On the Regulation of Networks as Complex Systems: A Graph Theory Approach

Daniel F. Spulber
Northwestern University

Christopher S. Yoo
University of Pennsylvania Carey Law School

Author ORCID Identifier:

Christopher Yoo 0000-0003-2980-9420

Copyright 2005 by Northwestern University, School of Law
Northwestern University Law Review

Printed in U.S.A.
Vol. 99, No. 4

Essay

ON THE REGULATION OF NETWORKS AS COMPLEX SYSTEMS: A GRAPH THEORY APPROACH

Daniel F. Spulber[*] & *Christopher S. Yoo*[**]

I. INTRODUCTION

One of the most salient economic developments of the last decade has been the transformation of the telecommunications industry. Competition

---

[*] Elinor Hobbs Distinguished Professor of International Business, Kellogg School of Management, and Professor of Law, Northwestern University School of Law.

[**] Professor of Law, Vanderbilt University Law School. We would like to thank Paul Edelman, Chris Guthrie, Todd Henderson, Lou Kaplow, Vik Khanna, Jeff MacKie-Mason, Richard Nagareda, Randy Picker, Bob Rasmussen, Sang-Seung Yi, Kiho Yoon and participants in a workshop conducted at the 15th Annual Meeting of the American Law and Economics Association, the University of Michigan Law School Workshop on Cyberlaw and Economics, and the Seoul National University Center for Law and Technology for comments on earlier drafts of this Essay. We plan to offer a more extended and formal treatment of this subject in a book entitled *Networks in Telecommunications: Economics and Law*, which is forthcoming on Cambridge University Press.

among conventional telephone technologies has steadily increased, as technological improvements have caused the fixed costs associated with many of the functions performed by conventional telephone providers to decrease sharply.[1] At the same time, existing network technologies have faced increasing competitive pressure from new transmission media. By the end of 2004, cellular telephones and more advanced wireless technologies, such as personal communications services ("PCS"), finally surpassed conventional wireline telephony as the leading platform for providing voice communications.[2] Cable modem systems, digital subscriber lines ("DSL"), and satellite broadband systems have also emerged as important means of communications, with a host of other broadband technologies waiting in the wings.[3] In addition, the shift toward packet-switched technologies made possible by the digital revolution has made different transmission technologies increasingly interchangeable. The impending arrival of Internet telephony, often called voice over Internet protocol ("VoIP"), represents the most prominent illustration of this phenomenon. Together these changes have rendered a sector that has long been dominated by regulated monopolies more competitive than ever before.[4]

This fundamental transformation in industry structure has been accompanied by a parallel (if somewhat incomplete) transformation in the way that communications networks are regulated. Long regarded by regulators as natural monopolies, telecommunications networks have been governed traditionally by a system of rate regulation that required network owners to provide service to all interested customers on nondiscriminatory and reasonable terms. Because rate regulation was targeted at the output of the en-

---

[1] For example, the development of microwave relay systems during the 1960s lowered fixed costs in long-distance telephone service to the point where competition became feasible. *See* Glen O. Robinson, *The Titanic Remembered: AT&T and the Changing World of Telecommunications*, 5 YALE J. ON REG. 517, 522–26 (1988) (reviewing GERALD R. FAULHABER, TELECOMMUNICATIONS IN TURMOIL: TECHNOLOGY AND PUBLIC POLICY (1987)). Similarly, the reduction in switching costs made possible by the dramatic improvements in computer processing capability has made local telephone service increasingly competitive. *See* United States Telecom Ass'n v. FCC, 359 F.3d 554, 568–71, 586–87 (D.C. Cir. 2004).

[2] *See* FCC INDUSTRY ANALYSIS TECHNOLOGY DIVISION, WIRELINE COMPETITION BUREAU, LOCAL TELEPHONE COMPETITION: STATUS AS OF DECEMBER 31, 2004, at 1, 3, 5 tbl.1, 17 tbl.13 (July 2005) (reporting 178 million wireline and 181 million wireless subscribers as of the end of 2004), *available at* http://www.fcc.gov/Bureaus/Common_Carrier/Reports/FCC-State_Link/IAD/lcom0705.pdf; Implementation of Section 6002(b) of the Omnibus Budget Reconciliation Act of 1993: Annual Report and Analysis of Competitive Market Conditions with Respect to Commercial Mobile Services, Eighth Report, 18 F.C.C.R. 14,783, 14,831–34 ¶¶ 101–106 (2003) [hereinafter Eighth Report on CMRS Competition] (detailing the increase in wireless-wireline competition).

[3] These include advanced mobile telephone technologies called third-generation wireless ("3G") that are capable of delivering Internet service, broadband provided via electric power lines, the widescale deployment of fiber optic cable to homes and businesses, and innovative unlicensed spectrum-based technologies such as ultra wideband and WiFi mesh networks.

[4] *See generally* Daniel F. Spulber, *Deregulating Telecommunications*, 12 YALE J. ON REG. 25 (1995).

tire network, regulatory authorities did not need an overarching theory of how networks are configured or how individual network components interact with one another. They only needed to ensure that the total revenues were sufficient to cover the total network cost.

The emergence of competition has forced regulatory authorities to abandon their traditional reliance on rate regulation in favor of a new approach known as access regulation. Unlike rate regulation, which focuses on the terms under which network owners sell *outputs* to *consumers*, access regulation instead focuses on the terms under which network owners must lease key *inputs* to *competitors*. The leading example of access regulation is the Telecommunications Act of 1996, which requires certain local telephone companies to provide competitors with access to key elements of their networks.[5] Access regulation has also emerged as a dominant feature in the regulation of a wide range of other network facilities, including cable television systems, broadband technologies, networks of utility poles, natural gas pipelines, and electric power distribution grids.[6]

The shift from rate regulation to access regulation has fundamentally altered the primary unit of regulatory analysis. Rather than setting prices charged for the services of the entire network, regulators must now determine the prices charged for individual network components. In the process, this shift has created the need for a better understanding of the interrelationships among network elements. Unfortunately, rather than developing more sophisticated theories of network behavior, regulatory authorities have implemented access regimes by continuing to focus on the cost of each network component.[7] This has had the effect of ignoring the interactions among network components and, instead, treats each element as if it existed in a vacuum.

---

[5] 47 U.S.C. § 251(c)(3) (2004). The statute requires that the accessed elements be "necessary" and that "the failure to provide access to such network elements would impair the ability of the telecommunications carrier seeking access to provide the services that it seeks to offer." *Id.* § 251(d)(2)(A) & (B). For a review of the regulatory antecedents to § 251(c)(3), see Daniel F. Spulber & Christopher S. Yoo, *Access to Networks: Economic and Constitutional Connections*, 88 CORNELL L. REV. 885, 960–65, 1005–09 (2003).

[6] *See* Joseph D. Kearney & Thomas W. Merrill, *The Grand Transformation of Regulated Industries Law*, 98 COLUM. L. REV. 1323, 1340–46 (1998) (calling the shift toward unbundling requirements part of the "great transformation of regulated industries law" and tracing the emergence of access regulation in the telephone and natural gas industries); Spulber & Yoo, *supra* note 5, at 960–70, 981–87, 1003–18 (tracing the development of access requirements for local telephony, networks of utility poles, and broadband networks); Michael O. Wise, *Overview: Deregulation and Antitrust in the Electric Power Industry*, 64 ANTITRUST L.J. 267 (1996) (discussing the emergence of access requirements in electric power).

[7] *See, e.g.*, 47 U.S.C. § 252(d)(1)(A)(i) (requiring that access prices for unbundled network elements of local telephone systems be "based on the cost"); *id.* § 224(d)(1) (requiring that prices for access to networks of utility poles be based on "the additional costs of providing pole attachments"). The FCC has implemented this provision by requiring that access prices be based on the historical cost associated with the portions of the network accessed. 47 C.F.R. § 1.1404(h)(2) (2004) (providing that "[d]ata and information should be based upon historical or original cost methodology, insofar as possible").

The existing academic commentary has yet to fill this void. The profusion of scholarly writing on the economics of networks that has appeared over the last two decades has focused almost exclusively on the phenomenon known as "network economic effects," which arise when a network's value is determined by the number of other users connected to it.[8] Although this literature has yielded a number of cogent insights, its scope is limited by the fact that it focuses on only one aspect of networks: their size. As a result, it is unable to shed much light on the relative benefits of different network configurations or how individual components can interact with one another when integrated into a complex system.

We believe that it is imperative that the debates about network policy shift to terms that better capture the key aspects of network behavior. Such an approach should reflect the fact that the design decisions regarding a particular network component are influenced in no small part by the relationship of that component to the other network components. Networks are constructed with a view toward the performance and structure of the system as a whole and, thus, cannot be understood solely by examining individual components in isolation. In short, network theory must reflect the extent to which the whole exceeds the sum of the parts. In addition, networks evolve over time through the actions of network users, changes in types of access to the networks, and interconnections between networks. The type of systemic approach we envision would reflect how changes to one part of the network can affect the performance and evolution of the network in ways that can be dramatic and unpredictable.

In this Essay, we would like to propose a new conceptual framework that promises to place network policy on a more analytically sound foundation. To illustrate some of the main insights of our approach, we draw on a branch of mathematics known as "graph theory."[9] Graph theory began as

---

[8] The literature on network externalities is vast. For the seminal articles, see Joseph Farrell & Garth Saloner, *Standardization, Compatibility, and Innovation*, 16 RAND J. ECON. 70 (1985); Michael L. Katz & Carl Shapiro, *Network Externalities, Competition, and Compatibility*, 75 AM. ECON. REV. 424 (1985). For a useful survey appearing in the legal literature, see Mark A. Lemley & David McGowan, *Legal Implications of Network Economic Effects*, 86 CAL. L. REV. 480 (1998).

[9] For relatively accessible introductions to the subject, see JOHN M. ALDOUS & ROBIN J. WILSON, GRAPHS AND APPLICATIONS: AN INTRODUCTORY APPROACH (2000); JONATHAN GROSS & JAY YELLEN, GRAPH THEORY AND ITS APPLICATIONS (1999); W.D. WALLIS, A BEGINNER'S GUIDE TO GRAPH THEORY (2000); DOUGLAS B. WEST, INTRODUCTION TO GRAPH THEORY (1996). For more sophisticated presentations of the subject, see, for example, HANDBOOK OF GRAPH THEORY (Jonathan L. Gross & Jay Yellen eds., 2004); REINHARD DIESTEL, GRAPH THEORY (2d ed. 2000); DIETER JUNGNICKEL, GRAPHS NETWORKS AND ALGORITHMS (Tilla Schade trans., 1999). For a historical survey of graph theory, see NORMAN L. BIGGS ET AL., GRAPH THEORY 1736–1936 (1976). The rapidly developing science of networks is chronicled in a number of popular works. *See* ALBERT-LASZLO BARABÁSI, LINKED: THE NEW SCIENCE OF NETWORKS (2002); MARK BUCHANAN, NEXUS: SMALL WORLDS AND THE GROUNDBREAKING SCIENCE OF NETWORKS (2002); DUNCAN J. WATTS, SIX DEGREES: THE SCIENCE OF A CONNECTED AGE (2003). For a rudimentary application of graph theory to telecommunications policy, see William W. Sharkey, *Representation of Technology and Production*, *in* 1 HANDBOOK OF TELECOMMUNICATIONS ECONOMICS 180, 186–202 (Martin E. Cave et al. eds., 2002).

an academic discipline in 1736 with legendary mathematician Leonhard Euler. Euler studied whether it was possible to cross each of the seven bridges interconnecting the two banks of the Pregel River and the island of Kneiphof, located within the city of Königsberg, without crossing any bridge more than once.[10] Other classic problems in graph theory include the "traveling salesman problem," which attempts to determine the shortest or cheapest route that passes through a series of destinations exactly once before ending where the journey had begun,[11] and the "four color problem," which focuses on the minimum number of colors needed to color a map so that no adjacent countries have the same color.[12] Mathematicians, physicists, biologists, and sociologists have also employed graph theory to develop an elaborate science of networks used to model such varied interactive phenomena as chemical and nuclear reactions, biological processes, the spread of epidemics, the structure of ecosystems, and the formation of social networks.[13] More recently, graph theory has been employed by computer and communications-network designers seeking insights into the complexities of network behavior.[14]

Despite the analytical power that graph theory provides, its insights have yet to be applied to broad issues of network policy. The oversight is

---

[10] Leonhard Euler, *Solutio Problematis ad Geometriam Situs Pertinentis* [Solution of a Problem Relating to the Geometry of Position], 8 COMMENTARII ACADEMIAE SCIENTIARUM IMPERIALIS PETROPOLITANAE 128 (1736), *reprinted in* BIGGS ET AL., *supra* note 9, at 3–11. The proof is quite simple. Euler realized that every node except for the beginning and ending nodes of the path must necessarily have an even number of links leading away from it if the type of path that Euler sought were to exist. Because the network created by the Königsberg bridges contained four nodes with an odd number of links, no such path existed. *See* BARABÁSI, *supra* note 9, at 12.

[11] The traveling salesman problem is a subset of the series of problems inspired by William Rowan Hamilton. A "Hamiltonian path" is a path that crosses each node of a network without crossing any node more than once. A "Hamiltonian cycle" is a Hamiltonian path that begins and ends at the same node. The traveling salesman problem seeks out the shortest or least-cost Hamiltonian cycle. The traveling salesman problem plays a prominent role in a leading patent case. *See In re* Trovato, 42 F.3d 1376 (Fed. Cir. 1994) (holding that a computer system for solving the traveling salesman problem is unpatentable subject matter), *withdrawn on reh'g*, 60 F.3d 807 (1995) (en banc).

[12] *See* Kenneth Appel & Wolfgang Haken, *Every Planar Map Is Four Colorable: Part I: Discharging*, 21 ILL. J. MATH. 429 (1977); Kenneth Appel et al., *Every Planar Map Is Four Colorable: Part II: Reducibility*, 21 ILL. J. MATH. 491 (1977).

[13] *See, e.g.*, STEPHEN N. SHORE, THE TAPESTRY OF MODERN ASTROPHYSICS 305 (2002); OLEG N. TEMKIN ET AL., CHEMICAL REACTION NETWORKS: A GRAPH-THEORETICAL APPROACH (1996); H. Jeong et al., *The Large-Scale Organization of Metabolic Networks*, 407 NATURE 651 (2000); Peter Yodzis, *Diffuse Effects in Food Webs*, 81 ECOLOGY 261 (2000). Graph theory also underlies the assertion that every person in the planet is connected through no more than six degrees of separation. *See* Stanley Milgram, *The Small World Problem*, 1 PSYCHOL. TODAY 60 (1967) (finding that letters sent to random residents of Omaha, Nebraska, only needed to pass through an average six people before reaching a target person in Boston, Massachusetts). The concept of six degrees of separation has been reflected in popular culture in a play that was subsequently turned into a popular film, *see* JOHN GUARE, SIX DEGREES OF SEPARATION (1990); SIX DEGREES OF SEPARATION (1993), as well as a popular parlor game involving Kevin Bacon, *see, e.g.*, BARABÁSI, *supra* note 9, at 58–60.

[14] *See* THOMAS G. ROBERTAZZI, PLANNING TELECOMMUNICATION NETWORKS (1999).

regrettable since graph theory provides powerful analytical tools capable of addressing the central shortcoming of the current regulatory framework by reflecting how interactions among network components can cause systemic effects that cannot be understood solely by studying individual network elements in isolation. The architecture of a network fundamentally affects the network's ability to handle communications traffic. Network performance can be measured in terms of the volume of traffic the network can handle, the reliability of operating systems, the accuracy of information transmission, and the speed of transmission. A network's usage patterns, much like the traffic flows on city streets, can create congestion and affect the performance of the network. The study of networks using graph theory, therefore, can help regulators and policymakers recognize how networks function as complex systems.

In addition, the graph theoretical approach that we propose offers starkly different policy implications than does the "network economic effects" view, an alternative approach to creating a regulatory framework. By focusing only on network size, discussions based on network economic effects argue that interconnection is a *benefit* for the host network and that interconnection should be provided free of charge.[15] In contrast, by considering networks as complex systems, we support the well-established understanding that interconnection creates *costs* for the host network that can require intercarrier payments.

The balance of this Essay is organized as follows: Part II lays out the basic terminology employed by graph theory to describe networks. Part III analyzes the insights that graph theory provides into the economics underlying decisions about network configuration. Part IV employs a graph theoretic principle known as the "max-flow/min-cut theorem" to explore the impact that interactions among different network components can have on network capacity. Part V demonstrates the power of our proposed approach by outlining how it would apply to a number of leading issues confronting policymakers. These issues include the decision to apply the same methodology to set prices for interconnection and access to unbundled network elements under the Telecommunications Act of 1996, the decision to base access rates on cost, proposals to base access prices on market benchmarks, whether to compel access to broadband networks, and the regulation of Internet telephony. We also explore the extent to which the framework can

[15] *See* David A. Balto, *Networks and Exclusivity: Antitrust Analysis to Promote Network Competition*, 7 GEO. MASON L. REV. 523 (1999); Adam Candeub, *Network Interconnection and Takings*, 54 SYRACUSE L. REV. 369, 371 (2004) ("[T]he takings/intercarrier payment assumption is, in fact, faulty. Interconnection does *not* necessarily involve one company 'using' another's network in a manner requiring one network to pay the other. Indeed, such assertion is economically suspect because interconnection confers a benefit to both networks—that of a larger calling universe, i.e., network effects—which renders each network more valuable."). Advocates stop short of suggesting the host network should pay for the supposed benefits it receives from providing the interconnection, but instead suggest that only the incremental cost of establishing the interconnection should be covered.

be employed to support a formal calculation of regulated rates. Even this preliminary sketch is more than sufficient to establish the benefits of shifting to an approach that is better able to capture the extent to which networks operate as integrated systems.

II. NETWORKS AS COMPLEX SYSTEMS

One of the central impediments to progress in the regulatory analysis of networks has been the lack of an established nomenclature for describing and analyzing the essential features of network structure. We expand upon the principles of graph theory to offer a methodology for describing communications networks that captures the complex manner in which individual network components interact with one another.[16]

The basic units of analysis under our approach consist of *nodes* and *links*.[17] Nodes are the junctions that represent the critical points of origin, routing, and termination. In a conventional wireline telephone system, nodes tend to be physical locations at which one or more specialized pieces of equipment are installed. Examples include customer premises, where calls originate and terminate, and central offices, where telephone companies maintain switching equipment. That said, nodes need not be confined to specific, physical locations. For example, mobile phones constitute nodes of a wireless telephone system despite their portability. Nodes can even jump from one network to another. This can occur, for example, when mobile phones roam across different wireless providers or when laptop computers with wireless local area network ("LAN") cards move between WiFi access points.

Links are any type of connection between nodes. Links can be fixed in location, such as telephone or fiber-optic communication lines. Links need not represent specific geographic corridors, however. For example, a transmitter within a wireless telephone system may be in a specific geographic location, but the communication links with mobile phones are temporary and vary with the location of the phones.

---

[16] We note as a preliminary matter that our analysis focuses exclusively on the branch of graph theory dealing with what we will call "managed networks," in which a single actor makes all of the decisions with respect to the network. We forego discussion of what we call "spontaneous networks," in which the particular network structure is the result of the decisions by a large number of decentralized and uncoordinated participants. *See* Matthew O. Jackson, *A Survey of Models of Network Formation: Stability and Efficiency*, *in* GROUP FORMATION IN ECONOMICS: NETWORKS, CLUBS AND COALITIONS 11, 12 (Gabrielle Demange & Myrna Holtz Wooders eds., 2005). Although it is arguable that the latter type may ultimately become the more important way to analyze communications technologies, for the time being, the dominance of a handful of infrastructure providers justifies analyzing them as managed networks.

[17] *Nodes* and *links* are in many ways analogous to the terms *vertices* and *edges* that appear in much of the literature on graph theory. We have modified the terminology to make them more specific to the particular technological contexts discussed here.

In practice, the *costs* of links can vary widely depending on the location and nature of the particular nodes being connected. In graph theory, these cost variations are represented by assigning different numerical values to the links and nodes. Links and nodes also vary in terms of *capacity*, which is also represented by assigning different numerical values to the links and nodes. The bandwidth of any particular transmission technology is typically limited. Nodes are also often subject to capacity limitations, since switches and routers in a telecommunications network can be limited in the number of calls they can route at any one time.

A system of nodes and links is called a *graph*. When links only operate in a particular direction, the graph is called a *directed graph*, and the links are depicted with arrows to indicate the direction of flow. A *network* is a graph with particular numerical values, such as cost or capacity, assigned to the links.

The *architecture* of a network refers to the set of nodes and the pattern of the links that connects them. A network functions as a system in the sense that it has various functionalities provided by various components as part of a larger set of interacting parts. For example, a telecommunications system provides services by transmitting communications acting as an integrated whole. The concept of a system is not new.[18] What is new, though, are better ways to understand the characteristics of complex systems. A *complex system* refers to a system in which its elements interact in ways that transcend any organizing principles being applied to the network, allowing the network to evolve and adapt to environmental changes.[19] The great interest in complex systems stems from the development of mathematical tools that provide insights into how networks perform and change.[20] Such techniques have been applied to study communications networks and the characteristics of connections in the World Wide Web.[21]

Analyzing networks through the lens of graph theory helps to explain network-architecture covering systems, such as telecommunications, which

[18] The systems approach was particularly fashionable beginning in the 1960s, having been applied widely in the natural sciences, mathematics, computer science, and the social sciences. The best-known classical applications are automated feedback and control systems. The application of systems theory to the study of social systems or organizations often is based on a biological metaphor. For example, von Bertalanffy states that "characteristic of organization, whether of a living organism or a society, are notions like those of wholeness, growth, differentiation, hierarchical order, dominance, control, competition, etc." This vision encompasses a wide array of approaches as "systems theories," including computer simulation, cybernetics, information theory, game theory, and decision theory. LUDWIG VON BERTALANFFY, GENERAL SYSTEM THEORY: FOUNDATIONS, DEVELOPMENT AND APPLICATIONS 47 (1968).

[19] *See, e.g.*, J.M. Ottino, *Complex Systems*, 49 AICHE J. 292, 293 (2003).

[20] As Ottino points out, the mathematical techniques used to study complex systems include "nonlinear dynamics, differential and difference equations and time series analysis, cellular automata, graph and network theory, and, depending on the problem, aspects of game theory, Markov processes, information theory, and genetic algorithms." *Id.* at 294.

[21] *See* Reka Albert et al., *Diameter of the World Wide Web*, 401 NATURE 130 (1999).

are the product of engineering design. Further, by applying the theory of random graphs, developments in the study of network evolution can provide insights into the forces that cause networks to evolve and adapt to their environment in various ways. The usage patterns of a telecommunications network depend on the decisions of individual subscribers who seek to make connections to each other. The connections to a network depend on the random arrival of customers seeking network services. Most importantly, for our purposes, regulators also affect the evolution of networks when they adopt rules that grant access to networks to firms that are competitors of network operators. Regulations such as those based on the Telecommunications Act of 1996 remove some of the control from the hands of network operators. The result of such access regulation is that a network's competitors exercise considerable discretion in determining at what points to connect to the network, what elements of the network to use, what additional switching equipment will be connected to the network, where the additional equipment will be collocated, and what types of traffic are added to the network. Network operators must alter network facilities and equipment to adapt to these physical modifications of the network and to changes in traffic patterns.[22] In short, under the regulated access regime, networks evolve through the access decisions of competitors. The result is to increase the complexity of network analysis still further.

III. GRAPH THEORY AND NETWORK CONFIGURATION

This Part explores the insights that graph theory provides into changes in network configuration. Section A introduces one of the fundamental cost-minimizing concepts of graph theory, the *minimum spanning tree*, and traces how changes in the cost of individual elements can have a dramatic impact on network design. Section B broadens the notion of network optimality to include reliability, analyzing how concerns about network failure and quality of service can lead to the inclusion of redundant capacity within a network. Section C explores how cost and capacity considerations interact to form economies of scale that affect technology choices as well as decisions about whether to aggregate traffic.

A. Cost Minimization as a Determinant of Network Configuration

Graph theory provides a framework for evaluating the performance of different network configurations. In this section, we introduce the *minimum spanning tree*, which offers a logical starting point for exploring the insights that graph theory provides into decisions about network configuration.

[22] In other work, we also analyze the transaction-cost implications of access and discuss how access regimes affect decisions about the proper boundaries of the network. *See* DANIEL F. SPULBER & CHRISTOPHER S. YOO, NETWORK REGULATION: THE MANY FACES OF ACCESS 40–60 (Vanderbilt Univ. Pub. Law Research Paper No. 05-19, 2005), *available at* http://ssrn.com/abstract=740297.

1. The Minimum Spanning Tree.—One basic network configuration directly links every pair of nodes. This type of network is represented by a *complete graph*. If a network consists of n nodes, then its complete graph will have $(1/2)n(n-1)$ links. For example, the network depicted in Figure 1 has six nodes and fifteen links. In the context of telecommunications, the complete graph corresponds to a point-to-point network. A prominent feature of point-to-point networks is that they contain numerous *cycles*, which are paths along which it is possible to pass through a succession of links and eventually return to the original node without crossing any link more than once.

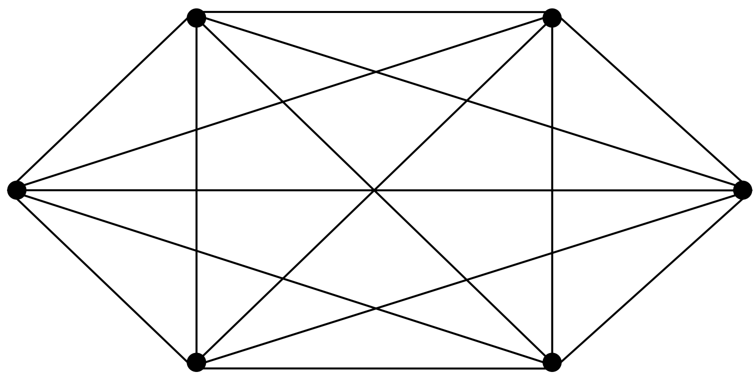

Figure 1. A point-to-point network of six nodes and fifteen links.

A point-to-point network is most desirable when the cost of switching is relatively high and the cost of transmission is relatively low. If links are relatively expensive, however, maintaining such an extensive set of dedicated connections can be inefficient. Under these circumstances, it may be more efficient to design a network that minimizes the number of links. The network architecture that minimizes the number of links is a *tree*, which is a graph that connects nodes without creating any cycles. A tree that connects all of the nodes in a network is known as a *spanning tree*. In a network with n nodes, such a spanning tree would consist of $n-1$ links. According to Cayley's Formula,[23] the number of spanning trees in a graph with n nodes is n^{n-2}. Thus, for four nodes, there are the sixteen possibilities depicted in Figure 2.[24] The number of spanning trees quickly becomes extremely large

[23] *See* Arthur Cayley, *A Theorem on Trees*, 23 Q.J. PURE APPLIED MATH. 376 (1889), *reprinted in* 13 THE COLLECTED MATHEMATICAL PAPERS OF ARTHUR CAYLEY 26 (A.R. Forsyth ed., University Press 1897). This is subject to the constraint that n be greater than or equal to two.

[24] Note that the foregoing discussion assumes that each of the four nodes is unique, either in terms of geographic location or cost. If the nodes are fungible, all of the trees are topologically identical (i.e.,

as the number of nodes increases. For ten nodes, there are 100 million possible configurations.

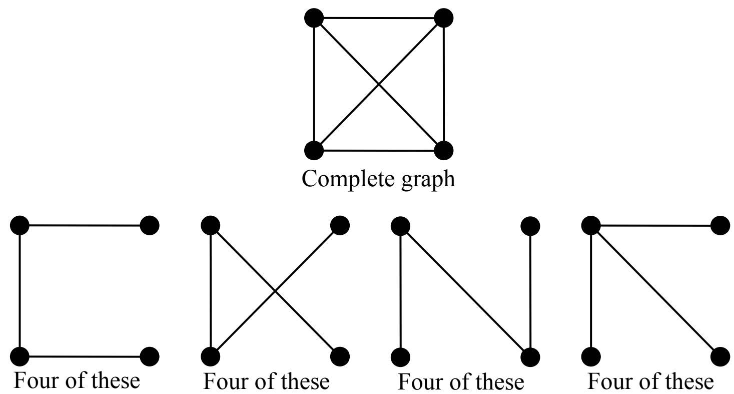

Figure 2. Spanning trees in a network with four nodes.

If the costs of connecting any two nodes were exactly the same, designing the least-cost network would simply be a matter of choosing one of the many available spanning trees. In practice, the costs of links can vary widely depending on the location and nature of the particular nodes being connected. In graph theory, these cost variations are represented by assigning different numerical values to the links and nodes that comprise the network.

Variations in the cost of links can provide network designers with a basis for determining the relative cost of different spanning trees. Given the manner in which the number of spanning trees expands exponentially as the number of nodes increases, one might think that identifying the minimum spanning tree for anything but the smallest network would require testing a large number of possibilities. Fortunately, a number of very efficient methods exist to identify minimum spanning trees. One example is Prim's algorithm, which proceeds as follows: Choose any node as the initial node and find the least-cost link connected to that node.[25] Add that link and its end node to form the start of the tree. Find the least-cost link connected to that tree and add that link and its end node to the tree. Continue until all nodes are included. What results is the minimum spanning tree (as depicted in Figure 3).

isomorphic) except for the four arrow-shaped trees depicted in the lower right-hand corner.

25 *See* R.C. Prim, *Shortest Connection Networks and Some Generalizations*, 36 BELL SYS. TECH. J. 1389 (1957).

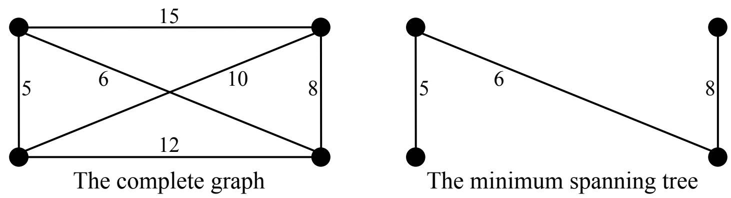

Figure 3. The minimum spanning tree.

2. Impact of a Change in the Cost of a Network Element.—Even this simple problem of identifying the minimum spanning tree demonstrates the extent to which networks constitute integrated systems. The links that comprise the cost-minimizing network are not chosen individually; they are instead selected in light of their relationship to the other network elements. The choices of which links to include and which to reject are thus interdependent. Whether a particular component should be part of the network depends not only on the cost of that component but also on the cost of the other components of the network. The decision also depends on the precise manner in which that component can be connected into the overall network.

Consider the impact of changing the cost associated with a single network component. Because the least-cost network configuration that connects all of the nodes of a network depends on the interrelationship of all of the network components, changing the cost associated with any network component has the potential to completely change what would be the least-cost network configuration that still connects every node. In other words, changes in the cost associated with a particular network component might lead to the creation of a completely different network. In the most extreme case, the particular element whose cost was being affected might even be omitted as a component of the resulting network.

Such an example is depicted in Figure 4. The only difference between the top and the bottom pair of graphs is a change in the cost associated in the diagonal link connecting the node in the upper left corner with the node in the lower right corner. This one change in cost causes a link that was previously part of the least-cost network configuration to be omitted from the network altogether.

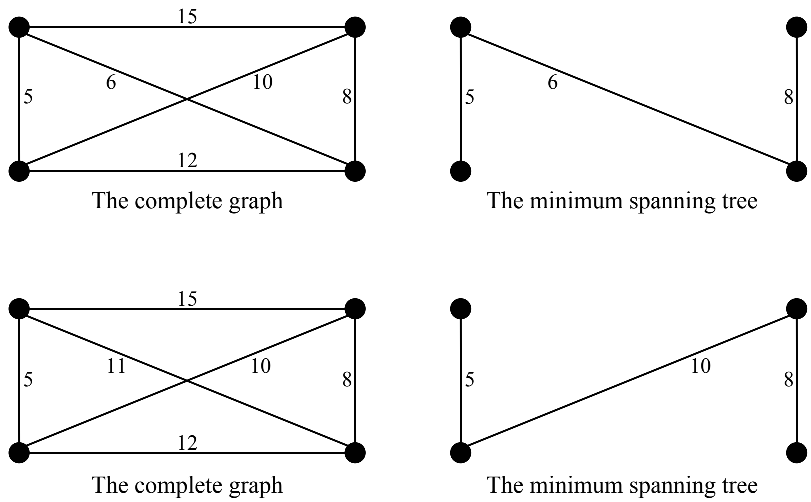

Figure 4. The impact of cost changes on the minimum spanning tree.

By capturing these interactions, graph theory provides intuitions about the ways that networks constitute integrated systems. The analytical tools provided by graph theory demonstrate how changes to one component of the network can affect the decisions about network configuration in surprising and unexpected ways. In so doing, it underscores that a network cannot be evaluated solely in terms of the costs associated with its individual components. Instead, the analysis of networks requires an appreciation for the complex manner in which different network components interact with one another. The role of each element can only be understood in the context of the system as a whole.

B. Reliability as a Determinant of Network Configuration

Cost is only one of the possible measures of network performance. It is not enough that a network be inexpensive; users also typically demand a certain degree of network reliability. We will discuss two different aspects of reliability: (1) protection against network failure and (2) maintenance of quality of service.

1. Network Failure.—Optimizing a network along any one dimension necessarily requires some degree of sacrifice along other dimensions. For example, because a minimum spanning tree contains no cycles, the path connecting any two nodes is unique. This fact leaves tree architectures vulnerable to network failure, since a breakdown in any of the links or nodes necessarily will disconnect a portion of the network.[26]

Reliability is accomplished by introducing a degree of redundancy so that the network can continue to satisfy communication demands even if some portion of the network fails. This redundancy is accomplished by introducing cycles into the network, which ensures that nodes are connected by more than one path. The most extreme case of redundancy is the complete graph with which we began our discussion.[27] The optimal level of redundancy designed into the network to protect against network failure depends on the tradeoff between cost and reliability.[28]

2. Quality of Service.—Reliability concerns can arise even when every element of a network is functioning properly. Reliability is also affected by the fact that network components vary not just in terms of cost but also in terms of capacity. When traffic arrives at a rate that exceeds the capacity of a particular network element, that traffic is forced into a queue. Saturation of the capacity of a particular network element, however, does not lead to outright network failure. Instead, it simply causes degradation in the network's quality of service. Many applications are not especially sensitive to brief delays. For example, a half-second delay may be imperceptible to users accessing e-mail and web content. Other applications, such as Internet telephony and streaming media, depend upon guaranteed throughput rates in order to function properly.

The uniqueness of the paths connecting any two nodes in a tree architecture turns every network component into a potential bottleneck. The saturation of the capacity of any particular network component thus has the inevitable effect of reducing the quality of service for all traffic that must pass through that component. The greater the variability in network demand, the more redundant capacity must be maintained. The saturation of network components may be caused by short-run, transient variability in demand, in which case reliability problems are likely to be temporary and geographically limited. Saturation can also result from more fundamental shifts in usage patterns, such as those that occur when network planners fail to properly anticipate the geographic development of neighborhoods, in which case reliability problems may be more enduring and may affect a larger area.

---

[26] Graph theory provides two measures of a graph's connectedness. One measure is to remove nodes and see whether the remaining graph still reaches all of the nodes. A graph with at least $k + 1$ nodes is *k-connected* if it is still connected after removing $k - 1$ or fewer nodes. The other measure is to remove links to see whether the remaining graph still reaches all of the nodes. A graph is said to be *k-link-connected* if it is still connected after removing $k - 1$ or fewer links.

[27] *See supra* notes 24–25 and accompanying text.

[28] *See* ALDOUS & WILSON, *supra* note 9, at 236–38; Prakash Mirchandani & David Simchi-Levi, *Communication Network Design Models*, *in* HANDBOOK OF GRAPH THEORY, *supra* note 9, at 1117, 1126–30.

As a result, network designers concerned with maintaining quality of service will employ cycles to design a degree of redundancy into the network, as depicted in Figure 5. Cycles can be created by adding redundant capacity along an existing link (indicated by link *A*) or by creating an alternative path between two nodes not previously linked together (indicated by link *B*). The presence of alternative pathways provides a safety valve that preserves quality of service.

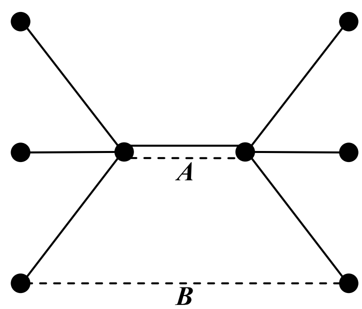

Figure 5. The use of cycles to maintain quality of service.

C. Interactions Between Cost and Capacity: Economies of Scale

The preceding sections have shown how cost and capacity independently affect the optimal network configuration. Cost and capacity interact when a higher capacity network technology exists that lowers the variable costs of providing network service but increases the fixed costs. When this occurs, the node or link technology is said to exhibit economies of scale.[29]

The presence of economies of scale can have a dramatic impact on network design. For example, if there are sufficient economies of scale in links, the network designer may choose a configuration that consolidates traffic. To illustrate this effect, consider a network with nodes that are located in two clusters, as depicted in Figure 6. Network designers must decide how best to link the two hubs: they can use a low-volume linking technology or a higher-capacity trunk line. If the volume of traffic is sufficiently large, the savings in variable costs associated with employing a trunk line will justify the incurrence of the larger fixed costs needed to do so. The minimum number of customers needed for a trunk line to be economically beneficial can be determined by dividing the fixed costs of establishing the trunk line by the per-customer cost savings resulting from trunk-

[29] *See generally* DANIEL F. SPULBER, REGULATION AND MARKETS 115–18 (1989) (discussing sources of economies of scale).

line transmission. In other words, the greater the cost savings, the smaller the number of customers needed for the trunk-line architecture to pay out.

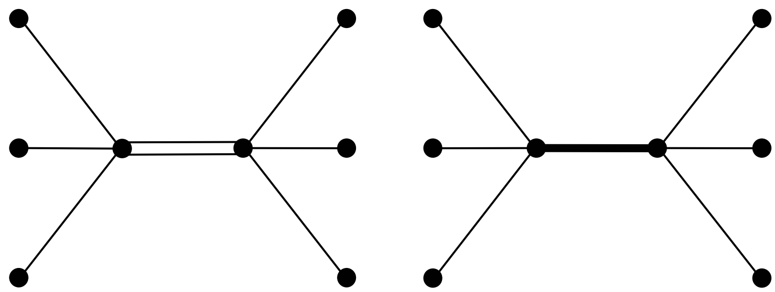

Figure 6. The choice between redundant links and trunk lines.

The preceding examples show that network designers may aggregate traffic to achieve economies of scale.[30] The use of central switching in hub-and-spoke configurations and the establishment of trunk lines take advantage of economies of scale in nodes and links through the use of high-capacity switches and transmission lines. Again, the optimal network configuration is highly sensitive to the costs of different switching and linking technologies.[31] In addition, the decision to employ a trunk line also implicates the tradeoff between cost and reliability, since placing reliance on a single, higher-volume link increases the risk of network failure.

The extent to which network designers deviate from a tree architecture thus depends upon a wide range of considerations, including the cost and capacity of individual components, the relative cost of switching and linking, the overall level and variability of network demand, the presence of economies of scale, and the benefits from aggregation provided by particular technologies. Decisions about network configuration also depend on the tradeoff between cost and reliability as well as a wide range of characteristics unique to the particular network. Most importantly, the analytical tools discussed above demonstrate the importance of conceptualizing networks as

---

[30] *See* ROBERTAZZI, *supra* note 14; B. Yaged, Jr., *Minimum Cost Routing for Static Network Models*, 1 NETWORKS 139 (1971).

[31] Another source of returns to scale is the economies of scope that arise when a single network handles multiple types of random traffic. The law of large numbers dictates that unless the demand for the product is completely positively correlated, bundling the products together typically reduces their cumulative variability. *See* Christopher S. Yoo, *Rethinking the Commitment to Free, Local Television*, 52 EMORY L.J. 1579, 1707–08 n.471 (2003). As a result, aggregating traffic can allow network owners to maintain quality of service in a more efficient manner by reducing the variability of the demands on network capacity.

integrated systems that are more than the sum of their parts and appreciating the extent to which network design hinges on understanding the complex interrelationship among network components.

IV. GRAPH THEORY AND NETWORK CAPACITY

In addition to providing insights into how the interrelationships among particular network components affect the network's overall cost and reliability, graph theory also reveals how network configuration affects the network's maximum capacity. Section A introduces a well-known principle in graph theory known as the max-flow/min-cut theorem, which provides a way to determine the maximum carrying capacity of any transportation network and the costs associated with operating it at that level.[32] Section B traces the implications of the max-flow/min-cut theorem for network policy.

A. The Max-Flow/Min-Cut Theorem

Although a formal proof of the max-flow/min-cut theorem exceeds the scope of this Essay, the basic intuitions underlying it can be easily explained in a relatively nontechnical manner. The theorem focuses on the maximum flow capacity of a network transporting traffic from a source (usually denoted *s*) to a destination (called a "sink" and usually denoted *t*).[33] A *cut* is a set of links that, if removed, divide the graph into two subgraphs. The two subgraphs contain distinct sets of nodes. Of greatest relevance, for our purposes, are those cuts in which the source (node *s*) and the sink (node *t*) lie in different sets. The set of nodes containing the source node *s* is typically called *S*, while the set of nodes containing the sink node *t* is typically called *T*. The key insight is that any traffic bound from node *s* to node *t* necessarily must travel from nodes contained in set *S* to nodes contained in set *T* and hence along the links in the cut. As a result, the total carrying ca-

---

[32] This theorem is prominently discussed in most graph theory texts. For a relatively accessible exposition, see Clifford Stein, *Maximum Flows*, *in* HANDBOOK OF GRAPH THEORY, *supra* note 9, at 1075.

[33] The network is portrayed as a one-way network from a single source to a single sink. Certain aspects of this simplification do not affect the generality of the analysis. For example, the insights provided by the analysis of *s-t* networks can be extended to networks with multiple sources and multiple sinks without any loss of generality. *See* WALLIS, *supra* note 9, at 160–61. In addition, the foregoing exposition focuses solely on the capacity of links without taking into account the capacity limits on nodes. A node with a capacity constraint can be reconceptualized as two uncapacitated nodes connected by a capacity-constrained link. As a result, the entire analysis can focus solely on the capacity of links without loss of generality.

Other aspects are potentially more problematic. The most significant concern is that telecommunications networks tend to be two-way rather than one-way. It is theoretically feasible to model these flows as two different networks, although if the capacity of certain nodes is used for traffic passing in both directions, some means for trading off capacity must be devised. The problem can also be analyzed as a multicommodity flow problem. Multicommodity problems pose additional difficulties that sometimes border on the intractable. *See* Stein, *supra* note 32, at 1083–84.

pacity of the links directly connecting one of the nodes in set *S* to the set of nodes in set *T* represents a constraint on the total carrying capacity of the network.

This insight is most easily understood by analyzing the simplest cuts in the context of the relatively simple network depicted in Figure 7. That network consists of eight nodes connected by a series of links; the capacity of each link is indicated on the graph.[34] Consider the cut represented by the dotted line labeled *A*. In this cut, which consists of the links connecting node *s* to nodes *a*, *b*, and *f*, set *S* consists of only node *s*, while set *T* consists of the remaining seven nodes (*a*, *b*, *c*, *d*, *e*, *f*, and *t*). Any traffic leaving node *s* must necessarily travel along one of the links connecting node *s* to nodes *a*, *b*, or *f*. As a result, it is clear that the total carrying capacity of the network from *s* to *t* can be no greater than the sum of the capacity of those links, which in this case is 38 + 1 + 2 = 41. A similar logic obtains with respect to cut *B*, which divides the network into a set *S* consisting of seven nodes (*s*, *a*, *b*, *c*, *d*, *e*, and *f*) and a set *T* consisting only of node *t*. Because only four links connect to node *t*, it is also clear that the total capacity of the network cannot exceed the total capacity of those four links, which in this case is 7 + 7 + 1 + 18 = 33. Because every traffic flow must necessarily pass through both of these cuts, the capacity of the links crossing these cuts represents a constraint on the flow capacity of the entire system. In other words, the network cannot possibly carry more than the lower of 41 or 33.

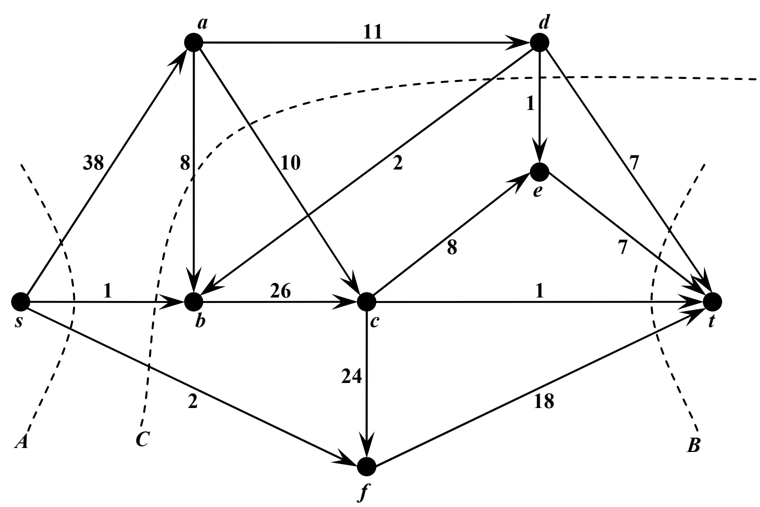

Figure 7. Illustration of the max-flow/min-cut theorem.

[34] Figure 7 is adapted from JUNGNICKEL, *supra* note 9, at 164–69.

One can generalize from this insight by analyzing the entire universe of possible cuts to the network. The cut with the smallest capacity (called the *minimum cut*) represents the maximum flow that one can push through the system.[35] The minimum cut in any network can be identified by using an iterative algorithm. In this case, the cut with the minimum capacity is the one depicted by cut *C*, which has a capacity of 31. Because all traffic traveling between nodes *s* and *t* must necessarily travel across cut *C*, the total capacity of the links crossed by cut *C* necessarily represents the maximum carrying capacity of the network. In addition, the algorithm for identifying the minimum cut also determines the flow passing along each noncritical link when the network is operating at full capacity. This makes it possible to determine the direct costs of the operating network when carrying its maximum volume.[36]

B. Implications of the Max-Flow/Min-Cut Theorem for Network Policy

Now consider the impact on the network of granting a competitor access to a network element. Access ties up some of the capacity of the leased elements, which in turn reduces the capacity available to the network owner. One of the most interesting insights of graph theory is that changing the capacity of any particular network component can have a surprising and somewhat unpredictable impact on the capacity of the network. Indeed, changes in capacity can cause the links that comprise the minimum cut (which represent the constraint on the total carrying capacity of the network) to change to a completely different set of links.

Consider, for example, the situation depicted in Figure 8, which is the same network represented in Figure 7, except now a competitor has obtained access to six units of the capacity of the link connecting nodes *c* and *e* (marked on the graph with a circle). From the perspective of the network owner, this has the effect of reducing the available capacity of that link by that amount. In this case, the magnitude of this reduction causes the links comprising the minimum cut to change to the one portrayed by the line marked *D*. As a result, granting a competitor access to a link that was not part of the minimum cut (and thus had slack in its capacity) nonetheless reduces the effective capacity of the entire network. The amount of the reduction is not necessarily equal to the magnitude of the access granted. Instead, the amount of reduction depends upon the configuration of the particular network involved. In the situation portrayed in Figure 8, the impact of allowing a competitor to occupy six units of capacity in the link connect-

[35] Although it is quite intuitive that the maximum carrying capacity of a transport network must be *less than or equal to* the capacity of the minimum cut, the max-flow/min-cut theorem actually establishes that the maximum flow of the network is *strictly equal to* the capacity of the minimum cut.

[36] If the network is not operating at full capacity, the optimization problem instead focuses on identifying the minimum cost flow and requires a slightly different approach. *See* Lisa Fleischer, *Minimum Cost Flows*, *in* HANDBOOK OF GRAPH THEORY, *supra* note 9, at 1087.

ing node *c* with node *e* is to reduce the effective capacity of the entire network by two. The value of carrying two units from node *s* to node *t* represents the opportunity cost of granting this competitor access to the link between nodes *c* and *e*. In addition, the change in direct costs associated with granting access can be determined by recalculating the flow passing along each noncritical link when the post-access network is operating at full capacity.

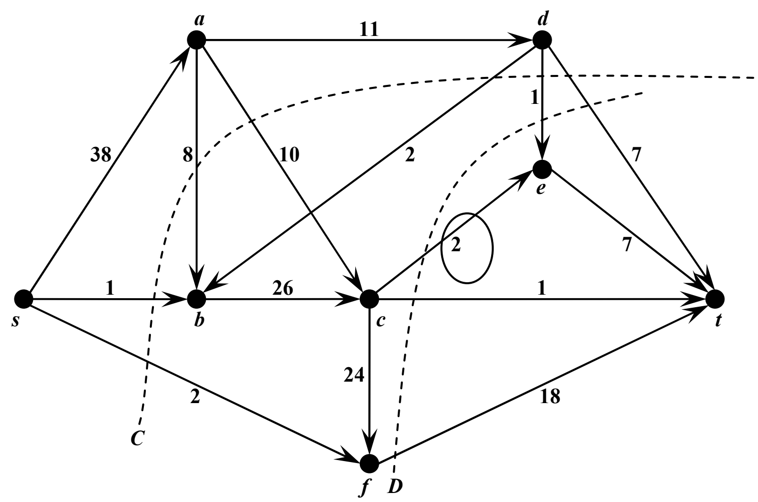

Figure 8. Impact of unbundled access on network capacity.

At the same time, it is quite possible that granting access to a particular link might have no effect on the total carrying capacity of the network. Whether that is the case depends on how much slack exists in the link to which access is being granted. If the link to which access is being granted is not part of the minimum cut and the additional reduction in capacity does not make it part of the minimum cut, granting access has no impact on the overall capacity or the full-capacity operating costs of the network. On the other hand, if the link is already part of the minimum cut or if granting access to that link places it on the minimum cut, granting access to that link reduces the total carrying capacity of the network.

Therefore, the retail value of the reduction in network capacity represents a good measure of the opportunity cost associated with granting access to competitors. Note, however, that the reduction in capacity is not necessarily equal to the size of access granted. The rest of the network may compensate for the reduction in the capacity of one link in ways that partially offset the impact of the reduction. In practice, access by competitors is not an isolated event but a pattern of usage that changes over time as more competitors access the network. The network's available carrying capacity will evolve over time as a consequence of the access decisions of competitors.

V. THE POLICY IMPLICATIONS OF A GRAPH THEORETICAL APPROACH

The ability of graph theory to reflect the interactions among network components offers insights that can guide policymaking. This Part will sketch how the graph theoretical approach that we propose might shed new light on a number of leading issues confronting policymakers. These include the decision to apply the same methodology to set prices for interconnection and access to unbundled network elements under the Telecommunications Act of 1996, the decision to base access rates on cost, proposals to base access prices on market benchmarks, compelled access to broadband networks, and the regulation of Internet telephony. We also explore the extent to which the framework can be employed to support a formal calculation of regulated rates. Although a complete analysis of these issues exceeds the scope of this Essay, the brief discussion we offer here on each policy choice should be sufficient to convey a general sense of graph theory's potential to inform network policy.

A. The Impact of Interconnection and UNE Access

Perhaps the key innovation of the Telecommunications Act of 1996 was the attempt to introduce competition into local telephone service. Congress envisioned that competition in local telephone markets might emerge through one of three paths.[37] First, a new entrant might obtain *all* of the necessary elements from the existing local telephone company, called an incumbent local exchange carrier ("LEC"), and resell them.[38] Second, a new entrant might construct an entirely new network.[39] Congress recognized, however, that not every facilities-based entrant would be able to have its entire network in place at the time it began to offer local service.[40] In order to allow competition to emerge before entrants had fully established their networks, Congress established a third path for entering local telephone markets. It required every incumbent LEC to provide other carriers with access to all of its network elements on an unbundled basis.[41] The

[37] *See* Verizon Communications, Inc. v. FCC, 535 U.S. 467, 491–92 (2002); Implementation of the Local Competition Provisions in the Telecommunications Act of 1996, First Report and Order, 11 F.C.C.R. 15,499, 15,509 ¶ 12 (1996) [hereinafter Local Competition Order].

[38] *See* 47 U.S.C. § 251(c)(4) (2004). Resale rates are based on retail rates charged to subscribers for the telecommunications service requested less the costs that will be avoided by the LEC, such as the costs attributable to marketing, billing, and collection. *Id.* § 252(d)(3).

[39] *Id.* § 251(c)(2).

[40] S. CONF. REP. NO. 104-230, at 147–48 (1996).

[41] 47 U.S.C. § 251(c)(3). In determining which network elements would be subject to the unbundled access requirement, the statute required the FCC to consider whether "access to such network elements as are proprietary in nature is *necessary*" and whether "the failure to provide access to such

statute requires that the terms of interconnection and access to unbundled network elements ("UNEs") be "just, reasonable, and nondiscriminatory."[42]

The Federal Communications Commission ("FCC") implemented the local competition provisions of the 1996 Act in a massive order issued just six months after the statute's enactment.[43] The order required that resale prices be based on retail prices less the costs that are avoided when telecommunications services are provided at wholesale rather than retail (such as the costs associated with marketing, billing, and collection).[44] Rates for both interconnection and access to UNEs would be determined under a methodology known as Total Element Long Run Incremental Cost ("TELRIC"). TELRIC bases rates on the sum of the incremental costs directly attributable to the specified element and a reasonable allocation of common costs, both determined on a forward-looking, replacement-cost basis.[45] The Supreme Court upheld TELRIC as a matter of statutory construction in *Verizon Communications, Inc. v. FCC*.[46] Although the FCC has begun to reconsider certain aspects of TELRIC, the notice of proposed rulemaking that initiated this proceeding made clear that the FCC envisioned only modest adjustments to the relevant cost categories and reaffirmed the basic decision to base both interconnection and UNE access pricing on forward-looking cost.[47]

Basing rates for both interconnection and access to UNEs on the TELRIC methodology leads to incorrect outcomes. In essence, TELRIC adopts a bottom-up approach that determines UNE access and interconnection rates by aggregating costs on an element-by-element basis. This approach neglects the effects of capacity utilization on the overall performance of the network.

---

network elements would *impair* the ability of the telecommunications carrier seeking access to provide the services that it seeks to offer." *Id.* § 251(d)(2)(A), (B) (emphasis added).

[42] *Id.* § 251(c)(2)(D), (c)(3); *see also* § 252(d)(1)(A) (requiring that interconnection and UNE access rates be based on cost and nondiscriminatory).

[43] Local Competition Order, *supra* note 37, at 15,509 ¶ 12.

[44] *Id.* at 15,955–64 ¶¶ 908–934.

[45] 47 C.F.R. § 51.505(a) (2004). As a formal matter, TELRIC properly refers only to the first of these two components. For simplicity, however, we will refer to both parts of the methodology collectively as TELRIC. In addition, the FCC did not formally require incumbent LECs to set interconnection and UNE access rates in accordance with TELRIC. Instead, the statute provides that parties remain free to establish the terms of interconnection and UNE access through private negotiations. That said, negotiations that fail are subject first to mediation and then arbitration by the state public utility commissions ("PUCs"). *See* 47 U.S.C. § 252. Because the FCC's rules require that any such arbitration be settled in accordance with TELRIC, TELRIC-based prices strongly influence the outcome of those negotiations and constitute a de facto ceiling for interconnection and UNE access rates.

[46] 535 U.S. 467 (2002).

[47] *See* Review of the Commission's Rules Regarding the Pricing of Unbundled Network Elements and the Resale of Service by Incumbent Local Exchange Carriers, Notice of Proposed Rulemaking, 18 F.C.C.R. 18,945, 18,957–60 ¶¶ 29–37 (2003) [hereinafter TELRIC NPRM].

Consider first interconnection. The 1996 Act gives competitors the right to request interconnection "at any technically feasible point."[48] Graph theory demonstrates the potential flaw in the idea that the costs of interconnection are confined to the network elements that are directly involved. Instead, interconnection by competitors is likely to introduce new sources and sinks into the network. Thus, substantial amounts of traffic may originate and terminate at points in the network that differ from the host network's initial points of origin and termination. This will alter traffic patterns. A network that is designed with a maximum flow/minimum cut pattern designed around particular sources and sinks will no longer be appropriate for traffic coming from new sources and sinks. The nodes at which interconnection occurs will not be the only nodes affected. Rather, the effects will be distributed across all nodes and links within the network. This invalidates the notion that only the incremental costs of providing the interconnection should be recovered. The interconnection affects the network's performance and creates costs throughout the network.

Moreover, graph theory shows that one should not expect the effects of UNE access to be confined to those elements. When individual elements are viewed in isolation, the TELRIC methodology seems quite reasonable. Typically, UNE access occupies only a few of the elements of an incumbent LEC's network. Those elements, however, can be critical to overall traffic patterns that connect the network's sources and sinks. The reduction of available capacity on critical links in the network will affect the network's maximum flow. Thus, UNE access can impose costs on the host network that extend well beyond the elements that are affected. In some cases, the costs of UNE access may even exceed interconnection costs. If usage patterns associated with interconnection are similar to those of the incumbent LEC's own traffic, absent any capacity constraints, there will be less of an impact on the network owner's decisions about network configuration. There will be no change in the network elements that comprise the minimum cut and, thus, in the components that constitute bottlenecks. The situation is quite different when usage patterns associated with interconnection differ from the patterns of the incumbent's own traffic. When that is the case, granting access to critical UNEs can create bottlenecks where none previously existed and can have a dramatic impact on the network's maximum flow. Under these circumstances, UNE access can have a dramatic impact on the cost, capacity, and configuration of networks.

B. The Shortcomings of the Cost-Based Pricing

Graph theory reveals problems with the prices charged for both interconnection and UNE access as well. As noted earlier, the FCC has effectively ruled that both interconnection and UNE access prices be determined

---

[48] 47 U.S.C. § 251(c)(2)(B).

in accordance with TELRIC, which builds rates from the bottom-up by aggregating the costs (determined on a forward-looking basis) of each individual network component. Other access regimes employ similar cost-based methodologies.[49]

Graph theory highlights a significant flaw in basing access prices on the costs associated with a particular network element. Cost-based approaches necessarily focus solely on the characteristics of the network element in isolation. As a result, they fail to take into account the manner in which that element interacts with the system as a whole. In so doing, the current approach fails to take into account one of the central characteristics of networks, which is how changes to one component can affect the entire system. Most importantly, allowing a competitor to occupy a portion of the capacity of a particular component can cause a cascade effect as the rest of the network attempts to compensate by rerouting traffic through other portions of the network.

An example based on the simple network portrayed in Figure 9 demonstrates the shortcomings of the current approach. It represents a classic ring configuration carrying traffic originating and terminating on nodes *a*, *b*, *c*, and *d*.

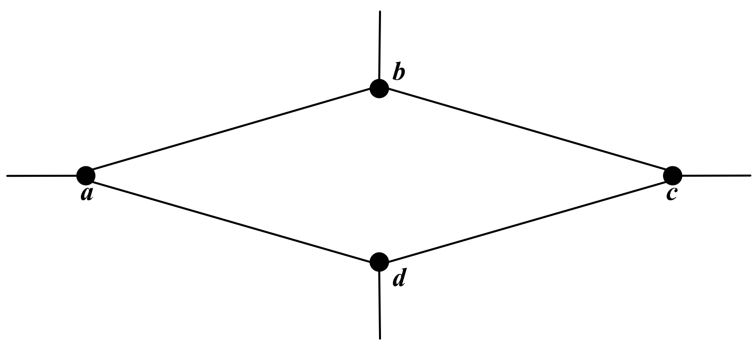

Figure 9. A ring network with four nodes.

Suppose that a competitor requested unbundled access to a portion of the capacity of the link connecting nodes *a* and *b*. According to the current approach, the price paid by the competitor for this access would be determined by the replacement cost of the capacity of that link plus a share of the common costs. Is this pricing regime likely to represent a reasonably accurate reflection of the true costs to the network owner?

---

[49] *See, e.g.*, 47 C.F.R. § 1.1409 (using a cost-based methodology to establish rates for access to networks of utility poles).

The answer is almost certainly no. From the perspective of the network owner, granting the competitor access to the link between nodes *a* and *b* has the effect of reducing the effective capacity of this link. If this link is operating well below capacity, granting access to the competitor would have a negligible impact on the system. The costs would be so low that allowing the network owner to recover on the basis of network cost would likely overcompensate the network owner.

A different situation arises if the link between nodes *a* and *b* is already part of the minimum cut or if the access request makes it part of the minimum cut. If the link between nodes *a* and *b* becomes saturated, the network will have to reroute calls between those nodes via nodes *c* and *d*. Because this traffic must pass through three links and three nodes, rather than one link and one node, the reduction in capacity will cause the network owner to incur three times the cost (assuming that the costs associated with each node and each link are symmetrical). Thus, even when only considering traffic between nodes *a* and *b*, it becomes clear that the overall impact to the network cannot be determined solely by analyzing the costs associated with the link between nodes *a* and *b*.

The impact of the system becomes even more dramatic when the traffic between the other nodes is taken into account. Not only does forcing the network to reroute calls between nodes *a* and *b* through nodes *c* and *d* affect the cost of servicing calls between nodes *a* and *b*, it also has a potential impact on traffic among the other nodes. Consider, for example, traffic passing between node *a* and node *c*. In the absence of an access request, the network owner could route some of the traffic between nodes *a* and *c* via node *b* and route the rest of the traffic via node *d*. The situation changes, however, if a request for access causes the link between nodes *a* and *b* to become saturated. In that event, the network owner has no choice but to increase the proportion of the traffic between nodes *a* and *c* routed through node *d*. This has the potential to degrade all of the traffic passing through nodes *a*, *c*, and *d*, including both transit traffic between nodes *a* and *c* and local traffic between nodes *a* and *d* and between nodes *c* and *d*. In addition, the link between node *b* and *c* is no longer carrying any traffic between nodes *a* and *c*, which essentially renders superfluous the capacity designed into this link to carry this traffic. The short-run impact of the UNE access requirement of the Telecommunication Act of 1996, thus, cannot be determined simply by looking at the affected elements in isolation.

Granting access to a particular link can have even more significant long-run effects by altering the design of a network to connect a set of nodes. Recall that decisions whether to aggregate traffic to take advantage of economies of scale turn on whether sufficient traffic exists to defray the costs. Reducing the effective capacity of a link can render a particular aggregation decision unprofitable. Alternatively, as noted earlier, a change in the cost or capacity of one or more links can change the optimal network configuration to the point where the particular link in question would not be part of the network at all.

The problem of determining rates is made all the worse by mandating interconnection "at any technically feasible point."[50] This requirement prevents the network owner from minimizing the adverse impact to its system by choosing which facilities to employ when fulfilling any particular request for service. In the worst case scenario, the right to designate the point of interconnection gives competitors the opportunity to act strategically by basing their access requests not on their needs, but rather on what would inflict the greatest harm on the network owner. A network owner may wish to hedge against this possibility by maintaining excess capacity in case one of its competitors decides to request access to a key portion of its network. This has the drawback of forcing the network owner to make capital investments that may never be used. Indeed, competitors that are acting strategically may well take into account whether the network owner maintains such excess capacity when deciding whether and where to request access. If so, the mere fact that the network owner has added excess capacity to hedge against the possibility of a strategic access request effectively guarantees that access will be sought elsewhere.

Graph theory has the potential to provide a better basis for determining the value of access to any particular set of network components. The algorithms discussed above provide the means for determining how much it costs to operate each network element, whether particular network elements are slack or saturated, and what impact granting access will have on the overall carrying capacity of the network.

That said, those seeking to turn graph theory into a regulatory system must confront a number of complications that may cause rates determined on the basis of graph theory to systematically undercompensate network owners. For example, determining the appropriate price becomes much more difficult once one takes into account that the volume of traffic typically varies over different portions of the day. During quiet times, the demands on the network are typically so low that no link is close to saturation, in which case the impact on the network owner would be negligible. When that is the case, access prices would be extremely low. The situation can be quite different during the peak times of the day, when granting access can saturate the overall carrying capacity of the network. Network saturation would justify charging a higher price for access. But at the same time, as queues begin to form behind the relevant bottlenecks, saturation also would degrade network performance.

A network owner confronted with such a possibility might well choose to help bolster the reliability of its network by adding capacity. Put another way, what appears at first blush to be excess capacity can more properly be regarded as a form of insurance that protects consumers against network

[50] 47 U.S.C. § 251(c)(2)(B).

failure when short-run surges in demand cause the network to be saturated. On an ex ante basis, such excess capacity is serving a useful purpose even if it is not subsequently used.

The analysis is complicated even further by the reality that, in most cases, the incurrence of fixed costs cannot be expanded incrementally. Instead, fixed costs are often indivisible, in that they can only be added in large, discrete quantities. In other words, because it is impossible to buy one-third of a switch, the network owner must either purchase an entire switch or no switch at all.[51] As a result, network owners interested in maintaining quality of service will necessarily add capacity before it is actually needed, which in turn will make it appear that networks never operate near full capacity. A false impression—that excess capacity is endemic to communications networks—would result. This false impression may cause graph theoretical analyses to systematically understate the adverse economic impact caused to the network owner by granting access.

Such problems are not unique to graph theory. Indeed, conventional ratemaking techniques are plagued with similar problems that typically revolve around whether a particular investment was either "prudent" when undertaken or is "used and useful" at the present time.[52] As a result, the problems confronting graph theory are not different in kind from those facing more conventional approaches to ratesetting (although they may be different in degree). Moreover, it is quite possible that graph theoretical models could be used to establish a baseline that could be adjusted to reflect these realities. Although such an approach would lack the elegance and accuracy of a formal solution, it may well provide a better foundation for policymaking if it offers a better second-best approximation of the efficient outcome than current ratemaking techniques.

C. Limitations on Basing Prices on Market Benchmarks

Other policymakers and scholars have proposed basing access rates on benchmark prices charged for comparable goods or on the final retail price of the network services.[53] Economic theory, as well, generally favors using comparable sales as the best indicator of value. The prices charged for similar goods and services provide observable evidence of the benefits to buyers and the costs to sellers. In addition, unlike cost-based approaches, which focus solely on the supply side of the transaction, market-based approaches take into account what is perhaps the fundamental insight of neo-

---

[51] *See* Spulber & Yoo, *supra* note 5, at 913.

[52] *See* William J. Baumol & J. Gregory Sidak, *The Pig in the Python: Is Lumpy Capacity Investment Used and Useful?*, 23 ENERGY L.J. 383 (2002).

[53] *See* TELRIC NPRM, *supra* note 47, at 18,959 ¶ 34 (describing proposals to base access rates on retail prices through a methodology known as the Efficient Component Pricing Rule ("ECPR")); Local Competition Order, *supra* note 37, at 15,837–39 ¶¶ 660–662, 15,859 ¶ 709 (same); Spulber & Yoo, *supra* note 5 (proposing such an approach for basing access prices on market benchmarks).

classical economics, which is that value is determined by an interaction of both supply and demand.[54] Institutional considerations also appear to favor adoption of a market-based regime. The decentralized decisionmaking inherent in markets is more likely to be able to incorporate the copious amounts of information needed to assess the value of particular network services than are regulatory authorities attempting to determine the correct measure of cost.[55] This is particularly the case given that network owners are the primary source of information for regulatory decisionmaking, since owners have every incentive to be chary about disclosing information about their networks. The advantages of decentralized decisionmaking also will be particularly pronounced in industries that are technologically dynamic. Although replacement cost tends to reflect market value over the long run, technological change and other exogenous shocks can causes prices to deviate widely from efficient levels in the short run.[56]

Even the seminal analyses that laid the foundation for the adoption of cost-based pricing explicitly recognized the preferability of taking a market-based approach. Despite this recognition that market-based benchmarks would better reflect economic value, the absence of external markets rendered market-based pricing a practical impossibility. This left regulators with little choice but to base their determination of a rate's reasonableness on the costs of providing network service, with the primary debate centering on whether such rates should be based on historical cost or replacement cost.[57] Whatever the validity of this logic when first offered, it carries little weight in an era in which robust competition for telecommunications services has emerged. The emergence of alternative network technologies has now made market benchmarking quite feasible. It has made it possible for policymakers to determine the value of network elements indirectly by using sales of comparable network services as market benchmarks. Thus, those seminal analyses supporting cost-based pricing implicitly acknowledged that the emergence of markets would cause the basic rationale supporting cost-based pricing to collapse.[58]

Consider the network element long regarded as the most likely subject of UNE access requests: the telephone wires connecting individual households to the main switching facilities maintained by local telephone companies, otherwise known as "local loops." Because of the high sunk costs associated with wiring individual neighborhoods with telephone wires, the local loop has long been regarded as the portion of the network least likely

---

[54] *See* Christopher S. Yoo, *New Models of Regulation and Interagency Governance*, 2003 MICH. ST. DCL L. REV. 701, 712.

[55] *See* FRIEDRICH A. HAYEK, THE FATAL CONCEIT 85–88 (W.W. Bartley III ed., 1989).

[56] *See* Spulber & Yoo, *supra* note 5, at 899.

[57] *See* Missouri *ex rel.* Southwestern Bell Tel. Co. v. Pub. Serv. Comm'n, 262 U.S. 276, 292–95 (1923) (Brandeis, J., concurring in the judgment) (offering a classic statement of this rationale).

[58] Indeed, the emergence of competitive alternatives undercuts the rationale for compelling access in the first place. *See* Spulber & Yoo, *supra* note 5, at 973, 992, 1020.

to be duplicated. Over the last few decades, however, wireless carriers have deployed spectrum-based voice networks that can perform the same functions as the local loop. Further, because wireless carriers often wish to place calls to wireline telephone customers or to customers of other wireless carriers, a market for services like those provided by the local loop has developed. Although the terms of wireless-to-*wireline* interconnection are determined by regulation,[59] the terms of wireless-to-*wireless* interconnection are settled through arms-length transactions mediated by a market.[60] In addition, competition among wireless providers has become quite robust, with eighty-three percent of the U.S. population being served by five or more wireless providers.[61] As a result, the outcomes of arms-length negotiations between wireless providers could arguably be regarded as relatively close approximations of efficient pricing.[62] Even if direct market benchmarks remain unavailable, methods exist for imputing prices for individual network elements from the price charged in the retail market for network services.[63]

As a result, even critics of market-based approaches concede that so long as competition is sufficiently robust, market benchmarks would represent the best basis for setting access rates.[64] The primary concern has been that the relevant markets that would be the source of such benchmarks are excessively concentrated. Under standard monopoly or oligopoly theory, the absence of robust markets can lead the prices charged in those markets to be inefficiently high.[65]

---

[59] For example, the terms of wireless-to-wireline interconnection are governed by the Telecommunications Act of 1996. *See* Local Competition Order, *supra* note 37, at 15,998–16,000 ¶¶ 1012–1013.

[60] *See* Interconnection and Resale Obligations Pertaining to Commercial Mobile Radio Services, Fourth Report and Order, 15 F.C.C.R. 13,523, 13,534 ¶ 28 (2000) (declining to regulate the terms of wireless-to-wireless interconnection).

[61] *See* Eighth Report on CMRS Competition, *supra* note 2, at 14,876 ¶ 217 (citing this fact as support for the conclusion that the market for wireless telephony is effectively competitive).

[62] The process of converting comparable transactions into actual prices does pose a number of complexities. For example, comparable transactions typically need to be adjusted to reflect differences in actual services provided. In addition, interconnection between wireless carriers is often accomplished indirectly through the incumbent LECs under terms dictated by state regulatory authorities. *See* Local Competition Order, *supra* note 37, at 15,514 ¶¶ 26–27. As such, wireless-to-wireless interconnection prices may not represent a truly independent point of reference.

[63] One such method is known as the Efficient Component Pricing Rule ("ECPR"). *See generally* J. GREGORY SIDAK & DANIEL. F. SPULBER, DEREGULATORY TAKINGS AND THE REGULATORY CONTRACT 283–392 (1997) (reviewing the history and criticisms of ECPR and offering a more refined, market-based version called M-ECPR).

[64] *See* Candeub, *supra* note 15, at 430 (calling the market-benchmark approach to setting UNE prices, "if capable of implementation, perfectly correct" before criticizing the proposal as unimplementable).

[65] *See id.* at 431; TELRIC NPRM, *supra* note 47, at 18,959 ¶ 35 (rejecting ECPR because existing retail prices may reflect monopoly rents and thus may cause access prices based on ECPR to be too high); Local Competition Order, *supra* note 37, at 15,859 ¶ 709 (same).

Graph theory suggests that this view of market-based pricing is too simplistic. Specifically, it suggests market-based pricing may yield access prices that are either too high or too low. The variability in the value of seemingly equivalent network components stems from the fact that the value of a network component varies with the particular network in which it is situated. In other words, an element in one network might have a very different value than the corresponding element in another network. Unlike the result suggested by conventional economic theory, which presumes that prices charged in concentrated markets will be too high, graph theory suggests that there may be no consistent bias in the direction of this effect. Thus, a graph theoretical approach surpasses present network theory by suggesting that regulators need to consider additional factors beyond market concentration in determining access pricing and access rules. Regulators should consider the effects of access regulations on network capacity and on the evolution of networks.

D. Compelled Access to Broadband Networks

Another striking development of recent years has been the displacement of narrowband services, in which users connect to the Internet through dial-up modems connected to conventional telephone lines, by broadband technologies, such as cable modem systems and digital subscriber lines ("DSL"). One notable difference between narrowband and broadband technologies is the role played by Internet service providers ("ISPs"). In the narrowband world, end users could use their telephone lines to contact their choice of any number of ISPs, such as America Online, MSN, Earthlink, Juno, or Netzero. In the broadband world, it is not unusual for cable modem or DSL providers to require their customers to use a designated, proprietary ISP.

The growth in the use of proprietary ISPs by broadband providers has prompted calls for the FCC to prohibit such exclusivity arrangements and to require that last-mile providers make their networks accessible to all unaffiliated ISPs on a nondiscriminatory basis.[66] This has emerged as a major issue in a number of recent mergers.[67] The FCC has sought comment on whether it should require cable modem and DSL providers to make their facilities available to unaffiliated ISPs on a nondiscriminatory basis,[68] a policy

---

[66] *See, e.g.*, Mark A. Lemley & Lawrence Lessig, *The End of End-to-End: Preserving the Architecture of the Internet in the Broadband Era*, 48 UCLA L. REV. 925 (2001).

[67] *See* Christopher S. Yoo, *Would Mandating Network Neutrality Help or Hurt Competition?: A Comment on the End-to-End Debate*, 3 J. ON TELECOMM. & HIGH TECH. L. 23, 39 (2004) (reviewing the fate of multiple ISP access proposals during AT&T's acquisitions of TCI and MediaOne, America Online's acquisition of Time Warner, and Comcast's acquisition of AT&T's cable holdings).

[68] *See* Inquiry Concerning High-Speed Access to the Internet Over Cable and Other Facilities, Declaratory Ruling and Notice of Proposed Rulemaking, 17 F.C.C.R. 4798, 4839–41 ¶¶ 72–74, 4843–47 ¶¶ 80–93 (2002) [hereinafter Cable Modem Declaratory Ruling and NPRM]; Appropriate Framework for Broadband Access to Internet over Wireline Facilities, Notice of Proposed Rulemaking, 17 F.C.C.R.

issue initially called "open access" and to which the FCC now refers as "multiple ISP access."[69]

To date, the debate has focused on the impact that mandating multiple ISP access would have on the levels of overall investment in new network technologies and on whether failing to mandate multiple ISP access would harm the markets for broadband content and applications.[70] What has been largely overlooked is how the interrelationships among individual components can cause multiple ISP access to have a complex and unpredictable impact on the overall network capacity and performance. To cite but one example, as noted earlier,[71] decisions about whether to use higher volume trunk-line technologies depend upon the ability to consolidate traffic. Only if the volume being carried is sufficiently large does the per-unit savings justify the incurrence of the higher fixed costs needed to deploy the trunk-line technology. Imposing multiple ISP access would have the effect of allowing an unaffiliated ISP to siphon off a portion of the traffic that would otherwise have flowed through the incumbent's network. The resulting reduction in traffic runs the risk of dramatically changing what would otherwise be the preferred network configuration. High volume technologies that would otherwise be cost-reducing become economically nonviable.

This effect is likely to be further exacerbated by the fact that once networks have been deployed, their configurations cannot easily be changed. As a result, broadband owners who deploy networks before multiple ISP access has been mandated might be left with an architecture they never would have created in the first place had they known that access would be mandated. This inevitably results in increasing the costs of providing network service. Higher costs, in turn, reduce the economic viability of various business models for Internet-based enterprises.

Graph theory's ability to reflect the interaction among cost, capacity, and reliability provides insights into the potential problems of multiple-ISP access that far exceed those provided by the current approaches to regulating communications network. It is this ability to capture the manner in which individual network elements interact with one another when placed

3019, 3042–43 ¶¶ 50–52 (2002); Inquiry Concerning High-Speed Access to the Internet Over Cable and Other Facilities, Notice of Inquiry, 15 F.C.C.R. 19,287, 19,298–306 ¶¶ 25–49 (2000); *see also* Yoo, *supra* note 67, at 40 (reviewing comments submitted during these proceedings asking the FCC to impose multiple ISP access).

[69] *See* Cable Modem Declaratory Ruling and NPRM, *supra* note 68, at 4839 ¶ 72.

[70] Our view is that the current debate has overstated the concerns underlying calls for multiple ISP access. The tendency for broadband providers to mandate the use of a proprietary ISP is more likely a reflection of the potential efficiencies inherent in the architecture of broadband technologies than of anticompetitive motives. *See* Yoo, *supra* note 67, at 33–34. Moreover, the overall structure of the market makes it implausible that any broadband provider would be able to adversely affect markets for broadband content and applications. *See id.* at 50–53. In addition, compelled access weakens investment incentives that are critical when emerging technologies are involved. *See id.* at 65; Spulber & Yoo, *supra* note 5, at 1020.

[71] *See supra* Part III.C.

within the context of an integrated system that remains one of the key aspects of network behavior.

E. The Regulation of Voice over Internet Protocol ("VoIP")

Internet telephony, sometimes called voice over Internet protocol ("VoIP"),[72] provides yet another illustration of graph theory's analytic potential. VoIP has been heralded as a transformative technological development. By offering the promise of turning every broadband platform into a potential provider of local telephone service, VoIP has the potential to increase the level of rivalry in an industry segment that has proven the most impervious to competition.

The advent of VoIP has inevitably raised questions about how it should be regulated. Although the FCC has issued a handful of early decisions in this area,[73] most of the key issues remain unresolved.[74] The area in which the FCC has offered a significant degree of guidance is with respect to the Communications Assistance for Law Enforcement Act ("CALEA"), which requires all telecommunications carriers[75] to design equipment, facilities, and services in such a way that allows properly authorized law enforcement agencies to conduct wiretaps.[76]

---

[72] Because this category of services encompasses services provided entirely over private data networks as well as over the public Internet, the FCC prefers the terms "voice over Internet protocol" or "IP telephony" to the term "Internet telephony." *See* Federal-State Joint Board on Universal Service, Report to Congress, 13 F.C.C.R. 11,501, 11,541 n.173 (1998) [hereinafter Stevens Report].

[73] *See* Petition for Declaratory Ruling that AT&T's Phone-to-Phone IP Telephony Services Are Exempt from Access Charges, Order, 19 F.C.C.R. 7457 (2004) (ruling that telephone calls initiated from a traditional handset using the standard telephone numbering system remains subject to traditional telephone regulation notwithstanding the fact that the long distance carrier converts the call into and transmits the call using an IP format); Petition for Declaratory Ruling that pulver.com's Free World Dialup Is Neither Telecommunications nor a Telecommunications Service, Memorandum Opinion and Order, 19 F.C.C.R. 3307 (2004) (ruling that services that simply facilitate peer-to-peer communication fall outside the scope of traditional telephone regulation); Stevens Report, *supra* note 72, at 11,541–45 ¶¶ 83–93 (distinguishing between "phone-to-phone" IP telephony, which remains subject to traditional telephone regulation, and "computer-to-computer" IP telephony, which constitutes an information service that is largely unregulated).

[74] The FCC recently began the process of determining how VoIP should be regulated by requesting comment on a wide range of VoIP-related issues. *See* IP-Enabled Services, Notice of Proposed Rulemaking, 19 F.C.C.R. 4863 (2004).

[75] "Telecommunications carrier" is a defined term under CALEA that differs from the definition established by the Communications Act of 1934. *See* 47 U.S.C. § 1001(8)(A)–(C) (2004).

[76] *Id.* § 1002(a). Consistent with its initial VoIP precedents, the FCC distinguished between "managed" VoIP services (in which the VoIP provider acts as an active mediator between end users by performing such critical functions as call set-up, switching, addressing, or routing) and "nonmanaged" VoIP services (in which the VoIP provider plays little or no role in routing and transmitting packets between end users and instead serves primarily as a centralized directory that helps users establish peer-to-peer communications). The FCC tentatively concluded that because managed VoIP services serve as "a replacement for a substantial portion of the local telephone exchange service," *id.* § 1001(8)(B)(ii), they are subject to CALEA. At the same time, the FCC tentatively concluded that nonmanaged VoIP services fall outside the scope of CALEA, either because they are similar to private networks, in that they

Graph theory provides several basic intuitions about the likely impact of applying CALEA to VoIP. Currently, the packets associated with VoIP traffic are indistinguishable from the packets associated with other applications and can travel by any of the available routes through the Internet. Subjecting VoIP to CALEA would force all Internet-based voice traffic to pass through a discrete number of gateways. The creation of such bottlenecks will have the inevitable effect of reducing the overall capacity of the network, since mandating that all VoIP traffic traverse designated gateways would preclude the network from employing a large percentage of the feasible paths available to interconnect two nodes. Graph theory also reveals that applying CALEA to VoIP would likely degrade the quality of VoIP service, since forcing traffic to pass through a limited number of discrete checkpoints would tend to create queues that would limit the ability of such networks to guarantee particular throughput rates. The resulting delays would weigh particularly heavily on voice communications, in which delays of as little as a quarter-second have been deemed unacceptable. CALEA requirements would also reduce the viability and profitability of VoIP services. In turn, this would frustrate the emergence of competition in local telephone service, which remains one of the central goals of modern telecommunications policy.[77]

F. A Formal Methodology for Calculating Regulatory Rates

Finally, graph theory offers the promise of providing a basis for calculating the economic impact of particular regulatory interventions. Although the problems posed by graph theory can be quite complex, algorithms and modern improvements in computer processing have rendered these problems more tractable. Simply applying a graph theoretical framework yields basic intuitions that offer a valuable contribution, even if the precise magnitude of that effect is never quantifiable, as demonstrated by the other policy examples discussed above. Understanding the nature of networks as complex systems is highly valuable in avoiding policy mistakes that stem from the representation of networks as collections of unconnected elements. At a

---

are confined to a limited universe of users, or because they constitute "information services." Either determination would be sufficient to take nonmanaged VoIP services outside the scope of CALEA. *See* Communications Assistance for Law Enforcement Act and Broadband Access and Services, Notice of Proposed Rulemaking and Declaratory Ruling, 19 F.C.C.R. 15,676, 15,693–95 ¶ 37, 15,708–10 ¶¶ 56–58 (2004) (explaining the distinction between managed and nonmanaged VoIP services).

[77] The FCC's tentative decision to distinguish between managed and nonmanaged VoIP services would have similar, albeit less sweeping, effects. The graph theoretical analysis that we propose suggests that by subjecting VoIP systems in which traffic flows in a more distributed manner to a lower degree of regulation than VoIP services that route traffic through centralized gateways, implementing CALEA in this manner would induce a regulatory bias toward nonmanaged systems that might have effects on network performance and capacity that go far beyond what conventional analyses might lead one to anticipate. Indeed, the CALEA-related requirements could interact with the configuration of particular networks or other aspects of the regulatory regime in a way that has effects that can be as dramatic as they are unpredictable.

minimum, the study of the behavior of complex systems suggests the need for regulatory forbearance and greater reliance on market solutions in network industries.

We acknowledge that a large number of practical obstacles remain before graph theory can become the basis for a regulatory regime. As an initial matter, the precise impact of any regulatory intervention would vary heavily depending on the particular network involved and the amount of capacity accessed. Indeed, each network and each access request would likely yield a different answer. Furthermore, the problems caused by the indivisibility of fixed costs and by variability in demand may make excess capacity appear endemic to network industries, which in turn may cause the prices calculated in accordance with graph theory to systematically underestimate the economic impact of granting access.[78]

In addition, regulatory authorities attempting to implement a graph theoretical pricing regime may depend on the network owner to provide information about the network. Disclosure of information about network architecture creates administrative costs for network operators and runs the risk of giving away competitive advantages. Access regulation also increases the likelihood that incumbent firms will be hurt by the strategic behavior of competitors.[79]

That said, these obstacles can prove surmountable. Indeed, the same problems long have confronted—and have been overcome by—regulatory authorities charged with establishing rates for telecommunications services. "Cost proxy models" are one example of a recent development in conventional ratemaking approaches that demonstrates how graph theoretical models may not be as analytically and informationally intractable as they may first appear.[80] These models, which have been used in a number of recent federal and state rate proceedings, avoid relying on network owners for information about how their networks are configured. Instead, cost proxy models base their calculations on computer simulations of how best to deploy a network given the particular technological processes employed. The more sophisticated versions of these models contain built-in optimizing routines that effectively rebuild the network from scratch each time in order to

---

[78] *See supra* note 51 and accompanying text.

[79] *See supra* note 50 and accompanying text.

[80] For the seminal work on cost proxy models, see D. GABEL & M. KENNET, ESTIMATING THE COST STRUCTURE OF THE LOCAL TELEPHONE EXCHANGE NETWORK (Nat'l Reg. Res. Inst., Working Paper No. 91-16, 1991); BRIDGER M. MITCHELL, INCREMENTAL COSTS OF TELEPHONE ACCESS AND LOCAL USE (Rand Corp., Working Paper No. R-3909, 1990). For an overview, see Sharkey, *supra* note 9, at 204–20. For more recent elaborations, see FARID GASMI ET AL., COST PROXY MODELS AND TELECOMMUNICATIONS POLICY (2002); DAVID M. MANDY AND WILLIAM W. SHARKEY, DYNAMIC PRICING AND INVESTMENT FROM STATIC PROXY MODELS (FCC Off. of Strategic Plan. & Pol'y Analysis, Working Paper No. 40, Sept. 2003), *available at* http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-238934A2.pdf.

determine its best configuration.[81] Once the optimal configuration has been determined, the costs provided by each element can be determined through an analysis of the applicable engineering processes. Using cost proxies in this manner obviates the need to rely on the network owner to provide all of the relevant cost data. For regulatory purposes, networks should not be compared to a hypothetical most-efficient network, but models of the network can be adjusted to reflect properties of the regulated firm's existing network.

In short, the increasingly widespread use of cost proxy models in more conventional ratemaking proceedings suggests that graph theoretical problems, while complex, can form a workable basis for network policy. Although the informational requirements needed to perform a graph theoretical are quite onerous, they are no more demanding than the information required to determine firm-specific tariffs under cost proxy models. The shift to a graph theoretical approach would simply change the optimization routines embedded in the relevant computer programs. The ultimate result may employ different algorithms and may require more iterations, but the problem would not seem to be any less tractable.

Thus, there is reason to think that graph theory could be directly operationalized into a fully working regulatory methodology that makes it possible to calculate access prices in a way that reflects the complexity of interactions between network components. Although the difficulties in implementing the graph-theoretical regulatory regimes discussed above may ultimately cause the calculation to fall short of a true, first-best outcome, it is quite possible that graph theory may provide a better second-best approximation than that provided by conventional approaches. If such a regime proves unworkable, the complexity of the interactions captured by our analysis counsels strongly in favor of regulatory forebearance. The individualized decisionmaking inherent in private ordering is better suited to implementing the context-specific approach required by graph theory than is the centralized decisionmaking process inherent in regulation.

VI. CONCLUSION

The conceptual approach set forth in this Essay has the potential to revolutionize the way policymakers and academics conceptualize network policy by providing a basis for describing and analyzing network architec-

[81] The reality that communications networks often require sunk costs in particular locations may require some feasibility limitations on some of the possible hypothetical networks. TELRIC takes such considerations into account by basing rates on "the lowest cost network configuration, *given the existing location of the incumbent LEC's wire centers*." 47 C.F.R. § 51.505(b)(1) (2004) (emphasis added). For purposes of cost recovery and some types of access regulation, however, it may be necessary to examine the existing configuration of loops and switches beyond the wire centers, as pointed out in J. Gregory Sidak & Daniel F. Spulber, *Givings, Takings, and the Fallacy of Forward-Looking Costs*, 72 N.Y.U. L. REV. 1068, 1141 (1997). Graph theory can help regulators analyze the effects on costs of changes in network architecture.

ture that captures the extent to which networks are more than just the sum of their component parts. It reveals how the relationship between different network elements causes them to interact with one another in ways that are often surprising and unpredictable. It also reveals that the current regulatory approach, which treats each network element as if it existed in a vacuum, fails to provide an accurate reflection of the impact that regulation has on the network as a whole. Most importantly, it demonstrates how access requirements can interfere with network owners' ability to design optimal network architectures. Although the application of graph theory to specific problems is not without its difficulties, we believe that the approach we propose provides sufficient insights and intuitions into network behavior to justify employing it as a tool of policy analysis. Our analysis demonstrates the importance of shifting to an analytical framework that can capture the complex interrelationships among the elements of networks. Understanding networks as complex systems should provide policymakers with a clearer picture of the effects of network architecture on usage patterns. This is useful for examining the effects of regulation on the evolution of networks.