

The Internet: A Future Tragedy of the Commons?[*]

Alok Gupta

Dale O. Stahl

Andrew B. Whinston

The University of Texas at Austin

Paper Presented at the Conference on

Interoperability and the Economics of Information Infrastructure

July 6-7, 1995

Rosslyn, VA

Abstract

In the future data communication networks interoperability becomes critical from both technological and business strategy perspective. Significance of interoperability has to be evaluated in terms of the overall economic performance of the system. In this paper we will present our view of the future of data communication networks, challenges in interoperability, and the economic challenges that will arise in this "real time" economy. We will provide insights derived from general equilibrium approach to these networks, e.g., what are the impacts of competition and interoperability on the competing entities which will own different parts of the network. We believe that potential excessive congestion is the single biggest obstacle in the feasibility of a global, interoperable network. We will discuss the simulation experiments we have carried out to determine approximate priority prices in real-time and discuss the potential benefits in managing congestion through such a pricing scheme. We define a framework for the policy research for an interoperable network which may facilitate electronic commerce. We also discuss the issues related to the market structure such as monopoly, duopoly, and more competitive ownership of the parts of the network and its impact on interoperability, efficiency, and economic performance of the system.

1.0 Introduction.

A "tragedy of the commons" arises when a common resource is degraded by over use. Well-known examples are ocean fisheries, urban roads, air and water. Whenever a public good (or "bad" as in the case of road congestion) does not belong to some legal entity empowered to manage the resource through usage

restrictions and/or fees, there are inadequate incentives for individual users to restrict usage to the socially optimal level. Consequently, the public good deteriorates (or the public bad swells), and the public suffers a loss relative to the potential social benefits of the commons.

Congestion on the Internet is a present and potentially paralyzing public bad. Today, Internet services include email, FTP, Telnet, IRC, Gopher, WWW, and limited real-time audio/video services such as MBONE. The evolution of these services indicates a desire for a perfectly interoperable communication network where any kind of information can be digitized and then shared, transmitted, or stored. To some extent rapid and easier access of the tools to provide technical interoperability has been the key factor in the growth rate of the Internet usage which has been estimated to be at an approximate annual rate of 100%. With the advent of WWW the access to information on the Internet has become easier and more user friendly - not to mention highly data intensive. The result of the increased availability of user friendly interfaces and the high speed of access coupled with higher awareness about these services has started creating serious congestion problems on the Internet. Consequently, the debate has started on the viability of a multi-service class, interoperable, data communication network in the long run. Several experts believe that there will be separate networks providing separate kinds of services with little interoperability - even though the infrastructure elements, such as communication lines, may be the same. We believe that an interoperable network can indeed be supported and designed; however, it will require a substantially different network management perspective. In this paper we discuss the issues related to the very subsistence of a multi-service class, interoperable, network. [\[1\]](#)

Let us first define the objective of a multi-service class network and our definition of interoperability - since there seems to be little consensus on that. The objective of an interoperable network is not simply to provide access to different forms of communication via a single interface, but also to provide the ability to customize a communication session according to need at any given time. The common definition of an interoperable network is a network which provides several services without any perceived deterioration in quality of any service, however, such a network may not ever be possible or even needed. We define an interoperable network as a network providing several services with acceptable deterioration in service quality, for example, if at any given time a full screen video conferencing is not available, users may choose to use a smaller screen, audio only, or they may even delay their conference for an acceptable amount of time. We think the benefits of providing interoperability may far outweigh the degradation in service quality - if it is managed properly. In this paper, we use the words service quality and service requirement to mean the response time delivered or needed for a particular service since once the data is digitized response time can cover practically all the aspects of service quality, for example video quality, audio clarity, and lag.

To deliver an appropriate service quality the most important factor is the bandwidth available to transmit and/or receive data. At present the Internet backbone is comprised of T1 and T3 lines with the data transmission speeds of 1.5 Mbps and 45 Mbps respectively; this is orders of magnitude greater compared to a couple of years ago when almost the entire backbone was running at 56 kbps. In the next 5 years this increase in capacity may be increased to gigabit ranges. However, the number of servers and users both have increased enormously in the last 3 years and, as mentioned earlier, the services provided on the network have become much more data intensive. Therefore, congestion will be a growing problem; in fact, congestion is the main culprit which may foil the emergence of a global, interoperable, data communication network.

We focusing on the issue of negative externalities in this paper in contrast to the much touted positive externalities. An interoperable network, by definition, has positive externalities. The users of this network have opportunities to use several different services in conjunction to customize their work environment. In many instances the value of usage is derived from the very fact that users have access to other users of this network. However, overuse of the network to such an extent that it deteriorates service quality will generate offsetting negative externalities, since just being connected is of little value unless it can be used in a productive manner. A network with sufficient interconnectivity to exhaust all the potential positive externalities will still be susceptible to a tragedy of the commons due to negative externalities. In our modeling of the Internet the positive externalities are embedded in the value of the services; we then look at the effect of congestion.

The instinctive economic solution to a tragedy of the commons is to assign property rights. Hence, it might appear to some observers that the transfer of the NSF backbone to the private sector, being an assignment of property rights over network hardware, is the solution to the problem of congestion. However, this assignment of property rights falls short of solving the problem. Each owner of network hardware will seek a pricing structure that maximizes its profits subject to the pricing structure of competing network owners and the price-sensitive demands of users. The resulting non-cooperative pricing structures will not necessarily be (and generally will not be) socially optimal [Scotchmer, 1985a,b]. This gives rise to an interoperability problem of a different kind; i.e., the problem of providing incentives to different entities to provide hardware/protocol interoperability. Note that this is simply not the question of standards but of economic incentives to design and implement standards.

An alternative approach to resolving congestion problems when private markets fail is price regulation or Pigouvian taxes. If a public entity has the necessary data on traffic and congestion, it could compute approximately optimal congestion tolls, and then impose them or levy a corresponding tax. Either way, this approach amounts to a restriction on the property rights of the network owners

which must be given a legal grounding.

In section 2 we present evidence that the Internet congestion problem is not merely theoretical nit-picking, but instead poses significant economic efficiency losses relative to the Gross Domestic Product. In section 3, we survey our existing understanding of how a privately-owned and operated Internet might function and perform relative to the benchmark of a social optimum. In section 4, we recommend a simulation modeling approach for policy research. Finally, in section 5, we outline a recommendation for NSF's involvement in the continual growth of the Internet.

2.0 Magnitude of the Problem.

Gupta, Stahl and Whinston (1995b) conducted a simulation study of the Internet to compare its performance under the current free-access policy and under optimal pricing. The model presented there is based on general equilibrium theory in economics, but departs from the Arrow-Debreu framework in a manner that makes the results computationally practical and employs the concept of a "stochastic equilibrium" in which (i) average stochastic flow rates of price-sensitive service requests are optimal for each user given the prices and anticipated delay, and (ii) the anticipated delays are the correct ex-ante expected delays given the average flow rates. An optimal stochastic equilibrium is one which maximizes the net social benefits. We derive a formula that characterizes the priority prices that support an optimal stochastic equilibrium. These prices are optimal congestion tolls. [\[2\]](#)

This equilibrium concept and associated results have significant informational and computational advantages. First, it allows the decentralization of the resource allocation process to the user level and reduces the information required for the user's decision problem to current rental prices and current expected delays. The administrative and communication costs of distributing this information pales in comparison to the associated costs of billions of Arrow-Debreu contingency markets (or even spot auction markets). Secondly, prices can be adjusted in real-time in a manner that pushes them into (and keeps them in) a neighborhood of the theoretically optimal prices, and this process can be decentralized as well. Again the computational and communication costs of this mechanism pales in comparison to that of fixed-point algorithms for Arrow-Debreu equilibrium prices. Although it might be impossible to achieve exact optimal pricing in practice for a volatile environment such as the Internet, we have demonstrated that it is possible to compute near-optimal prices in real-time. As a result of this near-optimal pricing, users with different values for the same service will choose different ways or time to obtain the same service. This, in turn, can provide substantial reduction in peak loads and will achieve better distribution of the load over time.

Computation of the optimal prices requires accurate information on the arrival rates expected waiting times. We used an iterative approach where the current estimates of the prices are computed given the historical information on flow rates and waiting times. This iterative approach can be implemented and analyzed by using simulation techniques where we estimate the prices using the transient information to guide the system towards a stochastic equilibrium. In the next subsection we present the simulation model which we used to estimate the prices and calculate the benefits.

2.1 The Simulation Model.

Figure 1 presents a conceptual model of The Internet. Essentially, we model the Internet infrastructure as a black-box, i.e., we aggregate the total delay at the server such that it appears that delay is only suffered at the server.[\[3\]](#) The users are connected to the Internet through some access providers (which we can consider as a service in itself). The access providers and the service providers, e.g., news, movies, video-conferencing, databases, etc., are "directly" connected to the Internet through a data-pipeline of a certain capacity. In this model, the capacity of the data-pipeline is essentially the bottleneck for the service providers.[\[4\]](#) In the absence of any pricing mechanism as more users demand a service, the quality of the service (in terms of data transfer rates) suffers.[\[5\]](#) Furthermore, as congestion increases at the data-pipeline, the backbone experiences more load also due to the resending of lost packets. The network service providers are able to monitor the loads at different servers, and they impose prices according to the load imposed by the servers on the backbone due to the congestion at their gateways. Since these prices are not estimated at the equilibrium conditions, they are approximate at any given time.[\[6\]](#)

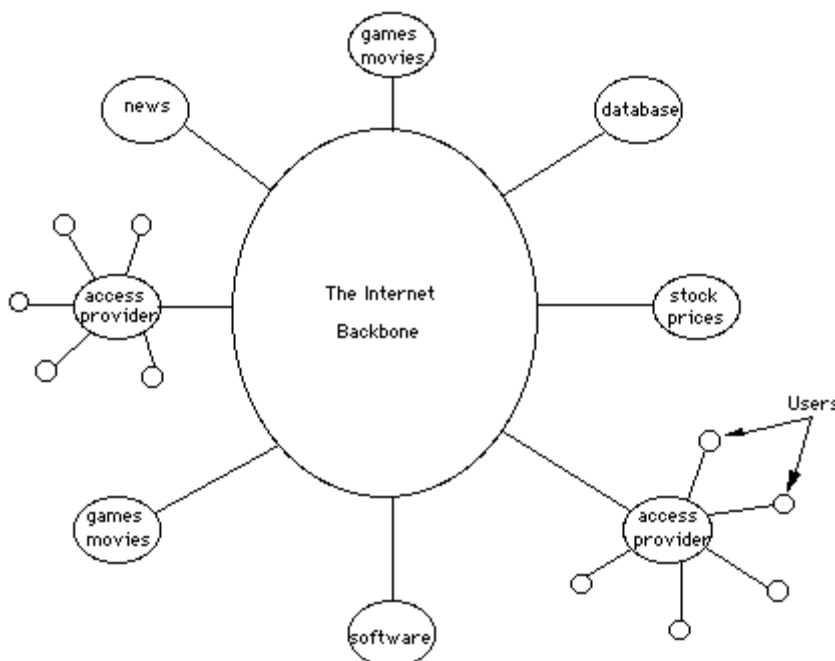


Figure 1 - A Conceptual Model of The Internet

Figure 2 provides a flow diagram of the simulation model. The arrival rates to the system are price/cost sensitive; to explore the effect of demand scaling (from exogenous growth trends), we vary a parameter X_0 . We can also interpret X_0 as the arrival rate to the system that would occur if there were free access and zero expected waiting times (i.e., the hypothetical uncongested arrival rate or the demand for network services). Note that realized arrivals into the system, being price and delay-sensitive, are always less than X_0 .

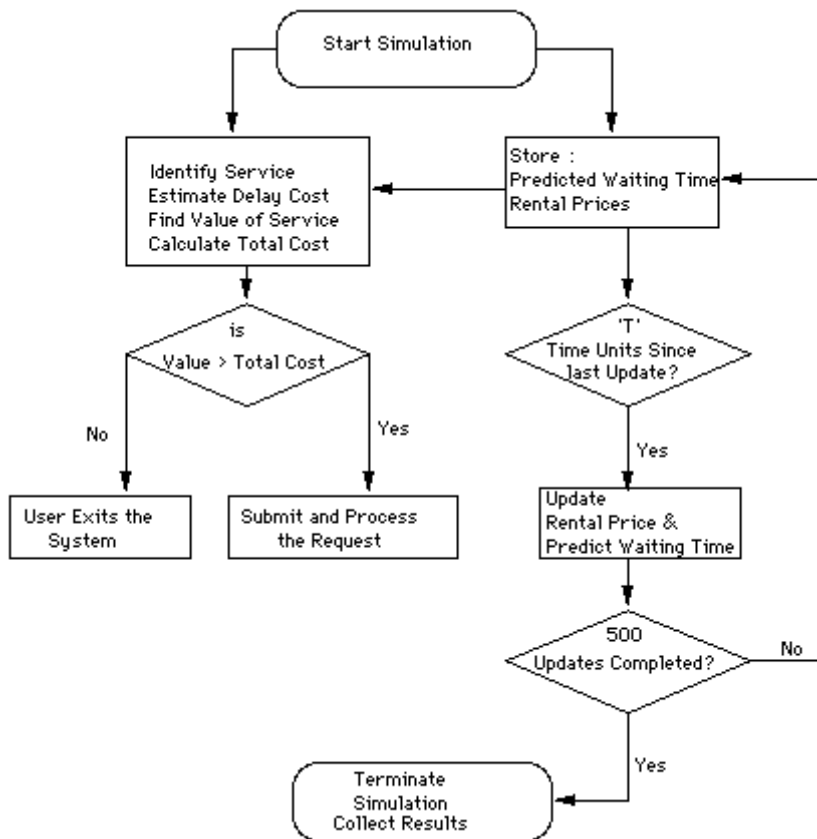


Figure 2 - Flow Chart of the Simulation Model

Upon the arrival of a potential service request the type of service required is identified; a service is characterized by the amount of computational cycles required at a server. Then, the current estimates of prices and predicted waiting times are obtained for all the servers offering the particular service (and updated every T units of time). We generate user values and delay costs from normal distributions for which the mean delay costs are set to be less than 1% of the mean job value. All the positive externalities are assumed to be embedded in the user value function. The user then evaluates the total expected cost of this service in terms of her delay cost and the service cost against her value of the service. If the total cost of the service is higher than her value for the service the user quits the system; otherwise, she submits the request for obtaining the service.[\[7\]](#)

A user's request is sent to the server which was chosen as the least cost server. If the server queue is empty, the request is immediately processed; however, if some job requests exist in the server queue, then the requests are handled in a FIFO manner.

The results presented here are based on a model which has 50 servers and 100 services. A server can provide several of the 100 services, and a service can be provided on up to 25 servers. A service "directory" was determined randomly and fixed throughout the simulation run. The capacity of the data pipelines at the servers are generated through a random process to be among the following: (i) 128 kbps (kilobits per second), (ii) 256 kbps, (iii) 384 kbps, (iv) 1.544 Mbps (megabits per second), (v) 4.0 Mbps, and (vi) 10.0 Mbps. The first three choices here represent 2, 4, or 6 multiplexed ISDN (Integrated Services Digital Network) data channels respectively, the fourth choice is the capacity of a T1 line, and fifth and sixth choices are typical of those achieved via Framerelay or SMDS (Switched Multimegabit Data Services) connections. The size of each service is also randomly generated to be in the range of 10 Kb - 15 Mb (or, 1.22 kilobytes - 1.8 megabytes); the distribution is chosen such that there are higher number of smaller services to simulate a more realistic service request distribution. The mean size of service is 2.4 Mb. The service directory and the network configuration, in terms of service sizes and server capacities, were kept constant for all the results reported

here.

We examine this system under a free access policy and optimal priority pricing. We compare these two pricing policies under different sets of load conditions by increasing the scaling parameter X_0 . A higher X_0 induces more load on the system and helps in understanding the behavior of a network with fixed capacity under increasing load.

2.2. Simulation Results

The results presented below involve two different information conditions. First, the results for the free-access policy is based on perfect information regarding the waiting times. [8] However, providing perfect information in a realistic situation is not practical because of excessive cost involved in computing new information for every new request; furthermore, several requests can be submitted virtually at the same time making this waiting time information invalid even if it was financially possible to provide perfect information.

In the more realistic condition of imperfect condition, users would need to generate (or access) estimates of expected waiting times during the current period. Because stochastic queues have the property that the variance is at least as large as the mean, the task of generating accurate estimates from a finite amount

of data in real-time is non-trivial. We have experimented with many statistical methods, and have had good results under the optimal pricing policy; however, the predictions under free-access at higher loads are sufficiently poor to result in negative net benefits. Therefore, to provide a conservative estimate of the cost of congestion, the reported net benefits from a free-access policy are based on perfect information.

The results for the optimal pricing policy are based on predicted waiting times instead of perfect information. In this case both prices and predicted waiting times are updated at the same time whereas in the former case only prices are updated after a fixed interval of time.

Table 1 displays net benefits and delay costs in dollars per month per server; these computations are based on calibrating the capacity cost according to the average cost of a T1 connection as follows. The current rental cost of a T1 line is about \$1500 per month, which implies that the cost of a 2.45 megabit/second capacity (the average capacity of servers in our simulation) is about \$2000 per month, or \$0.00077 per second. The average job size in our simulation program was 2.4 Mb, so an average server would handle one job/sec. Thus, it is reasonable to assume that the mean value of a job is at least the cost of processing: i.e. \$0.00077. In our simulation, the mean cost of delay was set to 0.008 times the mean value of a job (or only \$0.022 per hour).^[9] Table 1 compares the performance of a periodic update pricing case where delay information is not perfect with performance of free access with perfect information. This is an extremely conservative comparison, still pricing does significantly better than free access resulting in substantial gains. If we use the same information assumptions for free-access and optimal pricing, then the estimated Internet benefits of optimal pricing over free-access double.

To scale these per-server estimates to the U.S. Internet (last column), we multiplied by 250,000, as a conservative estimate of the number of servers on the Internet having an average capacity of an average server in our simulation.^[10] Thus, we arrive at a conservative estimate of the potential efficiency loss in the order of \$10 billion annually. Given the historical growth rate of the Internet (100%), the potential loss will exceed \$100 billion annually by the year 1999 or sooner.

Arrival Rate (X_0)	Benefits with prices (\$'s)	Benefits free access (\$'s)	Delay Cost with prices (\$'s)	Delay Cost free access (\$'s)	Internet Benefits* (\$'s)
50	1749.3276	1338.9921	123.9569	639.0016	103 Million
100	2776.1812	630.6611	266.6029	1732.0273	536 Million
250	3881.9087	515.7304	1629.5195	2041.7187	842 Million
500	4401.1711	483.7473	3261.6245	2145.9051	979 Million

Scenario	Benefit	Cost	Net Benefit	Capacity	Delay	Revenue
Baseline	10000	0	10000	100	100	10000
Optimal Pricing	15000	0	15000	100	100	15000
Capacity Expansion	13000	5000	8000	427	100	13000

Table 1 - Estimated Dollar Benefits per Month

with Perfect Information for Free access and Periodic Updates for Pricing

The results presented here are suggestive of the benefits of applying a near-optimal pricing scheme on the Internet. Essentially, without a pricing mechanism users with zero or low delay cost have nothing to discourage them from over utilizing the services; however, with a pricing mechanism they are forced to obtain only the services for which their value is higher than the cost. In addition they choose the appropriate service class dependent upon their service requirements. Service providers, on the other hand have incentives to provide multiple service classes because they can generate higher revenues and provide better service overall.

2.3 Capacity Expansion.

An alternative to optimal congestion tolls is to increase the capacity of the Internet so no one experiences congestion. We believe that the arguments in favor of simply over-providing the capacity on the Internet are in error. There are physical and cost limitation of providing capacity, whereas on the application level the desire for additional capacity seems boundless. Furthermore, the future applications on the Internet will have inherently different quality of service requirements. Thus, it is essential that appropriate resource management techniques be developed and tested for a multi-service class network.

To explore the cost of this capacity-expansion approach within our simulation model, we set $X_0 = 250$ and incrementally increased the capacity of every server in proportion to the aggregate delay costs experienced at that server until the aggregate net benefits rose to the level obtainable with optimal pricing and the original capacity. The required increase in capacity was 4.274 times the original capacity. The current monthly rental cost of this capacity expansion is \$6550 compared with a benefit increase of \$3184 per server. Thus, it would be uneconomical to increase capacity enough to achieve the same net benefits that could be obtained from optimal pricing.

3.0 Private Market Outcomes.

In the presence of externalities (such as congestion) it is well-known that non-cooperative private market outcomes are not socially optimal. Further, given the interoperability requirements of the Internet, the number of network competitors will be finite and of non-negligible size. In other words, the classic assumption of many small price-taking suppliers will be far from holding. Instead, the Internet market will be better described as a game with a small

number of strategic players. Further, the game has some characteristics of a Prisoners' Dilemma, which implies that the tragedy of the commons is not likely to be avoided by the non-cooperative outcome.

We know very little about how a privately owned Internet might function. The bulk of the theoretical results are confined to the unrealistic case of identical users, in which case two-part tariffs can support the social optimum. Intuitively, a monopolist who charges an access fee and a usage fee, since he can extract all the user surplus with the access fee, has the incentive to maximize that surplus by charging a usage fee equal to the optimal congestion toll [Oi, 1971]. Further, even if there are several (identical) network providers (and identical consumers), they will choose a usage fee equal to the optimal congestion toll [Scotchmer, 1985b].

Unfortunately, these results vanish in the realistic case with heterogeneous users. For example, if users differ in how they value delays, then the social optimum will involve segregation of users by delay cost into subnetworks, but some of the subnetwork owners have incentives to upset this optimal segregation. It is not hard to construct simple examples for which there does not exist a "pure-strategy" non-cooperative equilibrium.[\[11\]](#)

To illustrate this point, we have formally studied a very similar model with two independent identical privately owned networks A and B, and two types of users that differ only by their delay cost. We assume that expected waiting times depend on aggregate demand according to the standard Poisson formula. The social optimum calls for segregation of the consumers: one type uses network A exclusively, while the other type uses B exclusively. Each network owner chooses an access fee and a usage fee. It turns out that the social optimum cannot be supported by a non-cooperative equilibrium. Even a constrained social optimum in which users are indifferent between A and B cannot be sustained as a non-cooperative equilibrium. Indeed, there exists no pure-strategy non-cooperative equilibrium.

For this overly-simple model, we can prove that there exists a stationary non-cooperative equilibrium in mixed (i.e. probabilistic) strategies. However, we are skeptical about the relevance of that game-theoretic "solution" in the real-world, especially since it is beyond our ability to compute this solution even without real-time constraints. Furthermore, in a more realistic dynamic setting, there are potentially countless other game-theoretic solutions that involve complex inter temporal strategies.

To understand the complexity of this game, consider what is involved in forecasting the network outcome given some fixed pricing structure. Taking a hypothetical assignment of users to various components of the network, what are the expected waiting times throughout the network? The expected waiting times

are a function of the demand for services, which in turn depend on the expected waiting times. Thus, forecasting the expected waiting times involves the solution of a fixed point problem for each component of the network. But then given these expected waiting times, the users will choose the least cost alternative, so the tentatively assumed assignment of users is likely to be inconsistent with cost minimization. We need to solve another fixed point problem with the waiting time fixed point problem nested within just to solve for the network outcome at some fixed pricing structure.

Now the owners of the components need to consider alternative pricing structures, and for each alternative they need to solve the above two-level nested fixed point problem. A non-cooperative pricing equilibrium entails solving a higher level fixed point problem. Recognizing the rapid dynamic changing nature of the Internet, it is not clear that the players will converge to an equilibrium of the game before the game itself changes.

Formal game theory assumes that the payoff function that maps players' strategies to outcomes is given. But as we have just argued, the process that implicitly defines the payoff functions entails a two-level fixed point problem for each potential configuration of players' strategies. Thus, the real players will likely face considerable uncertainty about the payoff function, and will need to adopt statistical methods for revising their assessments.

On top of these problems, the organizational structure of the private Internet is complex. At the ground level, we have the phone lines, fiber optic cable, routers and servers. At the next level, we have service providers (data bases, shopping services, finances, etc.), and above this level, we have information brokers (such as AOL, Prodigy, and Microsoft). It is likely that the higher level brokerages will severely limit the ability of network entities to use access fees to extract surplus, in which case they will lack effective incentives to charge optimal congestion tolls.

In this environment, we may need an active public policy involving price regulation or Pigouvian taxes to avoid a tragedy of the commons. In the classic common resource situation, the imposition of a public fee (or tax) equaling the marginal social cost of use will avoid the tragedy. Such a fee is equivalent to optimal congestion tolls. The simulation results of Gupta, Stahl and Whinston strongly suggest that the computation of optimal taxes is feasible.

However, since the environment is so complex, the optimal policy is not obvious. We need to develop a model of the Internet that contains the essential and important characteristics of the Internet, which can serve as a test bed for conducting policy studies. How will alternative regulations or taxes affect the industry structure, the pricing schemes, the pattern of use across service and user classes, congestion, social benefits, and investment incentives?

4.0 A Model of the Internet for Policy Research.

The simulation model of Gupta, Stahl and Whinston is a first step towards developing an adequate model of the Internet for policy research. That model captures the network hardware and basic services, but must be extended to include ownership status, proprietary information services and for-profit service providers, and information brokers. With each of these levels, a pricing structure must be specified that could include bandwidth fees for servers and brokers, access fees and secure access code/protocols for user classes, and usage fees based on workload and priority class. These intertwined pricing structures must be coupled with an accounting/collection system that is cost-effective.

Pricing and performance (delay) data must be available to users. In a non-cooperative private market, entrepreneurs are likely to collect this data and make it available (for a fee). The presence of pricing structures will create significant incentives for the development of software (such as smart agents) to use this information to minimize the acquisition costs of services for individual users. Competitive pressures will force brokers to provide effective decision support services.

The frequency and timing of price changes will be important, but we have little a priori grounds for predicting these characteristics. For instance, will a leader emerge that sets prices first, with other (smaller) players following? Will players post prices simultaneously or sequentially (what order) as in Maskin and Tirole (1988)? Alternative regimes will need to be studied to determine the effectiveness of specific policies as a function of the price/timing regime.

Game theory, on the one hand, appears to be ideally suited to studying this market game. Unfortunately, classic game theory has virtually no predictive power in this complex dynamic environment. As illustrated in the previous section, even an extremely simplified competitive network model may have no pure-strategy non-cooperative equilibrium. On the other hand, permitting inter temporal strategies unleashes the "Folk Theorems" of game theory which say that virtually any behavior is possible.[\[12\]](#)

A more promising approach is evolutionary game theory[\[13\]](#), in which players are endowed with computable decision processes ranging from simple rules to sophisticated strategic thinking and who learn from experience.[\[14\]](#) Recognizing the immense complexity of the "game" (as discussed in section 3), considerable creativity and thought needs to go into the formulation of reasonable boundedly rational strategies. Even the partial task of estimating the outcome from the current period is daunting.

What will players assume about user expectations of waiting times? Will the player's models of users incorporate user forecasting based on published reports,

past observations and expectations, and current posted prices? How thoroughly will users search for the least-cost service provider of their anticipated bundle of needs, and how rapidly will users switch between alternative providers when anticipated net benefits change? How accurate or biased will the users' forecasts of the network status be for the upcoming period?

Will players assume that other players will continue their past pricing unchanged, or will they attempt to anticipate how the other players will react to recent experience, observed prices, and network status? In the former extreme myopia case, will the players attempt to estimate membership and demand elasticities and adopt myopic best-responses? What statistical methods will be used to uncover permanent structural variables from transient noise? How can these players combat potential bias in their parameter estimates stemming from the gross mis-specification of other players behavior. Under less extreme myopia, in which players assume that other players react to recent experience and network status (but not observed prices), how can such players estimate the other players response function (which depend on multi-dimensional continuous variables) given only finitely many observations in discrete time? How can these players combat potential bias in their parameter estimates stemming from the non-stationarity of the learning phase of the network operation?

The window of time over which data is used to calibrate the players' models of users is an important consideration. In a non-stationary environment, a window that is too long will yield parameter estimates that are slow to adjust to real changes and hence profits will suffer. On the other hand, a window that is too short will yield very unreliable parameter estimates.

The time horizon over which players attempt to optimize their strategies is another important consideration. Ideally, players should consider an infinite horizon with an appropriate discount factor, but that is unrealistic, especially considering the high level of uncertainty. Just as chess players "should" consider the entire finite game tree but the very best contemplate at most 7 moves ahead, so our network players will most likely only consider several periods, implicitly making the last period a synopsis of the uncertain future. The computational burden increases exponentially in the number of future periods considered, so given real computational costs, players are likely to use short horizons. On the other hand, if the horizon is too short, the estimates of the user's elasticities will miss effects which take longer to materialize, and these estimation errors could easily lead to "over-shooting" and other instability phenomena.

How quickly will players adjust their prices (or how long is a "period")? Rapid price changes may annoy users to the point of losing customers, and it will also increase administrative costs of computing and publishing price and accounting. On the other hand, since loads can change on an hourly basis, real-time pricing would lose much of its effectiveness if price changes were not at least on an

hourly or minute basis. Another disadvantage of infrequent price changes is potential instability that can occur in discrete-time dynamic systems as opposed to continuous-time dynamic systems. Also, frequent price changes facilitates learning through experimentation, but it also confounds efforts to discover longer-term elasticities.

The economic stakes are sufficiently high so we can safely predict that considerable private research will be devoted to devising profitable strategies for the real game. As academics, we are interested in anticipating how the real players might behave, and what public policies should be adopted to protect the common resource aspect of the Internet. A central research agenda item would be to determine what (if any) dynamically stable strategies arise. Simulation is the only practical way to pursue this question.

During such simulation runs, profits, delay costs, net benefits and efficiency losses can be measured. Alternative public policies concerning price regulation, taxation, vertical and horizontal mergers can be incorporated into the model and simulation runs can be used to determine the impacts. Does the policy tend to stabilize or destabilize the system? Who are the winners and losers? Are capital investment incentives distorted or

corrected?

A further promising methodological approach that would build on these simulation studies would be laboratory experimentation with human players. For example, a simulation program could be modified to take pricing decisions from a human subject for a specified subnetwork, and use boundedly rational strategies for the remainder of the network. Perhaps we would learn from the human behavior how to improve the performance of the simulated players, or perhaps the human subjects would learn more about the behavior of the system and the consequences of their actions. In the latter case, such a simulation/experimental platform could be valuable in the education and training of the players in the real game. Alternative experimental paradigms would include several human players, teams of players, and tournaments in which participants across the Internet could be invited to submit programmable strategies that would be entered in a multi-round tournament to see which strategies would emerge as winners.

5.0 Recommendations.

The National Science Foundation will have a continuing vital role in the future of the Internet. First, we recommend that NSF continually monitor the health of the Internet, measuring the flows, delays, prices and estimated net benefits, and publishing its reports on the Internet freely available to all users.

Second, we recommend that NSF fund research into the issues raised in this

paper. Since we have little historical experience to go on, it is vitally important that we build as quickly as possible a knowledge base concerning ownership structure, strategic pricing, and policy effectiveness. The structural changes already unleashed by the privatization of the NSF backbone will create repercussions that we may forever be reacting to. Future policy decisions should be based on sound analysis and testing prior to implementation.

Third, NSF should fund a study of the legal basis for alternative public policies. Does the FCC have sufficient jurisdiction to implement the policies necessary to avoid a tragedy of the commons? If State utility commissions become involved, how can their actions be coordinated to benefit the national interest?

References

- Abreu, D. (1988), "Towards a Theory of Discounted Repeated Games," Econometrica, 56, 383-396.
- Arthur, W. (1990), "A Learning Algorithm that Mimics Human Learning," Stanford University.
- Bohn, R., H-W Braun, and S. Wolff (1994), "Mitigating the Coming Internet Crunch: Multiple Service Levels via Precedence", mimeo.
- Dutta, P.K. (1991), "A Folk Theorem for Stochastic Games," J. of Econ. Theory, forthcoming.
- Friedman, J. (1971), "A Non-cooperative Equilibrium for Supergames," Review of Economic Studies, 38, 1-12.
- Fudenberg, D. and E. Maskin (1986), "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," Econometrica, 54, 533-556.
- Gupta, A., D. O. Stahl and A. B. Whinston (1995a), "An Economic Approach to Network Computing with Priority Classes," Working Paper, CISM, University of Texas at Austin.
- Gupta, A., D. O. Stahl and A. B. Whinston (1995b), "Pricing of Services on The Internet," in IMPACT: How ICsup4(2) Research Affects Public Policy and Business Markets, Fred Phillips and W. W. Cooper eds., Greenwood Publishing, CT, Forthcoming.
- Holland, J. (1976), "Adaptation," in Progress in Theoretical Biology IV, R. Rosen and F. Snell eds., Academic Press, 263-293.
- MacKie-Mason and Varian (1994), "Pricing Congestible Network Resources,"

mimeo.

Mailath, G. (1992), "Introduction: Symposium on Evolutionary Game Theory", J. of Econ. Theory, 57, 259-277.

Maskin, E. and J. Tirole (1988), "A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles," Econometrica, 56, 571-599.

Mendelson and Whang (1990), "Optimal Incentive-Compatible Priority Pricing in the M/M/1 Queue," Operations Research, 38, 870-883.

Oi, W. (1971), "A Disneyland Dilemma: Two-Part Tariffs for a Mickey Mouse Monopoly," Quarterly J. of Econ., 85, 79-96

Scotchmer (1985a), "Profit-Maximizing Clubs," J. of Public Economics, 27, 25-45.

Scotchmer (1985b), "Two-tier Pricing of Shared Facilities in a Free-Entry Equilibrium," Rand J. of Econ., 16, 456-472.

Shenker, S. (1995). "Service Models and Pricing Policies for an Integrated Services Internet," in Public Access to the Internet, B. Kahin and J. Keller eds., MIT Press, forthcoming.

Stahl, D. (1992), "Evolution