



University of Navarra

ZERO ENTRY BARRIERS IN AN
NP-COMPLETE WORLD:
TRANSACTION STREAMS AND THE
COMPLEXITY OF ELECTRONIC COMMERCE

Brian Subirana*

RESEARCH PAPER No 387
May, 1999

* Professor of Information Systems

Research Division
IESE
University of Navarra
Av. Pearson, 21
08034 Barcelona - Spain

Copyright © 1999, IESE
Do not quote or reproduce without permission

ZERO ENTRY BARRIERS IN AN NP-COMPLETE WORLD:

TRANSACTION STREAMS AND THE COMPLEXITY OF ELECTRONIC COMMERCE

Abstract

Transaction streams introduced in (Subirana 1998) model how transactions on the Internet* are being conducted and help to explain new types of electronic commerce intermediators that are appearing. Electronic markets on the Internet both lower some entry barriers and result in an increase of intermediators. For example, in 1998 there existed over 500,000 bookstores on the Internet. In this paper, we present a novel computational model of transaction streams and prove that in transaction-stream-based electronic markets, searching for the best price is NP-complete. This means that search costs in general can increase as a market matures. This also provides evidence against theories based on the notion that electronic markets lower search costs. Other implications of the findings presented and challenges for the further development of electronic commerce are also discussed.

(*) We use the Internet as an example of a relatively mature electronic commerce infrastructure. Equivalent conclusions can be derived in other networks.

**ZERO ENTRY BARRIERS IN AN
NP-COMPLETE WORLD:
TRANSACTION STREAMS AND THE
COMPLEXITY OF ELECTRONIC COMMERCE**

1. Introduction

Electronic Commerce is enabling new applications that are transforming many industries. In some cases, as for example in the publishing industry, the transformation is very radical. Encyclopedia publishers, book clubs, retailers are all pressed for change and many are disappearing. CD-ROM encyclopedias, virtual communities, on-line shopping are being led by new players which are displacing incumbents from their long-held market share pedestals. These new applications benefit from computer architectures such as the Internet.

Such computational infrastructures permit market agents scattered around the world to coordinate and exchange information at increasingly lower costs. This is changing the nature of electronic markets by reducing the entry costs for new entrants. Many new intermediators are appearing. Amazon alone claims to have over 300,000 associate stores (see www.amazon.com/associate). This, in turn, vastly increases the computational complexity of electronic markets. E.g. an exhaustive search for the source of a title with the best price requires visiting over 500,000 stores. Electronic commerce is creating additional computational complexity in other arenas such as advertising, security, digital notaries, and new digital currencies.

Here we are concerned with researching the connection between computational theory and the behavior of economic activities related with electronic commerce. We concentrate our analysis on modeling electronic markets. Transaction streams (Subirana 1998), (Subirana and Carvajal 1998b) model the electronic transaction of goods and services on the Internet. In this paper we provide a computational model of transaction streams and show that certain properties of transaction streams are NP-complete. In section 2 we review existing literature on electronic markets. In section 3 we define transaction streams and provide various examples. In section 4 we provide a computational model of transaction streams. In section 5 we show that the problem of finding the providers with the best price is, in general, an NP-complete problem. The final section discusses the implications of the findings presented.

2. Electronic Markets

Previous work has stressed the roles of markets and hierarchies as distinct mechanisms for coordinating the transactions related to the flow of intangible goods, materials or services through adjacent steps in the value-added chain (Malone, Yates & Benjamin 1987). Markets coordinate the flow through supply and demand forces between different individuals and firms. Malone et al. (1987) contend that the evolution of information technology, by reducing the costs of coordination, is leading to a shift toward proportionately more use of markets compared to hierarchies to coordinate economic activity. They also

argue that electronic markets are a more efficient form of coordination for certain classes of product transactions whose asset specificity is low and whose products are easy to describe.

In these authors' view, electronic markets will evolve from electronic single-source sales channels to biased markets where one of the providers uses the market transaction mechanisms in its favor, from there to unbiased markets, and finally to personalized markets. Personalized markets are those in which the customers can use customized aids in making their choices. For example, some airline reservation systems allow the user to set preferences such as departure time, seating assignment and rates, which are then used in subsequent transactions. The airline market is therefore customized to the users - different users have different options depending on their preferences.

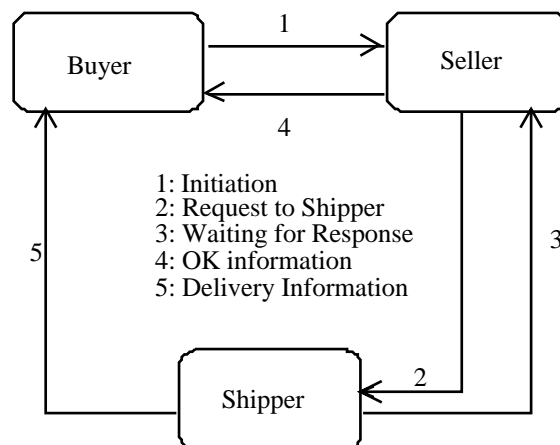
Bakos (1991, 1997, 1998) analyses the impact of electronic markets through the analysis of search costs. Buyers must, directly or indirectly, pay search costs to obtain information about prices and product offerings available in the market. Electronic markets have a vast impact on search costs because of the coordinating effect of information technology. Using economic theory, he shows that this reduction in search costs plays a major role in determining the implications of these systems for market efficiency and competitive behavior. This reduction results in direct efficiency gains from reduced intermediation costs, and in indirect but possibly larger gains in allocation efficiency from better-informed buyers. The benefits realized by market participants increase as more organizations join the system, leading to network externalities and resulting in increasing rents for the consumers (Katz & Shapiro 1985). Our findings suggest that on the Internet new dynamics are producing opposite effects: as the complexity of electronic markets increases, entry barriers are diminished and it is either not possible or very costly to compute the optimal price.

The logic of Bakos (1991) is based on 5 economic characteristics of electronic market systems (search costs, increasing returns, switching costs, entry costs, maintenance costs) that have been changed by recent market and technological developments around the Internet. The first two characteristics that he analyses have been enhanced. First, some search costs as stated above have been apparently reduced through browsing or robots; and second, the benefits realized increase as more organizations join the WWW electronic markets. However, the last three characteristics have been drastically changed by open standards such as TCP/IP, HTML and Java and the vast market adoption of the Internet as a platform for economic activities. Indeed, Bakos' (91) third argument was that "electronic marketplaces can impose significant switching costs on their participants". However, in most Internet enabled marketplaces, the switching costs are often reduced (or nil). This is because the Internet provides an open standard for information transmission (i.e. a dedicated line is no longer an entry barrier), and Java and XML can be used to create an interface that can reduce the switching costs by translating among alternative interface options (i.e. the interface software is no longer always a switching barrier). The entry costs (point four of Bakos' argument), defined as large system development and maintenance costs, have also been substantially reduced. Many open solutions exist that can be adopted and integrated into a solution with a fraction of the development cost that was necessary in the past. Pre-WWW network technology was expensive and proprietary. This explains why early success stories such as airline reservation and hospital supply took place in markets with a big demand for immediate and distributed coordination.

Other relevant work related to electronic markets also exists (Hagel & Armstrong 1997) (Rheingold 1993). Benjamin & Wigand (1995) evaluate how the emergence of a national infrastructure such as the Internet can change the different segments of an industry value chain. The analysis is centered around the National Information Infrastructure (NII) initiative and evaluates what was once a vision of an Office of Technology Assessment (OTA) report: "The network will, in many instances, serve as the market. When this occurs, market structure will

depend as much on network characteristics and the economies of networks as it does on relationships among firms” (OTA 1994). Another approach to analyze electronic markets is to model electronic transactions and then study the behavior of the resulting model. In the next section we will introduce one such approach based on transaction streams.

Figure 1: Exchange procedures according to Schmid and Lindemann (1994)

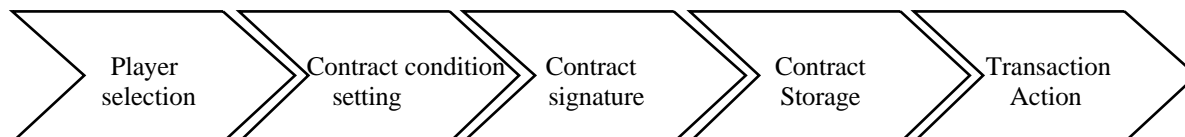


3. Transaction Streams and Entry Barriers

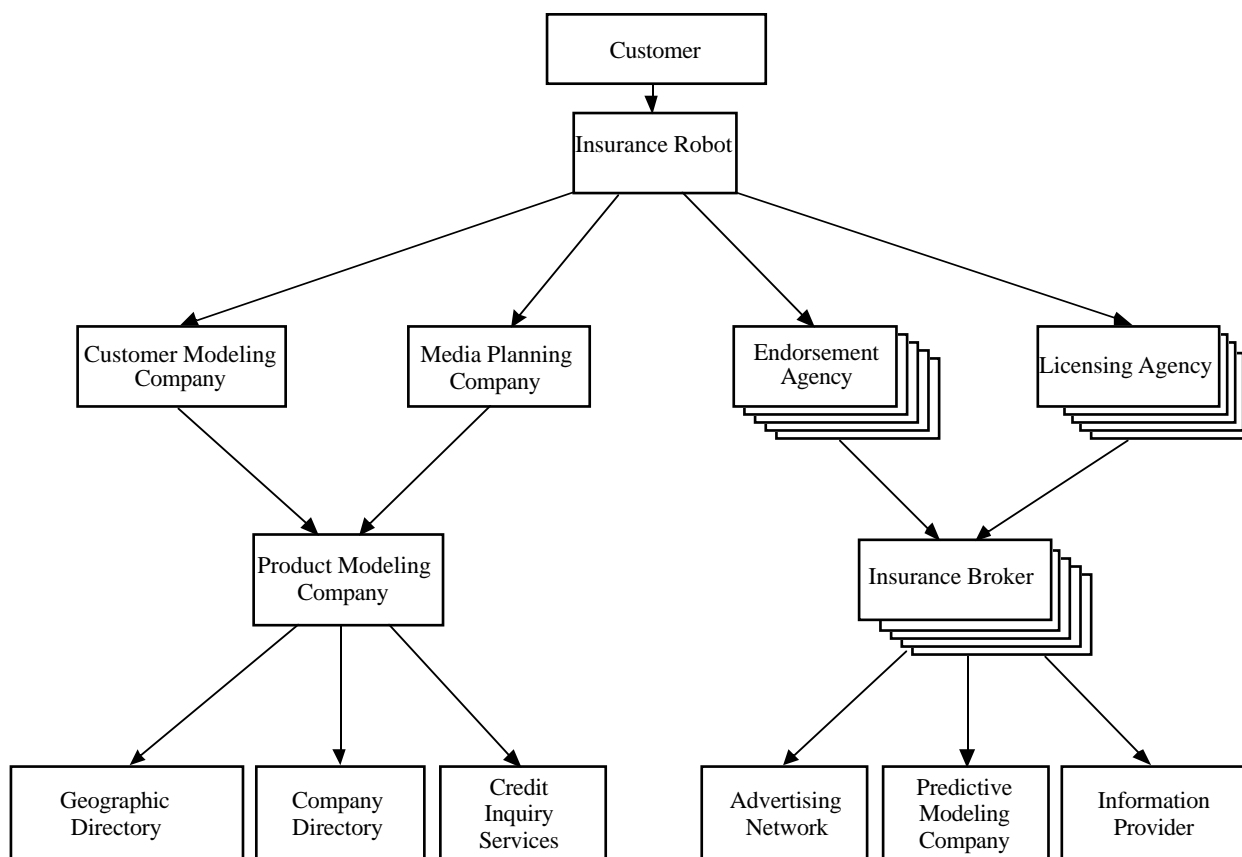
A transaction is the establishment of a contract between a set of agents, such as people and firms, to perform a given action. By "contract" we mean an agreement between a set of agents to perform a course of actions, usually with detailed implicit and explicit conditions and alternative paths (Lee 1998). Examples are purchasing a cinema ticket, placing an ad in a newspaper, purchasing a book, etc.

An on-line transaction involves at least three roles (see Figure 1): buyer (customer), seller (supplier) and shipper (distribution service). A market transaction corresponds to a finite number of interaction processes between market participants in various roles. Information technology systems hold the potential to change the interactions between participants by helping to leverage buying power, and to streamline complex and inefficient processes (Gebauer et al. 1998). The goal is to initiate, arrange and complete a contractual agreement for exchanging goods and services in the most efficient manner. The interaction processes involved can be grouped into classes to form the phases of a market transaction (Schmid and Lindemann 1994).

Before a transaction is completed, five processes need to be enacted (Subirana, 1998): player selection, contract condition setting, contract signature, contract storage and transaction action. We will refer to them as the transaction processes. These transaction processes can be represented graphically as shown in Figure 2.

Figure 2: Transaction processes

Transaction streams are electronic markets in which more than one player are involved in the five transaction processes. Many examples of transaction streams are given in (Subirana and Carvajal 1998a,b). Figure 3 provides a hypothetical example of transaction streams that illustrates the application of the first transaction process, player selection, to the insurance industry. In the example, we have plotted the players involved when a sample user “clicks” on an insurance robot to purchase an insurance product. The robot itself calls a set of licensing agencies, endorsement agencies (Lai et al. 1997) and other companies such as modeling and advertising firms. Licensing and endorsement agencies in turn search for insurance brokers who, through third-party predictive modeling and endorsement agencies, assess the risk premium that should be used. The robot then collects all the answers and provides the aggregated response to the user. Note that insurance brokers feed their systems with information provided by other companies as well. This enables them to optimize their hit-ratio while minimizing risk.

Figure 3: Cascading Effect in Transaction Streams

Entry costs are practically zero in many of these cases. For example, the Amazon Associate Program allows a user to configure a bookstore with almost no effort. Given that Amazon provides all the back office infrastructure, all that is needed to start a bookstore is a connection to the Internet and a Web editor. Amazon provides an economic incentive for this service in the form of the commission pay-back. Zero entry barriers (an Associate can be fully operational in less than 10 minutes) is one of the characteristics that explains the success of the Amazon Associate Program: over 300,000 bookstore Associates in less than two years. Amazon is just one example of over 1,000 companies offering their “front and back office” services to those who want to use them. This is populating the Internet with millions of URLs that offer a myriad of services. This, in turn, results in transaction streams that can involve simultaneously an increasing number of players as the Internet matures. In the next section, we will introduce a computational model of transaction streams. In the following section, we will use the model to prove that, in general, finding the best price is an NP-complete problem.

4. Modeling Transaction Streams and the Transaction Streams Price Search Problem (TSPSP)

In this section we will introduce a computational model for transaction streams that will allow us to prove that searching for the best price is an NP-complete problem. This means that finding the best price will be increasingly complex as the size of the data to consider increases.

The computational requirements for NP-complete problems is such that linear increases in the input data result in exponential increases in the time needed to solve the problem. Even with abundant computational power, exponential demands soon collapse any existing computer. A well known problem that is NP-complete is the Travelling Salesman Problem (TSP). Indeed, finding an optimal route (in terms of distance or time) to visit a number of cities is NP-complete. From a practical point of view this means that, in general, one needs a simplified or semi-optimal algorithm to establish the route. Observe that it would not be difficult to write a program that finds the optimal solution by trying all possible alternatives. The issue is not whether the algorithm exists to compute the solution but what are its computational requirements. If one has 100 cities, visiting all of the alternatives requires visiting the 100 factorial combinations (that is, $100! \approx 9.3 \times 10^{157}$) or 9.3×10^{157} –quite a large number. If we need 0.001 seconds to evaluate each alternative on a 500MHz machine, then the complete problem would be solved in 3×10^{147} years! What makes exponential problems hard to manage is that a little increase in the input data results in a much greater increase in the computational requirements. For example, if instead of 100 we have 150 cities (an increase of just 50%), the above example would yield computational requirements of 1.8×10^{252} years, an increase of 10^{95} times!. Observe that there may be cases in which a trivial solution exists (or one without exponential requirements), for example if all cities lie in a straight line. However, we do not know if there is an algorithm that can solve the most general traveling salesman problem without exponential requirements.

There are many other problems that are of this nature and for which we do not know if there are non-exponential solutions. In fact, all of these problems belong to a class of problems (or computational languages) that are called NP-complete. One can prove that if there is a non-exponential solution for one of these problems, then there is a non-exponential solution for all of them. This is one of the most challenging open research problems in computer science, which is often quoted as the P=NP question. In the rest of this section we introduce a

computational model of transaction streams that will lead us to prove, in the following section, that certain properties of transaction streams are NP-complete.

We will call "market agent" any participant in any of the transaction processes, and we will represent it with the notation N_i . A market agent service will be the digital good or service provided by the market agent. The market agent service price will be the price allocated to the good or service provided. Please note that this price will, in general, be different for each service provided. A given good will be provided through a transaction stream involving a set of market agents: $(N_1, \dots, N_i, \dots, N_n)$ (1).

From the point of view of a market agent, N_i , the examples shown in Figures 3 can be generalized with the scheme shown in Figure 4. A market agent, N_i , is providing a service, $SERV_i$, to another market agent (its client) which we call CLI_i (the client itself will be one of the N 's involved in the transaction stream). To provide this service, N_i requires input from a host of suppliers $SUP_{i,1}, \dots, SUP_{i,k}$ (note that each supplier will in turn be one of the N 's involved in the transaction stream).

To determine the market agent service price, we need to know the context in which the service is being provided, that is, the $(N_1, \dots, N_i, \dots, N_n)$ as well as the $SERV_i, CLI_i, SUP_{i,1}, \dots, SUP_{i,k}$ for each node N_i . We will call *an agent arrangement* the set $(N_i, SERV_i, CLI_i, SUP_{i,1}, \dots, SUP_{i,k})$. In general, not all agent arrangements of nodes will provide a feasible transaction stream. For example, doubleclick.com can not decide what ad to place until Altavista has contacted them with the page they should serve the ad on. We can define a function, the transaction stream feasibility function, which determines whether within a given service context or agent arrangement $f(N_i, SERV_i, CLI_i, SUP_{i,1}, \dots, SUP_{i,k}) \rightarrow (T, F)$ a service can be provided or not. If the result is true (the service can be provided) then we can define, for a given agent arrangement, a price function $f(N_i, SERV_i, CLI_i, SUP_{i,1}, \dots, SUP_{i,k}) \rightarrow \text{Price}$.

A *transaction stream arrangement* will be a set of agent arrangements which together deliver a given service. Thus, by transaction stream arrangement we mean the market agents that need to be involved, the service that each of the market agents needs to provide and the order in which these market agents need to proceed.

The *Transaction Streams Price Search Problem* (TSPSP) is: "given a set of market agents $(N_1, \dots, N_i, \dots, N_n)$, and a service to be provided, say $SERV_0$, what is the transaction stream arrangement that delivers the service with the minimum cost". In the next section we show that when the number of Nodes increases, the TSPSP problem becomes NP-complete.

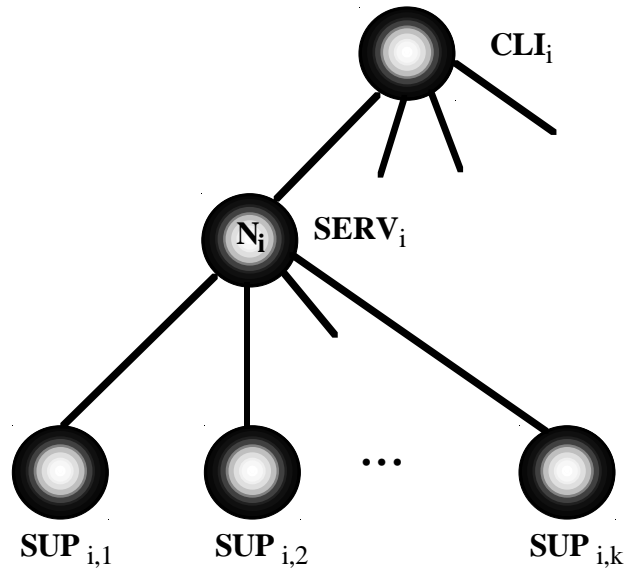
5. Transaction Streams are NP-Complete

To demonstrate that the transaction streams price search problem is NP-complete, we will reduce the traveling salesman problem (TSP) to the TSPSP problem. This is a common method of showing that problems are NP-complete and it consists in transforming any TSP instance into a TSPSP with two warrants (Hopcroft and Ullman 1979). The first warrant is that such transformation is done with polynomial computational requirements. The

(1) It is important that each node N_i correspond to a single market service agent transaction. If a company is involved in more than one service to provide a transaction, we will assume that it is a "different" market agent and we will include a new node N_j to represent it.

second warrant is that once the TSPSP problem is solved the TSP problem can be solved (using the TSPSP solution) with polynomial computational requirements. Observe that if the two warrants are satisfied and if there was a polynomial solution for the TSPSP problem, then there would be a polynomial solution for the TSP problem that would consist in transforming TSP problems into TSPSP problems.

Figure 4: Transaction streams feasibility network



We begin by introducing the TSP problem, and continue by formally presenting a simplified instance of the TSPSP that will lead us to the reduction of the TSP problem to the TSPSP problem. In introducing the TSP problem we follow the notation of Hopcroft and Ullman (1979).

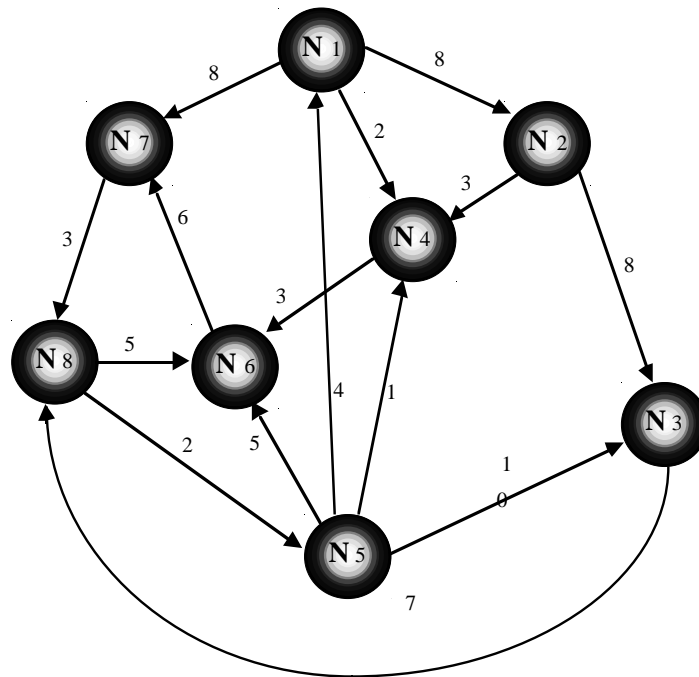
Given a graph G , a Hamilton circuit is a circuit that visits each vertex exactly once and returns to its starting point. The traveling salesman problem is: Given a complete graph with weights on its edges, what is the Hamilton circuit of minimum weight? It can be proven that the TSP problem is NP-complete. In fact, determining whether a graph has a single Hamilton circuit is an NP-complete problem.

Given a graph G , composed of a set of nodes $(x_1, \dots, x_i, \dots, x_n)$ and a set of edges $(e_1, \dots, e_j, \dots, e_m)$, each edge will consist of triples $e_i = (x_{ik}, x_{ij}, w_i)$, where x_{ik} and x_{ij} are the two nodes of an edge and w_i is the weight associated with the edge. We need to find a polynomial way of reducing the TSP problem on this graph to a TSPSP problem on a set of market agents or nodes $(N_1, \dots, N_i, \dots, N_n)$.

The set of market agents that we will construct will be one in which each graph node x_i corresponds to a market agent node N_i . We will simplify the transaction stream feasibility function so that, to complete a transaction stream, all nodes need to be involved once and only once in a linear order (not in a tree order as in Figure 4), following always the available edges. The cost will be equivalent to the weight of the edge (independently of the order chosen). Figure 5 shows one such graph in which all the market agents are represented and where arrows indicate whether a company can provide a service to another and at what price. Note that with this simplified version of the TSPSP problem we have reduced it to the TSP q.e.d.

The proof illustrates some of the inherent complexities in the TSPSP problem. Namely, that of finding an optimum combination of services. As soon as the Internet opens up the possibility of aggregating services, the complexity of finding the optimum combination explodes. However, there are other aspects of the TSPSP problem that are also complex. For example, that of deciding whether an agent arrangement is feasible or not as computed by the transaction stream feasibility function.

Figure 5: Transaction streams feasibility network



6. Conclusion

The Internet enables rapid coordination of various providers, resulting in an increase in the number of agents involved in a given product or service fulfillment process. In this paper we have shown that with such increase in the number of alternative market agent aggregations, the computational complexity of searching for the cheapest option becomes NP-complete. Perhaps the most important consequence of the proof presented in section 5 is that under certain circumstances the best price cannot be obtained in an efficient manner. Furthermore, the Internet is forging an environment in which such complexity is increasing, not decreasing. The proof sheds some light on the basic factors that affect complexity. These include the payment mechanisms, the exponential number of transaction stream arrangements, difficulties in automatically describing preferences, the transaction stream feasibility function, etc. There is one additional factor that is worth mentioning and that is not included in the model. The search cost, although apparently small for Web price searches, can be substantial if an extensive computational search is performed. The bandwidth requirements of an NP-complete search may have no bounds. If many shopping robots start looking for the best combination of products, we may soon see a collapse of the current infrastructure. If a mechanism is developed to establish the bandwidth costs, then optimising cost will become even more complex and possibly not deterministic.

Significant entry barriers can be built by understanding the forces that create complexity. Associate programs, interest links, yellow pages, fidelity programs (such as clickrewards, beenz) and advertising networks are some of the models companies are using to harness customer attention. This results in better information, which may be used to guide yield management techniques to determine optimum price quotes. The more complexity in the system, the more likely it is that a dominant competitor can use its information and computational resources to secure value appropriation from the market.

Since blind search will become complex and hardly ever useful, it is not surprising that after a phase of fast content growth, Internet developers are now getting ready for a world rich with metadata (data about data) that can better support search functions. For example, RDF and XML (see <http://www.w3.org/> for more details) are two standards being developed to assist in describing data. This means that the Internet may be evolving towards an environment where content providers can describe their data (for example, "This site is in English and is devoted to parents of high-school children") and users can describe their intentions (e.g., "I am interested in sites related to science education"). Metadata can be used for many purposes, among them that of facilitating price comparisons. For example, a user may request a page from a software distributor and indicate with its "metadescrptors" some preferences (e.g., "I am a registered user, I run programs on a Unix workstation, I have such and such credit cards and belong to such and such community"). The software distributor may use these preferences to customise its response and use "metadescrptors" in turn to inform the user about specific offer alternatives. (e.g. "You can buy this package for a 20% discount or this other one for a 30% discount if you become a beta tester and share the IPR of the code you develop with community X"). If the metadescrptors are standard, then the user may be able to search its most common suppliers with a robot.

Transaction streams model the trading of goods as the aggregation of services provided by different market agents. The model presented leaves room for future research. What is the computational complexity of building the transaction streams feasibility network? What is the complexity of the feasibility and cost functions? What if the service can be parameterized? For example, when you perform a wire transfer, you have the option of sending all the amount through the same bank, but it may be more cost-effective to take advantage of discounts and chop the amount into smaller pieces. How can we optimise cost then? What are the fiscal and regulatory implications of the complexity findings reported here? How can risk premiums and other insurance functions be best performed in an NP-complete world?

The availability of an electronic commerce Internet infrastructure that can support search functions is only one of the requirements for a fully functioning information society. Companies, individuals and governments will have to decide on what is the best way to handle this emerging complexity. Should prices be published in a standard metadata language? What should be done with offers, coupons and fidelity incentives that distort price comparisons? Should there be a simplified computational tax paradigm to prevent filing taxes from also becoming an NP-complete problem? □

References

- Armstrong, A. and Hagel, J. (1996) "The Real Value of On-Line Communities", *Harvard Business Review*, Vol. 73 (3), pp. 134-141.
- Bakos, J.Y. (1991) "A Strategic Analysis of Electronic Marketplaces", *MIS Quarterly*, September, pp. 295.
- Bakos, J.Y. (1997) "Reducing Buyer Search Costs: Implications for Electronic Marketplaces", *Management Science*, Vol. 43, No. 12, December 1997.
- Bakos, J.Y. (1998) "The Emerging Role of Electronic Marketplaces on the Internet", *Communications of the ACM* (Association for Computing Machinery), Vol. 41, Issue 8, pp. 35-42.
- Benjamin, R. and Wigand, R. (1995) "Electronic Markets and Virtual Value Chains on the Information Superhighway", *Sloan Management Review*, Winter, 62.
- Cronin, M.J. (1996) *The Internet Strategy Handbook: Lessons from the New Frontier of Business*. Boston, MA: Harvard Business School Press.
- Gebauer J. and Segev A. (1998) *Assessing Internet-based Procurement to Support the Virtual Enterprise*. Berkeley, CA: The Fisher Center for Information Technology and Management, University of California.
- Granoff, P. (1997) "Virtual Vineyards", MIT'S *Technology Review*, Aug-Sept.
- Hagel III, J. and Armstrong A.G. (1997) *Net Gain. Expanding markets through virtual communities*. Boston, MA: Harvard Business School Press.
- Hagel III, J. and Lansing, W.J. (1994) "Who owns the customer?", *McKinsey Quarterly*, No. 4.
- Hopcroft, J.E. and Ullman J.E. (1979) *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley.
- Lai, C.; Medvinsky, G. and Clifford, B. *Endorsement, Licensing and Insurance for Distributed Services*. Cambridge, MA: MIT Press, p. 417.
- Malone, T.W.; Yates, J. and Benjamin, R.I. (1987) "Electronic Markets and Electronic Hierarchies", *Communications of the ACM*, June, Vol. 30, Number 6, p. 484.
- OTA (Office of Technology Assessment) (1994) *Electronic Enterprises: Looking to the Future*. Washington D.C.: U.S. Government Printing Office, OTA-TCT-600, May.
- Schwarz, E.I. *Webnomics*. Broadway, New York, 1997.
- Subirana, B. (1998) "Transaction Streams and Value Added: Sustainable Business Models on the Internet". In Hitt, M.A.; Ricart, J.E. and Nixon R.D. (Eds.), *New Managerial Mindsets: Organizational and Strategy Implementation*.

Subirana, Brian and Carvajal, Patricia (1998a) "Mapping Insurance Transaction Streams". In Schmid, Beat F.; Selz, Dorian and Wittig, Dörte, "EM - Converging Media / EC in the Insurance Industry", *EM - Electronic Markets*, Vol. 8, No. 4, 12/98.

Subirana, Brian and Carvajal, Patricia (1998b) "Transaction Streams: Definitions and implications for trust in Internet-based electronic commerce", Research Paper No. 376, IESE.