Evaluating Tradeoffs of Congestion Pricing for Voice Calls

Jimmy S. Shih
University of California at Berkeley
Room 473 Soda Hall #1776
Berkeley, CA 94720-1776
1 (510) 642-8919
jshih@cs.berkeley.edu

Randy H. Katz
University of California at Berkeley
Room 637 Soda Hall #1776
Berkeley, CA 94720-1776
1 (510) 642-8778
randy@cs.berkeley.edu

ABSTRACT
We conducted user experiments and simulations to understand the tradeoffs of congestion pricing between system performance and user satisfaction for a large community of users. We found that congestion pricing can be effective for voice calls because it only needs to be applied occasionally and that users are responsive to occasional price increases.

1. INTRODUCTION
Congestion pricing, varying prices according to the load, can more efficiently allocate scarce resources by influencing user behavior. It can entice users to conserve resources or defer usage during congested periods. By involving users in resource allocation decisions, it offers operators an extra dimension of flexibility in making tradeoffs between system performance and user satisfaction. It can improve system performance by avoiding overload, supporting more users with a given capacity, and saving money in terms of the needed capacity. Operators can then pass along some of these benefits to users by reducing prices or improving quality-of-service. On the other hand, it can reduce user satisfaction by interrupting users with announcements of price changes and not allowing users to predict their costs ahead of the time.

The state-of-the-art in congestion pricing consists of papers [1,2] that analyze its advantages and disadvantages. Using simulations, other papers [3,4,5,6] have shown that congestion pricing can potentially be better or worse than flat-rate pricing. These simulation results are limited to qualifying its benefits and drawbacks because they are strongly dependent on how they model users' behavioral response to price changes. Without a realistic user model based on user experimentation, it is difficult to quantify the tradeoffs of congestion pricing.

2. PROBLEM AND APPROACH
Our goal is to estimate how well congestion pricing can work in a system designed to support a very large community of users. However, it is infeasible to perform a study involving many users. Thus we propose a methodology that combines small-scale user experiments with simulations to understand the tradeoffs of congestion pricing. We use the user experiments to understand whether users would react to changing prices, how sensitive they are to different price increases, and whether they would accept congestion pricing. Then using a user model derived from the user experiments, we perform simulations to understand the tradeoffs an operator would face when there are many users, and how it should set the appropriate parameters for managing congestion pricing. Afterwards, we combine user experiments with simulations to further verify the user model. We do this by having users reacting to price changes set by an operator who is responding to the load and the reactions of many simulated users (see Figure 1).

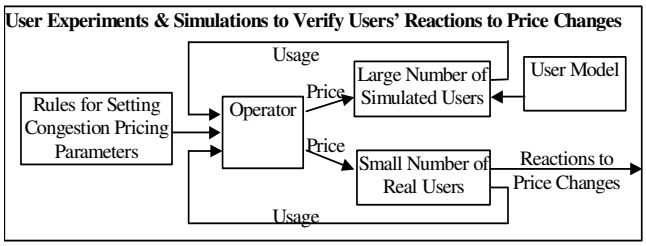

Figure 1. Evaluating the tradeoffs of congestion pricing.

We applied this methodology to voice calls by using an IP-telephony service where users can use their computers or phones to make and receive telephone calls. The service is ideal for applying congestion pricing because the Internet-to-PSTN gateway it uses is a natural bottleneck due to a limited number of phone lines to the PSTN. It is also ideal because we can easily perform load monitoring, admission control, accounting, and price settings at the gateway.

3. IMPLEMENTATION AND RESULTS
3.1 User Experiments
We evaluated congestion pricing for voice calls by first prototyping an IP-telephony service. We used a H.323 gateway, with a Primary Rate ISDN connection to the PSTN capable of supporting 23 simultaneous calls, as the Internet-to-PSTN gateway. We then developed a web interface and an Interactive Voice Response system for the gateway so that users can use their computers or phones to make and receive phone calls. When users make calls from their computers, they can see the current prices on the web interface and when they are on their phones, they will hear short announcements of price changes.

We offered the IP-telephony service to 100 dormitory students, of various majors, for over one year. We did not charge the students

real money, instead we limited each to 1000 tokens a week. This value is chosen so that an average user will run out of tokens by the end of the week. During phone calls, we charged users certain token rate per minute and occasionally adjusted the rate to observe how they would react to such change. We experimented with different ways of applying congestion pricing, measured how sensitive users are to price increases, and surveyed users about their acceptance of congestion pricing.

Based on extensive experimentation, we found that users are responsive to occasional price increases if prices change slowly, at most once every three minutes, and not during the first few minutes of a call (see Figure 2). We also found that users do not talk longer after a price decrease and that under congestion pricing, users would defer some of their usages from peak hours to off-peak hours (see Figure 3). From the user surveys, we found that users would accept congestion pricing if given a small discount.

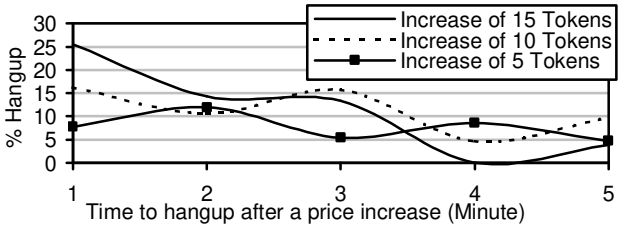

Figure 2. Prices change at most once every three minutes.

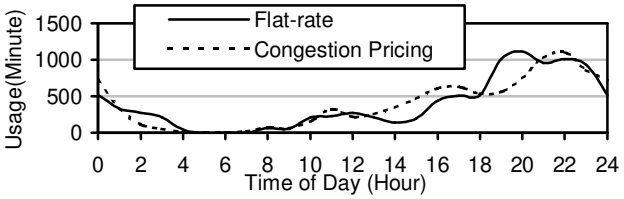

Figure 3. Calling pattern of flat-rate and congestion pricing.

3.2 Simulations

There are two questions we would like to answer with simulations. First, how should a telecommunication operator apply congestion pricing when there are many users. In particular, how should it set the following four parameters.

- *Capacity* – how many phone lines at the Internet-to-PSTN gateway it should purchase.
- *Threshold* – at what load level should it start applying congestion pricing.
- *Interval* – how quickly should prices be allowed to change.
- *Init* – how long to wait before applying congestion pricing to new calls.

Second, using the appropriate parameter values, what are the potential benefits and drawbacks of congestion pricing. The operator would like to know the expected improvements in call blocking rate or provisioning, and how often users would encounter a price change.

We used traces from flat-rate pricing to generate a user's workload model and observations from the user experiments to model his/her behavior. From the simulation results, we propose a set of rules-of-thumb operators can use to set the above four parameters. By setting them to appropriate values, we estimated that congestion pricing can reduce call blocking rate by 60% or provisioning by 20% while causing users to experience price changes in 4% of their usage.

3.3 Verification: Combining Simulations with User Experiments

Our estimates of the benefits and drawbacks strongly depend on our user model, so we further verified it by having a new group of 100 users using the service along with simulated users. The simulated users would make calls and react to price changes according to the model and the operator of the service would use the appropriate parameters to manage congestion pricing. We then observed real users' reactions to price changes under this setting. Our verification confirmed that we can easily entice users to terminate their calls after a price increase if prices change slowly, at most once every three minutes, and infrequently, at most 4% of the time.

4. CONCLUSION

Our findings indicate that congestion pricing can be very effective for voice calls. From the experiments, we found that prices should not change very quickly, the first few minutes of a call are very important to users, and users can easily be enticed to accept and choose congestion pricing. From the simulations, we determined that congestion pricing only needs to be applied during high utilization and to calls lasting several minutes or longer. Thus, operators can effectively use congestion pricing for voice calls because they can gain much in terms of reduction in call blocking rate or saving in provisioning while only needing users to experience price changes occasionally.

5. REFERENCES

[1] P.C. Fishburn and A.M. Odlyzko. Dynamic behavior of differential pricing and quality of service options for the Internet. *Decision Support Systems*, 28:123-36, March 2000.

[2] J.K. MacKie-Mason and H.R. Varian. Pricing the Internet. *JFK School for Government*, May 1993.

[3] L. Murphy, J. Murphy, and J.K. MacKie-Mason. Feedback and efficiency in ATM networks. *In the Proceedings of the 1995 IEEE International Conference on Communications. Converging Technologies for Tomorrow's Applications (ICC '96)*, New York, NY, USA, 2:1045-9, 1996.

[4] I.Ch. Paschalidis and J.N. Tsitsiklis. Congestion-dependent pricing of network services. *IEEE/ACM Transactions on Networking*, 8(2):171-84, April 2000.

[5] J.M. Peha. Dynamic pricing as congestion control in ATM networks. *In the Proceedings of the 1997 IEEE Global Telecommunications Conference (GLOBECOM '97)*, New York, NY, USA, 3:1367-72, 1997.

[6] X. Wang and H. Schulzrinne. An integrated resource negotiation, pricing, and QoS adaptation framework for multimedia applications. *IEEE Journal on Selected Areas in Communications*, 18(12):2514-29, Dec. 2000.