B purchase the same video twice as they compete with each other, which causes $E\Pi^{AS}(Q') > E\Pi_{CG}^{comp}$. Then, $E\Pi_{CG}^{coop} = E\Pi^{AS}(Q^*) \geq E\Pi^{AS}(Q') > E\Pi_{CG}^{comp}$. ■

From Proposition 4, we can see that cooperation is the better choice for the cloudlets, and it makes the cloudlet group get more profit, which also brings more expected profit for every cloudlet. It is also noted that the price curves of cloudlets A and B are $Q^A = (q_1, q_2, \cdots, q_{G_A})$ and $Q^B = (q_1, q_2, \cdots, q_{G_B})$ when they compete with each other. With $G_A \geq G_B$, we can get $q_{G_A} \leq q_{G_B}$, that is, the final deal price of cloudlet A is less than that of cloudlet B. For the cloudlet group with the two cloudlets cooperating with each other, the price curve is $Q^* = (q^*_1, q^*_2, \cdots, q^*_{G_A+G_B})$, and the final deal price is $q^*_{G_A+G_B}$. It is obvious that $q^*_{G_A+G_B} \leq q_{G_A} \leq q_{G_B}$, then the mobile users under the scenario that the two cloudlets cooperate with each other will get more surplus in the group-buying auction. Therefore, all AAs are willing to cooperate with the AS and the same type of group-buying auctions should not be initiated by multiple cloudlets at the same time.

V. Numerical Experiments

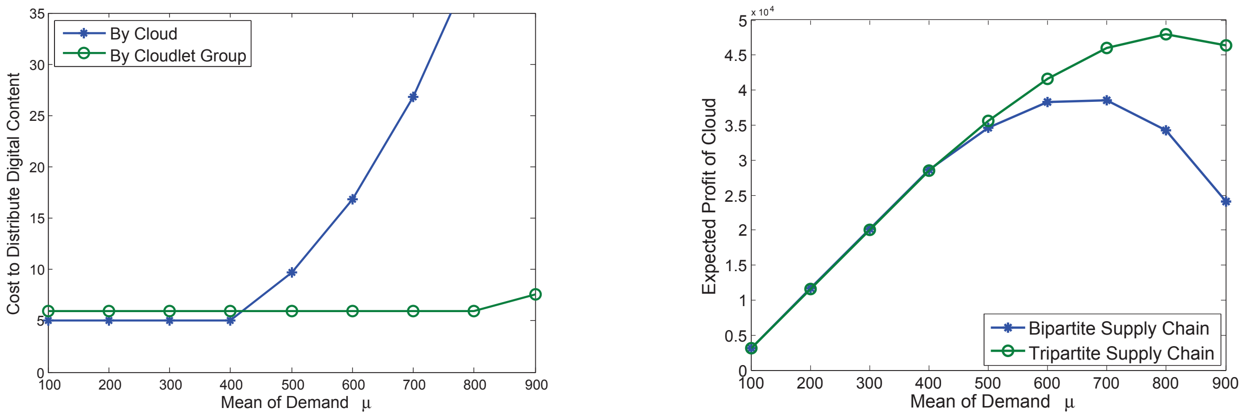
In this section, we will analyze the performance of the community clinic using numerical experiments, which includes economics analysis and energy analysis.

A. Economics Analysis

1) Supply Chain: We take the cumulative distribution function of demand D to be the normal distribution as an example to compare the profit of the cloud in the bipartite supply chain model and tripartite supply chain model. With $H(Q) = \int_{D=0}^{Q} Df(D)dD = \mu\Phi(x) - \sigma\phi(x)$ [9], we can evaluate the profit of the cloud easily. Here μ and σ are the mean and standard deviation of demand D, $\phi(x)$ and $\Phi(x)$ are the density function and cumulative distribution function of the standard normal distribution, and $x = \frac{Q-\mu}{\sigma}$.

We illustrate the cost of the cloudlet group C_{CG}, which includes the CAPEX (capital expenditure) and OPEX (operational expenditure), with an example. We consider an application scenario that a cloudlet group comprised of 100 cloudlets is deployed in an airport lounge for a total coverage of $20000m^2$. Each cloudlet unit is implemented by a WiFi-enabled PC "Lenovo ThinkCentre M78" (retail price: 499\$; power consumption: $300W$), which connects to a powerful switch "Cisco WS-C4506-E" (retail price: 3000\$; power consumption: $3000W$). With 3 years' service lifetime, the CAPEX for one cloudlet per day can be estimated as $(3000/100 + 499)/(365 \times 3) = 0.48\$$. The OPEX for one cloudlet is mainly the power consumption (average retail price: $< 0.1\$/kWh$ [10]), considering that the cloudlets are connected with each other through the switch. Then the OPEX for one cloudlet per day is around $(300W + 3000W/100) \times 24h \times 0.1\$/kWh = 0.79\$$. Supposing that a cloudlet sells 20 digital contents per day, the cost of the cloudlet group C_{CG} is nearly $(0.48\$ + 0.79\$)/20 \approx 0.06\$$, e.g., the cost is around 6 cents per day.

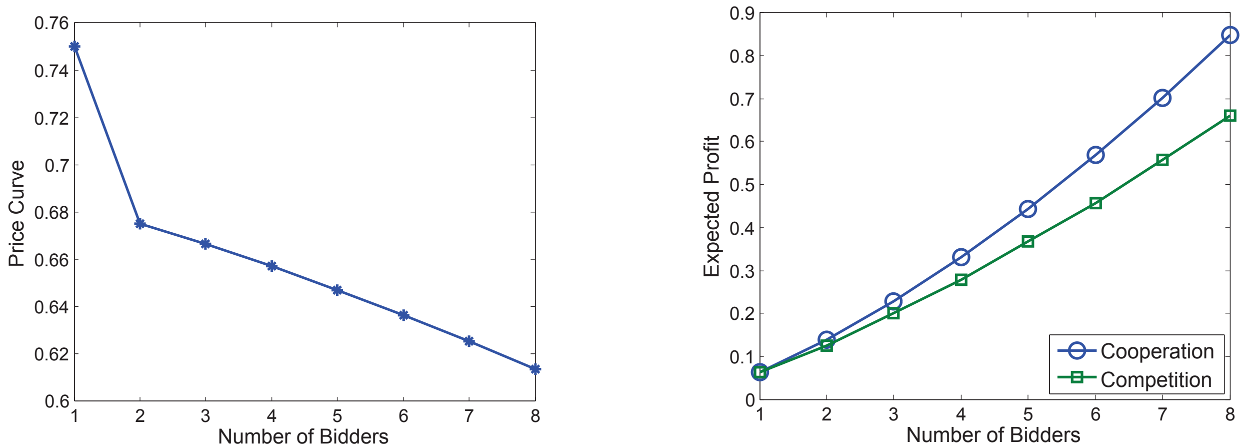
Then we deploy numerical experiments to compare the profit of the cloud in the bipartite model and tripartite model (Fig. 5). Let the price $P = 100$, the salvage value is $S = 10$, the goodwill cost is $G = 5$, the license cost is $C_L = 15$, the cost to distribute videos by the cloud and cloudlet group are $C_C = 5$ and $C_{CG} = 6$ (All these variables are in cents). Here



(a) Distribution Cost

(b) Expected Profit of Cloud

Fig. 5. The effect of users' demand on the cost to distribute unit digital content and the expected profit of the cloud. ($P = 100$, $S = 10$, $G = 5$, $C_L = 15$, $C_C = 5$, $C_{CG} = 6$, $g = 2$, $\delta = 0.5$, $\tilde{Q} = 400$, $Q = 0$ $\sigma = 20$)



(a) Price Curve

(b) Expected Profit of Cloudlet Group

Fig. 6. The effect of number of bidders on the price curve and the expected profit of the cloudlet group. ($c_1 = \frac{1}{2}$, $c_v - c_{v-1} = \frac{1}{v+1}$, $v = 2, \cdots, 8$)

we assume that $C_C < C_{CG}$ because the cost to distribute videos by the cloud is, due to the economies of scale, lower than that by cloudlet group. We denote the cost gain factor $G(x) = x^3$ when the mean of demand exceeds the capacity of the cloud and cloudlet group. The standard deviation of demand D is $\sigma = 20$. With the capacity of the cloud $\tilde{Q} = 400$, we consider the tripartite supply chain could support 2 times mobile users than the bipartite supply chain, that is $g = 2$. We also assume that all the digital contents are offered by the cloudlet group (i.e., $Q = 0$), and the cloud shares the surplus equally with the cloudlet group (i.e., $\delta = 0.5$). According to Eq. (8) and Eq. (11), we obtain the change of the cost to distribute unit digital content by the cloud and cloudlet group in Fig. 5(a). The distribution cost of the cloud increases rapidly as soon as the users' demand exceeds the cloud's capacity (when $\mu = 400$) and surpasses the cost for the cloudlet group to distribute videos (when $\mu = 500$).

The expected profit of cloud on two types of supply chains are also evaluated in Fig. 5(b). It shows that the expected profit of the cloud is nearly the same under the condition that the mobile users' demand is not more than the cloud's capacity (when $\mu < 400$). However, the cloud will earn more on the tripartite supply chain than on the bipartite supply chain when the users' demand exceeds the cloud's capacity (when $\mu > 400$). The numerical experiments reveal that the cloud will earn more with the participation of cloudlet group. It also shows that the expected profit of the cloud will decrease on the bipartite supply chain when $\mu > 600$, which suggests that the explosive growth of the mobile users has the system cost so large that the profit of the cloud actually decreases. The profit of the cloud on the tripartite supply chain also decreases when the mean of demand exceeds the capacity of the tripartite supply chain ($\mu > g\tilde{Q} = 800$).

2) Real-time Group-Buying Auction: From the nonlinear programming problem described in Eq. (15), we can see that the price curve and expected profit of AS are totally determined by the cost to serve the successful bidder. In order to satisfy the condition that the marginal cost to serve the successful bidder is diminishing, $\frac{c_v}{v} - \frac{c_{v+1}}{v+1} \le \frac{c_v}{v} - \frac{c_{v-1}}{v-1}$, ($v = 1, 2, \cdots$), we deploy the marginal cost as $\frac{c_v}{v} - \frac{c_{v+1}}{v+1} = \frac{1}{(v+1)(v+2)}$, $c_1 = \frac{1}{2}$ in our

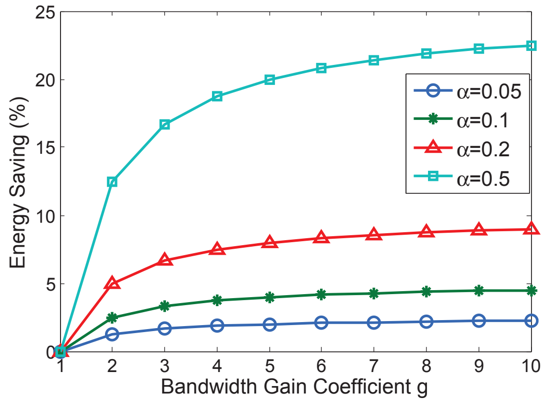


Fig. 7. Energy Saving (%) with $g \in [1, 10]$ and $\alpha = 0.05, 0.1, 0.2, 0.5$.

numerical experiments. By solving the nonlinear programming problem, the price curve and expected profit of AS are shown in Fig. 6. In Fig. 6(a), the price curve is decreasing as the number of the mobile users increases. This clearly shows that the price of the video becomes lower when more mobile users are bidding successfully. The lower price will reduce the mobile users' cost on acquiring the digital content, which motivates more mobile users to use the service of the cloudlet group. From the AS's expected profit curve shown in Fig. 6(b), we can see that, the AS's expected profit is growing as it serves more mobile users no matter the members of the cloudlet group cooperate or compete with each other, and the cloudlet group has incentives to provide more videos to the mobile users. However, cooperation can bring more profit to the AS than competition. The gap in the expected profit between cooperation and competition increases greatly with the increase of the mobile users. In general, the real-time group-buying auction could lower the price of the digital content and increase the benefit of the AS. Such strategy can have the mobile users regard the cloudlet group as the first choice if the requested video can be found in the cloudlet group. If the AS would like to maximize its profit, it should cooperate with other members of the cloudlet group.

B. Energy Analysis

According to Eq. (5), the cloudlet group can assist the mobile device to save its energy consumption by providing high hit ratio α and bandwidth gain ratio g. We deploy numerical experiments with $g \in [1, 10]$ and $\alpha = 0.05, 0.1, 0.2, 0.5$ and acquire the ratio of energy saving in Fig. 7. The bandwidth gain between the cloudlet group and mobile device makes the ratio of energy saving grow fast at beginning ($g = 2, 3$), then the rate of change is slow gradually as the bandwidth gain increases. That means only increasing the bandwidth between the cloudlet group and mobile device would not help much to save the energy consumption. At that time, we have to raise the hit ratio to meet the mobile users' requests on energy saving. Fig. 7 shows that the larger hit ratio it is, the more energy it will be saved. That is why we bring the cloudlet group into our design. The cloudlet group can offer a variety of videos which would satisfy the mobile users' requests as many as possible. If the cloudlet group is able to hit half of the the mobile users' requests ($\alpha = 0.5$), the mobile device still could save considerable energy even though the bandwidth gain ratio g is small ($\eta_T = 25\%$ and 35% when $g = 2$ and 3). From the above discussion, we could see that distributing video programmes through the cloudlet group will consume less energy than directly transmitting the videos (either download or streaming service) from the cloud through the WiFi/3G connection.

VI. Related Work

Recently, mobile cloud computing, as a combination of mobile computing and cloud computing, has been fiercely debated [11]. Several frameworks, including MAUI [12], Cuckoo [13], CloneCloud [14], ThinkAir [15], and Where-Store [16], have offloaded the tasks from mobile devices to the cloud. The cloud leverages effective resource allocation [17] to have the mobile application executed in the geographically distributed data centers [18], for example, CloudFront [19]. However, having the mobile application executed in the geographically distributed data centers enforces the cloud provider to build the massive date centers. As the investment of the massive data center becomes large economies of scale, minimizing the cost of a data center can achieve a high payback. The expenses that data centers cost go mainly to servers (45%), infrastructure (25%), power draw (15%), and networks (15%) [20]. One way to reduce the cost of the data center is to save energy consumption. ElasticTree [21], which has an effective network traffic pattern, can save up to 50% energy cost of the data center. Power saving was also considered in [22] where the proposed models decrease the total electricity cost of the data center with guaranteed quality of service. Decreasing the network traffic is another effective approach to cut down the cost of the data center. Inter-datacenter traffic was studied in [23] which reveals that up to 45% of total traffic goes through data center egress routers. This work motivates the researchers to minimize the cost on inter-datacenter traffic. Jetway [24] is one of the effective algorithms to minimize the expense of inter-datacenter's video traffic through optimal video flows in an online fashion.

All these approaches try their best to minimize the cost on power draw and networks, but the total cost on these two parts is only 30%. A more effective way is to build up less servers and infrastructure because the cost on servers and infrastructure is nearly 70%. Our approach leverages the cloudlet group to reduce the overall cost on building up data centers. In addition, we design an effective real-time group-buying auction to motivate more mobile users to use the service of the cloudlet group, which further lessens the burden of data centers.

VII. Conclusion

In this paper, we have proposed the community clinic scheme to economize the mobile cloud service cost under the condition that the service demands exceed the capacity of the cloud. The unique economic and technic characteristics of the scheme bring benefit to the cloud, the cloudlet and mobile users. We have firstly analyzed the energy consumption of the mobile device and found that the mobile devices can save more energy with the service of the cloudlet group. Then we have proven that the cloudlet group can assist to increase the profit of the cloud through modeling the system with and without the cloudlet group as two types of supply chains. Moreover, we present the real-time group-buying auction to attract more mobile users to be served by the cloudlet group with a lower price, and design the effective price curve for the unique subgame-perfect. The real-time group-buying auction also promotes the cooperation among the members of the cloudlet group and maximizes the expected profit for the cloudlet group. Numerical experiments are conducted to demonstrate the effectiveness of our tripartite model compared with the bipartite model. In general, the community clinic achieves a win-win-win outcome among the cloud, cloudlet group and mobile users.

Acknowledgment

This work was supported in part by Hong Kong GRF grants (PolyU 523207, PolyU 521312) and NSFC grant No. 61272463.

References

[1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2013-2018," *Cisco white paper*, 2014.

[2] A. Rath, S. Hua, and S. S. Panwar, "Femtohaul: Using femtocells with relays to increase macrocell backhaul bandwidth," in *IEEE INFOCOM*, 2010.

[3] M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies, "The case for vm-based cloudlets in mobile computing," *IEEE Pervasive Computing*, vol. 8, no. 4, pp. 14–23, 2009.

[4] A. V. Hill, *The newsvendor problem*. Clamshell Beach Press, 2011.

[5] J. Zhang, T. Xiong, and W. Lou, "Community clinic: Economizing mobile cloud service cost via cloudlet group," Technical Report, 2014. [Online]. Available: http://www.comp.polyu.edu.hk/˜csweilou/Publications/hkpu-comp-tr-2014-1-zhang.pdf

[6] J. Chen, X. Chen, and X. Song, "Bidder's strategy under group-buying auction on the internet," *IEEE Transactions on Systems, Man and Cybernetics. Part A. Systems and Humans*, vol. 32, no. 6, pp. 680–690, 2002.

[7] J. Chen, X. Chen, and X. Song, "Comparison of the group-buying auction and the fixed pricing mechanism," *Decision Support Systems*, vol. 43, no. 2, pp. 445–459, 2007.

[8] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[9] R. L. Winkler, G. M. Roodman, and R. R. Britney, "The determination of partial moments," *Management Science*, vol. 19, no. 3, pp. 290–296, 1972.

[10] U.S. Energy Information Administration, "Electric power monthly," *http://www.eia.gov/electricity/monthly/pdf/epm.pdf*, 2014.

[11] M. Satyanarayanan, "Mobile computing: The next decade," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 15, no. 2, pp. 2–10, 2011.

[12] E. Cuervo, A. Balasubramanian, D.-K. Cho, A. Wolman, S. Saroiu, R. Chandra, and P. Bahl, "Maui: Making smartphones last longer with code offload," in *ACM MobiSys*, 2010.

[13] R. Kemp, N. Palmer, T. Kielmann, and H. Bal, "Cuckoo: A computation offloading framework for smartphones," in *Mobile Computing, Applications, and Services*. Springer, 2012, pp. 59–79.

[14] B. G. Chun, S. Ihm, P. Maniatis, M. Naik, and A. Patti, "Clonecloud: Elastic execution between mobile device and cloud," in *ACM EuroSys*, 2011.

[15] S. Kosta, A. Aucinas, P. Hui, R. Mortier, and X. Zhang, "Thinkair: Dynamic resource allocation and parallel execution in the cloud for mobile code offloading," in *IEEE INFOCOM*, 2012.

[16] P. Stuedi, I. Mohomed, and D. Terry, "Wherestore: Location-based data storage for mobile devices interacting with the cloud," in *ACM Workshop on Mobile Cloud Computing & Services: Social Networks and Beyond*, 2010.

[17] M. Alicherry and T. Lakshman, "Network aware resource allocation in distributed clouds," in *IEEE INFOCOM*, 2012.

[18] F. Wang, J. Liu, and M. Chen, "Calms: Cloud-assisted live media streaming for globalized demands with time/region diversities," in *IEEE INFOCOM*, 2012.

[19] Amazon, "Amazon cloudfront," *http://aws.amazon.com/cloudfront*.

[20] A. Greenberg, J. Hamilton, D. Maltz, and P. Patel, "The cost of a cloud: Research problems in data center networks," *ACM SIGCOMM Computer Communication Review*, vol. 39, no. 1, pp. 68–73, 2008.

[21] B. Heller, S. Seetharaman, P. Mahadevan, Y. Yiakoumis, P. Sharma, S. Banerjee, and N. McKeown, "Elastictree: Saving energy in data center networks," in *NSDI*, 2010.

[22] L. Rao, X. Liu, L. Xie, and W. Liu, "Minimizing electricity cost: Optimization of distributed internet data centers in a multi-electricity-market environment," in *IEEE INFOCOM*, 2010.

[23] Y. Chen, S. Jain, V. Adhikari, Z. Zhang, and K. Xu, "A first look at inter-data center traffic characteristics via yahoo! datasets," in *IEEE INFOCOM*, 2011.

[24] Y. Feng, B. Li, and B. Li, "Jetway: Minimizing costs on inter-datacenter video traffic," in *ACM MM*, 2012.