Robust Resource Reservation in Virtual Wireless Networks

Ali Abbasi and Majid Ghaderi
Department of Computer Science, University of Calgary
{aabbasi, mghaderi}@ucalgary.ca

***Abstract*—In this paper, we study resource reservation in virtual wireless networks with the aim of minimizing the operational cost. With this regard, the main constraint facing the operator is that only limited information about future traffic demand is typically available to the operator. To address this issue, we investigate reservation policies that are robust to the worst-case traffic demand which fits the available information *i.e.*, the policies that minimize the worst-case expected operational cost.**

The problem is formulated for several resource reservation options that are commonly offered in practice. For each case, convexity of the problem is discussed and the its dual form is presented as a semidefinite program. While, semidefinite programs can be solved in polynomial time, the optimal closed-form reservation policies are obtained for several practical cases. Moreover, the worst-case cost of these policies are analytically compared to the expected cost of the algorithm that has full knowledge of the future demand.

The theoretical analysis is supplemented with numerical results to demonstrate the behavior of our algorithms in terms of cost in some example traffic scenarios.

I. Introduction

Over the past few years, due to proliferation of high-end hand-held devices *e.g.*, smartphones and tablets, the amount of data traffic on cellular networks has grown exponentially [1]. This trend is expected to continue in the coming years as the amount of traffic in 2019 is predicted to increase tenfold over 2014 [2]. To keep up with the traffic demand, mobile operators are constantly enhancing the cellular network capacity by moving towards denser deployment of base stations (BSs) [3] in addition to adopting more advanced communication techniques [4]. Doing so, fewer users are associated to each BS which in turn provide them with higher data rates.

Evolving towards dense small cell deployment of the cellular network imposes higher capital and operational expenditure on mobile operators. Furthermore, as cellular traffic changes from voice-dominated to data-dominated the revenue per bit transferred is decreasing at a fast rate [1]. This motivates designing new mechanisms to save and reduce the operational cost of cellular networks [5], [6]. Among these mechanisms, sharing the network infrastructure is particularly interesting [3], [6] as it allows to statistically multiplex multiple isolated and customized services or operators on the same physical infrastructure which results in better utilization of resources and hence reduced capital and operational expenditure.

The key enabling technology that could facilitate network sharing is *virtualization* [7], [8] which refers to providing an abstraction of hardware, computing, or spectrum resources of the network and slicing them to *virtual resources*. This abstraction mechanism allows a virtual operator to utilize the cellular network and provide wireless access service to its subscribers without becoming involved in the physical and operational details of the underlying infrastructure. However, this model presents new challenges that need to be addressed. Most importantly, as virtual operators typically provide users with customized or low-cost services, their operational cost should be low to make their service economical. Thus, the essential challenge is to find policies for acquiring virtual resources that lead to minimizing the monetary cost that is paid by the mobile virtual network operators (MVNOs) to the cellular Infrastructure Provider (InP). These policies would naturally depend on the resource pricing mechanisms that are implemented by the cellular InPs. In this paper, we consider a pricing mechanism where the cellular InP offers the option of reserving virtual resources at a cost that is lower than the cost of acquiring resources online. Reservation-based pricing mechanisms are common in other domains and are currently offered by many Infrastructure-as-a-Service (IaaS) providers including Amazon EC2 and IBM SmartCloud Enterprise.

To make accurate resource reservation, an MVNO needs to have some information about future traffic demand, which normally is not available in a live operational network. However, as highlighted in [5], [9], the cellular traffic exhibits periodic behavior. In [9], it is shown that a multi-order Markov model is able to predict the day-ahead aggregate traffic at per hour granularity with reasonably small mean squared error. Hence, we can assume that the MVNO has access to simple statistical information regarding the future traffic demand distribution *e.g.*, its mean and variance. Accordingly, the uncertainty in the future demand can be described by the set of traffic distributions that are compatible with such prior information [10]. A sensible approach for an MVNO to hedge against this uncertainty is to follow the so-called *ditributionally robust optimization* [10], [11] framework that considers the worst-case distribution. This way, the operator chooses the reservation policy that minimizes the worst-case expected cost of the operation. An interesting feature of

This work is supported in part by the Smart Applications on Virtual Infrastructure (SAVI) project funded under the National Sciences and Engineering Research Council of Canada (NSERC) Strategic Networks grant number NETGP394424-10.

this approach is that no assumption is made about the real distribution of traffic.

To the best of our knowledge, reservation-based resource acquisition has not been investigated in the context of mobile virtual networks. Most related to our work are [12]–[14]. In [12] reservation of computing instances in IaaS providers without any statistical information is investigated. The problem is shown to be a generalization of the well-studied ski-rental problem [15] which it admits to constant-factor approximation. The effect of various statistical information on the optimal solution to the ski-rental problem is investigated in [13], [14]. In contrast to the above works that allow the decision maker to buy the ski (reserve the required resources) at any point during the time frame of the operation, we allow the MVNO to reserve virtual resources only at the beginning of the time frame. This model is similar to the one employed in electricity markets [16] where a utility company can purchase electricity either from the day-ahead market or the real-time market. Consequently, the algorithms and performance guarantees presented in [12]–[14] do not apply to our problem.

Our contributions in this work can be summarized as follows:

- We present several virtual resource reservation models that are offered in practice. For each model, we formulate the minimax optimization problem that is associated with minimizing the worst-case expected cost of the system.
- Solving the dual problem associated with the worst-case expected cost of each model, we present the corresponding closed-form optimal reservation policy where the statistical information is limited to the mean and variance of the traffic distribution.
- We study the effect of having higher order statistics on the performance of the proposed policies and show that even having access to variance information in addition to the mean information could significantly improve the performance of the presented algorithms.

The paper is organized as follows. In Section II, the system model is introduced and the minimax reservation problem is formulated. In Sections III, IV, and V, the problem is solved for each of the considered reservation models. Section VI discusses extension of the model to higher order statistics. Sample numerical results are presented in Section VII. Finally, Section VIII concludes the paper.

II. System model and problem formulation

We consider a system consisting of a cellular InP and a MVNO. The InP owns a set of BSs covering the network and their backhaul connection to the Internet. The MVNO acquires virtual resources from the InP and provides service to its users. We focus on the operation of the MVNO on a single base station. Time is divided into equal size frames. The frame duration is chosen so that statistical characteristics of the traffic demand remain fairly consistent *e.g.*,an hour [9]. A frame is further divided into equal size times *e.g.*, 10-minute intervals.

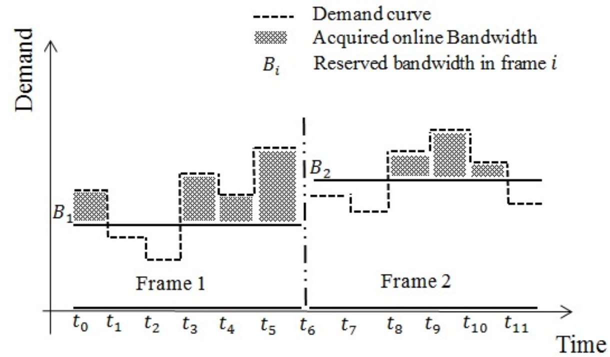

Fig. 1. Illustration of demand and reservation in each frame..

Reservation can only take place at the beginning of each frame. However, if the reserved bandwidth [1] does not satisfy the demand at a specific timeslot, the MVNO can acquire additional bandwidth at the online price. We assume that there exists a bound D on the traffic demand in every timeslot.

Since we assume the traffic distribution remains the same during a frame, the optimal reserved bandwidth is the same for all the timeslots of the frame. This model is exemplified in Figure. II. The figure demonstrates two frames of size 6 timeslots. As can be seen, in the first frame, the reserved bandwidth B_1 does not satisfy the demand in timeslots t_0-t_1, t_3-t_4, t_4-t_5, and t_5-t_6, thus, the operator is forced to purchase more bandwidth in these timeslots. In addition, demand is predicted to be higher in the next frame, thus, the reserved bandwidth (B_2) is larger in the second frame.

A. Reservation models

We consider three reservation models that are offered in practice:

- **Reservation with No Usage Fee (NUF)**. In this model, the virtual operator pays an upfront fee for the reserved bandwidth B. The usage up to this limit is free of charge. However, if additional bandwidth is required, it can be purchased at the online price. This model is similar to electricity market pricing where the utility company can procure the required electricity either in the day-ahead or real-time markets [16], [17]. Typically, the average price in the day-ahead market is lower than the real-time market. In this model, the cost incurred to satisfy demand x at a timeslot where bandwidth B is reserved in the corresponding frame is given by

$$\begin{aligned} cost^{NUF}(x, B, p_B, p_O) &= p_B B + p_O \max(x - B, 0) \\ &= p_B(B + \rho \max(x - B, 0)), \end{aligned} \tag{1}$$

  where p_B and p_O are the base (reservation) and average online prices per timeslot, respectively, and ρ denotes the ratio of these prices *i.e.*, $\rho = \frac{p_O}{p_B}$.

[1]We use the terms virtual resource and bandwidth interchangeably from now on.

TABLE I
AMAZON EC2 PRICING FOR AN INSTANCE OF LINUX MACHINE AT US EAST AS OF FEB 23, 2015. RESERVATION PRICES ARE FOR 1-YEAR INSTANCE RESERVE.

Instance Type	Payment Option	Upfront	Hourly
Small	On demand	0	0.026
	Partial upfront	102	0.0059
	All upfront	151	0
Medium	On demand	0	0.026
	Partial upfront	218	0.004
	All upfront	303	0

- **Reservation with Discounted Usage Price (DUP)**. In this model, the virtual operator pays a smaller upfront fee for the reserved bandwidth B compared to reservation with NUF. However, the operator also pays for utilizing the reserved bandwidth albeit at a discounted price. As before, additional bandwidth can be acquired at the online price. This model is particularly common among Infrastructure-as-a-Service (IaaS) providers *e.g.*, Amazon EC2 [18]. Table I shows how Amazon EC2 charges users for a medium-sized Linux machine instance. As demonstrated, other than on demand pricing, there is an all upfront option which is similar to reservation with NUF and there is a partial upfront option which is similar to reservation with DUP. In this model, the cost of satisfying demand x where bandwidth B is reserved is as follows

$$cost^{DUP}(x, B, d, p_B, p_O) = p_B B + \mathbf{1}_{\{x \leq B\}} d p_O x + \mathbf{1}_{\{x > B\}} p_O (dB + (x - B)), \tag{2}$$

where $\mathbf{1}_E$ is the indicator function for event E and $0 < d < 1$ denotes the price discount factor (i.e., the ratio between the charged and online price). Note that if $x \leq B$, other than the reservation cost, the operator only pays the discounted price for usage. However, if $x > B$, up to B, the usage price is discounted while the additional bandwidth *i.e.*, $x - B$ is purchased at the online price. This cost function can be simplified as

$$\begin{aligned} cost^{DUF}(x, B, d, p_B, p_O) &= p_B B + p_d x + \\ &\quad (p_O - p_d)\max(x - B, 0) \\ &= p_B(B + \alpha x + \beta \max(x - B, 0)), \end{aligned} \tag{3}$$

where $p_d = dp_O$ and $\beta = \rho - \alpha = \frac{p_O - p_d}{p_B}$.

- **Reservation with Discounted Online Price (DOP)**. In this model, usage of the reserved bandwidth is free of charge. In addition, as more bandwidth is reserved, the virtual operator receives more discount on the online price of the additionally required resources. This model is motivated by the data plans offered by mobile operators *e.g.*, Koodo mobile [19] and Fido [20] in Canada. For example, Table II depicts the monthly plans offered by Koodo mobile [19]. As demonstrated, going from a 35\$/month plan to a 60\$/month plan reduces the cost of additional data from 5\$/100 MB to 5\$/250 MB. In this model, the cost of satisfying demand x can be formulated as

$$cost^{DOF}(x, B, p_B, p_O) = p_B(B + \rho\frac{D - B}{D}\max(x - B, 0)). \tag{4}$$

The online price is proportional to $\frac{D-B}{D}$. Thus, for every two reservations B_1 and B_2 such that $B_2 > B_1$, we have $\frac{D-B_2}{D} < \frac{D-B_1}{D}$.

B. Problem formulation

At the beginning of a frame, the operator can only observe limited information about the future traffic distribution. This information includes first k moments $\boldsymbol{\mu} = [\mu_1, \mu_2, \ldots, \mu_k]$ of the distribution. The moment vector $\boldsymbol{\mu}$ defines the set $\boldsymbol{f(\mu)}$ of traffic distributions that fit this vector *i.e.*,

$$\boldsymbol{f(\mu)} = \{f(x) | f(x) \geq 0, \int f(x) = 1, \int x^i f(x) = \mu_i\}. \tag{5}$$

As the operator is not aware of the actual distribution, it should plan for the worst-case distribution. Assume that reservation pricing is performed according to some model R (one of the previously introduced models). Then, the worst-case expected cost of the system for the reserved bandwidth B and the moment vector $\boldsymbol{\mu}$ is defined as

$$COST^R(B, \boldsymbol{\mu}, \boldsymbol{R_p}) = \max_{\boldsymbol{f(\mu)}} \int_0^D f(x) cost^R(x, B, \boldsymbol{R_p}) dx, \tag{6}$$

where $cost^R(x, B, \boldsymbol{R_p})$ denotes the cost of satisfying demand x based on reservation model R given the reserved bandwidth B. $\boldsymbol{R_p}$ denotes the associated parameters vector *e.g.*, base price, *etc*.Then, the goal of the operator is to find the optimal reserved bandwidth B^* such that

$$B^* = \arg\min_B \; COST^R(B, \boldsymbol{\mu}, \boldsymbol{R_p}). \tag{7}$$

In the following sections, we solve this problem for all of the above reservation models where the available information is limited to the first and second moment. To do so, we utilize the following observations.

Observation 1. *If $f(x)$ is a probability distribution i.e., $f(x) \geq 0, \int f(x)dx = 1$, $\phi_1(B)$ is a function of B, and $\phi_2(x, B)$ is a function of x and B, then*

$$\int f(x)\big[\phi_1(B) + \phi_2(x, B)\big]dx = \phi_1(B) + \int f(x)\phi_2(x, B)dx. \tag{8}$$

Observation 2. *If $f(x)$ is a probability distribution, $\boldsymbol{f(x)}$ is a set of distributions, $\phi_1(B)$ and $\phi_3(B)$ are functions of B, and $\phi_2(x, B)$ is a function of x and B, then*

$$\max_{\boldsymbol{f(x)}}\{\int f(x)\big[\phi_1(B) + \phi_3(B)\phi_2(x, B)\big]dx\} = \phi_1(B) + \phi_3(B)\max_{\boldsymbol{f(x)}}\{\int f(x)\phi_2(x, B)dx\}. \tag{9}$$

TABLE II
KOODO MOBILE'S CANADA-WIDE DATA PLANS AS OF FEB 23, 2015.

Mnthly Fee	**35$**	**40$**	**45$**	**50$**	**60$**
Anytime Minutes	200	300	500	750	UNLIMITED
Data	UP TO 50 MB	UP TO 300 MB	UP TO 500 MB	UP TO 750 MB	UP TO 1 GB
Additional Data	$5/50 MB	$5/100 MB	$5/100 MB	$5/100 MB	$5/2500 MB

C. Evaluation metrics

Following the literature on online algorithms [15], the performance of the proposed reservation policies is evaluated w.r.t the optimal offline algorithm that knows the exact amount of future traffic during the frame. Let $cost^R(x, B^*, \boldsymbol{R_p})$ denote the cost satisfying demand x based on reservation model R given the optimal reserved bandwidth B^* and associated parameters vector $\boldsymbol{R_p}$. The metrics are as follows:

- **Expectation of ratio (EoR).** The worst-case expected ratio of the cost incurred by the optimal reservation policy B^* and the cost of the offline algorithm *i.e.*,

$$EoR^R = \max_{\boldsymbol{f}(\boldsymbol{\mu})} \mathbb{E}\big[\frac{cost^R(x, B^*, \boldsymbol{R_p})}{cost_{off}(x, \boldsymbol{R_p})}\big]\,. \tag{10}$$

- **Ratio of expectation (RoE).** The worst-case ratio of the expected cost incurred by the optimal reservation policy B^* and the expected cost of the offline algorithm *i.e.*,

$$RoE^R = \max_{\boldsymbol{f}(\boldsymbol{\mu})} \frac{\mathbb{E}[cost^R(x, B^*, \boldsymbol{R_p})]}{\mathbb{E}[cost_{off}(x, \boldsymbol{R_p})]}\,. \tag{11}$$

In the following sections, we find the optimal reservation policy for each of the above-mentioned reservation models and discuss their characteristics w.r.t the evaluation metrics.

III. RESERVATION WITH NUF

Without loss of generality, we assume $p_B = 1$. We aim to find the optimal reserved bandwidth B^* for reservation with NUF model which is defined as

$$B^* = \arg\min_B \; COST^{NUF}(B, \rho, \boldsymbol{\mu}), \tag{12}$$

where $COST^{NUF}(B, \rho, \boldsymbol{\mu})$ is given by

$$\begin{aligned} &\max_{\boldsymbol{f}(\boldsymbol{\mu})}\{\int_0^D f(x)[B + \rho\max(x - B, 0)]dx\} \\ &= B + \rho \max_{\boldsymbol{f}(\boldsymbol{\mu})} \int_0^D f(x)\max(x - B, 0)dx\}\,. \end{aligned} \tag{13}$$

First, we prove some of the properties of (13).

Theorem 1. *The cost of reservation with NUF* (13) *is convex w.r.t B.*

Proof: **First,** we show that for a specific $f(x) \in \boldsymbol{f}(\boldsymbol{\mu})$, the objective of (13) is convex w.r.t B. B is linear w.r.t B and ρ is positive so it suffices to show that

$$\int_0^D f(x)\max(x - B, 0)dx$$

is convex w.r.t B. To do so, we confirm that its second derivative is nonnegative w.r.t B [21]. First, based on Leibniz integral rule [22], we have

$$\begin{aligned} &\frac{d}{dB}\int_0^D f(x)\max(x - B, 0)dx \\ &= \frac{d}{dB}[\int_0^B 0dx] + \frac{d}{dB}[\int_B^D f(x)(x - B)dx] \\ &= -\int_B^D f(x)dx + f(D)(D - B)\cdot 0 - f(B)(B - B)\cdot 1 \\ &= -\int_B^D f(x)dx\,. \end{aligned}$$

Then, we find the second derivative as follows

$$\begin{aligned} \frac{d}{dB^2}\int_0^D f(x)\max(x - B, 0)dx = \\ &= -\frac{d}{dB}[\int_B^D f(x)dx] \\ &= -[\int_B^D 0dx + f(D)\cdot 0 - f(B)\cdot 1] \\ &= -[-f(B)] = f(B)\,. \end{aligned}$$

Thus, $\frac{d}{dB^2}\int_0^D f(x)\max(x - B, 0)dx = f(B) \geq 0$, which proves the convexity of the objective.

Second, we take $\max$ operation over the set of all traffic distribution functions that satisfy the moment vector $\boldsymbol{\mu}$. Note that the $\max$ function is a convex function [21] *i.e.*, if $f_1(x), f_2(x), \ldots, f_n(x)$ are all convex functions then

$$\max\{f_1(x), f_2(x), \ldots, f_n(x)\}$$

is a convex function. This completes the proof. ∎

Problem (12) has some integral constraints. In the next section, we show how they are eliminated through duality.

A. Dual formulation

To find $COST^{NUF}(B, \rho, \boldsymbol{\mu})$, we should solve the following problem

$$\begin{aligned} \max \quad & \int_0^D f(x)[B + \rho\max(x - B, 0)]dx \\ \text{s.t:} \quad & \int_0^D f(x)dx = 1, \\ & \int_0^D x^i f(x)dx = \mu_i, \quad 1 \leq i \leq k\,. \end{aligned} \tag{14}$$

Associating the Lagrange multiplier λ_0 to the first constraint and multiplier λ_i to each μ_i constraint, the Lagrangian associated with the problem (14), is found as

$$L(\boldsymbol{\lambda}, x) = \lambda_0 + \sum_{i=1}^{k} \lambda_i \mu_i + \int_0^D f(x)\big[B + \rho \max(x - B, 0) - \sum_{i=0}^{k} \lambda_i x^i\big] dx\,. \tag{15}$$

Also, the dual function is defined as

$$g(\boldsymbol{\lambda}) = \sup_x L(\boldsymbol{\lambda})\,. \tag{16}$$

Note that if $\sum_{i=0}^{k} \lambda_i x^i < B + \rho \max(x - B, 0)$ for some x_1, we can set $f(x_1) = \infty$ and achieve $g(\boldsymbol{\lambda}) = \infty$ [2]. Therefore, the dual function is defined as follows

$$g(\boldsymbol{\lambda}) = \begin{cases} \lambda_0 + \sum_{i=1}^{k} \lambda_i \mu_i, & \text{if } \sum_{i=o}^{k} \lambda_i x^i \geq \\ & B + \rho \max(x - B, 0) \\ \infty, & \text{otherwise,} \end{cases}$$

which results in the following dual problem for (14)

$$\begin{aligned} \min_{\boldsymbol{\lambda}} \quad & \lambda_0 + \sum_{i=1}^{k} \lambda_i \mu_i \\ \text{s.t:} \quad & \sum_{i=0}^{k} \lambda_i x^i \geq B + \rho \max(x - B, 0), 0 \leq x \leq D\,. \end{aligned} \tag{17}$$

With a similar argument, we can state the following observation.

Observation 3. *The solution to* $\max_{\boldsymbol{f(\mu)}} \mathbb{E}\{\phi(x)\}$ *for* $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_k]$ *is given by the following program*

$$\begin{aligned} \min \quad & \lambda_0 + \sum_{i=1}^{k} \lambda_i \mu_i \\ \text{s.t:} \quad & \sum_{i=0}^{k} \lambda_i x^i \geq \phi(x), \quad 0 \leq x \leq D\,. \end{aligned} \tag{18}$$

The dual problem has an infinite-dimensional constraint that includes the polynomial $\sum_{i=0}^{k} \lambda_i x^i$. These constraints typically could be transformed to constraints over semidefinite matrices [23]. Therefore, the dual problem is a semidefinite program.

B. No reservation policy results in $EoR^{NUF} < \rho$.

According to the definition of EoR (10), we would like to find B such that the expected ratio of the optimal reservation policy and the optimal offline algorithm is minimized. Suppose that x denotes the demand. The optimal algorithm always knows the demand exactly, thus, it can reserve the required bandwidth at the base price which is 1. Therefore its cost would be x while the cost of reservation with NUF is $B + \rho \max(x - B, 0)$. Therefore, we aim to solve the following

$$\min_B \big\{ \max_{\boldsymbol{f(\mu)}} \{ \int_0^D f(x) \frac{B + \rho \max(x - B, 0)}{x} dx \} \big\}\,. \tag{19}$$

Based on the duality observation 3, the inner maximization equals to

$$\begin{aligned} \min \quad & \lambda_0 + \sum_{i=1}^{n} \lambda_i \mu_i \\ \text{s.t:} \quad & \sum_{i=0}^{n} \lambda_i x^i \geq \frac{B + \rho \max(x - B, 0)}{x} \quad 0 \leq x \leq D\,. \end{aligned}$$

Consequently, the dual constraint can be expressed as

$$\sum_{i=0}^{n} \lambda_i x^{i+1} \geq B + \rho \max(x - B, 0), \quad 0 \leq x \leq D\,.$$

Satisfying the constraint for $x = 0$, we obtain $B \leq 0$ which along with the constraint $0 \leq B \leq D$ results in $B = 0$. Therefore, maximizing the inner problem would lead to no reservation. In this case, the optimal reservation policy always buys the required resources online which consistently cost ρ times the offline algorithm. This completes the proof. Note that this result holds irrespective of the amount of information (the number of moments) that is available to the virtual operator. Following a similar argument, the same result ($EoR \not< \rho$) could be obtained for the DUP and DOP models.

C. Reservation with NUF: first moment constraint

In this section, we solve the problem (12) assuming that the operator is only aware of the traffic mean at the beginning of th frame. In this case, based on the dual problem formulation (17), $COST^{NUF}(B, \rho, \mu)$ is given by

$$\begin{aligned} \min_{\boldsymbol{\lambda}} \quad & \lambda_0 + \lambda_1 \mu \\ \text{st:} \quad & \lambda_0 + \lambda_1 x \geq B + \rho \max(x - B, 0), 0 \leq x \leq D\,. \end{aligned} \tag{20}$$

The infinite-dimensional constraint of (20) can be separated into infinite-dimensional constraint

$$\lambda_1 x + \lambda_0 \geq B, \quad 0 \leq x \leq B \tag{21}$$

and

$$\lambda_1 x + \lambda_0 \geq \rho x + (1 - \rho) B, \quad B \leq x \leq D\,. \tag{22}$$

Based on the endpoint values of x in (21) *i.e.*, $x = 0$ and $x = B$, the former constraint (21) is reduced to the following linear constraints

$$\begin{aligned} & \lambda_0 \geq B \\ & \lambda_1 B + \lambda_0 \geq B\,. \end{aligned} \tag{23}$$

For any other point $0 < x < B$, constraint (21) can be represented as a convex combination of $\lambda_0 \geq B$ and $\lambda_1 B + \lambda_0 \geq B$. In addition, based on the end point values, constraint (22) is reduced to the following set of linear constraints

$$\begin{aligned} & \lambda_1 B + \lambda_0 \geq B \\ & \lambda_1 D + \lambda_0 \geq \rho D + (1 - \rho) B \end{aligned} \tag{24}$$

[2] Recall that the only constraint on $f(x)$ is that it should be nonnegative and the other constraints of the probability distribution function $f(x)$ have already been included in the formulation.

in which the constraint $\lambda_1 B + \lambda_0 \geq B$ is redundant. Overall, by the sets of constraints (23) and (24), the following holds on λ_0 and λ_1

$$\begin{aligned}&\lambda_0 \geq B\\&\lambda_1 \geq \max\{1-\frac{\lambda_0}{B}, \rho + \frac{(1-\rho)B-\lambda_0}{D}\}\,.\end{aligned}$$

Accordingly, problem (20) is transformed to

$$\begin{aligned}\min_{\lambda_0} \quad & \lambda_0 + \mu \max\{1-\frac{\lambda_0}{B}, \rho + \frac{(1-\rho)B-\lambda_0}{D}\}\\ \text{st:} \quad & \lambda_0 \geq B\,.\end{aligned} \tag{25}$$

Observation 4. *Given $\rho > 1$, $\lambda_0 \geq B \geq 0$, and $B \leq D$, the following holds*

$$1-\frac{\lambda_0}{B} \leq \rho + \frac{(1-\rho)B-\lambda_0}{D}\,.$$

Proof: Since $0 \leq B \leq D$, we have $(\frac{1}{D}-\frac{1}{B}) \leq 0$ and $(1-\frac{B}{D}) \geq 0$. Also, based on the assumption, we have $1-\rho \leq 0$. Therefore, since $\lambda_0 \geq 0$, we have

$$\lambda_0(\frac{1}{D}-\frac{1}{B}) + (1-\rho)(1-\frac{B}{D}) \leq 0$$

which implies

$$1-\frac{\lambda_0}{B} - \rho - \frac{(1-\rho)B-\lambda_0}{D} \leq 0$$

or, alternatively;

$$1-\frac{\lambda_0}{B} \leq \rho + \frac{(1-\rho)B-\lambda_0}{D}$$

■

Now, assume that

$$\lambda_0(\frac{1}{D}-\frac{1}{B}) + (1-\rho)(1-\frac{B}{D}) > 0$$

which according to the observation might happen where $\rho < 1$. This expression is simplified to

$$\lambda_0 > \frac{(1-\rho)(1-\frac{B}{D})}{(\frac{1}{D}-\frac{1}{B})} = (\rho-1)B$$

which is a looser constraint in comparison to $\lambda_0 \geq B$ since $\rho - 1 < 0$. Therefore, for all $\rho > 0$, the following constraints

$$\begin{aligned}&\lambda_0 \geq B\\&\lambda_1 \geq \rho + \frac{(1-\rho)B-\lambda_0}{D}\end{aligned} \tag{26}$$

completely specify the domain of the possible solutions of (20). Thus, the objective of (20) is represented as follows

$$\lambda_0 + \mu(\rho + \frac{(1-\rho)B-\lambda_0}{D}) = \mu\rho + (1-\rho)\frac{B\mu}{D} + \lambda_0(1-\frac{\mu}{D})$$

Since the derivative of the objective *i.e.*, $(1-\frac{\mu}{D})$ is positive w.r.t λ_0, its minimum is attained at $\lambda_0 = B$. Thus, $COST^{NUF}(B,\rho,\mu)$ is given by

$$COST^{NUF}(B,\rho,\mu) = \mu\rho + B(1-\frac{\mu\rho}{D})$$

Next, we would like to find B^* that minimizes this cost. The solution depends on the derivative of the cost w.r.t B which is $1-\frac{\mu\rho}{D}$. Then, based on the sign of the derivative, B^* is obtained as follows

$$B^* = \begin{cases}0, & D > \mu\rho\\ D, & D \leq \mu\rho.\end{cases} \tag{27}$$

Remark. *If we have no limit on the maximum demand i.e., $D = \infty$, no bandwidth is reserved.*

Remark. *If the average online price is lower than the base price i.e., $\rho < 1$, no bandwidth is reserved.*

Theorem 2. *Based on the first moment information, RoE^{NUF} is no less than ρ.*

Proof: The set of possible distributions is limited to those that have the mean value of μ. The optimal offline algorithm knows exactly the amount of traffic demand and reserves it at the base price at the beginning of the frame. Therefore for all distributions, the expected cost of the optimal offline algorithm is μ. Therefore, the denominator in (11) is constant and for achieving the maximal ratio, the worst-case expected cost should be considered. On the other hand, based on the optimal policy (27), where $D > \mu\rho$, no bandwidth is reserved, thus, all the required resources are acquired at the online price which expectedly costs $\mu\rho$. Therefore, $RoE^{NUF} = \rho$. ■

D. Reservation with NUF: first and second moment constraints

Here, we assume that the operator could observe both mean μ and variance σ^2 of the traffic distribution ($\boldsymbol{\mu} = [\mu, \mu^2+\sigma^2]$) and would like to find B^* that minimizes the following

$$COST^{NUF}(B,\rho,\boldsymbol{\mu}) = B + \rho \max_{\boldsymbol{f}(\boldsymbol{\mu})}\{\int_0^D f(x)\max(x-B,0)dx\}\,. \tag{28}$$

The formulation of (28) corresponds to the European call option pricing investigated in economics literature [24]–[26]. The problem is defined as follows. An option provides the holder with the right to buy a specified quantity of an underlying asset at a fixed price (called a strike price) at the expiration date of the option. Then, the expected option price is defined as follows

$$\mathbb{E}_f\{\max(x-k,0)\}$$

where k is the constant strike price and f is the distribution of the underlying value of the asset x. Lo [25] found the upper bounds on the European option price where only mean and variance of the asset distribution is known. The bounds are as follows

$$\begin{aligned}&\max_{\boldsymbol{f}(\mu,\sigma)}\{\mathbb{E}\{\max(x-k,0)\}\} =\\ &\begin{cases}\frac{1}{2}\left[(\mu-k)+\sqrt{\sigma^2+(\mu-k)^2}\right], & k \geq \frac{\mu^2+\sigma^2}{2\mu}\\ \mu-k+k\frac{\sigma^2}{\mu^2+\sigma^2}, & k < \frac{\mu^2+\sigma^2}{2\mu}.\end{cases}\end{aligned} \tag{29}$$

To find the optimal reserved bandwidth B^*, in the rest of this section we replace $\max\{\int_0^D f(x)\max(x-B,0)dx\}$ in (28) with the above bounds and find the optimal reservation policy for each bound and compare the results.

First case: $B \in [\frac{\mu^2+\sigma^2}{2\mu}, D]$. According to (29) we have

$$COST_1^{NUF}(B,\rho,\boldsymbol{\mu}) = B + \frac{\rho}{2}\left[(\mu - B) + \sqrt{\sigma^2 + (\mu - B)^2}\right]. \tag{30}$$

To minimize $COST_1^{NUF}(B,\rho,\boldsymbol{\mu})$, we take its derivative w.r.t B which gives

$$\frac{dCOST_1^{NUF}(B,\rho,\boldsymbol{\mu})}{dB} = 1 - \frac{\rho}{2}\left(1 + \frac{\mu - B}{\sqrt{(\mu-B)^2+\sigma^2}}\right). \tag{31}$$

By setting the derivative to 0, the optimal reserved bandwidth is found as follows

$$B^* = \mu + \sigma\frac{\rho-2}{2\sqrt{\rho-1}}. \tag{32}$$

Let $\kappa = \frac{\rho-2}{2\sqrt{\rho-1}}$. Then, $COST_1^{NUF}(B^*,\rho,\boldsymbol{\mu})$ is obtained as

$$COST_1^{NUF}(B^*,\rho,\boldsymbol{\mu}) = \mu + \sigma[\kappa + \frac{\rho}{2}[\sqrt{1+\kappa^2} - \kappa]]. \tag{33}$$

Theorem 3. *Given $B^* = \mu + \sigma\kappa$, RoE^{NUF} is of order $\theta(\sqrt{\rho})$.*

Proof: Recall from the proof of theorem 2 that the expected value of the optimal offline algorithm is μ. To show that for B^*, RoE^{NUF} is asymptotically equal to $\sqrt{\rho}$, we take the limit of ratio of both expressions when $\rho \to \infty$ *i.e.*,

$$\begin{aligned} &\lim_{\rho\to\infty} \frac{\frac{\mu+\sigma\sqrt{\rho}[\frac{\kappa}{\sqrt{\rho}} + \frac{\rho}{2}[\sqrt{1+\kappa^2}-\kappa]]}{\mu}}{\sqrt{\rho}} = \\ &\lim_{\rho\to\infty} \frac{1 + \frac{\sigma}{\mu}\sqrt{\rho}}{\sqrt{\rho}} = \frac{\sigma}{\mu}, \end{aligned} \tag{34}$$

where the simplification comes from the following

$$\lim_{x\to\infty} \frac{(x-2)}{2\sqrt{x-1}\sqrt{x}} + \frac{\sqrt{x}}{2}\left[\sqrt{1 + \frac{(x-2)^2}{4(x-1)}} - \frac{x-2}{4\sqrt{x-1}}\right] = 1. \tag{35}$$

Since the achieved ratio is a constant $\frac{\sigma}{\mu} > 0$, the proof is complete. ■

Second case: $B \in [0, \frac{\mu^2+\sigma^2}{2\mu}]$. According to (29), $COST_2^{NUF}(B,\rho,\boldsymbol{\mu})$ is found as follows

$$\begin{aligned} COST_2^{NUF}(B,\rho,\boldsymbol{\mu}) &= B + \rho[\mu - B + B\frac{\sigma^2}{\mu^2+\sigma^2}] \\ &= \rho\mu + B[1 + \rho(\frac{\sigma^2}{\sigma^2+\mu^2} - 1)] \\ &= \rho\mu + B[1 - \rho\frac{\mu^2}{\sigma^2+\mu^2}]. \end{aligned} \tag{36}$$

To obtain the minimum value of $COST_2^{NUF}(B,\rho,\boldsymbol{\mu})$, we take its derivate w.r.t B which is given by

$$\frac{dCOST_2^{NUF}(B,\rho,\boldsymbol{\mu})}{dB} = 1 - \rho\frac{\mu^2}{\sigma^2+\mu^2}.$$

Based on the sign of the derivative, there are two options for the optimal reserved bandwidth B^* as follows

- $\rho < 1 + \frac{\sigma^2}{\mu^2}$. In this case, $\frac{dCOST_2^{NUF}(B,\rho,\boldsymbol{\mu})}{dB} > 0$ and B should be set to the minimum possible value *i.e.*, $B^* = 0$ which results in $COST_2^{NUF}(B^*,\rho,\boldsymbol{\mu}) = \rho\mu$.
- $\rho \geq 1 + \frac{\sigma^2}{\mu^2}$. In this case B^* is set to the maximum possible value *i.e.*,

  $$B^* = \frac{\mu^2+\sigma^2}{2\mu}$$

  which results in $COST_2^{NUF}(B^*,\rho,\boldsymbol{\mu}) = \rho\frac{\mu}{2} + \frac{\mu^2+\sigma^2}{2\mu}$.

In both cases, the expressions of $COST_2^{NUF}(B^*,\rho,\boldsymbol{\mu})$ are linearly dependent on ρ *i.e.*, are of the order $\theta(\rho)$ which is asymptotically greater than the result obtained in the analysis of $COST_1^{NUF}(B^*,\rho,\boldsymbol{\mu})$ *i.e.*, $\theta(\sqrt{\rho})$. Therefore, the optimal reservation policy is to choose B^* according to (32).

IV. Reservation with DUP

The cost function for reservation with DUP model is defined in (3). In this section, our goal is to find $COST^{DUP}(B,\alpha,\beta,\boldsymbol{\mu})$ for the cases where $\boldsymbol{\mu} = [\mu]$ and $\boldsymbol{\mu} = [\mu, \mu^2+\sigma^2]$. As will be demonstrated in the rest of the section, although the cost function is different from the reservation with NUF model, the optimal reservation policies are very similar.

First, $COST^{DUP}(B,\alpha,\beta,\boldsymbol{\mu})$ can be simplified as follows

$$\begin{aligned} &COST^{DUP}(B,\alpha,\beta,\boldsymbol{\mu}) \\ &= \max_{\boldsymbol{f(\mu)}}\{\int_0^D f(x)[B + \alpha x + \beta\max(x-B)]dx\} \\ &= B + \max_{\boldsymbol{f(\mu)}}\{\alpha\int_0^D xf(x)dx + \beta\int_0^D f(x)\max(x-B)dx\}\cdot \\ &= B + \max_{\boldsymbol{f(\mu)}}\{\alpha\mu + \beta\mathbb{E}\{\max(x-B)dx\}\} \\ &= \alpha\mu + B + \beta\max_{\boldsymbol{f(\mu)}}\mathbb{E}\{\max(x-B)dx\} \end{aligned} \tag{37}$$

The simplification comes from the fact that the virtual operator at least knows the mean of traffic distribution *i.e.*, $\int_0^D f(x)x = \mu$. As could be noticed, $COST^{DUP}(B,\alpha,\beta,\boldsymbol{\mu})$ is similar to (13) except for the addition of the term $\alpha\mu$ and the change of ρ to β. Therefore, we can use the results of section III to obtain optimal policies for reservation with DUP model. Note that the term $\alpha\mu$ is constant w.r.t B, therefore, its addition to the cost function (13) does not change the proposed optimal reservation policy. Thus, the optimal reservation policy for ***DUP with the first moment constraint*** is as follows

$$B^* = \begin{cases} 0, & D > \beta\mu \\ D, & D \leq \beta\mu. \end{cases}$$

Moreover, if $D > \beta\mu$, no bandwidth would be reserved which results in the total cost of $\alpha\mu + \beta\mu = \rho\mu$. The following holds on RoE^{DUP}.

Theorem 4. *Based on the first moment information, RoE^{DUP} is no less than ρ.*

For the case where the available moment vector is $\boldsymbol{\mu} = [\mu, \mu^2 + \sigma^2]$, the worst-case expected cost is found according to (37) and Lo's bounds (29). The cost is as follows

$$COST^{DUP}(B, \alpha, \beta, \boldsymbol{\mu}) = \alpha\mu + B + \frac{\beta}{2}\Big[(\mu - B) + \sqrt{(\mu - B)^2 + \sigma^2}\Big] . \tag{38}$$

Similar to (32), the optimal reservation policy for **DUP with the first and second moment constraints** is given by

$$B^* = \mu + \sigma\frac{(\beta - 2)}{2\sqrt{\beta - 1}} . \tag{39}$$

The similarity of the above reservation policy to the one for the NUF model allows us to state the following. The proof is omitted due to lack of space.

Theorem 5. *Given $B^* = \mu + \sigma\frac{(\beta-2)}{2\sqrt{\beta-1}}$, RoE^{DUP} is of order $\theta(\sqrt{\beta})$.*

V. Reservation with DOP

The reservation cost function is defined in (4). In this section, we find $COST^{DOP}(B, \rho, \boldsymbol{\mu})$, for the cases where the known moment vector $\boldsymbol{\mu}$ is limited to the first and second moments.

A. Reservation with DOP: first moment constraint

Based on the dual formulation (3), $COST^{DOP}(B, \rho, \mu)$ is given by the following optimization program

$$\begin{aligned} \min_{\boldsymbol{\lambda}} \quad & \lambda_0 + \lambda_1\mu \\ \text{s.t:} \quad & \lambda_0 + \lambda_1 x \geq B + \rho\frac{D - B}{D}\max(x - B, 0), 0 \leq x \leq D . \end{aligned}$$

The constraint can be reduced to

$$\lambda_0 \geq B, \lambda_1 B + \lambda_0 \geq B, \lambda_1 D + \lambda_0 \geq B + \rho\frac{(D - B)^2}{D},$$

or equivalently the following constraints on λ_0 and λ_1,

$$\begin{aligned} &\lambda_0 \geq B \\ &\lambda_1 \geq \max\{1 - \frac{\lambda_0}{B}, \frac{B}{D} + \rho\frac{(D - B)^2}{D^2} - \frac{\lambda_0}{D}\} \end{aligned}$$

Observation 5. *For $\lambda_0 \geq B$ and $B \leq D$, the following holds*

$$\frac{B}{D} + \rho\frac{(D - B)^2}{D^2} - \frac{\lambda_0}{D} > 1 - \frac{\lambda_0}{B} .$$

Therefore, the problem can be represented as

$$\begin{aligned} \min_{\lambda_0} \quad & \lambda_0(1 - \frac{\mu}{D}) + \mu[\frac{B}{D} + \rho\frac{(D - B)^2}{D^2}] \\ \text{s.t:} \quad & \lambda_0 \geq B . \end{aligned} \tag{40}$$

Since $(1 - \frac{\mu}{D}) > 0$, the minimum objective is attained at $\lambda_0 = B$. Therefore, the solution to (40) is given by

$$COST^{DOP}(B, \rho, \mu) = B + \rho\frac{(D - B)^2}{D^2}\mu . \tag{41}$$

The optimal reservation policy is obtained by taking the derivative of (41) *i.e.*, $\frac{1}{D^2}(D^2 - 2\rho(D - B)\mu)$ and setting it to 0. The optimal reserved bandwidth is as follows

$$B^* = \frac{D(2\rho\mu - D)}{2\rho\mu} . \tag{42}$$

B. Reservation with DOP: first and second moment constraints

Here, the observed moment vector is $\boldsymbol{\mu} = [\mu, \mu^2 + \sigma^2]$. According to observation 2, the worst-case expected cost of reservation with DOP is given by

$$B + \rho\frac{D - B}{D}\max_{\boldsymbol{f}(\boldsymbol{\mu})}\mathbb{E}\{\max(x - B, 0)\}$$

This allows to use Lo's maximal bounds (29) on $\max(x - B, 0)$ directly. Doing so, a closed-form solution for the worst-case expected cost is obtained as

$$\begin{aligned} COST^{DOP}(B, \rho, \boldsymbol{\mu}) =& B + \rho\frac{D - B}{2D}[(\mu - B) + \\ & \sqrt{(\mu - B)^2 + \sigma^2}] \end{aligned} \tag{43}$$

The minimal value of (43) is achieved at B^* which satisfies the following relation

$$\frac{dCOST^{DOP}(B^*, \rho, \boldsymbol{\mu})}{dB} = 0$$

However, the derivative of 43 is given by the following complicated expression

$$\frac{dCOST^{DOP}(B, \rho, \boldsymbol{\mu})}{dB} = 1 - \frac{(D - B)\rho\left(1 + \frac{\mu - B}{\sqrt{(\mu - B)^2 + \sigma^2}}\right)}{2D} - \frac{\rho\left(\mu - B + \sqrt{(\mu - B)^2 + \sigma^2}\right)}{2D} \tag{44}$$

which does not easily admit to a closed-form solution. However, there are standard techniques for solving such equation numerically, e.g. gradient descent method [21].

VI. Discussion and Extension

A. Higher order statistics

While the information that is available to the operator at the time of reservation is typically limited to low order statistics *e.g.*, mean and variance of the traffic distribution, in general information about the higher moments can also be incorporated in decision making process. To do so, the dual problem (3) is transformed to a semidefinite optimization program [26]. For example, having the k-moment vector $\boldsymbol{\mu}$, $COST^{NUF}(B, \rho, \boldsymbol{\mu})$ is given by the following program

$$\begin{aligned} \min_{\boldsymbol{\lambda}} \quad & \lambda_0 + \sum_{i=1}^{k}\lambda_i\mu_i \\ \text{s.t:} \quad & \sum_{i=0}^{k}\lambda_i x^i \geq B + \rho\max(x - B, 0), \quad 0 \leq x \leq D \end{aligned}$$

where the constraint can be decomposed as

$$(\lambda_0 - B) + \sum_{i=1}^{k} \lambda_i x^i \geq 0, \quad 0 \leq x \leq B$$

$$(\lambda_0 + (\rho - 1)B) + (\lambda_1 - \rho)x + \sum_{i=2}^{k} \lambda_i x^i \geq 0, \quad B < x < D. \tag{45}$$

Each of the constraints is a polynomial constraint that can be expressed as a condition on a positive semidefinite matrix. Let $\lambda_0'' = \lambda_0 - B$ and $\lambda_i'' = \lambda_i$ for $1 \leq i \leq k$. Moreover, Let $\lambda_0' = \lambda_0 + (\rho - 1)B$ and $\lambda_1' = (\lambda_1 - \rho)$ and $\lambda_i' = \lambda_i$ for $2 \leq i \leq k$. Then, the first constraint is satisfied [23], [26] if there exists a positive semidefinite matrix $\boldsymbol{X} = [x_{ij}]$ such that

$$\begin{aligned} \sum_{i+j=2l-1} x_{ij} &= 0, \quad 1 \leq l \leq k \\ \sum_{i+j=2l} x_{ij} &= \sum_{r=2}^{l} \lambda_r'' \binom{k-r}{l-r} B^r, \quad 0 \leq l \leq k \end{aligned} \tag{46}$$

Also the second constraint is satisfied [26] if there exists a positive semidefinite matrix $\boldsymbol{Z} = [z_{ij}]$ such that

$$\begin{aligned} \sum_{i+j=2l-1} z_{ij} &= 0, \quad 1 \leq l \leq k \\ \sum_{i+j=2l} z_{ij} &= \sum_{r=l}^{k} \lambda_r' \binom{r}{l} B^r, \quad 0 \leq l \leq k \end{aligned} \tag{47}$$

All in all, the value of $COST^{NUF}(B, \rho, \boldsymbol{\mu})$ for a specific B can be obtained by solving a semidefinite program. Therefore, the optimal reserved bandwidth B^* can be found following a numerical search method *e.g.*, Golden section search [27] in which the above-mentioned semidefinite program is called in each iteration.

VII. Numerical results

In this section, we conduct a numerical study to further investigate the properties of the proposed reservation policies.

First, we demonstrate how the costs incurred by the proposed optimal reservation policies are compared to the optimal offline cost *i.e.*, where the operator knows the exact distribution of the traffic. We assume that pricing is done based on the NUF model. The frame size is considered to be 1000 timeslots. The traffic demand in each timeslot is generated from a Poisson distribution with mean of $\lambda = 1000$ request where the maximum demand is considered to be $D = 5000$ per timeslot. The total cost of a frame is the cost of bandwidth reservation plus the cost of acquiring additional required bandwidth averaged over all the timeslots. The optimal reservation policies in the case of availability of the mean or the mean and variance are developed in section III. The resulting total cost incurred by these policies are denoted by COST-mean and COST-var respectively.

To provide a baseline for our comparison, we also consider the optimal offline algorithm. When the traffic distribution is known to the operator the reservation policy that lead to minimal expected cost is to reserve bandwidth as follows[3]

$$B^* = F^{-1}(1 - \frac{1}{\rho})$$

where F is the CDF of the traffic distribution and ρ is the ratio of the online price and base price. The cost incurred by the optimal algorithm is denoted by COST-optimal. We assume the base price per unit of bandwidth per timeslot is 1. The online price is increased from 1 to 10, therefore, the ratio of the online to the base price *i.e.*, ρ is increased from 1 to 10 as well. The result is demonstrated in Figure 2.

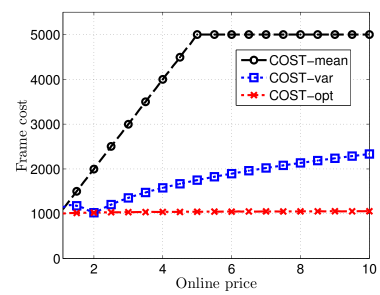

Fig. 2. Frame cost vs the online price in reservation with NUM model.

As can be seen in the figure, for small values of ρ, COST-mean is linearly dependent on ρ, while, for large values of ρ, it reserves all the possible allocable bandwidth. Including the variance in the decision-making process has a significant impact on the cost as COST-var shows a square root dependency in terms of ρ. In addition, the cost incurred by the optimal algorithm remains indifferent of the changes in the value of ρ.

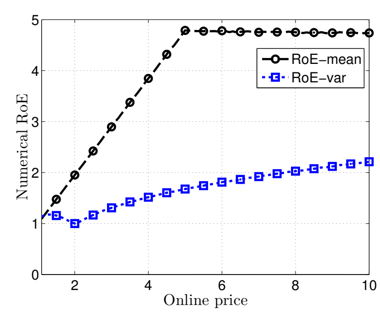

Fig. 3. Ratio of frame cost to the optimal cost.

Figure 3 depicts the ratio of $\frac{COST-mean}{COST-opt}$ and $\frac{COST-var}{COST-opt}$ which we call them numerical RoE here. An interesting observation is that even for large values of ρ (up to $\rho \leq 8$), COST-var incurs at most twice the cost of the optimal offline algorithm.

In the second set of results, we demonstrate the reservation behavior of the proposed policies by changing the online price.

[3]This can be easily shown by taking the derivative of $B + \rho \int f(x) \max(x - B, 0) dx$ and solving the equation achieved by setting to 0.

The base price for the three reservation models NUF, DUP, and DOP are considered to be 1, 0.6, and 1.2 respectively. As before, we assume that the traffic mean is $\mu = 1000$ arrival per timeslot and the maximum demand is $D = 5000$. Also, in DUP the users receive 80% discount over the online price if the used bandwidth is already reserved. The optimal reserved bandwidth for these models are demonstrated in Figures 4 and 5. As can be seen in Figure 4, the optimal policy in NUF and DUP models has an ON/OFF nature, that is, while the ratio of the online price and the base price is small, no bandwidth is reserved. On the other hand, if this ratio is large, all the available bandwidth is reserved. However, in the DOP model, the reserved bandwidth scales almost linearly w.r.t ρ except for the smaller values of this ratio. Moreover, Figure 5 asserts that when information about the variance is available, the optimal bandwidth reservation scales as $\sqrt{\rho}$.

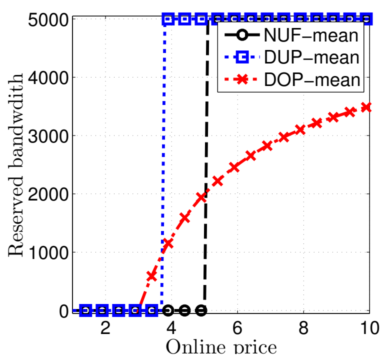

Fig. 4. Reservation based on the traffic mean.

VIII. Conclusion

In this paper, we studied resource reservation problem in mobile virtual networks. We presented a minimax optimization framework that aims to minimize the worst-case expected operational cost of the system knowing only low order statistics of the future traffic demand. Given the first and second moment statistics, the closed-form reservation policies are proposed. Our numerical results shows that given the variance information, for practical pricing scenarios (where the ratio of average online price and base price is lower than 10), the performance of the proposed policies are within a small factor (≈ 2) of that of the optimal offline policy.

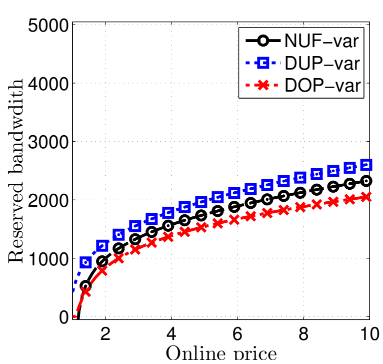

Fig. 5. Reservation based on the traffic mean and variance.

References

[1] "Ericsson Mobility Report," http://www.ericsson.com/res/docs/2015/ericsson-mobility-report-feb-2015-interim.pdf, Feb 2015.

[2] "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update," http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf, Feb 2015.

[3] E. Obregon, K. W. Sung, and J. Zander, "On the Sharing Opportunities of Ultra-Dense Networks in the Radar Bands," in *Proc. IEEE DYSPAN*, MCLEAN, USA, Apr 2014, pp. 215–223.

[4] S. Parkvall, A. Furuskar, and E. Dahlman, "Evolution of LTE toward IMT-advanced," *IEEE Commun. Mag.*, vol. 49, no. 2, pp. 84–91, 2011.

[5] C. Peng, S.-B. Lee, S. Lu, H. Luo, and H. Li, "Traffic-Driven Power Saving in Operational 3G Cellular Networks," in *Proc. ACM MobiCom*, Las Vegas, USA, Sep 2011, pp. 121–132.

[6] R. Mahindra, M. A. A. Khojastepour, H. Zhang, and S. Rangarajan, "Network-Wide Radio Access Network Sharing in Cellular Networks," in *Proc. IEEE ICNP*, Goettingen, Germany, Oct 2013, pp. 1–10.

[7] R. Kokku, R. Mahindra, H. Zhang, and S. Rangarajan, "NVS: A Virtualization Substrate for WiMAX Networks," in *Proc. ACM MobiCom*, Chicago, USA, Sep 2010, pp. 233–244.

[8] ——, "CellSlice: Cellular Wireless Resource Slicing for Active RAN Sharing," in *Proc. COMSNETS*, Bangalore, India, Jan 2013, pp. 1–10.

[9] M. Z. Shafiq, L. Ji, A. X. Liu, and J. Wang, "Characterizing and Modeling Internet Traffic Dynamics of Cellular Devices," in *Proc. ACM SIGMETRICS*, San Jose, USA, Jun 2011, pp. 305–316.

[10] W. Wiesemann, D. Kuhn, and M. Sim, "Distributionally robust convex optimization," *Operations Research*, vol. 62, no. 6, pp. 1358–1376, 2014.

[11] E. Delage and Y. Ye, "Distributionally Robust Optimization and Its Tractable Approximations," *Operations Research*, vol. 58, no. 3, pp. 595–612, 2010.

[12] W. Wang, B. Li, and B. Liang, "To Reserve or Not to Reserve: Optimal Online Multi-Instance Acquisition in IaaS Clouds," in *Proc. USENIX ICAC*, San Jose, USA, Jun 2013, pp. 13–22.

[13] A. Khanafer, M. Kodialamy, and K. P. N. Puttaswamy, "The Constrained Ski-Rental Problem and its Application to Online Cloud Cost Optimization," in *Proc. IEEE INFOCOM*, Turin, Italy, Apr 2013, pp. 1492–1500.

[14] C. Dong, H. Zeng, and M. Chen, "A Cost Efficient Online Algorithm for Automotive Idling Reduction," in *Proc. ACM DAC*, San Fransisco, USA, Jun 2014, pp. 1–6.

[15] A. Madry and D. Panigrahi, "The Semi-stochastic Ski-rental Problem," in *Proc. FSTTCS*, Mumbai, india, Dec 2011, pp. 300–311.

[16] L. Huang, J. Walrand, and K. Ramchandran, "Optimal Power Procurement and Demand Response with Quality-of-Usage Guarantees," in *Proc. IEEE PES-GM*, San Diego, USA, Jul 2012, pp. 1–8.

[17] M. Ghamkhari, H. Mohsenian-Rad, and A. Wierman, "Optimal Risk-aware Power Procurement for Data Centers in Day-Ahead and Real-Time Electricity Markets," in *Proc. INFOCOM Smart Data Pricing (SDP) Workshop*, Toronto, Canada, May 2014, pp. 610–615.

[18] "Amazon EC2 Pricing," http://aws.amazon.com/ec2/pricing/.

[19] "Koodo Lightweight Canada-wide plans," https://www.koodomobile.com/.

[20] "Fido monthly plans," http://www.fido.ca/web/page/portal/Fido/MonthlyPlans/Group12.

[21] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge university press, 2004.

[22] D. V. Widder, *Advanced Calculus*. Courier Corporation, 1989.

[23] A. Nemirovski, *Lectures on Modern Convex Optimization*. SIAM, 2001.

[24] H. E. Scarf, "A Min-Max Solution of an Inventory Problem," *Studies in the mathematical theory of inventory and production*, vol. 10, pp. 201–209, 1958.

[25] A. W. Lo, "Semi-parametric upper bounds for option prices and expected payoffs," *Financial Economics*, vol. 19, no. 2, pp. 373–387, 1987.

[26] D. Bertsimas and I. Popescu, "On The Relation Between Option and Stock Prices: A Convex Optimization Approach," *Operations Research*, vol. 50, no. 2, pp. 358–374, 2002.

[27] P. Wheatley and C. F. Gerald, *Applied Numerical Analysis*. Addison-Wesley Publishing Company, 1984.