A

Pricing Data: Past Proposals, Current Plans, and Future Trends

SOUMYA SEN, Princeton University
CARLEE JOE-WONG, Princeton University
SANGTAE HA, Princeton University
MUNG CHIANG, Princeton University

Traditionally, network operators have used simple flat-rate broadband data plans for both wired and wireless network access. But today, with the popularity of mobile devices and exponential growth of apps, videos, and clouds, service providers are gradually moving towards more sophisticated pricing schemes, including dynamic pricing. This decade will therefore likely witness a major change in ways in which network resources are managed and the role of economics in allocating these resources. This survey reviews some of the well-known past broadband pricing proposals (both static and dynamic), their current realization in various consumer data plans around the world, and discusses several research challenges and developments. It also provides a brief overview of the relationships between various pricing ideas in broadband data plans and their counterparts in other markets, such as electricity and transportation networks. By exploring the benefits and challenges of pricing data, this paper attempts to facilitate both the industrial and the academic communities' efforts in understanding the existing literature, recognizing new trends, and shaping an appropriate and timely research agenda.

Categories and Subject Descriptors: C.2.0 [**General**]: Data Communications—*Internet, Pricing*; C.2.3 [**Network Operations**]: Network Management—*Congestion Management*; K.6.0 [**Management of Computing and Information Systems**]: General—*Economics*

General Terms: Broadband Networks, Network Economics, Congestion Management

Additional Key Words and Phrases: Congestion, Broadband pricing, Data plans, Economics

ACM Reference Format:
Sen, S., Joe-Wong, C., Ha, S., Chiang, M. 2012. A Survey of Broadband Data Pricing: Past Proposals, Current Plans, and Future Trends ACM V, N, Article A (February 2012), 34 pages.
DOI = 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

1. INTRODUCTION

In November 1996, AOL, the largest U.S. Internet Service Provider (ISP) of the time, switched from hourly pricing to a flat-rate monthly data plan of $19.95 [Lewis 1996]. One week later, Pacific Telesis Enterprises CEO M. Fitzpatrick made an apocalyptic prediction [Marshall 1996]:

> *My indicators point to a genuine data tsunami, a tidal wave, in the next 18 months. And, ladies and gentlemen, while we can surf the net, nobody can surf a tidal wave.*

Fortunately, that devastating tsunami did not come, but the tides of growth in the demand for data did help the Internet grow.

However, today, the fear of a data tsunami is slowly coming back [Carew 2012]. The Cisco Visual Networking Index [Cisco Systems 2011] predicts that global IP traffic (including managed IP traffic[1]) will reach 890.5 exabytes/year and mobile data will reach 75 exabytes/year by the end of 2015. Thus, mobile data traffic[2] and global consumer IP traffic are expected

[1]Cisco's definition of "consumer" includes fixed IP traffic generated by households, university populations, and Internet cafes, "mobile" includes mobile data and Internet traffic generated by handsets, notebook cards, and mobile broadband gateways, "Internet" denotes all IP traffic that crosses an Internet backbone, and "managed IP" includes corporate IP WAN traffic and IP transport of TV and VoD.
[2]"Mobile data traffic" includes handset-based data traffic, such as text messaging, multimedia messaging, and handset video services. Mobile Internet traffic is generated by wireless cards for portable computers and handset-based mobile Internet usage.

to grow between 2010 and 2015 at a CAGR (Compound Annual Growth Rate) of 92% and 34%, respectively. With the growing popularity of iPhones, iPads, bandwidth-hungry applications, and cloud-based services, ISPs have been increasingly turning to pricing as the ultimate congestion management tool to mitigate this growth –a feat often achieved by imposing harsh overage penalties and "throttling" the very customers who drive this demand.

The basic idea of congestion-based pricing has existed for several decades in many different utility markets, e.g. transportation, electricity distribution networks. As for the Internet, even as early as 1997, when it was just evolving into a commercial service and its revenue models were hotly debated, MacKie-Mason et al. [1997] wrote:

> *We argue that a feedback signal in the form of a variable price for network service is a workable tool to aid network operators in controlling Internet traffic. We suggest that these prices should vary dynamically based on the current utilization of network resources.*

They were proposing a form of dynamic congestion pricing for the Internet. As a mathematically oriented research topic, congestion pricing has been extensively studied in transportation networks, energy markets, telephone networks, ATM and IP networks, etc. Thus, in some sense, there are not many new statements to be made about the generic theory of congestion pricing. And yet, wireless and wireline data networks have so far used the most rudimentary forms of congestion pricing, e.g., usage-based charges. This can partly be explained by Clark's [1997] observation:

> *Whatever the providers may prefer to do, competition may force some forms of fixed-fee pricing in the marketplace.*

While providers may have preferred flat rates in the 1990s, the situation is quite different now: the problem has worsened and the operators are more aggressive in pursuing new pricing schemes. The acceptance of this fact is perhaps nowhere more clearly evident than in the "New Rules for an Open Internet" [Genachowski 2010] announced on December 21, 2010, by FCC chairman J. Genachowski:

> *The rules also recognize that broadband providers need meaningful flexibility to manage their networks to deal with congestion.... And we recognize the importance and value of business-model experimentation.*

Therefore, the interesting question is: given the rapid rise in capacity demand in the age of apps and clouds, how will pricing policies change over this decade?

The shift in pricing trends is more easily noticeable in growing economies, e.g. India and Africa, where dynamic congestion pricing for voice calls is already being practiced. The next logical step in this pricing evolution is dynamic pricing for data. But pricing data has several unique challenges (and opportunities), arising from both technological and social considerations. A clear understanding of these issues requires consideration of both the past and the present developments. To this end, this paper reviews some of the best-known pricing proposals of the last two decades and report on some very interesting pricing schemes that are in use today.

1.1. Contributions

In 1996, Breker [1996] conducted a survey of computer network pricing schemes that covered a few of the known protocols of the time. A similar review of pricing concepts for broadband IP networks was undertaken about a decade ago by Falkner [2000]. Since then, researchers have proposed several new schemes for pricing data and some variations of these plans have been adopted by wireless ISPs for their consumers. Ezziane [2005] surveyed on the charging and billing mechanisms used by 3G wireless service providers. In particular, it reported on new approaches needed for mediation, billing, and charging to these new services. More recently, Gizelis [2011] provided a detailed overview of the literature on pricing schemes for wireless networks. An annotated bibliography of several papers related to various aspects of Internet economics, including pricing, was compiled by Klopfenstein [2009]. Our work not only complements these earlier surveys and reports on broadband pricing, but also attempts to understand the kinds of pricing plans currently offered by wired and wireless ISPs around the world, while noting some of the weaknesses in their approaches. Also, it should be noted that the focus of this work is on pricing for the end-user/consumer and not the economics of peering or transit-fee agreements between service providers. The key features of this work are summarized below:

— A broad overview of several past and recent proposals for pricing data is provided, along with a classification of static and dynamic pricing plans.
— Several illustrative examples of real pricing plans implemented by ISPs in different parts of the world.
— A comprehensive overview of the relationships between ideas of broadband data pricing and earlier congestion pricing ideas in road and electricity networks.
— Emerging trends and future social, economic, and technological challenges, particularly for dynamic pricing and app-based pricing.
— A brief description of prepaid and postpaid plans, and their adoption in various parts of the world.
— Satellite broadband pricing plans for consumers.

Networking researchers and network managers, caught in their conscientious efforts to find the best solutions to this growing problem, often tend to overlook many of the innovative practices already deployed by real network operators. However, appraising the past several years of theoretical foundations on pricing and identifying their varied realizations in the present world will likely be key to predicting future trends and shaping an appropriate research agenda. The information presented in this work draws from a wide range of sources, including research proposals, existing data plans, news articles, consumer forums, and reports from experimental field trials.

This paper is organized as follows: Section 2 provides an overview of pricing plans by subscription type, i.e., prepaid and postpaid plans and reports on their adoption in different parts of the world. A detailed review of proposals and realizations of several static and dynamic pricing plans are provided in Sections 3 and 4, respectively. The need and challenges for innovating pricing plans, particularly in the dynamic pricing of mobile data are discussed in Section 5. Section 6 reports on some of the promising trends in pricing innovations, including the availability of satellite broadband for end-users. Section 7 concludes the paper. The Appendix provides additional reference materials on similar pricing policies that have been used in electricity grids and road networks, highlighting their applicability to designing new broadband congestion pricing schemes for the future.

2. PRICING PRACTICES

Broadly speaking, broadband data plans can be divided either by their subscription type (prepaid or postpaid), or by pricing mechanism, i.e., static or dynamic pricing. We begin with a brief discussion of prepaid and postpaid plans and their predominance in various parts of the world.

2.1. Prepaid and Postpaid Plans

Prepaid plans are those in which a consumer pays for his/her data usage beforehand, while for postpaid plans, a consumer is billed for his/her monthly usage at the end of the billing cycle. In this section, we provide an overview of prepaid and postpaid plan adoption in different parts of the world and its relationship to consumer market demographics.

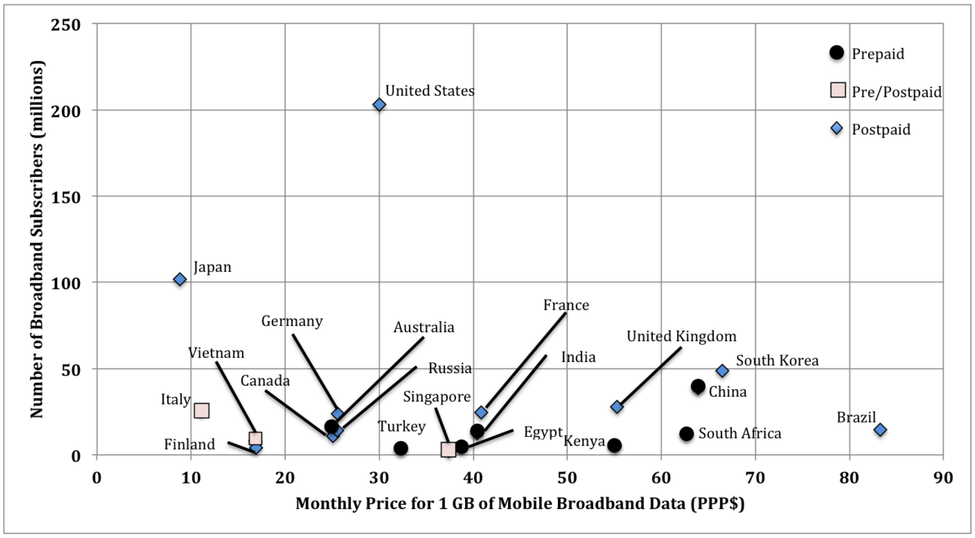

Fig. 1. Prepaid and postpaid mobile broadband data plans around the world. Data were taken from operator websites as well as [86CallChina 2011; BBC 2011; CIA 2011; Communications Commission of Kenya 2011; EMISR 2011; IDA Singapore 2011; ITU 2011; Lancaster 2011; Malik 2011; NTT DoCoMo 2011; OECD 2011; Telecomandinternet 2010; Telecompaper 2010b; TeleGeography 2010; TWA Netwerk 2011; VietnamNet Bridge 2011].

Figure 1 shows whether prepaid or postpaid plans are the dominant form of data plans in several countries of the world. To determine this data, we examine the standard mobile subscriptions for the largest mobile operator in each country and plot the monthly cost (in $ after accounting for purchasing power parity) of 1GB of data usage. We note that the data plans in different countries have different broadband caps, which are not reflected in the graph, but may be found in [ITU 2011]. The cost of 1GB of data/month has been plotted against the number of mobile broadband subscribers in each country; we observe that there is a large variation in cost for 1GB of mobile data. Additionally, we see that the countries of Africa (e.g., Kenya, South Africa), the Middle East (e.g., Turkey, Egypt), India and China tend to prefer prepaid plans, while postpaid plans dominate in Europe and the Americas.

Figure 2 more explicitly shows the correlation between prepaid mobile broadband data plans and lower per capita gross national incomes (GNI): we see that countries with a lower GNI per capita tend to have prepaid plans as the dominant subscription mode (with the exception of Australia). Moreover, two of the three countries offering both pre- and postpaid plans (Italy and Vietnam) have GNI per capita below the wealthiest 9 countries.

This observation may be partly explained by studies examining the popularity of prepaid voice call plans in lower-income groups. Castells et al. [2007] found a strong correlation between the availability of prepaid plans for voice calls and adoption rates of mobile telephone subscribers, which is corroborated by Hauge et al. [2009]. The authors attribute this correlation to cost-conscious consumers' desire to control their expenditures. Indeed, Donovan and Donner's study [2010] of mobile broadband availability in Africa focuses exclusively on the availability of prepaid data plans, indicating their importance for mobile broadband adoption in Africa.

Within the U.S., prepaid data plans have gained popularity due to their lower, shorter-term financial commitment and relative simplicity compared to postpaid usage-based pricing plans with data caps. Moreover, operators are offering more attractive prepaid devices with data-centric plans that appeal to young, price-sensitive users [Kaputska 2010].

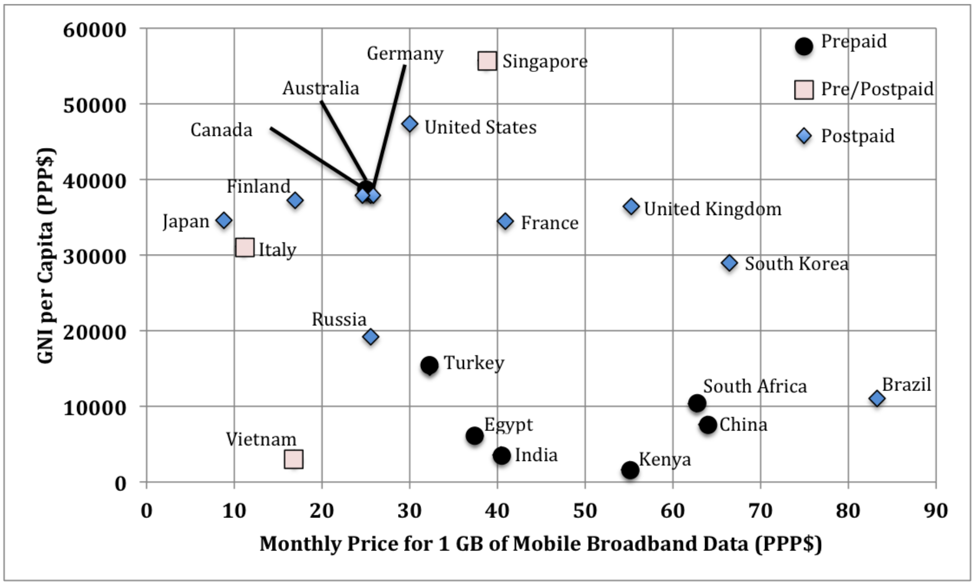

Fig. 2. GNI per capita versus the monthly cost of 1GB of mobile broadband data. Data sources as in Fig. 1.

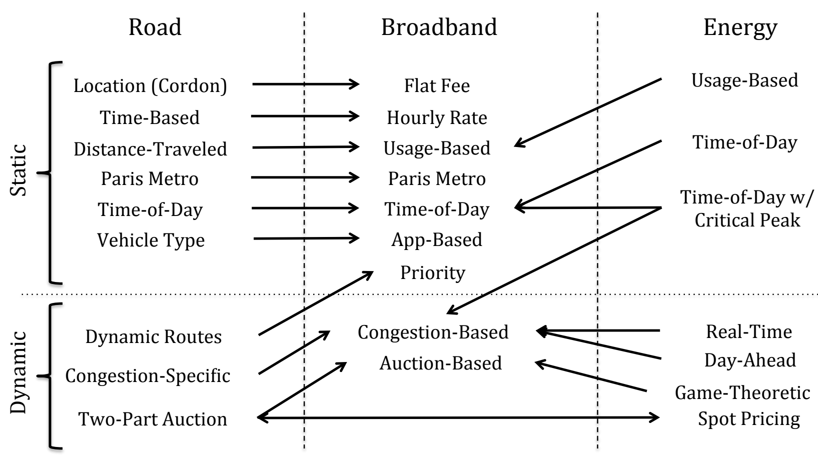

Fig. 3. Relationships between some of the pricing plans proposed for broadband, electricity and road networks.

2.2. Schemes for Pricing Data

Next, we focus on the pricing schemes, namely the different forms of static and dynamic plans, that have been explored within the networking community. Many of these broadband pricing plans have been heavily influenced by previous pricing proposals for other utility markets, like transportation and energy. Figure 3 gives an overview of the logical relationships between pricing principles from these markets that have led to similar plans for data networks. While this paper focuses mainly on proposals and realizations of pricing plans for data networks, readers interested in static and dynamic pricing for electricity and road networks are referred to the Appendix for additional details on each of these pricing schemes.

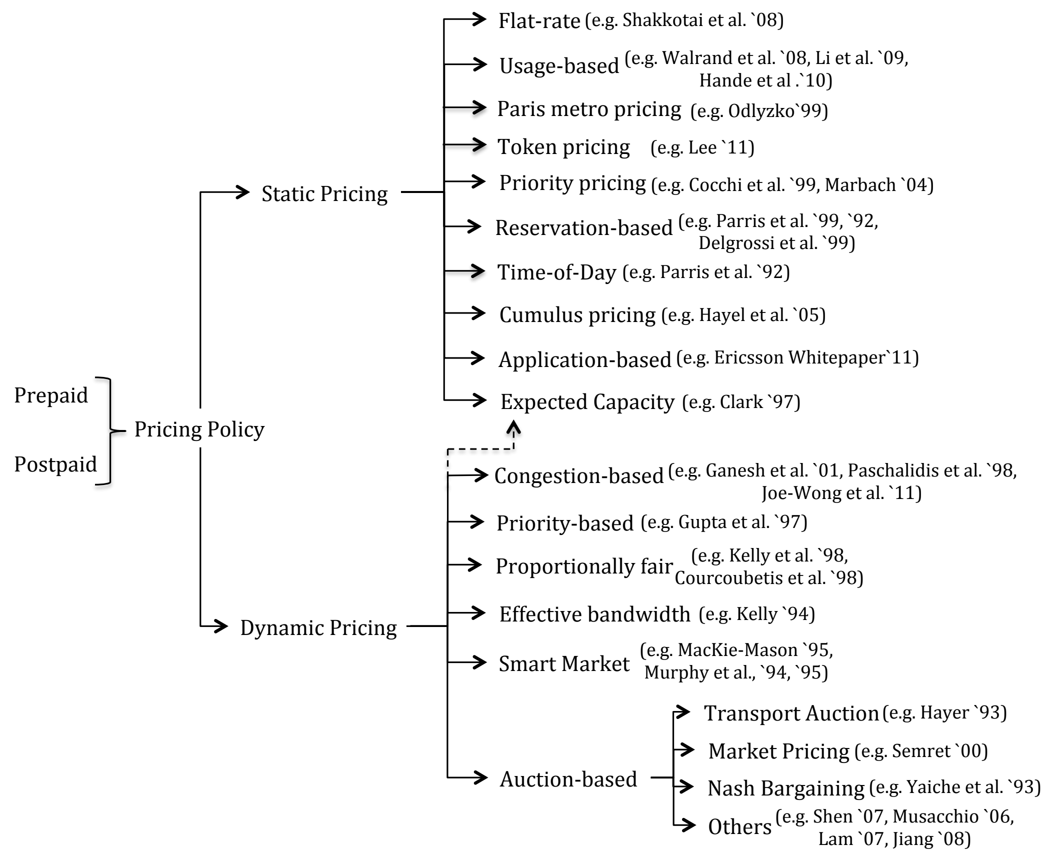

Fig. 4. Taxonomy of selected pricing strategies and some related proposals.

Sections 3 and 4 describe some of the best-known static and dynamic pricing practices, respectively, and their realizations in the real world as broadband data plans, wherever applicable. In general, most of the proposed pricing schemes, although of much academic interest, have so far seen little adoption by ISPs [Wiseman 2000]. Figure 4 gives a visual taxonomy of the pricing plans discussed in this paper and the related literature.

3. STATIC PRICING

3.1. Fixed Flat-Rate Pricing

3.1.1. Monthly Rate. In the past, ISPs have charged users a flat monthly fee for broadband access, irrespective of the actual time spent on the network or data usage. Such a simple pricing model was praised by Anania and Solomon [1997] and Odlyzko [2001] based on historical and political precedence. A more theoretical argument was made by Shakkotai et al. [2008], showing that simple flat-rate pricing is quite efficient in extracting revenue for elastic traffic if all users run an identical application, but with different valuations for the application.

There are several variations of the monthly flat-rate pricing model in the real world. For example, "unlimited" data plans put no cap on the bandwidth used every month. If a maximum usage limit is predetermined according to a flat price, the plan is called "flat up to a cap." Exceeding the limit usually incurs penalty costs that are proportional to the usage above the cap, i.e. "metered." "Tiered" data plans with different flat-rate pricing for different usage caps are often used to provide a range of choices to the consumers.

Usually, users on flat-up-to-a-cap plans are shifted to usage-based pricing or a higher-priced tier upon exceeding their data limit. But another emerging trend is "flat to a cap, then throttle." Orange Spain offers the Delfin plan at £40/month for smartphones, which gives unlimited access up to 500 MB at full speed, and thereafter limits users' bandwidth to128 kbps [Sephton 2011b].

Other forms of penalty for overshooting the cap include service discontinuation; Comcast, the largest U.S. cable TV operator, has introduced a "flat up to a cap" plan of 250 GB/month. Any user exceeding the cap twice within six months can be subjected to one year termination of service [Comcast 2010].

Without such penalties, flat-rate plans will simply become unviable for ISPs. As if to drive home that point, TelstraClear, New Zealand's second largest telco, experimented with a free weekend plan during which they switched off their usage metering and removed data caps from Friday evening to Sunday midnight of December 2-5, 2011. When all their data hungry customers simultaneously descended on the net with whetted appetite for the big weekend buffet, many were dismayed to find their speeds down to one-fifth of the usual [Keall 2011].

3.1.2. Hourly Rate. Some providers have begun offering flat-rate plans in which mobile internet services are billed by the hour, i.e., the cap is specified in terms of time instead of usage. For example, USB modem customers of the Egyptian mobile operator Mobinil can choose from packages of 30 hours for EGP 80/month or 60 hours for EGP 125/month. An additional five hours can be bought for EGP 20 [Telecompaper 2010a]. Other plans add a maximum usage cap for each day. NTT DoCoMo in Japan offers similar, prepaid plans, designed for Sony's Playstation devices [NTT DoCoMo 2011]. DoCoMo charges 980 yen for 20 hours of 3G access over one month or 4980 yen for 103 hours of access over three months.

Although flat-rate billing is cheap to implement and operate, encourages user demand, and creates simple and predictable monthly fees for customers, it suffers from several disadvantages. First, flat-rate plans lead to inefficient resource allocation and market segmentation, with low-usage customers typically subsidizing the heavy users (i.e., the bandwidth hogs) [Hendershott 2006]. Second, while ISP revenues depend on the median user, peak load costs are driven by the heavy users, which creates a price-cost mismatch.

These disadvantages of flat-rate pricing were well summarized by Varaiya while motivating his vision of the Berkeley INDEX project [1999]:

> *Although flat-rate continues to be the pre-dominant form in which Internet access is sold, that form of pricing is unviable. Flat-rate pricing encourages waste and requires 20% of users who account for 80% of the traffic to be subsidized by other users and other forms of revenue. Furthermore, flat-rate pricing is incompatible with quality-differentiated services.*

Consequently, as the demand for bandwidth grows, most ISPs are replacing their flat-rate plans with usage-based or "metered" data plans.

3.2. Usage-Based Pricing

New Zealand conducted an early experiment with usage-based pricing in 1989, when six participating universities agreed to pay by the volume of traffic sent in either direction through the NZGate gateway [Brownlee 1997]. Such a pricing plan was necessitated by New Zealand's geographically isolated location, expensive trans-Pacific links, and the lack of subsidies from the New Zealand government [McKnight and Bailey 1997]. This arrangement provided an early demonstration that metering traffic and charging users by daily volume is feasible.

"Metered" implies that a user is charged in proportion to the actual volume of data usage [Walrand 2008; Li et al. 2009]. In practice, operators often use "cap then metered" plans (also known as "usage-based" pricing), for which a user pays a flat price up to a predetermined volume of traffic, beyond which the user is charged in proportion to the volume of data consumed [Hande et al. 2010].

Tiered, cap then metered is currently the dominant pricing model in the U.S. On June 7, 2010, AT&T introduced a $15/month plan for 200 MB and $25/month for 2 GB of data, along with different rates of overage charges for the two tiers [Kang 2010; Frakes 2010]. Following in the footsteps of T-Mobile and AT&T, Verizon Wireless also introduced monthly plans of $30 for 2 GB, $50 for 5 GB, or $80 for 10 GB, with a $10/GB overage charge [Segall 2011]. Moreover, AT&T has introduced caps even for its wireline service: users pay a flat fee for up to 150 GB for DSL and 250 GB for U-Verse per month, and are then charged $10 for each 50 extra GB [Taylor 2011].

This move towards usage-based pricing has occurred outside of the U.S. market as well. In June 2010, the U.K.'s second largest operator, Telefonica-owned O2, announced an end to its "all-you-can-eat" data plan, with similar moves suggested by Orange and T-Mobile, Vodafone, and Hutchison Whampoa's 3UK [Parker 2010]. Similarly, tiered usage-based data plans are seen as the pricing scheme of choice for LTE networks, and have been adopted by Korea's KT and LG U+, Japan's NTT DoCoMo, and Hong Kong's CSL [Morgan 2011].

While proponents of usage-based pricing see it as a means to create incentive compatibility for efficient network resource utilization, such pricing plans create a new set of challenges regarding user adoption and demand loss, increase the complexity of billing and monitoring network performance, and charge customers irrespective of congestion levels in the network. Additionally, usage-based pricing still fails to address the problem of large peak-load costs incurred from many users accessing the network resources at the same time. Clark succinctly summarized these issues in [1997]:

> *The fundamental problem with simple usage fees is that they impose usage costs on users regardless of whether the network is congested or not.*

3.3. Paris Metro Pricing

Paris Metro Pricing (PMP) was proposed by Odlyzko [1999] as a simple and elegant solution for creating differentiated service classes. It proposed partitioning the network resources into several logical traffic classes, each of which is identical in its treatment of data packets but charges users differently. Thus, users willing to pay more will select the more expensive, and hence less congested, logical traffic class.

PMP is designed to enable maximum simplicity for the end-user in expressing his/her user preference through self-selection of the desired service level. But Odlyzko also identified some potential problems to overcome to implement PMP, such as finding ways to set prices and capacities of the logically separate channels, and to have a predictable performance of different channels to avoid network instability. PMP may also require better designs for the user interface to let users dynamically alter their preferences and assign application sessions to different traffic classes.

3.4. Token Pricing

Lee et al. [2011] introduce a token pricing scheme in which users pay a fixed falt-rate monthly fee for Internet access. Each user receives a certain number of "tokens" from the service provider, who offers two service classes of different qualities. The basic quality class requires no tokens for access but may become congested. The higher quality class requires users to redeem some tokens, and hence has lesser congestion and a better service because of a Paris Metro Pricing effect. Thus, users self-prioritize their sessions and implicitly pay for the higher QoS with their tokens for particularly urgent sessions during peak network congestion. Each session, regardless of size, costs the same number of tokens, and users continually receive more tokens (e.g. one every ten minutes). The benefits of such a system in reducing peak network congestion in a real network is yet to be explored.

3.5. Priority Pricing

Cocchi et al. [1991; 1993] studied a pricing scheme in a multiple service network with priority classes, but without any resource reservation. Users can request different qualities of service (QoS) by setting bits in their packets. A higher priority class charges a higher per-byte fee but is assumed to receive better service from the network. Thus, users who pay a greater per-byte fee for higher priority are in effect paying for the negative externality imposed on traffic from other, lower priority users. The authors showed that such quality sensitive pricing is more efficient (in a Pareto sense) than flat pricing. However, this result depends on the reservation-less assumption.

A non-cooperative game-theory framework was used to analyze a static priority pricing by Marbach [2004]. In his single-link model, users assign a priority class to their packets and are charged accordingly. These charges are based on the packets submitted to the network rather than their actual delivery. Marbach shows that there always exists a Wardrop equilibrium bandwidth allocation, but it is not necessarily unique. This equilibrium allocation and the associated link revenue do not depend on the prices of the different priority classes, but the prices do have a simple relation with the packet loss probability due to congestion in that class.

One disadvantage of priority pricing is users' inability to express their desired levels of delay and bandwidth share; priority pricing's consistent preferential treatment of higher priority classes might drive lower priority classes to little or no usage. Priority pricing is largely absent today, but some ISPs are considering the idea of creating a "priority data plan" in which users of the premium service are prioritized during periods of network congestion. Recently, SingTel of Singapore introduced such an option, called "priority pass," for its top-tier dongle customers [Kwang 2011].

3.6. Reservation-Based Pricing

Parris et al. [1992] were one of the first to study pricing in a reservation-oriented network. They considered the issues of network utilization, ISP revenue, and blocking probability under per-packet, setup, and peak-load pricing schemes. In their work, users are characterized by their connection durations, budgets, and chosen classes of service (with a higher per-byte fee for a higher priority service class). The network decides to either accept or block the connection, depending on the sufficiency of user's budget and availability of network resources. Using simulations, the authors show that for a given pricing scheme, price increases will at first increase and then eventually decrease the net revenue, but will always decrease the blocking probability and network utilization. Setup pricing decreases the blocking probability and increases revenue for the ISP, and more generally performs better than per-packet pricing (in that the blocking probability from admission control is lower under setup pricing than per-packet pricing for the same level of revenue generated).

Despite these advantages, Parris et al.'s [1992] form of reservation pricing suffers from some shortcomings. First, the idea of having to pay a flat-rate setup cost is unfair towards those users with shorter conversations. Second, poorer users may not be able to afford a connection under a high set up cost, thus leading to a greater digital divide. Third, average network utilization is typically lower in presence of setup costs, thus demonstrating a tradeoff between network efficiency and revenue maximization.

Parris and Ferrari [1992] presented another reservation pricing scheme for real-time channel establishment. Under this plan, users are charged based on the type of service requested, as measured by factors such as the bandwidth, buffer space, and CPU time resources reserved, and the delay imposed on other users. The total charge a user pays is a product of the type of service measure, channel duration, and a time of day factor. However, the work does not provide any clear guidelines on how economic considerations are to be mapped to a single time-of-the-day factor or the impact of the overhead associated with estimating the network parameters in real-time.

In a later work, Delgrossi and Ferrari [1999] consider a pricing scheme based on the portion of resource capacity used by a reserved data channel in a multiple-service network. They introduce a charging formula with different reservation and transport cost components for real-time and non-real-time traffic, along with a discussion on computing resource capacity requirements for the channel as a function of the buffer, processing power, and schedulability.

3.7. Time-of-Day Pricing

Time-of-day or ToD pricing schemes charge peak and off-peak hours differently, so as to disperse user demand more uniformly over time. Parris et al. [1992; 1992] considered a form of ToD combined with reservation-based pricing, which divides a day into peak and off-peak periods and incorporates the time elasticity of user demand. They show that peak-load pricing reduces peak utilization and the blocking probability of all traffic classes, and increases revenue by inducing a more even distribution of demand over peak and off-peak periods.

The most basic form of ToD in practice is a two-period plan that charges different rates during the daytime and night time. For example, BSNL in India offers unlimited night time (2-8 am) downloads on monthly data plans of Rs 500 ($10) and above. Other variations of ToD pricing are offered elsewhere; for instance, the European operator Orange has a "Dolphin Plan" for £15 ($23.58 USD) per month that allows unlimited web access during a "happy hour" corresponding to users' morning commute (8-9 am), lunch break (12-1 pm), late afternoon break (4-5 pm), or late night (10-11 pm).

3.8. Expected Capacity Pricing

In 1997, Clark [1997] wrote

> *In the future it will be desirable to provide additional explicit mechanisms to allow users to specify different service needs, with the presumption that they will be differentially priced.*

He proposed expected capacity pricing as a mechanism to allow users to explicitly specify their service expectation (e.g., file transfer time), while accounting for differences in applications' data volume and delay tolerance. The idea is that by entering into profile contracts for expected capacity with the operator, different users should receive different share of network resources *only* at times of congestion [Songhurst 1999].

One specific proposal to realize this service involved traffic flagging (i.e., each packet is marked as being *in* or *out* of the user's purchased profile, irrespective of network congestion level) by a traffic meter at access points where the user's traffic enters the network. This is followed by congestion management at the switches and routers where packets marked as *out* are preferentially dropped during congested periods, but are treated in an equal best-effort manner at all other times. The expected capacity is thus not a capacity guarantee from the network to the user, but rather a notion of the capacity that a user expects to be available and a set of mechanisms that allow him or her to obtain a different share of the resource at congested times. This pricing can be simply enforced at the router and switches of the network, and allows service providers to have more stable estimates of the future necessary capacity based on the total expected capacity sold, rather than the sum of peak rates of all users' access links. A dynamic pricing version of the scheme is also explored in [Clark 1997]. However, in order to implement this pricing scheme, the issue of assigning price value to the expected capacity profiles requires further study.

3.9. Cumulus Pricing

Cumulus pricing schemes (CPS) consist of three stages: specification, monitoring, and negotiation. A service provider initially offers a flat-rate contract to the user for a specified period based on the user's estimate of resource requirements. During this time the provider monitors the user's actual usage and provides periodic feedback to the user (by reporting on "cumulus points" accumulated from their usage) to indicate whether the user has exceeded the specified resource requirements. Once the cumulative score of a user exceeds a predefined threshold, the contract is renegotiated. Hayel and Tuffin [2005] study such a scheme and use simulated annealing to optimize the total network revenue in terms of the renegotiation threshold.

CPS is a simple pricing scheme that can be easily implemented at the network edge. Some ISPs have been experimenting with similar ideas. For example, Vodafone in the U.K. announced a new "data test drive" plan that allows customers joining any of the monthly pay contracts to have unlimited data access (including tethering, but excluding roaming) for the first three months. The data usage report is then fed back to the user to negotiate whether the chosen plan is appropriate for them. The user can choose to either continue with existing plan, possibly incurring overage charges, or switch to an alternative plan suggested by Vodafone [Sephton 2011c].

3.10. Application and Content-Based Pricing

Several mobile service providers have been experimenting with various forms of application or content-based pricing. While the currently available app-based data plans are mostly designed to attract and "lock in" customers by bundling content and data plans, they highlight an emerging trend of operators experimenting with pricing structures that charge (or subsidize) differently based on application type [Ericsson 2011; Higginbotham 2010]. In 2009, Three in the U.K. offered its customers access to two years of Spotify premium (music streaming on-demand) with HTC Hero Android phones for a £35/month plan [Sephton 2011a]. Similarly in 2011, Telus, an operator in Canada offered a free six-month subscription to Rdio (music streaming) to its subscribers who purchased a Rdio-supported smartphone and data plan [Vardy 2011] and the Danish operator, TDC (Tele-Danmark Communications), bundled the cost of accessing its streaming music service, TDC Play, into its mobile data plans. Another innovative operator, Orange France, has been actively bundling access to pay TV, Video on Demand, and streaming music services in partnership with Deezer. Orange UK has introduced "swapables," offering two free mobile media services to its £35/month Panther data plan users. These services may be swapped for others every month [Owoseje 2011].

Besides subsidized content streaming, another new trend in app-specific pricing for mobile social network services[3] is also emerging. With the introduction of standards like the "Traffic Detection Function" in 3GPP, vendors providing Deep Packet Inspection (DPI) and Policy and Charging Rules Function (PCRF) functionalities can help ISPs with traffic analysis to enable such app-based data plans. For example, Mobistar in Belgium offers plans of up to 1GB/month with "zero-rated" (free) access to Facebook, Twitter, Netlog (a local social network) and its own mobistar.be domain [Bubley 2011b]. In May 2011, Allot Communications announced at the LTE World Summit that it is deploying its ChargeSmart solution with a multinational mobile operator, covering about 30 million users, that will enable the creation of more personalized, content-based pricing plans [Allot Communications 2011]. Similarly, Japan's Softbank Mobile chose Ericsson to provide it with technology that "*enables truly flexible charging options and paves the way for the operator to develop*

[3]Another variation on this trend is Gemalto's "Facebook for SIM," which allows users to access Facebook from any device, even without a data plan or app download, using SIM-embedded software (rather than software installed on the phone's OS) and uses SMS to exchange updates.

differentiated service offerings based on individual subscriber needs and the service or application used" [Ericsson 2010]. Additionally, "toll-free app" pricing plans in which application providers (or advertisers) either subsidize the user's bandwidth costs or include access costs in the end-user subscription fees have received some attention [Sachson 2011].

In spite of the growing interest in these experimental data plan trials, the feasibility of such application-based pricing plans is still doubtful for several main reasons [Bubley 2011a]. First, mobile operators cannot fine-tune their pricing plans and policies at the short timescale of mobile application upgrades, thus making app-specific data plans difficult to implement. Applications update their features, functionalities, and even protocols very frequently, sometimes on the order of months or even weeks, and hence different mobile customers will often have different versions of the app, which may impact the network in very different ways. Without substantial help from the major content providers, ISPs will find it technically difficult to know and take preemptive measures against the resulting network impact of new application releases and upgrades.

Second, users often do not know which services offered by a mobile application are to be viewed as "inside" the app and which count towards the bandwidth quota. Moreover, it is often difficult for users to know whether the links, photos, and videos shared by friends on a social network page are linked to on-site content within the app or to outside web services. These data plan subtleties form another major hindrance towards adoption of app-based pricing.

Third, personalization of mobile applications and the rise of HTML5 will result in each user having a different amount and type of ad content for the same app, thus making it difficult for operators to create a uniform app-based data plan. Even with DPI capabilities, it is technically challenging for ISPs to account for the inherent differences in the app content and features of different individuals, which, moreover, change on a daily or hourly basis.

Fourth, if applications interact with each other and provide data transfer services under mutual agreements, it would be harder to account for and implement any zero-rated offerings and other forms of app-based data plans. Lastly, there are network neutrality related policy issues that must to be considered for such differential pricing of sponsored content. Given these technical, social, and regulatory challenges, application and content-based pricing, although interesting, is still in its nascent stage.

4. DYNAMIC PRICING

4.1. Dynamic Priority Pricing

Gupta et al. [1997] present a dynamic priority pricing mechanism in which the prices serve as a congestion toll for network access. The authors model user service requests as a stochastic process and network nodes as priority queues. A user's incoming request has an instantaneous value for the service and a linear rate of decay for this value to capture the delay cost. Users' requests can be fulfilled using different alternatives, each with its corresponding price and waiting time. The user trades off between the total cost of service and the cost of delay to choose an optimal alternative in a particular priority class. The benefits of such a scheme is that it achieves resource allocation in real time in a completely decentralized manner and users' myopic decisions lead to socially optimal choices.

Dynamic priority pricing is based on general equilibrium theory, but since computing Arrow-Debreu equilibria in the volatile environment of the Internet is expensive, the authors introduce the notion of a stochastic equilibrium and derive optimal prices that support the unique stochastic equilibrium. They also develop a decentralized real-time mechanism to compute near-optimal prices for the stochastic equilibrium: the cost of service (i.e., the price) is updated after each time period by taking a weighted average of the price in the previous period and the new optimal price, calculated from updated estimates of the waiting times for different priority classes. The authors also demonstrated the convergence prop-

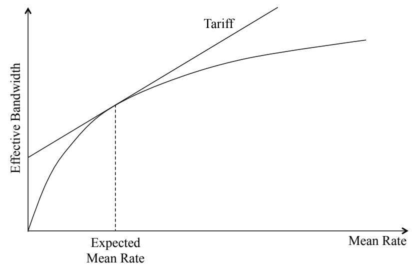

Fig. 5. Illustration of the tariff charged in effective bandwidth pricing [Kelly 1994].

erties of this dynamic pricing mechanism with simulations that allow the system to adapt to changing loads on the network. However, implementing dynamic priority pricing will require modifications to network topology management applications, databases, accounting systems, and the end-user interface.

4.2. Proportional Fairness Pricing

Kelly et al. [1998] proposed proportional fairness pricing as a means to allocate resources (which determine user rates) in proportion to the user's willingness to pay. The global optimization of maximizing net utility across all users, given resource capacity constraints, can be decomposed into a user and a network optimization problem. Kelly shows that there exists a price vector and a rate vector that optimize both the user and the network's optimization problems. Alternatively, if each user chooses a price per unit time according to his or her willingness to pay, and if the network allocates rates per unit price that are proportionally fair, then a system optimum is achieved when users' choices of prices and the network's choice of rate allocation are in equilibrium. Courcoubetis et al. [1998] extend this idea by replacing end-users with intelligent agents that can decide the willingness to pay on behalf of the user while maximizing the user's utility. However, this element introduces overhead in installing such agents on the users' devices or machines and adding network servers to compute the optimal rate allocation vectors at a short timescale.

4.3. Effective Bandwidth Pricing

Kelly's effective bandwidth pricing [1994] is a variant on usage-based pricing in which users are charged based on self-reported peak and mean traffic rates as well as the observed mean rate and duration of each connection. Before a user's connection is accepted, the user is required to provide mean and peak rates for the connection. Given a formula describing effective bandwidth as a function of the peak and mean rates, the user is charged a tariff given by the tangent line to this effective bandwidth formula (as a function of the mean rate) at the self-reported mean and peak rates. Evaluating this tariff at the observed mean rate, the result is multiplied by the connection duration to give the total charge to the user. Figure 5 shows a schematic illustration of such a tariff.

Kelly shows that under this pricing scheme, users minimize their expected cost by accurately reporting the connection's mean and peak rates. Thus, the final charge to the user consists of a term proportional to the connection duration and another term proportional to the connection volume. Users may renegotiate the tariff for a flat fee, e.g., for highly variable traffic.

This pricing scheme can also be extended to connection acceptance control–a connection is accepted if the network's effective load, as calculated from the tariffs charged to existing connections, falls below a certain threshold value. While effective bandwidth pricing is compatible with user incentives and fairly simple, it does require an explicit effective bandwidth formula, and it requires users to know, or at least estimate, the peak and mean rates of each connection. Moreover, further validation is needed to understand whether the benefits of this pricing scheme justify the accounting overhead associated with charging each connection based on its duration and volume.

4.4. Responsive Pricing

MacKie-Mason et al. [1997] describe the concept for responsive pricing in the following words:

> *By associating a cost measure with network loading, all users can be signaled with the prices necessary to recover the cost of the current network load. Price-sensitive users–those willing and able to respond to dynamic prices–increase economic efficiency by choosing whether or not to input traffic according to their individual willingness to pay the current price.*

In other words, a user's price sensitivity and time sensitivity for different applications can be exploited by networks to dynamically set prices to alleviate congestion. This process broadly encompasses the philosophy behind different forms of dynamic time-dependent pricing. In the case of the Internet, MacKie-Mason [1997] argue that such a responsive pricing system is required for network efficiency.

The network can set responses prices either in a closed-loop feedback [Murphy et al. 1994; Murphy and Murphy 1994] or a "Smart Market" approach [MacKie-Mason and Varian 1995]. In a closed-loop setting, the network state, measured in terms of the buffer occupancy at the gateway, is converted to a price per packet for users' adaptive applications, which then decide how much data to transmit. This closed-loop feedback is thus similar to Gupta et al.'s priority pricing [1997], discussed in Section 4.1: the basic idea is that dynamic prices are set so as to incentivize users to behave in a way that enhance network efficiency. However, due to the adaptive nature of the prices, there is a slight delay in the feedback loop: prices are set based on network conditions in the previous time period.

This delay in the feedback loop is tightened in the Smart Market approach proposed by MacKie-Mason and Varian [1995]. In this approach, each user places a "bid" on the packets that reflects their willingness to pay to send the packet onto the network at a given time. The gateway admits packets in the descending order of their bids as long as the network performance remains above a desired threshold. Users are charged according to the minimum bid on a packet admitted into the network at the time, and thus pay only for the congestion cost at the market clearing price. While such auction schemes encourage network and economic efficiency, they require substantial changes in users' application interfaces and providers' billing systems, with additional concerns in the case of billing contention, etc. Moreover, as discussed by Sarkar [1997], such a pricing scheme would require extensive coordination among different nodes of the network, and is open to abuse by those controlling bottleneck facilities. Sarkar argues that these concerns may be overcome with appropriate governmental regulation.

4.5. Dynamic Congestion Pricing

Dynamic congestion pricing is a particular realization of the idea of responsive pricing, in which the network announces prices based on current congestion levels and the user response to these prices is fed back into the control loop to compute new prices. Ganesh et al. [2001] use congestion prices as a mechanism to provide both feedback and incentives to end-systems for rate adaptation in a decentralized manner, and study the resulting dynamics. Paschalidis and Tsitsikilis [1998] address the issue of revenue and welfare maximization for dynamic congestion pricing of customer calls by using a dynamic programming formulation. In their model, users initiate calls that differ in their service class, resource requirements, and call duration. Based on the current congestion level, the service provider charges a fee per call, which in turn influences users' demand. Their findings additionally corroborate the usefulness of time-of-day pricing in reducing network congestion problems.

Such pricing innovations have been adopted in many markets outside of the U.S. In recent years, network operators in highly competitive and lucrative markets, such as those in India and Africa, have adopted innovative congestion-dependent dynamic pricing for *voice calls* [The Economist 2009], although not yet for *mobile data* plans. The African operator MTN pioneered "dynamic tariffing," a congestion-based pricing plan in which the cost of a call is adjusted every hour, in each network cell, depending on the level of usage. Using this pricing scheme, instead of a large peak demand around 8 am, MTN Uganda found that many of customers' were waiting to take advantage of cheaper call rates, thus creating an additional peak at 1 am [The Economist 2009]. A similar congestion pricing for voice calls called "location-based tariff" was launched in India by Uninor. It offers discounts to a customer's calls based on the network traffic condition in the call's originating location; these discounts are visible to customers on their handsets [Sen 2011]. Tango Telecom for Airtel Africa also offers real-time charging and dynamic pricing solutions to mobile operators in India for voice calls based on factors like cell load, the time of day, location, and network traffic patterns.

In the future, similar trends for dynamic pricing for data can also be expected to emerge. A recent pilot trial in the U.S. [TUBE 2012] has already demonstrated the feasibility, benefits, and architecture of such a system.

4.6. Game-Theoretic Pricing

Several authors have used game-theoretic models for pricing data, some of which are briefly discussed here. Hayer [1993] proposed transport auctions as a way to distribute excess capacity across users with delay tolerant traffic. A decentralized auction-based approach to pricing of edge-allocated bandwidth, called "market pricing," was explored by Semret et al. [2000] in a differentiated services Internet.

Yaïche et al. [2000] introduced a cooperative game-theory framework that used Nash bargaining to compute bandwidth allocation for elastic services and pricing in broadband networks. This framework provides rate allocations for users that are not only Pareto-optimal from the viewpoint of the whole system, but also consistent with fairness axioms of game theory.

Other researchers have used game-theoretic formulations to investigate the effect of pricing on user adoption and fair use. Pricing models to induce participation and collaboration in a public wireless mesh network were studied by Lam et al. [2007]. Shen and Basar [2007] investigated optimal nonlinear pricing policy design as a means of controlling network usage and generating profits for a monopolistic service provider. Dynamic game models have also been used to determine WiFi access point pricing by Musacchio and Walrand [2006], which is relevant in the context of congestion management through WiFi offloading.

Jiang et al. [2008] introduce a model to study the role of time preferences in network pricing. In their model, each user chooses his/her access time based on his/her preference, the congestion level, and the price charged. The authors show that maximization of both the social welfare and the revenue of a service provider is feasible if the provider can differentiate its prices over different users and times. However, if the prices can only be differentiated over the access times and not across users due to insufficient information, the resulting social welfare can be much less than the optimum, especially in the presence of many low-utility users.

Despite these theoretical works, game-theoretic models have found little traction among real network operators so far, perhaps due to the stylized nature of the theoretical models and the challenges in estimating user utility and system parameters in the real world.

The different real-world examples of various static and dynamic pricing plans for the wired and wireless services discussed in this paper are summarized in Table I.

Table I. List of key example pricing plans discussed.

Pricing Practice		Example Pricing Plan (see the paper for details)		
Type	Category	Description	Network	Country
Static	Fixed Flat-Rate	Monthly fee; unlimited	Both	Vanishing worldwide
		Monthly; flat to a cap, then usage-based	Wired/wireless	U.S. (AT&T, Verizon)
		Monthly; flat to a cap, then throttle	Wireless	Spain (Orange)
		Monthly; flat to a cap; then disconnect	Wired	U.S. (Comcast)
		Monthly; shared	Wired	Canada (Rogers)
		Hourly rate	Wireless	Egypt (Mobinil)
	Usage-Based	Cap then metered	Wired/wireless	Worldwide (e.g. U.S., U.K.)
	Priority Pricing	Priority pass (for dongle users)	Wireless	Singapore (SingTel)
	Time of Day	Day-time & Night-time rates	Wireless	India (BSNL)
		Users choose happy hours	Wireless	U.K. (Orange)
	Cumulus Pricing	Usage-based contract negotiation	Wireless	U.K. (Vodafone)
	App-Based Pricing	Free access to select apps; bundling	Wireless	U.K. (Orange) Denmark (TDC)
Dynamic	Congestion-Based	Hourly price changes	Wireless (voice calls)	Uganda (MTN)
		Location and cell-load based	Wireless (voice calls)	India (Uninor)
		Time- and usage-based	Wireless (data)	U.S. (pilot trial)

5. CHALLENGES IN PRICING INNOVATION

As discussed earlier, pricing based on monthly bandwidth usage leaves a timescale mismatch: ISP revenue is based on monthly usage, but peak-hour congestion dominates its cost structure. Static usage-based pricing schemes use overage penalties to limit network congestion by reducing demand from individual users, but they cannot prevent the peak demands across users from concentrating at the same time. Consequently, simple usage-based models do not mitigate the ISP's congestion problem; there needs to be a time-dependent component of pricing to avoid crowding of users at the same time.

The previously explored, simple peak/off-peak (two period) time-dependent pricing schemes are also inadequate because they can incentivize only the highly price-sensitive users to shift some of their non-critical traffic to off-peak periods. Hence such pricing plans often end up creating two peaks - one during the daytime (from time-sensitive transactions) and one at night (from elastic sessions that wait for large discounts), instead of leveling out the demand profile [The Economist 2009]. In general, all static pricing schemes suffer from their inability to adapt prices in real time to respond to the usage patterns, and thus fail to exploit most users' varying degrees of delay tolerance.

Dynamic pricing is better equipped to overcome these issues and does not need to pre-classify hours into peak and off-peak periods. But many of the dynamic congestion-dependent pricing schemes are myopic and reactive to network conditions; that is, the prices fluctuate depending on current cell load, thus inconveniencing users. Moreover, such pricing schemes have been explored mainly for mobile voice traffic, which is typically different from mobile data traffic in its delay sensitivity, duration, and real-time interactivity needs. For example, unlike voice calls, much broadband data traffic could be automatically scheduled to low-congestion periods from the client device without requiring any user involvement, e.g., downloading queued movies late at night or synchronization of cloud-based apps. Such features of data traffic provide new opportunities for dynamic pricing. But dynamic pricing for data, particularly for mobile data, creates challenges that need to be accounted for in designing future policy and pricing initiatives. Some of these important challenges are discussed next.

5.1. Social Considerations

5.1.1. User Convenience. Dynamic real-time pricing plans for data can be tricky to realize in practice. On the one hand, real-time price adjustments (e.g. every half-hour) can lead to much inconvenience for users as well as uncertainty about their monthly bill. On the other hand, fully-automated user responses, as in the case of a "Smart Market" with "bids" on resource reservation, require better prediction of users' reservation intervals, new software installation in the client devices, improved QoS solutions, and upgrades in the ISP's billing infrastructure, with additional issues such as billing contention management. Thus, dynamic pricing at very short timescales can be inconvenient for both ISPs and their customers.

A potential direction going forward would be to borrow ideas from dynamic pricing realizations that have worked in other markets, such as day-ahead pricing in electricity networks. With day-ahead pricing, the future prices for all 24 hours are available to the users so that they are able to plan for their next day's usage and set a budget that they are willing to spend each day, thus providing an easier way of dealing with such uncertainties.

5.1.2. Deployment Precedence. Congestion-based pricing has been practiced in various forms, including dynamic ones, in several utility markets, such as electricity and road networks (discussed in greater details in the Appendix). Electricity markets in the U.S. have implemented both real-time and day-ahead dynamic pricing [U.S. Energy Information Administration 2012], which with the proliferation of smart meters are proving to be successful business models. Similarly, in the case of road networks, congestion pricing has been accepted by the public and legislators in cities like Oslo and Rotterdam [Harsman 2001], where residents agree with the need for tolls to decrease peak traffic in crowded pathways. Moreover, the ability to collect tolls with electronic road pricing has enabled even dynamic, congestion-dependent tolls. The successful adoption of dynamic pricing in these markets indicates that similar pricing ideas may be viable for data networks.

5.1.3. Ethical Issues. Today's flat-rate pricing essentially forces subsidization of heavy users by the median and low bandwidth users. Dynamic pricing thus enhances fairness by eliminating this subsidization, while also reducing the ISPs' costs and enhancing economic efficiency for users. However, many suffer from the misconception that dynamic pricing schemes will force users to avoid accessing the Internet at all peak times and may even lead to disconnecting low-income customers.

In reality, under dynamic pricing, users only need to reduce their mobile video usage by some discretionary amount at peak times, without compromising or threatening their basic connectivity. Consumer advocates have argued that such pricing schemes are unfair to low-income groups, as they are unable to reduce their usage in peak periods because they have very little usage to begin with. However, this notion is arguably flawed: dynamic pricing actually creates the option for all consumers to choose how much they are willing to spend by deciding when and how much bandwidth they want to consume. This view is shared by economists, who describe dynamic pricing for the Internet as a natural extension of existing network control feedback mechanisms that bring "*users back into the loop and thereby ensure that performance measures are user-oriented.*" [MacKie-Mason et al. 1997]. Recent studies in electricity markets have further strengthened this view by confirming that 80%-90% of low-income households actually stand to gain from dynamic pricing of electricity [Faruqui 2010].

5.1.4. Policy Initiatives. While policymakers in the U.S. and several other countries have already approved congestion pricing schemes for electricity and road tolls, the regulatory environment for the Internet has been more stringent so far, partly due to net-neutrality concerns.

Academics in law schools have already cautioned that the ongoing debate on network neutrality in the U.S. often overlooks the need for allowing service providers to have flexibility in exploring different pricing regimes [Yoo 2009]:

> *Restricting network providers' ability to experiment with different protocols may also reduce innovation by foreclosing applications and content that depend on a different network architecture and by dampening the price signals needed to stimulate investment in new applications and content.*

However, since 2010, there has been a monumental shift in this paradigm. This sentiment is highlighted in FCC Chairman J. Genachowski's 1 December 2010 statement, [Schatz and Ante 2010] which recognizes *"the importance of business innovation to promote network investment and efficient use of networks, including measures to match price to cost."* Genachowski states:

> *Broadband providers need meaningful flexibility to manage their networks... to address the effects of congestion.*

5.2. Economic Considerations

5.2.1. Optimal Pricing. One of dynamic pricing's challenges from an economic modeling perspective is to determine the optimal way for ISPs to set prices that they are ready to offer and their consumers are willing to accept. In doing so, the ISP tries to minimize its cost of overshooting capacity on the bottleneck link and the cost of offering discounts over a baseline usage charge at different times of the day. The solution also needs the ISP to estimate the deferral behavior of their customers by monitoring their responses to the offered prices and readjusting these estimates over time. A mathematical formulation for calculating day-ahead,[4] time-dependent usage-based prices in a scalable and efficient manner using a convex optimization framework was recently proposed in [Joe-Wong et al. 2011]. Other theoretical models for reducing congestion and achieving efficiency from an unselfish social planner's perspective have been studied as well [Jiang et al. 2008].

5.2.2. User Behavior Estimation. Offering time-varying prices requires ISPs to estimate their users' delay tolerance for different traffic classes and their deferral patterns in response to offered prices [Joe-Wong et al. 2011]. The key challenges in doing so are choosing the right

[4] "Day-ahead" pricing means that the prices for the next day are announced in advance to the users.

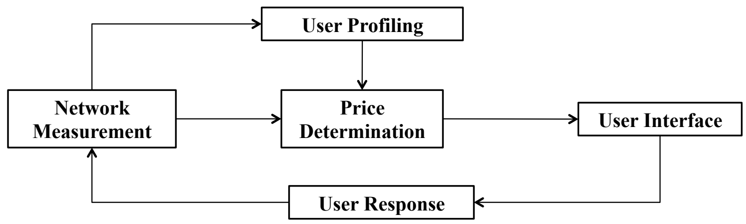

Fig. 6. Schematic of a feedback-control loop between users and ISPs.

form for utility functions, finding a suitable model that allows computational tractability, and identifying methods to profile usage behavior and estimation of user's price-delay tradeoffs. Moreover, for scalability reasons, the ISP may need to perform these estimations without monitoring each individual household's demand pattern.

5.3. Technological Considerations

5.3.1. Feedback-Control Functionality Separation. Implementing dynamic pricing for data requires creating a feedback-control loop between the ISP server, which computes the prices to offer to users, and the users who respond to these prices [MacKie-Mason et al. 1997]. The ISP-side functionality will typically consist of near real-time traffic monitoring, estimation of user delay tolerances for various traffic classes, and computing the prices to offer based on aggregate traffic measurements. On the user side, functionalities like the ability to view future prices and usage history can be provided by installing applications on users' mobile devices. Figure 6 provides a schematic showing the components of this loop.

5.3.2. System Scalability. Maintaining scalability of a dynamic pricing system requires developing models and algorithms that allow ISPs to compute future prices to offer based on aggregate congestion levels, rather than monitoring each individual customer's usage and response. As ISPs maintain large networks of many users, any dynamic pricing system must scale well to multiple users.

5.3.3. Privacy and Security. The design of dynamic pricing systems should account for user privacy, particularly in terms of the feedback that ISPs record from their users and the security of the data. A potential way to avoid privacy issues is to minimize the need to monitor individual users' usage patterns – a feature that also contributes to scalability.

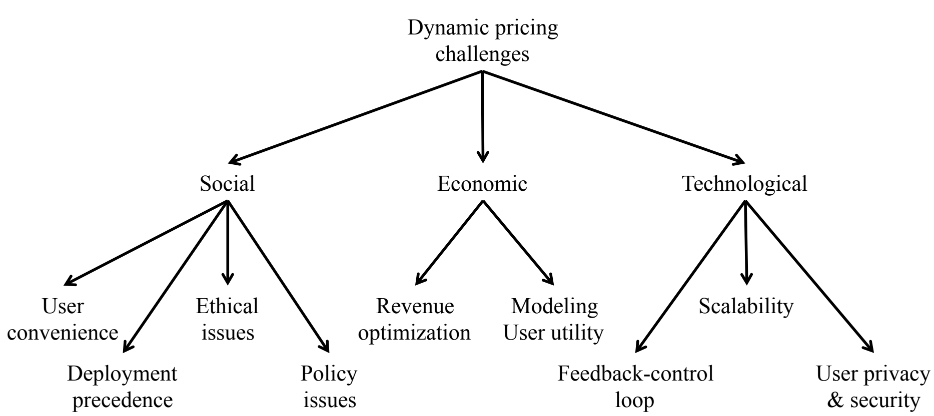

Fig. 7. Research challenge areas to overcome for dynamic pricing of data.

To realize the full potential of dynamic pricing for broadband data, the future research agenda should account for and overcome these various challenges, as summarized in Figure 7.

6. NEW DIRECTIONS

The recent exponential growth in data demand has catalyzed major changes within broadband pricing research and practice. In this section, we discuss the new directions that are emerging from these recent changes.

6.1. Satellite Broadband

This survey focuses mainly on wired and wireless pricing, as these are the mediums most impacted by users' increasing demands for data. Recently, however, many companies have begun to offer satellite broadband as an alternative to wired or wireless Internet access. While satellite broadband solutions have existed for over twenty years, satellite Internet access has only recently become a popular offering for end-users. Satellite is becoming especially prevalent as a way to reach users in rural and sparsely populated regions, such as in Africa, where installing more traditional wired or wireless infrastructures is not cost-effective [Svitak 2012].

Since satellite broadband has not yet experienced serious congestion problems, pricing for satellite broadband remains fairly simple, though the growth in demand for high-bandwidth services may open up possibilities for new pricing schemes in the future. Technological capabilities that allow for fine granular capacity provisioning and dynamic channel allocation will also facilitate the adoption of more innovative pricing practices. Sun and Modiano [2006] propose a pricing-based solution to the problem of channel allocation in satellite networks. They use a slotted Aloha-based model to characterize the Pareto-efficient throughput region in single- and multiple-satellite networks, and prove that under certain conditions, there exists a unique set of user prices associated with a competitive equilibrium in this Pareto-efficient region.

Today, satellite broadband consumers mostly pay in proportion to the uplink and downlink data rates by subscribing to a chosen tier of available data plans. Each tier has specified maximum upload and download speeds and a monthly cap on bandwidth. Beyond this cap, a user's speed is either throttled down or charged by overage, depending on the plan (see Figure 8 for a schematic of these factors). For example, Viasat provides satellite broadband services to the U.S. with a service called Exede, which gives 12 Mbps downlink and 3 Mbps uplink speeds. Three different consumer plans are offered; they offer the same speeds, but come with either 7.5GB, 15GB, or 25GB data caps, and cost $49.99, $79.99, and $129.99 per month, respectively [D'Orazio 2012]. Viasat's child company, WildBlue, also offers similar services with lower downlink speeds of 512 Kbps - 1.5 Mbps. Users can upload or download a "threshold" amount of data in a 30-day period before their speeds are reduced.

Similarly, in the U.K., the Hylas 1 satellite by Avanti PLC combines high-speed satellite broadband with U.K.-based ground stations. Services of up to 8 Mbps downlink and 2 Mbps uplink speeds are provided for consumers and businesses. Tiered monthly plans in the range of £33-£66 offer data allowances of 6-15 GB and maximum uplink and downlink speeds of 0.5-2 Mbps and 2-10 Mbps, respectively. If a user exceeds the allowed data cap, (s)he can either pay for excess usage at £6/GB, or the provider can set the system to throttle down to a low speed [Rural Broadband 2012].

Some satellite plans have daily caps in addition to monthly data caps. For example, Vox Telecom of South Africa, which launched its YahClick satellite broadband service in January 2012, offers monthly tiered plans of R 261 for a 512 Kbps downlink speed, 3 GB monthly limit and 100 MB/day data cap; R 747 for a 2 Mbps downlink speed, 7.5 GB monthly limit and 250 MB/day data cap; and R 1188 for a 5 Mbps downlink speed, 15 GB monthly limit and 500MB/day data cap. On the other hand, Vox's competitor Liquid Telecom provides

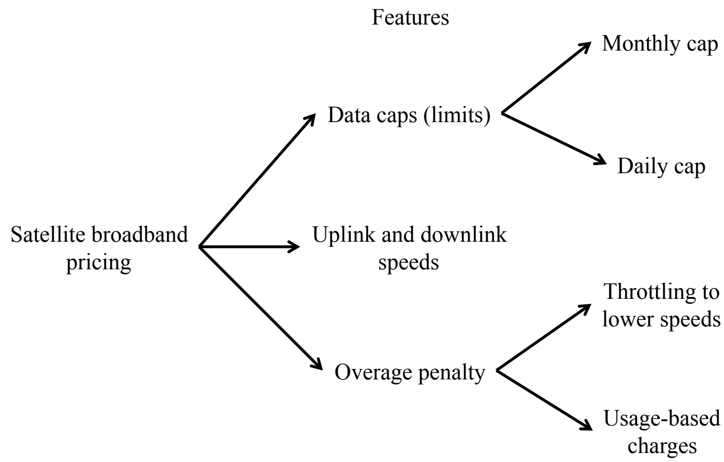

Fig. 8. Satellite plan features and congestion control mechanisms.

monthly tiered plans offering unlimited data at 512 Kbps for R 798, or 2 Mbps with a 1 GB or 10 GB monthly limit and no daily data cap for R 1265 and R 3340 respectively [Muller 2012].

IPNet JSC, a provider and systems integrator of the Hughes Network Systems (HNS) solutions in Russia, also provides various tiered pricing plans with different uplink and downlink speeds and data caps. Upon exceeding the cap, users are charged based directly on excess usage [IPNet JSC 2012].

6.2. New Pricing Architectures

The U.S. is witnessing a transition from flat-rate fees towards usage-based pricing by the major carriers, AT&T and Verizon [Frakes 2010; Segall 2011]. This move has generated much debate among policy makers, consumer advocates, and the industry. But academics have often argued that architectural issues in Internet pricing are more important than the form or basis of the actual prices charged: for instance, Shenker et al. [1996] argue that the flat- versus usage-based pricing debate oversimplifies the issue, as in reality there is a continuum between the two. The authors make the case that research into Internet pricing has overly focused on optimality, and in particular on matching prices to the marginal congestion costs. Instead, they advocate an architectural focus, i.e., designing the network architecture to facilitate various pricing plans, such as allowing receivers rather than senders to be charged for usage.

Architecture and systems enabling such pricing innovations for broadband data are arguably becoming an active area of research, and are central to shaping future research agenda. Understanding these interactions between economics and technology in order to create the overall design of future networks has been identified as a priority in the NSF's Future Internet Design initiative [Fisher 2007].

6.3. New Pricing Plans

A new pricing option expected to become more popular in the coming years is that of a "shared data plan," which allows users to share their data cap across multiple devices, at a premium for each additional device [Buckley 2012; Galbraith 2011]. Rogers in Canada has been offering such plans on a promotional basis since 2011 [Electronista 2011]. Orange of France Telecom has also started such a plan in Austria that allows iPad owners to share one allotment of data with a phone, and 38% of iPad owners on its network now subscribe to

Table II. Summary of some emerging trends.

Trend	Description	Reference
Flat-rate & usage-	Coexistence of both types of plans	
based pricing	Emergence of shared data plans	[Buckley 2012; Fried 2011]
	Focus on architectural issues in	[Shenker et al. 1996]
	Internet pricing is needed	
Time-dependent pricing	Traffic shaping with off-peak usage incentives	[El-Sayed et al. 2011]
	Static discounts (two-period ToD)	
Dynamic pricing	Practiced for voice calls (e.g., location-based)	[The Economist 2009]
	Adopted in Africa & India	
	Needs new theory for dynamic data pricing	[Joe-Wong et al. 2011]
	New architecture & systems implementation	[Ha et al. 2012]
Differentiated,	Tiered data plans with various QoS	[Ericsson 2011]
app-based pricing	App-specific “zero-rating,” app bundles	[Sachson 2011]

this plan [Fried 2011]. Sharing a data cap across multiple devices may induce more efficient usage of the quota per consumer.

In dealing with the growing problem of bandwidth demand, researchers from Alcatel-Lucent have recently proposed static time-dependent discounts for incentivizing off-peak capacity usage [El-Sayed et al. 2011]. They also present policy and quality of service (QoS) mechanisms to implement traffic shaping with off-peak usage incentives to users.

Another interesting trend emerging in several parts of the world, particularly in Asia and Africa, is that of a larger scale adoption of dynamic pricing for voice calls [The Economist 2009]. As the demand for data, and in particular mobile broadband, grows, there will be a need for new architectures and systems that meet the dynamic pricing challenges described in Section 5. Some initial steps in this direction are reported in [Joe-Wong et al. 2011; Ha et al. 2012], which report on initial U.S. pilot-trial results of their proposed dynamic time-dependent usage-based pricing system, called TUBE [TUBE 2012]. This approach is similar in essence to the objectives of the Berkeley INDEX project [Varaiya et al. 1996], but in the context of understanding consumers’ responses to dynamic pricing of mobile data.

Differentiated pricing for broadband [Ericsson 2011], e.g., application-based pricing of data, is also a growing trend. As reported in Section 3.10, ISPs are exploring new ways to attract customers by “zero-rating” or bundling certain applications. In the future, we will likely see more extensions of similar practices for managing high-bandwidth and delay-sensitive applications.

A summary of these potential future trends is given in Table II.

7. CONCLUSIONS

In this work, we draw attention to the growing problem of network congestion and highlight some of the recent steps taken by ISPs to mitigate its effects. The projected growth in demand for data, especially from mobile data and video traffic, is far more than what can be supported even by the latest technological advances, e.g., 4G/LTE and WiFi offloading, due to expensive backhauling and increasing wired network congestion. Consequently, ISPs have been aggressively using pricing as a congestion control tool.

The basic idea of congestion pricing has been known in the networking community for several decades, but only now have conditions demanded that it be put into practice. We review many known data network pricing proposals, both static and dynamic, and discuss the extent to which some of these have been adopted by ISPs around the world. We also discuss existing pricing practices in electricity markets and road networks that have influenced many of the ideas of broadband pricing.

In addition to reviewing past plans and present deployments, we have identified various shortcomings to ISP's current approaches, as well as several opportunities arising from mobile data's very characteristic usage patterns, which need to be exploited further. To this end, we discuss the predominant trends in pricing data and identify challenges that must be met, particularly in realizing more innovative dynamic pricing solutions such as day-ahead time- and usage-dependent pricing, app-based pricing, and others. The material presented in this paper seeks to inform researchers and operators keen on understanding ongoing worldwide developments in pricing plans, and will help to shape a new research agenda in network economics.

8. ACKNOWLEDGEMENTS

This work has benefited immensely from our discussions with several colleagues and collaborators. We would like to specially thank Victor Glass, Prashanth Hande, Raj Savoor, Steve Sposato, Rob Calderbank, Danny Tsang, Sundeep Rangan, Yuan Wu, Junshan Zhang, T. Russell Hsing, Keith Cambron, Andrew Odlyzko, Roch Guérin, Augustin Chaintreau and Jennifer Rexford for their valuable comments, viewpoints, and information that have helped to shape this paper. Additionally, we acknowledge the support of our industrial collaborators, in particular, Reliance Communications of India, MTA of Alaska, SES, AT&T, Qualcomm, Comcast, and NECA (the National Exchange Carrier Association).

REFERENCES

86CallChina. 2011. China GPRS/EDGE data plans. http://www.86callchina.com/gprs-cdma-data-cards.htm.

Allot Communications. 2011. Allot's intelligent charging solution selected by multinational mobile operator to deploy new pricing. Press Release. May 17, http://tinyurl.com/SmartCharging.

Anania, L. and Solomon, R. J. 1997. Flat – the minimalist price. In *Internet Economics*, L. W. McKnight and J. P. Bailey, Eds. The MIT Press, Cambridge, MA, 91–118.

BBC. 2011. Africa's mobile phone industry 'booming'. November, http://www.bbc.co.uk/news/world-africa-15659983.

Borenstein, S. 2005. The long-run efficiency of real-time electricity pricing. *The Energy Journal 26,* 3, 93–116.

Borenstein, S., Jaske, M., and Rosenfeld, A. 2002. Dynamic pricing, advanced metering, and demand response in electricity markets. Tech. rep., Center for the Study of Energy Markets. Working Paper.

Breker, L. P. 1996. A survey of network pricing schemes. In *Proceedings of the 8th Symposium on Computer Science*. University of Saskatchewan.

Brownlee, N. 1997. Internet pricing in practice. In *Internet Economics*, L. W. McKnight and J. P. Bailey, Eds. The MIT Press, Cambridge, MA, 77–90.

Brunekreeft, G. 2000. Price capping and peak-load pricing in network industries. *Discussion Papers 73*, 1–9. University of Freiburg, Institute for Transport Economics and Regional Policy.

Bubley, D. 2011a. Another reason why app/service based pricing for mobile broadband will fail. Disruptive Wireless. September 7, http://tinyurl.com/AppPricing.

Bubley, D. 2011b. Belgian MNO tries app-specific zero rating - bad idea in my view. Disruptive Wireless. November 1, http://tinyurl.com/6urunl3.

Buckley, S. 2012. Verizon shared data plans show up in employee training materials, still on track? Engadget. January 30.

Carew, S. 2012. Users complain, AT&T blames data tsunami. Reuters Media File. February 14.

Caron, S. and Kesidis, G. 2010. Incentive-based energy consumption scheduling algorithms for the smart grid. In *First IEEE International Conference on Smart Grid Communications*. IEEE, 391–396.

Castells, M., Qiu, M., and Linchuan, J. 2007. *Mobile Communication and Society: a Global Perspective : a Project of the Annenberg Research Network on International Communication.* The information revolution and global politics. MIT Press, Cambridge, MA.

CHAO, H. 1983. Peak load pricing and capacity planning with demand and supply uncertainty. *The Bell Journal of Economics 14,* 1, 179–190.

CHARLES RIVER ASSOCIATES. 2005. Impact evaluation of the California statewide pricing pilot. Tech. rep., Charles River Associates.

CIA. 2011. Country comparison: Population. CIA World Factbook. https://www.cia.gov/library/publications/ the-world-factbook/rankorder/2119rank.html.

CISCO SYSTEMS. 2011. Cisco visual networking index: Forecast and methodology, 2010-2015'. June 1, http://tinyurl.com/VNI2011.

CLARK, D. D. 1997. Internet cost allocation and pricing. In *Internet Economics*, L. W. McKnight and J. P. Bailey, Eds. The MIT Press, Cambridge, MA, 215–252.

COCCHI, R., ESTRIN, D., SHENKER, S., AND ZHANG, L. 1991. A study of priority pricing in multiple service class networks. *ACM SIGCOMM Computer Communication Review 21,* 4, 123–130.

COCCHI, R., SHENKER, S., ESTRIN, D., AND ZHANG, L. 1993. Pricing in computer networks: Motivation, formulation, and example. *IEEE/ACM Transactions on Networking 1*, 614–627.

COMCAST. 2010. About excessive use of data. October, http://customer.comcast.com/help-and-support/ internet/common-questions-excessive-use/.

COMMUNICATIONS COMMISSION OF KENYA. 2011. Quarterly sector statistics report. http://www.cck. go.ke/resc/downloads/SECTOR_STATISTICS_REPORT_Q1_11-12.pdf.

COURCOUBETIS, C., STAMOULIS, G. D., MANOLAKIS, C., AND KELLY, F. P. 1998. An intelligent agent for optimizing QoS-for-money in priced ABR connections. Tech. rep., Institute of Computer Science–Foundation for Research and Technology Hellas and Statistical Laboratory, University of Cambridge. Preprint.

DELGROSSI, L. AND FERRARI, D. 1999. Charging schemes for reservation-based networks. *Telecommunication Systems 11,* 1, 127–137.

DONOVAN, K. AND DONNER, J. 2010. A note on the availability (and importance) of pre-paid mobile data in Africa. In *2nd International Conference on Mobile Communication Technology for Development.* Karlstad University, 263–267.

D'ORAZIO, D. 2012. Viasat prices satellite broadband plans, calls them Exede. TheVerge. January 10, http://www.theverge.com/2012/1/10/2694708/viasat-exede-pricing-availability-satellite-broadband-internet.

DU, P. AND LU, N. 2011. Appliance commitment for household load scheduling. *IEEE Transactions on Smart Grid 2,* 2, 411–419.

EL-SAYED, M., MUKHOPADHYAY, A., URRUTIA-VALDÉS, C., AND ZHAO, Z. J. 2011. Mobile data explosion: Monetizing the opportunity through dynamic policies and QoS pipes. *Bell Labs Technical Journal 16,* 2, 79–100.

ELECTRONISTA. 2011. Rogers intros data sharing for 6GB plans with hefty premium. February 8, http://www.electronista.com/articles/11/02/08/rogers.6gb.data.plan.gets.expensive.sharing/.

EMISR. 2011. National broadband plan. http://www.tra.gov.eg/emisr/Presentations/Plan_En.pdf.

ERICSSON. 2010. Japan's Softbank Mobile chooses Ericsson mobile broadband solution. August 2, Press Release.

ERICSSON. 2011. Differentiated mobile broadband. White Paper, http://www.ericsson.com/res/docs/ whitepapers/differentiated_mobile_broadband.pdf.

EZZIANE, Z. 2005. Charging and pricing challenges for 3G systems. *IEEE Communications Surveys & Tutorials 7,* 4, 58–68.

FALKNER, M., DEVETSIKIOTIS, M., AND LAMBADARIS, I. 2000. An overview of pricing concepts for broadband IP networks. *IEEE Communications Surveys 3*, 2–13.

FARUQUI, A. 2010. The ethics of dynamic pricing. *The Electricity Journal 23,* 6, 13–28.

FARUQUI, A., HLEDIK, R., AND SERGICI, S. 2009. Piloting the smart grid. *The Electricity Journal 22,* 7, 55–69.

FARUQUI, A. AND WOOD, L. 2008. Quantifying the benefits of dynamic pricing in the mass market. Tech. rep., Edison Electric Institute.

FISHER, D. 2007. US National Science Foundation and the Future Internet Design. *ACM SIGCOMM Computer Communication Review 37,* 3, 85–87.

FRAKES, D. 2010. AT&T announces tethering details and new plans for iPhone, iPad. InfoWorld.

FRIED, I. 2011. The Data Plan of the Future is Available Now, at Least in Europe. All Things D. http://tinyurl.com/FT-shared.

GALBRAITH, C. 2011. Shared data plans set to make global impact. Billing World. October 24.

GANESH, A., LAEVENS, K., AND STEINBERG, R. 2001. Congestion pricing and user adaptation. In *Proceedings of IEEE INFOCOM*. Vol. 2. IEEE, 959–965.

GENACHOWSKI, J. 2010. New rules for an open Internet. FCC. December 21, http://www.fcc.gov/blog/new-rules-open-internet.

GIZELIS, C. AND VERGADOS, D. 2011. A survey of pricing schemes in wireless networks. *IEEE Communications Surveys & Tutorials 13,* 1, 126–145.

GOMEZ-IBANEZ, J. A. AND SMALL, K. A. 1994. *Road pricing for congestion management: a survey of international practice*. Transportation Research Board, Washington, DC. National Cooperative Highway Research Program.

GUPTA, A., STAHL, D., AND WHINSTON, A. 1997. Priority pricing of integrated services networks. In *Internet Economics*, L. W. McKnight and J. P. Bailey, Eds. The MIT Press, Cambridge, MA, 323–352.

HA, S., JOE-WONG, C., SEN, S., AND CHIANG, M. 2012. Pricing by timing: Innovating broadband data plans. In *SPIE OPTO Broadband Access Communication Technologies VI Conference*. Paper 8282-12.

HANDE, P., CHIANG, M., CALDERBANK, R., AND ZHANG, J. 2010. Pricing under constraints in access networks: Revenue maximization and congestion management. In *Proceedings of IEEE INFOCOM*. IEEE, 1–9.

HARSMAN, B. 2001. Urban road pricing acceptance. In *IMPRINT-EUROPE Seminar*.

HAU, T. D. 1990. Electronic road pricing: Developments in Hong Kong 1983-1989. *Journal of Transport Economics and Policy 24,* 2, 203–214.

HAUGE, J., CHIANG, E., AND JAMISON, M. 2009. Whose call is it? Targeting universal service programs to low-income households' telecommunications preferences. *Telecommunications Policy 33,* 3-4, 129–145.

HAUSMANN, J., KINNUCAN, M., AND MCFADDDEN, D. 1979. A two-level electricity demand model: Evaluation of the Connecticut time-of-day pricing test. *Journal of Econometrics 10,* 3, 263–289.

HAYEL, Y. AND TUFFIN, B. 2005. A mathematical analysis of the cumulus pricing scheme. *Computer Networks 47*, 907–921.

HAYER, J. 1993. Transportation auction: a new service concept. M.S. thesis, University of Alberta. TR-93-05.

HENDERSHOTT, T. 2006. *Economics and Information Systems, Volume 1*. Elsevier, Amsterdam, Chapter 2, 4–9.

HERTER, K. 2007. Residential implementation of critical-peak pricing of electricity. *Energy Policy 35,* 4, 2121–2130.

HIGGINBOTHAM, S. 2010. Mobile operators want to charge based on time and apps. GigaOm. December 14.

HOLGUIN-VERAS, J., CETIN, M., AND XIA, S. 2006. A comparative analysis of us toll policy. *Transportation Research Part A: Policy and Practice 40,* 10, 852–871.

I.B.M. 2007. Ontario energy board smart price pilot final report. Tech. rep., Ontario Energy Board. I. B. M. Global Business Services and eMeter Strategic Consulting.

IDA SINGAPORE. 2011. Telecommunications. http://www.ida.gov.sg/Publications/20070822130650.aspx.

IPNET JSC. 2012. Price and rates. http://www.zaoipnet.ru/en/price/connectionservice/.

ITU. 2011. Measuring the information society. http://www.itu.int/ITU-D/ict/publications/idi/2011/ Material/MIS_2011_without_annex_5.pdf.

JIANG, L., PAREKH, S., AND WALRAND, J. 2008. Time-dependent network pricing and bandwidth trading. In *IEEE Network Operations and Management Symposium Workshops*. IEEE, 193–200.

JOE-WONG, C., HA, S., AND CHIANG, M. 2011. Time-dependent broadband pricing: Feasibility and benefits. In *31st International Conference on Distributed Computing Systems*. IEEE, 288–298.

JOE-WONG, C., SEN, S., HA, S., AND CHIANG, M. 2012. Optimized day-ahead pricing for the smart grid with device-specific scheduling flexibility. *IEEE Journal on Selected Areas in Communications*. To appear.

JOKSIMOVIC, D., BLIEMER, M. C. J., AND BOVY, P. H. L. 2005. Dynamic road pricing optimization with heterogeneous users. In *Proceedings of 45th Congress of the European Regional Science Association*. European Regional Science Association.

KANG, C. 2010. AT&T wireless scraps flat-rate Internet plan. The Washington Post. June 2.

KAPUTSKA, P. 2010. Prepaid plans heat up: Why no-contract wireless is on the rise. PC World. September 21.

KEALL, C. 2011. TelstraClear calls limitless weekend 'very successful', InternetNZ 'disastrous'. The National Business Review. December 5.

KELLY, F. 1994. On tariffs, policing, and admissions control for multiservice networks. *Operations Research Letters 15*, 1–9.

KELLY, F., MAULLOO, A. K., AND TAN, D. H. K. 1998. Rate control for communication networks: Shadow prices, proportional fairness, and stability. *Journal of the Operational Research Society 49*, 237–252.

Klopfenstein, B. C. 2009. Internet economics: An annotated bibliography. *Journal of Media Economics 11,* 1, 33–48.

Kwang, K. 2011. SingTel unveils priority mobile broadband access. June 13, http://www.zdnetasia.com/singtel-unveils-priority-mobile-broadband-access-62300719.htm.

Lam, K. R., Chiu, D. M., and Lui, J. C. S. 2007. On the access pricing and network scaling issues of wireless mesh networks. *IEEE Transactions on Computing 56*, 140–146.

Lancaster, H. 2011. Italy–telecoms, IP networks, digital media and forecasts. BuddeComm. http://www.budde.com.au/Research/Italy-Telecoms-IP-Networks-Digital-Media-and-Forecasts.html.

Lee, D., Mo, J., Walrand, J., and Park, J. 2011. A token pricing scheme for internet services. *Proc. of Seventh ICQT*.

Lewis, P. H. 1996. An "all you can eat" price is clogging Internet access. New York Times. December 17, p. A1.

Li, N., Chen, L., and Low, S. 2011. Optimal demand response based on utility maximization in power networks. In *IEEE Power and Energy Society General Meeting*. IEEE, 1–8.

Li, S., Huang, J., and Li, S. Y. R. 2009. Revenue maximization for communication networks with usage-based pricing. In *IEEE Global Telecommunications Conference*. IEEE, 1–6.

Littlechild, S. 2003. Wholesale spot price pass-through. *Journal of Regulatory Economics 23,* 1, 61–91.

MacKie-Mason, J. and Varian, H. 1995. Pricing the Internet. In *Public Access to the Internet*, B. Kahin and J. Keller, Eds. Prentice-Hall, Englewood Cliffs, NJ.

MacKie-Mason, J. K., Murphy, L., and Murphy, J. 1997. Responsive pricing in the Internet. In *Internet Economics*, L. W. McKnight and J. P. Bailey, Eds. The MIT Press, Cambridge, MA, 279–303.

Malik, O. 2011. In India, broadband means a 3G connection. GigaOm. July 22, http://gigaom.com/broadband/in-india-broadband-means-a-3g-connection/.

Marbach, P. 2004. Analysis of a static pricing scheme for priority services. *IEEE/ACM Transactions on Networking 12*, 312–325.

Marshall, J. 1996. Economics, not engineering, will unclog the Internet. San Francisco Chronicle. November 4.

Matsukawa, I. 2001. Household response to optional peak-load pricing of electricity. *Journal of Regulatory Economics 20,* 3, 249–267.

May, A. 1992. Road pricing: an international perspective. *Transportation 19,* 4, 313–333.

May, A. and Milne, D. 2000. Effects of alternative road pricing systems on network performance. *Transportation Research Part A: Policy and Practice 34,* 6, 407–436.

May, A. and Milne, D. 2004. The impact on network performance of drivers' response to alternative road pricing schemes. In *Road Pricing: Theory and Practice*, G. Santos, Ed. Research in transportation economics. Elsevier JAI, Chapter 3, 61–86.

McKinnon, A. 2006. A review of european truck tolling schemes and assessment of their possible impact on logistics systems. *International Journal of Logistics 9,* 3, 191–205.

McKnight, L. W. and Bailey, J. P. 1997. An introduction to Internet economics. In *Internet Economics*, L. W. McKnight and J. P. Bailey, Eds. The MIT Press, Cambridge, MA, 91–118.

Mohsenian-Rad, A. H. and Leon-Garcia, A. 2010. Optimal residential load control with price prediction in real-time electricity pricing environments. *IEEE Transactions on Smart Grid 1,* 2, 120–133.

Mohsenian-Rad, A. H., Wong, V. W. S., Jatskevich, J., and Schober, R. 2010. Optimal and autonomous incentive-based energy consumption scheduling algorithm for smart grid. In *Innovative Smart Grid Technologies 2010*. IEEE, 1–6.

Morgan, M. 2011. Pricing schemes key in LTE future. Telecomasia.net. September 12, http://www.telecomasia.net/content/pricing-schemes-key-lte-future.

Morrison, S. A. 1986. A survey of road pricing. *Transportation Research 20,* 2, 87–97.

Muller, R. 2012. Satellite broadband pricing comparison. January 25, http://mybroadband.co.za/news/broadband/.

Murphy, J. and Murphy, L. 1994. Bandwidth allocation by pricing in ATM networks. *IFIP Transactions C: Communications Systems C-24*, 333–351.

Murphy, J., Murphy, L., and Posner, E. C. 1994. Distributed pricing for embedded ATM networks. In *Proceedings of the International Teletraffic Conference*. Elsevier, 1053–1063.

Musacchio, J. and Walrand, J. 2006. WiFi access point pricing as a dynamic game. *IEEE/ACM Transactions on Networking 14*, 289–301.

NTT DoCoMo. 2011. DOCOMO to introduce two prepaid data billing plans. http://www.nttdocomo.com/pr/2011/001543.html.

ODLYZKO, A. 1999. Paris metro pricing for the Internet. In *Proceedings of the 1st ACM Conference on Electronic Commerce*. ACM, 140–147.

ODLYZKO, A. 2001. Internet pricing and the history of communications. *Computer Networks 36,* 5, 493–517.

OECD. 2011. OECD broadband portal. June, http://tinyurl.com/oecd-mobile-data.

OWOSEJE, T. 2011. Orange refreshes Panther tariff with Swapables benefits. Mobile Magazine. September 7, http://tinyurl.com/Swapables.

PAHAUT, S. AND SIKOW, C. 2006. History of thought and prospects for road pricing. *Transportation Policy 13,* 2, 173–176.

PARKER, A. 2010. O2 scraps unlimited data for smartphones. Financial Times.

PARRIS, C. AND FERRARI, D. 1992. A resource based pricing policy for real-time channels in a packet-switching network. Tech. rep., Tenet Group, ICSI, UC Berkeley. TR-92-018.

PARRIS, C., KESHAV, S., AND FERRARI, D. 1992. A framework for the study of pricing in integrated networks. Tech. rep., Tenet Group, ICSI, UC Berkeley. TR-92-016.

PASCHALIDIS, I. C. AND TSITSIKILIS, J. N. 1998. Congestion-dependent pricing of network services. *IEEE/ACM Transactions on Networking 8*, 171–184.

PENDLEY, E. 1996. Deregulation of the energy industry. *Land and Water Law Review 31*, 27.

POOLE, R. AND ORSKI, C. K. 2003. HOT networks: A new plan for congestion relief and better transit. Tech. rep., Reason Foundation. February.

RICHARDS, M., GILLIAM, C., AND LARKINSON, J. 1996. The London congestion charging research programme. 1. the progamme in overview. *Traffic Engineering and Control 37,* 2, 66–71.

ROOZBEHANI, M., DAHLEH, M., AND MITTER, S. 2010. Dynamic pricing and stabilization of supply and demand in modern electric power grids. In *First IEEE International Conference on Smart Grid Communications*. IEEE, 543–548.

RURAL BROADBAND. 2012. Avanti Hylas 1 satellite broadband services. http://www.ruralbroadband.co.uk/avanti-hylas-1-8mbps.html.

SACHSON, T. 2011. Subsidized data delivery via toll-free apps. eComm 2011.

SAMADI, P., MOHSENIAN-RAD, A., SCHOBER, R., WONG, V. W. S., AND JATSKEVICH, J. 2010. Optimal real-time pricing algorithm based on utility maximization for smart grid. In *First IEEE International Conference on Smart Grid Communications*. IEEE, 415–420.

SARKAR, M. 1997. Internet pricing: a regulatory imperative. In *Internet Economics*, L. W. McKnight and J. P. Bailey, Eds. The MIT Press, Cambridge, MA.

SCHATZ, A. AND ANTE, S. E. 2010. FCC chief backs usage-based broadband pricing. Wall Street Journal. December 2.

SEGALL, L. 2011. Verizon ends unlimited data plan. CNN Money. July 6.

SEMRET, N., LIAO, R. R.-F., CAMPBELL, A. T., AND LAZAR, A. 2000. Pricing, provisioning and peering: Dynamic markets for differentiated internet services and implications for network interconnections. *IEEE Journal on Selected Areas in Communications 18*, 2499–2513.

SEN, S. 2011. Bare-knuckled wireless. Business Today. March 10.

SEPHTON, J. 2011a. The changing pricing model for mobile content. Strategy Analytics. November 24, http://tinyurl.com/ContentPricing.

SEPHTON, J. 2011b. Spain - New Orange Delfin price plan. Strategy Analytics. July 13, http://tinyurl.com/OrangeSpain.

SEPHTON, J. 2011c. Trialling mobile data - a new trend for mobile plans? Strategy Analytics. October 28, http://tinyurl.com/mobiletrials.

SHAKKOTAI, S., SRIKANT, R., OZDAGLAR, A., AND ACEMOGLU, D. 2008. The price of simplicity. *IEEE Journal on Selected Areas in Communication 26,* 7, 1269–1276.

SHEN, H. AND BASAR, T. 2007. Optimal nonlinear pricing for a monopolistic network service provider with complete and incomplete information. *IEEE Journal on Selected Areas in Communication 25*, 1216–1223.

SHENKER, S., CLARK, D., ESTRIN, D., AND HERZOG, S. 1996. Pricing in computer networks: Reshaping the research agenda. *Telecommunications Policy 20,* 3, 183–201.

SONGHURST, D. 1999. *Charging Communication Networks: from Theory to Practice*. Elsevier.

STARKIE, D. 1986. Efficient and politic congestion tolls. *Transportation Research Part A: General 20,* 2, 169–173.

SUN, J. AND MODIANO, E. 2006. Channel allocation using pricing in satellite networks. In *40th Annual Conference on Information Sciences and Systems*. IEEE, 182–187.

SVITAK, A. 2012. Demand for bandwidth drives commercial satcom. Aviation Week. January 30.

TAYLOR, P. 2011. AT&T imposes usage caps on fixed-line broadband. Financial Times. March 14.

TELECOMANDINTERNET. 2010. China's 3G mobile broadband subscribers: China Telecom, China Mobile and China Unicorn. http://tinyurl.com/7sgsk9w.

TELECOMPAPER. 2010a. Mobinil offers time-based mobile data plans. July 19, http://www.telecompaper.com/ news/mobinil-offers-timebased-mobile-data-plans.

TELECOMPAPER. 2010b. Russia targets 60-80% broadband penetration in 5 yrs. http://www.telecompaper.com/news/russia-targets-60-80-broadband-penetration-in-5-yrs.

TELEGEOGRAPHY. 2010. Fixed and mobile broadband base increases 71% in 2010 to 34.2m (Brazil). http://tinyurl.com/brazil-3g.

THE ECONOMIST. 2009. The mother of invention: Network operators in the poor world are cutting costs and increasing access in innovative ways. Special Report, September 24.

TUBE. 2012. TUBE Website. http://scenic.princeton.edu/tube/.

TWA NETWERK. 2011. Dark outlook of unlimited data plan for Korean smartphones. http://www.twanetwerk.nl/upl_documents/smart%20phone%20data%20plan%20ZK.pdf.

U.S. ENERGY INFORMATION ADMINISTRATION. 2012. Wholesale market data. http://www.eia.gov/ electricity/wholesale/.

U.S. OFFICE OF HIGHWAY POLICY INFORMATION. 2011. Toll facilities in the United States. Publication No: FHWA-PL-11-032.

VARAIYA, P. 1999. Pricing and provisioning of quality-differentiated services. In *Information Theory and Networking Workshop*. IEEE, 38.

VARAIYA, P. P., EDELL, R. J., AND CHAND, H. 1996. INDEX Project proposal.

VARDY, M. 2011. Telus tuneage: Now powered by Rdio. The Next Web. August 3, http://thenextweb.com/ca/ 2011/08/03/telus-tuneage-now-powered-by-rdio/.

VIETNAMNET BRIDGE. 2011. Vietnam telecom market remains unsustainable. http://tinyurl.com/vietnam-3g.

VYTELINGUM, P., RAMCHURN, S. D., VOICE, T. D., ROGERS, A., AND JENNINGS, N. R. 2010. Trading agents for the smart electricity grid. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1*. International Foundation for Autonomous Agents and Multiagent Systems, 897–904.

WALRAND, J. 2008. Economic models of communication networks. In *Performance Modeling and Engineering*, Z. Liu and C. H. Xia, Eds. Springer Publishing Company, New York, Chapter 3, 57–90.

WELLS, J. AND HAAS, D. 2004. *Electricity Markets: Consumers Could Benefit from Demand Programs, but Challenges Remain*. DIANE Publishing, Darby, PA.

WEN, C. AND TSAI, C. 2005. Traveler response to electronic tolls by distance traveled and time-of-day. *Journal of the Eastern Asia Society for Transportation Studies 6*, 1804–1817.

WENDERS, J. AND TAYLOR, L. 1976. Experiments in seasonal-time-of-day pricing of electricity to residential users. *The Bell Journal of Economics 7,* 2, 531–552.

WISEMAN, A. E. 2000. Economic perspectives on the internet. *Bureau of Economics, Federal Trade Commission*.

WOLAK, F. 2006. Residential customer response to real-time pricing: the Anaheim critical-peak pricing experiment. Tech. rep., Stanford University. http://www.stanford.edu/wolak.

YAÏCHE, H., MAZUMDAR, R. R., AND ROSENBERG, C. 2000. A game theoretic framework for bandwidth allocation and pricing in broadband networks. *IEEE/ACM Transactions on Networking 8*, 667–678.

YOO, C. S. 2009. Network neutrality, consumers, and innovation. *University of Chicago Legal Forum 25*, 179. U of Penn Law School, Public Law Research Paper No. 08-40.

A. ELECTRICITY MARKET PRICING

Much like today's data networks, the electricity market has experienced a capacity shortage over the past decade. To deal with this shortfall, electricity providers have explored many different pricing plans, and in particular have carried out several trials of time-of-day pricing. These trials and the accompanying literature suggest that similar plans may be adopted by network operators. Indeed, the electricity and broadband networks are quite similar in that data traffic can be compared to electricity usage, while different data applications are analogous to different energy-consuming household appliances. In light of this analogy, the discussion below gives a brief overview of the existing literature on pricing for electricity industries, focusing on time-of-day and dynamic pricing. A summary of all these pricing policies and their related literature is provided in Table III.

A.1. Static Pricing

Prior to 2000, the U.S. energy market was heavily regulated by both the state and federal governments [Pendley 1996]. These regulations imposed unique constraints on market pricing policies which are not present in the telecommunications industries; thus, in the ensuing discussion we focus on pricing proposals from the period after the deregulation of energy companies.

The electricity market has practiced static, time-of-day pricing for many years, a move accelerated by power shortages over the past decade. Thus, many works have studied empirical data from consumer trials of time-of-day pricing [Charles River Associates 2005; Herter 2007; Matsukawa 2001; I.B.M. 2007; Wells and Haas 2004; Wolak 2006]. Faruqui et al. [2009] and Wells and Haas [2004] give an overview of pricing studies in the United States. Generally, such trials use two periods (peak and off-peak) per day. These trials were conducted in a variety of areas, from California to Japan.

A.1.1. Time-of-Day Pricing. Time-of-day pricing for electricity was discussed as early as 1976, when Wenders and Taylor [1976] examined several ongoing experiments with seasonal time-of-day pricing in the U.S. In these experiments, the price varied not only with the time of the day (i.e., "peak" and "off-peak" prices) but also with the time of the year. Wenders and Taylor focus on the goals of these experiments: determining the price- and time-elasticities of different consumers, and the number of time periods per day that can be feasibly implemented. They propose experimental designs that can be used to achieve these goals, and suggest metrics for evaluating the costs and benefits of ToD pricing.

Many papers on time-of-day pricing use real data to validate theoretical models, with the goal of forecasting user demand in different periods so that the electricity provider can set prices accordingly. For example, Hausmann et al. [1979] develop a theoretical model to predict user demand in different periods as a function of the prices offered, as well as factors such as the appliance mix and weather. Real data from a Connecticut pricing trial is used to demonstrate the accuracy of this demand-prediction model. Similarly, Faruqui and Wood [2008] use real data to quantitatively estimate users' responses to offered time-of-day prices. The models used include both climate and demographic factors; Faruqui and Wood [2008] also examine the benefits of offering ToD pricing, given parameters for the electricity operator's cost structure and consumer demand and participation forecasts.

In Japan, ToD rates have been offered on a voluntary basis; Matsukawa [2001] considers the usage behavior of consumers who do and do not opt into ToD pricing. Matsukawa develops a model of electricity demand and uses it to analyze real data, finding that the household response to ToD pricing is relatively small. However, offering voluntary ToD pricing does constitute a Pareto improvement in terms of user expenditure and operator costs.

A.1.2. Critical Peak Pricing. Critical-peak pricing extends ToD pricing by allowing higher prices during some critical hours or days, e.g., especially hot days during the summer. At such critical times, demand is expected to spike more than usual, so a higher price is offered to help counteract this large peak.

In the mid 2000s, a large ToD pilot study with two different prices (peak and off-peak) as well as occasional critical peak periods was conducted in California. Charles River Associates [2005] and Herter [2007] analyze the results of this study, finding that consumers' electricity demand in higher-priced periods does indeed decline, despite constant overall demand. Charles River Associates [2005] analyzes the effect of *shifting* the peak-price period to different times of the day, a form of dynamic pricing. The results indicate that consumers do respond to the price signals despite daily shifts in the timing of the peak-price period.

Consumers with smart thermostats reduce their usage at almost twice the rate of those without such automated, adaptive thermostats, illustrating the potential for automating consumer response to time-dependent pricing. Herter [2007] considers the same data but distinguishes between high- and low-usage consumers, finding that high-usage consumers reduce their usage by a larger percentage, but low-usage consumers save more as a percentage of their bill before time-dependent pricing.

Wolak [2006] analyzes the results of a separate study conducted in Anaheim, California. In this study, ToD pricing was only implemented on critical peak days, and was effective in reducing usage during peak hours. However, users in the study received a monetary rebate for reducing their usage relative to their average usage in these peak hours, and thus had some incentive to artificially inflate their baseline usage on non-peak days. Despite this flaw, however, Wolak concludes that critical ToD pricing is a politically acceptable pricing scheme that, with the right pricing incentives, can reduce the costs of electricity providers on critical usage days.

Finally, IBM Global Services and eMeter Consulting [2007] consider the results of an Ontario-based pricing trial. Three different ToD prices were offered each day, with different rates on up to nine critical peak days, designated as such one day in advance. In hot or cold weather (i.e., summer or winter), the critical peak prices were effective in reducing usage, though usage was not significantly reduced under more mild temperatures. Surprisingly, no significant effect in usage reduction during peak periods was observed on non-critical days, though overall electricity consumption declined slightly. The small number of periods and long (several hours) period duration likely contributed to consumers' unwillingness to shift their usage to low-price periods.

A.2. Dynamic Pricing

Borenstein [2005] argues that the gains from real-time or *dynamic* pricing of electricity far exceed those of simple, static ToD pricing, and supports his arguments with simulations based on real data. In particular, long-term efficiency gains exist even under inelastic demand: users do reduce their electricity consumption during peak periods in response to real-time price adjustments. The literature on dynamic pricing in electricity markets can be roughly divided into game-theoretic pricing schemes and more traditional responsive pricing schemes.

A.2.1. Game-Theoretic Pricing. Vytelingum et al. [2010] treat the electricity market as an auction, with dynamic offers from electricity distributors and real-time responses from users (i.e., households) buying electricity. The authors propose a new auction scheme to take into account the market's varying capacity, and shows that it can achieve high market efficiency. Caron and Kesidis [2010] similarly consider market efficiency, but instead use a game-theoretic framework to focus on reduction of peak demands. In this work, users schedule their energy usage in a cooperative game with either full or limited information, so as to minimize the total load on the network.

A.2.2. Spot Pricing. Littlechild [2003] studies a form of dynamic pricing in which energy utilities are required to offer (time-varying) wholesale market prices directly to residential consumers, thus allowing greater freedom of choice to end users. An empirical trial of this pricing plan in San Diego was not successful, causing a sharp rise in spot prices. Thus, Littlechild proposes an alternative: *translational maximal price caps* [2003]. This pricing scheme, briefly implemented in the U.K., caps prices at their current levels for a certain amount of time, e.g. one or two years, in order to reduce excess profit margins over the cost of electricity distribution. In the U.K., this pricing scheme helped develop retail competition in the electricity industry, to the point where price controls are no longer deemed necessary.

Table III. Summary of previous papers on pricing in electricity markets.

Work	Pricing Type	Pilot Trial	Description
[Borenstein 2005]	ToD (2 periods)	–	Social welfare analysis of simulations with real data
[Brunekreeft 2000]	ToD (2 periods)	–	Theoretical social welfare analysis
[Caron and Kesidis 2010]	Game-theoretic	–	User scheduling as a cooperative game
[Chao 1983]	Dynamic	–	Models capacity planning w/ uncertain demand
[Charles River 2005]	ToD; critical peaks	California	Analysis of trial results
[Faruqui et al. 2009]	ToD (2 or 24 periods)	Proposed	Review of previous studies
[Faruqui and Wood 2008]	ToD (2 periods)	California	Quantitative user behavior prediction
[Hausmann et al. 1979]	ToD (2 periods)	Connecticut	Demand function estimation with real data
[Herter 2007]	ToD (2 periods)	California	Analysis of trial results
[I.B.M. 2007]	ToD (3 periods); critical peaks	Ontario	Analysis of trial results
[Joe-Wong et al. 2012]	Day-ahead	–	Price computation algorithm, simulations with real data
[Li et al. 2011]	Real-time	–	Derives equilibrium user- and provider-optimal prices
[Littlechild 2003]	Spot prices	San Diego	Cost-benefit analysis using trial results
[Matsukawa 2001]	Opt-in ToD (2 periods)	Japan	Analysis of trial results
[Mohsenian-Rad et al. 2010]	Real-time	–	Distributed scheduling of user appliances
[M.-R. and L.-G. 2010]	Real-time	–	Price prediction and appliance scheduling
[Roozbehani et al. 2010]	Real-time	–	Derives stabilizing price-feedback control law
[Samadi et al. 2010]	Real-time	–	Derives social-welfare-maximizing dynamic prices
[Vytelingum et al. 2010]	Auction	–	Auction mechanisms for time-varying supply
[Wells and Haas 2004]	ToD (24 periods)	–	Cost-benefit analysis of case studies
[Wolak 2006]	ToD (2 periods)	Anaheim	Analysis of trial results

A.2.3. Responsive Pricing. Many papers have studied responsive dynamic pricing from a user's perspective of predicting future prices and scheduling devices accordingly. For instance, Mohsenian-Rad and Leon-Garcia [2010] propose an algorithm for predicting prices one or two days in advance. The authors then provide an algorithm for scheduling devices according to these prices, with the residential user balancing impatience with the desire to save money by delaying some appliances. Mohsenian-Rad et. al.'s related paper [2010] considers the same problem, but with an emphasis on several users sharing a power source and simultaneously scheduling energy consumption in a distributed manner. More recently, Du and Lu [2011] introduced an appliance commitment algorithm that schedules thermostatically controlled household loads based on price and consumption forecasts to meet an optimization objective.

Other papers consider users' actions in conjunction with the provider's price determination. Borenstein et al. [2002] review the literature up to 2002 on modeling responses to dynamic prices and real studies of dynamic pricing. Samadi et al. [2010] consider such a pricing system from the perspective of social welfare, extending Brunekreeft's social welfare analysis of static time-of-day pricing [2000] and Chao's social welfare analysis of dynamic pricing under supply and demand uncertainty [1983].

An appliance-specific model of user demand is considered in Li et al. [2011], in which the authors suppose that consumers can store energy in batteries for future use and shows the existence of dynamic, real-time prices which jointly optimize user utility and social benefit. Roozbehani et al. [2010] more explicitly consider a feedback loop between users and provider, proposing a real-time pricing algorithm from the perspective of price stability. Joe-Wong et al. [2012] also consider a feedback loop with appliance-specific models of user demand, but propose a day-ahead pricing scheme in which users can view the prices one full day in advance.

B. SURVEY OF ROAD PRICING

> *Suppose there are two roads, ABD and ACD both leading from A to D. If left to itself, traffic would be so distributed that the trouble involved in driving a "representative" cart along each of the two roads would be equal. But, in some circumstances, it would be possible, by shifting a few carts from route B to route C, greatly to lessen the trouble of driving by those still left on B, while only slightly increasing the trouble of driving along C. In these circumstances a rightly chosen measure of differential taxation against road B would create an "artificial" situation superior to the "natural" one. But the measure of differentiation must be rightly chosen.* – Pigou, 1920. Economics of Welfare.

Transportation networks are arguably the first networks to see some form of congestion pricing. As early as the 1920s, Pigou and Knight studied the social cost of road pricing [Pahaut and Sikow 2006]. Since then several variants of congestion pricing have been proposed and adopted in the form of road tolls by transport authorities across the world, especially in busy cities, e.g., London and Hong Kong [Hau 1990]. Many of the ideas of broadband congestion pricing can in fact be traced back to these earlier pricing models.

As with electricity market pricing, the public acceptance and success of road congestion pricing strongly suggests that network researchers and ISPs may benefit from adapting similar measures. Detailed surveys of pricing in transportation networks can be found in [Morrison 1986; Gomez-Ibanez and Small 1994]; here, we briefly review the evolution of commonly used road pricing policies and their analogues in broadband networks. Table IV gives examples of actual road pricing plans that were tried out in several different countries.

B.1. Static Pricing

B.1.1. Location-Based and Time-of-Day Pricing. Location-based pricing schemes are usually implemented as point, cordon or zone pricing. The toll charges imposed by these schemes apply to vehicles passing through a designated area. Qualitatively, they are similar to static, flat-rate pricing in broadband networks.

Table IV. Summary of road pricing deployments.

Pricing Practice		Description	Example Deployments
Type	Category		
Static	Location-Based	Flat fee at fixed cordons.	U.S.
		ToD-varying fee at fixed cordons.	Singapore, U.S.
	Flat-Rate	Distance-traveled	Japan, Taiwan
		Time spent in the network	London trial
	Vehicle Type	Charge by axle number	U.S., Taiwan
Dynamic	Congestion	Combine distance and time spent	Cambridge, U.K. trial
	Dynamic	Congestion of different routes	Not implemented
	Two-Part Auction	Market trading of entry permits	Not implemented

A common variation of these pricing schemes incorporates a time-dependent factor, analogous to static time-of-day pricing. Such pricing schemes have been deployed in the U.S. on new expressways, e.g., Routes 57 and 91 in Orange County, California in 1995 [Gomez-Ibanez and Small 1994]. Other countries have also deployed location-dependent tolls; since 1975, Singapore has charged a toll to enter downtown regions of busy metropolitan areas at peak congestion times. Bergen, Oslo, and Trondheim of Norway, and Stockholm of Sweden also instituted time-of-day congestion fees to enter downtown areas in the 1990s. Harsman [2001] reports that these tolls are widely accepted by the public.

B.1.2. Distance-Traveled Pricing. Distance-traveled pricing is a static charging policy in which drivers pay in proportion to the distance traveled on a road, irrespective of the congestion condition on the road or an expressway. This is similar to static usage-based pricing in the context of the Internet, in which users are charged in proportion to their data volume without any consideration of whether the usage imposes significant negative network externalities on other users.

Distance-traveled pricing has become a more popular alternative to fixed cordons [May and Milne 2004] in implementations of real pricing schemes. For instance, Taiwan has implemented a version of distance-traveled pricing with charges varying by the time of day has been implemented [Wen and Tsai 2005]. Distance-traveled pricing is also prevalent in the U.S. [Holguin-Veras et al. 2006] and has been implemented in Switzerland, Austria and Germany for trucks [McKinnon 2006].

B.1.3. Time-Based Pricing. Time-based pricing charges users based on how long it takes to travel from one location to another; it is thus analogous to a flat hourly rate in broadband pricing. May and Milne [2000] compare distance-traveled and time-based pricing to location-based and congestion-specific schemes (see Section B.2.1) in Cambridge, U.K. They conclude that time-based pricing generally performs best overall. However, a London study showed that time-based pricing encourages reckless driving [Richards et al. 1996]. For this reason, time-based pricing has not been implemented.

B.1.4. Paris Metro Pricing. Odlyzko's Paris metro pricing proposal for broadband networks takes its name from an actual pricing scheme on the Paris metro that was used up to the 1980s [1999]. In order to regulate congestion, the Paris metro had identical first and second class coaches, but with different ticket prices. Users then self-selected into different classes, so that the higher-priced coaches were less congested. A similar concept is used today in the U.S. for roads with high-occupancy vehicle (HOV) lanes [Poole and Orski 2003]. These are identical to all other lanes of the road, except that HOV lanes are restricted to vehicles with multiple passengers. Users can self-select to take advantage of the less-congested HOV lanes by paying the higher "price" of carpooling with other passengers.

B.1.5. Pricing by Vehicle Type. Many of the tolls implemented in the United States charge based on vehicle type, e.g., trucks are charged more than passenger vehicles [U.S. Office of Highway Policy Information 2011]. Indeed, charging by vehicle type was mentioned as a criterion for efficient road pricing in a report by the U.K. Ministry of Transportation [May 1992]. Such pricing plans are analogous to app-based pricing for broadband, as both differentiate based on the intrinsic nature of the traffic considered.

B.2. Dynamic Pricing

B.2.1. Congestion-Specific Pricing. This pricing policy combines distance traveled and the time spent to travel that distance by making the price rate per mile dependent on the speed with which the vehicle travels. Congestion-specific pricing thus resembles dynamic congestion-based pricing for broadband and electricity markets. However, though it was considered for Cambridge, U.K., it was never implemented [Gomez-Ibanez and Small 1994].

B.2.2. Dynamic Road Pricing. Proposals for this pricing scheme use dynamic origin-destination models and route generation models to compute route choices that result in a dynamic network loading model, which is then used to compute route costs [Joksimovic et al. 2005]. Thus, dynamic road pricing is somewhat similar to dynamic priority pricing, in that the prices force users to trade off between a longer travel time (a more indirect route) or a more expensive, more direct route. However, this model is hard to generalize to networks with interaction between pricing and route selection.

B.2.3. Two-Part Auctions. Two-part or permit auctions are somewhat similar to spot pricing for electricity markets in that they also rely on a secondary market. As discussed by Starkie [1986], governments can sell permits to pass through congested areas, which can then be resold among users. Thus, while spot pricing takes advantage of the existing market divide of electricity providers and distributors, permit auctions in transportation networks *create* a secondary market: the government sells permits to some network users, who then redistribute the permits among themselves and other users in an auction-like process.