이종 이동통신망에서의 네트워크 혼잡 변동 과금 정책

Congestion-Dependent Pricing in Heterogeneous Wireless Access Networks

요 약

Mobile data services grow rapidly in few years mainly due to prevalence of smart handheld devices. Under conventional pricing schemes such as flat, volume, and tiered pricing, network congestion occurs frequently in particular time intervals and locations. Meanwhile, WiFi networks have abundant capacity, but small coverage areas compared to cellular networks. In this paper, we study congestion-dependent pricing in heterogeneous wireless access networks and compare it with conventional pricing schemes. In congestion-dependent pricing, more mobile data traffic can be offloaded through WiFi networks by giving users incentives to deviate. A simple numerical result on temporally variable demand verifies that total revenue is increased about 13% in congestion-dependent pricing.

1. Introduction

1.1. Motivation

Recently, the amount of mobile data traffic grows enormously as high performance handheld devices such as Smartphones and pads get popular. In [1], it is reported that global mobile data traffic grew 2.6-fold in 2010 and 60% of the traffic was generated from the top 10% of mobile data subscribers. The total global mobile data traffic will increase 26-fold between 2010 and 2015. The average smartphone will generate 1.3GB of traffic per month in 2015. However, it is anticipated that the cellular network capacity will not keep up with the explosive data traffic growth due to theoretical bound in spectral efficiency and inter-cell interference. As a result, a cell gets congested frequently in specific time interval and location as the data traffic demand varies in time and location.

WiFi offloading is studied in recent works [2, 3] as a solution to mobile data explosion. WLAN technologies such as WiFi networks provide abundant data rates and small coverage area so that network congestion rarely occurs unless many fixed devices are connected to the network (e.g. conference hall). There are two types of WiFi offloading: *on-the-spot* and *delayed offloading*. *On-the-spot offloading* is to use spontaneous connectivity to WiFi and transfer data on the spot. In *delayed offloading*, each data transfer is associated with a specified deadline and as users come in and out of WiFi coverage areas, it repeatedly resume data transfer until the transfer is complete. If the data transfer does not finish within its deadline, cellular networks finally complete the transfer. These techniques seem viable to reduce cellular network congestion, but users may have little incentive to use WiFi networks due to a pricing structure (e.g. unlimited data plan) so that these techniques may not be effective. In this paper, we focus on the pricing method that leads to mobile data offloading through WiFi networks and alleviates network congestion.

Widely used pricing schemes in mobile data services are flat, volume, and tiered pricing. In flat pricing or unlimited data plan, a fixed fee is charged to a subscriber regardless of usage. In volume pricing, a charge is proportional to the usage of a subscriber, i.e. the price per unit data is fixed. In tiered pricing, network providers offer multiple price points and corresponding usage limits for users. A subscriber chooses a price point that he or she prefers. These pricing schemes are simple and succeeded in gathering subscribers in the start-up stages, but they fail to control temporal and spatial congestion. Even in flat pricing, a few users with large demand may aggravate the whole network without spending additional fee. In the meantime, as a method of congestion control, network providers regulate quality of service (throughput, latency, etc.) for users who are in a congested area or generate large data traffic. However, this leads to user dissatisfaction and fails to attract users who are willing to pay more to get high quality of service.

1.2. Related work

Congestion-dependent pricing has got much attention in many public utility services (e.g. road, electricity, subway) which have limited supply and variable demand. The road and electricity industry has explored congestion-dependent pricing over decades. W. Vickrey first proposed the idea of congestion pricing [4]. His idea was implemented in several systems. In

London, a congestion pricing system has been in effect on frequently congested areas since 2003 [5]. All motorists traveling within the congestion charge zone (CCZ) between 7:00 am and 6:00 pm must pay a fee. As a result, the amount of traffic in the areas in 2006 became 16% lower than that in 2002. The surplus reinvested into London's transport system. The time-dependent pricing in electricity has been studied both in theory and application [6, 7].

Congestion-dependent pricing in capacity sharing networks is proposed in several works. In [8], authors studied congestion-dependent pricing in multi-class call services. A wireless service provider charges a call fee that depends on the current congestion level. They proved that the revenue of dynamic pricing is always greater than that of static pricing. Also, dynamic pricing can be more robust to errors in the demand estimation. Joe-Wong et al. [9] developed time-dependent pricing for an Internet service provider. The time-dependent pricing is proven to alleviate the maximum demand at peak hours and increase the demand at under-utilized hours by providing incentives to deviate from peak hours.

Most previous works on congestion-dependent pricing in access networks assume a single capacity sharing network. In this paper, we assume heterogeneous networks which have different characteristics in terms of throughput and coverage and study congestion-dependent pricing and conventional pricing mechanisms in heterogeneous networks.

2. System model

In this section, we describe a system model for the operation of a service provider and users. We assume that a cell of the cellular network provider has a total capacity, C_{3g} in Mbps. Subscribers in a cell share the fixed cellular capacity. Also, they can access WiFi networks when they come in WiFi coverage areas. We assume that WiFi networks are provided by the cellular network provider freely or users have authentication to several WiFi access points (APs). The WiFi capacity of users who enter WiFi coverage areas is assumed to be fixed to C_w in Mbps. We assume that there are N users in the cell. Mobile data traffic of users arrives according to a Poisson process and the file size B is exponentially distributed with μ. We only consider best-effort data services for simplicity. If a file is served successfully, the user pays a fee of u_hB where u_h is nonnegative unit price (in \$/Byte) at time interval h and B is the file size. We assume a known *demand function* $\lambda_{i,h}(u_h)$ at time interval h, as a function of the unit price. The *demand function* determines the traffic arrival rates of users. We denote $\boldsymbol{\lambda}_h(u_h) = \sum_{i=1}^{N} \lambda_{i,h}(u_h)$. Intuitively, we assume that $\lambda_{i,h}(u_h)$ decreases as u_h increases. We assume that the probability that a file is delay-tolerant of 1-α and the maximum allowable deadline of it is D.

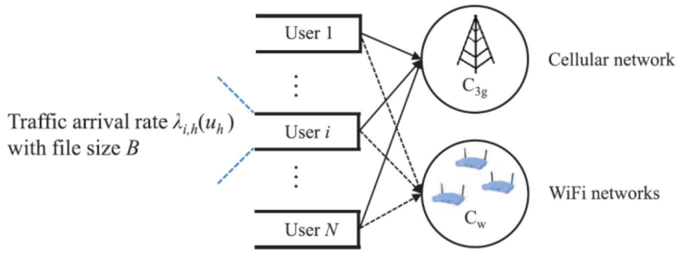

Figure 1: An illustration of traffic arrival and service in heterogeneous networks

Table 1: Summary of major notation

Variable	Definition
N	Number of users
C_{3g}	The cellular capacity
C_{3g}	The WiFi capacity
$m(t)$	The number of active users at time t
$u(h)$	Unit price at time interval h (\$/Byte)
$\lambda_{i,h}(u)$	Traffic arrival rate of of user i at time interval h when the unit price is u
$U_{i,h}$	Random variable of utility of a file
B	File size (exponentially distributed with μ)
α	Fraction of real-time traffic
D	Deadline of insensitive traffic

Let $m(t)$ be the number of active users whose data transfer is in progress at time t. The cellular capacity of active users at time t is assumed to be $C_{3g}=m(t)$. We assume that mobile data traffic has the same bandwidth requirement, r, so that an incoming file is admitted if and only if $(m(t)+1)r < C_{3g}$. We denote the maximum number of active users as $N_{max} = \lceil C_{3g}/r \rceil < N$.

A pricing policy determines the unit price u_h at each time interval h. We assume that the arrival process in L time intervals $s_1,\ldots, s_L$ is periodic (e.g. $L = 24$, $s_1 = [0 : 00, 1 : 00],\ldots, s_{24} = [23 : 00, 24 : 00]$) such that for all $k = 1,\ldots,L$,

$$\lambda_{i,h}(u_h) = \lambda_i^k(u_k), h \in s_k. \tag{1}$$

The expected revenue of the access network provider in a period or L time intervals is the following.

$$\sum_{h=1}^{L} \boldsymbol{\lambda}_h(u_h)u_h\mu^{-1}\mathbf{P}\{m(t) < N_{max}\} \tag{2}$$

where $\mathbf{P}\{m(t) < N_{max}\}$ is the probability that the network can serve the data traffic and μ^{-1} is the average file size. We further assume that a potential data file with the arrival rate $\lambda_{i,h}(0)$ of user i in time interval h results in a user utility of $U_{i,h}$ if it is transmitted through a cellular network. $U_{i,h}$ is a nonnegative random variable and the unit is the same as price. We assume that a potential data file will go through if and only if the utility $U_{i,h}$ is greater than the price u_hB, i.e., $\lambda_{i,h}(u_h) = \lambda_{i,h}(0)\, \mathbf{P}\{U_{i,h}(t) \geq u_hB\}$. To model delayed offloading, we assume that the utility of a file decreases as the delay increases. We denote $V(d)$ as a disutility function where d is the delay of the file such that $U_{i,h}(d) = U_{i,h}\ V(d)$. We assume that delay is dependent on a specified deadline for a file and WiFi contact probability.

3. Numerical results

In this section, we consider a scenario that demand is variable in time and numerically compare the performance of volume pricing and congestion-dependent pricing. In volume pricing, a network provider chooses a single unit price that is always in effect and maximizes the expected revenue in (2). In congestion-dependent pricing, a network provider chooses a unit price vector with the same objective. We set $N = 100$, C_{3g} = 8Mbps, C_w = 2Mbps, $\mathbf{E}[B]$ = 1MB. We assume that $s_{daytime} = [9:00, 21:00]$, $s_{night} = [21:00, 9:00]$ span a 24-h period and the arrival process is periodic. The utility $U_{i,h}$ is assumed to be uniformly distributed in [0, 0.1] and [0, 0.05] for daytime and night, respectively. The total potential demand is set to be 0.02 and 0.005 file per second for daytime and night, respectively. We suppose that users can choose whether they set the deadline as 1 hour or 0. When a user set the deadline as 1 hour, the WiFi contact probability is assumed to be 0.7 for all users. This is a result from our quantitative study in [3]. Table 1 summarizes the numerical results. The expected revenue of congestion-dependent pricing is 13% greater than that of volume pricing. Expensive unit price at peak hours enforces users to choose WiFi offloading or deviate from peak hours. As a result, network congestion is alleviated while the total revenue increases.

Table 2: Parameter values in numerical examples

Parameter	Value
N	100
C_{3g}	8Mbps
C_w	2Mbps
$\mathbf{E}[B]$	1MB
$U_{i,h}$	Uniform[0,0.1] at daytime Uniform[0,0.05] at night
$\boldsymbol{\lambda}_h$	0.02 files/sec at daytime 0.005 files/sec at night
α	0

Table 3: Numerical comparison between volume and congestion-dependent pricing

	Volume	Congestion-dependent
Optimal unit price at daytime	0.092 \$/MB	0.096 \$/MB
Optimal unit price at night	0.092 \$/MB	0.025 \$/MB
Expected revenue per day (24-h)	2055\$	2323\$

4. Conclusion and future work

A famous economist, W. Vickrey, proposed principles of efficient congestion pricing in 1992. He addresses that charges should reflect the marginal social cost in terms of the impacts on others as closely as possible. In this point of view, widely used pricing schemes in mobile data services such as flat, volume, and tiered pricing are not efficient in terms of congestion control since they fail to reflect the marginal social cost.

While congestion pricing results in better performance than other pricing schemes, there are some drawbacks in congestion pricing. Congestion pricing requires complex control to estimate user response (potential demand, user utility) and calculate the optimal unit price vector in each time interval. Also, it may be unattractive to users who prefer fixed or predictable charge. A step-by-step approach (e.g. from the most congested time and location) may be helpful in introducing congestion pricing.

In this paper, we see the possibility of congestion-dependent pricing in heterogeneous networks. As a future work, we will derive some theoretical and analytical results about congestion-dependent pricing in heterogeneous networks. A unified pricing system that embraces demand profiling and network investment is also of interest.

5. References

[1] Cisco Systems Inc. Cisco visual networking index: Global mobile data traffic forecast update, 2010-2015, 2011. http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html

[2] Aruna Balasubramanian, Ratul Mahajan, and Arun Venkataramani. Augmenting mobile 3g using WiFi. In *Proceedings of ACM MobiSys*, 2010.

[3] Kyunghan Lee, Injong Rhee, Joohyun Lee, Yung Yi, and Song Chong. Mobile data offloading: How much can WiFi deliver? In *Proceedings of ACM CoNEXT*, 2010.

[4] William S. Vickrey. Congestion theory and transport investment. *American Economic Review*, 59(2):251-260, 1969.

[5] Todd Litman. London congestion pricing, 2004. Victoria Transport Policy Institute.

[6] Severin Borenstein. The long-run effects of real-time electricity pricing. *The Energy Journal*, 26(3):93-116, 2005.

[7] Frank Wolak. Residential customer response to real-time pricing: the anaheim critical-peak pricing experiment, 2006.

[8] Ioannis Ch. Paschalidis and John N. Tsitsiklis. Congestion-dependent pricing of network services. *IEEE/ACM Transactions on Networking*, 8(2):171-184, 2000.

[9] Carlee Joe-Wong, Sangtae Ha, and Mung Chiang. Time-dependent broadband pricing: Feasibility and benefits, In *Proceedings of IEEE ICDCS*, 2011.