Supernetworks: The Origins, Some Applications, and Possibilities

Anna Nagurney ▷
nagurney@gbfin.umass.edu

Anna Nagurney is the John F. Smith Memorial Professor in the Department of Finance and Operations Management in the Isenberg School of Management at the University of Massachusetts at Amherst. She is the first female to be appointed to a named Professorship in the University of Massachusetts system. She is the Founding Director of the Virtual Center for Supernetworks and the Supernetworks Laboratory for Computation and Visualization at UMass Amherst. She received her AB, ScB, ScM, and PhD degrees from Brown University in Providence, Rhode Island. She has published more than 125 papers and eight books. Her most recent book is *Supply Chain Network Economics: Dynamics of Prices, Flows, and Profits* [Edward Elgar Publishing, 2006].

In an invited essay, *Navigating the Network Economy* [*OR/MS Today*, June 2000 ▷], I argued that we were in a new era of *Supernetworks*. Since that time the world has been transformed through events of historical proportions which have dramatically and vividly reinforced the dependence of our societies and economies on different networks from *physical* networks; i.e., transportation and logistical networks, communication networks, energy and power networks, to more *abstract* networks comprising: financial networks, environmental networks, social, and knowledge networks, and combinations thereof. Indeed, historical events over the past several years have graphically illustrated the interconnectedness, interdependence, and vulnerability of organizations, businesses, and other enterprises on network *infrastructure* systems. The decisions made by the users of the networks, in turn, affect not only the users themselves but others, as well, in terms of safety and security, profits and costs, timeliness of deliveries, the quality of the environment, etc. Hence, the understanding of the impacts of human decision-making on such networks is of paramount importance.

In this essay, I argue that Supernetworks are the *paradigm* for the modeling, analysis, and solution of complex problems in the information-based *Network Economy*. In particular, the supernetwork paradigm, as evidenced by my book [20], along with many articles and applications (see: http://supernet.som.umass.edu), is sufficiently general and yet elegantly compact to formalize the modeling and analysis associated with network systems on which humans interact. ***Super*** networks are

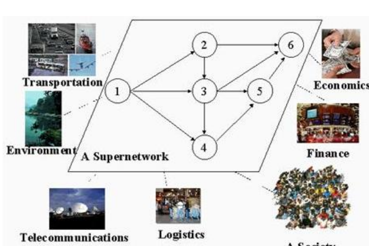


Figure SN-1. A supernetwork (▷ Larger picture)

networks that are **above and beyond** existing networks, which consist of nodes, links, and flows, with nodes corresponding to locations in space, links to connections in the form of roads, cables, etc., and flows to vehicles, data, etc. Supernetworks are conceptual in scope, graphical in perspective, and, with the accompanying theory, which is networked-based and predictive in nature.

The supernetwork framework, captures, in a unified fashion, decision-making facing a variety of decision-makers including consumers and producers as well as distinct intermediaries, such as financial brokers, electric power distributors, and electronic retailers. The decision-making process may entail weighting trade-offs associated with the use of transportation versus telecommunication networks. The behavior of the individual decision-makers is modeled as well as their interactions on the complex network systems with the goal of identifying the resulting flows and prices. By being able to predict the various flows based on network topologies and interactions amongst the decision-makers one gains deep insights into the vulnerabilities as well as the strengths of various linkages and network structures.

The Origins of Supernetworks

The origins of supernetworks can be traced to the study of transportation networks, telecommunication networks, and, interestingly, to biology, as reviewed in [20]. Below I highlight the origins of the term *supernetwork*.

In Transportation

In 1972, Dafermos [6] demonstrated, through a formal model, how a *multiclass* traffic network could be cast into a single-class traffic network through the construction of an expanded (and abstract) network consisting of as many copies of the original network as there were classes. She clearly identified the origin/destination pairs, demands, link costs, and flows on the abstract network. The applications of such networks she stated, "arise not only in street networks where vehicles of different types share the same roads (e.g., trucks and passenger cars) but also in other types of transportation networks (e. g., telephone networks)." Hence, she not only recognized that abstract networks could be used to handle multimodal transportation networks but also telecommunication networks! Moreover, she considered both user-optimizing and system-optimizing behavior, terms which she had coined with Sparrow [8] in 1969 (and which correspond, respectively, to Wardrop's (1952) first and second principles of travel behavior [37]). Her research was motivated, in part, by Beckmann, McGuire, and Winsten's 1956 book, *Studies in the Economics of Transportation* (see also [1]). In 1976, Dafermos [7] proposed an integrated traffic network equilibrium model in which one could visualize and formalize the entire transportation planning process (consisting of origin selection, or destination selection, or both, in addition to route selection, in an optimal fashion) as path choices over an appropriately constructed abstract network. The genesis and formal treatment of decisions more complex than route choices as path choices on abstract networks, that is, supernetworks, were hence reported as early as 1972 and 1976.

The importance and wider relevance of such abstract networks in decision-making, with a focus on transportation planning were accentuated through the term "hypernetwork" used by Sheffi [34], which he later [35] redefined as a "supernetwork." He recognized Dafermos' 1976 contributions and considered probabilistic-choice models. Thus, decision-making in a transportation context could be modeled as a "route" selection over an abstract network. The route, henceforth, referred to as a "path" to emphasize the generality of the concept, would correspond to a choice and the links to parts and pieces of the complete decision.

In Telecommunications

In his 1985 *American Scientist* article, Denning [9] discussed the internal structure of computer networks and emphasized how "protocol software can be built as a series of layers. Most of this structure is hidden from the users of the network." He then proceeded to ask the question, "What should the users see?" Denning answered the question in the context of the then National Science Foundation's Advanced Scientific Computing Initiative to make national supercomputer centers accessible to the entire scientific community. He said that such a system would be a network of networks, that is, a "supernetwork," and a powerful tool for science. Interestingly, he emphasized the importance of location-independent naming, so that if a physical location of a resource would change, none of the supporting programs or files would need to be edited or recompiled. Hence, in a sense, his view of supernetworks is in concert with that of ours in that nodes do not need to correspond to locations in space and may have an abstract association.

In 1979, Schubert, Goebel, and Cercone [33] had used the term in the context of knowledge representation as follows: "In the network approach to knowledge representation, concepts are represented as nodes in a network. Networks are compositional: a node in a network can be some other network, and the same subnetwork can be a subnetwork of several larger supernetworks. . . "

In 1997, the Illinois Bar Association considered the following to be an accepted definition of the Internet [11]: "the Internet is a supernetwork of computers that links together individual computers and computer networks located at academic, commercial, government and military sites worldwide, generally by ordinary local telephone lines and long-distance transmission facilities. Communications between computers or individual networks on the Internet are achieved through the use of standard, nonproprietary protocols." The reference to the Internet as a supernetwork was also made in The *Atlantic Monthly* in 1996 by Fallows [10], who noted that "The Internet is the supernetwork that links computer networks around the world."

Vinton G. Cerf, the co-developer of the computer networking protocol TCP/IP, in his keynote address to the Internet/Telecom 95 Conference [36], noted that at that time there were an estimated 23 million users of the Internet, and that vast quantities of the US Internet traffic "pass through internet MCI's backbone." He then went on to say, "Just a few months back, MCI rolled out a supernetwork for the National Science Foundation known as the very broadband network service or VNBS . . . VBNS is being used as an experimental platform for developing new national networking applications."

Decision-making on transportation and telecommunication networks can be done simultaneously through the supernetwork concept. For example, as demonstrated in [20], supply chain networks with electronic commerce, financial networks with intermediation, teleshopping versus shopping, telecommuting versus commuting, as well as transportation and location decisions in the Information Age formulated and solved within the supernetwork theoretical umbrella.

A variety of abstract networks in economics were studied in my 1999 book [16], which also contains extensive references to the subject. In [20] we have demonstrated that the abstract network concept also captures the interactions between/among the underlying networks of economies and societies. As noted in [17]: "The interactions among transportation networks, telecommunication networks, as well as financial networks is creating supernetworks. . . "

In Genetics

Interestingly, the term supernetworks has also been applied in biology, notably, in genetics. According to Noveen, Hartenstein, and Chuong [30], many interacting genes give rise to a gene network, with many interacting gene networks giving rise to a gene "supernetwork." They further state: "The function of a gene supernetwork is more complicated than a gene network. A gene supernetwork, for example, may be involved in determining the development of an entire limb while a gene network, working within the supernetwork, may be involved in setting up one of the axes of the limb bud." According to the same source, a gene supernetwork is defined as "a collection of gene networks which participate with each other during the morphogenesis of a specific structure, for example an organ, a segment, or an appendage." The authors then go on to discuss duplication, divergence, and conservation of a gene supernetwork and note that, as with gene networks, gene supernetworks can be duplicated during evolution, "thus giving rise to new structures which are the same as or different from the original structure."

Clearly, one of the principal facets of network systems today is the interaction among the networks themselves. For example, the increasing use of electronic commerce, especially in business to business transactions, is changing not only the utilization and structure of the underlying logistical networks but is also revolutionizing how business itself is transacted and the structure of firms and industries. Cellular phones are being used as vehicles move dynamically over transportation networks resulting in dynamic evolutions of the topologies themselves. Power outages in one part of the world may affect transportation and financial systems around the globe as the August 14, 2003 blackout demonstrated. The unifying concept of supernetworks with associated methodologies (optimization theory, network theory, variational inequalities, projected dynamical systems, etc.) allows one to explore the interactions among such networks as transportation networks, telecommunication networks, as well as financial networks, to capture the dynamic interactions and also to measure the associated risks and gains/losses.

Supernetworks and Applications

Supernetworks may be comprised of such networks as transportation, telecommunication, logistical and financial networks, among others. They may be multilevel as when they formalize the study of supply chain networks or multitiered as in the case of financial networks with intermediation. Furthermore, decision-makers on supernetworks may be faced with multiple criteria and, hence, the study of supernetworks also includes the study of multicriteria decision-making. In the Table below, some specific applications of supernetworks are given, for which results have been published in the literature. Subsequently, I elaborate upon several of the applications. For publications and additional references, see http://supernet.som.umass.edu

Examples of Supernetwork Applications
Telecommuting/Commuting Decision-Making
Teleshopping/Shopping Decision-Making
Supply Chain Networks with Electronic Commerce
Financial Networks with Electronic Transactions
Environmental and Energy Networks
Knowledge and Social Networks
Integrated Social and Supply Chain Networks
Electric Power Supply Chains and Transportation Networks

In particular, the supernetwork framework allows one to formalize the alternatives available to decision-makers, to model their individual behavior, typically, characterized by particular criteria which they wish to optimize, and to, ultimately, compute the flows on the supernetwork, which may consist of product shipments, the number of travelers between origins and destinations, the volumes of financial flows, energy flows, as well as the associated "prices." Hence, the concern is with human *decision-making* and how the supernetwork concept can be utilized to crystallize and inform in this dimension.

Supply Chain Networks and Electronic Commerce

The study of supply chain network problems through modeling, analysis, and computation is a challenging topic due to the complexity of the relationships among the various decision-makers, such as suppliers, manufacturers, distributors, and retailers as well as the practical importance of the topic for the efficient movement of products. The topic is multidisciplinary by nature since it involves particulars of manufacturing, transportation and logistics, retailing/marketing, as well as economics. In today's world, there is growing uncertainty and risk due to various threats and even illnesses which have affected dramatically the timely delivery of goods and have impacted transportation of humans as well.

The introduction of electronic commerce has, however, unveiled new opportunities in terms of research and practice in supply chain analysis and management since electronic commerce (e-commerce) has had a huge effect on the manner in which businesses order goods and have them transported with the major portion of e-commerce transactions being in the form of business-to-business (B2B). Estimates of B2B electronic commerce ranged from approximately 0.1 trillion dollars to 1 trillion dollars in 1998 and with forecasts reaching as high as $4.8 trillion dollars in 2003 in the United States. It has been emphasized that the principal effect of business-to-business (B2B) commerce, estimated to be 90% of all e-commerce by value and volume, is in the creation of new and more profitable supply chain networks.

In Figure SN-2 I depict a four-tiered supply chain network in which the top tier consists of suppliers of inputs into the production processes used by the manufacturing firms (the second tier), who, in turn, transform the inputs into products which are then shipped to the third tier of decision-makers, the retailers, from whom the consumers can then obtain the products. Here we allow not only for physical transactions to take place but also for virtual transactions, in the form of electronic transactions via the Internet to represent electronic commerce. In the supernetwork framework, both B2B and B2C can be considered, modeled, and analyzed. The decision-makers may compete independently across a given tier of nodes of the network and cooperate between tiers of nodes. In particular, my colleagues and I in a 2002 article in *Netnomics* applied the supernetwork framework to supply chain networks with electronic commerce in order to predict product flows between tiers of decision-makers as well as the prices associated with the different tiers. We assumed that the manufacturers as well as the retailers are engaged in profit-maximizing behavior whereas the consumers seek to minimize the costs associated with their purchases. The model therein determines the volumes of the products transacted electronically or physically.

As mentioned earlier, supernetworks may also be multilevel in structure. In [21], we demonstrated how supply chain networks can be depicted and studied as multilevel networks in order to identify not only the product shipments but also the financial flows as well as the informational ones. In Figure SN-3, I illustrate how a supply chain can be depicted as a multilevel supernetwork in which the financial network as well as the actual physical transportation network are also represented.

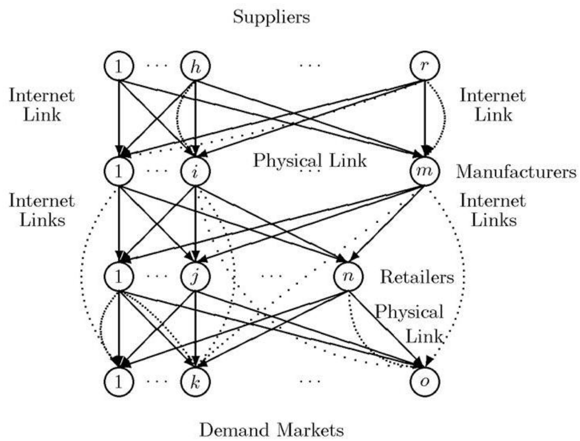


Figure SN-2. A Supply Chain Network (▷ Larger picture)

For example, in the supernetwork depicted in Figure SN-3, the logistical network affects the flows on the actual transportation network whereas the financial flows are due to payments as they proceed up the chain and as the transactions are completed. The information flows, in turn, are in the form of demand, cost, and flow data at the instance in time.

Obviously, in the setting of supply chain networks and, in particular, in global supply chains, there may be much risk and uncertainty associated with the underlying functions. Some research along these lines has already yielded promising results [19]. Continuing efforts to include uncertainty and risk into modeling and computational efforts in a variety of supernetworks and their applications is of paramount importance given the present economic and political climate.

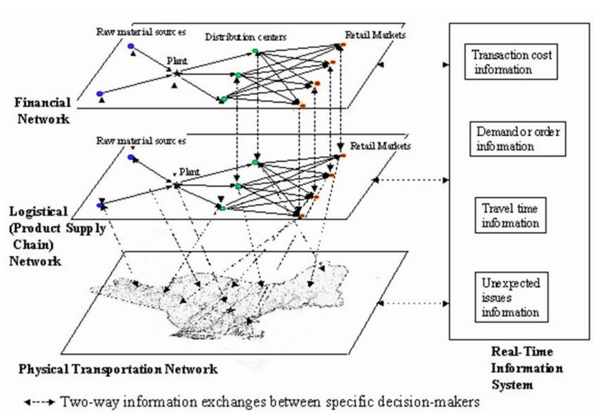


Figure SN-3. Supply Chain-Transportation Supernetwork Representation (▷ Larger picture)

In addition, I emphasize that the inclusion of environmental variables and criteria is also an important topic for research and practice in the context of supply chain networks, as has been demonstrated recently by Nagurney and Toyasaki [28] and is being presently investigated by my group in the context of electric power networks [38] and is generating much interest internationally.

Financial Networks and Electronic Transactions

Financial networks have been utilized in the study of financial systems since the work of Quesnay [31] in 1758, who depicted the circular flow of funds in an economy as a network. His conceptualization of the funds as a network, which was abstract, is the first identifiable instance of a supernetwork.

Quesnay's basic idea was subsequently applied in the construction of flow of funds accounts, which are a statistical description of the flows of money and credit in an economy. However, since the flow of funds accounts are in matrix form, and, hence, two-dimensional, they fail to capture the behavior on a micro level of the various financial agents/sectors in an economy, such as banks, households, insurance companies, etc. Moreover, the generality of the matrix tends to obscure certain structural aspects of the financial system that are of continuing interest in analysis, with the structural concepts of concern including those of financial intermediation.

Advances in telecommunications and, in particular, the adoption of the Internet by businesses, consumers, and financial institutions have had an enormous effect on financial services and the options available for financial transactions. Distribution channels have been transformed, new types of services and products introduced, and the role of financial intermediaries altered in the new supernetworked landscape. Furthermore, the impact of such advances has not been limited to individual nations but, rather, through new linkages, has crossed national boundaries.

The topic of *electronic* finance has been a growing area of study, as described in [18]. This is due to its increasing impact on financial markets and financial intermediation, and the related regulatory issues and governance. Of particular emphasis has been the conceptualization of the major issues involved and the role of networks is the transformations.

Now, I briefly describe a supernetwork framework for the study of financial decision-making in the presence of intermediation and electronic transactions. Further details can be found in [22,23]. The framework is sufficiently general to allow for the modeling, analysis, and computation of solutions to such problems.

The financial network model consists of: agents or decision-makers with sources of funds, financial intermediaries, as well as consumers associated with the demand markets. In the model, the sources of funds can transact directly electronically with the consumers through the Internet and can also conduct their financial transactions with the intermediaries either physically or electronically. The intermediaries, in turn, can transact with the consumers either physically in the standard manner or electronically. The depiction of the network at equilibrium is given in Figure SN-4.

It is assumed that the agents with sources of funds as well as the financial intermediaries seek to maximize their net revenue (in the presence of transaction costs) while, at the same time, minimizing the risk associated with the financial products. The solution of the model yields the financial flows between the tiers as well as the prices. We also allow for the option of having the source agents not invest a part (or all) of their financial holdings.

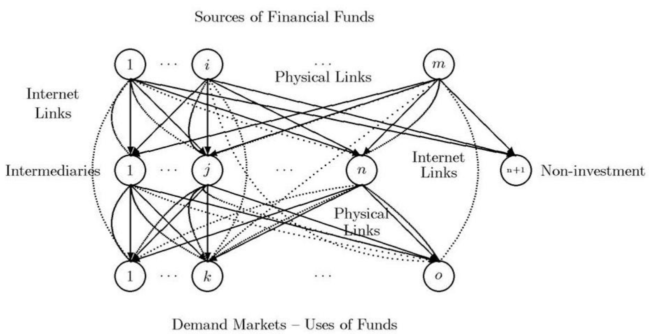


Figure SN-4. Financial Network (▷ Larger picture)

The Supernetwork Structure Reveals Answers to Questions Dating Back Half a Century

More recently, Liu and Nagurney [14] demonstrated that the supernetwork framework can also be used to show that financial equilibrium problems with intermediation can be reformulated and solved as transportation network equilibrium problems. Similarly, Wu et al. [38] proved that, as hypothesized in Chapter 5 of [1], electric power distribution networks can be reformulated and solved as transportation network equilibrium problems over an appropriately constructed abstract network or supernetwork. Copeland [5] asked whether "money flows like water or electricity?" In [14] we established that money flows like transportation flows and in [38] that electric power flows like transportation flows. Such reformulations of financial network problems and electric power supply chains have yielded, through the supernetwork concept, new interpretations of the governing equilibrium conditions in terms of path flows and associated costs. Because of such reformulations we now can apply additional computational procedures that yield path flow information (for a recent large-scale empirical application [15]).

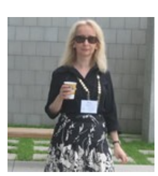
The Possibilities

In this essay, I have argued for the Supernetwork paradigm as a powerful tool for the study of network systems, emphasizing that it can capture not only the interrelationships among networks but, most importantly, the effects of human decision-making on the induced flows and prices. Through the computation of the flows and prices one can determine the optimal/equilibrium network designs and structures as well as the associated vulnerabilities. Hence, supernetworks provide not only powerful engineering and operations research/management science tools but also financial and economic ones. Finally, the supernetwork paradigm uniquely captures the human aspects and brings a richness to the conceptualization and the understanding of the underlying processes.

For example, the complex network literature [29] initiated principally by physicists, is only about a decade old, whereas the publications in OR/MS on networks date back a half century. An important aspect of the complex network physics literature concerns network efficiency measurement and vulnerability analysis. By showing that it is not just the topology that matters, but, also the associated costs and flows on links subject to the decision-makers' behavior, we have been able to generalize the results of Latora and Marchiori [12, 13]. In particular, we have been able to show the importance and the rankings of network components, i.e., the nodes and links in a more coherent and reasonable manner [26, 27].

In addition, the connections that have been established through the supernetwork paradigm of various applications are now being exploited to address the dynamics of such network systems, leading not only to new theories associated with equilibria, but with disequilibria, as well as time-dependent equilibria [14, 25]. For example, the well-known Braess paradox [2] (see also [3]), which has been the subject of much attention recently by the computer scientists [32] has been extended to a time-dependent version in which it was demonstrated that the paradox occurs only for a certain range of demands and, after a certain demand the new route is never used! Hence, if one considers network design issues in either telecommunication networks or in transportation networks, operating in a user-optimizing manner, then the addition of a new road/link may lead to increased user costs within a range of demands and the new route may not even be used past a certain demand; suggesting that the new link, which induces a new route, should not have been built since, over time, we may expect an increase in demand.

Supernetworks have, thus far, enabled the identification of similar structures and relationships between financial networks and transportation networks; electric power networks and transportation networks; supply chains and transportation networks, as well as transportation and telecommunication networks. Supernetworks have also been used to extend network-based results in other disciplines, including physics. Future research, I suspect, will include new concepts for dynamics, besides the recently formulated double-layered dynamics, new research into contagion, as well as into resiliency and robustness of network systems.



Anna takes a break at the Puerto Rico meeting.

References

[1] D. Boyce, H.S. Mahmassani, and A. Nagurney, A Retrospective on Beckmann, McGuire, and Winsten's *Studies in the Economics of Transportation*, *Papers in Regional Science* 84:1 (2005), 85–103.

[2] D. Braess, Uber ein Paradoxon der Verkehrsplanung, *Unternehmenforschung* 12 (1968), 258–268.

[3] D. Braess, A. Nagurney, and T. Wakolbinger, On a Paradox of Traffic Planning, Translation of the (1968) Original D. Braess Paper from German to English, *Transportation Science* 39:4 (2005), 446–450.

[4] M.-J. Cojocaru, P. Daniele, and A. Nagurney, Double-Layered Dynamics: A Unified Theory of Projected Dynamical Systems and Evolutionary Variational Inequalities, *European Journal of Operational Research* 175:1 (2006), 494–507.

[5] M.A. Copeland, *A Study of Moneyflows in the United States*, National Bureau of Economic Research, New York, 1952.

[6] S. Dafermos, The Traffic Assignment Problem for Multimodal Networks, *Transportation Science* 6:1 (1972), 73–87.

[7] S. Dafermos, Integrated Equilibrium Flow Models for Transportation Planning, in M.A. Florian (Ed.), *Traffic Equilibrium Methods, Lecture Notes in Economics and Mathematical Systems* 118, Springer-Verlag, NY, 1976, 106–118.

[8] S. Dafermos and F.T. Sparrow, The Traffic Assignment Problem for a General Network, *Journal of Research of the National Bureau of Standards* 73B (1969), 91–118.

[9] P.J. Denning, The Science of Computing: Supernetworks, *American Scientist* 73 (1985), 127–129.

[10] J. Fallows, The Java Theory, *The Atlantic Monthly* 277:3 (1996), 113–117.

[11] Illinois State Bar Association, Opinion No. 96–10, Topic: Electronic Communications; Confidentiality of Client Information; Advertising and Solicitation, May 16, 1997.

[12] V. Latora and M. Marchiori, Efficient Behavior of Small-World Networks, *Physical Review Letters* 87:19 (2001), Article No. 198701 (4 pp).

[13] V. Latora and M. Marchiori, Economic Small-World Behavior in Weighted Networks, *European Physical Journal B* 32:2 (2003), 249–263.

[14] Z. Liu and A. Nagurney, Financial Networks with Intermediation and Transportation Network Equilibria: A Supernetwork Equivalence and Reinterpretation of the Equilibrium Conditions with Computations, *Computational Management Science* 4:3 (2007), 243–281.

[15] Z. Liu and A. Nagurney, *An Integrated Electric Power Supply Chain and Fuel Market Network Framework: Theoretical Modeling with Empirical Analysis for New England*, Isenberg School of Management, University of Massachusetts, Amherst, MA, 2007.

[16] A. Nagurney, *Network Economics: A Variational Inequality Approach*, 2nd edition, Kluwer Academic Publishers, The Netherlands, 1999.

[17] A. Nagurney, Navigating the Network Economy, *OR/MS Today* 27:3 (2000) 74–75.

[18] A. Nagurney, Editor, *Innovations in Financial and Economic Networks*, Edward Elgar Publishing, UK, 2003.

[19] A. Nagurney, J. Cruz, J. Dong, and D. Zhang, Supply Chain Networks, Electronic Commerce, and Supply Side and Demand Side Risk, *European Journal of Operational Research* 164:1 (2005), 120–142.

[20] A. Nagurney and J. Dong, *Supernetworks: Decision-Making for the Information Age*, Edward Elgar Publishing, UK, 2002.

[21] A. Nagurney, K. Ke, J. Cruz, K. Hancock, and F. Southworth, Dynamics of Supply Chains: A Multilevel (Logistical/Informational/Financial) Network Perspective, *Environment and Planning B* 29:6 (2002), 795–818.

[22] A. Nagurney and K. Ke, Financial Networks with Intermediation, *Quantitative Finance* 1:4 (2001), 441–451.

[23] A. Nagurney and K. Ke, Financial Networks with Electronic Transactions: Modeling, Analysis, and Computations, *Quantitative Finance* 3:2 (2003), 71–87.

[24] A. Nagurney, J. Loo, J. Dong, and D. Zhang, D., Supply Chain Networks and Electronic Commerce: A Theoretical Perspective, *Netnomics* 4:2 (2002), 187–220.

[25] A. Nagurney, D. Parkes, and P. Daniele, The Internet, Evolutionary Variational Inequalities, and the Time-Dependent Braess Paradox, *Computational Management Science* 4:4 (2007), 355–375.

[26] A. Nagurney and Q. Qiang, A Network Efficiency Measure for Congested Networks, *Europhysics Letters* 79:3 (2007), Article 38005 (5 pp).

[27] A. Nagurney and Q. Qiang, A Network Efficiency Measure with Application to Critical Infrastructure Networks, *Journal of Global Optimization*, in press.

[28] A. Nagurney and F. Toyasaki, Reverse Supply Chain Management and Electronic Waste Recycling: A Multitiered Network Equilibrium Framework for E-Cycling, *Transportation Research E*, 41:1 (2005), 1–28.

[29] M. Newman, The Structure and Function of Complex Networks, *SIAM Review* 45:2 (2003), 167–256.

[30] A. Noveen, V. Hartenstein, C.-M. Chuong, Gene Networks and Supernetworks: Evolutionarily Conserved Gene Interactions, in C.-M. Chuong (Ed.), *Molecular Basis of Epithelial Appendage Morphogenesis*, Landes Bioscience, Austin, TX, 1998, pp. 371–391.

[31] F. Quesnay, *Tableau Économique*, 1758. Reproduced in Facsimile with an Introduction by H. Higgs by the British Economic Society, 1895.

[32] T. Roughgarden, *Selfish Routing and the Price of Anarchy*, MIT Press, Cambridge, MA, 2005.

[33] L. Schubert, R. Goebel, and N. Cercone, The Structure and Organization of a Semantic Net for Comprehension and Inference, in N.V. Findler (Ed.), *Associative Networks: Representation and Use of Knowledge by Computers*, Academic Press, New York, NY, 1979.

[34] Y. Sheffi, *Transportation Network Equilibrium with Discrete Choice Models*, Ph. D. Thesis, Civil Engineering Department, Massachusetts Institute of Technology, Cambridge, MA, 1978.

[35] Y. Sheffi, *Urban Transportation Networks —- Equilibrium Analysis with Mathematical Programming Methods*, Prentice-Hall, Englewood Cliffs, NJ, 1985.

[36] Telecom 95 (1995), MCI and the Internet: Dr. Vinton G. Cerf, Keynote Address, http://www.itu.int/TELECOM/wt95/pressdocs/papers/cerf.html.

[37] J.G. Wardrop, Some Theoretical Aspects of Road Traffic Research, *Proceedings of the Institute of Civil Engineers, Part II*, 1952, 325–378.

[38] K. Wu, A. Nagurney, Z. Liu, and J.K. Stranlund, Modeling Generator Power Plant Portfolios and Pollution Taxes in Electric Power Supply Chain Networks: A Transportation Network Equilibrium Transformation, *Transportation Research Part D*, 11:3 (2006), 171–190.

Data Mining in Industrial Applications and Innovation

Andrew Kusiak ▹
andrew-kusiak@uiowa.edu

Andrew Kusiak is Professor of Mechanical and Industrial Engineering at The University of Iowa, and the Research Coordinator of their Intelligent Systems Laboratory. He has championed *Innovation Science* in the manufacturing environment, having just published a primer in *International Journal of Computer Applications in Technology* 28:2-3, 2007, 140–149. Dr. Kusiak serves on journal editorial boards, edits book series, and is the Editor-in-Chief of the *Journal of Intelligent Manufacturing*. He is also the Past Chair of the INFORMS Data Mining Section.

About Data Mining

Data mining [23] has more than a decade-old history, yet it remains a mystery to many. There is a range of intuitive opinions of data mining, ranging from considering it a branch of statistics to a somewhat deterministic tool. These views are in essence correct as data mining has different shades by encompassing algorithms that are deeply versed in statistics, computational intelligence, and those resembling deterministic approaches.

How Different is Data Mining from Traditional OR/MS Methods?

If statistics is your basic tool of interest, it is worthy remembering that statistical models describe populations of objects. Data mining models could describe a population as well; however, the population-based field might be too crowded for data mining to enter. Data mining has a better chance to distinguish itself as an individual-based approach rather than a population-based science.

Statistical models, e.g., regression, use all parameters specified by the user. It is true that not all of them are equally significant and some may not be incorporated in the model. However, while using the model all independent parameter values need to be provided to determine the decision (the value of the dependent variable). A decision tree or decision rules extracted with a decision-tree or a decision-rule algorithm do not require all values of independent variables for making decisions. Data mining algorithms usually have the ability to select important parameters; however, an entire branch of data mining known as feature selection deals in refined ways of feature (parameter) selection.

While statistical models largely evolve around independent and statistically significant variables, data mining algorithms can make good use of insignificant and even correlated variables. Combinations of insignificant variables often turn into powerful decision-making models.

For those who enjoy using operations research, data mining brings a fresh perspective. An operations research analyst carefully studies the problem on hand and depending on her/his background recommends a particular model and an algorithm. The model could be deterministic or stochastic. By recommending (fitting) a model or a collection of models the OR analyst fixes the number of variables, e.g., a transportation model uses specific variables, constraints, and the objective function. It is true that the selected model can be modified by adding new variables, constraints, and so on. Often the model is built from scratch; however, when fully developed, usually the structure resembling the traditional OR models can be identified. As statisticians enjoy curve fitting, operations researchers fit models or their elements and we call it modeling. In some cases, e.g., the formal modeling phase is replaced with algorithm-fitting or its development.

Another issue deserving attention is that operations research deals almost exclusively with quantitative variables. The discussion on combining quantitative and qualitative variables in operations research modeling has been lasting decades without much success. Data mining might have solved the quantitative and qualitative variable dilemma of operations research. Many data mining algorithms handle with great ease strings of numbers and qualitative values. In fact, an entire sentence could be one of many variables considered in data mining. Data mining allows casting a wide net over all possible quantitative and qualitative variables. In addition, unlike operations research, where modeling is in essence top-down (model fitting), in data mining a bottom-up approach is followed. A typical data mining algorithm derives a model from the data rather than trying to fit a preconceived model. One could say that a traditional modeler comes with luggage (e.g., dependent on previous training and experience), while a data miner does not. A machine learning algorithm extracts the model.

Where Data Mining Works Best?

A data mining algorithm does only a partial job, i.e., it creates a model. In cases when a user is looking for patterns the results produced by the data mining algorithm may suffice. This type of data mining is called *descriptive*. The latter