THE ROLE OF THE REAL CITY IN CYBERSPACE: UNDERSTANDING REGIONAL VARIATIONS IN INTERNET ACCESSIBILITY[†]

Mitchell L. Moss
Taub Urban Research Center
New York University
mitchell.moss@nyu.edu

Anthony M. Townsend
Department of Urban Studies and Planning
Massachusetts Institute of Technology
amt@mit.edu

I. INTRODUCTION

Since 1993, when the first graphical web browser, Mosaic, was released into the public domain, the Internet has evolved from an obscure academic and military research network into an international agglomeration of public and private, local and global telecommunications systems. Much of the academic and popular literature has emphasized the distance-shrinking implications and placelessness inherent in these rapidly developing networks. However, the relationship between the physical and political geography of cities and regions and the virtual (or logical) geography of the Internet lacks a strong body of empirical evidence upon which to base such speculation.

This chapter presents the results of a series of studies conducted from June 1996 to August 1998. Our research suggests there is a metropolitan dominance of Internet development by a handful of cities and regions. We identity and describe an emerging structure of "virtual" hubs and pathways which are linking a set of major cities in the United States, suggesting that there is a complex emerging inter-urban communications network that goes far beyond Castells' (1989) informational mode of development.

More importantly, we analyze the utility and relevance of three measurements of Internet development for urban planning and policy research. The first measure is the number of computers connected to the Internet on a full-time basis, based on data collected by Matrix Information and Demography Services of Austin, Texas. We find this measure to be of limited usefulness, as it only measures location-specific hardware installations with little reference to their purpose or function. The second measurement system we explore is based on the Internet's addressing scheme, known as the domain name system. Unique to individual organizations (business, education, non-profit, government), domain names are registered with InterNIC, an administrative clearinghouse contracted by the National Science Foundation. Associated with each domain name is a unique billing address, which permits the localization of the organization using that name. This is the most informative measure for assessing variations in Internet use across a variety of geographic units, from states to individual ZIP codes. Finally, we examine the capacity and topography of nationwide Internet backbone networks that transport data between metropolitan areas.

Based on these three measurements, we find that a limited number of cities am metropolitan areas dominate the rapidly emerging telecommunications landscape of the United States, leading in the development of increasingly sophisticated applications and technologies. Accessibility to the most highly developed real and virtual Internet infrastructure is a metropolitan phenomenon, and highly stratified among regions and cities. Furthermore, we describe each of the three measurements used – host counts, domain counts, and backbone network capacity – and their unique advantages and disadvantages for research. Successfully applying them to urban analysis requires not only an understanding of urban and regional development processes, but also the purpose, design, and function of these complex technical systems.

[†] Originally Published in *Information, Place, and Cyberspace: Issues in Accessibility*. D.G. Janelle and D.C. Hodge (eds.).

II. CITIES, REGIONS, AND TELECOMMUNICATIONS

The Internet and other telecommunications advances pose a serious challenge to the study of urban life. Electronic commerce and the decline of distance-sensitivity in telecommunications pricing have encouraged speculation that the dispersion of human settlement is imminent. However, as Peter Hall (1997, 316) states, 'the urban world of the 1990's…is a world in which cities deconcentrate and spread to become complex systems of cities linked together by flows of people and information.' Thus, while advanced telecommunications permits the evolution of an increasingly complex urban system with multiple linkages and hierarchies, place and centrality remain extremely important in the information economy. Gaspar and Glaeser (1996) present strong evidence that telecommunications and travel are synergistic, suggesting the intrinsic value of face-to-face interaction beyond what can be communicated at a distance. The Internet has the potential to subsume other communications media such as telephony, print, television and radio, as well routine activities like shopping, learning, entertaining, and socializing.

We have sought to quantity and localize various measurements of Internet development in an effort to understand the relationship of the Internet to urban and regional development. Hall (1997, 318) asserts that cities' competitiveness in the global economy 'depends on their capacity to generate, process, and exchange information'. To address those concerns, our measurements of Internet development seek to describe the emerging structure of systems by which cities exchange information through electronically mediated communications. Just as the Interstate Highway System transformed urban development in 20th century America, the Internet will help shape urban activity patterns in the 21st century.

In the past, researchers have studied the flow of information in colonial American cities through newspapers (Pred 1973), by telephone in the urban complex of the Northeastern United States (Gottmann 1961), and in the information economy by means of office buildings and overnight letter delivery (Sui and Wheeler 1993; Mitchelson and Wheeler 1994). More recently, others have tried to quantify the informational capacity of cities and regions, identifying clusters of both human (Nunn and Warren 1997) and physical (Greenstein, Lizardo, and Spiller 1997) information age capital. However, with the notable exception of Dodge and Shiode's (2000) recent work on Internet real estate in the United Kingdom and Murnion's work (see Chapter 12), little research has yet to directly address the geography of the evolving Internet.

There are several reasons why the Internet has eluded urban scholars, at least in the United States. One is that the telecommunications industry in general, and particularly the Internet sector, is extremely competitive. Because the Internet services sector emerged during a period of rapid deregulation in the American telecommunications industry, it has grown unsupervised by federal regulators, and there is a lack of systematically gathered data on its operations. While individual companies may possess substantial information that could be used to advance scholarly research or public understanding of these new networks, there are few available sources of geographic data. Second, the Internet operates primarily over pre-existing telephone network infrastructure. With the exception of Qwest. Inc.'s new national fiber optic network, there is little physical construction activity solely associated with the deployment of new Internet infrastructure. Rather, the flip of a switch to light a dark strand of glass fiber is all that is required to deploy new capacity. Compare this to the spread of the cellular telephone network, which can be measured either visually by counting towers or by consulting the U.S. Federal Communications Commission's (FCC) extensive public database of transmission antennas.

This chapter describes three basic measurements used over a two-year period to attempt to assess Internet accessibility among cities and metropolitan areas in the United States. The difficulty of obtaining reliable data at comparable levels of geographic aggregation limit the usefulness of presenting the results together, further highlighting the need for new empirical tools and methods to be developed. However, we seek to identify methods of measuring Internet development and diffusion to guide further empirical research in this area, and stimulate debate about the implications of these observations and methods for urban theory.

III. MEASURING REGIONAL VARIATIONS IN INTERNET DEVELOPMENT

The primary question this series of studies seeks to address is which cities and metropolitan areas exhibit a rapid buildup of Internet-related telecommunications infrastructure, and secondly, how can we effectively measure the geographic distribution of the Internet? Measuring this phenomenon proved very difficult. First, as noted earlier, telecommunications providers have strong incentives to keep data on their operations closely guarded. Second, from a very early stage in its development the Internet was designed to subvert geography by using a packet-switched message routing system, which operates more like the postal service than the telephone network. As a result, messages can be rerouted around faulty switches or stations. Thus, the few aspects of the Internet that can be localized and measured have only limited relevance to physical geography. Third, each of the three data sources measured aspects of the Internet that did not always have overlapping geography. Finally, we needed to consider what these measurable criteria could indicate about the level of human activity on the internet in different cities and regions and the socioeconomic consequences of these information flows. This section describes the three measurements we used and the strengths and limitations of each. The final section presents a composite picture of the North American system of cities we have derived from these observations.

Computers Connected to the Internet on a Full-time Basis

Matrix Information and Directory Services of Austin, Texas has measured the geography of the Internet since the early 1990's, and is considered a leader in analyzing what it calls the matrix of inter-connected global computer networks. MIDS' methods permit the localization of Internet hosts (computers connected to the Internet on a full-time basis) to geographic areas as specific as a street address. Our study was based on data provided by MIDS for January 1996. and indicated the number of Internet hosts per county for all 50 states. The data included a substantial number of manual corrections to the automated survey that generates the data, based on MIDS' proprietary knowledge of the known geographic location of large clusters of host computers in corporate research and development centers. for example.

This measurement was a useful first step towards visualizing the emerging telecommunications landscape of the United States. As anticipated, Santa Clara County, California (Silicon Valley) and Middlesex County, Massachusetts (Route 128) were the largest clusters of permanently connected Internet hosts. Table 1 shows the top 25 counties by number of host computers. However, we did not anticipate the large number of central cities that appear on this list. Furthermore, 12 of the 25 top counties were located within just four metropolitan areas: San Francisco, Los Angeles, Washington DC, and New York. An asterisk indicates these twelve counties. While these results confirmed our suspicions regarding the concentration of high levels of telecommunications activity in a select group of metropolitan areas, the overwhelming presence of Silicon Valley, Route 128, and several academic clusters in Michigan forced us to re-evaluate the relevance of these findings. We believed these high figures to be artifacts of these regions' role as the birthplace of the Internet in academic and industrial settings. Our interest thus turned to identifying locations where a broad-based adoption of these technologies had rapidly occurred across a wide range of industries and population groups.

Table 1. Top 25 counties by number of Internet hosts, January 1996

County, State	Description	Hosts
*Santa Clara, CA	Silicon Valley	554,967
Middlesex, MA	Route 128	243,765
*Los Angeles, CA	Central City	159,944
*New York, NY	Central City	146,371
*Fairfax, VA	Edge City	131,874
*Orange, CA	Edge City	123,685
*San Diego, CA	Central City	111,981
Cook, IL (Chicago)	Central City	110,726
Hennepin, MN (Minneapolis)	Central City	109,047
*San Mateo, CA	Edge City	92,781
Salt Lake, UT	Central City	90,693
*Alameda, CA	Edge City	89,851
Washtenaw, MI	Universities	82,790
King, WA (Seattle)	Central City	79,142
Allegheny, PA (Pittsburgh)	Central City	64,616
Philadelphia, PA	Central City	62,387
Travis, TX (Austin)	Central City	62,371
Dallas, TX	Central City	61,811
Bay, MI	Universities	57,726
*District of Columbia, DC	Central City	55,755
*San Francisco, CA	Central City	53,183
*Prince George's, MD	Edge City	46,450
*Montgomery, MD	Edge City	42,156
Mecklenburg, NC (Charlotte)	Central City	37,369
Fulton, GA (Atlanta)	Central City	34,103

Source: Moss and Townsend 1996
* indicates that county is located within the metropolitan areas of New York, Los Angeles, San Francisco or Washington, DC.

At this point, we encountered the first of many instances in which the technological realities of the Internet drastically affected the interpretation of our results. The measurement of host counts was extremely coarse because it failed to differentiate between various types of computer equipment. For example, if a financial services company in Manhattan uses a highly centralized computer system based on a mainframe and dumb terminals, this method might only count a single host – the mainframe being the only machine directly connected to the Internet. Conversely, a small software company in Silicon Valley that uses a Local Area Network to connect its dozen microcomputers to the Internet individually would have a much higher host count. In some cases, networked printers might even be counted by this method. Finally, the increased use of firewalls, computers designed to mediate external Internet connections and shield institutional networks from intruders, excludes a significant number of hosts from detection, especially those of large corporations. These factors generate significant variations in the effectiveness of comparing host counts across regions and industrial sectors.

It also seemed counter-intuitive to use a measurement based purely on the technical organization of the Internet to infer some understanding of the rate of adoption of these technologies across regions. The nature of the modern American economy and the production systems of the Internet services industry further complicates this approach. Many organizations do not physically house their Internet-accessible information at their physical location, preferring to hire contractors who provide expertise and equipment.

While the information-producing jobs and economic activity associated with a website may take place at a centralized office in a dense urban area. it is just as likely that the fruits of this labor are electronically disseminated from a remote location, which could conceivably be located anywhere. These limitations in the host count measurement led us to seek other indicators of Internet use, which proved more useful in understanding the spread of the Internet among cities.

The Location of Organizations Using the Internet

Domain names are one of the basic forms of Internet addressing, which map groups of numeric Internet addresses to intuitive names like `www.nyu.edu` or `www.att.com`. Each domain name is registered with Network Solutions, an organization chartered by the National Science Foundation to administer the domain name system. As Figure 1 shows, each name is registered to an individual or organization. and the publicly available registration record contains a billing address for that entity. From this information, it is possible to localize the location of the entity that owns that domain name as specifically as the postal code (ZIP) level. This geographical specificity of the domain name makes it a highly attractive measure for Internet activity. Network Solutions has enjoyed a monopoly over domain registrations for the most popular commercial, non-profit, educational, and government domains in the United States, leaving only a small portion of the American Internet beyond the scope of these data.

The strengths of this measurement for urban research stem from its representation of a social phenomenon, rather than a technical one. Because each domain name roughly corresponds to a corporate, government, or educational entity, this measurement indicates spatial variations in the adoption and use of Internet-based communications by organizations. Since nearly 90 percent of Internet growth over the period between 1994 and 1997 was from the addition of commercial domain names, these results primarily measure the extent to which businesses deployed these new technologies. Domain registrations also indicate the date each domain was first registered, permitting us to identify those regions that had the most rapid growth in Internet use.

Figure 1. A typical domain name registration record

Registrant:	`Five Points Internet Solutions`

Address:	`45 Havemeyer St. #2R`
	`Brooklyn, NY 11211`
	`United States`

Domain Name:	`www.fivepoints.net`

However, the domain measurement is seriously handicapped as well. First, it does not take into account an organization's size or its dependence upon or capacity to generate flows of information over the Internet. There is no practical method for sorting through the hundreds of thousands of records and assigning weights to entities of differing sizes, revenues, or information processing and production capacity. As a result, this technique weights Microsoft's immense Internet presence little more than the small website maintained by Redmond, Washington's municipal government. Also, the geographic data associated with domain registrations do not always correspond to the true physical location of a domain's primary users (who may be dispersed over multiple continents), but rather to an administrative or MIS headquarters location. For example, while AT&T is headquartered in New York City, its domain registration for `www.att.com` is in Florida. Furthermore, most of its data networking operations are controlled from a center in the St. Louis, Missouri area. Finally, companies are increasingly registering multiple domain names: those of the company's products, or variations upon the company's name. This practice may be responsible for some distortion in the overall results.

The primary source of data for our research on the geographic distribution of domain names was Imperative! of Pittsburgh, Pennsylvania. Imperative! sells mailing lists to direct marketers who offer Internet-related products to the owners of domain names and maintains a database of currently registered domain names that can be aggregated by nearly any geographical unit. We analyzed regional variations in domain name registration at three geographic levels; (1) U.S. Census regions, (2) 85 major U.S. cities, and (3) postal code areas in New York City (Moss and Townsend 1997a, 1997b, 1998). Among major U.S. cities, the largest clusters (Manhattan and San Francisco), were also the most densely networked. Furthermore, in the 15 cities with the largest number of Internet domains, which accounted for 12.6 percent of all U.S. domain registrations in April 1994, new domains were registered faster than other areas. By 1997, these 15 cities accounted for nearly one-fifth (19.7 %) of all U.S. domain registrations. Clearly, Internet technologies were being more rapidly deployed in major urban areas. Outside the dense urban areas of the Atlantic, Pacific, and Gulf coasts, only Phoenix and Chicago had a significant number of domains.

We also computed the density of domains with respect to population, or domains per 1000 persons. Domain density was typically highest among cities whose primary function was as a resort, government, or education center. Nodal cities also showed high concentrations (Moss and Townsend 1998).

Comparing these results to a limited set of domain counts for approximately a dozen cities in April 1994 permitted us to track the rate of domain growth over a 3-year period. These cities registered domains far faster than the national average. adjusted for population. The growth rate of domains between 1994 and 1997 was linked to a city's relative position in the national urban system, with the advanced service centers growing most rapidly. This indicates a strong relationship between information-intensive economic activity and early adoption of the Internet among businesses. The cities primarily fall into four broad categories of growth rates. summarized in Table 2.

Table 2. Trends in growth of domain name registrations, 1994-1997.

Description	Examples	Domain growth rate (Multiple of national avg.)
Global information centers	Manhattan (NYC), San Francisco	6x
Mid-sized information centers	Atlanta, Boston, Miami, Seattle	4-5x
Regional Centers	Denver, Dallas, Phoenix	2-3x
World cities	New York, Los Angeles, Chicago	1-2x

Source: Moss and Townsend 1998

The fact that the set of cities most commonly referred to as world cities or global cities recorded the slowest growth rates among large Internet clusters is disturbing, for it indicates an averaging function. As our detailed analysis of domain registrations in Manhattan indicates, adoption of the Internet is not a widespread phenomenon across urban populations. Rather, it is almost entirely limited to the central business districts, with moderate adoption rates in the more successful immigrant communities (Moss and Townsend 1997a). The fact that smaller cities such as Austin or Boston exhibit higher growth rates is most likely due to a more even spread of technological opportunities among their more homogeneous populations. On the other hand, world cities appear to be characterized by a digital elite co-existing with a vast, largely disconnected information ghetto. As one example of this disturbing trend, a preliminary survey of domain registrations in the Los Angeles area indicates a much slower diffusion of Internet technologies among Spanish-speaking and immigrant communities.

Jed Kolko, a doctoral student at Harvard University, is currently conducting research that will extend the analysis of domain name registrations to all 285 U.S. Metropolitan Statistical Areas (MSAs) over a four-year period from 1995-1998. This work addresses a shortcoming of our research, in that we neglected to explore Internet adoption and use in suburban areas surrounding the major cities. In fact, the bulk of new office space in recent decades has emerged not in the central cities that we focused upon. but rather in the edge cities that surround them. (Garreau 1991) Some preliminary results, in Table 3, show that even where including the surrounding metropolitan areas there is strong evidence of a select group of cities that dominate Internet activity.

Table 3. Domain name registrations by metropolitan area: January 1998.

Consolidated Metropolitan Stastical Area	Domains, January 1998	Percent of all U.S. Domains
New York	112,524	8.6
Los Angeles	109,917	8.4
San Francisco/Silicon Valley	89,584	6.8
Washington, DC	43,766	3.3
Boston, Route 128	41,736	3.2
Chicago	38,447	2.9
Philadelphia	28,693	2.2
Miami	27,993	2.1
Dallas	26,520	2.0
Seattle	25,238	1.9
These 10 Metropolitan Areas	544,418	41.4
Rest of U.S.	771,393	58.6
Entire U.S.	1,315,811	100.0

Source: Kolko 1998

Once again, the four metropolitan areas that contained the largest clusters of Internet hosts, and the densest concentrations of domains, also account for the largest nodes of Internet activity by this measure.

Internet Backbone Networks

While it was important to identify centers of Internet activity and variations in the concentration of Internet indicators among American cities, we also need to quantify the flows of information between cities to understand how these networks are developing within the American urban system. Communications on the Internet is primarily carried over fiber optic networks, portions of which have been adapted from their original use (the transmission of voice telephone calls), although the first dedicated networks optimized for the TCP/IP protocol are now under construction. Traversing the country along traditional rights-of way, such as railroad tracks, interstate highways, and even abandoned canals, these networks are the physical manifestation of the 'information superhighway.' Like the host and domain-name measurement, our analysis of the conglomeration of national data networks collectively known as the Internet 'backbone' indicates a high degree of centralization in the deployment and use of Internet technologies.

Again, the technology in question dictated careful interpretation of the results. While the host count described a physical, real phenomenon (the connection of computers to the Internet), and the domain count a thoroughly virtual, logical one (the organization of Internet Protocol addresses into convenient hierarchies), the measurement of backbone capacity is a hybrid. Although some networks operate on isolated fiber optic cables, many are merely virtual networks operated over lines leased from national and regional telephone companies. Often, a backbone provider's only capital equipment are the powerful routing computers that manage the flow of data packets at network junctions (Rickard 1997).

While the geography and topography of these networks has received more attention from scholars than the identification of nodes of Internet activity, no other studies have analyzed the aggregate topology and capacity of the major backbone networks (Moss and Townsend 1998). Using maps and data from Boardwatch Magazine's Quarterly Directory of Internet Service Providers, we compiled a list of the capacity and endpoints of every major backbone link for 29 major backbone operators in the United States. We estimate that these 29 wholesale providers supply at least 95 percent of long-haul Internet data transport services in the United States. Based on the assumption that barriers to backbone accessibility were likely to be found at the inter-metropolitan, rather than intra-metropolitan level, we further aggregated these links by metropolitan area. The existence of networks, such as Metropolitan Fiber System's Metropolitan Area Ethernets, and the rapid proliferation of the baby Bells' high-speed regional fiber networks underlies this assumption. This aggregation allowed us to focus on the largest capacity fiber optic networks, constructed of DS-3 (45 Mbps), OC-3 (155 Mbps) and OC-12 (622 Mbps) technology. One megabit per second (Mbps) indicates a data transfer rate equal to approximately 128 pages of text per second.

The analysis is also confined to direct network connections. Theoretically, any city on a given network has access to all locations served by all networks. However, there is significant traffic congestion of data packets at inter-network gateways (bridges), and there are strong indications that providers have established direct links on the most highly trafficked inter-metropolitan routes. For example, several providers have established direct links between New York and Washington, D.C., even though they already operate a route connecting these two metropolitan areas through intermediate cities, such as Baltimore, Philadelphia or Wilmington, Delaware. The San Francisco Bay Area is directly linked to almost every metropolitan area in the United States, although many could presumably have been served indirectly through another node. These patterns of investment indicate the superiority of direct connections and their importance to a metropolitan area's ability of to import and export information via the Internet. The results of the backbone analysis were the most striking and conclusive of the three measurements we used. Summarized in Table 4, the data show that a group of seven metropolitan areas (San Francisco/Silicon Valley, Washington, D.C., Chicago, New York, Dallas, Los Angeles, and Atlanta) form a core group of urban areas that dominate the Internet in the United States.

Table 4. Top 10 metropolitan areas by backbone capacity

Metropolitan Area	Percent of total national backbone capacity	Total inter-metropolitan backbone capacity (in Mbps)
San Francisco/Silicon Valley	11.6	7,506
Washington, DC	10.4	7,826
Chicago	9.8	7,663
New York	9.7	6,766
Dallas	7.1	5,646
Los Angeles	6.7	5,056
Atlanta	6.6	5,196
Denver, CO	3.7	2,901
Seattle, WA	2.5	1,972
Houston, TX	2.4	1,890

Source: Moss and Townsend 1998

These seven metropolitan areas each have the capacity for over 5,000 megabits per second (Mbps) of Internet data throughput, sufficient to transfer text across the Internet at a rate of over 640,000 pages per second. No other metropolitan areas approach this level of capacity. Denver, ranked eighth, has only 60 percent of the backbone capacity of seventh-ranked Atlanta. As a result, these seven metropolitan areas share 62.0 percent of the nation's backbone capacity. The next 14 metropolitan areas together account for an additional 25.5 percent of the nation's backbone capacity, while the remainder of the United States houses the remaining 12.5 percent. Outside the major metropolitan backbone hubs, communities are linked to the Internet backbone via less robust data pipelines such as DS-1 lines (1.5 Mbps, also known as T-l), frame relay (0.05 to 0.25 Mbps), ISDN (0.125 Mbps), and modem (0.025 to 0.5 Mbps) lines, which substantially limits their ability to move large amounts of information quickly. Many of these technologies. especially the popular DS-1 lines, have notoriously distance-sensitive price structures that have severely restricted their proliferation in non-metropolitan areas. These seven metropolitan Internet hubs also connect directly to a great variety of other metropolitan areas and cities. The San Francisco Bay Area has the largest number of external connections, with 153 links to dozens of other metropolitan areas. Washington, Chicago, and New York have over 100 links maintained by various network companies to other cities. Dallas, Los Angeles and Atlanta each have 75 or more direct external links. By contrast, the next 14 metropolitan areas each have 42 or fewer external linkages. The variety of linkages associated with the top seven metropolitan areas is yet further evidence of their key role as central switching centers for flows of information on the Internet.

The most striking finding of this analysis is that among the top seven metropolitan areas, the majority of backbone capacity was used to link these regions to each other or to provide service within the region, rather than to connect less important cities and outlying areas. This aspect of the emerging backbone network merits further research as it suggests that these metropolitan areas do not serve as conduits for a national system of data distribution, but instead have coalesced into a separate, highly networked urban system containing both the major producers and most consumers of Internet services. This configuration is very similar to observations about so-called 'global cities' like New York, London, and Tokyo, which have become largely disconnected from their national economies while being increasingly integrated with each other.

The backbone analysis was the most complete and accurate of these three analyses because the data set involved was more limited and manageable with existing analytic techniques. However, like the domain and host count measures, it should serve only as a general indicator of relative differences in magnitude of Internet accessibility and adoption among cities, regions, and metropolitan areas. Notwithstanding, the stark difference in backbone capacity between the top seven metropolitan areas and the rest of the nation was the clearest conclusion drawn from this series of studies.

IV. CONCLUSIONS

The research on the geography of the Internet summarized in this paper was conducted over a two-year period from 1996 to 1998. Despite substantial difficulty in obtaining accurate, timely, and comprehensive data sets, our results indicate several important trends, as well as questions for future research. When compared, the three sets of data indicate a consistent metropolitan dominance of the Internet in the United States. This section summarizes our collective findings and observations.

Most significantly, the results of this research demonstrate conclusively that a select group of metropolitan areas, and in particular their central cities, overwhelmingly dominate the Internet in the United States. Above and beyond their status as centers of population and employment, these regions consistently lead the nation in the magnitude, density, and growth of Internet clusters. This fact calls into question many assumptions regarding the spatial effects of information technology and telecommunications, as many influential authors have promulgated a deterministic view that new technologies will lead to a radical decentralization of population and economic activity (Gilder 1995, Negroponte 1995, Toffler 1980). This view is not borne out by our research. The New York, San Francisco, Washington, and Los Angeles metropolitan areas that are the largest clusters of Internet activity appear to be employing these technologies to reassert their economic importance in the American urban system. Furthermore, the presence of a set of smaller, Internet-savvy metropolitan areas, including

Austin. Boston, Miami, and Seattle, suggest that rather than causing decentralization, the Internet may permit the development of a more complex, networked urban system.

These studies also indicate the lingering importance geographical location in determining accessibility to advanced telecommunications services. Cities like Atlanta, Chicago, and Dallas are extremely important hubs for national Internet backbone, networks out of proportion to their share of other measures of Internet activity. We infer that their geographic centrality is an important factor in this allocation of network capacity. As a result, these cities have access to more, faster, and less expensive Internet infrastructure than cities of comparable size that are less ideally located. However, the absence of comparable levels of Internet development in other centrally located cities, such as Detroit and Philadelphia, lead us to believe that a certain set of socioeconomic prerequisites influence the allocation of these technologies.

In our effort to measure and describe the emerging telecommunications landscape of the United States, we anticipated being able to measure flows of information across the Internet to gain a greater understanding of which regions were net information producers, and which were net information consumers. Abler (1970) and Mitchelson and Wheeler (1994) used this methodology in the past to determine the relative position of cities within an urban hierarchy. While we were unable to gather such data, future research may benefit from a proposal to integrate geographic information in the Internet's domain name system. The Request For Comment 1976 proposal (RFC 1976). *Putting Locations in the Domain Name System*, would theoretically permit researchers to measure the flow of information across monitoring stations along the Internet in real time and to more accurately determine the location of host computers than is currently possible (Davis 1998).

Perhaps the most important conclusion we have made is that researchers studying the Internet need to be very sensitive to the technical intricacies of the systems they investigate. The relevance of the host count measurement was seriously undermined by several factors discussed earlier, such as the use of firewalls to mask corporate networks. Without a clear understanding of exactly how the Internet is constructed, we probably would have accepted those results beyond their actual significance. Furthermore, because much of this infrastructure is actually an array of intangible data and logical constructs (domains, virtual backbones) or easily reconfigured electronic equipment (host computers and fiber optic networks), means it can be reallocated almost instantly in response to market shifts, natural disasters, etc. By definition, the Internet is highly volatile and in constant flux. Research runs the risk of being an anachronism before it is ever published. For cities, this means that unlike the televised urban riots that launched Lyndon Johnson's War on Poverty or the graphic images of Rust Belt cities in decay that helped put Ronald Reagan in the White House, Internet communities can disappear with the flip of a switch.

Finally, there is a strong need for systematic data gathering on the geography of Internet development and information flows. The RFC 1976 proposal is a good candidate, and we hope its key points will be incorporated in the new version of the basic Internet Protocol, IPv6, now being designed.

One of the many interesting ideas to emerge from the NCGIA Project Varenius specialist meeting, from which the chapters in this volume are drawn, was the concept of a federally funded Internet Census. Historically, Congress has often charged the Census Bureau to conduct exhaustive surveys on the emerging telecommunications industries. A full twelve pages of the 1852 economic Census were dedicated to the telegraph (Standage 1998), and two special studies of the telephone industry in 1902 and 1907 each contained over 500 pages of statistics about the structure and geography of the fledgling industry. These important documents were very influential in highlighting the uneven geographic distribution of telecommunications capabilities and paved the way for the universal service provisions that governed most national telecommunications policies for the remainder of the 20th century.

As this chapter has shown, the Internet in the United States is spreading in a highly uneven fashion, strongly favoring profitable markets and punishing economically distressed cities and remote areas. If the Internet will be the primary vehicle for commerce and communications over the next 50 years, as many in the Clinton-Gore Administration believe, accurate data about the location of Internet infrastructure and users will be essential for making sound policy decisions. Public officials have chosen to defer to the market in regulating the development of the Internet, but this research shows that a laissez-faire approach has not led to an equitable distribution of access and technology among and within cities and metropolitan areas. It follows that further systematic collection of data is necessary to more precisely determine the cause of these disparities.

REFERENCES

Abler, R. 1970. What makes cities important? *Bell Telephone Magazine*.
Castells. M. 1989. *The Information City: Information Technology, Economic Restructuring, and the Urban-Regional Process*. Cambridge MA: Blackwell.
Davis, C. 1998 *RFC 1976 Resources: Putting Locations in the Domain Name System*. [http://www.kei.com/homepages/ckd/dns-loc/]. Viewed 28 July 1998.
Dodge, M. and Shiode. N. 2000. Where on the Earth is the Internet?: An empirical investigation of the spatial patterns of Internet 'real estate' in relation to geospace in the United Kingdom. In Wheeler, J.O.. Aoyama, Y. and Warf, B. (eds.). *Cities in the Communications Age: The Fracturing of Geographies*. New York: Routledge.
Garreau, J. 1991. *Edge City: Life on the New Frontier*. New York: Doubleday.
Gaspar. J. and Glaeser, E. 1996. Information technology and the future of cities. *National Bureau of Economic Research Working Paper*, No. 5562. Cambridge MA.
Gilder, G. 1995. *Forbes ASAP*. 27 February, p.56.
Gottmann. J. 1961 *Megalopolis*. Cambridge MA: MIT Press.
Greenstein. S.. Lizardo. M. and Spiller, P. 1997.The evolution of the distribution of advanced large scale information infrastructure the United States. *National Bureau of Economic Research Working Paper*, No. 5929. Cambridge MA.
Hall, P. 1997. Modeling the post-industrial city. *Futures* 29: No. 4/5
Kolko. J. 1998. Personal communication. J. Kolko can be reached at email: kolko@nber.org
Matrix Information and Directory Services. Inc. 1996. *Matrix Maps Quarterly*, No. 303 (January). Austin TX.
Mitchelson. R.L.. and Wheeler, J.O. 1994. The flow of information in a global economy: the role of the American urban system in 1990. *Annals of the Association of American Geographers*. 84:87-107.
Moss. M.L. and Townsend. A.M. 1996. Leaders and losers on the Internet. Taub Urban Research Center, New York University. September [http://www.informationcity.org/research/]
Moss. M.L. and Townsend, A.M. 1997a. Manhattan leads the 'net nation. Taub Urban Research Center, New York University. August. [http://www.informationcity.org/research/past-research/manhattan-leads-net-nation.pdf]
Moss. M.L. and Townsend. A.M. 1997b. Tracking the 'net: Using domain names to measure the growth of the Internet in U.S. cities. *Journal of Urban Technology*. Vol. 4: No. 3. [http://www.informationcity.org/research/]
Moss, M.L.. and Townsend, A.M. 1998. Spatial analysis of the Internet in U.S. cities and states. Paper presented at the *Urban Futures - Technological Futures Conference*, Durham UK, 23-25 April, [http://urban.nyu.edu/research/newcastle/newcastle.html]
Negroponte, N. 1995. *Being Digital*. New York: Knopf.
Nunn, S. and Warren, R. 1997. Metropolitan telematics infrastructure and capacity for economic development in the information society. Center for Urban Policy and the Environment, Indiana University.
Pred, A.R. 1973. *Urban Growth and the Circulation of Information*. Cambridge MA: Harvard University Press.
Rickard, J. 1997. The Internet - what is it? *Boardwatch Magazine - Internet Service Providers Quarterly Directory*. Littleton CO, Fall.
Standage, T. 1998. *The Victorian Internet*. New York: Walker and Co.
Sui. D.Z., and Wheeler. J. 1993. The location of office space in the metropolitan service economy of the United States. 1985-1990, *The Professional Geographer* 45:33-43.
Toffler. A. 1980. *The Third Wave*. New York: William Morrow & Co.