Is Packet-Usage Pricing Inevitable?

by Tatsuo Tanaka

Columbia Institute for Tele-Information
Graduate School of Business
Columbia University
809 Uris Hall
New York, NY 10027
(212)854-4222

RESEARCH PAPER #GLOCOM-10

Is Packet-Usage Pricing Inevitable?

Tatsuo TANAKA

ver 1.0
December 13, 1994

Center for Global Communications
International University of Japan
6-15-21, Roppongi, Minato-ku, Tokyo 160, Japan
Tel +81-3-5411-6677, Fax +81-5412-7111,
E-mail:tatsuo@glocom.ac.jp

<Address from Septemper 1995 to June 1996>
Center on Japanese Economy and Business
521 Uris Hall
Columbia University
New York, NY, USA, 10027
Tel:+1-212-854-3976, Fax:+1-212-678-6958
E-mail:tatsuo@glocom.ac.jp

Abstract

Packet-usage pricing has been proposed to cope with the problem of Internet congestion. This is a apparently natural extension of the analogy between automobile's highways and the so-called information superhighway. There are some important differences between the two, however. In this paper, by focusing on those differences, the author shows that flat-fee pricing can lead to a public planner's optimal solution. The main assumptions behind this conclusion are: (1) information should be evaluated by time consumed, not by number of packets, and (2) the marginal cost of capacity expansion is constant.

Key Word: Internet, Pricing, Congestion
JEL Cassification: D4, L86, L96

Note: Revised version of this paper will be appeared at http://tenjin.glocom.ac.jp/tanaka/ by summer in1996.

1. Introduction

The recent explosion of Internet usage has encouraged some economists to analyze this promising telecommunications medium from the standpoint of economics. Among various issues, the problem of pricing is one of the most important in terms of either economics or business. Currently, most Internet users pay a fixed fee to service providers, independent of actual usage of telecommunications lines. In other words, the cost of the Internet is zero for any specific access activity. Recent studies have argued that such flat access fees cause congestion, and that pricing in terms of usage is necessary to reduce that congestion.

For example, MacKie-Mason and Varian (1994b) analyze the problem of Internet congestion using the standard model of public good. Their conclusion is that the Internet should introduce a pricing system based on packet usage, to reduce congestion. Their proposal would have users pay a clearing-market price per packet. They also argue that pricing should raise capital for service providers, in order to expand network capacity. Similarly, Gupta, Stahl and Whinston (1994), using simulation analysis of network activity, propose coping with congestion by means of the pricing system at each server. Their proposed pricing system is based on the load (packets) which the server handles, and thus can be interpreted as a form of packet-based pricing.

The advocates of packet-usage pricing for Internet access have suggested this is necessary to promote so-called multimedia industries. The reasoning is that most multimedia services, such as real-time video transmission, require immediate transmission of data packets, whereas services such as electronic mail do not not need immediate transmission. In view of the public welfare, therefore, real-time video should take transmission priority over e-mail. But If pricing is not by usage, it is argued, e-mail might crowd out real-time video service because e-mail will continue to flow even when the network is congested, interrupting (and thus destroying the utility of) real-time video transmission. Thus, to promote emerging multimedia industries, packet-usage pricing should be introduced in order to assign priorities to various Internet services

In above these two models, flat access fees necessarily cause congestion and reduces the overall public welfare. This is a natural consequence of free access to public goods, referred to as "the tragedy of commons."(Hardin[1986]) These can be interpreted as typical application of the public goods model, such as highways. As far as the Internet obeys such a model, it should be priced in the same way as highway tolls.

However,the analogy between the Internet and highways may not hold with respect to pricing, because the information superhighway presents characteristics not seen in the model of automobile highways. If we focus on these new characteristics, we can propose alternative pricing schemes.

In this paper, we will demonstrate the possibility that flat-fee pricing, now the prevalent scheme, can expand the Internet to all potential users and realize a socially optimal solution. In other words, packet-usage pricing is not necessary. The main assumptions behind this conclusion are: (1) users measure information in terms of time consumed, not data packets, and (2) the marginal cost of capacity expansion is constant (or decreasing).

In Chapter 2, we will consider the problems of packet-usage pricing. In Chapter 3, we will develop a simple model of flat-fee pricing, and demonstrate how it can expand Internet usage to a socially optimal level. Chapter 4 will offer discussions and conclusions.

2. Problems of Packet-usage Pricing and an Alternative Solution

There are at least three problems with packet-usage pricing.

First, packet-usage pricing raises to an extraordinarily level the cost of transmitting still or video images, compared with standard electronic mail, because images require far more packets -- a still image typically demands more than 100 times the number of packets used for an e-mail message. Thus, services transmitting still or video images will have exceptionally high costs, making it difficult for so-called multimedia industries to use the Internet as a platform for development.

Moreover, packet-usage pricing will not realize the intentions of its advocates. They argue that real-time video should take transmission priority over e-mail because, technically speaking, it cannot cope with delay (effective two-way voice communication is said to be possible only at delays of less than 0.025 second). They expect packet-usage pricing to result in the desired allocation of priorities; that is, real-time video would "bid" higher than e-mail in order to obtain priority. This is an analogy to postal communications services, where users opt for first-class mail to transmit priority messages.

This analogy with the postal service does not hold up, however. Because real-time video consists of far more packets than e-mail, the video consumer will "bid" a lower price-per-packet than the e-mail user unless the real-time transmission is enormously important --

as in remote equipment operation. Note that a still image requires more than 100 times more packets than an e-mail message, and 10 minutes of video will need another 100 times more packets. Users will not pay a 100 to 10,000 times higher price, they will bid a lower price-per-packet for image data than for e-mail. As a result, e-mail would take transmission priority -- precisely opposite the intended result.

This paradox is caused by the fact that, in the models cited, users evaluate the value of information in terms of packets (actually, in the Mackie-Mason & Varian model, user's utility is an increasing function of packets). In reality, however, users do not evaluate information in terms of packets. Consider the data content of television, radio and books, for example. While a TV program consists of far more "packets" than a radio broadcast, and radio far more packets than books, "users" have not abandoned books or radio. Should an individual allocate one hour each to watching TV, listening to radio and reading, we should consider he evaluates these three media as offering equal value. We should not say he values the book's information more highly simply because he allocates the same one hour to this packet-poor medium. This example suggest that, from the view of users, information should be better measured by time consumed than by number of packets.

Second, packet-usage pricing assumes infinite use of Internet services if there is no cost for specific activity. Congestion is thus inevitable, making usage-based pricing necessary to impose budget constraints on users and reduce network activity.

We should remember, however, that there are also constraints due to available time and physical/mental capacity. For example, it is both impossible and undesirable to read and reply to 1000 e-mail messages each day. It is also unlikely the typical user would access interactive video service for more than 10 hours per day. Hence, when measuring information consumption in terms of time, we must acknowledge an upper limit to the demand for information.

In other words, there may be a saturation point in utility function because the human brain has an upper limit to its processing capability, whereas hardware develops capability virtually without limit. It is well known that heavy users of Internet service often succumb to "information overload" -- which can easily be experienced by signing on to 10 relatively active mailing lists, which will result in 50-plus e-mail messages daily, consuming most of any individual's time and mental capacity.

Given this, demand function for packets intercepts the horizontal line as shown in figure 1(a). The vertical line indicates price or cost per packet, and the horizontal line indicates the number of packets. Demand function d is a horizontal sum of demand functions d1, d2 and d3, which represent, respectively, demand for e-mail, still images and video. Since each demand curve intercepts the horizontal line, their sum (curve d) also intercepts, as shown in figure1(b). On the other hand, supply function remains horizontal until congestion emerges, because the marginal cost of additional use is zero if there is no congestion. Then, given the supply curve S, as shown in figure1(b), we need not pay a fee per packet usage (K is the point from which congestion starts). In this case, what we should pay is a fixed fee independent of the number of used packets. Thus a flat-fee pricing is a possible solution, and the opposite of packet-usage pricing.

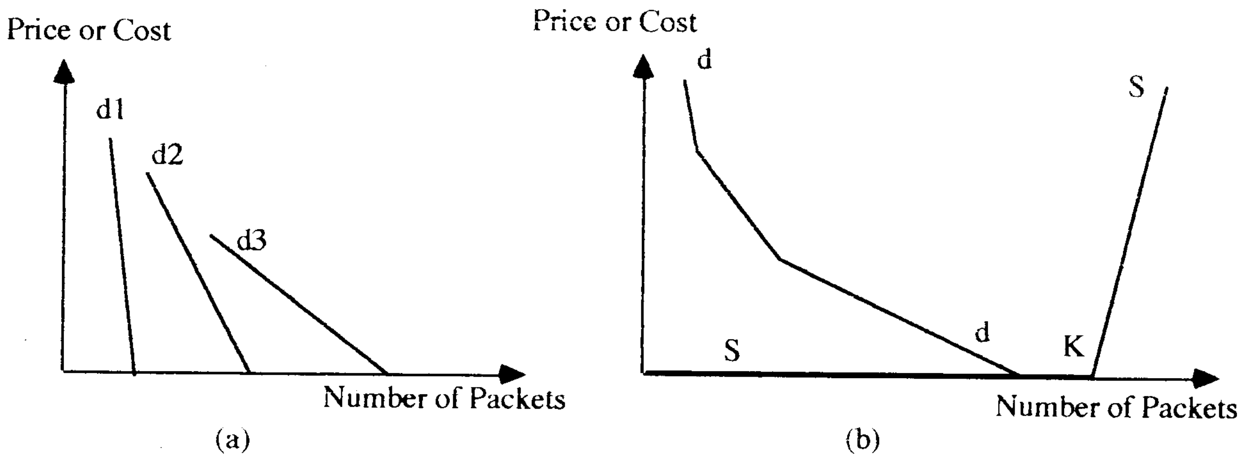


figure 1

Third, packet-usage pricing assumes implicitly the cost of network expansion is so high, or increasing so rapidly, that capacity expansion is difficult or requires considerable time. In fact, neither model gives sufficient consideration to the potential for capacity expansion.

In case of automobile highways this is valid, because expanding the width of a highway requires considerable time and money. Thus, we are forced to make due with limited highway availability for fairly long times, resulting in congestion.

The highway analogy breaks down here, however, due to the clear difference with respect to the ease of capacity expansion for the information superhighway. Expanding Internet capacity does not require the purchase of land nor the hiring of a huge work force. An even more important point is the difference with respect to potential technological progress. Telecommunications technology is progressing so quickly that the cost of expanding communications capacity has been decreasing rapidly. For example, the cost of updating from T1 to T3 is much lower than that of updating form ?? to T1 if we evaluate cost in terms of increased performance.[1]

In other words, because technical progress is limited, the marginal cost of expanding highway capacity is increasing. For the information superhighway, however, thanks to technological progress, we can expect the marginal cost of capacity expansion to decrease steadily. Thus, Internet service providers will invest more readily than highway contractors, so that congestion may be more readily alleviated.[2] In such a case, packet-usage pricing is not inevitable, and flat-fee pricing offers a viable alternative.

In the next section we will examine the potential for this solution using a simple static model.

3. Model

Let us assume there exist only two services on the Internet -- electronic mail and real-time video transmission. The amounts of user consumption of these services are indicated by x and y respectively, measured in time consumed. For example, one e-mail unit is the amount of e-mail read or written in one hour, and one video unit is the amount of transmission completed in one hour. In other words, x and y indicate time consumed for e-mail and real-time video.

Let "a" and "A" denote the number of packets required for transmitting one unit of e-mail and one unit of real-time video, respectively. Hence, when a user consumes x e-mail and y real-time video, he consume $ax + Ay$ packets. Since the bandwidth of real-time video is much higher than that of e-mail, A is far greater than a -- provisionally, we assume A is 100 times greater than a.

Let $u = u(x,y)$ serve as a utility function for a representative individual. Utility is measured in terms of value. Individuals are assumed to be identical. Note that utility is not a function of consumed packets, but of consumed time for services.

---

1 According to Mackie-Mason and Varian [1994a], while the number of packets delivered in the US's backbones had increased 128 times from 1988 to 1991, the cost had increased by a factor of about 3.2. In Swiss also, traffic of packets has increased by a yearly factor of 2.5, whereas the necessary cost to meet this traffic is estimated to increase by about only 35%. (Harms, 1994, p.341-2)

2 If the average cost is decreasing, the competitive equilibrium tends to be unstable. One possible consequence of such condition is the aggressive investment of the providers in order to drive out other firms (Murakami, chapter 5, forthcoming). This consequence means the excess capacity, not shortage of the capacity.

<User side>

To contrast, first we examine the case of packet-usage pricing. Under a packet-usage pricing scheme, the constraint is budget. Hence, individuals solve the following maximization problem.

$$\max_{x,y} u(x, y) - D, \quad \text{sub to } p(ax+Ay)=M$$

Here, p is price per packet and M is budget. D is cost of delay for each individual, supposed to be independent from the behavior of the individual because any individual is too small to affect the network as a whole. [3]

If p=0 then, optimal x and y are naturally infinite as far as u is a strictly increasing function. This is the reason congestion is inevitable when the marginal cost of access is zero, making usage-based pricing necessary.

The first order condition is

$$(\partial u/\partial y)/(\partial u/\partial x) = A/a. \qquad (1)$$

As stated above, A/a would equal 100. This equation demonstrates that, at equilibrium, the marginal utility of real-time video must be far (i.e. 100 times) higher than that of e-mail in terms of time. To put this another way, people would use real-time video over the Internet only when that service is almost 100 times more beneficial than e-mail. This is unlikely for most users. Consequently innovative usage such as video transmission will be crowded out of the Internet. This is the first problem outlined in Chapter 2.[4]

Now let's change the constraint from budget to available time. After this, we will assume users send and receive packets freely when paying a fixed fee, F.

Then, individuals solve the following maximization problem.

---

[3] Hence, the consumer's choice is not optimum from the view of social optimum, because she does not take into account the external effect of delay. I will consider this problem in the next chapter. But in any rate, it is plausible assumption that individual user does not care about her own contribution to the delay.

[4] Here I neglected the time constraint. But when the time constraint is taken into account, the conclusion is similar. Let the constraint be p(ax+Ay)=M and x+y=T. T is total available time. Then the condition for the maximization is $(\partial u/\partial y - \partial u/\partial x) = \mu p(A-a)$. Here μ is the marginal utility of the income M. This equation means that marginal utility of time consumed for real time video y is larger than that for e-mail x.

$$\max_{x,y} u(x,y) - D \quad \text{sub to } x + y = T$$

Here, T is the amount of available time for the individual. The first order condition is

$$(\partial u/\partial x) = (\partial u/\partial y). \qquad (2)$$

This demonstrates that the marginal utility of e-mail is equal to the marginal utility of video transmission in terms of time. In other words, the benefit from additional time for e-mail equals the benefit from additional time for video. This condition sounds more natural than the earlier state.

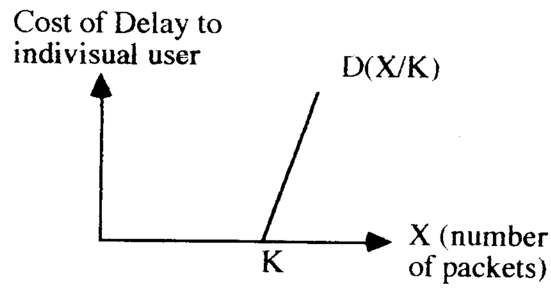
Let the solution of this maximization problem be x^* and y^*. Then $u^*=u(x^*,y^*)$ is the utility the user gains from the Internet, and u^*-D is the net utility (benefit) of joining the Internet. If this net utility exceeds the fixed fee ($u^*-D > F$), the individual will join the Internet. u^*-D is the maximum price an individual is willing to pay for Internet access. If the service provider is a total monopoly, it will impose this price, u^*-D, on users.

The packet demand of the individual user is ax^*+Ay^*. Note that congestion is not inevitable, because the individual's demand for packets is finite. If the capacity of the network is sufficiently large, users are faced with no severe cost penalty as a delay. Let n indicate the number of users connected to the network. Then $X=n(ax^*+Ay^*)$ is total packet demand. If this demand does not awfully exceed network capacity, serious delay will not occur. Let us consider this possibility from the view of the provider.

<Provider side: social optimum>

A service provider operates a network bearing fixed costs; these costs depend only on capacity K, and not on the number of transmitted packets. The function c(K) denotes cost. Capacity K is defined as the number of transmitted packets up to the point delays begin. In other words, if demand for packets, $X=n(ax^*+Ay^*)$, exceeds K, then delays occur. If demand remains below capacity K, there is no delay. Thus, we assume delay D is a function of X/K in the following way:

$$\begin{aligned} D &= 0 && \text{if } X <= K \\ &= D(X/K) \quad [D'>0] && \text{if } X > K \end{aligned}$$



Before examining service provider behavior, we will examine the public planner's optimal solution as a base for comparison. Total social benefit W is,[5]

$$W = n(u^*-D) - c(K). \qquad (3)$$

Note that public planner's maximization is concerned about only the supply side because the individual user's choice of consumptin x and y is left to each individual. Differentiating this by n and setting it to zero gives the following condition.

$$u^*-D = n(ax^*+Ay^*)D'/K \qquad (4)$$

The left side of this equation is a decreasing function of n starting from the positive value u*, and the right side is an increasing function of n starting from zero. Therefore, there exists a unique positive solution to equation (4).

The left side of (4) is the marginal and average benefit of a new user to the network. The right side can be rewritten as $n(\partial D/\partial n)$, with $\partial D/\partial n$ the marginal cost of delays to the existing individual user base caused by adding a new user. Thus, $n(\partial D/\partial n)$ is the marginal cost of delay to all users caused by adding users. Hence equation (4) means the number of users should be determined in such a way that the benefit from adding users equals the induced cost in terms of delays to the existing user base. To put this result in another way, rationing use is a rational choice to achieve a social optimum. It is well known that the rationing is an alternative means of coping with congestion. [6]

[5] We assume the utility of individuals who do not obtain access to the network is zero.
[6] If rationing is not allowed, the number of subscribers continues to increase and congestion will become more and more severe. The number of subscribers will finally reach the level at which the net benefit u*-D is reduced to zero by the cost of congestion. This is a typical example of the 'tragedy of commons.' However, since service providers usually control the number of subscribers; such a consequence is avoided.

Whether the delay D is zero or not at equilibrium depends on the slope of D. If the slope of D is large, then the delay can be zero.[7]

The above discussion assumes capacity K is fixed. In reality, however, K is also variable. Differentiating W by K gives the following equation.

$$\partial W/\partial K = nD'X/(K*K) - c' \qquad (5)$$

If the value of this equation is positive, investment should be made to raise capacity K and increase the public welfare. By substituting (4) into (5), we get

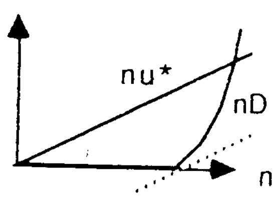
$$(\partial W/\partial K) = n(u^*-D)/K - c'$$

The sign of this expression depends on the shape of function c(K). As stated in chapter 2, however, expanding telecommunications capacity is easier than expanding automobile highway capacity, and progress in telecommunications technology will continue to reduce the cost of expansion. Thus, c'(K) will decrease, or at worst remain constant. Here we assume c(K) is a linear function without interception as $c(K) = c^*K$. Then, the above equation reduces to:

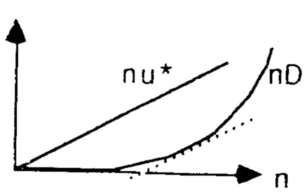
$$(\partial W/\partial K) = n(u^*-D)/K - c^* \qquad (6)$$

This form is easy to interpret. $n(u^*-D)$ is the total benefit to all users, so the first term is average benefit per capacity. Because n changes in proportion to K from the equation (4), D(n/K) remains unchaged relative to K.[8] Thus the first term of the right hand of (6) remains constant. Since the average benefit of the capacity is constant, the marginal benefit of the capacity is also constant and euall to the average benefit. The second

[7] Let us assume liner function as for D, like $D=b(X/K-1)$ when $X/K>1$. Then we can show that if $b > u^*$, the delay will be zero.



b > u*; optimum delay is zero



b < u*; optimum delay is posituve

[8] When we double K and n in equation (4), both the left and right sides remain unchanged. Thus, K and n move proportionally until n reaches the number of all potential users.

term of the right hand is the marginal cost. Thus, from the standpoint of social welfare, if the marginal benefit of capacity exceeds marginal cost, there should be investment in expanding capacity. This is a natural consequence.

An important point is that, once (6) is positive and such expansion begins, it continues until all potential users have access to the network, because marginal benefit and marginal cost are constant. Consequently, the public planner's optimal solution is the state in which all potential users have Internet access. Below, it will be assumed that (6) is positive.

Maximized social welfare is

$$n_0(u^*-D) - c(K^*) \tag{7}$$

Here, n_0 is the number of all potential users, and K* denotes optimum value. Under a certain assumption, delay D will be zero at final equilibrium.[9]

<Monopoly>

Next we will consider whether this optimal solution can be achieved by private industry. We consider two cases: one a case of monopoly and the other case of competition with free market entry.

First, we will consider the case where the service provider is a monopoly. In this case, the monopoly provider will set the flat fee F equal to the user's maximum willingness to pay, that is u*-D. Hence, the monopoly provider maximizes profit.

$$\max_{n,K} \quad \text{Profit} = n(u^*-D) - c(K)$$

Note that this maximization equation is identical to that for maximization of social welfare in (3). Consequently, the selected values of n and K by the monopoly provider are equal to the public planner's optimal values. Thus, a public planner's solution can be realized under monopoly conditions.

This consequence results from the horizontal demand curve which the monopoly provider faces. Since users are identical, all users will join the network if u*-D exceeds flat fee F. Hence, the demand for packets is horizontal at the price F. Generally speaking,

[9] Let us assume linearity, that is, $D=b(X/K-1)$ when $X/K>1$ and $c(K)=c^*K$. Then we can show that if $b/(ax^*+Ay^*) >= c^*$ then the delay D is zero at equilibrium.

when demand function is horizontal, monopoly equilibrium equals the social optimum because marginal revenue follows the demand curve.

The profit of the monopoly provider is calculated by $n_0(u^*-D) - c(K^*)$. Here, n_0 is the number of potential users, and K^* denotes an optimum value for capacity. Note that this profit equals the maximized total social benefit in (7). Hence, in the case of monopoly, all benefit goes to the monopoly provider and no benefit to users.

<Competition with free entry>

Second, we consider the case where providers compete, with free market entry. We assume new providers immediately enter the market if existing providers are earning profits. In addition, as stated above, we assume the cost function is linear as $c(K)=c^*K$ so that newly entering providers do not suffer any disadvantage from economies of scale.

Under this assumption, if existing providers want to prevent new entry, they must reduce the fee F to a cost-per-user leading to zero profits. Hence:

$$F = c(K_i)/n_i = c^*K_i/n_i.$$

in which i denotes the i-th provider. Hence K_i/n_i is the same among all providers, which means that delay D is the same among all providers.

Note that the situation with $\Sigma n_i < n_0$ can not result in equilibrium because potential users who have yet to procure access necessarily offer higher fees in order to do so. The reason is that users gain a net benefit from this low pricing; the individual user's net benefit $(u^*-D) - F$ is positive. Accordingly, potential users are willing to pay a higher fixed fee F than F. Such bidding will prompt new providers to enter the market, or push existing providers to expand capacity. In either case, investment continues as long as there are potential users. At final equilibrium, all potential users have network access. The number of providers and their size are not determined because the cost and demand function is linear except for the delay function D.

Whether the delay D disappears or not at final equilibrium depends on the slope of the functions D and c, as well as on the case of monopoly. If the delay D depends on the provider's own capacity and transmitted packets, we can expect delay will reach the same value in the case of monopoly.[10] In other words, a public planner's optimal solution is again achieved.

[10] Let us assume delay function depends on the provider's capacity and packets, that is, D is written as $D(X_i/K_i)$.

The only difference from the monopoly case is that all benefit goes to users. From the assumption of new entry, the provider's profit is zero. User's benefit is the difference between utility u*-D and the cost of flat fee F. Thus,

$$\text{User's total benefit} = n_0(u^*\text{-}D) - n_0F$$
$$= n_0(u^*\text{-}D) - c^*K.$$

This is what a provider earns in the monopoly case. Note that, contrary to the monopoly case, all benefit goes to the users. From the standpoint of users, the competitive case is better because users gain far greater benefit.

To sum up, if we assume each user's utility function is based on consumed time, rather than packets, and the marginal cost of capacity expansion is constant, we can show that flat-fee pricing will lead to a public planner's optimal solution whether as a monopoly or in market competition. Packet-usage pricing is unnecessary because severe congestion can be avoided due to the constraint of available time and the rapid expansion of capacity. Although both monopoly and free-market cases will achieve a public planner's optimal solution, the latter is preferable because it distributes more benefit to users.

4. Discussion and conclusion

The conclusion drawn herein differs from those of other authors. We advocate flat-fee pricing rather than packet-usage pricing. This difference derives from different assumptions regarding the utility function of users and the marginal cost of capacity expansion. As stated repeatedly, we assume user utility should be a function of time consumed for information rather than a function of packet quantity for information. We

---

Let us consider the case that all potential users have access, so Xi is fixed. Because the provider's profit is maintained at zero by the assumption of free market entry, the incentive to expand capacity comes from the benefit to users. User's net benefit is

$$u^* - D(X_i/K_i) - c^*K_i.$$

Users will push the provider to maximize this in terms of K_i. This maximization problem is a miniature of the maximization of profit in the monopoly case, after having involved all potential users. Accordingly, roughly speaking, we can expect a solution identical to that of monopoly. In other words, whether delay disappear or not depends on functions D and c in the same way as for monopoly.
If D depends on capacity and packets in the whole network like D(X/K), then delay is more unlikely to disappear because delay depends on other provider's capacity. Each provider's investment has a positive external effect, so they surely tend to refrain from investment, trying instead to become a free rider. Total investment will become too small and market failure occurs.

also assume the marginal cost of capacity expansion is constant thanks to technical progress. These assumptions are the major causes of our differing conclusion.

To put this another way, our conclusion stems from rejection of the analogy with automobile highways when analyzing the information superhighway. Such analogy measures the utility of the superhighway in terms of the quantity of transmitted information (packets = cars), and assumes capacity expansion will be capital-intense, as is highway construction. Under such conditions, the conclusion drawn by other authors can be valid. However, by rejecting this analogy and focusing instead on the unique characteristics of the information superhighway, this paper provides a model for alternative solutions.

Which model is valid? This question should be solved by empirical study. Do users evaluate information in terms of packets? Has congestion increased in recent years? Has the cost of capacity expansion decreased? Which is more rapid, the increase in packet quantity or the expansion of capacity? These questions remain for further study.

As final remarks, I will consider two ponts. One is a defect of the flat-fee pricing, and the other is the role of government.

The defect of this pricing system is that the public planner's solution is not social optimum because the user's choice itself is not optimal. I neglected the external effect of user's choice in the above analysis. Let us consider this external effect explicitly. Maximization of social welfare is written as:

$$\text{Max } W = n\{ u(x,y) - D(X/K) \} - c(K)$$
$$\text{sub to } X = n(ax + Ay), \; x+y=T.$$

The first order condition for x is

$$\partial W/\partial x = n\{ \partial u/\partial x - \partial u/\partial y + D'(n/K)(A-a) \} = 0$$

If delay is positive, D' is positive, therefore, D'(n/K)(A-a) is also positive. Hence $\partial u/\partial x < \partial u/\partial y$ should be held to realize the social optimum. But according to the equation (2), $\partial u/\partial x$ is equal to $\partial u/\partial y$ from the individual user's optimization. Therefore individual user's choice of x and y is not compatible with the social optimum.

The consumption y is larger than social optimum and x is smaller than social optimum. This incompatibility is because users don't consider their choice's external effect on other users through the delay. From the view of the social optimum, users should refrain the time for the real time video y which causes congestion more than the e-mail x.

In other words, public planner's optimal solution discussed in the previous chapter was the second best, not the first best.

One solution to cope with this problem is to introduce a pricing based on usage or some kind of priority system as some researchers propose.[11] In this point, packet-pricing system has an advantage over the flat-fee pricing system.

But the welfare loss of flat-fee's inefficiency will be not so large compared to the loss of "tragedy of commons", because thisinefficiency is cased by only a misallocation of available time. Note that in the case of "the tragedy of commons," user's benefit eventually reach to zero. Moreover, if the social optimal of capacity K is such that the delay D becomes zero as shown in the footnote 9 , marginal addition of the capacity makes D' equal to zero. If D' is zero, the social optimum condition becomes $\partial u/\partial x=\partial u/\partial y$, which is the same to individual user's optimal condition.

Besides, we should remember that the packet-usage pricing needs considerable cost of accounting the individual usages.[12] If the cost of accounting exceed the loss of this misallocation of time, the flat fee pricing is preferable from the view of social optimum. This is a question to be tested empirically.

Next let us consider the role of government suggested by this model. The model indicates Internet capacity will expand basically through private providers. But when the following conditions hold, government can contribute to the Internet by subsidy or regulation:

Below, delay D is neglected for simplicity. Then, the benefit to all users is nu* and the cost of construction is c*K. Dividing by n, we get benefit and cost per user, u* and c*(K/n). Since K and n move proportionally according to equation (4), K/n is fixed and c* determines the behavior of c*(K/n). So, for simplicity of notation, let C* denote c*(K/n).

In this model we assume u* and C* are constant and benefit exceeds cost (u* > C*) as in figure2(a). In this case, private firms will continue expanding capacity K and reach a socially optimal solution. But in reality, u* and C* could behave differently.

(1)Network Externality or Scale Economy

u* may represent an increasing function of K owing to network externality (Katz and Shapiro, 1985), and C* may represent a decreasing function of K due to scale economy. Let us assume either is the case.

---

[11] Varian's "smart market" and Cocchi's "priority pricing" are examples of such pricing.(Mackie-Mason and Varian, 1994b, Cocchi,Estrin, Shenker, and Zhang, 1993)
[12] Mackie-Mason and Varian (1994c) discussed the accounting problem.

Then, first, if C* exceeds u* at the initial stage, as in figure2(b), initial stimulus may be needed because private firms can not earn profit for small K(<K'). Although private industry may attempt such initial stimulus, the cost may be too great. If so, government can assume the risk of funding such initial stimulus. In the case of the U.S., ARPA net can be interpreted as having played that role, though unintentionally.

Second, an anti-trust policy is needed, because larger firms have an advantage over smaller companies. As figure2(b) shows, a firm with larger capacity K1 can supply greater benefit u*1 at lower cost C*1 than a rival firm with smaller capacity K2. Thus, larger firms usually win such competition, which creates a tendency toward monopoly. As shown in the previous chapter, free-market competition, which distributes benefit to users, is preferable to monopoly, which distributes benefit to the provider. To prevent monopolization of the market, government should apply anti-trust policy against excessively large Internet providers.[13]

[13] Regarding network externality, another solution to cope with monopoly is "free connectivity". This means all providers must connect with each other uniformly. In other words, subscribers to small service providers can enjoy the same service as subscribers to large providers. There is no difference as to service between inside and outside the net. If this is the case, network externality affect all users uniformly, and larger firms do not enjoy any advantage.

Almost all Internet service providers now connect to each other uniformly. So as a fact, not as a rule, free connectivity has been maintained so far.

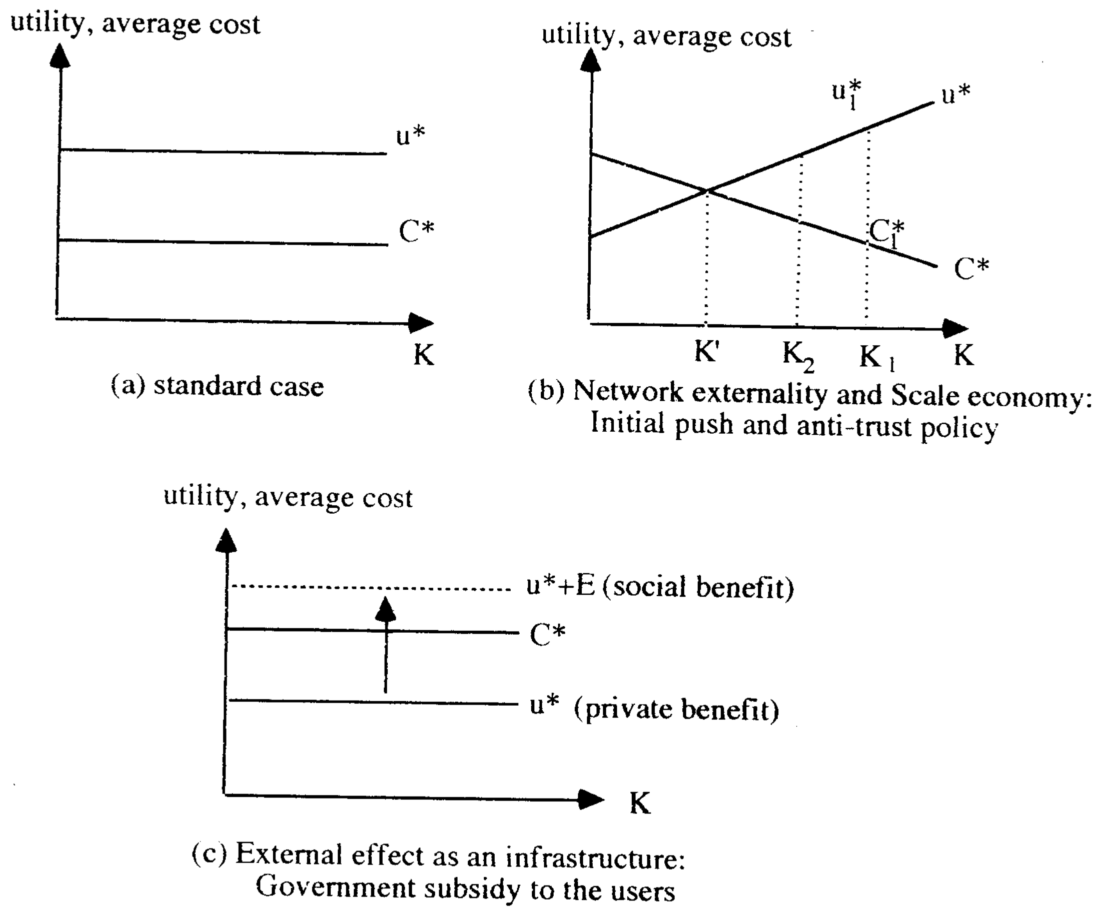


Figure 2

(2)External Effect as Infrastructure

The model presented in this paper assumes the Internet has no external effect as an infrastructure. It is sometimes claimed, however, that the Internet is an infrastructure for future industry or society (Gore, 1993). Business use of the Internet has accelerated in recent years, and educational or medical use will develop in the near future.

Let us denote these external effects per user as E. If (i) $u^* < C^*$ and (ii) $u^*+E > C^*$ as in figure2(c), then government subsidy is justified because social benefit exceeds cost whereas individual benefit is lower than cost. Government should subsidize users or providers.

The choice of users or providers makes no difference with respect to efficiency, but a subsidy to users is the better choice because a subsidy to providers could promote monopoly. Any subsidy to users should parallel the public health care system.

Note that the role of government as subsidizer is rather limited in the above discussion, because there are numerous conditions for justifying government subsidy of users and R&D, and they are unlikely to hold easily. Although network externality and scale economy are the more likely to hold, one consequence of these two factors -- the initial stimulus -- is a one-time-only subsidy. Thus, the most certain role of the government would be the different consequence of preventing monopoly.

In conclusion: Packet-based pricing appears to represent a natural proposal from the standpoint of economics when we view the information superhighway through analogy with automobile highways. The characteristics of these media differ, however. This paper has focused on two characteristics of the information superhighway. First, we assume user utility is a function of time consumed for information content, not of quantity of packets for information content. Second, we assume the marginal cost of capacity expansion is constant or decreasing. These two assumptions do not hold for automobile highways, but do hold for the information superhighway.

Under these assumptions, we can demonstrate that flat-fee pricing works well and leads to a socially optimal solution. That is to say, under a flat-fee scheme, all potential users will have access to the Internet at final equilibrium. Packet-usage pricing is not necessary.

Though both monopoly and free-market cases attain the public planner's optimum, the free-market solution is preferred because it distributes benefit to users. The most strongly expected role for government is to prevent monopoly, though there are some limited case where government may contribute to the social welfare.

This study is very tentative, because the Internet is a novel object for economists and presents many new or unknown factors. Especially, basic empirical study is needed to judge whether or not packet-usage pricing is a proper solution. This study is an attempt to show that other solution may be preferable. As stated above, empirical analysis is now necessary.

References

Cocchi, Ron, Scott Shenker, Deborah Estrin, and Lixia Zhang, 1994, "A Study of Priority Pricing in Multiple Service Class Networks," Gore Al 1993, Remarks by the vice President at the national press club newsmaker luncheon, National Press Club, Washington, DC, December 21, 1993

Gupta, Stahl and Whinston 1994 "Managing The Internet as an Economic System," proceedings of conference "Making Money in the Internet," May 8-10, 1994 University of Texas, Austin. (http://cism.bus.utexas.edu/)

Hardin, G. 1968, The Tragedy of Commons. *Science*, pp.1243-1247.

Harms, Jurgen, 1994, "From SWITCH to SWITCH* -extrapolating from a case study," in Proceedings of INET' 94, pp.341-1 to 341-6.

Katz and Shapiro, 1985, "Netwok Externalities, Competition and Compatibility," *American Economie Review*, 75, pp.424-440.

MacKie-Mason, Jeffrey and Hal Varian 1994a "Some Economics of the Internet," Feb. 17, 1994. (http://gopher.econ.lsa.umich.edu/)

MacKie-Mason, Jeffrey and Hal Varian 1994b "Pricing the Internet," paper prepared for the conference "Public Access to the Internet," JFK School of Government, May 26-27, 1993. (http://gopher.econ.lsa.umich.edu/)

MacKie-Mason, Jeffrey and Hal Varian 1994c "Some FAQs about Usage-Based Pricing," Nov. 4, 1994. (http://gopher.econ.lsa.umich.edu)

Murakami, Yasusuke, 1996, *An Anti Classical Political Economic Analysis*, translated by Kozo Yamamura, Stanford University Press, forthcomming.